

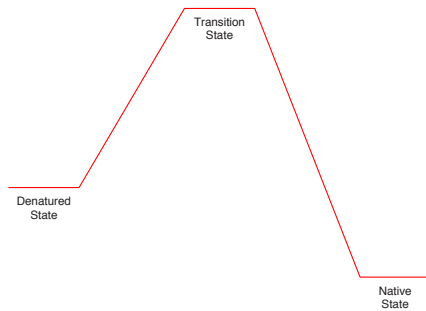
Statistics Schmatistics

On the folded, the unfolded, and the transition state

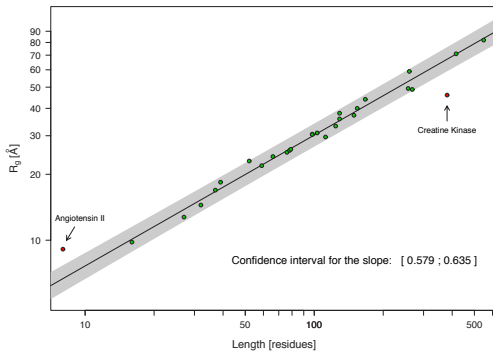
Ingo Ruczinski

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health

Energy Profile



Radius of Gyration of Denatured Proteins



Biostatistics for Laboratory Scientists

140.615 Introduction to Biostatistics for Laboratory Scientists I
3rd term (January 16 - March 9, 2007).

140.616 Introduction to Biostatistics for Laboratory Scientists II
4th term (March 19 - May 11, 2007).

MWF 10:30 - 11:50.

East Baltimore Campus
Bloomberg School of Public Health
615 N Wolfe St

Radius of Gyration of Denatured Proteins

Do chemically denatured proteins behave like random coils?

- The radius of gyration R_g of a protein is defined as the root mean square distance from each atom of the protein to their centroid.
- For an ideal (infinitely thin) random-coil chain in a solvent, the average radius of gyration of a random coil is a simple function of its length n :

$$R_g \propto n^{0.5}$$

- For an excluded volume polymer in a solvent, the average radius of gyration of a random coil is given by

$$R_g \propto n^{0.588}$$

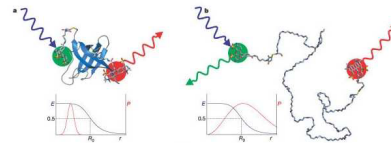
→ The radius of gyration can be measured using small angle x-ray scattering.

Reference:
Flory PJ (1953). Principles of Polymer Chemistry, Cornell University Press.

Deviations from Random Coil Behavior

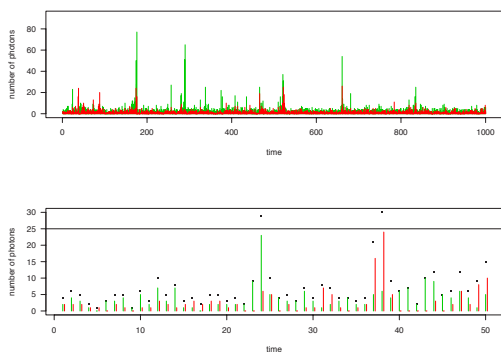
Are there site-specific deviations from random coil dimensions?

Förster Resonance Energy Transfer enables us to measure the distance between two dye molecules within a certain range. This can be used to study site-specific deviations from random coil dimensions in denatured peptides.

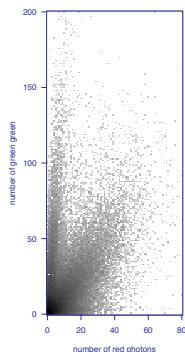


Reference:
Schuler B, Lipman EA, Eaton WA (2002). Abstract Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature*, 419 (6908): 743-7.

Deviations from Random Coil Behavior



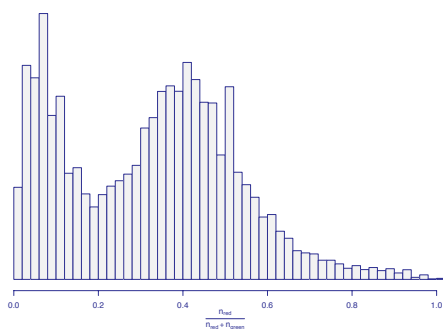
Deviations from Random Coil Behavior



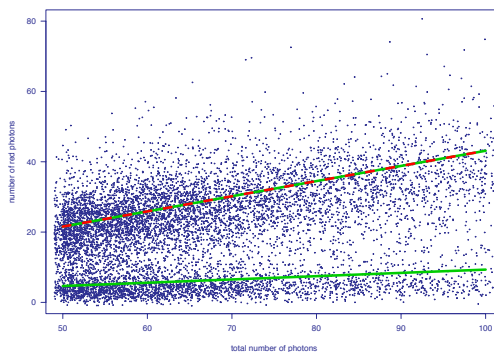
We have two underlying distributions for the green and red photons:

- One stemming from a peptide only having a donor dye.
- One stemming from a peptide being properly tagged with a donor and an acceptor dye.

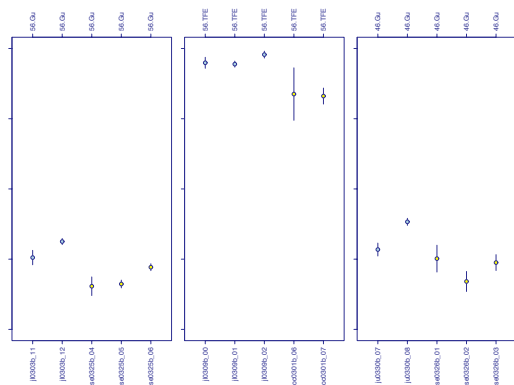
Deviations from Random Coil Behavior



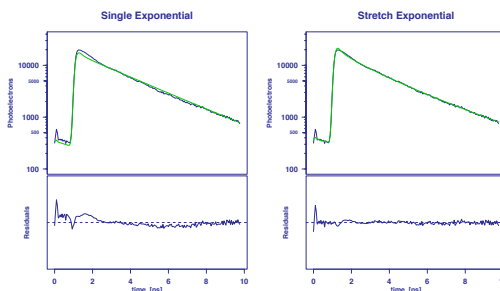
Deviations from Random Coil Behavior



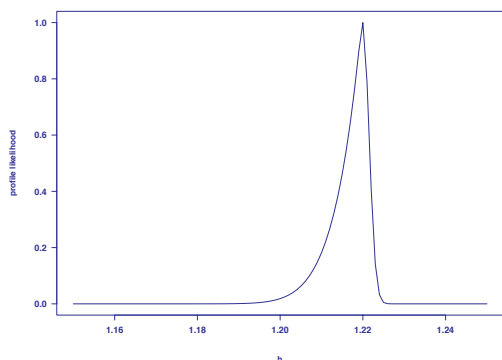
Replication



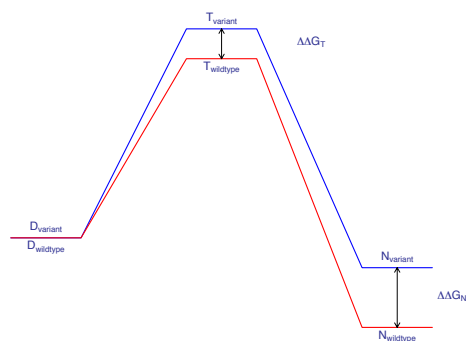
Photon Lifetime



Photon Lifetime

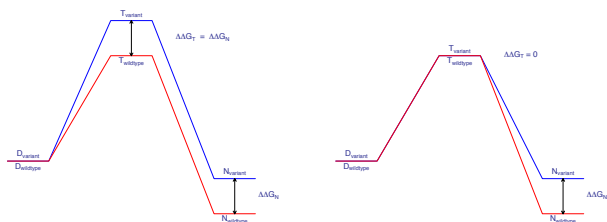


Energy Profile



→ The Φ -value is defined as the ratio $\Delta\Delta G_T / \Delta\Delta G_N$.

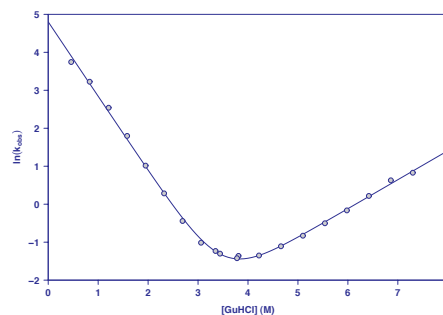
Energy Profile



- If the part of the protein that contains the variant amino acid is fully structured in the transition state, we have $\Delta\Delta G_T \approx \Delta\Delta G_N$, and hence $\Phi \approx 1$.
- If the part of the protein that contains the variant amino acid is equal in denatured and the transition state, we have $\Delta\Delta G_T \approx 0$, and hence $\Phi \approx 0$.

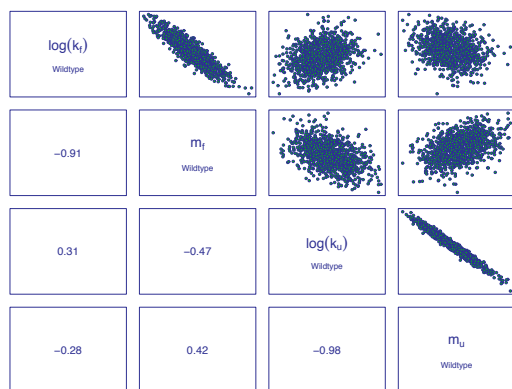
At least this is the idea ...

Standard Error

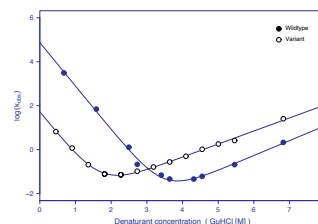


$$\ln(k_{\text{obs}}) = \log \left(\exp \left[\log(k_f) + m_f \times \frac{[\text{GuHCl}]}{RT} \right] + \exp \left[\log(k_u) + m_u \times \frac{[\text{GuHCl}]}{RT} \right] \right)$$

Standard Errors



Phi-Value Estimation



$$\log(k_{\text{obs}}) = \log \left(\exp \left[\log(k_f) + m_f \times \frac{C_{\text{den}}}{RT} \right] + \exp \left[\log(k_u) + m_u \times \frac{C_{\text{den}}}{RT} \right] \right)$$

$$\Delta\Delta G_T = RT \times \left[\log(k_f^{\text{wildtype}}) - \log(k_f^{\text{variant}}) \right]$$

$$\Delta\Delta G_N = RT \times \left[\log(k_u^{\text{wildtype}}) - \log(k_u^{\text{wildtype}}) - \log(k_u^{\text{variant}}) + \log(k_u^{\text{variant}}) \right]$$

Standard Error

$$se(\hat{\Phi}) = |\Phi| \times \sqrt{\left(\frac{\sigma_T}{\Delta\Delta G_T}\right)^2 - 2\rho_{\Delta\Delta G} \left(\frac{\sigma_T}{\Delta\Delta G_T}\right) \left(\frac{\sigma_N}{\Delta\Delta G_N}\right) + \left(\frac{\sigma_N}{\Delta\Delta G_N}\right)^2}$$

Web Server

Ruczinski, Sonnick, and Plaxco

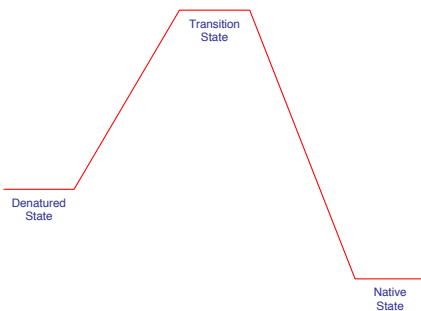
Methods for the accurate estimation of confidence intervals of experimental values

Please read important the format information for uploading your data. For options and output information, please read the help file.

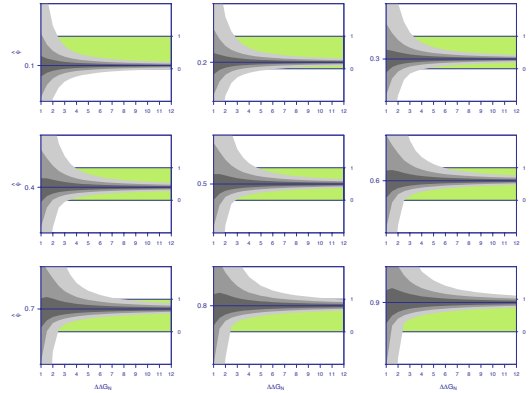
1. Upload your data: example.csv
2. Specify the temperature: T = 35 °C
3. Specify the energy units: kcal/mol
4. Specify the comparisons: wildtype only all comparisons
5. Specify the type of fit: individual fits parallel arms
6. Specify the denaturant concentrations: isotopy rate: M - unfolding rate: M
7. Specify the coverage for the 0-value confidence intervals: % coverage
8. Specify the number of significant digits in the output: significant digits
9. And they are off:

Mail Inqo

Energy Profile



Standard Error



Web Server

Ruczinski, Sonnick, and Plaxco

Methods for the accurate estimation of confidence intervals of experimental values

Please read important the format information for uploading your data. For options and output information, please read the help file.

Results

Table 1: Estimates for the kinetic parameters. The energies are given in kJ.

| | logN | M _{app} | n | M ₀ | logN ₀ | M _{app0} | n ₀ | M ₀₀ | ΔΔG _T | M _{appΔG_T} | n _{ΔG_T} | M _{0ΔG_T} |
|------|------|------------------|-------|----------------|-------------------|-------------------|----------------|-----------------|------------------|--------------------------------|-----------------------------|------------------------------|
| WT | 4.80 | 0.06 | -4.88 | 0.08 | -4.70 | 0.13 | 1.89 | 0.06 | 23.56 | 0.20 | -6.75 | 0.87 |
| EMA | 1.75 | 0.06 | -4.97 | 0.16 | -2.87 | 0.08 | 1.45 | 0.04 | 10.87 | 0.12 | -4.42 | 0.15 |
| DEL | 4.35 | 0.05 | -5.05 | 0.08 | -4.27 | 0.10 | 1.88 | 0.04 | 21.10 | 0.15 | -6.74 | 0.87 |
| DRV | 4.10 | 0.06 | -4.68 | 0.06 | -4.34 | 0.12 | 1.75 | 0.06 | 20.84 | 0.10 | -6.61 | 0.84 |
| VSA | 4.41 | 0.04 | -4.28 | 0.10 | -3.81 | 0.08 | 1.75 | 0.04 | 12.84 | 0.12 | -6.01 | 0.89 |
| VSSM | 4.84 | 0.08 | -4.80 | 0.12 | -3.22 | 0.15 | 1.77 | 0.06 | 19.99 | 0.22 | -6.57 | 0.10 |
| VST | 4.18 | 0.06 | -4.33 | 0.14 | -3.78 | 0.10 | 1.80 | 0.05 | 12.30 | 0.16 | -6.13 | 0.13 |
| VSG | 4.09 | 0.10 | -4.48 | 0.20 | -3.16 | 0.18 | 1.41 | 0.14 | 8.47 | 0.27 | -5.86 | 0.42 |

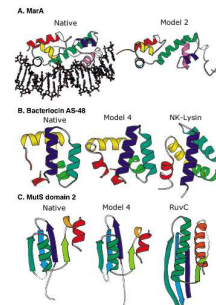
A spreadsheet with the above data for download is here. The file is in csv format, which can directly be used, for example, with R and Excel.

Table 2: Estimates for the 0 values and the changes in free energy. The energies are given in kJ.

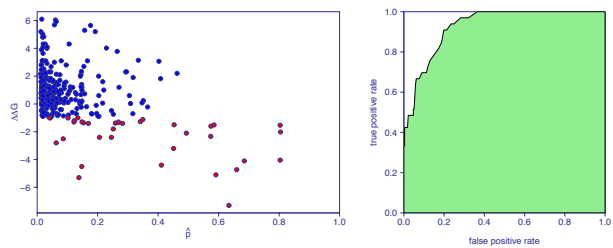
| | M ₀ | lower | upper | ΔΔG _T | M _{app0} | ΔΔG ₀ | M _{appΔG₀} | P _{0/native} |
|---------|----------------|-------|-------|------------------|-------------------|------------------|--------------------------------|-----------------------|
| WT-EMA | 0.00 | 0.02 | 0.86 | 0.84 | 1.56 | 0.30 | 10.88 | 0.37 |
| WT-DEL | 0.56 | 0.11 | 0.33 | 0.80 | 1.36 | 0.20 | 2.43 | 0.40 |
| WT-DRV | 0.66 | 0.13 | 0.40 | 0.80 | 1.74 | 0.21 | 2.63 | 0.43 |
| WT-VSA | 0.09 | 0.02 | 0.26 | 0.12 | 0.86 | 0.18 | 10.62 | 0.37 |
| WT-VSSM | -0.03 | 0.02 | -0.16 | 0.11 | -0.10 | 0.24 | 3.07 | 0.47 |
| WT-VST | 0.44 | 0.02 | 0.18 | 0.12 | 1.86 | 0.20 | 11.26 | 0.40 |
| WT-VSG | 0.05 | 0.02 | 0.09 | 1.02 | 0.33 | 18.09 | 0.52 | 0.20 |

A spreadsheet with the above data for download is here. The file is also csv format.

Structure Prediction / Functional Annotation



Stability / Affinity



Acknowledgments

- UC Santa Barbara Department of Chemistry Biochemistry
Miguel de los Rios, Kevin Plaxco.
- University of Chicago Department of Chemistry
Tobin Sosnick.
- Rice University Department of Biochemistry and Cell Biology
BK Muralidhara, Pernilla Wittung-Stafshede.
- UC Berkeley Department of Molecular and Cell Biology
David Wildes, Susan Marqusee.
- University of Washington Department of Biochemistry
David Baker & Lab.
- New York University Department of Biology and Department of Computer Science
Richard "Chill" Bonneau.
- University of North Carolina Department of Biochemistry and Biophysics
Deanne Sammond, Brian Kuhlman.

<http://biostat.jhsph.edu/~iruczins>