Transition States in Protein Folding

Thomas Weikl

Max Planck Institute of Colloids and Interfaces
Department of Theory and Bio-Systems
Overview

- Mutational $\Phi$-value analysis of the folding kinetics
- Modeling $\Phi$-values for $\alpha$-helices
- Modeling $\Phi$-values for small $\beta$-sheet proteins
Protein folding problems

- **The structure problem**: In which native structure does a given sequence fold?

- **The kinetics problem**: How does a protein fold into its structure?
How does a protein fold?

- The **Levinthal paradox**: How does a protein find its folded conformation as "needle in the haystack"?

- The **"old view"**: Metastable folding intermediates guide a protein into its native structure

- The **"new view"**: Many small proteins fold without detectable intermediates (2-state proteins)
2-state folding: Single molecules

- Donor and acceptor dyes at chain ends
- State-dependent transfer efficiency

Schuler et al., Nature 2002
2-state folding: Protein ensemble

- **rapid mixing** to initiate folding

- **single-exponential relaxation** for 2-state process:
  
  denatured \( \leftrightarrow \) native
  
  state \( D \) \( \leftrightarrow \) state \( N \)
Mutational analysis of 2-state folding

- **Transition state** theory:
  
  \[ k \propto \exp(-G_{T-D}) \]

- **Mutations change** the folding rate \( k \) and **stability** \( G_{N-D} \)

- **Central quantities:** \( \Phi \)-**values**
  
  \[ \Phi \equiv \frac{\Delta G_{T-D}}{\Delta G_{N-D}} \]
Traditional interpretation of $\Phi$

$\Phi = 1$: mutated residue is native-like structured in $T$

$\Phi = 0$: mutated residue is unstructured in $T$
Traditional interpretation of $\Phi$

• $\Phi$: degree of structure formation of a residue in T

• Inconsistencies:
  - some $\Phi$'s are $<0$ or $>1$
  - different mutations of the same residue can have different $\Phi$-values

Goldenberg, NSB 1999
**Example: α-helix of Cl2**

<table>
<thead>
<tr>
<th>Mutation</th>
<th>$\Phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S12G</td>
<td>0.29</td>
</tr>
<tr>
<td>S12A</td>
<td>0.43</td>
</tr>
<tr>
<td>E15D</td>
<td>0.22</td>
</tr>
<tr>
<td>E15N</td>
<td>0.53</td>
</tr>
<tr>
<td>A16G</td>
<td>1.06</td>
</tr>
<tr>
<td>K17G</td>
<td>0.38</td>
</tr>
<tr>
<td>K18G</td>
<td>0.70</td>
</tr>
<tr>
<td>I20V</td>
<td>0.40</td>
</tr>
<tr>
<td>L21A</td>
<td>0.25</td>
</tr>
<tr>
<td>L21G</td>
<td>0.35</td>
</tr>
<tr>
<td>D23A</td>
<td>-0.25</td>
</tr>
<tr>
<td>K24G</td>
<td>0.10</td>
</tr>
</tbody>
</table>

- $\Phi$-values for mutations in the helix range from -0.25 to 1.06

- **Our finding:**

Itzhaki, Otzen, Fersht, 1995
- Mutational $\Phi$-value analysis of the folding kinetics
- **Modeling $\Phi$-values for $\alpha$-helices**
- Modeling $\Phi$-values for small $\beta$-sheet proteins
Helix cooperativity

- we assume that a helix is either fully formed or not formed in transition-state conformation $T_i$

- we have two structural parameters per helix:
  - the degree of secondary structure $\chi_\alpha$ in $T$
  - the degree of tertiary structure $\chi_t$ in $T$
Splitting up free energies

- we split up mutation-induced free energy changes into **secondary** and **tertiary components**:

\[ \Delta G_N = \Delta G_\alpha + \Delta G_t \]

\[ \Delta G_T = \chi_\alpha \Delta G_\alpha + \chi_t \Delta G_t \]

- general form of \( \Phi \)-values for mutations in an \( \alpha \)-helix:

\[
\Phi \equiv \frac{\Delta G_T}{\Delta G_N} = \chi_t + \left( \chi_\alpha - \chi_t \right) \frac{\Delta G_\alpha}{\Delta G_N}
\]
Φ-values for α-helix of Cl2

general formula:  \[ \Phi = \chi_t + (\chi_\alpha - \chi_t) \frac{\Delta G_\alpha}{\Delta G_N} \]

mutational data for Cl2 helix:
Φ-values for helix 2 of protein A

general formula: \[ \Phi = \chi_t + (\chi_\alpha - \chi_t) \frac{\Delta G_\alpha}{\Delta G_N} \]

mutational data for helix 2:
Summary

Consistent interpretation of $\Phi$-values for helices:

- with **two structural parameters**: the degrees of secondary and tertiary structure formation in $T$
- by **splitting up** mutation-induced **free energy changes** into secondary and tertiary components

C Merlo, KA Dill, TR Weikl, PNAS 2005
TR Weikl, KA Dill, JMB 2007
• Mutational $\Phi$-value analysis of the folding kinetics
• Modeling $\Phi$-values for $\alpha$-helices
• Modeling $\Phi$-values for small $\beta$-sheet proteins
WW domains are 3-stranded \( \beta \)-proteins with two \( \beta \)-hairpins

- we assume that each hairpin is fully formed or not formed in the transition state
Evidence for hairpin cooperativity

- $\beta 3s$ is a designed 3-stranded $\beta$-protein with 20 residues
- transition state rigorously determined from folding-unfolding MD simulations
- result: either hairpin 1 or hairpin 2 structured in T

Rao, Settanni, Guarnera, Caflisch, JCP 2005
A simple model for WW domains

- We have two transition-state conformations with a single hairpin formed.

- The folding rate is:

  \[ k \approx \frac{1}{2} \left( e^{-G_1/RT} + e^{-G_2/RT} \right) \]

- \( \Phi \)-values have the general form:

\[
\Phi \equiv \frac{-RT \Delta \log k}{\Delta G_N} = \frac{\chi_1 \Delta G_1 + \chi_2 \Delta G_2}{\Delta G_N}
\]
Φ-values for FBP WW domain

- general formula: \( Φ_{\text{theo}} = \frac{χ_1ΔG_1 + χ_2ΔG_2}{ΔG_N} \)
- a first test: Φ’s for mutations affecting only hairpin 1 should have value \( χ_1 \)
Φ-values for FBP WW domain

- general formula: \( \Phi_{\text{theo}} = \frac{\chi_1 \Delta G_1 + \chi_2 \Delta G_2}{\Delta G_N} \)

- single-parameter fit:

\[
\begin{align*}
\chi_1 &\approx 0.77 \\
\chi_2 &= 1 - \chi_1 \approx 0.23
\end{align*}
\]
Summary

Reconstruction of transition states from mutational $\Phi$-values based on:

- substructural cooperativity of helices and hairpins
- splitting up mutation-induced free energy changes

C Merlo, KA Dill, TR Weikl, PNAS 2005
TR Weikl, KA Dill, J Mol Biol 2007
TR Weikl, Biophys J 2008