Protein Unfolding Data Analysis

Spectroscopy

Adrian Velazquez-Campoy

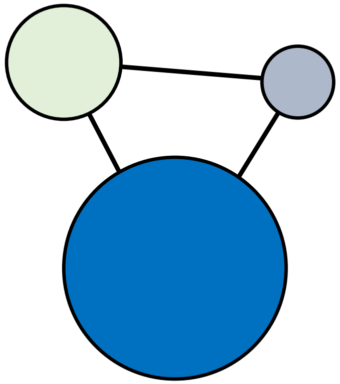
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

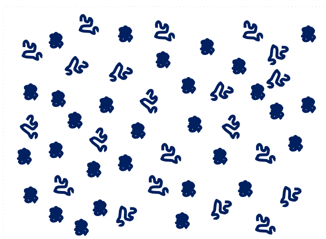
MOSBRI
Molecular-Scale Biophysics Research Infrastructure

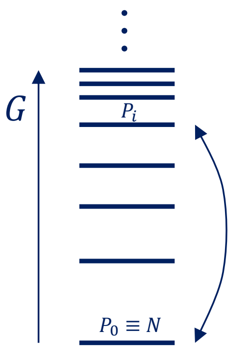
Stability

ΔG Gibbs energy of stabilization or unfolding

Difference in Gibbs energy between the native state and the unfolded state







$$\chi_i \propto e^{-(\Delta G_i/RT)}$$

$$\Delta G_i = G_i - G_0 = -RT \ln \frac{[P_i]}{[P_0]} = -RT \ln \beta_i$$

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101004806

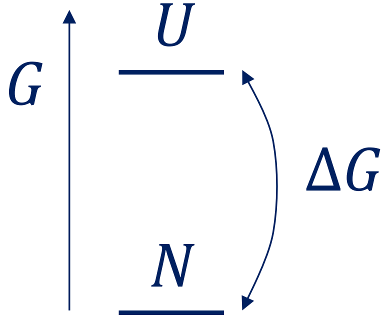
$$\Delta G = \Delta G(T, P, pH, \mu, [D], [L])$$

$$\Delta G([D]) = \Delta G_0 - m[D]$$

$$\Delta G(T) = \Delta H(T_m)\left(1 - \frac{T}{T_m}\right) + \Delta C_P\left(T - T_m - T\ln\left(\frac{T}{T_m}\right)\right)$$

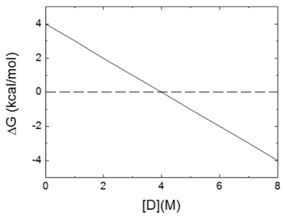
- **Goal:** **Increase environmental stress and displace the equilibrium in order to make populations of comparable size, exploring different population ratios between N and U**
  Calculate/estimate K

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

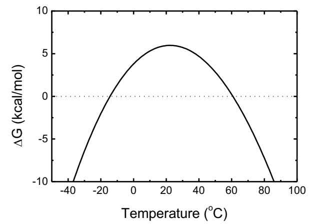


$$\Delta G = \Delta G(T, P, pH, \mu, [D], [L])$$

$$\Delta G([D]) = \Delta G_0 - m[D]$$

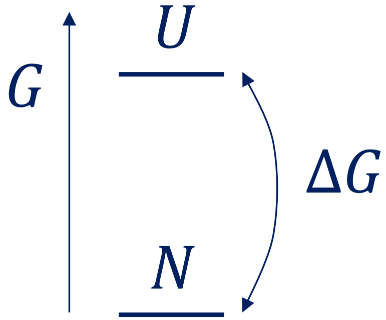


$$\Delta G(T) = \Delta H(T_m)\left(1 - \frac{T}{T_m}\right) + \Delta C_P\left(T - T_m - T\ln\left(\frac{T}{T_m}\right)\right)$$



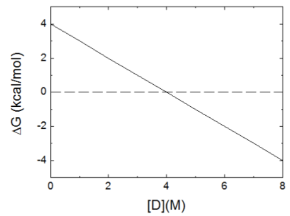
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

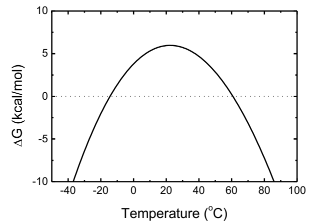


$$\Delta G = \Delta G(T, P, pH, \mu, [D], [L])$$

$$\Delta G([D]) = \Delta G_0 - m[D]$$



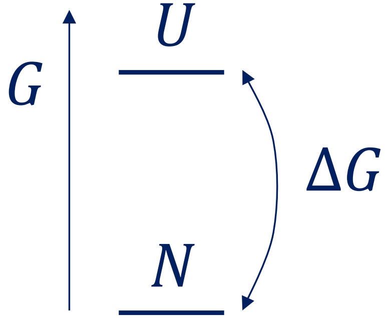
$$\Delta G(T) = \Delta H(T_m)\left(1 - \frac{T}{T_m}\right) + \Delta C_P\left(T - T_m - T\ln\left(\frac{T}{T_m}\right)\right)$$



ΔG_0	**Stability energy**
m	**Susceptibility against denaturant concentration**
$\Delta H(T_m)$	**Unfolding enthalpy**
T_m	**Unfolding temperature**
ΔC_P	**Unfolding heat capacity**

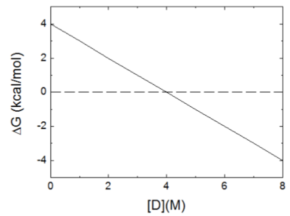
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

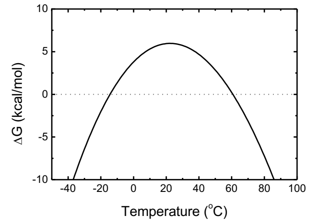


$$\Delta G = \Delta G(T, P, pH, \mu, [D], [L])$$

$$\Delta G([D]) = \Delta G_0 - m[D]$$



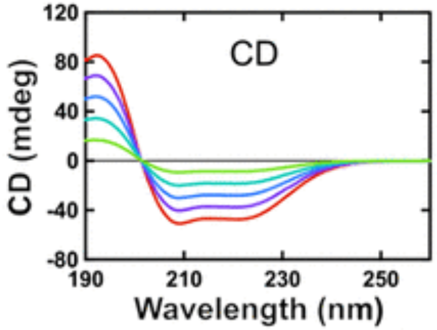
$$\Delta G(T) = \Delta H(T_m)\left(1 - \frac{T}{T_m}\right) + \Delta C_P\left(T - T_m - T\ln\left(\frac{T}{T_m}\right)\right)$$

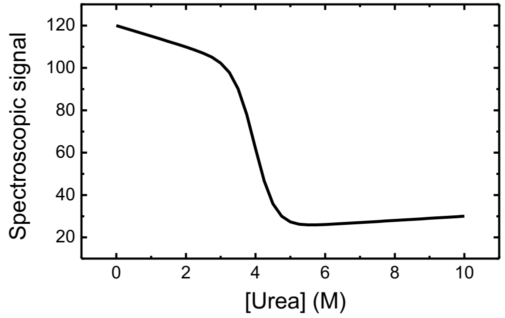


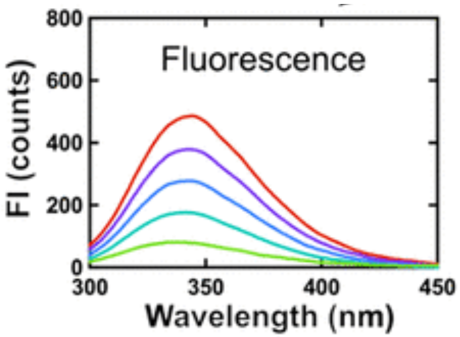
Parameters define the stability profile or diagram: ΔG([D]) y ΔG(T)
Parameters can be determined experimentally

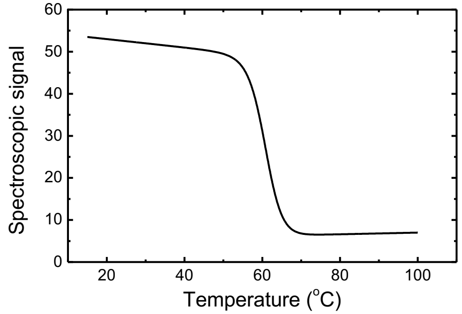
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

Unfolding experiment



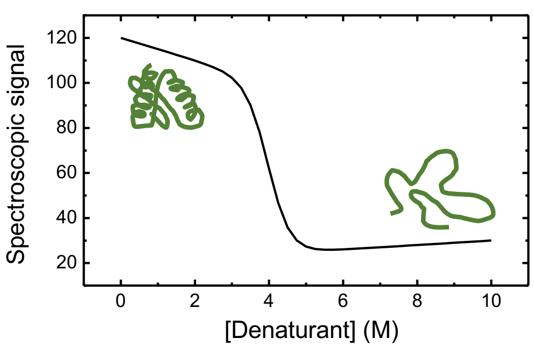


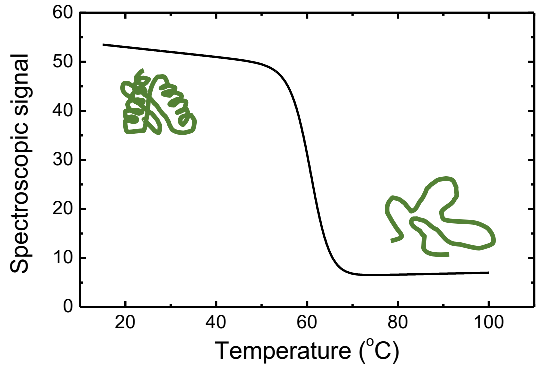




This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure





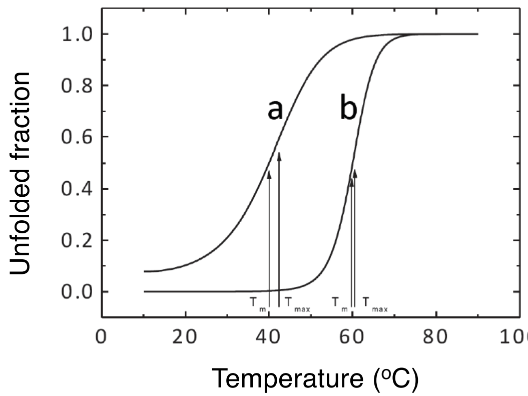
Cooperative unfolding → Reduction in the number of accessible states

Partially folded states are not significantly populated

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

▶ Data analysis?

- Unless you are doing a screening evaluation, never use sigmoidal fittings (e.g., Boltzmann function)
- Never apply "naïve" normalization
- The inflection point in chemical denaturation provides $[D]_m$
- The inflection point in thermal denaturation DOES NOT provide T_m ($T_m \lesssim T_{max}$)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

▶ Standard approach based on the partition function *Q*

$$P_0 \leftrightarrow P_1 \leftrightarrow P_2 \leftrightarrow \cdots \leftrightarrow P_i \leftrightarrow \cdots \leftrightarrow P_n$$

$$[P_i] = \beta_i [P_0]$$

$$[P_i] = \left(\prod_{r=1}^{i} K_r \right) [P_0]$$

$$Q = \sum_{i=0}^{n} \frac{[P_i]}{[P_0]} = \sum_{i=0}^{n} \beta_i = \sum_{i=0}^{n} e^{-(\Delta G_i/RT)}$$

Wyman & Gill. "Binding and Linkage", University Science Books, 1990

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

▶ Standard approach based on partition function *Q*

$$Q = \sum_{i=0}^{n} \frac{[P_i]}{[P_0]} = \sum_{i=0}^{n} \beta_i$$

$$\chi_i = \frac{[P_i]}{[P]_T} = \frac{\beta_i}{Q}$$

$$\beta_i = exp(-\Delta G_i/RT)$$

$$\Delta G_i = \Delta G_i(T, pH, \mu, [D], [L], a_w, \dots)$$

$$\Delta G_i([D]) = \Delta G_{0,i} - m_i[D]$$

$$\Delta G_i(T) = \Delta H_i(T_{m,i})\left(1 - \frac{T}{T_{m,i}}\right) + \Delta C_{P,i}\left(T - T_{m,i} - T\ln\frac{T}{T_{m,i}}\right)$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

▶ Standard approach based on the partition function *Q*

$$Q = \sum_{i=0}^{n} \frac{[P_i]}{[P_0]} = \sum_{i=0}^{n} \beta_i$$

$$\chi_i = \frac{[P_i]}{[P]_T} = \frac{\beta_i}{Q}$$

$$\beta_i = exp(-\Delta G_i/RT)$$

$$\langle S \rangle = \sum_{i=0}^{n} \chi_i S_i = \frac{\sum_{i=0}^{n} \beta_i S_i}{\sum_{i=0}^{n} \beta_i}$$

Average of any system property (measurement)

$$NLSF \rightarrow \{T_{mi}, \Delta H_i, \Delta C_{Pi}; i = 1, \ldots, n\}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

$$N \leftrightarrow U$$

i	G	ΔG	$\exp(-\Delta G/RT)$
	G_N	0	1
	G_U	ΔG	K

$$Q = 1 + \beta = 1 + K$$

$$\chi_N = \frac{1}{1+K}$$

$$\chi_U = \frac{K}{1+K}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

Chemical Denaturation

$$\Delta G([D]) = \Delta G_0 - m[D]$$

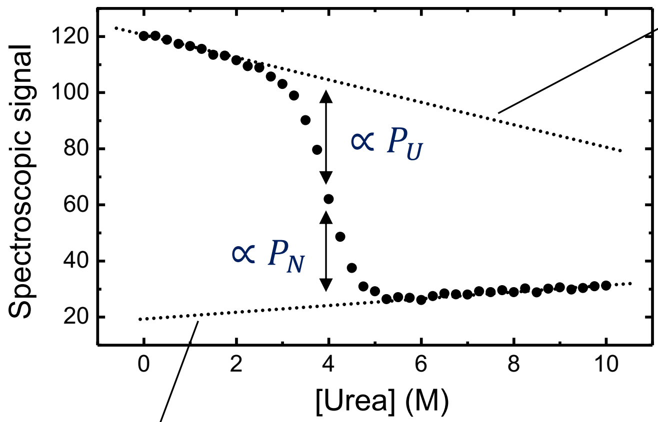
$$\Delta G_0 = \Delta G([D])|\Big|_{[D]=0} \qquad m = -\left(\frac{\partial \Delta G([D])}{\partial [D]}\right)_{T,P}$$

$$\Delta G([D]) = -RT \ln K([D])$$

$$\Delta G([D]_m) = 0 \longrightarrow [D]_m = \frac{\Delta G_0}{m}$$

$$\Delta G([D]_m) = 0 \longrightarrow \begin{cases} K = 1 \\ \chi_N = \chi_U = 0.5 \end{cases}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806



$$S_N = a_N + b_N[D]$$

$$S_U = a_U + b_U[D]$$

$$S = \chi_N S_N + \chi_U S_U$$

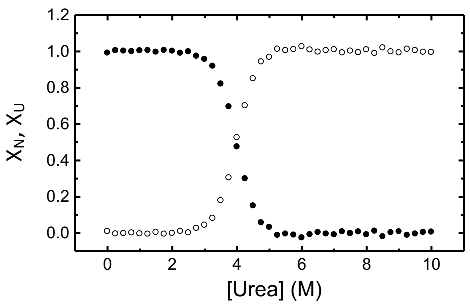
$$\left.\begin{aligned} \chi_N &= \frac{S - S_U}{S_N - S_U} \\ \chi_U &= \frac{S_N - S}{S_N - S_U} \end{aligned}\right\}$$

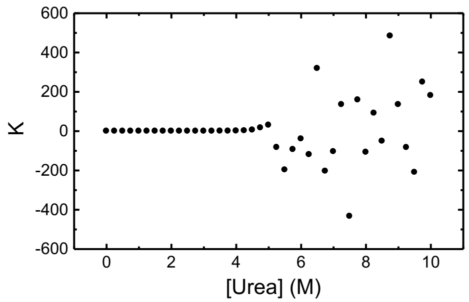
$$K = \frac{\chi_U}{\chi_N}$$

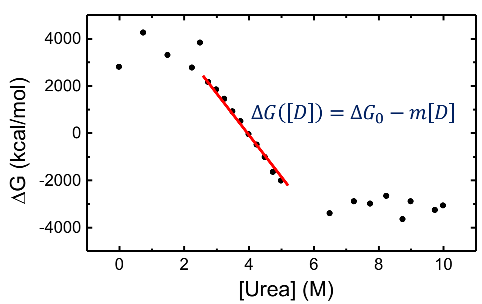
$$\Delta G = -RT \ln K$$

ΔG_0 7600 ± 200 cal mol^{-1}

m 1920 ± 40 cal mol^{-1} M^{-1}

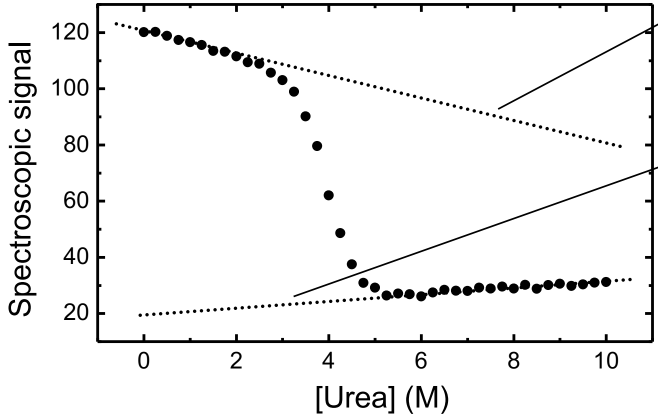






This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

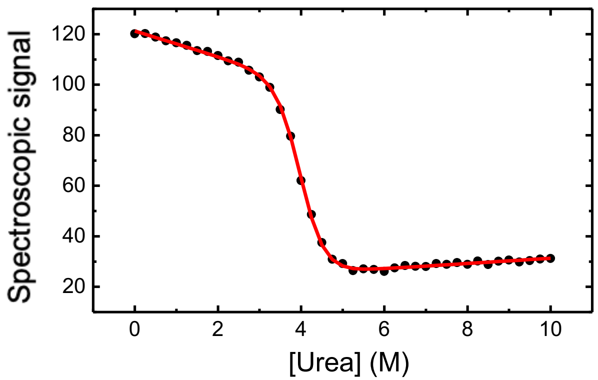


$$S_N = a_N + b_N[D]$$

$$S_U = a_U + b_U[D]$$

$$S = \chi_N S_N + \chi_U S_U$$

$$S = \frac{(a_N + b_N[D]) + e^{-(\Delta G_0 - m[D])/RT}(a_U + b_U[D])}{1 + e^{-(\Delta G_0 - m[D])/RT}}$$



ΔG_0 8200 ± 50 cal mol^{-1}
m 2050 ± 10 cal mol^{-1} M^{-1}

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

Thermal Denaturation

$$\Delta G(T) = \Delta H(T_m)\left(1 - \frac{T}{T_m}\right) + \Delta C_P\left(T - T_m - T\ln\left(\frac{T}{T_m}\right)\right)$$

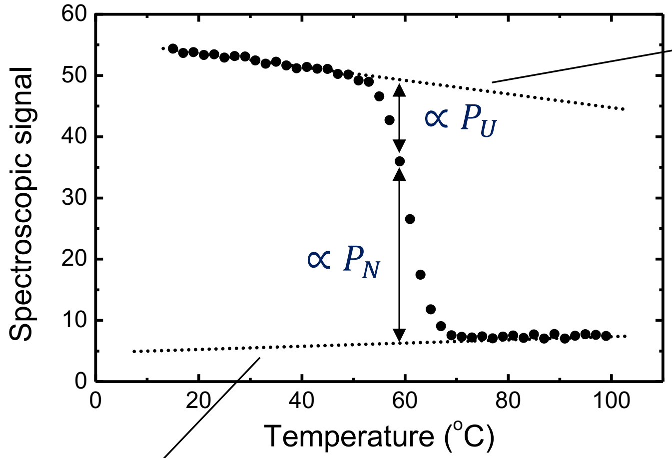
$$\Delta H(T_m) = \Delta H(T)|\Big|_{T=T_m}$$

$$\Delta C_P = \left(\frac{\partial \Delta H(T)}{\partial T}\right)_P = T\left(\frac{\partial \Delta S(T)}{\partial T}\right)_P$$

$$\Delta G(T) = -RT\ln K(T)$$

$$\Delta G(T_m) = 0 \longrightarrow \begin{cases} K = 1 \\ \chi_N = \chi_U = 0.5 \\ \Delta S(T_m) = \dfrac{\Delta H(T_m)}{T_m} \end{cases}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806



$$S_N = a_N + b_N T$$

$$S = \chi_N S_N + \chi_U S_U$$

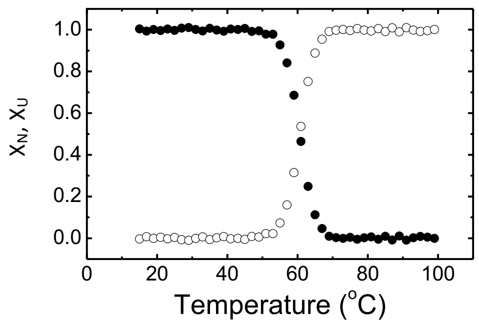
$$S_U = a_U + b_U T$$

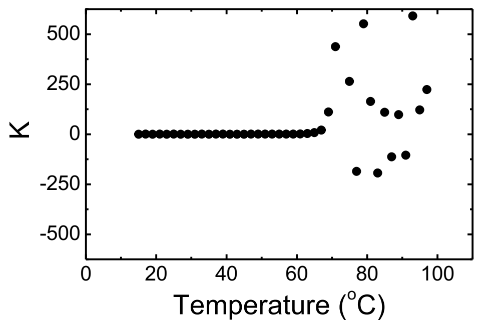
$$\left.\begin{aligned}\chi_N &= \frac{S - S_U}{S_N - S_U}\\ \chi_U &= \frac{S_N - S}{S_N - S_U}\end{aligned}\right\}$$

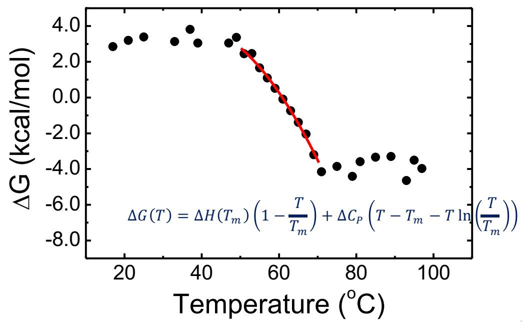
$$K = \frac{\chi_U}{\chi_N}$$

$$\Delta G = -RT \ln K$$

$\Delta H(T_m)$	109 ± 2 kcal mol^{-1}
T_m	60.9 ± 0.2 °C
ΔC_p	4 ± 1 kcal K^{-1} mol^{-1}

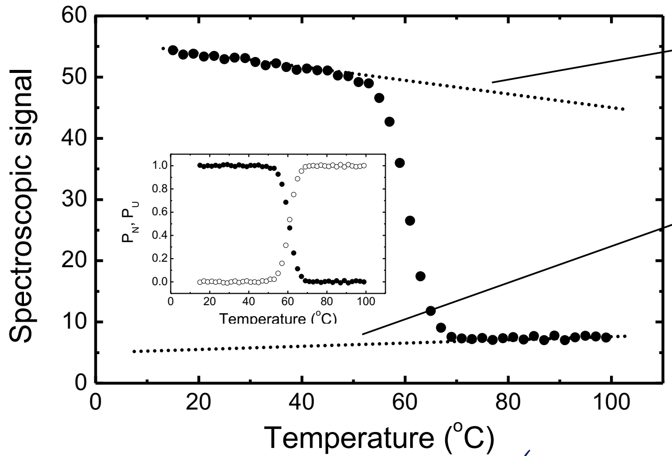






This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure



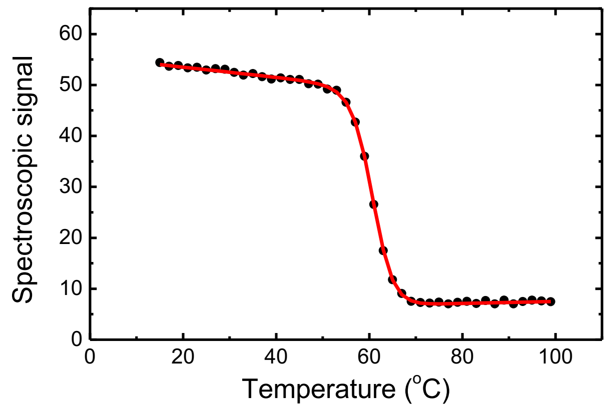
$$S_N = a_N + b_N T$$

$$S_U = a_U + b_U T$$

$$S = \chi_N S_N + \chi_U S_U$$

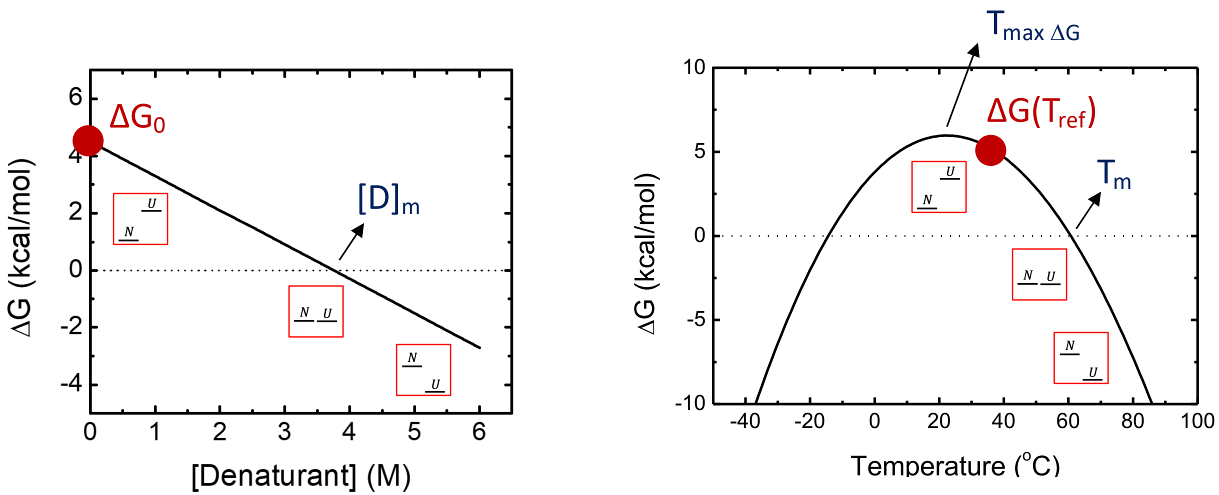
$$S = \frac{(a_N + b_N T) + e^{-\left(\Delta H(T_m)\left(1-\frac{T}{T_m}\right)+\Delta C_P\left(T-T_m-T\ln\left(\frac{T}{T_m}\right)\right)\right)/RT}(a_U + b_U T)}{1 + e^{-\left(\Delta H(T_m)\left(1-\frac{T}{T_m}\right)+\Delta C_P\left(T-T_m-T\ln\left(\frac{T}{T_m}\right)\right)\right)/RT}}$$

$\Delta H(T_m)$	101 ± 2 kcal mol^{-1}
T_m	60.69 ± 0.04 °C
ΔC_P	2.2 ± 0.7 kcal K^{-1} mol^{-1}



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

▶ Stability profile

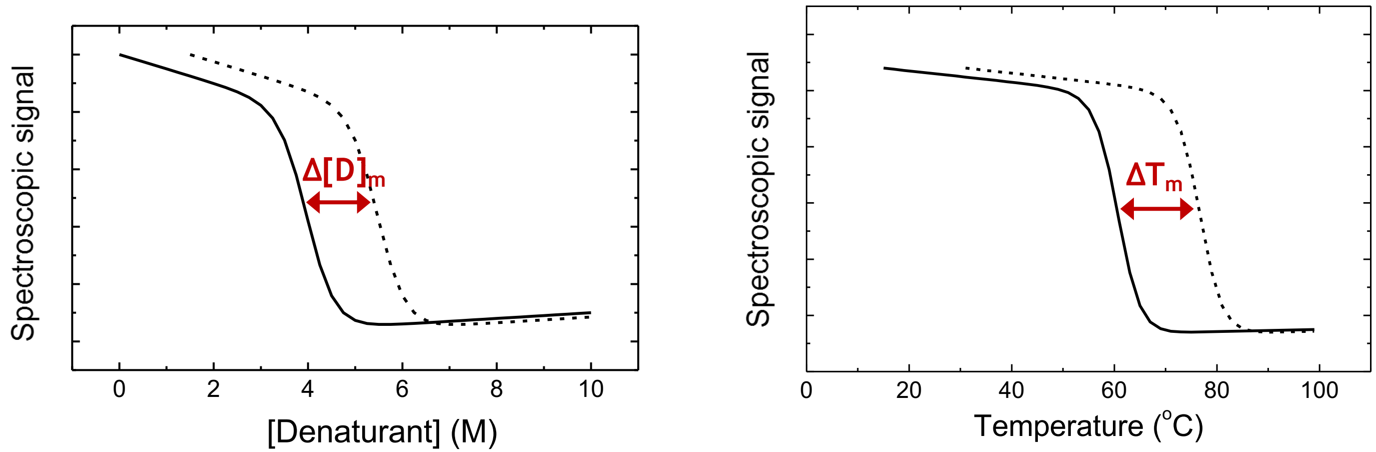


$$\Delta G([D]) = \Delta G_0 - m[D]$$

$$\Delta G(T) = \Delta H(T_m)\left(1 - \frac{T}{T_m}\right) + \Delta C_P\left(T - T_m - T\ln\frac{T}{T_m}\right)$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

▶ Stability differences?

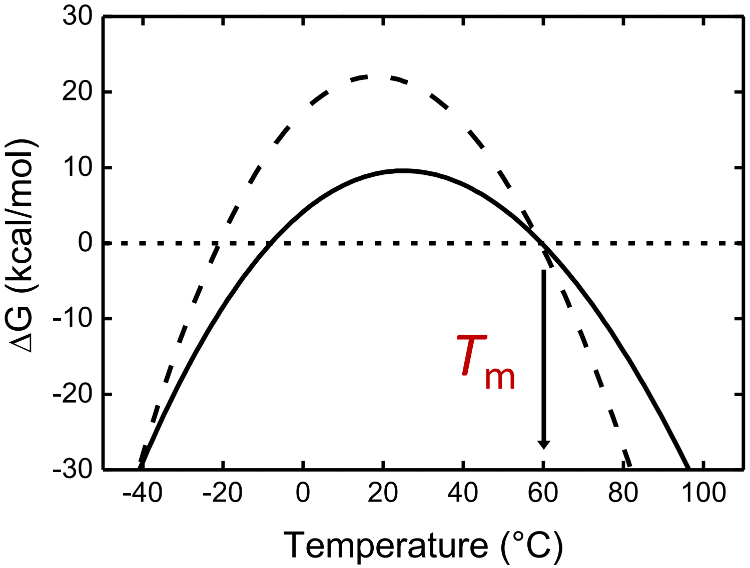


Changes in T_m and $[D]_m$ are easily to observe quantify, but they are NOT reliable indexes for stability or stability changes

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

High and low temperature stability

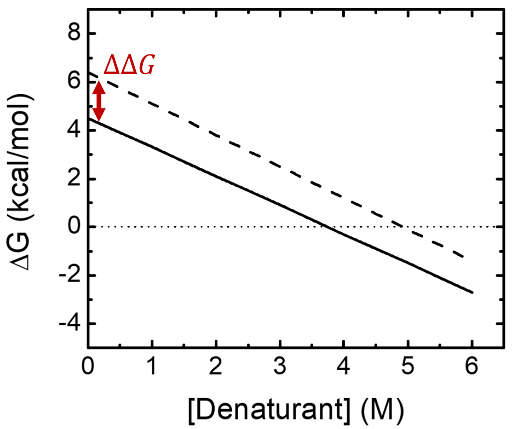


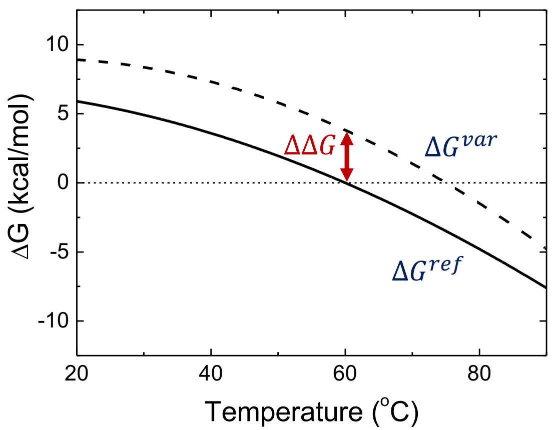
$\Delta H(T_m)$	180 kcal mol^{-1}
$\Delta S(T_m)$	0.54 kcal K^{-1} mol^{-1}
ΔC_P	5.0 kcal K^{-1} mol^{-1}
$\Delta H(T_m)$	350 kcal mol^{-1}
$DS(T_m)$	1.05 kcal K^{-1} mol^{-1}
ΔC_p	8.0 kcal K^{-1} mol^{-1}

Which protein is more stable?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

▶ Stability differences?





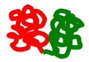
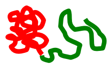
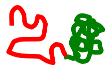
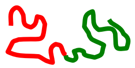
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

Two-transition unfolding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

$$N \leftrightarrow I_1 \leftrightarrow I_2 \leftrightarrow U$$

i	ΔG_i	$\exp(-\Delta G_i/RT)$
	0	1
	ΔG_1	K_1
	ΔG_2	K_2
	$\Delta G_1+\Delta G_2$	$K_1 K_2$

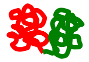
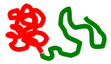
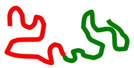
$$Q = 1 + \beta_1 + \beta_2 + \beta_3 = 1 + K_1 + K_2 + K_1 K_2$$

$$\chi_N = \frac{1}{1 + K_1 + K_2 + K_1 K_2} \qquad \chi_{I_2} = \frac{K_2}{1 + K_1 + K_2 + K_1 K_2}$$

$$\chi_{I_1} = \frac{K_1}{1 + K_1 + K_2 + K_1 K_2} \qquad \chi_U = \frac{K_1 K_2}{1 + K_1 + K_2 + K_1 K_2}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

$$N \leftrightarrow I \leftrightarrow U$$

i	ΔG_i	$\exp(-\Delta G_i/RT)$
	0	1
	ΔG_1	K_1
	$\Delta G_1 + \Delta G_2$	$K_1 K_2$

$$Q = 1 + \beta_1 + \beta_2 = 1 + K_1 + K_1 K_2$$

$$\chi_N = \frac{1}{1 + K_1 + K_1 K_2} \qquad \chi_{I_2} \approx 0$$

$$\chi_{I_1} = \frac{K_1}{1 + K_1 + K_1 K_2} \qquad \chi_U = \frac{K_1 K_2}{1 + K_1 + K_1 K_2}$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

▶ Selecting an appropriate unfolding model

- Understanding "names" of models, as well as physical idealization and mathematical formalism
- Make sense of estimated parameter values
- Make sense of parameters uncertainties

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

▶ Best model for analysis?

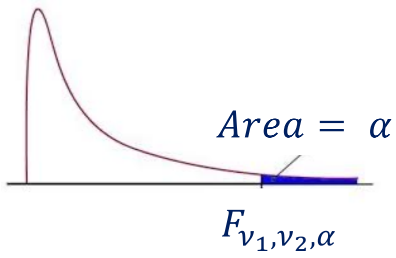
- Employ structural/functional data, if available
- Revise carefully experimental data
- Identify unusual features
- Start with simplest compatible model (Ockham's razor)
- Increase complexity until unnecessary (no statistical improvement or no convergence)
- Understand assumptions/constraints of model
- Judge critically the estimated parameters

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

▶ When to stop increasing model complexity?

- **Select the (statistically) best model**

$$\chi^2 \equiv RSS = \sum_{i=1}^{n} \left(y(x_i) - f(\theta; x_i) \right)^2$$



Model 1: $n - p_1, \chi_1^2$

$(p_2 > p_1)$ **Model 2:** $n - p_2, \chi_2^2$

$$F = \frac{(\chi_1^2 - \chi_2^2)/(p_2 - p_1)}{\chi_2^2/(n - p_2)} \sim F_{p_2-p_1, p_2-p_1}$$

$F \leq F_{p_2-p_1, p_2-p_1, \alpha}$ **select Model 1 (H_0)**

$F > F_{p_2-p_1, p_2-p_1, \alpha}$ **select Model 2 (H_1)**

$$AIC = n \ln \left(\frac{\chi^2}{n - p} \right) + 2p$$

$$BIC = n \ln \left(\frac{\chi^2}{n - p} \right) + p \ln n$$

Select model with minimum AIC or BIC

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Global fitting

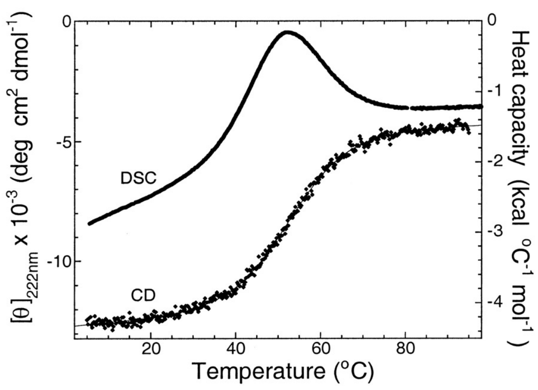
- **Strengthen model selection**
- **Strengthen parameter estimation**
- **Reduce uncertainties of estimated parameters**
- **Overcome experimental limitations**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

▶ Global fitting?

- Replicate experiments
- Concentration ranges
- Together with other biophysical techniques

...



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure