Physics of Life

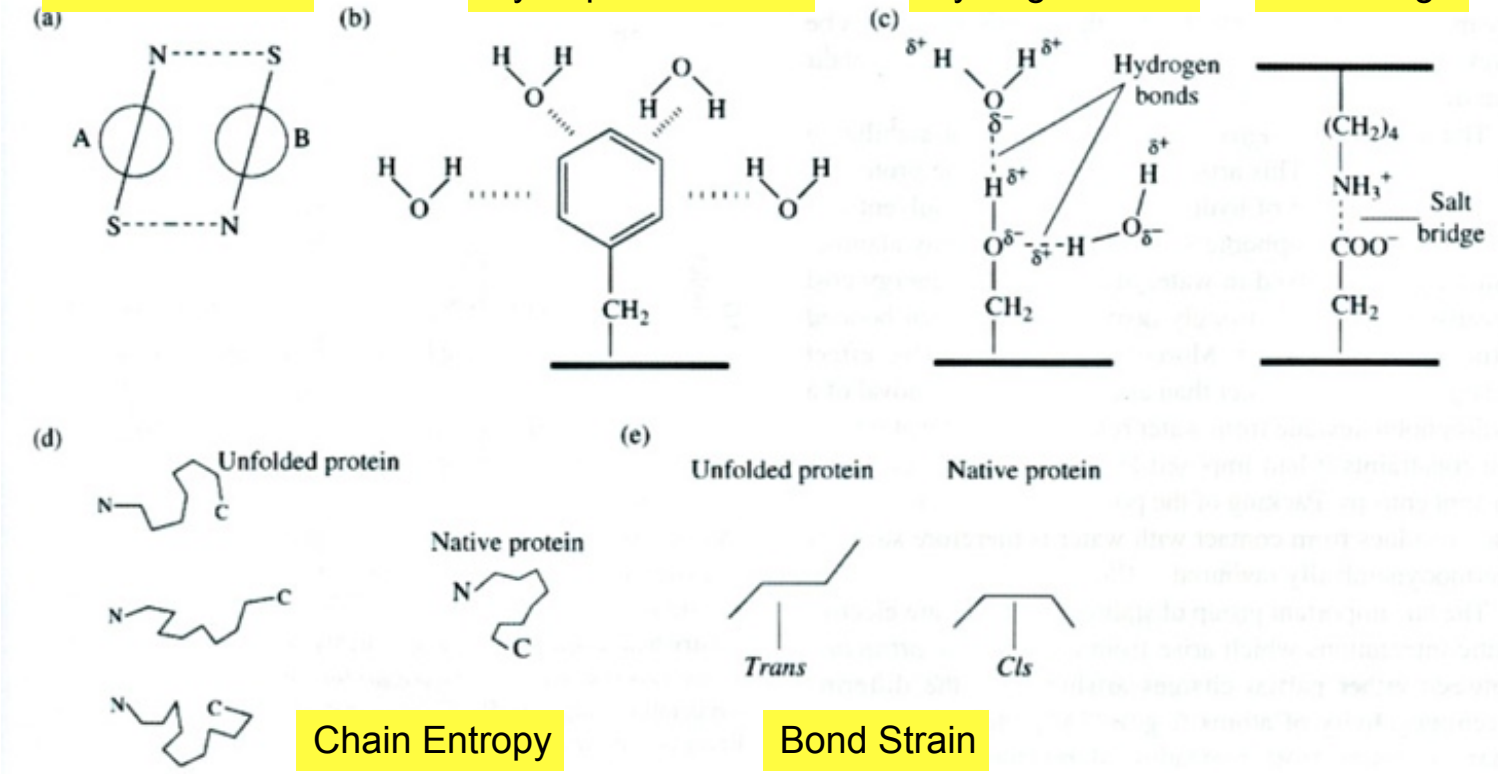
PHYS-468

Protein Folding

Henning Stahlberg,
LBEM, IPHYS, SB, EPFL

Forces involved in protein folding

Van der Waals | Hydrophobic Effect | Hydrogen Bond | Salt Bridge



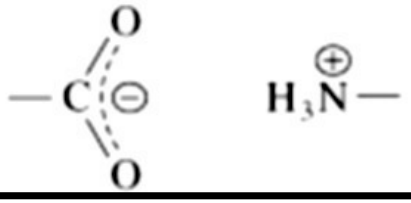
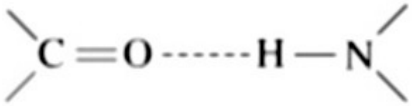
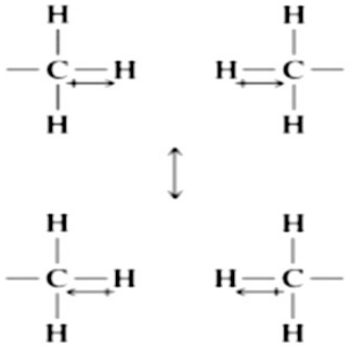
Chain Entropy | Bond Strain

Figure 6.1. The main forces involved in protein folding. Stabilizing forces include (a) van der Waals interactions. These are short-range attractive forces between any pair of atoms due to transient magnetic dipoles which they induce in each other. North (N) and south (S) poles of the dipoles are shown. If two atoms come too close together, repulsion between electron clouds can overcome this attraction. Each atom has a van der Waals radius and the sum of these radii for a pair of atoms is the van der Waals contact distance. (b) The hydrophobic effect. Hydrophobic amino acids such as phenylalanine 'prefer' to be buried in the interior of a protein rather than be exposed to solvent water. Repulsion of water by phenylalanine is shown (''''). (c) Electrostatic interactions. Hydrogen bonds stabilize interactions between polar side-chains and solvent water. Salt bridges arise from electrostatic attraction between side-chains of opposite charge. The main destabilizing forces are (d) Chain entropy. Many more conformations are available for the protein in the unfolded state than in the folded state. (e) Bond strain due to unfavourable interactions in folded state (e.g. constraint into a *cis* rather than the more favoured *trans* configuration).

2

Interactions in biomolecules

1 eV = 23.06 kcal/mol = 96.48 kJ/mol
0.0136 eV = 0.239 kcal/mol = 1 kJ/mol

		Approx. Value
	• Charge-charge interactions (Salt bridge) ~40 … 200 kJ/mol = 0.4 … 2.1 eV	1 eV
	• Hydrogen bonds ~2 … 20 kJ/mol = 0.021 … 0.21 eV	0.1 eV
	• Van der Waals interactions (London dispersion force) ~0.4 … 4 kJ/mol = 0.004 … 0.04 eV	0.01 eV
	• Hydrophobic interactions ~3 … 10 kJ/mol = 0.03 … 0.1 eV	0.05 eV
	• Covalent Bond ~200 … 1000 kJ/mol = 2.1 … 10 eV	5 eV

Forces involved in protein folding

Table 6.1. Estimates of net magnitudes of stabilizing and destabilizing forces in protein folding.

Interaction (protein of 100 residues)	ΔG at 25 °C (kcal/mole)
Destabilizing	
Chain entropy	330–1000
Unfavourable folding interactions	200
Stabilizing	
Increased van der Waals bonds due to close packing	−227
Hydrogen bonds	−49 to −719
Hydrophobic effect	−264
Disulfide bridges	−4

Note that stabilizing forces contribute a negative ΔG while destabilizing force ΔG values are positive.

Chain Entropy

Bond Strain

Van der Waals

Hydrogen Bond

Salt Bridge

Hydrophobic Effect

C=C bridges

4

Thermal Melting of Proteins

Standard State:

$$G^0 = H^0 - TS^0$$

$$\Delta G^0 = \Delta H^0 - T\Delta S^0$$

PMF := 2.3 RT ΔpH + F ΔΨ

R = Gas Constant = 8.315 J/mol/K

T = Temperature (= 300 Kelvins at room temperature)

F = Faraday Constant = 96.5 kJ/V/mol

ΔΨ = electrical membrane potential = 200 mV

H^o = Enthalpy = $U + pV$ = Thermodynamic Potential

U = inner energy; pV = Pressure * Volume

S^o = Entropy = $k_B \cdot \ln \Omega$

A solute *A* in a pure solution:

$$G = G^0 + RT\ln([A])$$

$[A]$ is the concentration of *A*.

Here, it is $[A]/(1\text{M})$, meaning it is in molar units

Chemical reaction from *R* to product *P* with multiple states R_i and P_i:

$$\Delta G = \Delta G^0 + RT\ln\left(\frac{\Pi[P_i]}{\Pi[R_i]}\right)$$

R = Gas Constant = 8.315 J/mol/K

T = Temperature (= 300 Kelvins at room temperature)

$$\Delta G = -30.5\frac{kJ}{mol} + RT\ln\left(\frac{[ADP][P_i]}{[ATP]}\right)$$

$$\Delta G = -30.5\frac{kJ}{mol} + RT\ln\left(\frac{0.25\cdot10^{-3}\cdot2.0\cdot10^{-3}}{2.5\cdot10^{-3}}\right)$$

$$\Delta G = -30.5\frac{kJ}{mol} - RT\ln(0.0002)$$

$$\Delta G = -30.5\frac{kJ}{mol} - 8.315\frac{J}{mol\cdot K}\cdot 300K\cdot(-8.5)$$

$$\Delta G = -30.5\frac{kJ}{mol} - 21.2\frac{kJ}{mol} = -51.7\frac{kJ}{mol}$$

$[ADP] = 0.25mM$

$[P_i] = 2.0mM$

$[ATP] = 2.5mM$

In calculations, pay attention to using only SI units, i.e., in this case "M" for Molar. Do not calculate with "mM".

ATP-Energy = ΔG(ATP → ADP+P_i) = -51.7 kJ/mol

5

Thermal Melting of Proteins

Chemical reaction from *R* to product *P* with multiple states R_i and P_i:

$$\Delta G = \Delta G^0 + RT \ln\left(\frac{\Pi[P_i]}{\Pi[R_i]}\right)$$

At equilibrium, there is no free energy G driving the system into either direction:

$$\Delta G = 0 \qquad \text{and} \qquad K = \frac{\Pi[P_i]}{\Pi[R_i]}$$

Therefore:

$$\Delta G^0 = -RT \ln(K)$$

$$\ln(K) = -\frac{\Delta G^0}{RT} = -\left(\frac{\Delta H^0 - T\Delta S^0}{RT}\right)$$

$$\ln(K) = \frac{1}{R}\left(\Delta S^0 - \frac{\Delta H^0}{T}\right)$$

van't Hoff relationship:
ln(K) is linear in 1/T

$$\frac{d\ln(K)}{d(1/T)} = \frac{-\Delta H^0}{R}$$

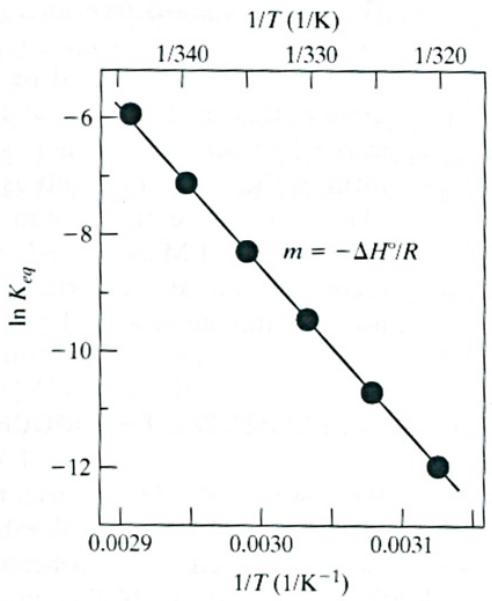


Figure 2.6 Van't Hoff plot for the thermal unfolding of the coiled coil domain of the GCN4 transcription factor. [Adapted from Holtzer et al. (2001), *Biophys. J.* **80**, 939–951.]

6

Thermal Melting of Proteins

$$\ln(K) = \frac{1}{R}\left(\Delta S^0 - \frac{\Delta H^0}{T}\right)$$

van't Hoff relationship:
ln(K) is linear in 1/T

$$\frac{d\ln(K)}{d(1/T)} = \frac{-\Delta H^0}{R}$$

$$\Delta H^0 = -R \cdot \frac{d\ln(K)}{d(1/T)}$$

The enthalpy H can be measured by measuring the equilibrium constant K at various temperatures T

$$\frac{d\ln(K)}{d(T)} = \frac{\Delta H^0}{RT^2}$$

$$\Delta H^0 = RT^2 \frac{d\ln(K)}{d(T)}$$

Another form of the van't Hoff relationship

7

Thermal Melting of Proteins

ΔG values of Folded-Unfolded state show stability of proteins at different temperatures.

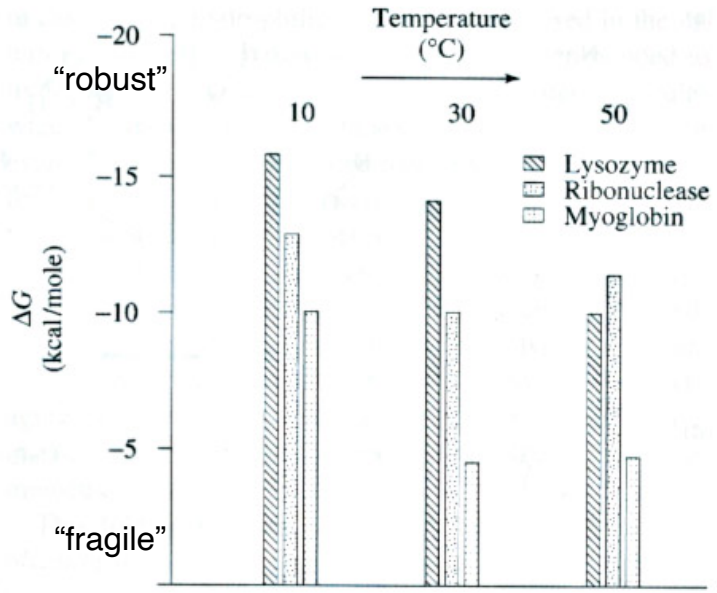


Figure 6.2. Folded proteins are highly-sensitive to temperature. ΔG values (i.e. $G_{folded} - G_{unfolded}$) for some small proteins in the temperature-range 10–50 °C. Temperature affects individual stabilizing/destabilizing interactions in a complex way which may result in the protein being more or less stable at higher temperature (e.g. compare ribonuclease and myoglobin at 30° and 50 °C). However, most proteins are unstable at high temperatures. Note the small magnitude of the energy difference which maintains the protein in a folded state.

Protein unfolding profile when temperature is increasing

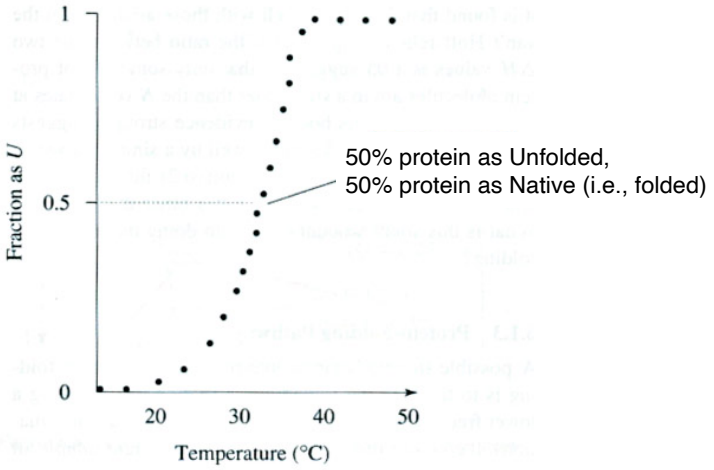


Figure 6.3. Protein denaturation profile of ribonuclease A. Temperature-induced denaturation was determined spectroscopically and is expressed as the fraction of unfolded (U) protein present. The situation when this fraction is 0.5 is described in the box.

$$K = \frac{[N]}{[U]}$$

van't Hoff relationship:

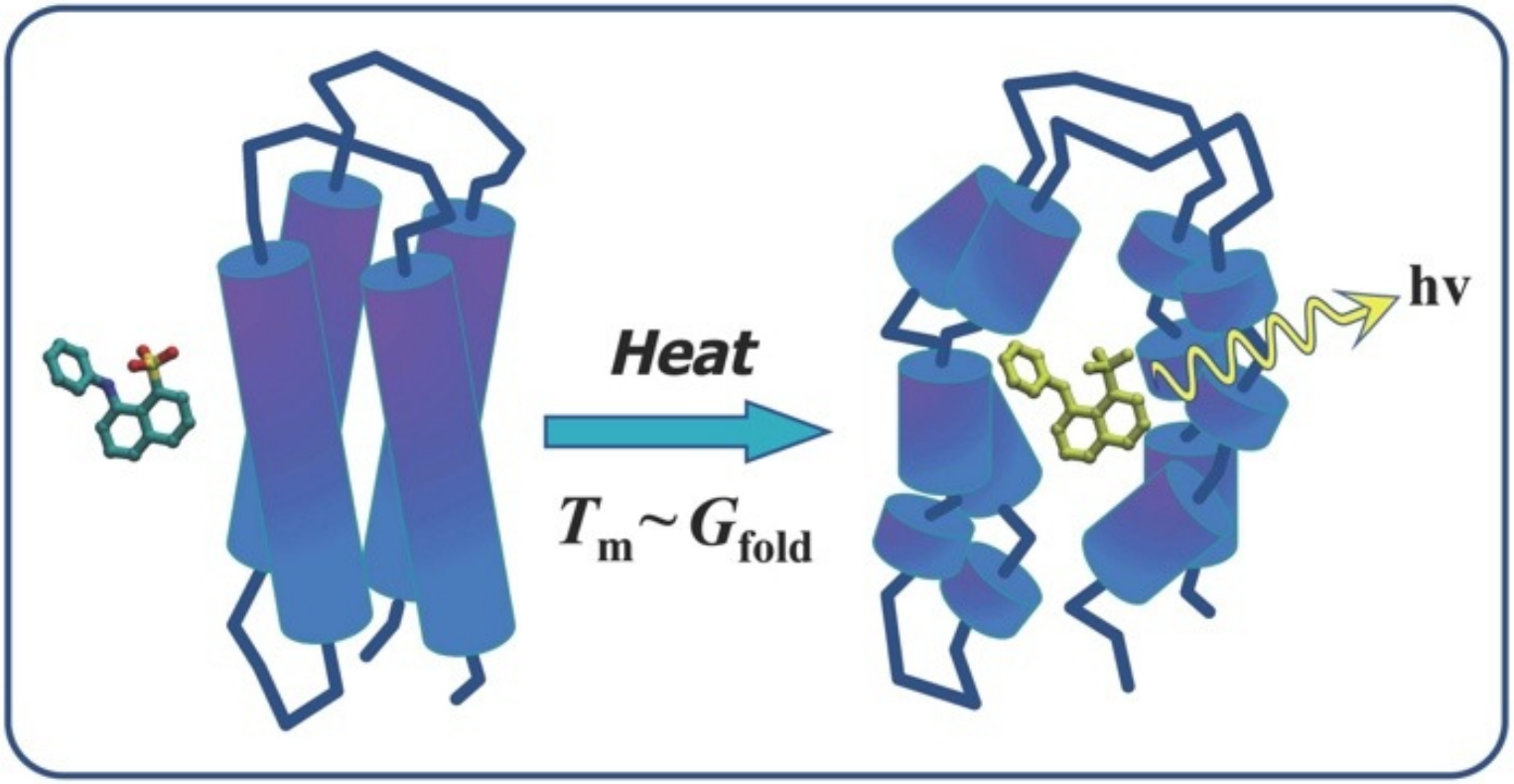
$$\Delta H = RT^2 \frac{d \ln K}{dT}$$

N = Native folded state
U = Unfolded state
R = Gas constant
T = Absolute temperature
H = *Enthalpy*

Compare:
$\Delta G^0 = \Delta H^0 - T\Delta S^0$

8

Thermofluor

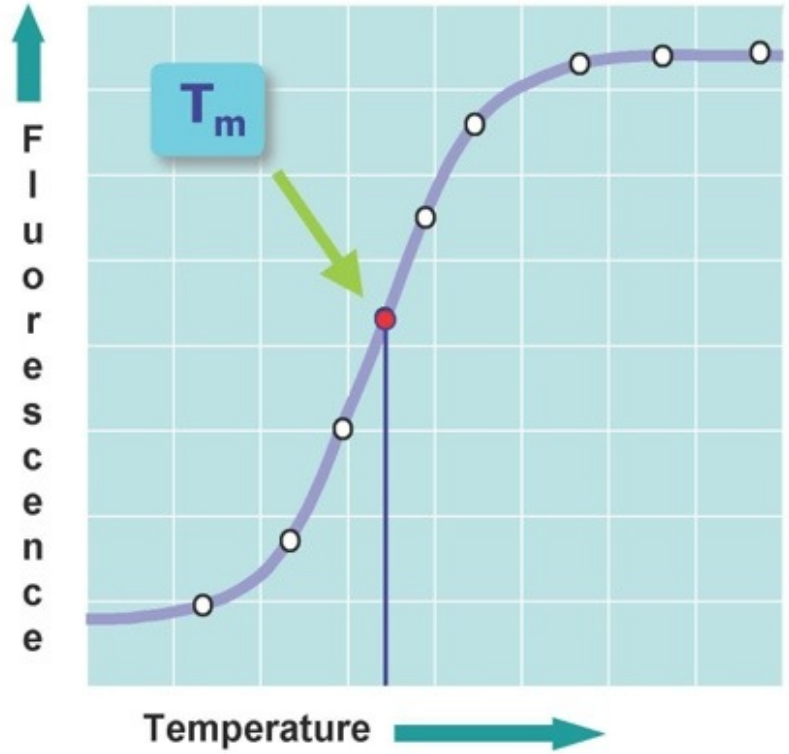


- Thermofluor gives an optical readout of protein melting as e.g. observed with DSC.
- Optical readout is much more sensitive than direct thermal measurement.
- TF consequently allows parallel measurements in 384 well-plates using fluorescent imaging plate readers.

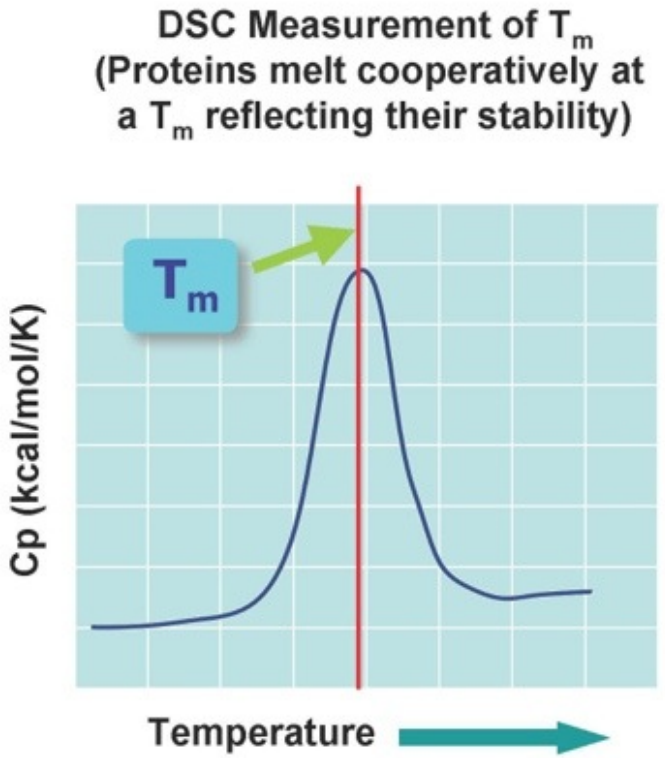
www.thermofluor.org

9

ThermoFluor & DSC



ThermoFluor Fluorescence Signal



Differential Scanning Calorimetry Signal

10

ThermoFluor Applications

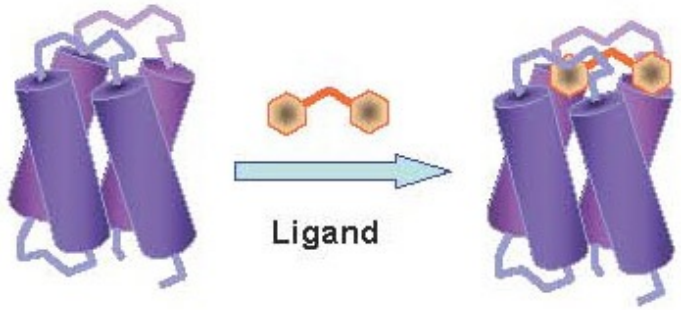
ThermoFluor can be used for:

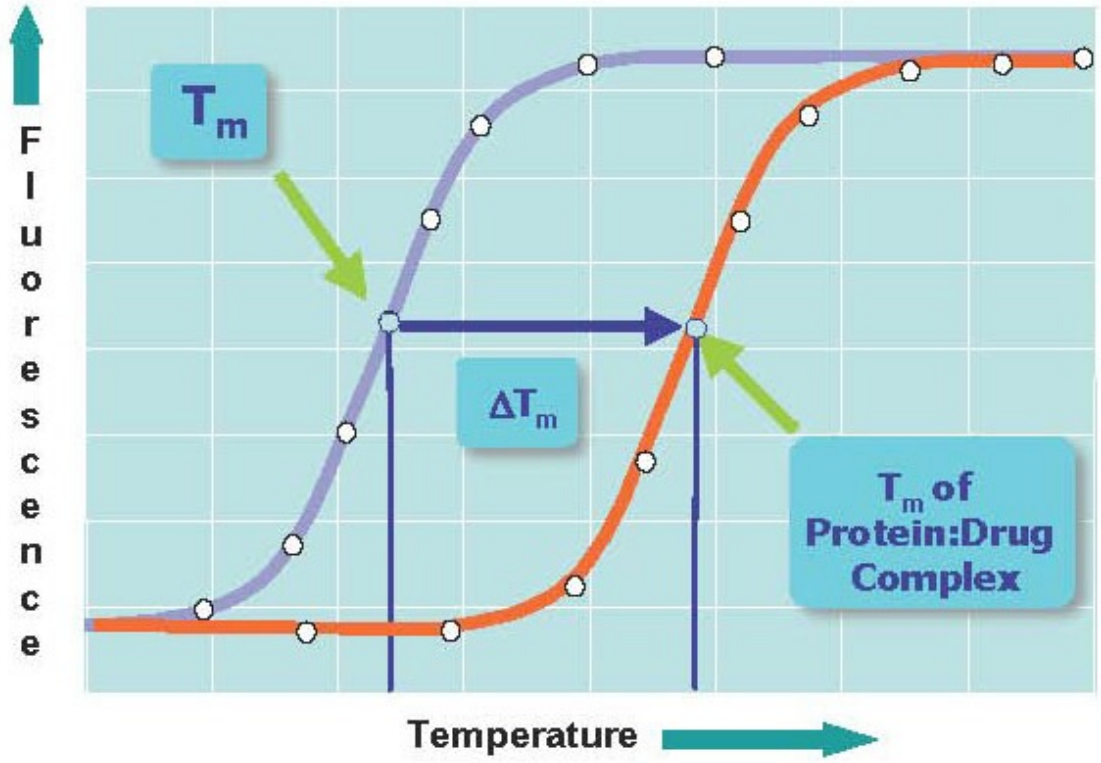
- Optimization of sample conditions (pH, type of buffer, salt concentration, cofactors, ligands, …) to improve sample stability, e.g., to increase the yield in protein purification in biochemistry, or for growing protein crystals for structure determination experiments.

- Measurement of ligand binding activity (if the ligand binding increases the temperature stability of the sample):
  - for drug screening in pharmaceutical drug development
  - to measure the binding cooperativity between multiple ligands (e.g., inhibitor plus substrate or cofactor, or agonist vs. antagonist
  - to characterize a protein of unknown function through binding experiments with libraries of “functional probes” (e.g., incorporating possible substrates, metabolic intermediates, cofactors, etc..)

11

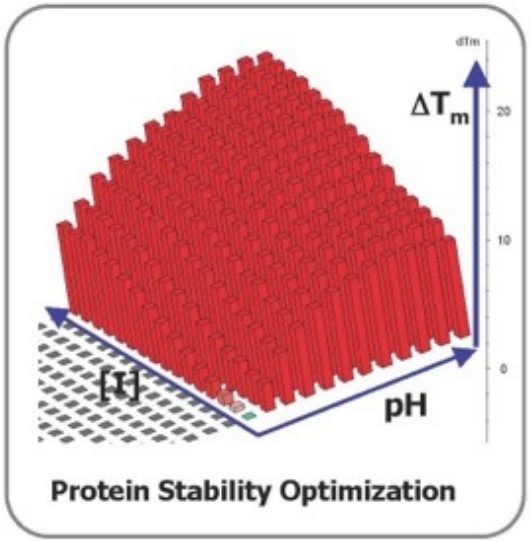
Ligand Binding Can Stabilize Protein Structure



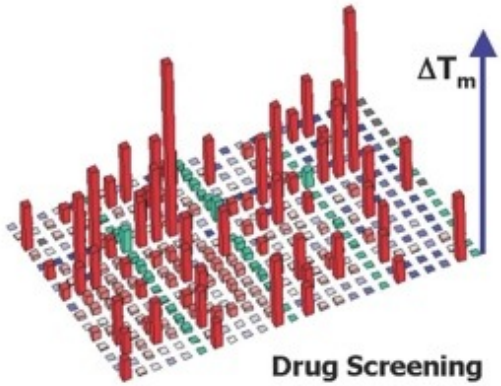


12

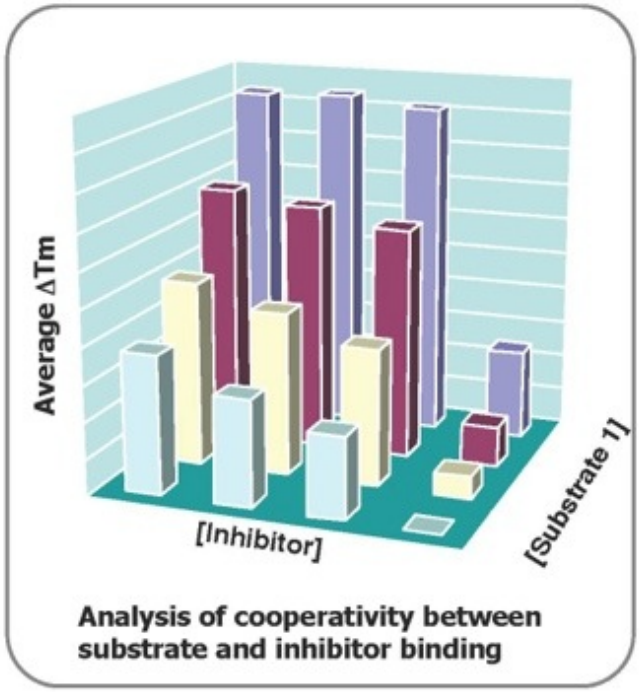
Thermofluor Applications



Protein Stability Optimization



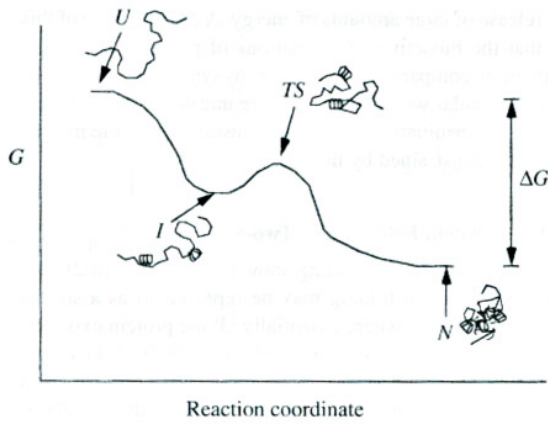
Drug Screening



Analysis of cooperativity between substrate and inhibitor binding

13

Folding of Proteins



U = Unfolded state

I = Intermediate state

TS = Transition state

N = Native folded state

Framework Model:

Hydrophobic Collapse:

Nucleation Model:

Nucleation-Condensation Model:

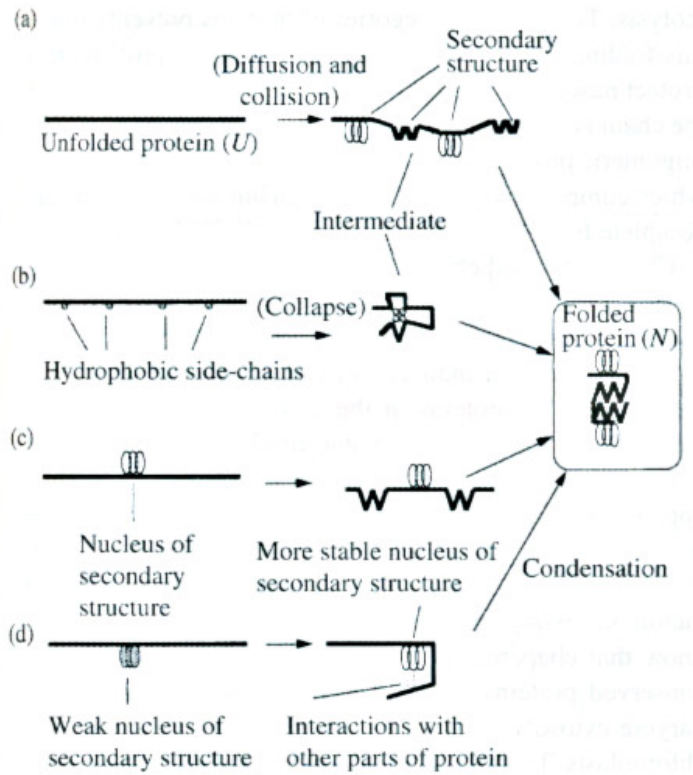
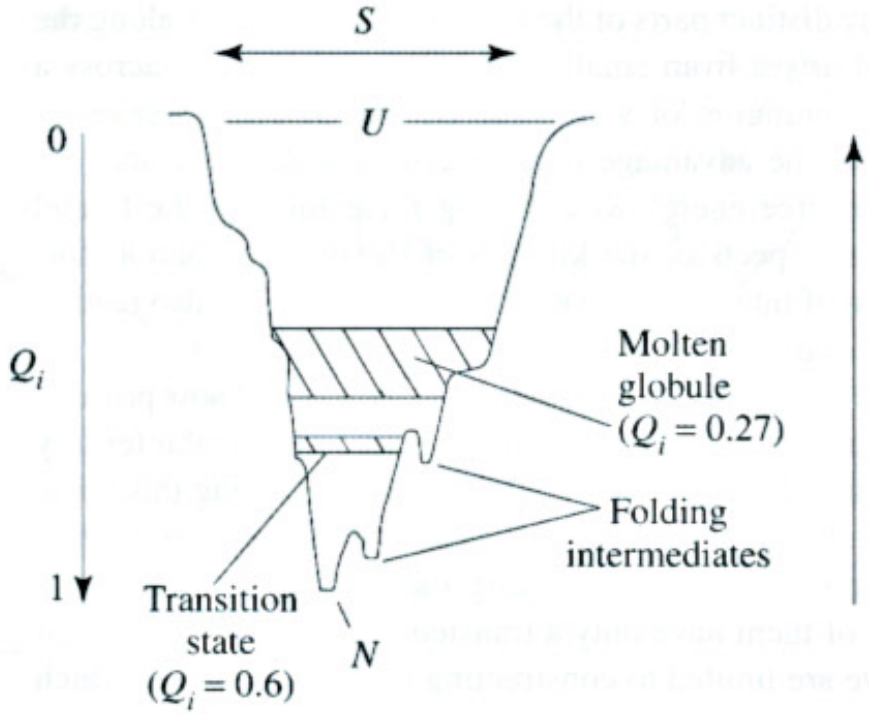


Figure 6.5. Models of protein folding. (a) Framework model; local elements of secondary structure form independently of tertiary structure. (b) Hydrophobic collapse model; protein chain rapidly collapses around hydrophobic side-chains. Secondary and tertiary structure form in restricted conformational space around this. (c) Nucleation model; secondary structure elements act as nuclei around which tertiary structure forms. (d) Nucleation-condensation model; A weak central nucleus is first formed followed by stabilization with sequentially distant residues leading to a larger and more stable nucleus. This condenses with the rest of the structure until folding is complete.

In General: Protein Folding happens along a path. It cannot happen by randomly testing all possible conformations. If protein folding would be a random trial-and-error process, then there would never be enough time for a protein to "test" all possible conformations.
Instead, some energy-driven process should lead to the folded state.

14

Folding of Proteins



$$Q_i = \frac{\#(\text{pairwise contacts in state } i)}{\#(\text{contacts in state } N)}$$

G = Gibbs free energy

Figure 6.6. A protein folding funnel. The funnel represents incremental changes in conformation as the protein folds from U to N. Entropy (S) decreases as the number of conformations available reduces. The parameter Q_j reflects gradual forming of the correct structural contacts in N. Experimentally identifiable species such as molten globules, folding intermediates and transition states are shown.

15

Chaperones

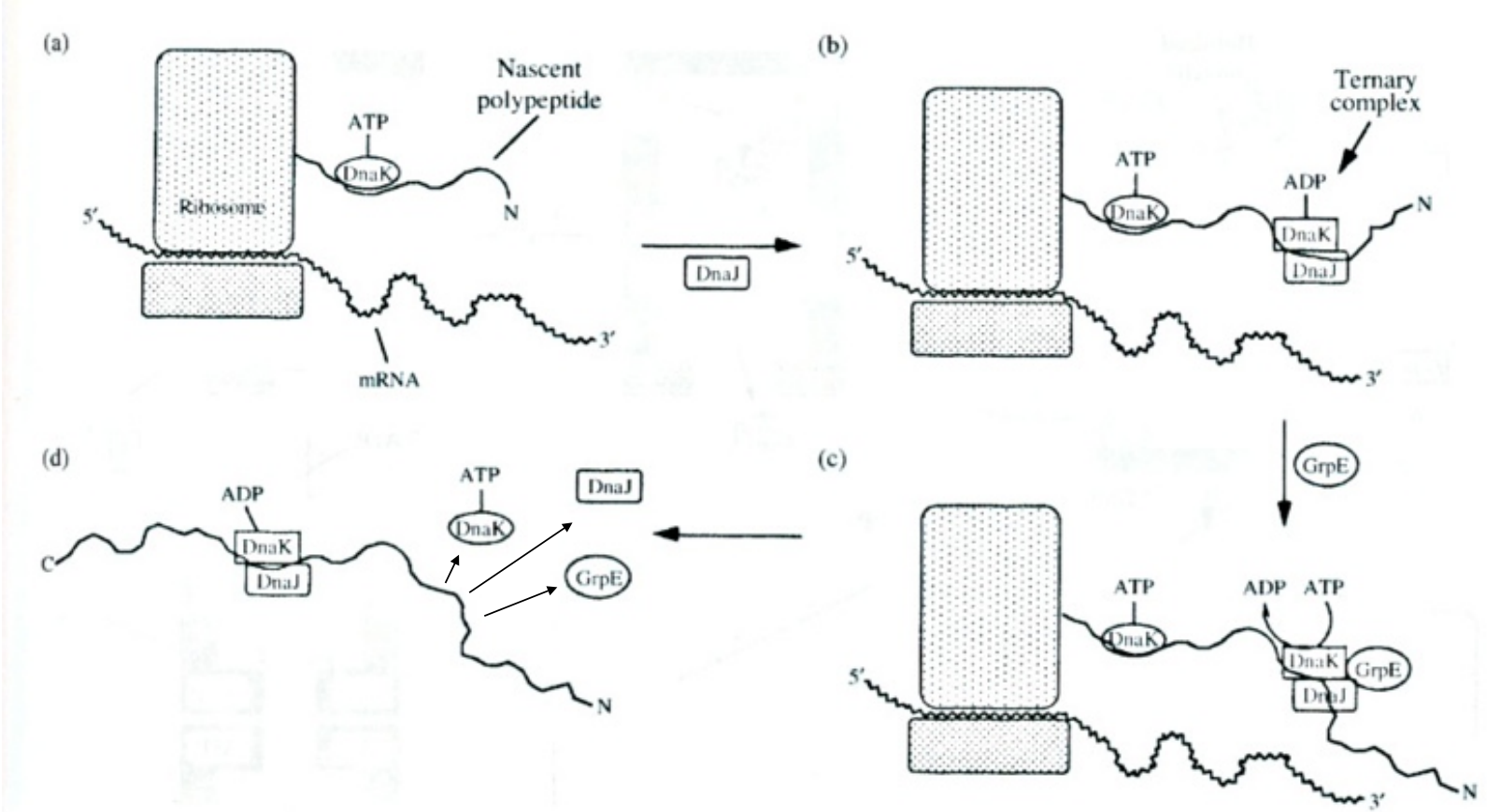


Figure 6.7. Functioning of Hsp 70 (DnaK) in folding of nascent polypeptides. (a) DnaK binds to hydrophobic stretches of nascent polypeptide chains during protein synthesis. (b) The co-chaperone, DnaJ, binds to DnaK, forming a ternary complex. This stimulates the ATP-ase activity of DnaK resulting in ATP being hydrolyzed to ADP. Note that the structure of DnaK-ADP differs from that of DnaK-ATP (c) A further protein, GrpE binds to the ternary complex, causing release of ADP and binding of ATP at DnaK. (d) The complex dissociates from the polypeptide. Note that several cycles of binding and dissociation can be gone through before peptide synthesis is complete. Moreover, several hydrophobic sites on a single polypeptide can act as substrates for the Hsp70 system.

GroEL

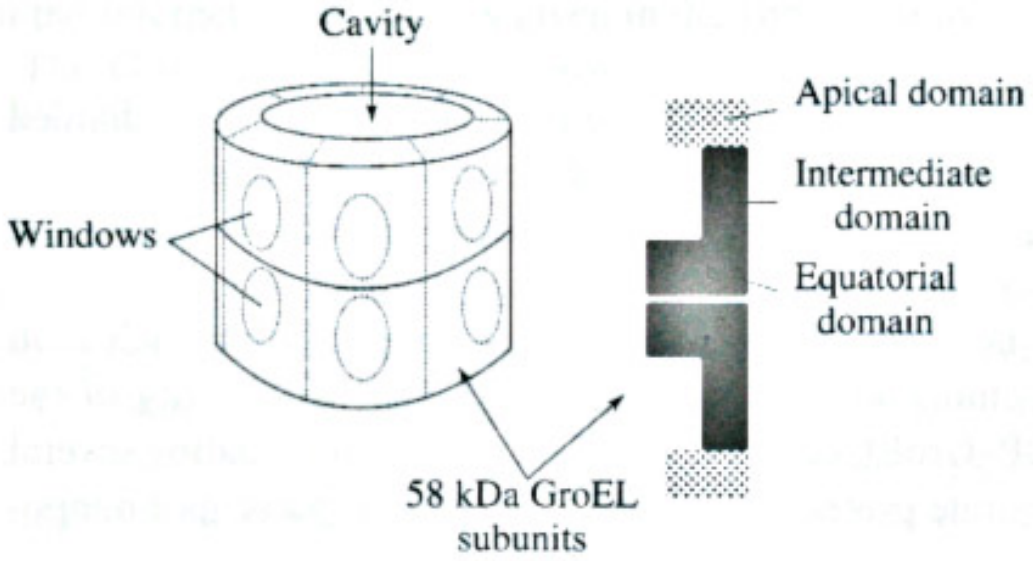
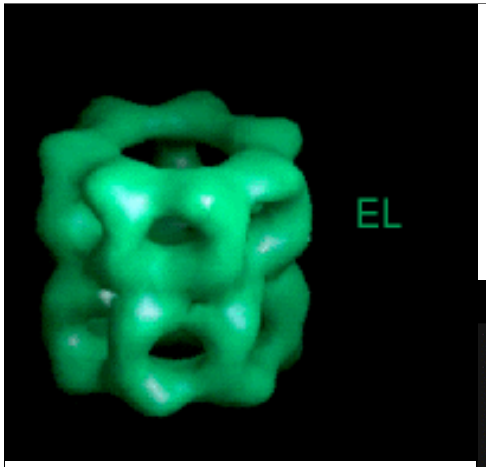
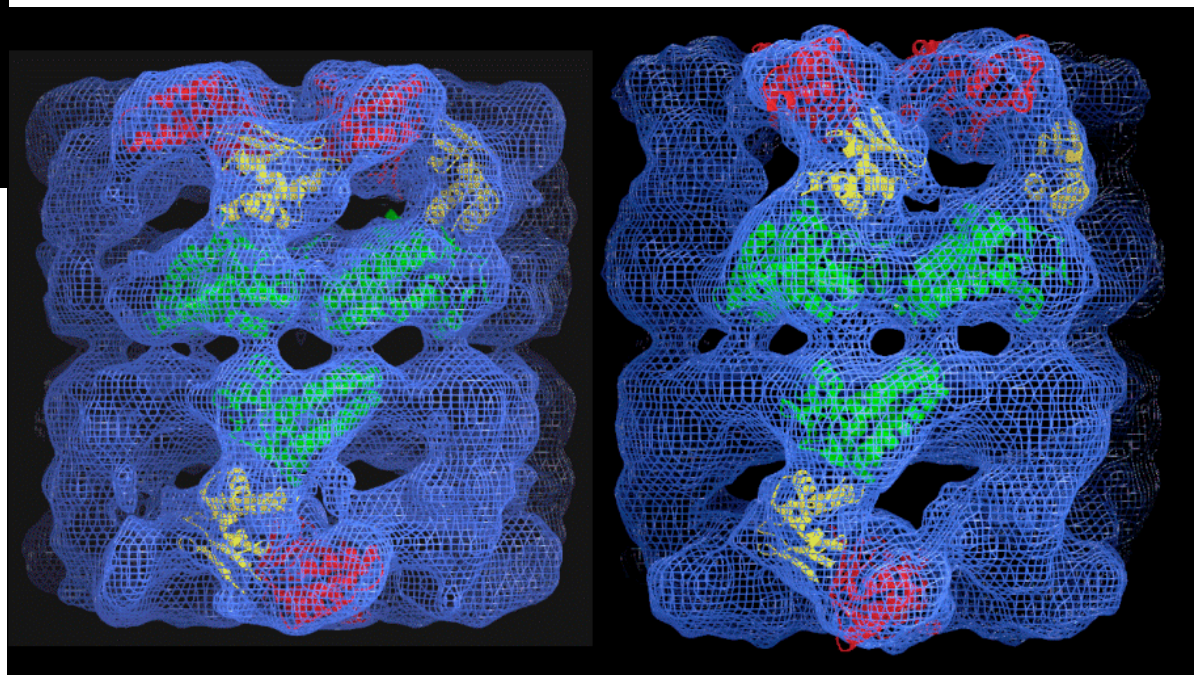


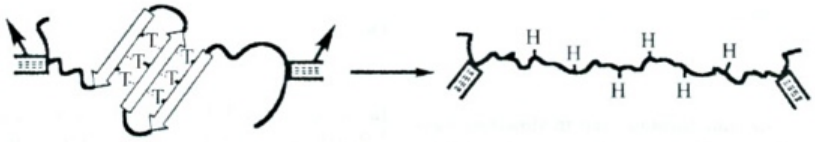
Figure 6.8. Organization of Hsp 60 chaperonin. Fourteen 58 kDa GroEL subunits are arranged in two circular rows of seven as shown. This forms a cavity within which folding can occur. Both the individual subunits and the complex they form are referred to as GroEl. Each subunit consists of three distinct structural domains; apical, intermediate and equatorial (which contain the ATP-ase site). Each equatorial domain also contains a 'window' sufficiently large for small molecules to diffuse freely through.

17

GroEL







Forced unfolding by elongation of the binding chamber upon binding of GroES

http://people.cryst.bbk.ac.uk/~ubcg16z/cpn/elmovies.html

18

Cell *149*, 113–123, March 30, 2012 ©2012 Elsevier Inc.

ATP-Triggered Conformational Changes Delineate Substrate-Binding and -Folding Mechanics of the GroEL Chaperonin

Daniel K. Clare,[1] Daven Vasishtan,[1] Scott Stagg,[2,6] Joel Quispe,[2] George W. Farr,[4,5,7] Maya Topf,[1] Arthur L. Horwich,[3,4,5] and Helen R. Saibil[1,*]

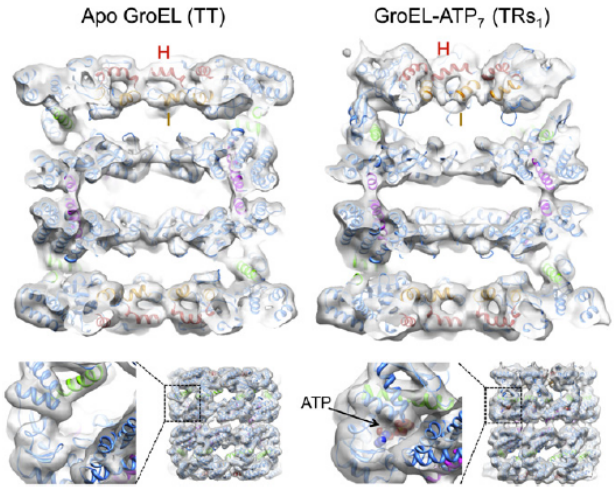


Figure 3. Cut-Open Views of the EM Maps and Fits of Apo GroEL and the GroEL-ATP$_7$ Complex with the Top Ring in the Rs1 State and the Bottom Ring in the T State

The EM density is in white, and apical domain helices H and I, defining the substrate-binding site, are in red and orange, respectively. Intermediate domain helix M (green) contains the catalytic aspartate that contacts the nucleotide, and equatorial domain helix D (magenta) runs from the γ-phosphate to an inter-ring contact. The fitting shows that α-helical secondary structures are largely resolved in these maps. Shown below each complex is a view of the region around the ATP-binding pocket, which is empty in the T state but filled with density in the Rs1 state. The ATP molecule inside the Rs1 density is shown as spheres with CPK coloring. See also Figure S3.

EMDB entries #1997 - #2003

http://www.ebi.ac.uk/pdbe/entry/EMD-1997

Display of GroEL with Chimera: Try this out:

http://www.cgl.ucsf.edu/chimera/data/tutorials/groel/groel.html

19

Cell *149*, 113–123, March 30, 2012 ©2012 Elsevier Inc.

ATP-Triggered Conformational Changes Delineate Substrate-Binding and -Folding Mechanics of the GroEL Chaperonin

Daniel K. Clare,[1] Daven Vasishtan,[1] Scott Stagg,[2,6] Joel Quispe,[2] George W. Farr,[4,5,7] Maya Topf,[1] Arthur L. Horwich,[3,4,5] and Helen R. Saibil[1,*]

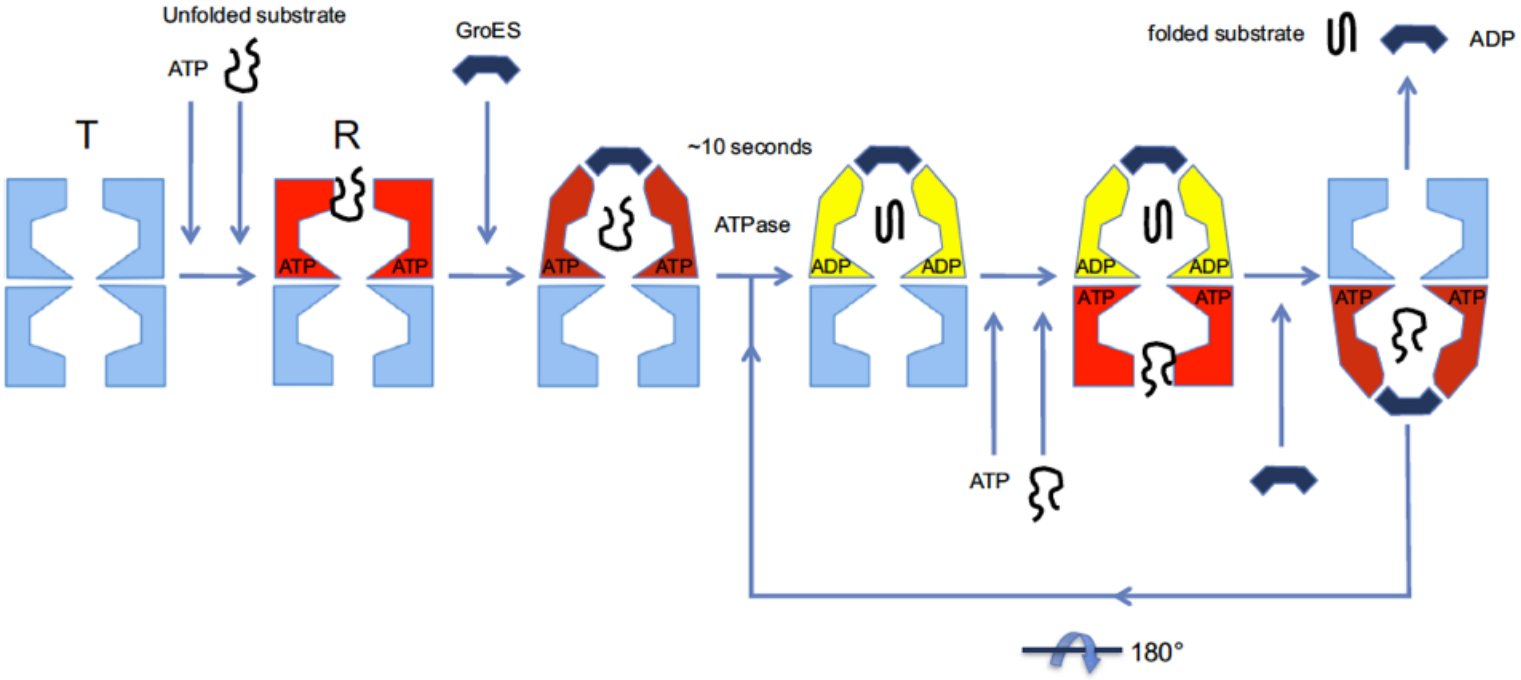


Figure 1. Diagram of the GroEL-GroES ATPase Cycle, Showing the Steps of Substrate Binding, Encapsulation, Folding, and Release
ATP hydrolysis in one ring is required to enable subsequent ATP binding to the opposite ring, but hydrolysis is not required for folding to proceed within the chamber.

20

Cell *149*, 113–123, March 30, 2012 ©2012 Elsevier Inc.

ATP-Triggered Conformational Changes Delineate Substrate-Binding and -Folding Mechanics of the GroEL Chaperonin

Daniel K. Clare,[1] Daven Vasishtan,[1] Scott Stagg,[2,6] Joel Quispe,[2] George W. Farr,[4,5,7] Maya Topf,[1] Arthur L. Horwich,[3,4,5] and Helen R. Saibil[1,*]

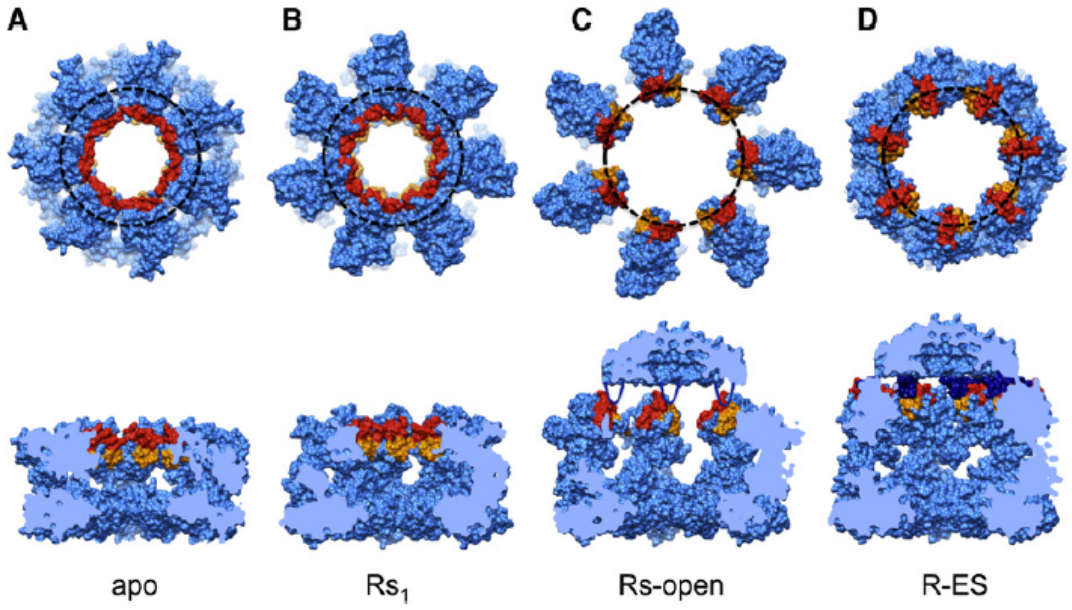


Figure 6. Footprint of GroES-Binding Sites on GroEL

The rings are seen in space-filling format from above and in cut-away side view in (A) apo, (B) Rs1, (C) Rs-open, and (D) R-ES (PDB ID 1SVT) rings. Helices H and I are in red and orange, respectively, and the binding residues on the mobile loops of GroES are shown in dark blue. The black dotted circle shows the radial distribution of GroES-binding sites. In the R-open state, the sites are at the same radius as in R-ES, but they are rotated by ~100°. GroES is schematically docked onto the R-open state to illustrate that the binding sites are readily accessible to the GroES mobile loops, unlike the situation in the Rs state. See also Figure S6.

21

Cell *149*, 113–123, March 30, 2012 ©2012 Elsevier Inc.

ATP-Triggered Conformational Changes Delineate Substrate-Binding and -Folding Mechanics of the GroEL Chaperonin

Daniel K. Clare,[1] Daven Vasishtan,[1] Scott Stagg,[2,6] Joel Quispe,[2] George W. Farr,[4,5,7] Maya Topf,[1] Arthur L. Horwich,[3,4,5] and Helen R. Saibil[1,*]

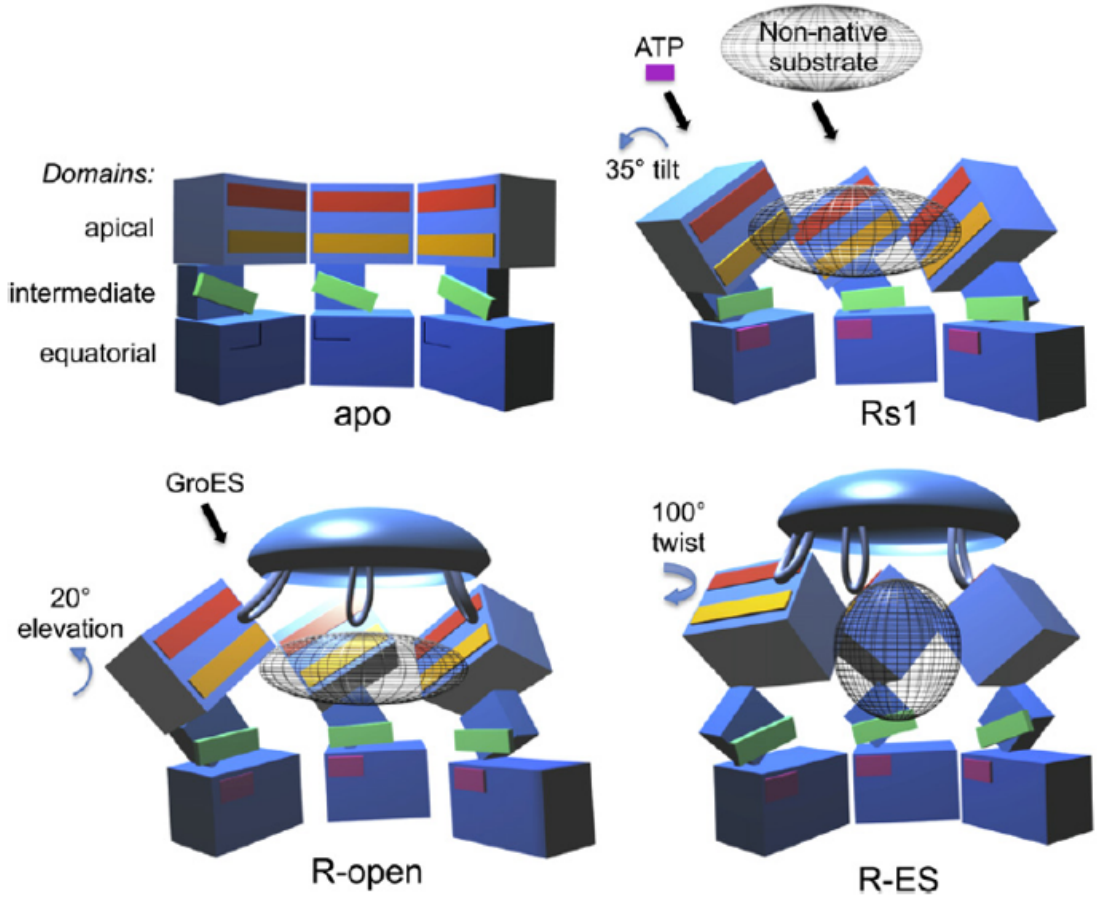


Figure 7. Cartoon of Domain Movements with Helices H, I, and M Highlighted in Apo, Rs, R-open, and R-ES States

Helices H (red), I (orange), and M (green) are highlighted. The folding substrate polypeptide is shown in gray mesh. GroES is shown docked onto the Rs-open state as in Figure 6C. The R-open apical domains have undergone about 70% of their elevation from apo to R-ES. This figure was generated in Blender 3D.

22

Cell *149*, 113–123, March 30, 2012 ©2012 Elsevier Inc.

ATP-Triggered Conformational Changes Delineate Substrate-Binding and -Folding Mechanics of the GroEL Chaperonin

Daniel K. Clare,[1] Daven Vasishtan,[1] Scott Stagg,[2,6] Joel Quispe,[2] George W. Farr,[4,5,7] Maya Topf,[1] Arthur L. Horwich,[3,4,5] and Helen R. Saibil[1,*]

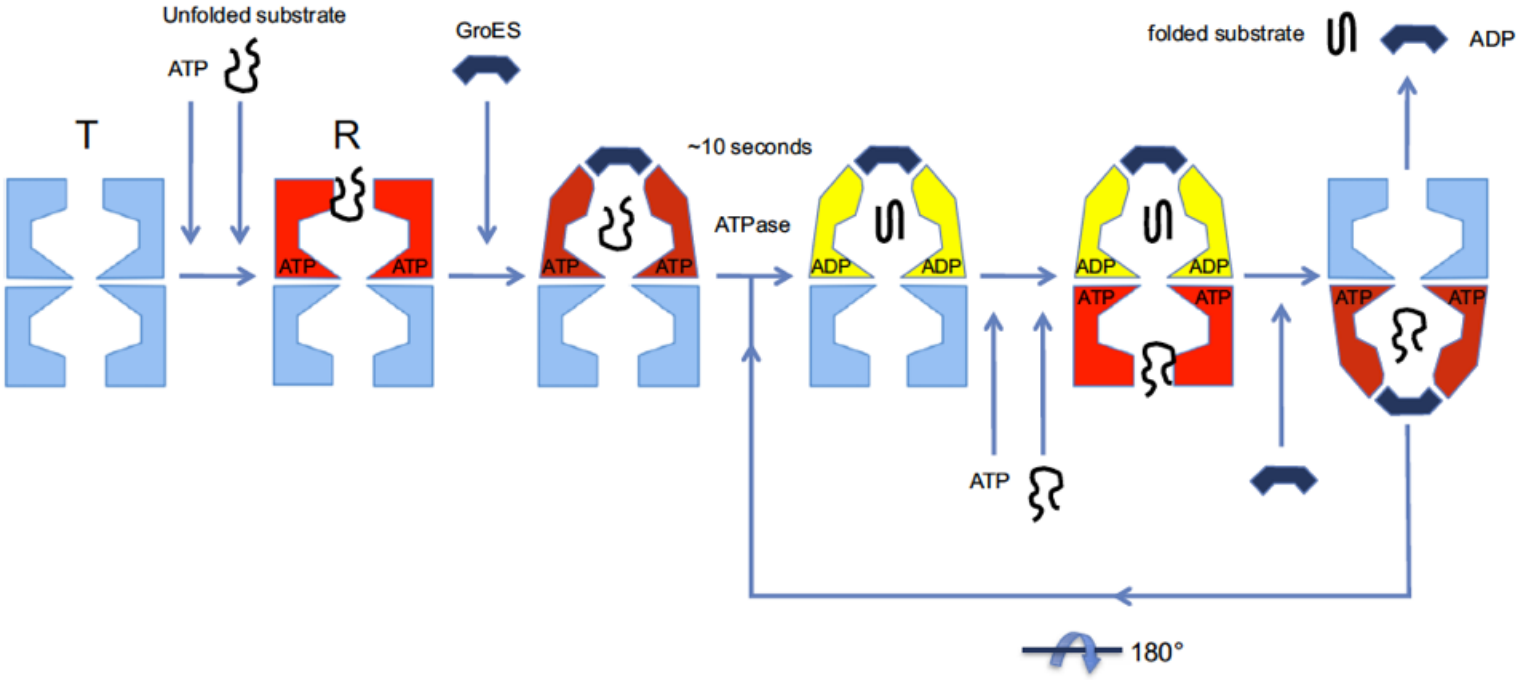


Figure 1. Diagram of the GroEL-GroES ATPase Cycle, Showing the Steps of Substrate Binding, Encapsulation, Folding, and Release
ATP hydrolysis in one ring is required to enable subsequent ATP binding to the opposite ring, but hydrolysis is not required for folding to proceed within the chamber.

23