9. Appendices

9.1. Model analysis of denaturant-induced transitions

Equilibrium denaturation studies have been very useful in understanding the structure, stabilization, and folding of numerous proteins. In this section I would like to discuss some of the denaturation models used in our study.

9.1.1. Denaturation of monomer

9.1.1. (A) *Two state monomer unfolding*

The basis for the thermodynamic analysis of the equilibrium denaturation of monomeric protein is the assumption that only 2 states, the native (N) and unfolded (denatured) (U) monomer, exist at equilibrium so that the reaction in the presence of denaturant or even at elevated temperatures can be described by **equation 9.1**.

$$N \xleftrightarrow{K} U \qquad \text{Equation 9.1}$$

where, K is equilibrium constant for the reaction.

At a given concentration of denaturant [D], the free energy for conversion of the native (N) to the unfolded (U) state, at any given denaturant concentration [D] was assumed to vary according to the empirical relationship as given by **equation 9.2** (*176, 203*):

$$\Delta G = \Delta G_0 - m \cdot [D], \qquad \text{Equation 9.2}$$

where ΔG_0 is the free energy for converting the native to the unfolded state extrapolated to zero denaturant and *m* is an empirical constant corresponding to the slope of a plot of ΔG against [D].

If the fraction of denatured species can be measured by suitable spectroscopic technique (e.g., CD, fluorescence), then the free energy of the reaction at any denaturant concentration or any temperature can be calculated from the equilibrium constant (*K*).

since $f_N + f_U = 1$ Equation 9.3

$[P] = [U] + [N]$ Equation 9.4

[*P*] is the total protein concentration of monomer

$f_N = \frac{[N]}{[P]}$; and $f_U = \frac{[U]}{[P]}$ Equation 9.5

$K = \frac{[U]}{[N]}$ or $\frac{f_U}{f_N}$ Equation 9.6

where and f_N and f_U represent the fraction of protein present in the native and unfolded states.

After solving; f_U in terms of *K* can be written as

$f_U = \frac{[K]}{[1+K]}$ Equation 9.7

At any denaturant concentration, the observed signal intensity (CD or fluorescence), S_{obs}, is given by

$S_{obs} = S_N f_N + S_U f_U$, Equation 9.8

where S_N and S_U represent the signal intensities of the native and unfolded protein at any concentration of denaturant [D].

And, $\Delta G = -RT \ln(K)$ Equation 9.9

or $\Delta G = -RT \ln (f_U/f_N)$, Equation 9,10

whereas,

$$f_N = \frac{S_{obs} - (S_U + m_U[D])}{(S_N + m_N[D]) - (S_U + m_U[D])}$$ Equation 9.11

where m_N and m_U are the dependence of the signal due to native and unfolded state of protein on denaturation concentration respectively

$$f_N = \frac{1}{1 + \exp^{-(\Delta G_0 - m[D])/RT}}$$ Equation 9.12

$$S_{obs} = \frac{S_N + S_U \exp^{-(\Delta G_0 - m[D])/RT}}{1 + \exp^{-(\Delta G_0 - m[D])/RT}}$$ Equation 9.13

where R is the gas constant, T is the absolute temperature and m is the measure of dependence of ΔG on denaturant concentration, [D]. Keeping 6 parameters variable and 3 as constant; the denaturation profiles of monomeric proteins in urea or Gdn-HCl monitored by CD or fluorescence can be fitted directly to **equation 9.11** by a nonlinear least squares routine using the Marquart algorithm. Additional terms like dependence of the signal due to native and unfolded protein on denaturant concentration should also be included.

9.1.1. (B) *Three state monomer unfolding*

Few proteins found to be having additional modes of stabilization at the quarternary structural level. The contribution of intrachain and interchain interactions to the overall protein conformational stability can make significantly different contributions. A general 3-state model of the equilibrium unfolding of the monomeric protein involving native monomer (N), partially structured monomeric intermediate (I) and unfolded or denatured monomer (U) can be described by **equation 9.14**.

$$\mathrm{N} \xleftrightarrow{K_1} \mathrm{I} \xleftrightarrow{K_2} \mathrm{U} \qquad \text{Equation 9.14}$$

where K_1 and K_2 are equilibrium constants for the first and second step of reaction defined in **equation 9.14**.

Total Protein concentration [*P*] and fraction of protein present in folded, intermediate and unfolded state can be defined as

$$[P] = [U] + [\mathrm{I}] + [N]; \qquad \text{Equation 9.15}$$

$$f_N + f_I + f_U = 1; \qquad \text{Equation 9.16}$$

$$f_N = \frac{[N]}{[P]};\ f_I = \frac{[I]}{[P]};\ \text{and}\ f_U = \frac{[U]}{[P]} \qquad \text{Equation 9.17}$$

Similar to two-state monomer unfolding model, here instead of one, two equilibrium constants K_1 for N↔I and K_2 for I↔U reaction should be calculated.

$$K_1 = \frac{[I]}{[N]}\ .or\ \frac{f_I}{f_N} \qquad \text{Equation 9.18}$$

$$K_2 = \frac{[U]}{[I]} \text{ or } \frac{f_U}{f_I} \qquad \text{Equation 9.19}$$

After solving; f_U and f_I in terms of K_1 and K_2 can be written as

$$f_U = \frac{[K_1 K_2]}{[K_1 K_2 + 1 + K_1]}; \qquad \text{Equation 9.20}$$

$$f_I = \frac{f_U}{K_2} or \frac{[K_1]}{[K_1 K_2 + 1 + K_1]}; \qquad \text{Equation 9.21}$$

At any denaturant concentration, the observed signal intensity (e.g. CD or fluorescence), S_{obs}, is given by

$$S_{obs} = \frac{S_N + S_I \exp\{-(\Delta G_{N\to I} - m_{N\to I}[D])/RT\} + S_U \exp\{-(\Delta G_{N\to I} - m_{N\to I}[D])/RT\}}{1 + \exp\{-(\Delta G_{N\to I} - m_{N\to I}[D])/RT\} + \exp\{-(\Delta G_{N\to I} - m_{N\to I}[D])/RT\}}$$
$$\frac{\exp\{-(\Delta G_{I\to U} - m_{I\to U}[D])/RT\}}{\exp\{-(\Delta G_{I\to U} - m_{I\to U}[D])/RT\}} \qquad \text{Equation 9.22}$$

where S_N, S_I and S_U are the negative molar ellipticities of the native, intermediate, and unfolded states, respectively. [*D*] is the denaturant concentration, $\Delta G_{N\to I}$ and $\Delta G_{I\to U}$ are the free energies for the N $\rightarrow$ I and I $\rightarrow$ U conversions, respectively, extrapolated to [*D*] = 0, and $m_{N\to I}$ and $m_{I\to U}$ are the *m* values for the same concentrations. *R* is the gas constant and *T* is the experimental temperature, usually 298 K. Keeping 10 parameters variable and 3 as constant; data can be fitted directly to **equation 9.22** by nonlinear least squares analysis (*191*). However *Morjana et al.* made a correction in the later issue of PNAS that PDI is dimeric and the thermodynamic parameters are independent of protein concentration. We believe statement is contradictory, as transition profiles are strictly dependent on protein concentration so the calculated thermodynamic parameters, especially when protein posses

concentration dependence therefore **equation 9.22** could be only true for unfolding of monomer via an intermediate state, not for dimer.

9.1.2. Denaturation of dimer

9.1.2. (A) *Two-state dimer denaturation*

Many proteins known to be dimer and unfold in two-state mechanism. The energetics of two-state dimer denaturation model can be determined in a manner similar to that for two-state monomer denaturation model. Here, the 2 states are now a native dimer (N_2), and a denatured or unfolded monomer (U) **Table 9.1.**

$$N_2 \xleftrightarrow{K} 2U \qquad \text{Equation 9.23}$$

where, K is equilibrium constant for the reaction.

Similar to two-state monomer unfolding,

$f_N + f_U = 1$ Equation 9.24

$[P] = [U] + 2[N_2]$ Equation 9.25

$[P]$ is the total protein concentration of monomer

$$f_N = \frac{2[N_2]}{[P]}; \text{ and } f_U = \frac{[U]}{[P]} \qquad \text{Equation 9.26}$$

$$K = \frac{[U]^2}{[N_2]} \text{ or } \frac{2[P]f_U^{\,2}}{f_N}; \qquad \text{Equation 9.27}$$

where and f_N and f_U represent the fraction of protein present in the native and unfolded states.

After solving; f_U in terms of K can be written as

$$f_U = \frac{\sqrt{K^2 + 8K[P]} - K}{4[P]}$$ Equation 9.28

Concentrations of the folded protein [N_2] (in dimer units) and the unfolded protein [U] (in monomer units) at different denaturant concentrations and either at fixed or variable temperature can be calculated by following relations:

$$[N_2] = P/2 \times \frac{S_{obs} - S_U + m_U T[D]}{S_N + m_N T[D] - S_U + m_U T[D]}$$ Equation 9.29

$$[U] = P \frac{S_N + m_N T[D] - S_{obs}}{S_N + m_N T - S_U + m_U T[D]}$$ Equation 9.30

where P is the total protein concentration in monomer units; S_{obs} is the experimentally measured signal value at a given temperature T or given denaturant concentration [D]; S_N and S_U, are the intercepts; and m_N, and m_U are the slopes of the native and unfolded baselines, respectively. The apparent equilibrium constant (K) and the corresponding free energy (ΔG) at temperature T or denaturant concentration [D] can be calculated according to:

$$K = \frac{2 \times P}{S_N + m_N T[D] - S_U + m_U T[D]} \times \frac{[S_N + m_N T[D] - S_{obs}]^2}{S_{obs} - S_U + m_U T[D]}$$ Equation 9.31

ΔG can be calculated using **equation 9.9**; $\Delta G = -RT \ln(K)$
where R is the gas constant and T is the absolute temperature.

All equilibrium unfolding data can be fitted by nonlinear least-squares analysis using a set of user-defined equations in the Nonlinear Curve Fitting window with the Origin 6.1 data analysis software package.

9.1.2. (B) *Three-state dimer denaturation*

In case of dimeric (or oligomeric) proteins additional modes of stabilization are available at the quaternary structural level. The contribution of intrachain and interchain interactions to the overall protein conformational stability can make significantly distinct contributions. A general three-state model of equilibrium dissociation and unfolding of the dimeric protein involves native dimer (N_2), monomeric intermediate (I) or dimeric intermediate (I_2), and unfolded monomer (U) **Table 9.1.**

$$N_2 \xleftrightarrow{K_1} 2 \cdot I \xleftrightarrow{K_2} 2 \cdot U$$ Equation 9.32

with monomeric intermediate;

$$N_2 \xleftrightarrow{K_1} I_2 \xleftrightarrow{K_2} 2 \cdot U$$ Equation 9.33

with dimeric intermediate.

9.1.2. (B) (I) *Three state dimer denaturation assuming monomeric intermediates*

In the first step protein dimers (N_2) dissociate in intermediates (I) that unfold in a second reaction step into unfolded monomers (U) according to **equation 9.33**:

$$N_2 \xleftrightarrow{K_1} 2 \cdot I \xleftrightarrow{K_2} 2 \cdot U$$

where $K_1 = [I]^2/[N_2]$ Equation 9.34

And $K_2 = [U]/[I]$ Equation 9.35

The dissociation free energy ΔG_1 between states N_2 and I, and unfolding free energy ΔG_2 between states I and U are given by:

$$\Delta G_1 = -RT\ln K_1 \text{ and } \Delta G_2 = -RT\ln K_2 \qquad \text{Equation 9.36}$$

where R and T are the gas constant and the temperature in Kelvin, respectively. The global stability of the dimer can be described by the equilibrium constant K and the unfolding free energy ΔG for the total unfolding reaction between states N_2 and U as

$$K = K_1K_2^2 \text{ and } \Delta G = -RT\ln K = -RT\ln(K_1K_2^2)=\Delta G_1 + 2\Delta G_2 \qquad \text{Equation 9.37}$$

For the calculation of the free energy at 0 M denaturant concentration one assumes that the variation of free energy between two conformational states is a linear function of denaturant concentration, [D] **Table 9.1.**

Therefore,

$$\Delta G_1 = \Delta G_1(H_2O) - m_1\cdot[D] \qquad \text{Equation 9.38}$$

$$\text{and } \Delta G_2 = \Delta G_2(H_2O) - m_2\cdot[D] \qquad \text{Equation 9.39}$$

where $\Delta G(H_2O)$ is the free energy change in the absence of denaturant, and m is the slope of the ΔG versus [D] plot.

The global signal S of the unfolding mixture after subtraction of background is composed from contributions of dimers (S_nF_n), intermediates (S_iF_i) and unfolded monomers (S_uF_u):

$$S_{obs} = S_nF_n + S_iF_i + S_uF_u \qquad \text{Equation 9.40}$$

where S_n, S_i, and S_u are the ellipticity signals for states N_2, I and U, and F_n, F_i and F_u are the fractions of protein in the dimeric, intermediate and unfolded state, respectively. F_n, F_u and F_i sum up to 1 in accordance with the law of mass conservation. In principle, S_n, S_i, and S_u are linear functions of [D], however, S_i can be considered as a constant incase I exists over a narrow range of denaturant concentration.

The experimental data should be normalized:

$$S_{obs} = S_n + F_{app}\,(S_u - S_n) \qquad \text{Equation 9.41}$$

with $$F_{app} = F_u + F_i[(S_i - S_n)/(S_u - S_n)] \qquad \text{Equation 9.42}$$

The global spectroscopic signal (S) results as

$$S = S_n + \{F_u + F_i[(S_i - S_n)/(S_u - S_n)]\} * (S_u - S_n) \qquad \text{Equation 9.43}$$

with $$F_i = F_u/K_2,\ F_u = K_2F_i, \text{ and } F_n = 2F_u^2[P]/K_1K_2^2 \qquad \text{Equation 9.44}$$

The fractions of unfolded monomer, F_u, and of intermediate, F_i, have to be solved in terms of K_1 and K_2:

$$F_u = \frac{K_1K_2\left[-(K_2+1)+\sqrt{(K_2+1)^2+8[P]/K_1}\right]}{4P} \qquad \text{Equation 9.45}$$

$$F_i = F_u/K_2 = \frac{K_1K_2\left[-(K_2+1)+\sqrt{(K_2+1)^2+8[P]/K_1}\right]}{4P\cdot K_2} \qquad \text{Equation 9.46}$$

where P is the total molar concentration of protein in monomer equivalents.

$P = 2[N_2] + [I] + [U]$. Equation 9.47

All equilibrium unfolding data were fitted by nonlinear least-squares analysis using a set of user-defined equations in the NonLinear Curve Fitting window with the Origin 6.1 data analysis software package.

9.1.2. (B) (II) *Three state dimer denaturation with dimeric intermediate*

From **equation 9.33 and 9.40;**

$$N_2 \xleftrightarrow{K_1} I_2 \xleftrightarrow{K_2} 2 \cdot U$$

$S_{obs} = S_nF_n + S_iF_i + S_uF_u$

Total protein concentration [*P*] in monomer units and fraction of protein present in different folded, intermediate and unfolded state can be defined as

$[P] = [U] + 2[I_2] + 2[N_2]$; Equation 9.48

$f_N + f_I + f_U = 1$;

$f_N = \frac{2[N_2]}{[P]}; f_I = \frac{2[I_2]}{[P]}$; and $f_U = \frac{[U]}{[P]}$ Equation 9.49

Similar to three-state monomer unfolding model, here, two equilibrium constants K_1 for $N_2 \leftrightarrow I_2$ and K_2 for $I_2 \leftrightarrow U$ reaction should be calculated and all equilibrium unfolding data should be fitted in same way as **9.1.2. (B) (I)** considering equations defined below.

$$K_1 = \frac{[I_2]}{[N_2]} or \frac{f_I}{f_N}$$ Equation 9.50

$$K_2 = \frac{[U]^2}{[I_2]} or \frac{2f_U{}^2[P]}{f_I}$$ Equation 9.51

After solving; f_U and f_I in terms of K_1 and K_2 can be written as

$$F_N = 2F_U{}^2[P]/K_1K_2$$ Equation 9.52

$$F_U = \frac{-K_1K_2 + \sqrt{K_1K_2(K_1K_2 + 8[P](1+K_1))}}{4[P](1+K_1)};$$

or $$\frac{K_1K_2\left(-1+\sqrt{(1+8[P](1+K_1)/K_1K_2)}\right)}{4[P](1+K_1)}$$ Equation 9.53

$$f_I = \frac{2f_U{}^2[P]}{K_2}; \text{ or } 2[P]/K_2\left(\frac{K_1K_2\left(-1+\sqrt{(1+8[P](1+K_1)/K_1K_2)}\right)}{4[P](1+K_1)}\right)^2$$ Equation 9.54

Table 9.1: Equilibrium unfolding of dimeric proteins monitored by spectrometry as expressed in dimeric equivalents. The abbreviations used are: N_2, native dimeric state; I_2, dimeric intermediate state; I, is monomeric intermediate state; U, unfolded state; P, total concentration of protein (M), expressed as dimer equivalent; [D], concentration of denaturant (M); S_{obs}, measured global signal; S, corrected signal after blank subtraction,; S_N, S_I, and S_U are molar signals of the corresponding protein states; T, Temperature (K); R, gas constant. The other parameters defined below.

	Without intermediate	Monomeric intermediate	Dimeric intermediate
Equilibrium	$N_2 \leftrightarrow 2U$	$N_2 \leftrightarrow 2I \leftrightarrow 2U$	$N_2 \leftrightarrow I_2 \leftrightarrow 2U$
Law of mass action	$K=[U]^2/[N_2]$	$K_1=[I]^2/[N_2]$ $K_2=[U]/[I]$	$K_1=[I_2]/[N_2]$ $K_2=[U]^2/[I_2]$
Conservation of mass	$[N_2] + [U]/2 = P$	$[N_2]+[I]/2+[U]/2= P$	$[N_2]+[I_2]+[U]/2= P$
Additivity of the signals	$S_{obs}= S_N[N_2] + S_U[U]$	$S_{obs}=S_N[N_2]+S_I[I]+S_U[U]$	$S_{obs}=S_N[N_2]+S_I[I_2]+S_U[U]$
Variations of the signals with denaturant[D]	$S_N = S_N + m_N$[D] ; $S_U = S_U + m_U$[D]	S_I = constant	S_I = constant
Variation of ΔG with [D]	$\Delta G= \Delta G_0 - m\cdot$[D] $= -RT \ln(K)$	Similar laws for ΔG_1 and ΔG_2	Similar laws for ΔG_1 and ΔG_2
Definition of the molar fractions	$f_N= [N_2]/P$; $f_U= [U]/2P$	$f_I= [I]/2P$;	$f_I= [I_2]/P$;
Solving equations	$4Pf_U^2 + K\cdot f_U - K =0$; $f_N = 1- f_U$ $S= \{S_N + (2S_U - S_N) f_U\}$	$4Pf_I^2+K_1(1+K_2)f_I-K_1= 0$; $f_N = 1- (1+K_2)f_I$ $f_U = K_2\cdot f_I$ $S= P\{S_N+[2S_I-S_N+K_2(2S_U - S_N)] f_I\}$	$4P(1+K_1)f_U^2+K_1K_2\cdot f_U-K_1K_2=0$; $f_N = (1- f_U)/(1+K_1)$ $f_I = (1- f_U) K_1/(1+K_1)$ $S=P\{2S_U+[S_N-2S_U+K_1(S_I - 2S_U)] f_N\}$
Total equilibrium constant and ΔG	K; ΔG	$K= K_1K_2^2$; $\Delta G= \Delta G_1+2\Delta G_2$	$K= K_1K_2$; $\Delta G= \Delta G_1+\Delta G_2$
Number of equations and parameters	5 equations, 5 parameters	8 equations, 8 parameters	8 equations, 8 parameters

9.2. Amino acid sequences (wild type)

9.2.1. *Streptococcus pyogenes* β recombinase

MAKIGYARVSSKEQNLDRQLQALQGVSKVFSDKLSGQSVERPQLQAMLNYIRE
GDIVVVTELDRLGRNNKELTELMNAIQQKGATLEVLNLPSMNGIEDENLRRLINN
LVIELYKYQAESERKRIKERQAQGIEIAKSKGKFKGRQHKFKENDPRLKHAFDLF
LNGCSDKEVEEQTGINRRTFRRYRTRYNVTVDQRKNKGKRDS

9.2.2. *Escherichia* coli RelB

MGSINLRIDDELKARSYAALEKMGVTPSEALRLMLEYIADNERLPFKQTLLSDED
AELVEIVKERLRNPKPVRVTLDEL

9.2.3. *Escherichia* coli RelE

MAYFLDFDERALKEWRKLGSTVREQLKKKLVEVLESPRIEANKLRGMPDCYKIK
LRSSGYRLVYQVIDEKVVVFVISVGKRERSEVYSEAVKRIL

9.2.4. *Methanococcus jannaschii* RelB

MRLKKRFKKFFISRKEYEKIEEILDIGLAKAMEETKDDELLTYDEIKELLGD

9.2.5. *Methanococcus jannaschii* RelE

MKVLFAKTFVKDLKHVPGHIRKRIKLIIEECQNSNSLNDLKLDIKKIKGYHNYYRIR
VGNYRIGIEVNGDTIIFRRVLHRKSIYDYFP

Table 9.2: Calculated physico-chemical parameters and available details for proteins. (www.expasy.org)

Protein	β recombinase	RelB	RelE+ 6x His.[a]	RelBE +6x His.[a]	RelBE
Organism	*S. pyogenes*	*E. coli*	*E. coli*	*E. coli*	*M. jannaschii*
UniProtKB/ TrEMBL; Primary accession number	Q57437	P0C079	P0C077	-	RelB (N.A.) ; RelE Q58503
Number of amino acids	205	79	95+6	173[a]+6	139 (51+88)
Molecular weight (Da)	23841.1	8940.3	11225.2+822.9	20988.4 [a]	16801 (6239+10562)
Theoretical pI	9.86	4.81	9.67	8.8	9.68
Number of Cysteine	1	0	1	1	1
Number of Tryptophan	0	0	1	1	0
Number of Tyrosine	6	2	5	7	8
Total number of negatively charged residues	28	16	14	30	21
Total number of positively charged residues	41	12	21	33	32
Molar extinction coefficient (at 278 nm); $M^{-1} \cdot cm^{-1}$	9702	3037 [a]	13291	16328 [a,b]	12739 [b]

[a] : In absence of first amino acid methionine.
[b] : calculated for hetero-dimer.
N.A.: Not available