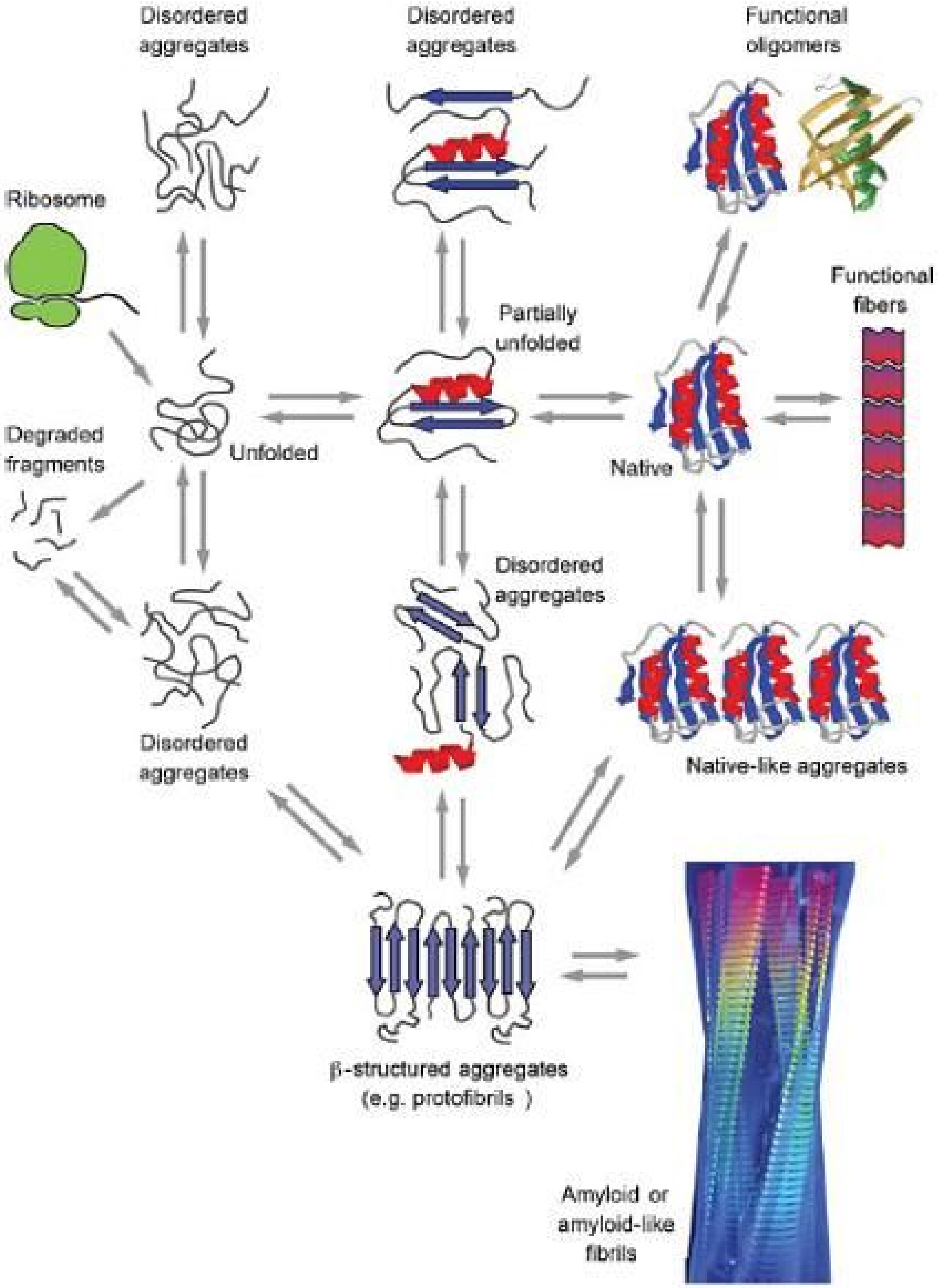
July 30, 2010

Collapse and cooperativity in protein folding

Jayant B. Udgaonkar

National Centre for Biological Sciences
Tata Institute of Fundamental Research
Bangalore, India

How do proteins fold, unfold and misfold?



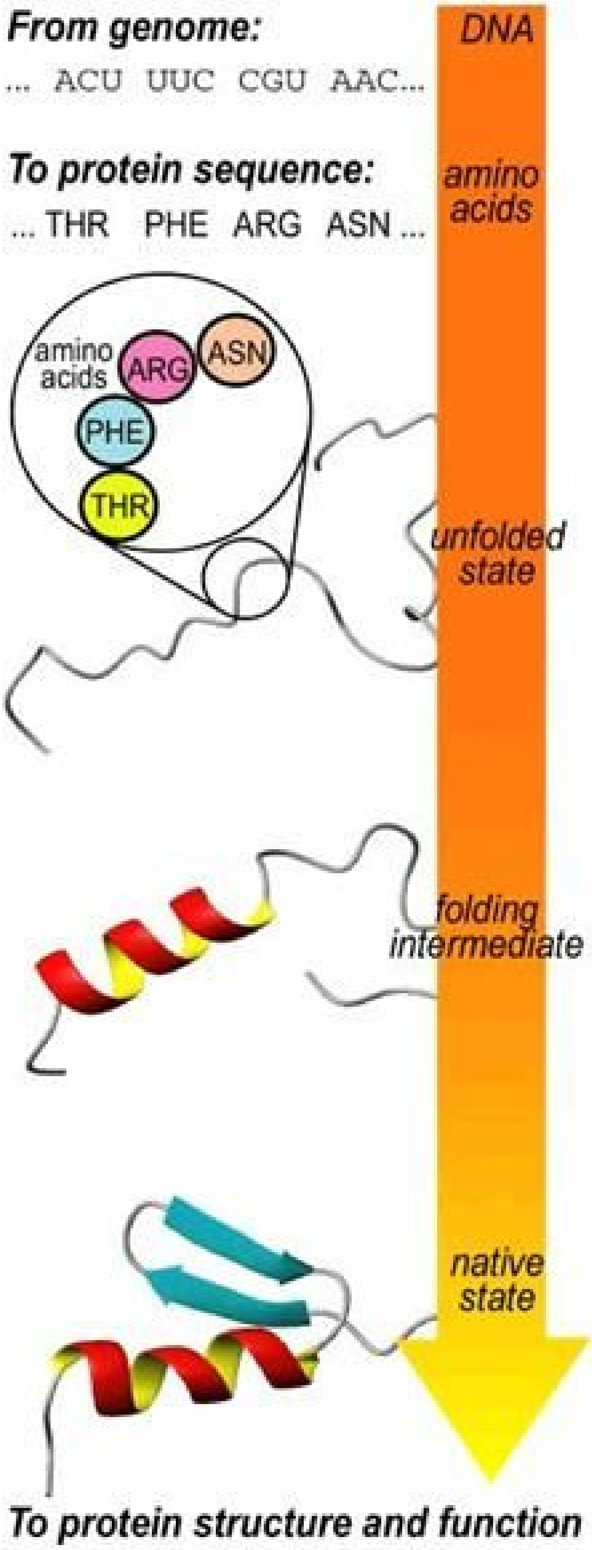
MOLECULAR PLAYERS IN THE LABORATORY

barstar
monellin
SH3 domain
GroEL

mouse prion protein
human tau
α-synuclein

ribonuclease A
cytochrome *c*
GFP
barnase
thioredoxin

Protein folding: A central unsolved problem in biology

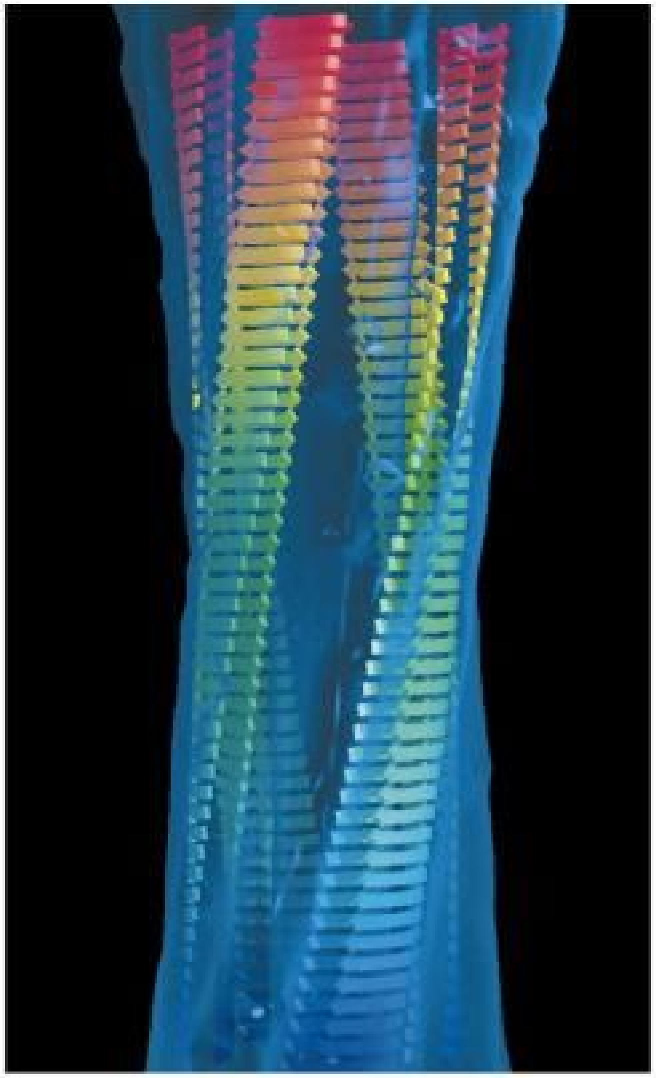


protein folding

$$U \leftrightarrow N$$

unfolded protein

folded protein



Challenge for both experiment and theory

Experiment: A fast, dynamic, kinetic process accompanied by large conformational changes.

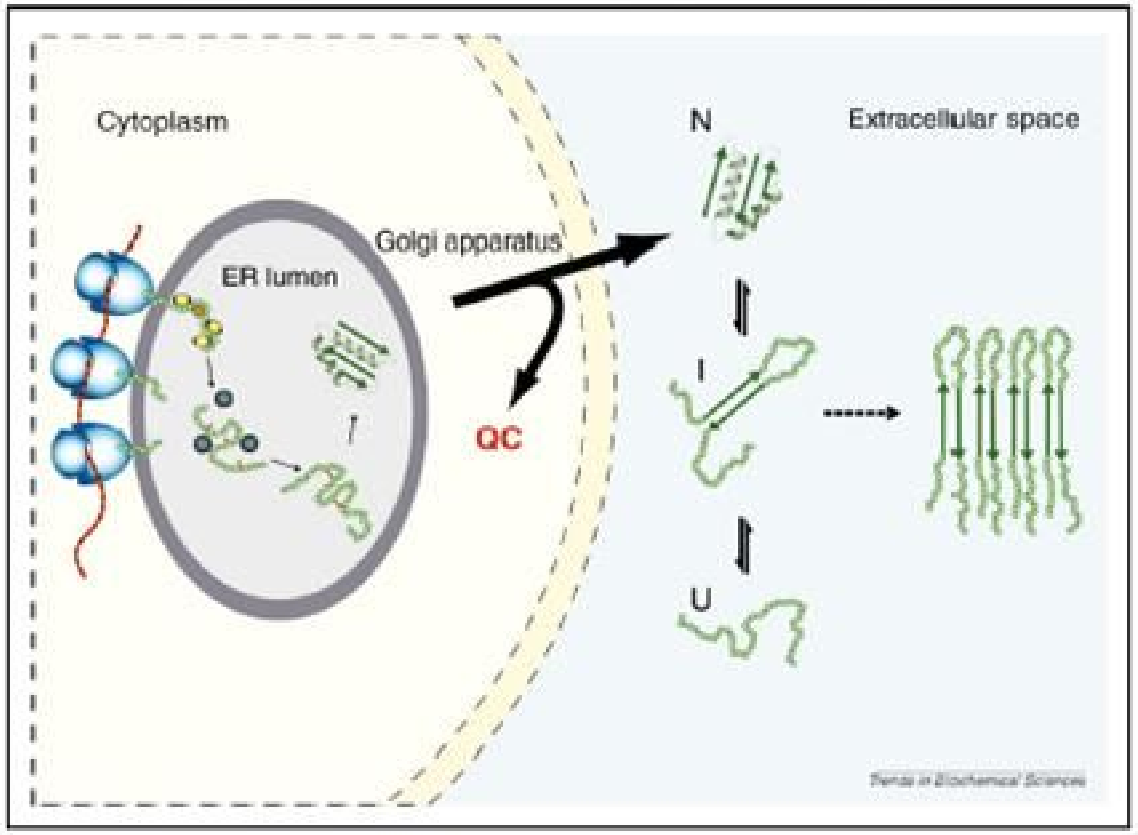
Theory: A closed problem with a solution. It should be possible to go from sequence to coordinates

Why is protein folding important?

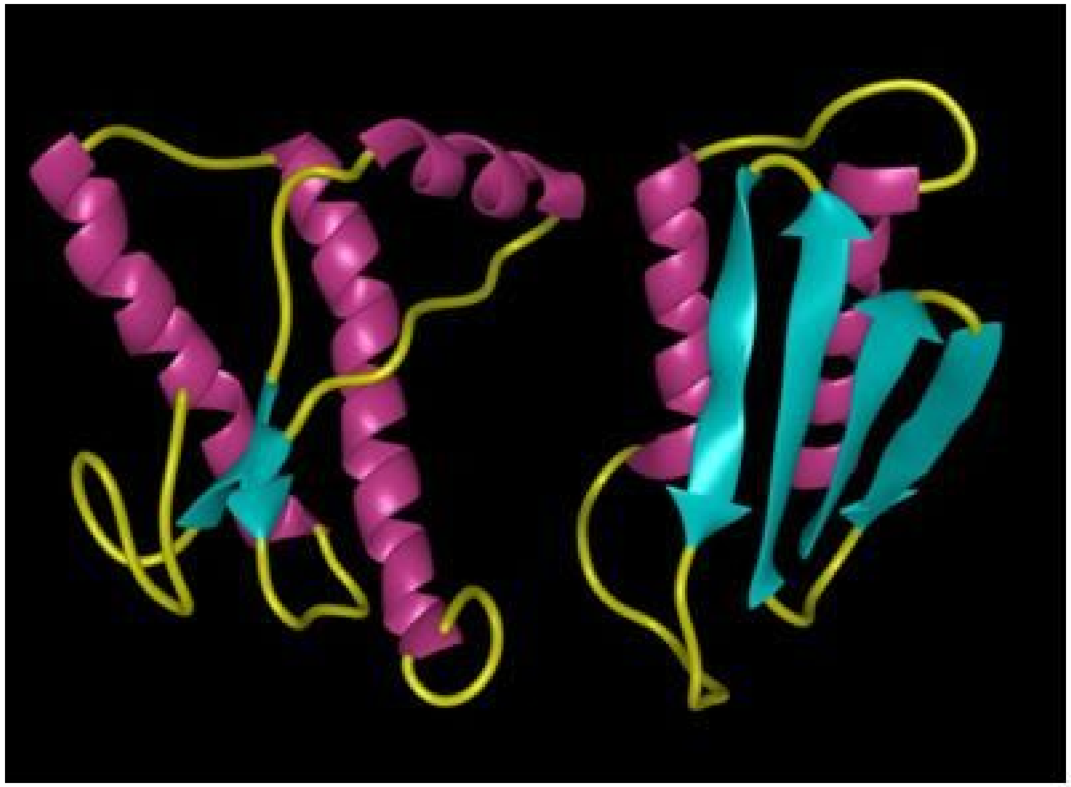
- Deduction of structure from sequence
- Rational Alteration of Existing Proteins
- Design of Novel Proteins
- Understanding the molecular basis of protein structure-function relationships
- Understanding the molecular basis of folding diseases

The highest operation in nature and in art is the attainment of significant form - Goethe

Protein misfolding and amyloid formation

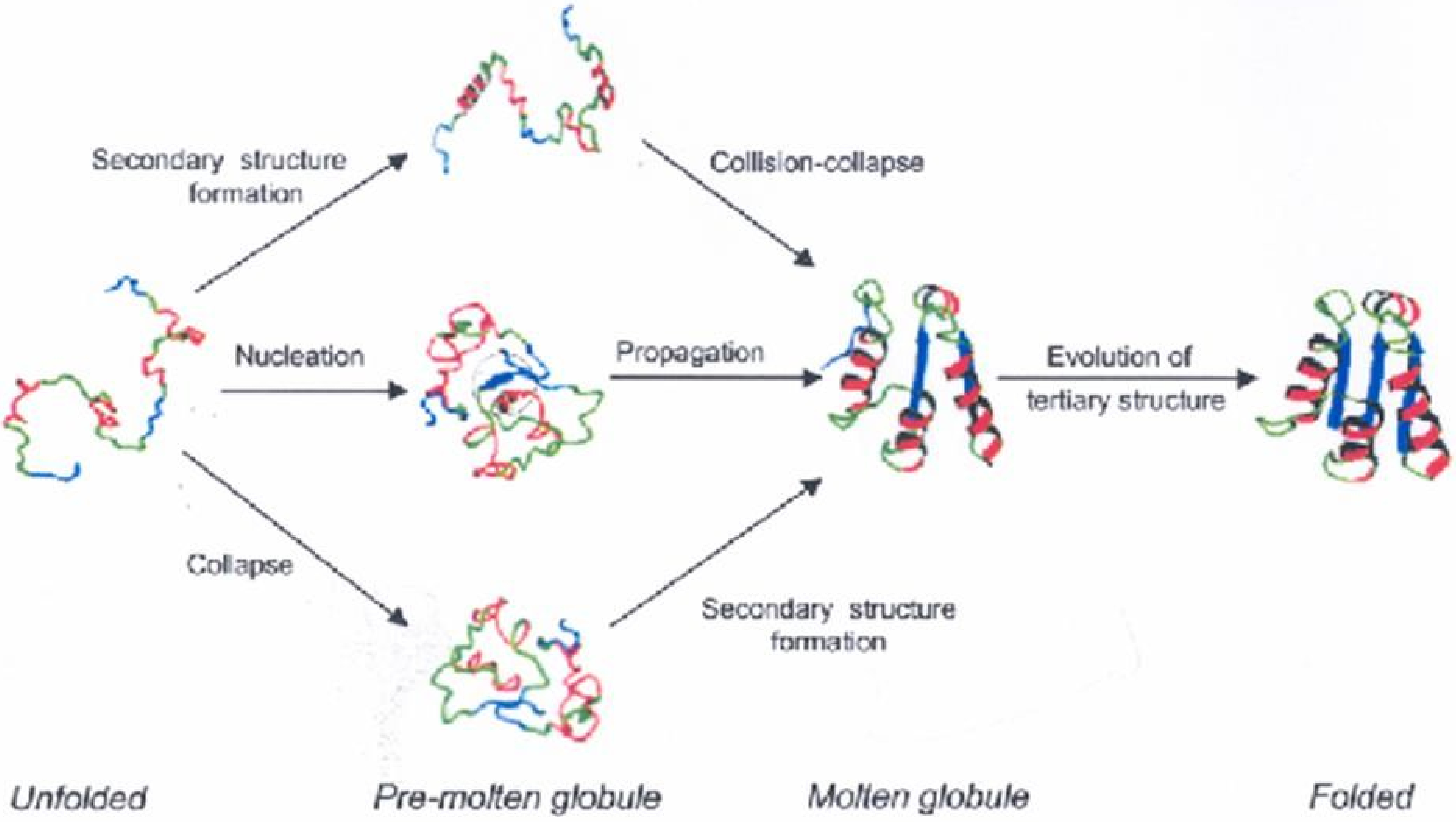


Prion protein



Cellular form Scrapie form

Pathways of Protein Folding



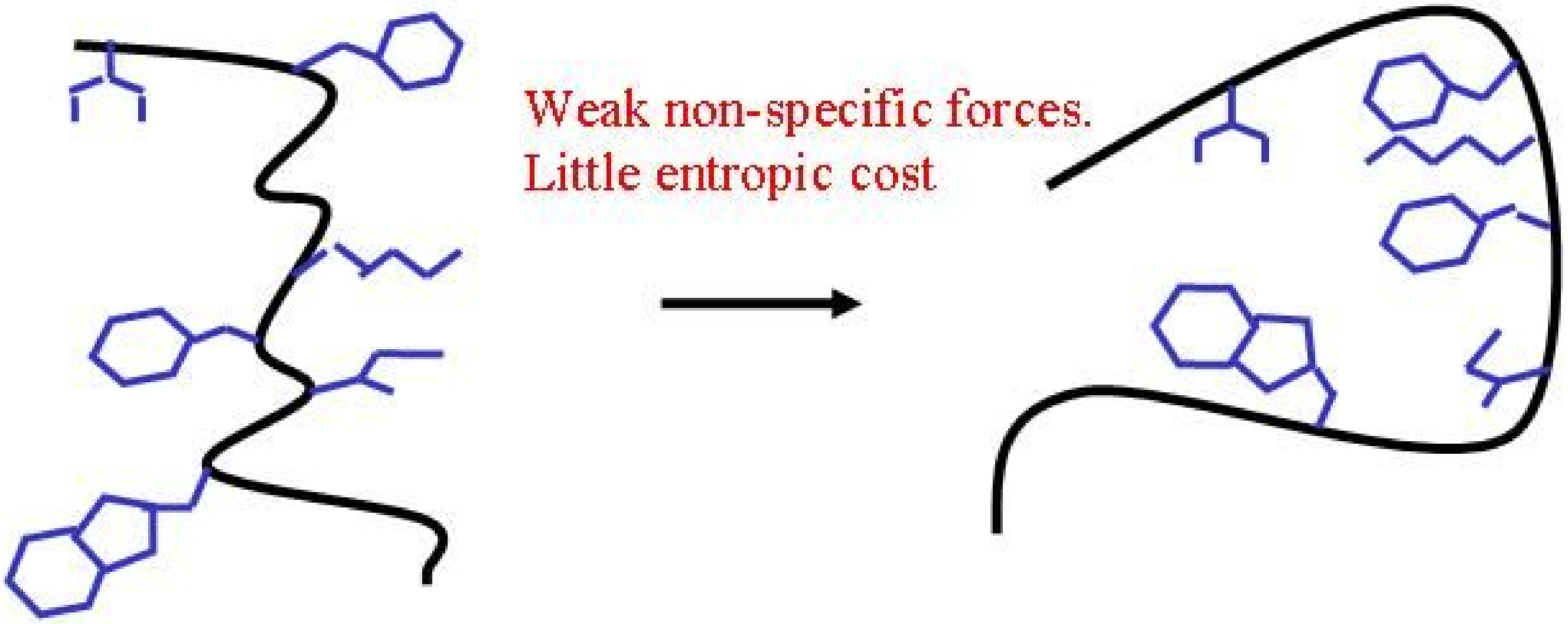
Udgaonkar (2008) Annu Rev Biophysics

Hydrophobic Collapse

Many proteins fold via an initial collapse of the polypeptide chain.

A polypeptide chain will collapse when intra-chain interactions dominate over chain-solvent interactions.

A more hydrophobic chain is more likely to collapse upon a change from good to bad solvent.



Unfolded chain

Exposed non-polar side chains

Weak non-specific forces.
Little entropic cost

Loose compact structure

Non-polar side chains partially buried

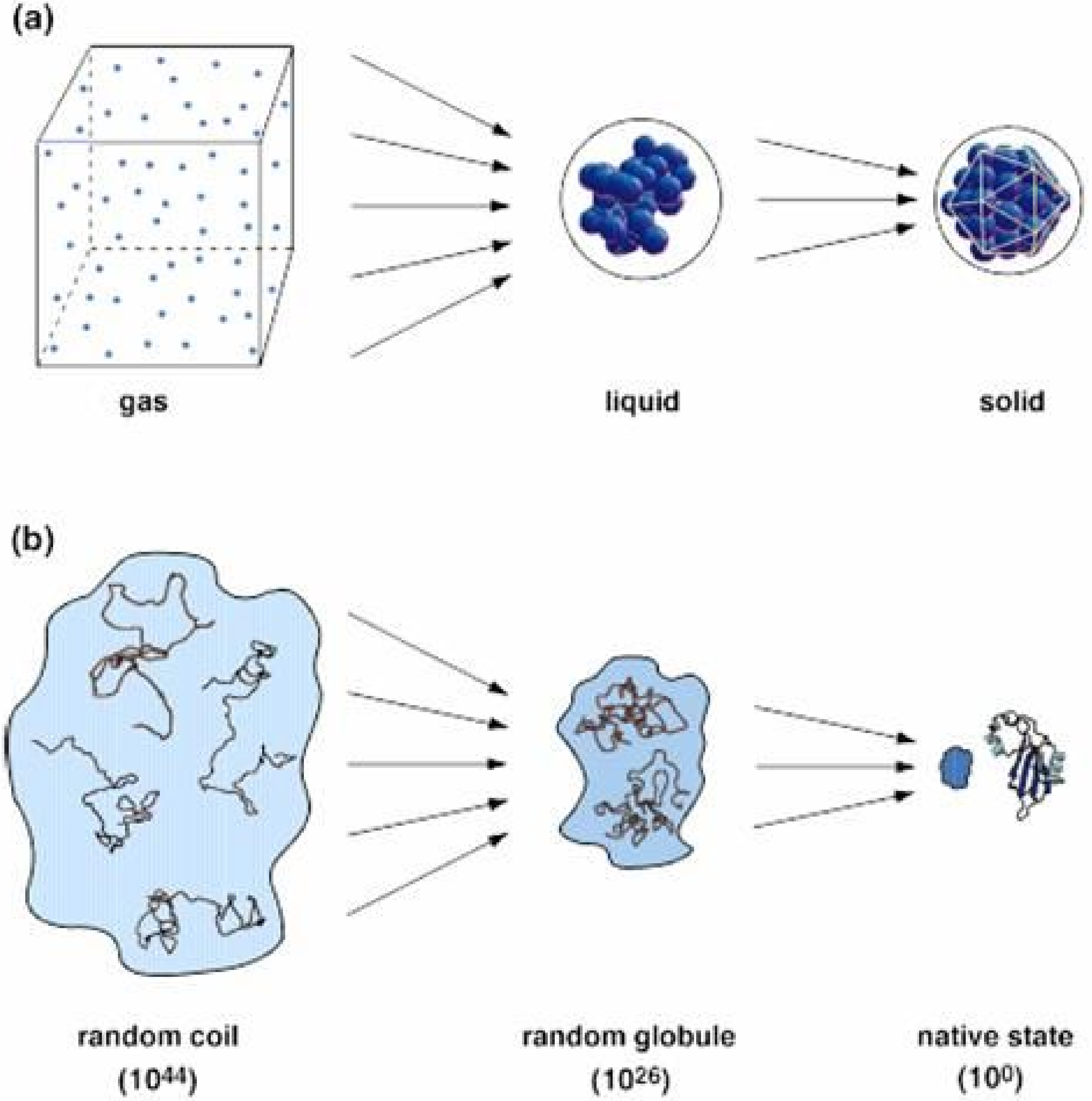
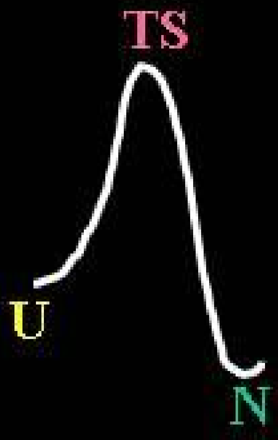


Figure 3. a) Schematic illustration of the phase transitions of a stable cluster consisting of 55 Lennard–Jones particles. The solid cluster has the structure of a Mackay icosahedron.[19] b) Schematic illustration of the corresponding transitions during protein folding. The protein barnase, composed of 110 amino acid residues, is used as an example. Carbon α traces are shown (prepared by the program Molscript).[160] The random coil and disorganized globule conformations were constructed to have radii of gyration of approximately 31 and 15 Å, respectively; the native structure has a radius of gyration of 13.4 Å. The numbers of possible main-chain conformations were calculated from the estimates of the main-chain entropy per residue; there are 2.57 and 1.7 conformations per residue in the random coil and random semicompact states, respectively.[55a] Therefore, the number of random coil states is $2.57^{109} \approx 10^{44}$, and the number of compact random globule states is $1.7^{109} \approx 10^{26}$.

Heterogeneity in protein folding reactions

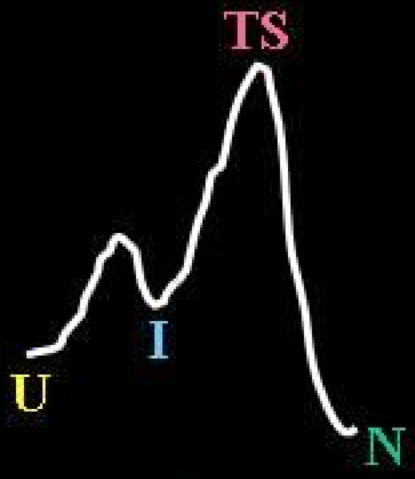
Two-state *versus* multi-state folding

Two-state folding

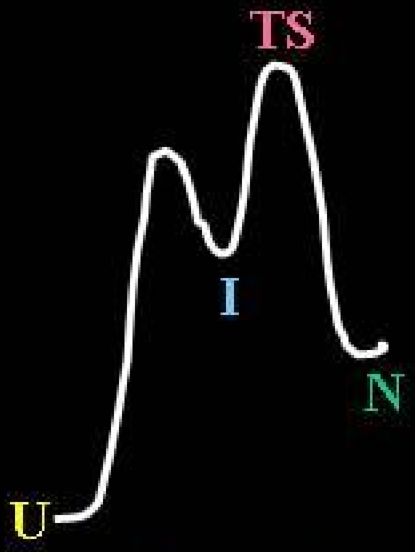


Apparent two-state folding

Strongly stabilizing

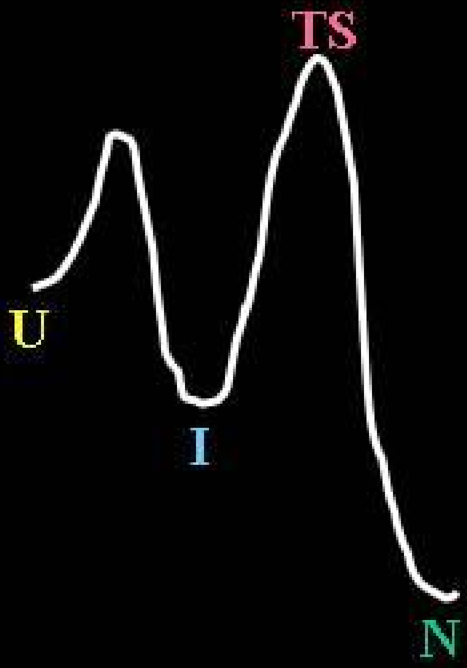


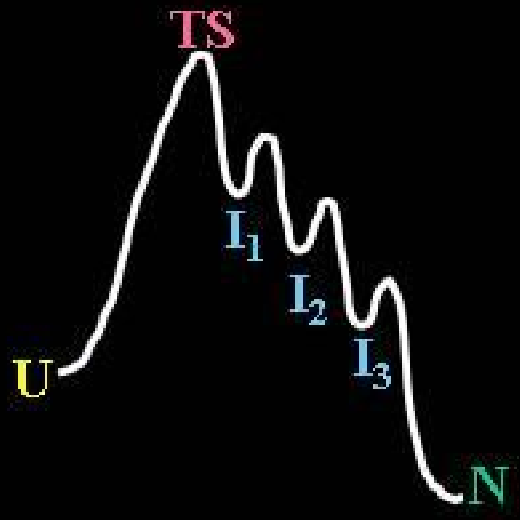
Marginally stabilizing



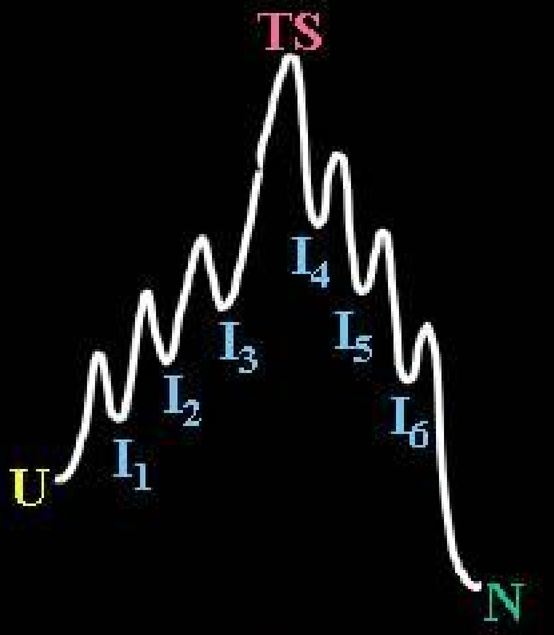
Destabilizing

Three-state folding



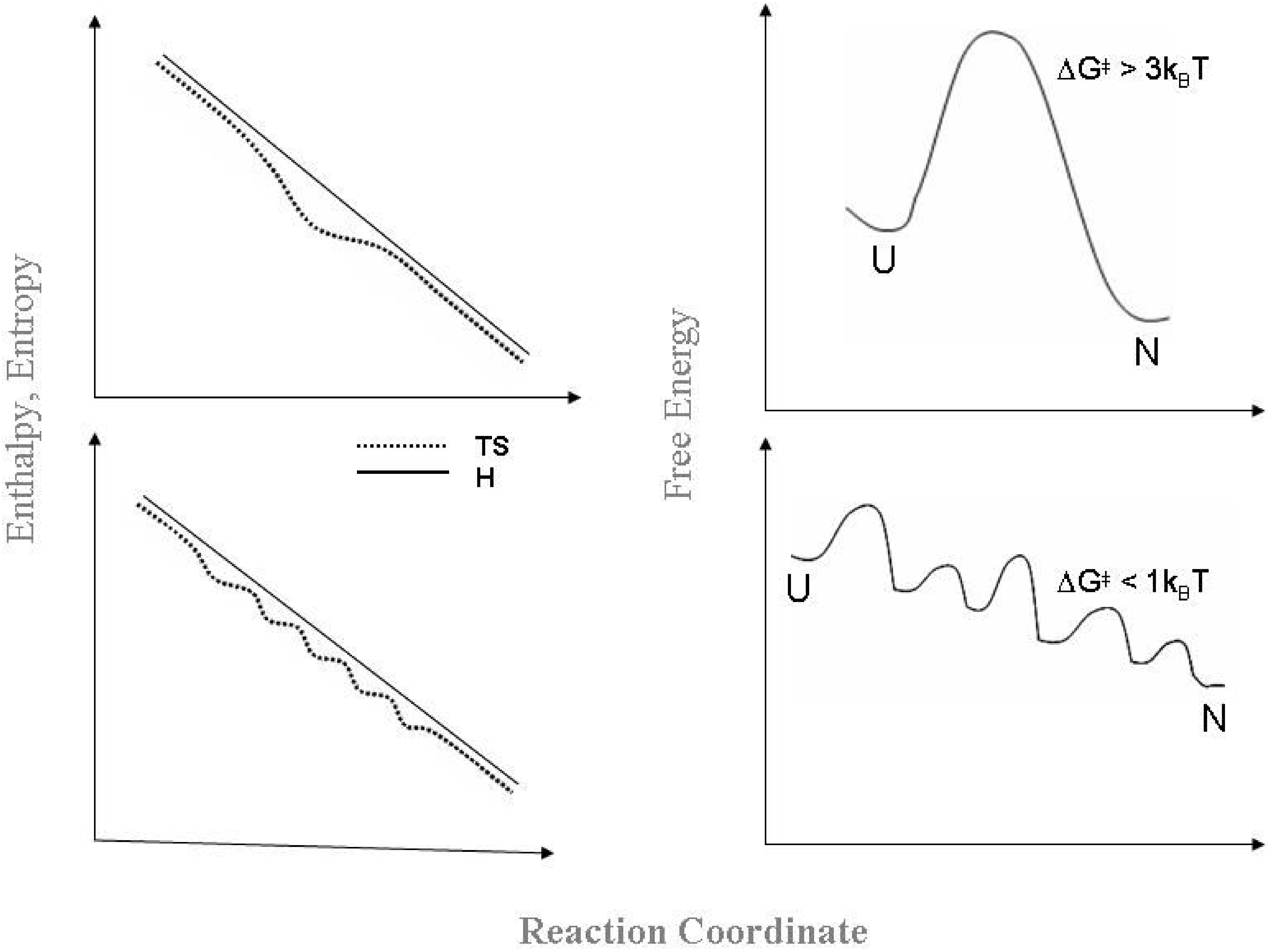


Intermediates after rate-limiting step



Continuum of intermediates

Dominant Free Energy Barrier versus Many distributed Barriers



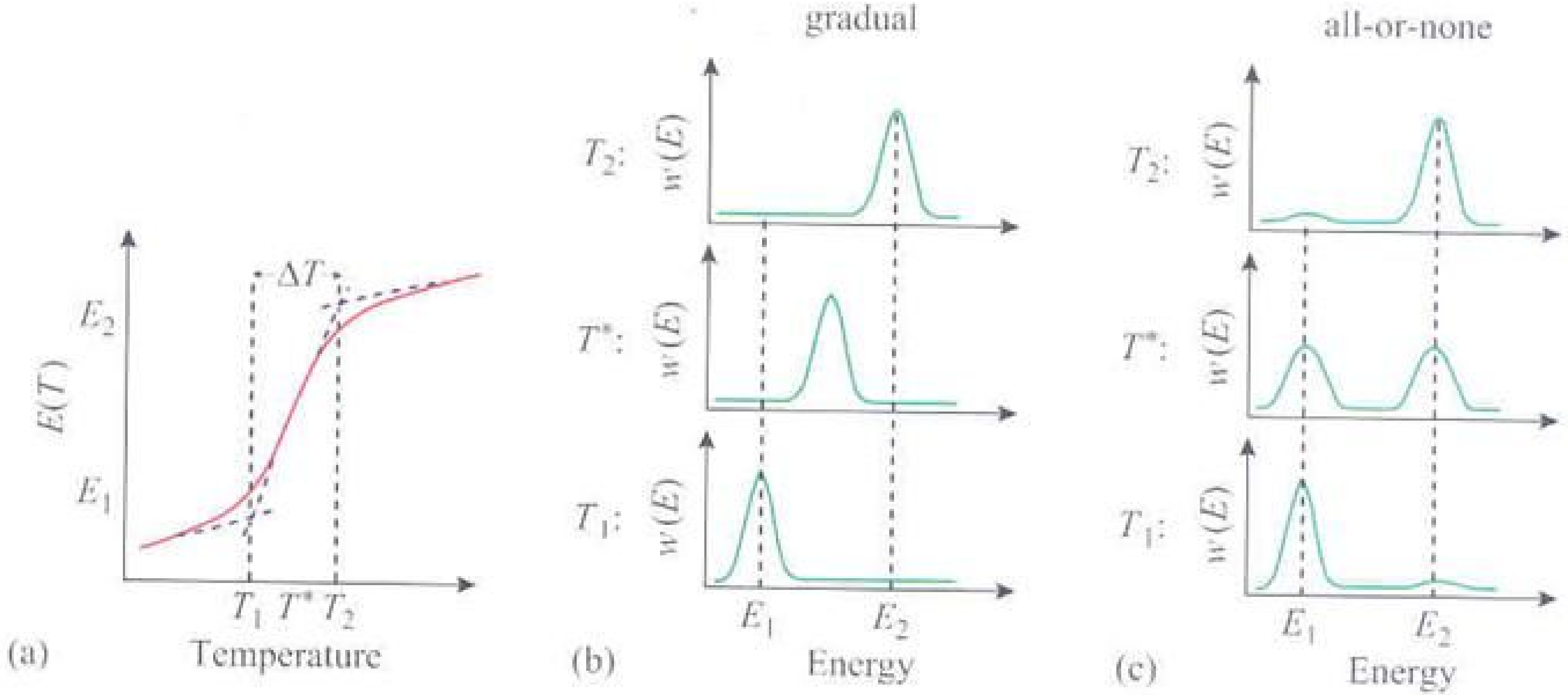
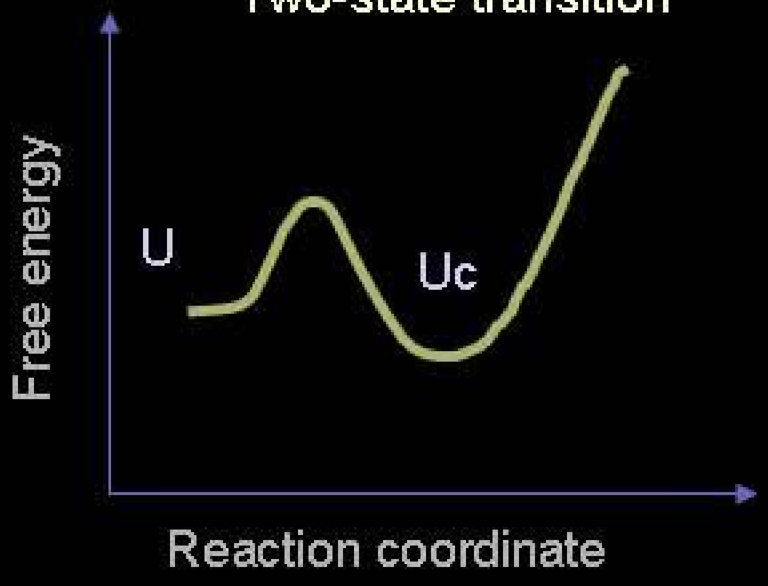


Figure 17.3. With the same temperature dependence of energy E (or any other observable parameter), the cooperative ("S-shaped") transition can be either of the "all-or-none" type (example: protein denaturation), or gradual (example: helix–coil transition in polypeptides). The difference is displayed in the shape of the function $w(E)$ showing the distribution of molecules over energy (or over any other observable parameter) rather than in the shape of the curve $E(T)$. Dashed lines in (a) show a graphical determination of ΔT, the width of the temperature transition.

Protein Physics, Finkelstein and Ptitsyn,

Is chain collapse a two-state or continuous structural transition?

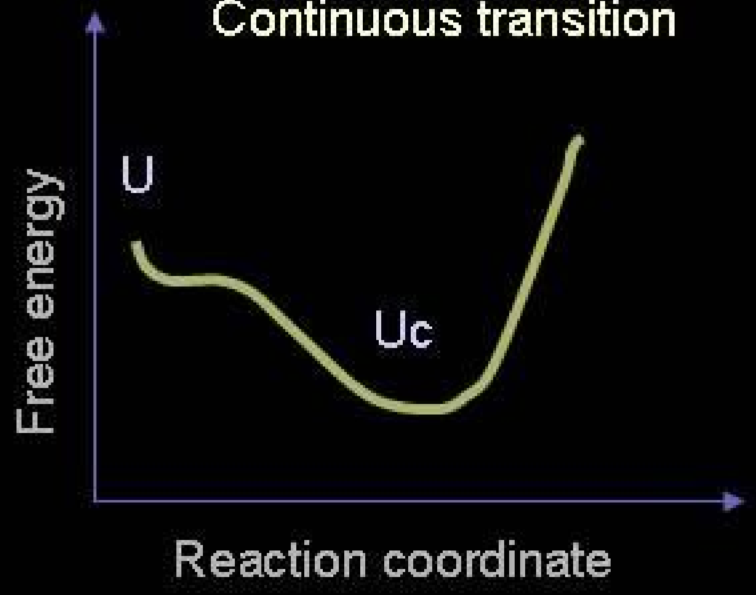
U and Uc are the unfolded forms in unfolding and refolding conditions.


Two-state transition

U and Uc are separated by an energy barrier.

U and Uc coexist and are in constant exchange.

Does Uc possess specific structure or is chain collapse a completely nonspecific event?


Continuous transition

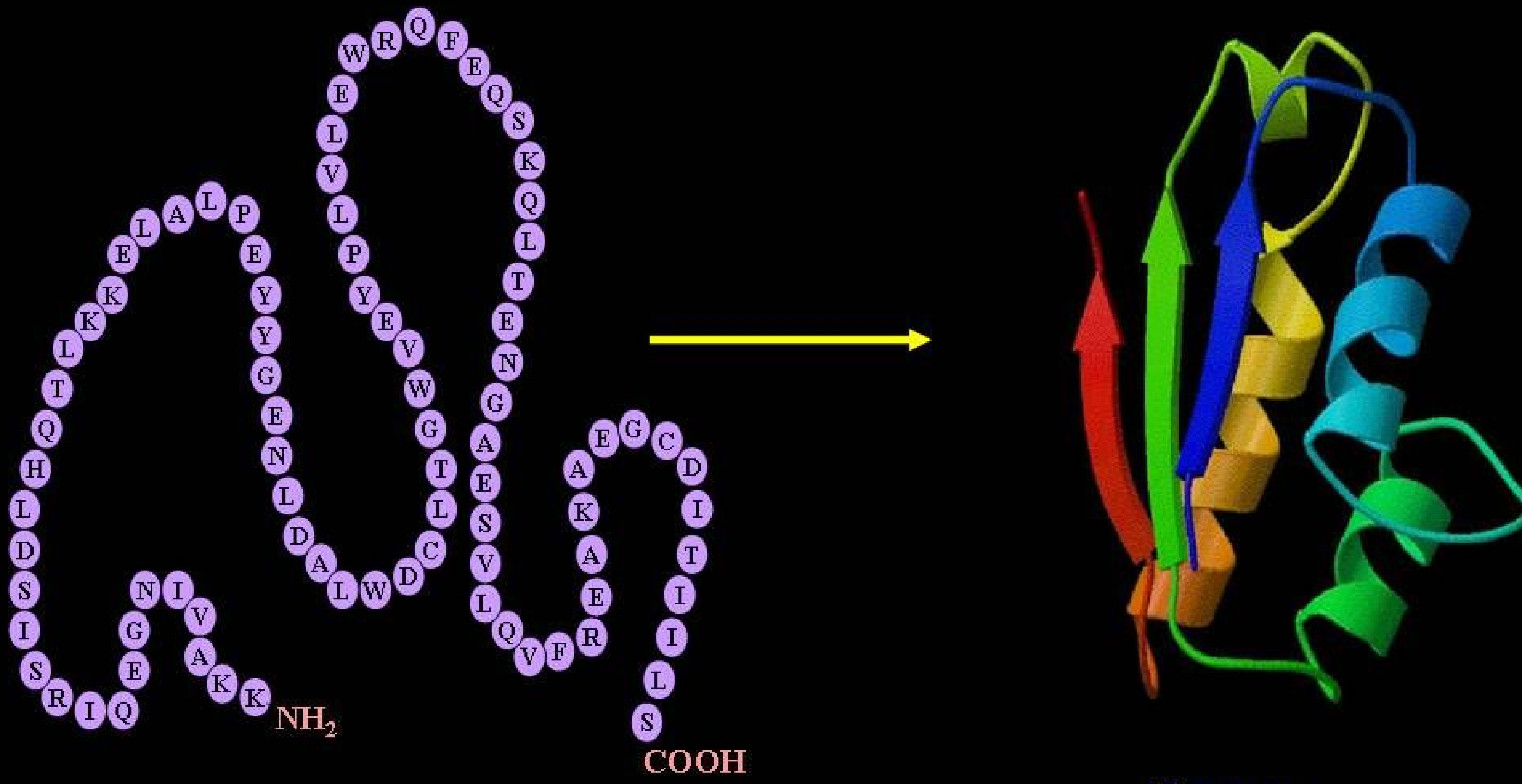
No energy barrier → downhill folding

U collapses to Uc through a continuum of intermediate structures.

Do homopolymer and heteropolymer (protein) chains collapse in a similar manner?

How fast is chain collapse?

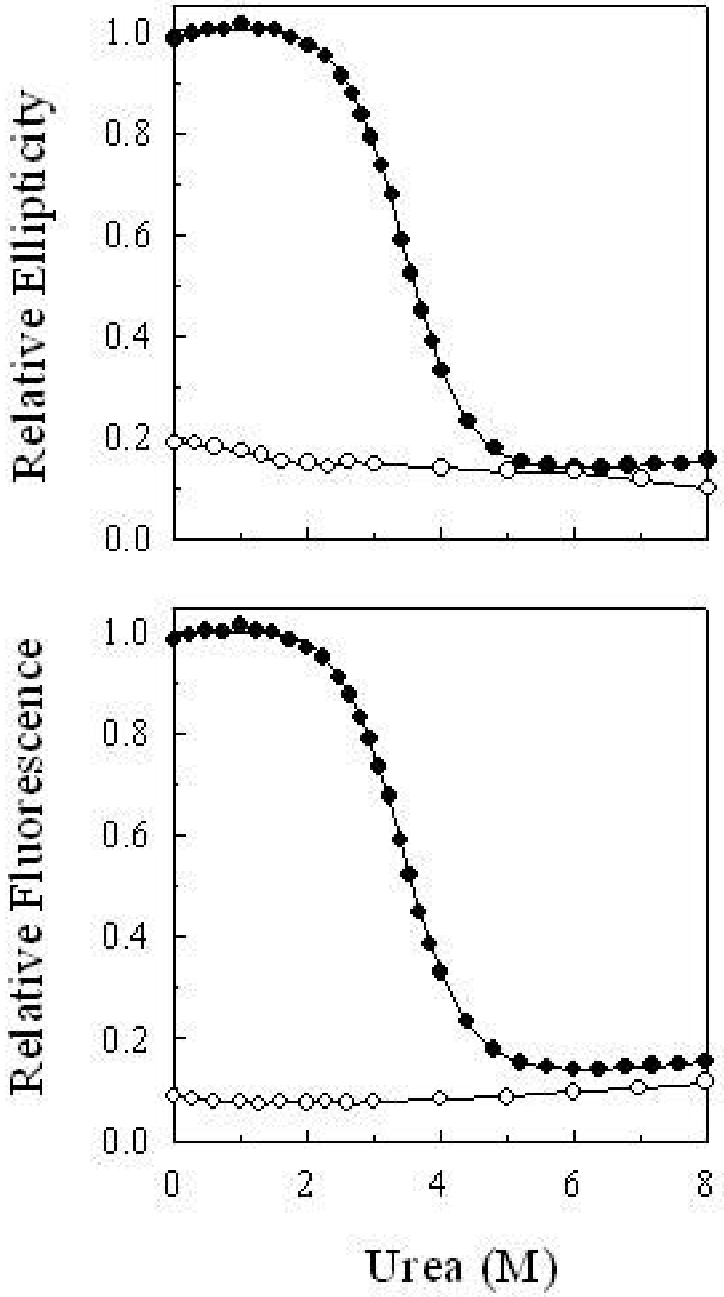
The amino acid sequence codes for a specific and unique structure



Barstar

The pH 12-denatured and urea-denatured states of barstar are similar

Equilibrium studies



condition	R_H (nm)	Elution Vol. (ml)
pH 7	1.90	13.4
pH 12	2.98	10.2
6 M GdnHCl, pH 7	2.85	9.8

The spectroscopic and hydrodynamic properties of both states are identical.

pH-induced unfolding is completely reversible

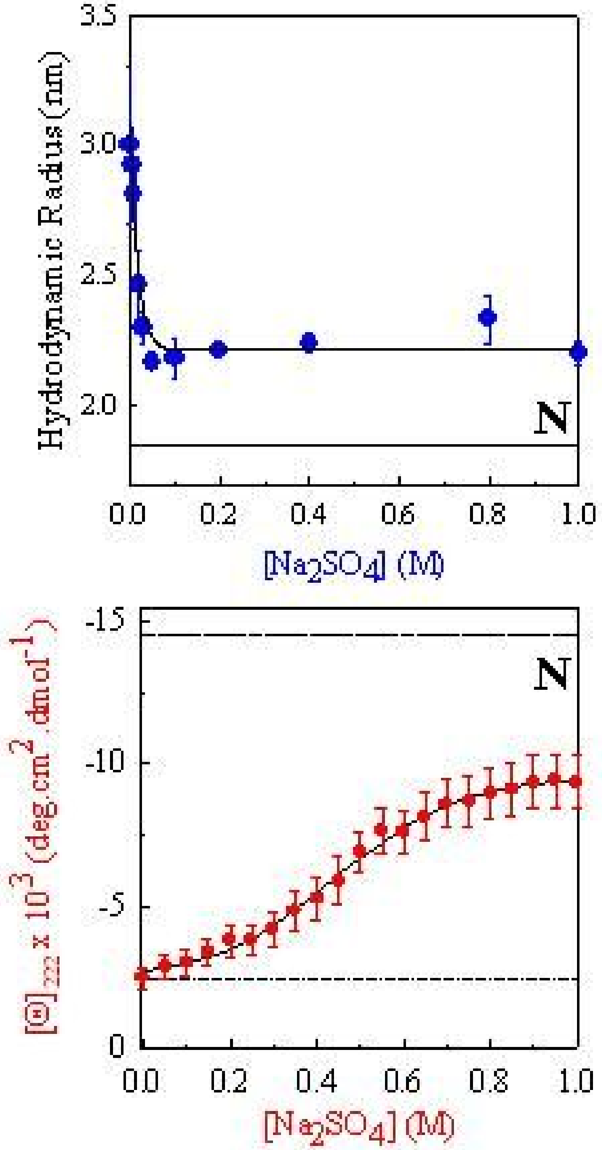
Bhadresh R. Rami and Jayant B. Udgaonkar*
Biochemistry **2001,** *40,* 15267−15279

Specific structure acquisition occurs in compact forms

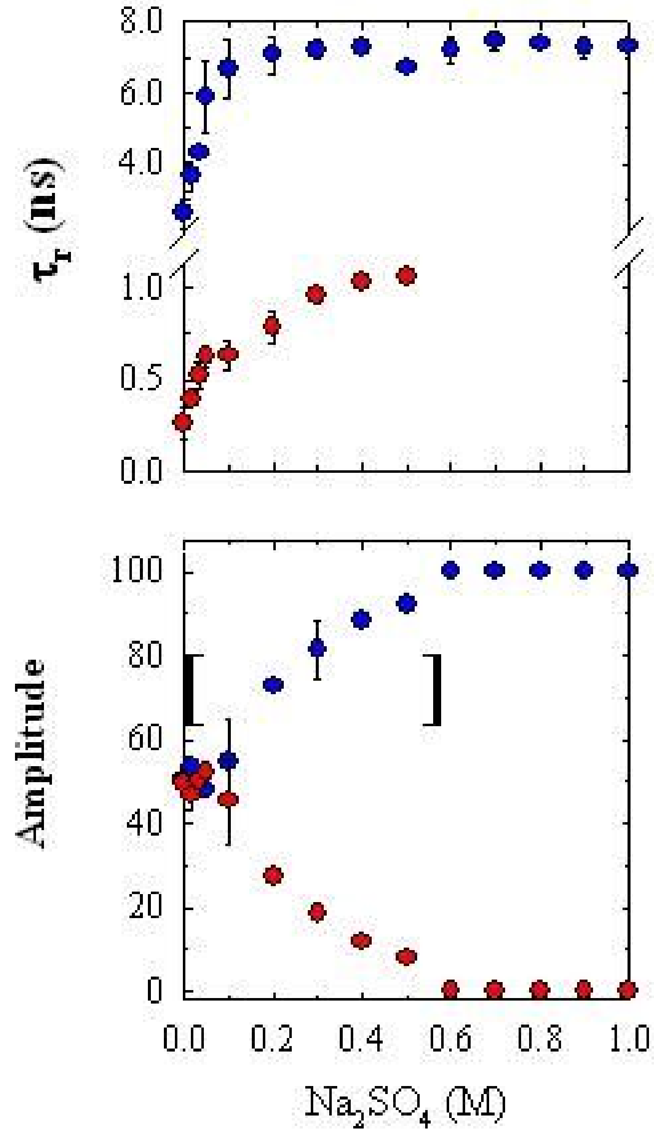
Bhadresh R. Rami and Jayant B. Udgaonkar* *Biochemistry* **2002,** *41,* 1710−1716

Bhadresh R. Rami,[‡] G. Krishnamoorthy,[§] and Jayant B. Udgaonkar* *Biochemistry* **2003,** *42,* 7986−8000

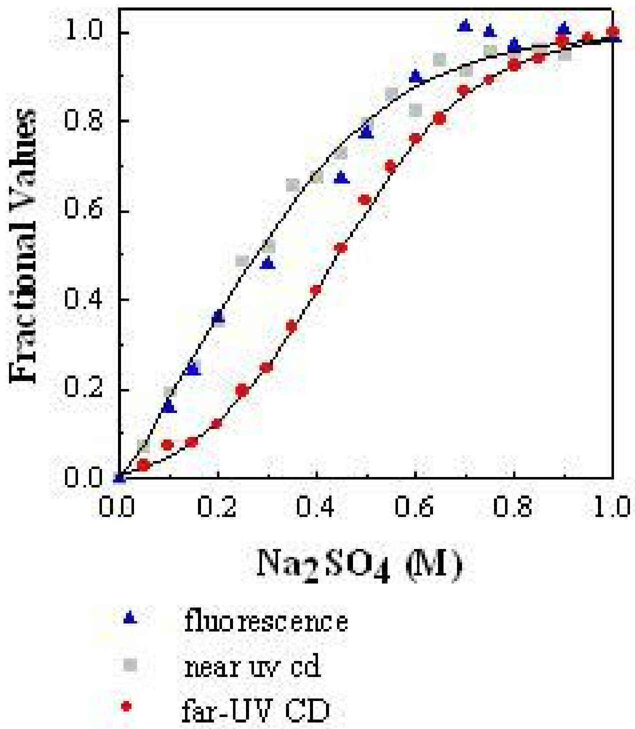
Collapse precedes secondary structure formation



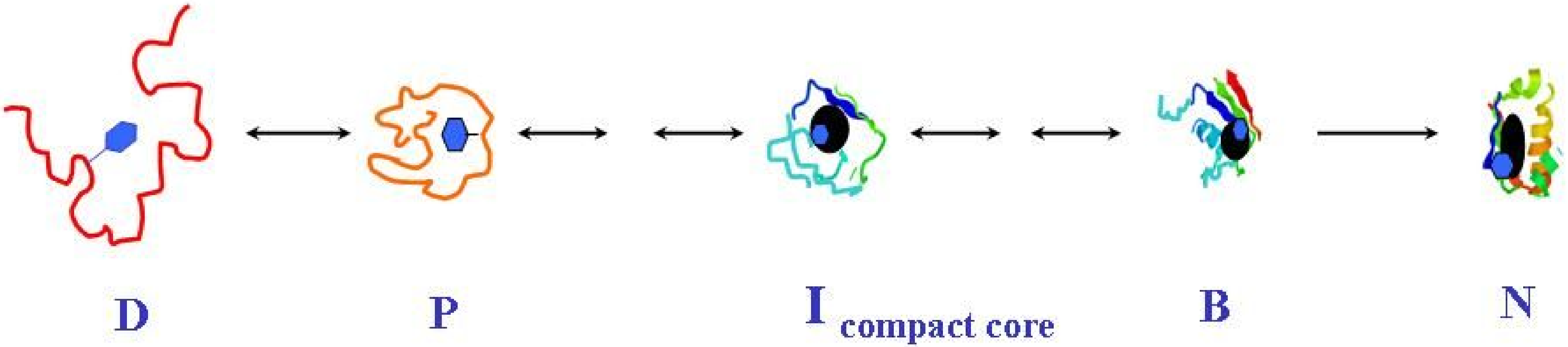
Collapse precedes core consolidation



Development of structure is not cooperative



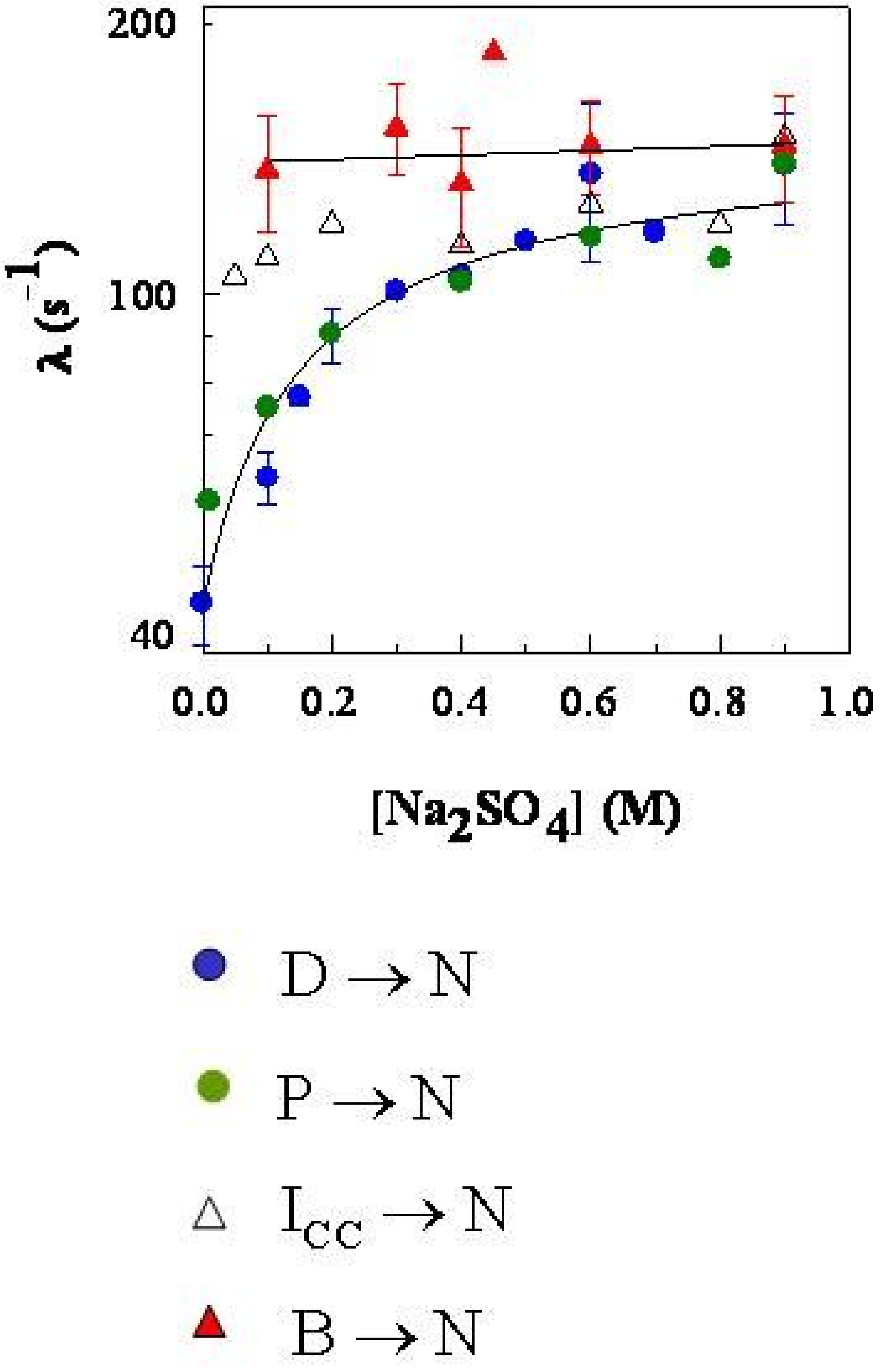
Progressive formation of structure during folding

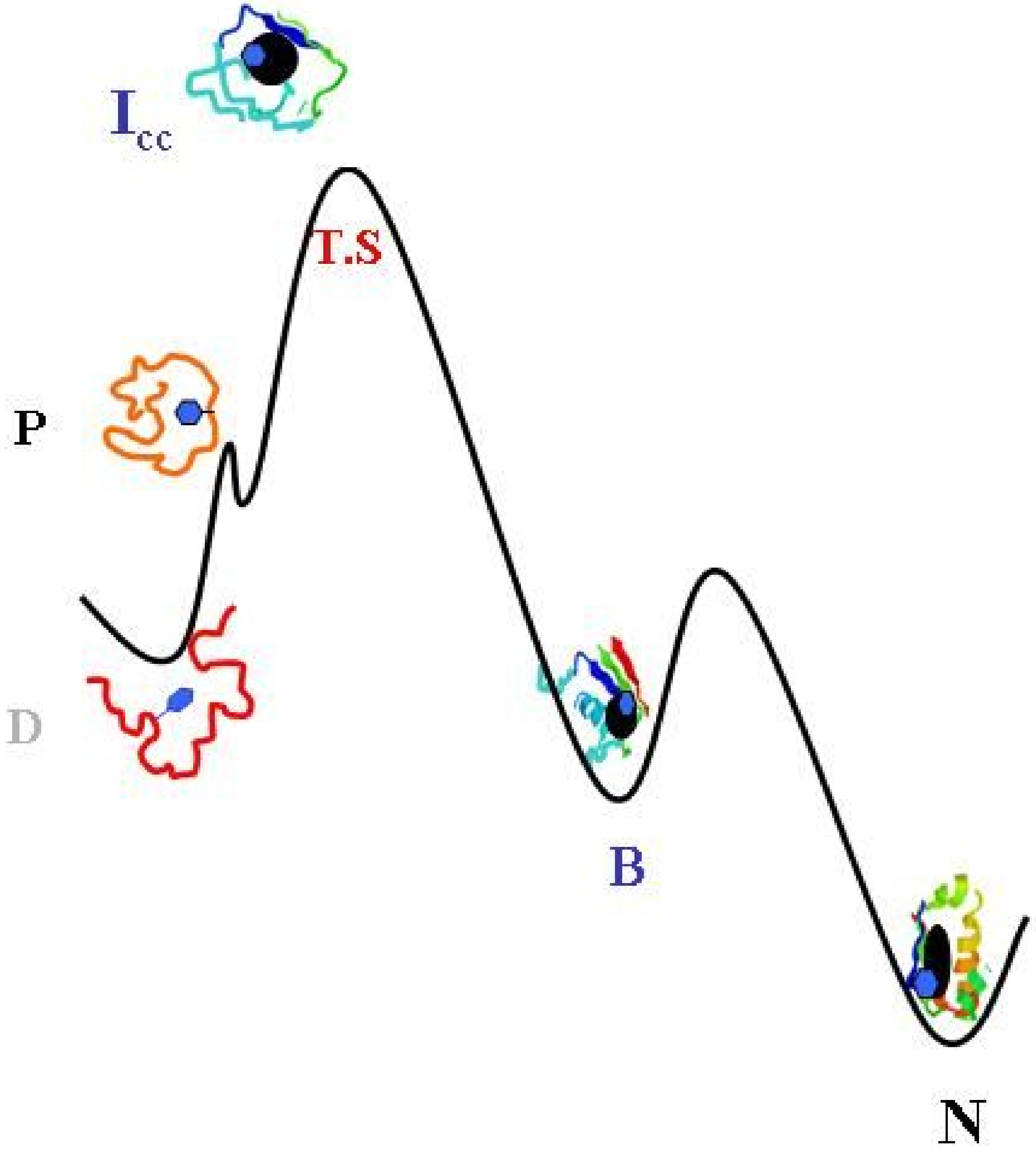


1. The denatured D form undergoes a global collapse to a pre molten globule P form
2. The P form is a compact, structure-less globule with a solvated core.
3. The P form undergoes a highly non-cooperative transformation to the molten globule B form.
4. The B form is a dry molten globule.
5. The P ↔ B transition may be an essentially continuous structural transition.
6. The B form is a folding intermediate that populates the direct folding pathway from the D form to the N state.

Bhadresh R. Rami and Jayant B. Udgaonkar*
Biochemistry **2002,** *41,* 1710−1716

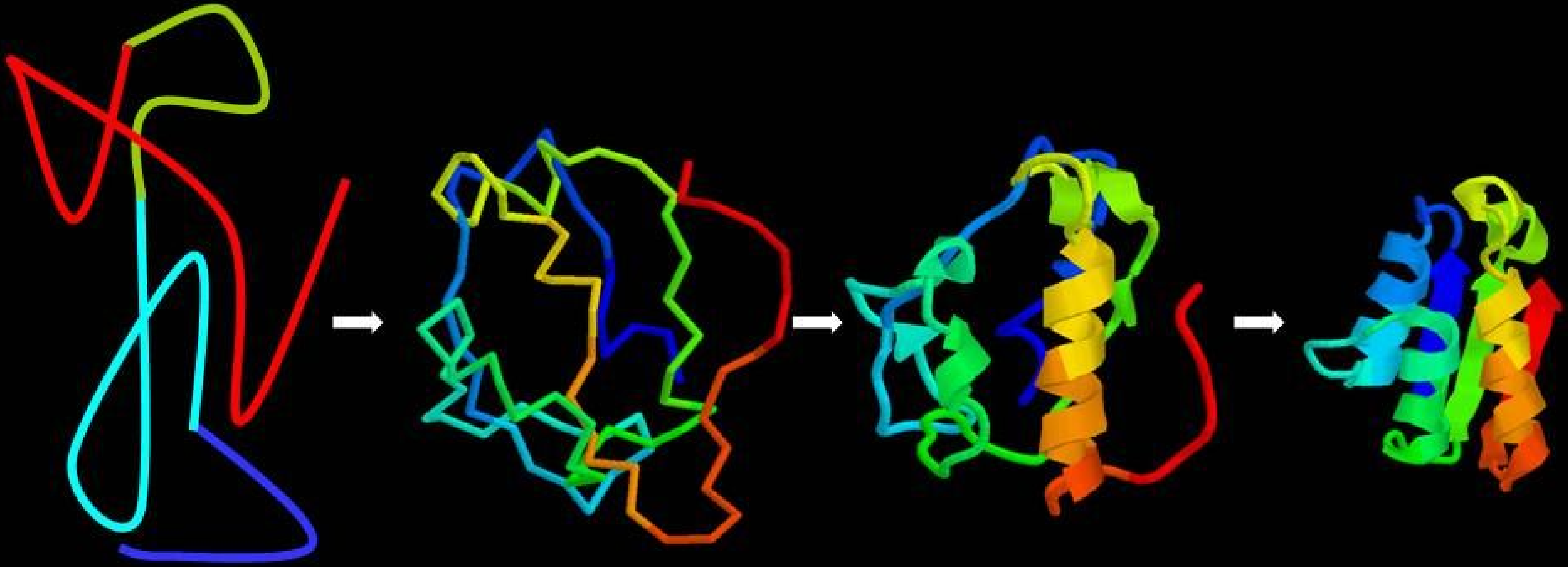
Consolidation of the core occurs in the rate-limiting step of folding





Bhadresh R. Rami,[‡] G. Krishnamoorthy,[§] and Jayant B. Udgaonkar[*]
Biochemistry **2003,** *42,* 7986−8000

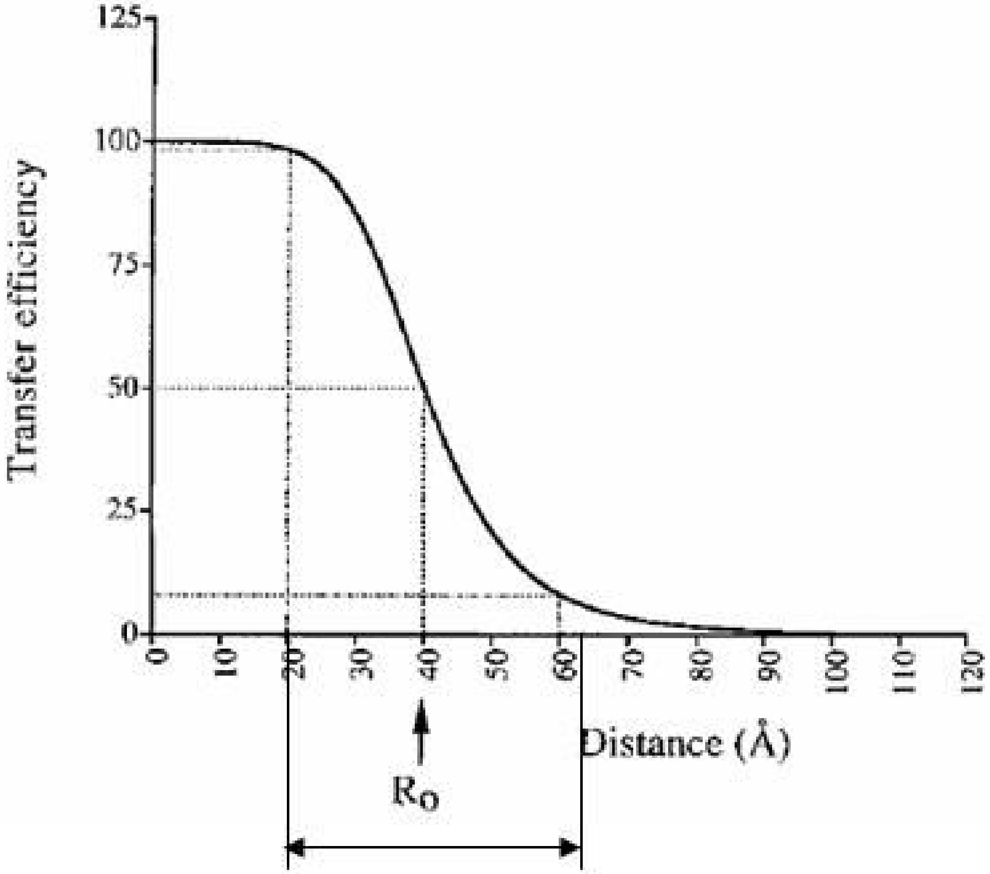
Hydrophobic collapse precedes specific structure formation during folding

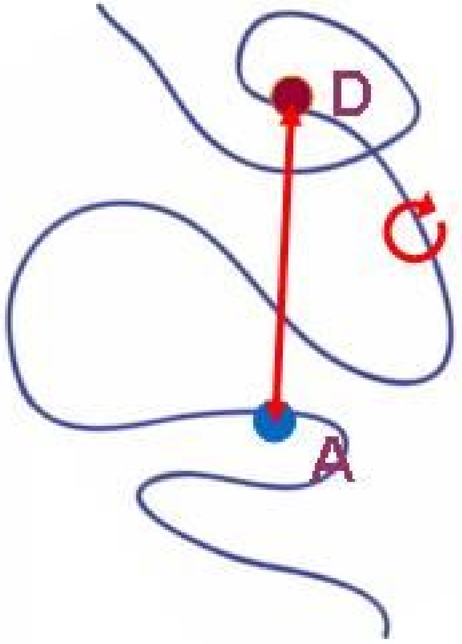


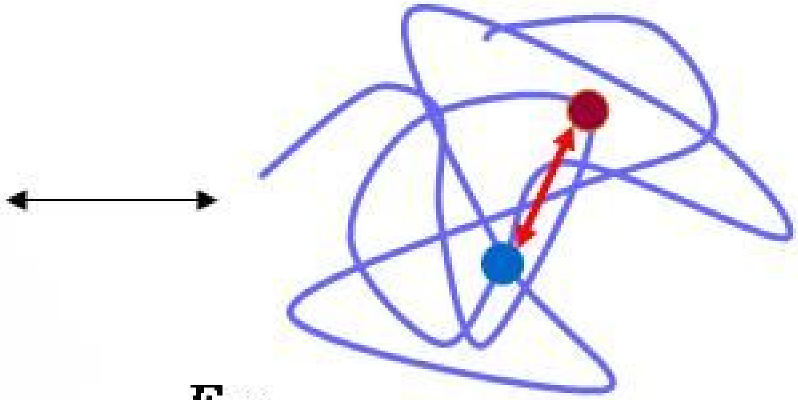
Unfolded form Pre-molten globule Molten globule Native state

Agashe, Shastry & Udgaonkar (1995) Nature 377, 754-759.
Rami & Udgaonkar (2002) Biochemistry 41, 1710-1716.
Pradeep & Udgaonkar (2004) J. Biol. Chem.

FRET as a spectroscopic ruler of molecular dimensions







$$E = 1 - \frac{F_{DA}}{F_D}$$

↓

$$E = [1 + \frac{R^6}{R_0^6}]^{-1}$$

FRET measurements are reliable in $0.5\ Ro < R < 1.5\ Ro$ range

$$Ro = 0.211(Q_D J \kappa^2 n^{-4})^{1/6}$$

Q_D = quantum yield of the donor

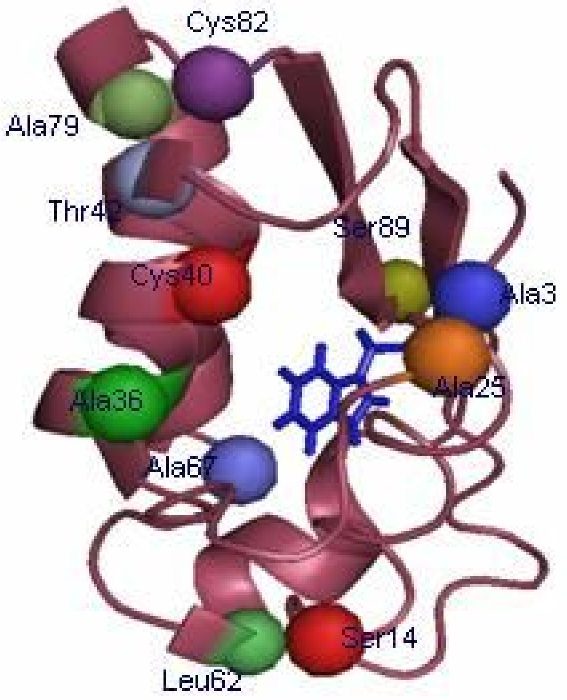
J = spectral overlap between the donor’s emission and acceptor’s absorption spectra

κ^2 = orientation factor

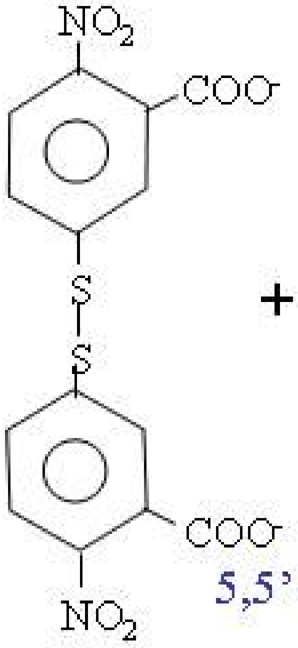
n = refractive index of the medium

Multi-site FRET enables determination of conformational changes in different regions of the protein

- Trp53 in barstar serves as the donor. Cys residues can be engineered at different sites in the protein.
- CysTNB is a good acceptor for Trp emission.
- Mutant forms of barstar containing a single Trp and a single Cys residue were generated by site-directed mutagenesis.
- The mutant proteins all had the same stability, unlabeled as well as labeled.



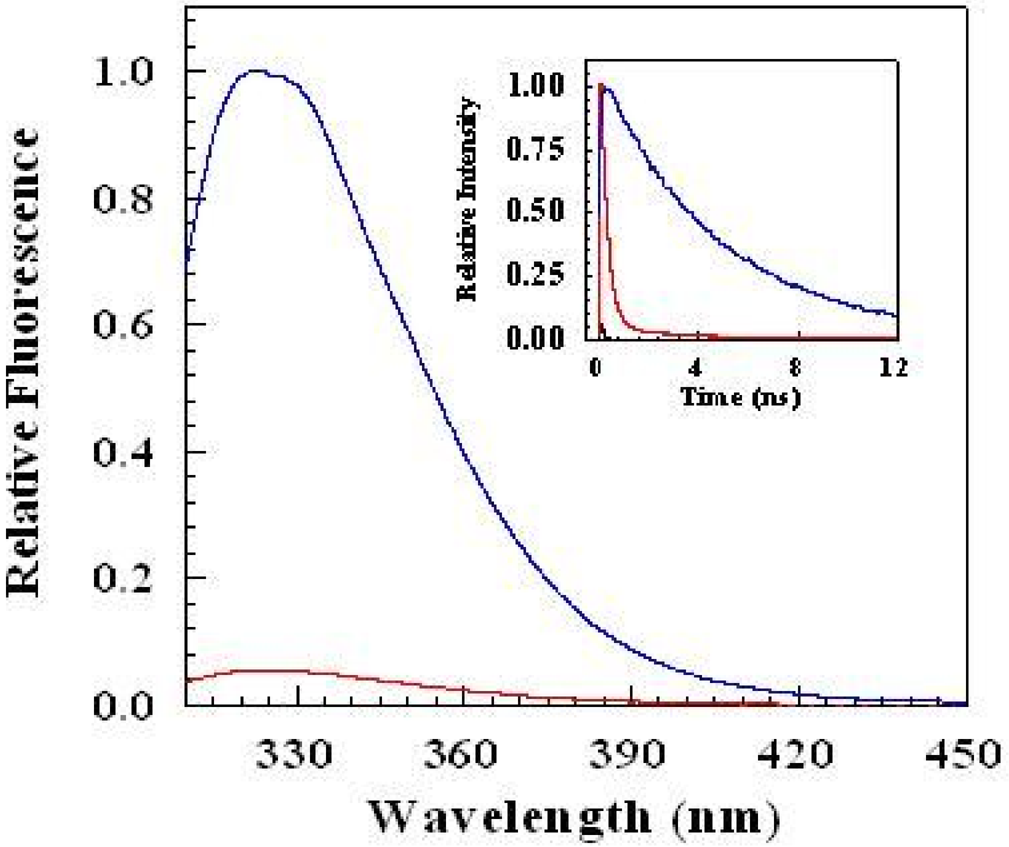
Labeling the protein with the acceptor TNB group

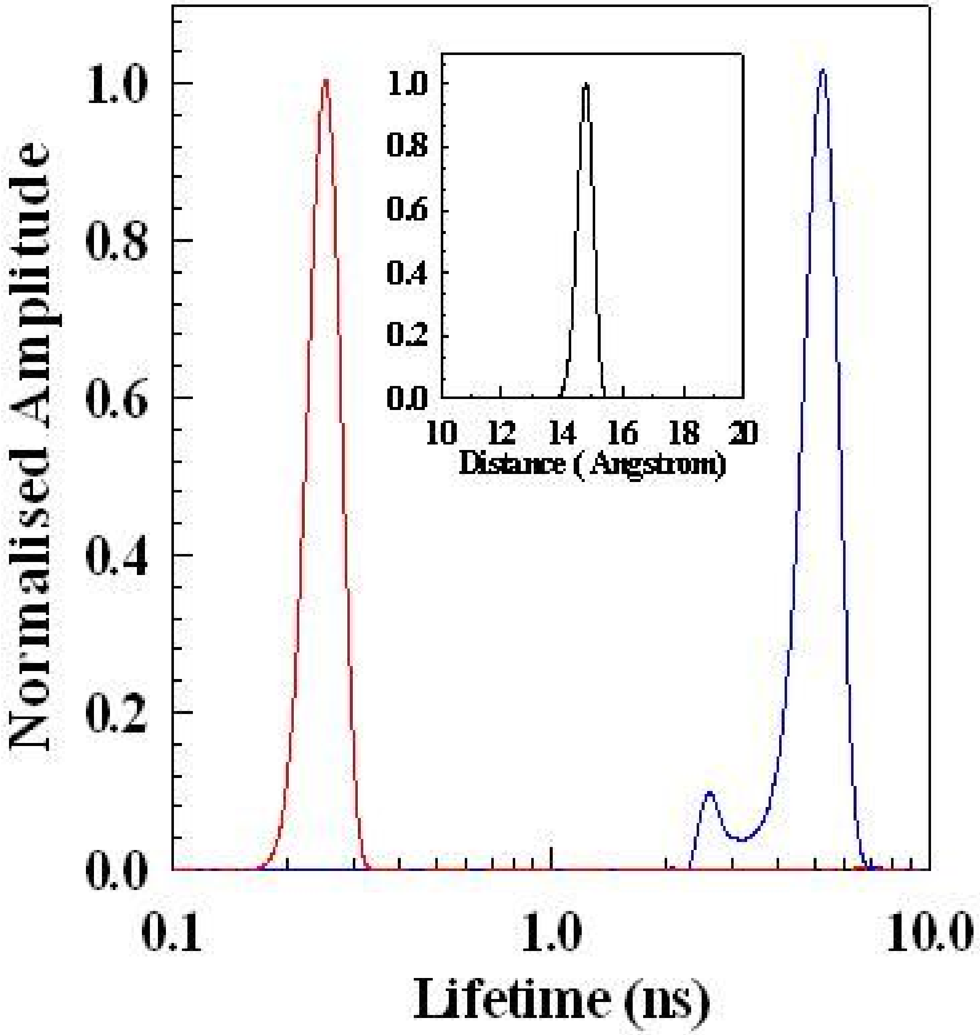


5,5'-dithio bis -(2-nitrobenzoic acid)

+

Distance measured by FRET is a good measure of structure during folding





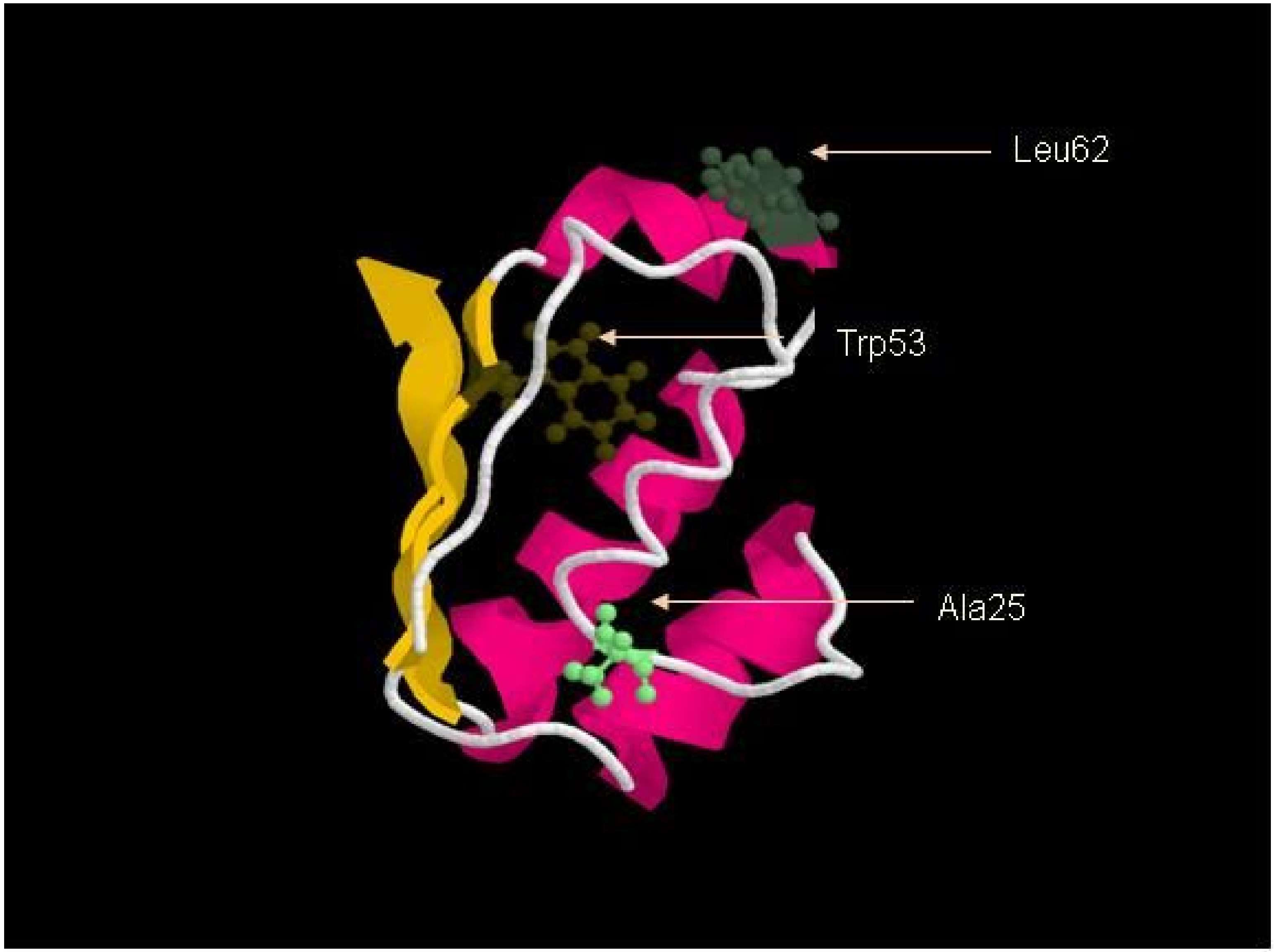
🟦 Unlabeled protein

🟥 Protein labeled with TNB

Lakshmikanth et al (2001) Nat. Struc. Biol. 8 , 799-804

FRET measurements of two core to surface distances

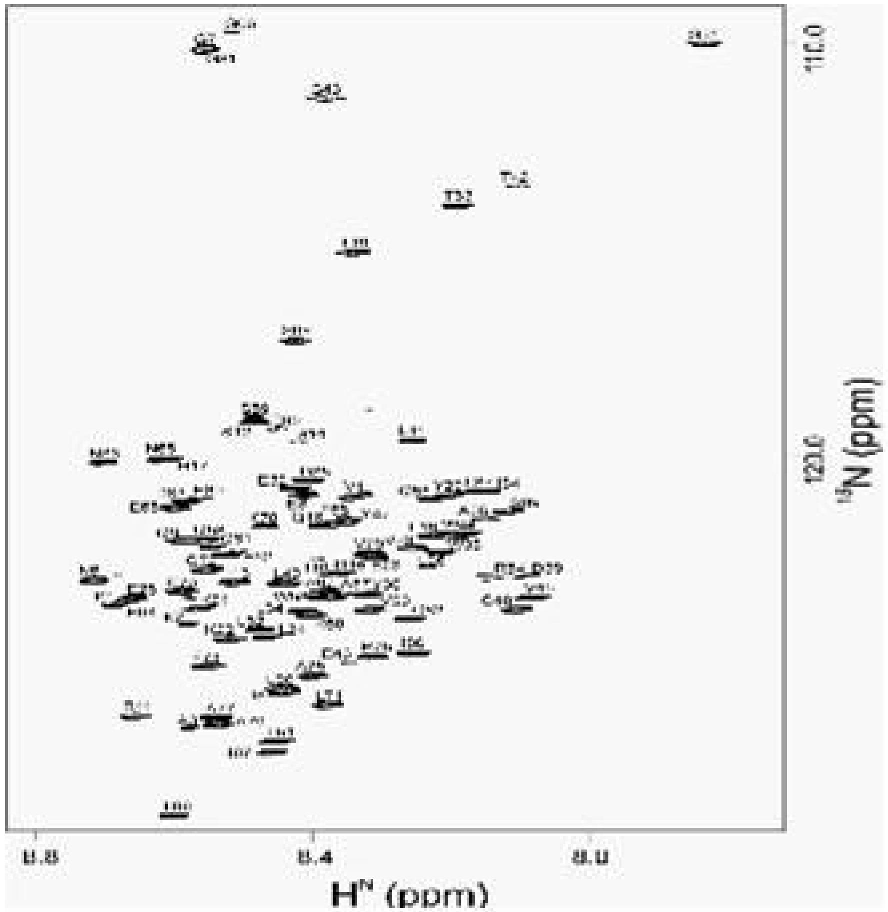
The Trp53-Cys25 and Trp53-Cys62 distances were determined



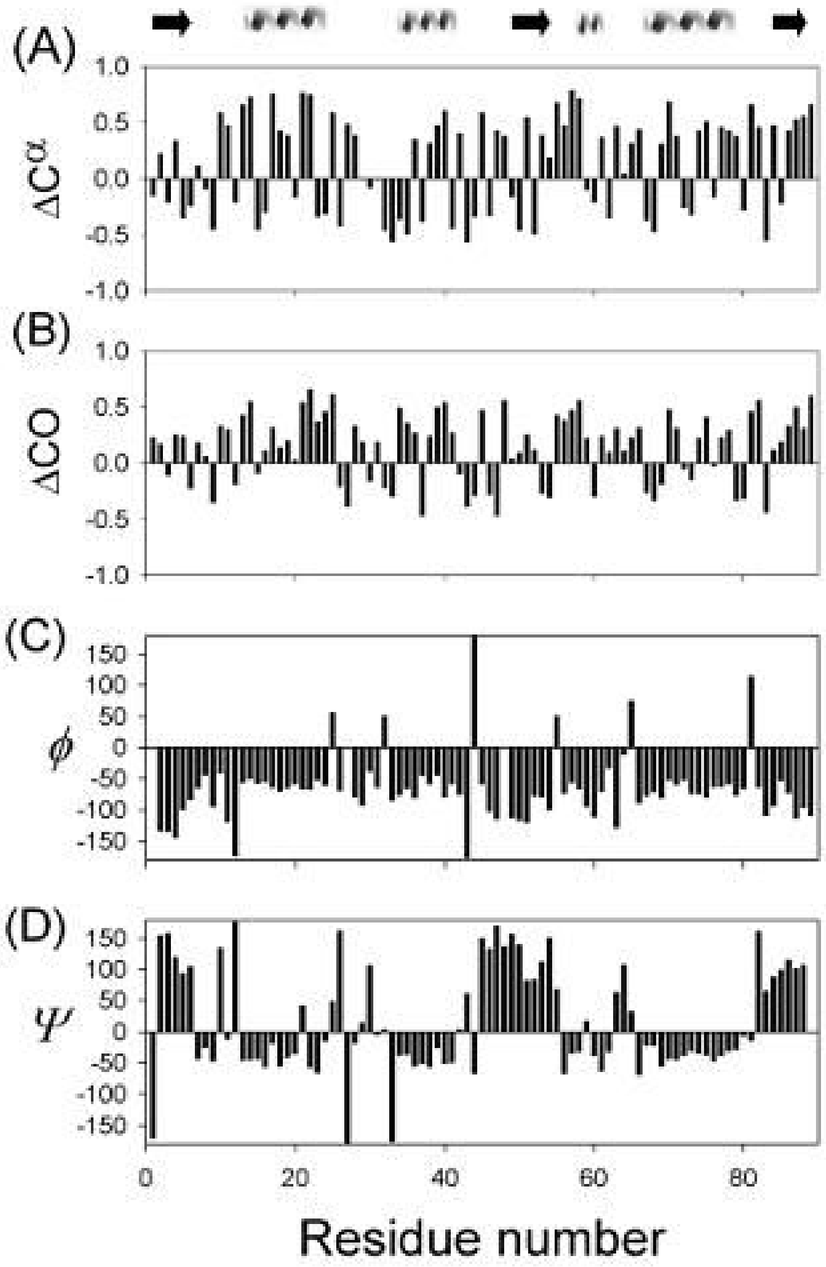
**Trp53 is completely buried in the core.
Cys25 and Cys62 are at surface positions.**

Barstar is devoid of native-like structure in 8 M urea

Bhavesh, Juneja, Udgaonkar & Hosur (2004) Protein Science



^{1}H-^{15}N HSQC spectrum of barstar in 8 M urea at pH 6.5 and 25 °C.



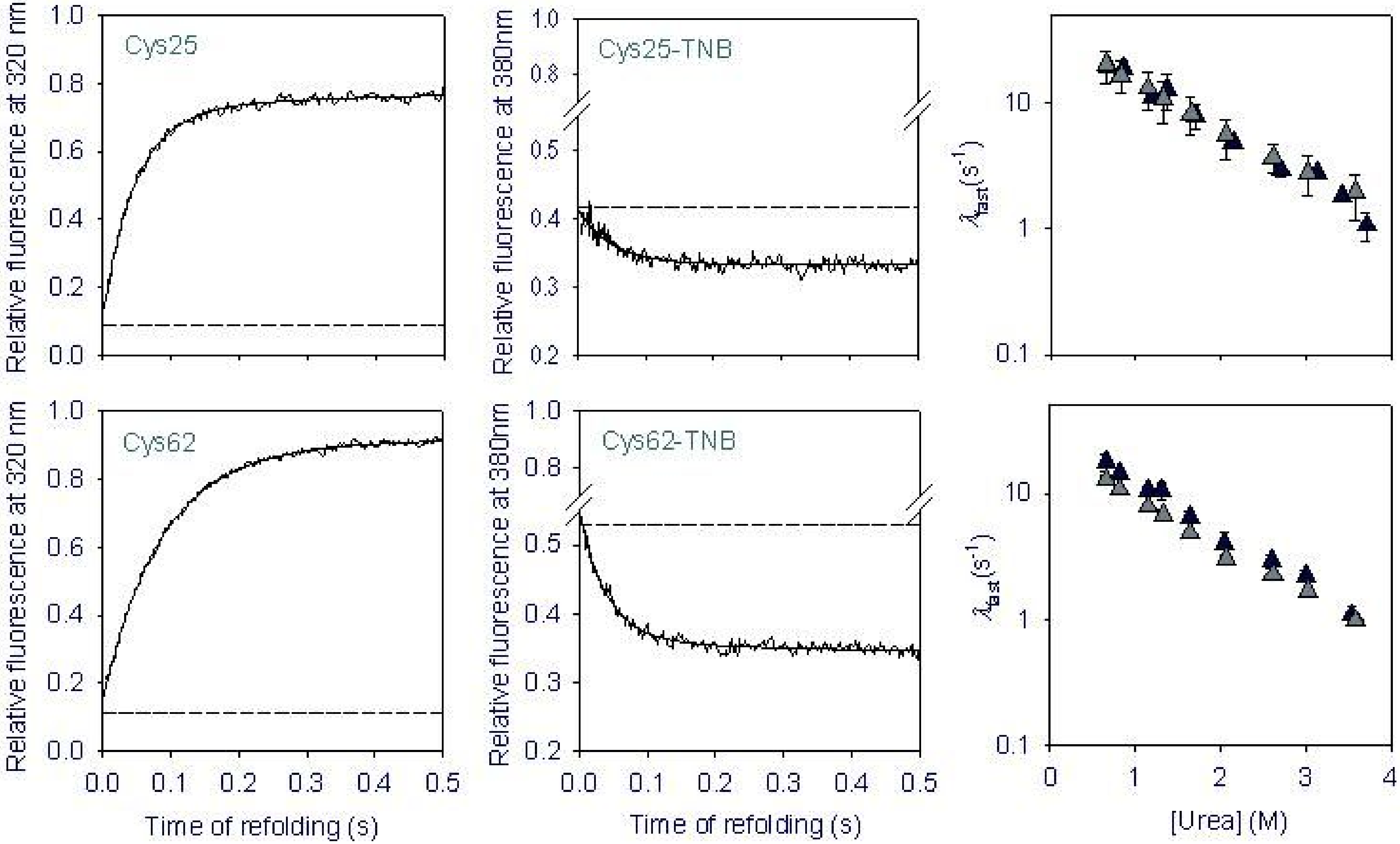
Secondary chemical shifts

Kinetics of folding of unlabeled and labeled proteins

Trp fluorescence

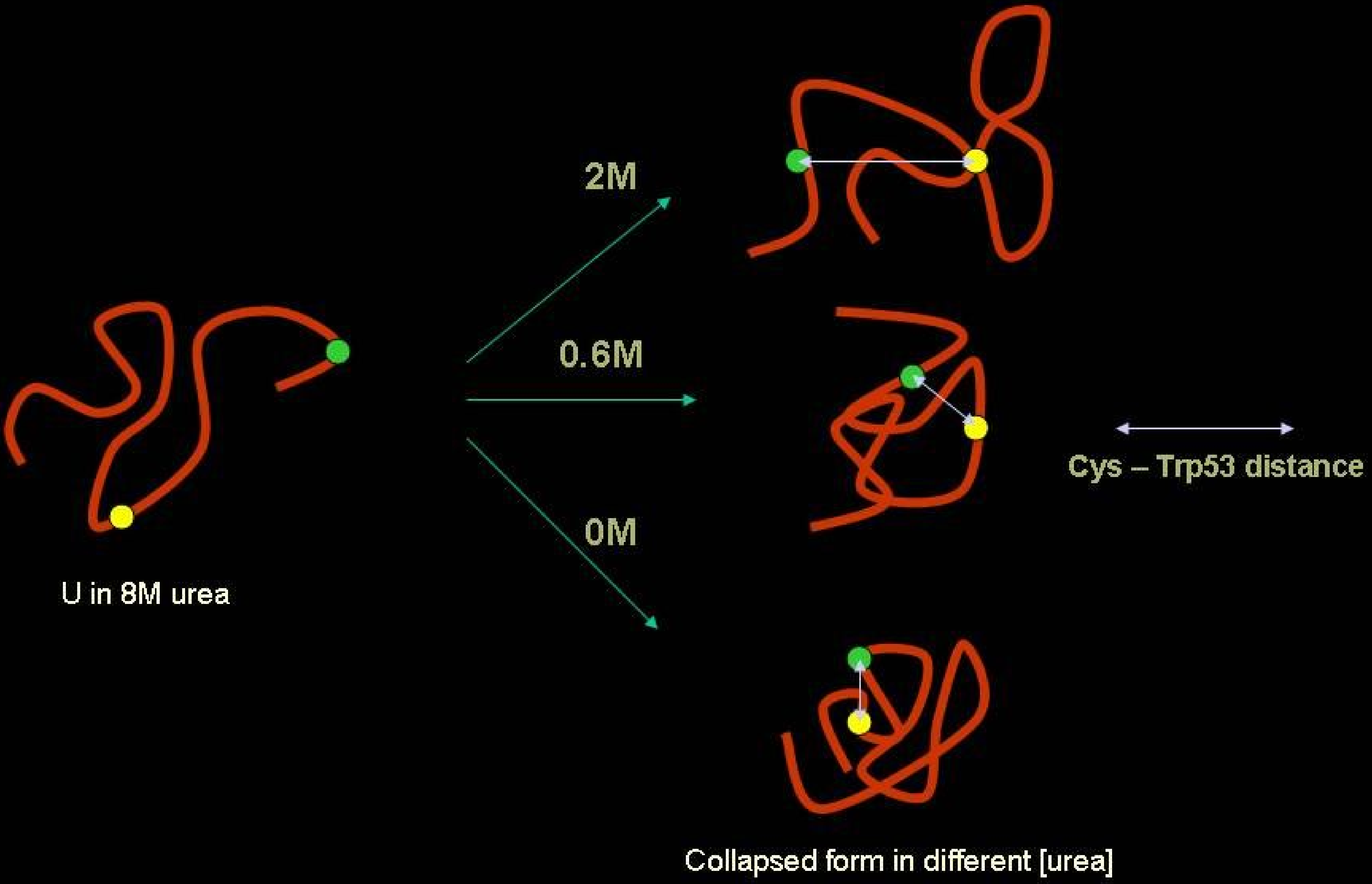
FRET efficiency
(1–E)

Rate constants

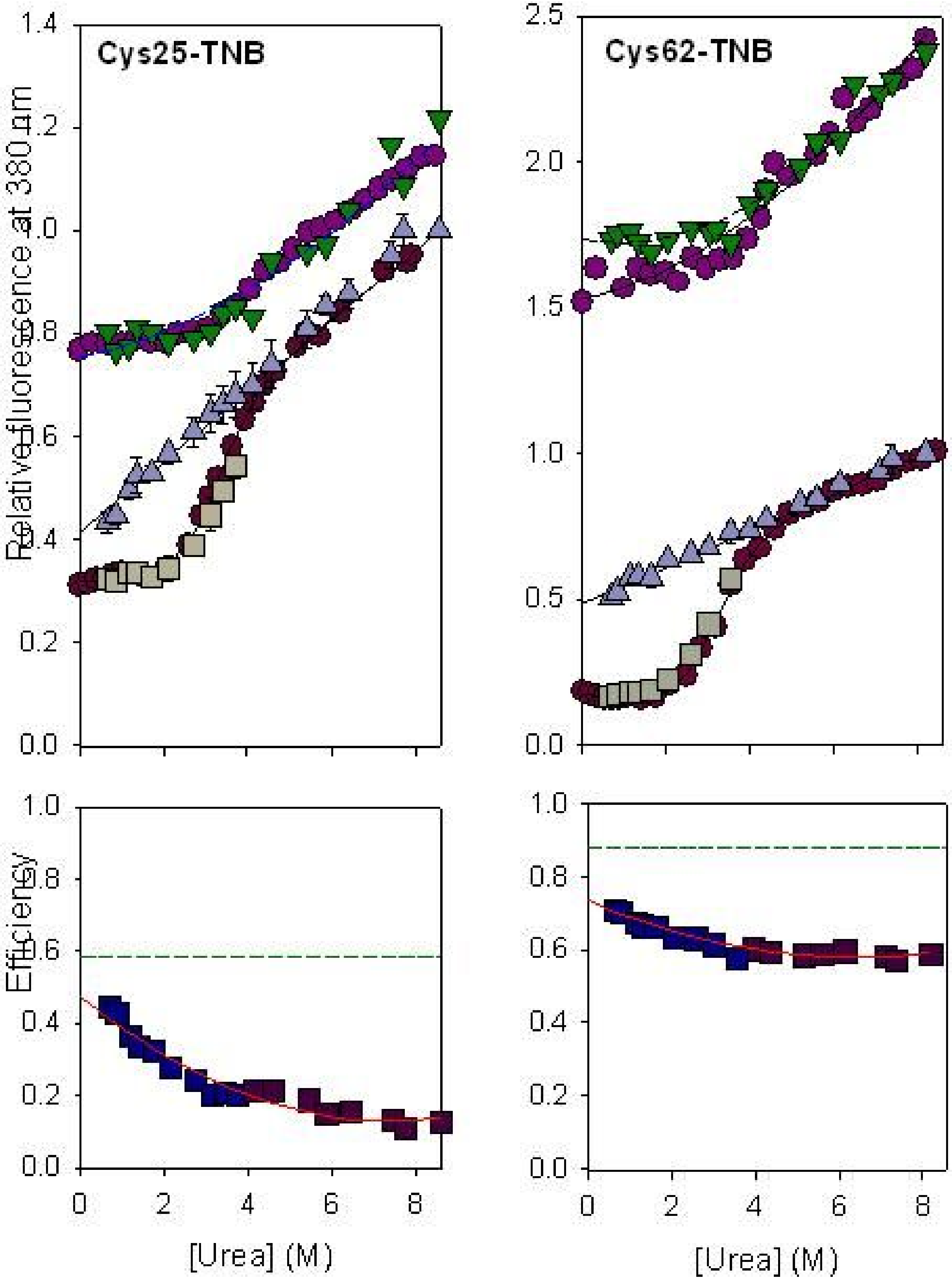


How does the collapsed state appear in different conditions?

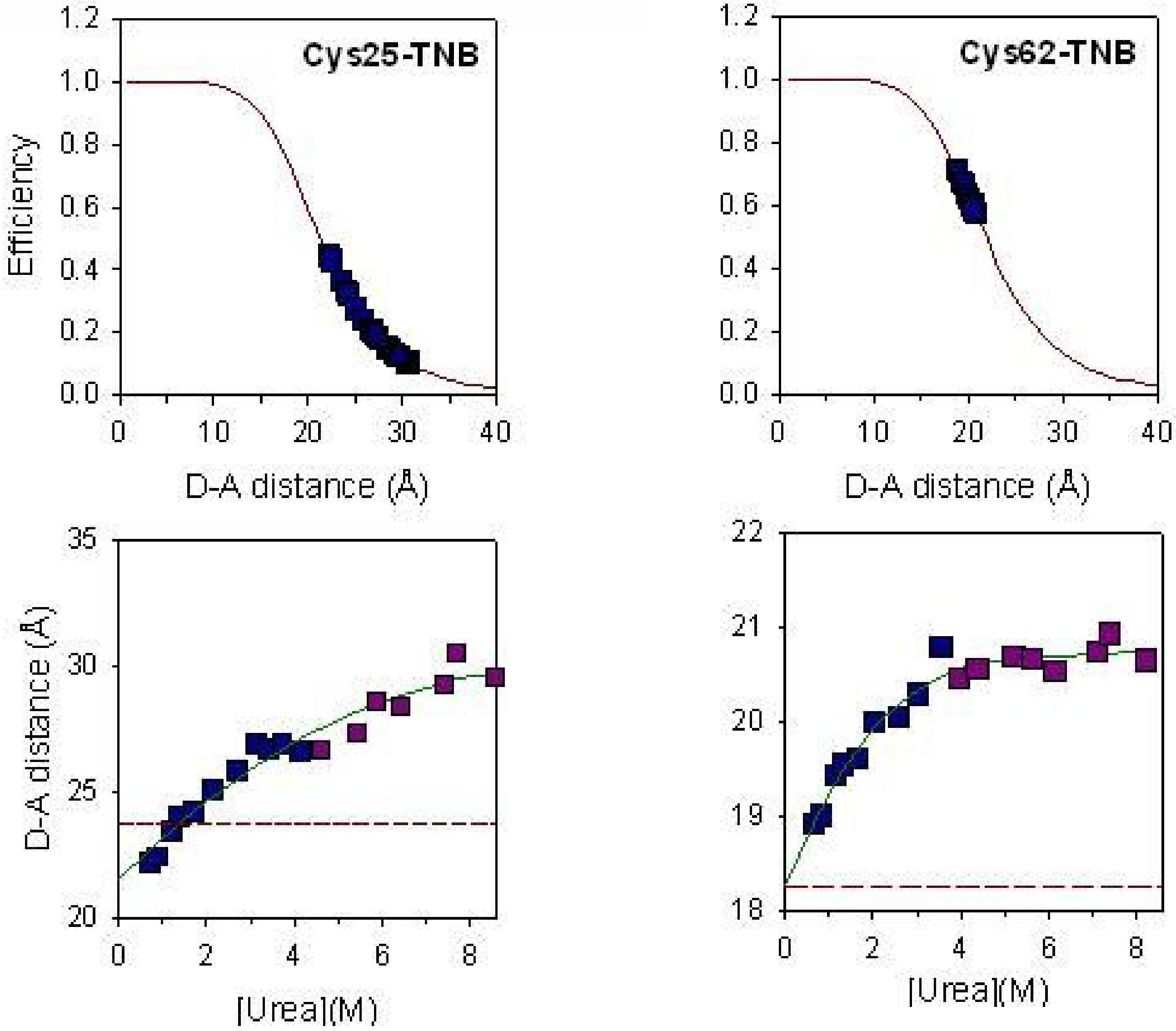
FRET measurements can give information on the separation of different regions in the collapsed state at 1 ms of folding



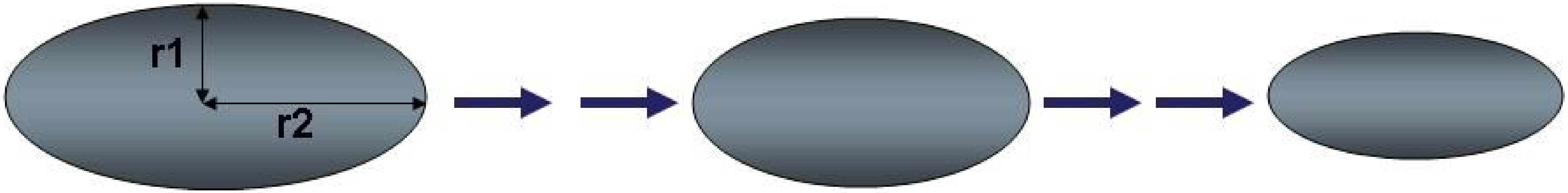
FRET efficiencies in the burst phase products



Intra-molecular distances determined by FRET

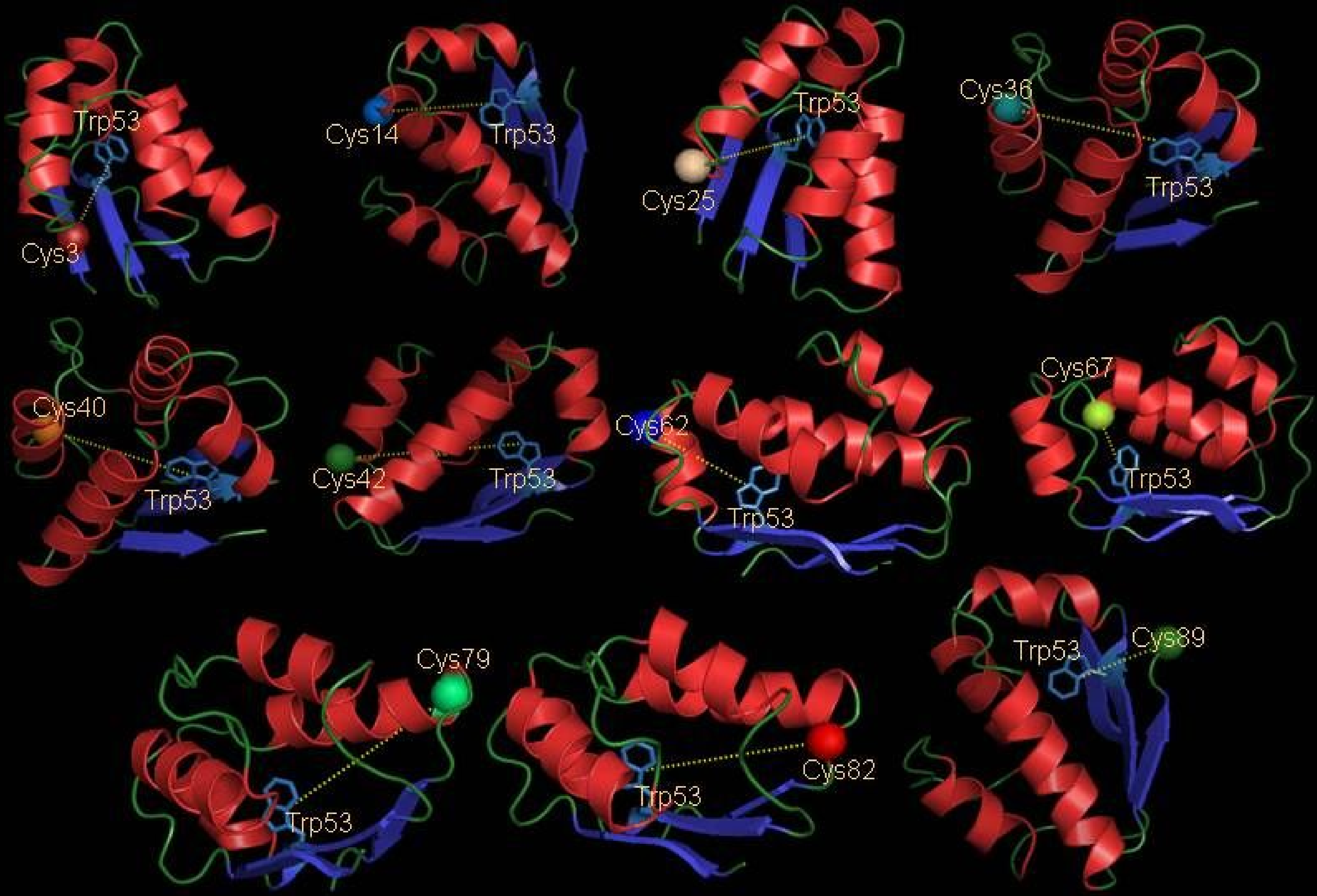


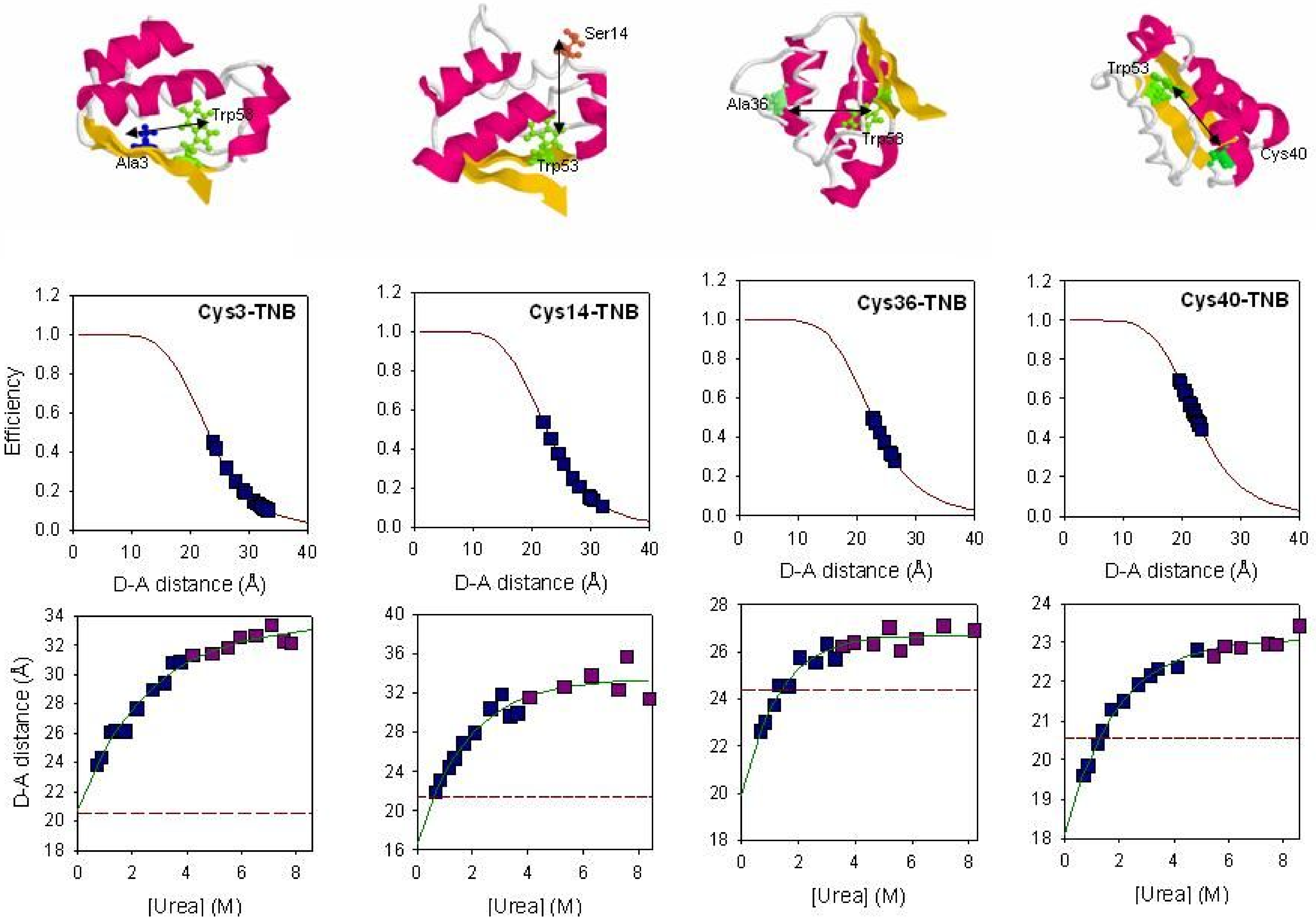
Initial collapse of the polypeptide chain is a gradual, continuous process

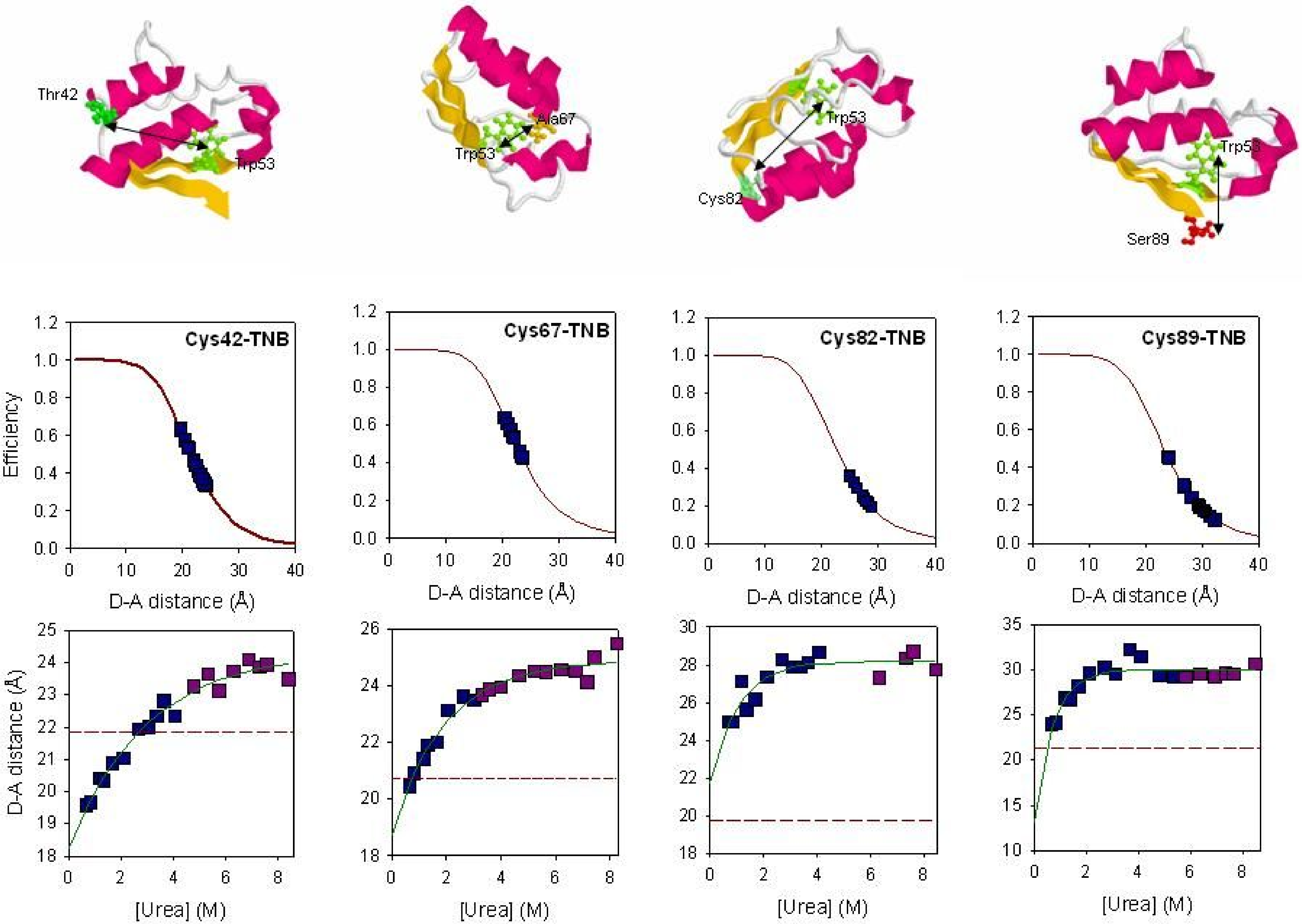


Sinha, K.K. & Udgaonkar, J.B. (2005) J. Mol. Biol. 353, 704-718.

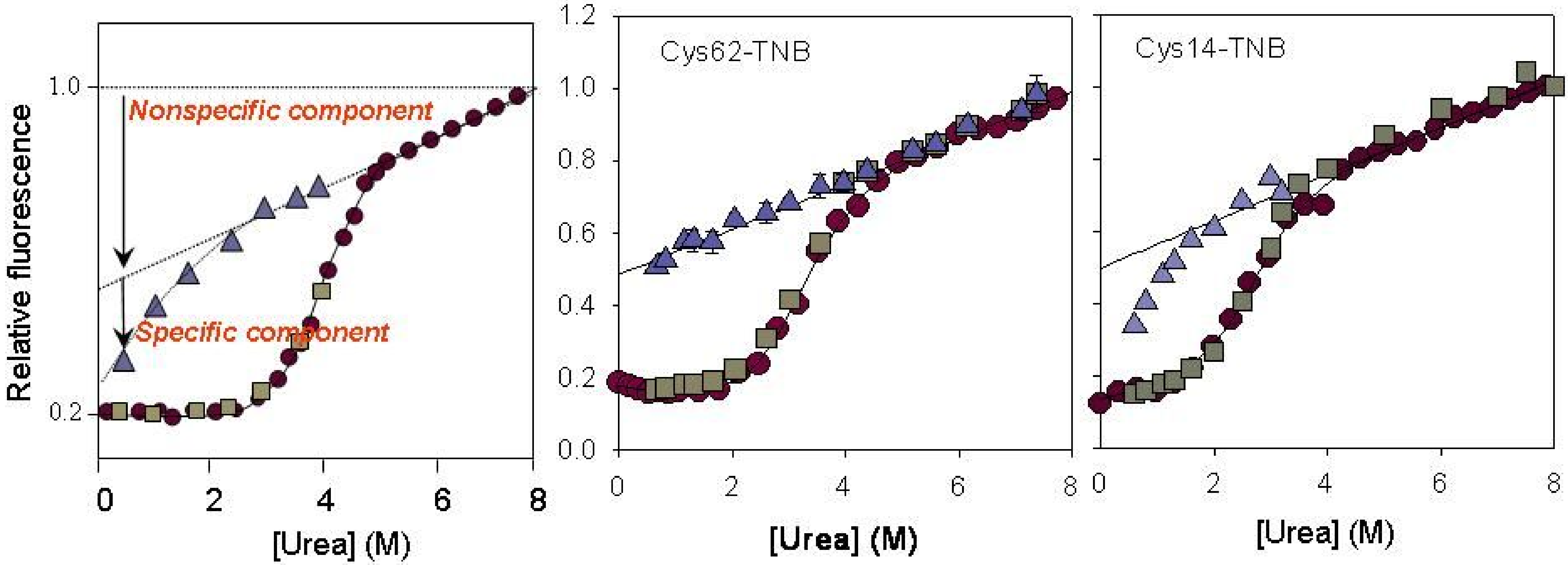
Different single cysteine, single tryptophan containing variants used for FRET




Ala3
Trp53
Ser14
Trp53
Ala36
Trp53
Trp53
Cys40
Cys3-TNB
Cys14-TNB
Cys36-TNB
Cys40-TNB
Efficiency
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
D-A distance (Å)
D-A distance (Å)
34
32
30
28
26
24
22
20
18
40
36
32
28
24
20
16
28
26
24
22
20
18
24
23
22
21
20
19
18
0
2
4
6
8
[Urea] (M)


Thr42
Trp53
Ala67
Trp53
Trp53
Cys82
Trp53
Ser89
Cys42-TNB
Cys67-TNB
Cys82-TNB
Cys89-TNB
Efficiency
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
D-A distance (Å)
D-A distance (Å)
25
24
23
22
21
20
19
18
26
30
28
35
15
10
0
2
4
6
8
[Urea] (M)

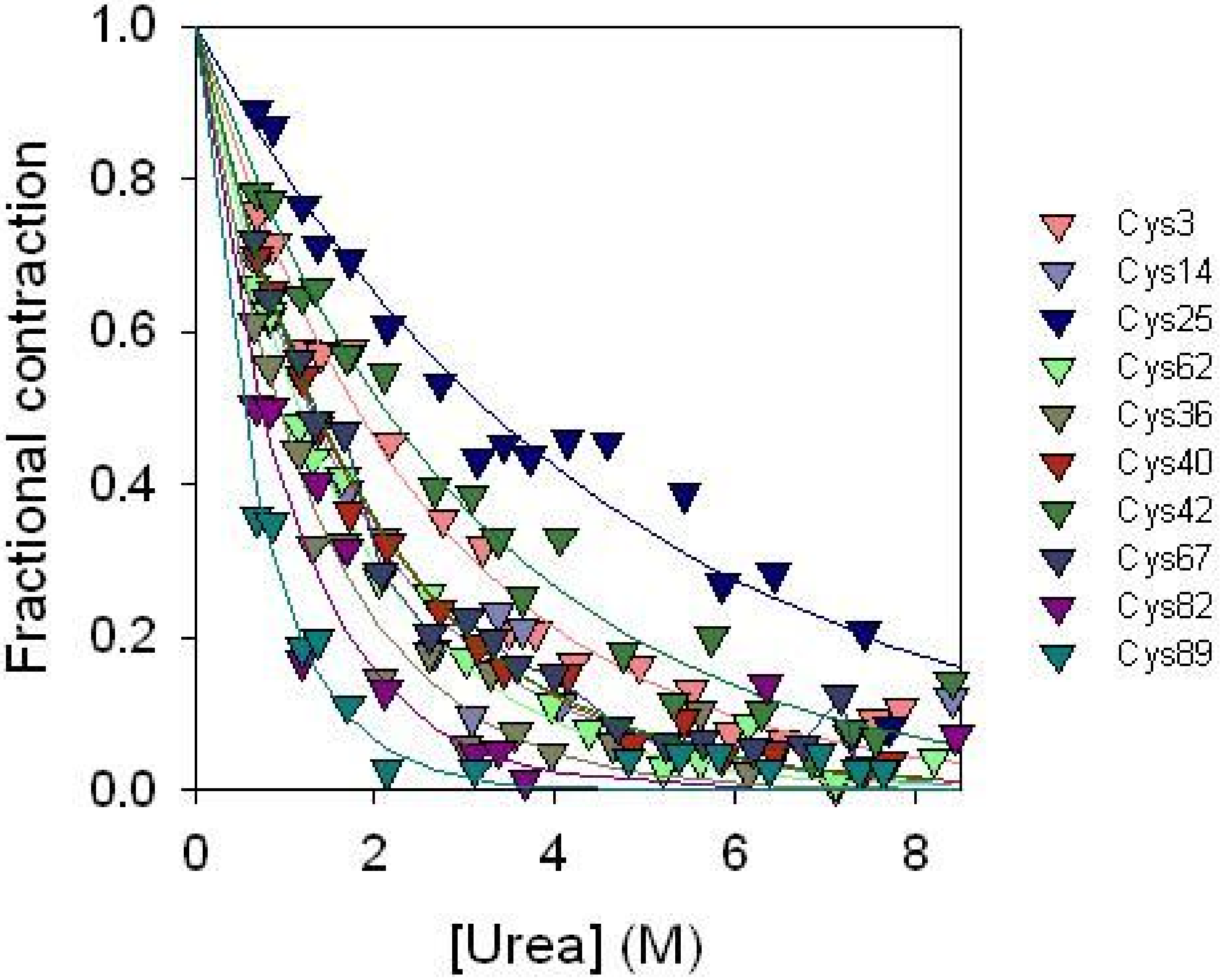
The initial collapse reaction in protein folding has nonspecific and specific components

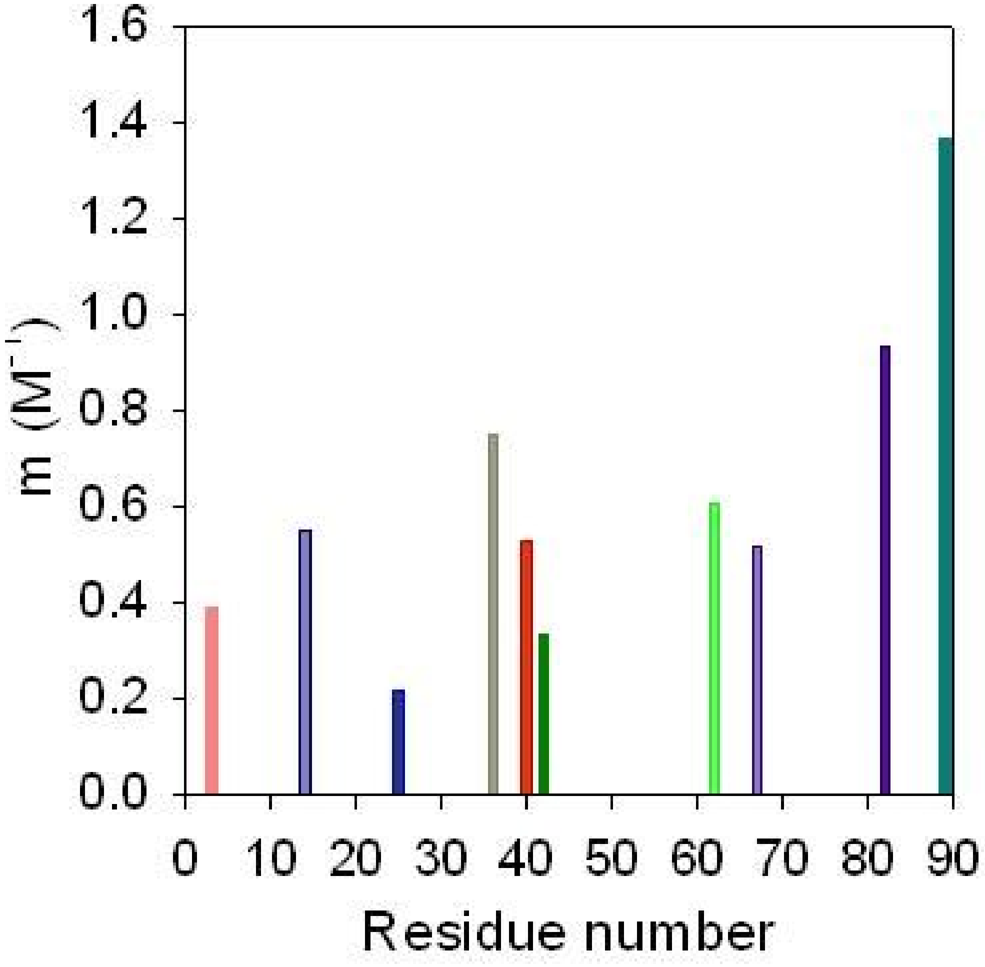


1. All intra-molecular distances show the same extent of non-specific chain contraction. This is attributed to solvent change, and is predicted by the behavior of unfolded protein.
2. Some but not all distances show a specific component to collapse: the extent of contraction is more than that induced by solvent change.

Collapse is continuous and not cooperative

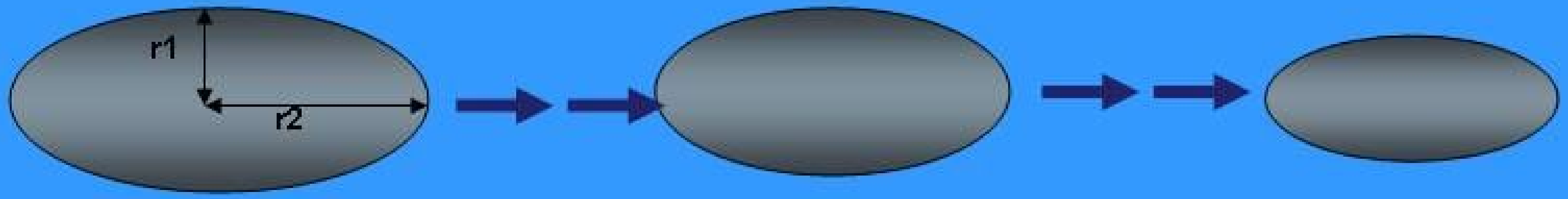
Collapse of different regions of the protein is not synchronized

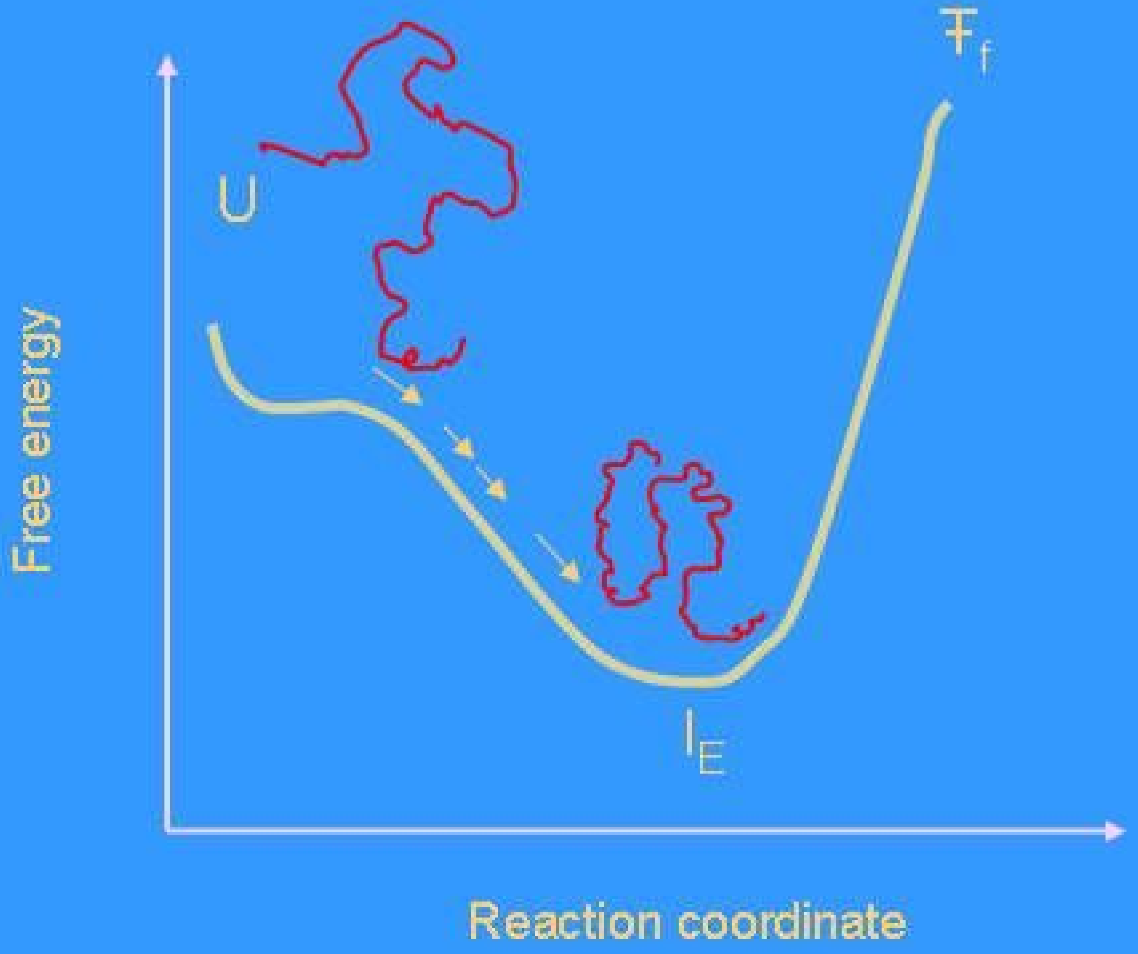




Kalyan K. Sinha and Jayant B. Udgaonkar
J. Mol. Biol. (2007) 370, 385–405

The unfolded polypeptide chain collapses in a continuous barrierless manner

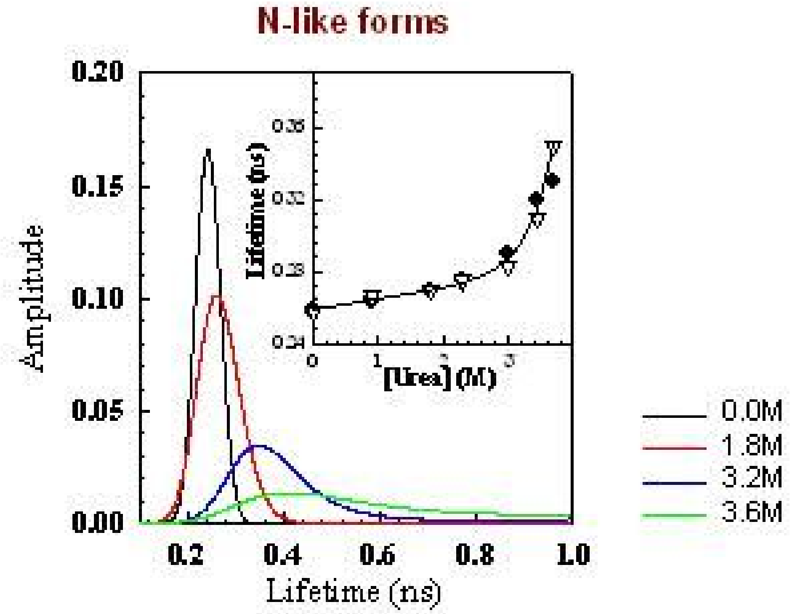


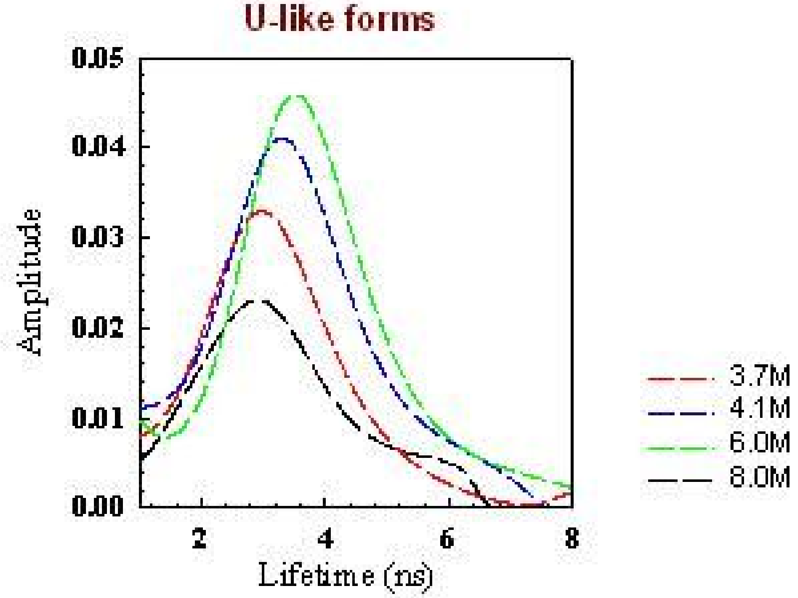


Sinha, K.K. & Udgaonkar, J.B. (2005) *J Mol Biol* **353**, 704-718
Sinha, K.K. & Udgaonkar, J.B. (2007) *J Mol Biol* **370**, 385-405

Incremental loss of structure during unfolding of barstar

Fluorescence lifetime distributions suggest that the protein expands continuously during unfolding





Unfolding occurs through a continuum of forms

<u>Higher-order transitions</u>

No energy barriers separate the N-like forms

N ↔ N1 ↔ N2 ↔ N3 ↔. . . . . .

No energy barriers separate the U-like forms

U ↔ U1 ↔ U2 ↔ U3 ↔. . . . . .

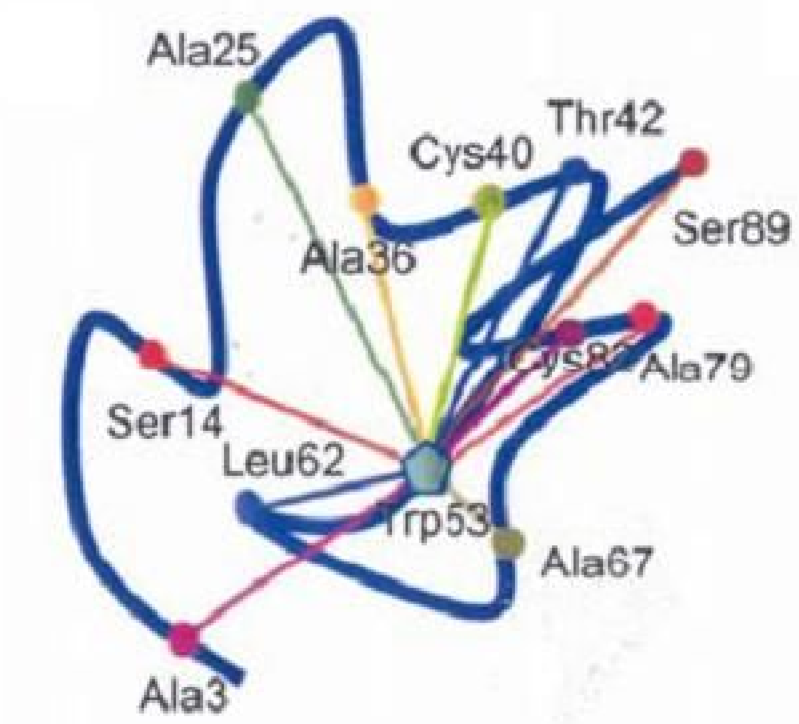
<u>First order transition</u>

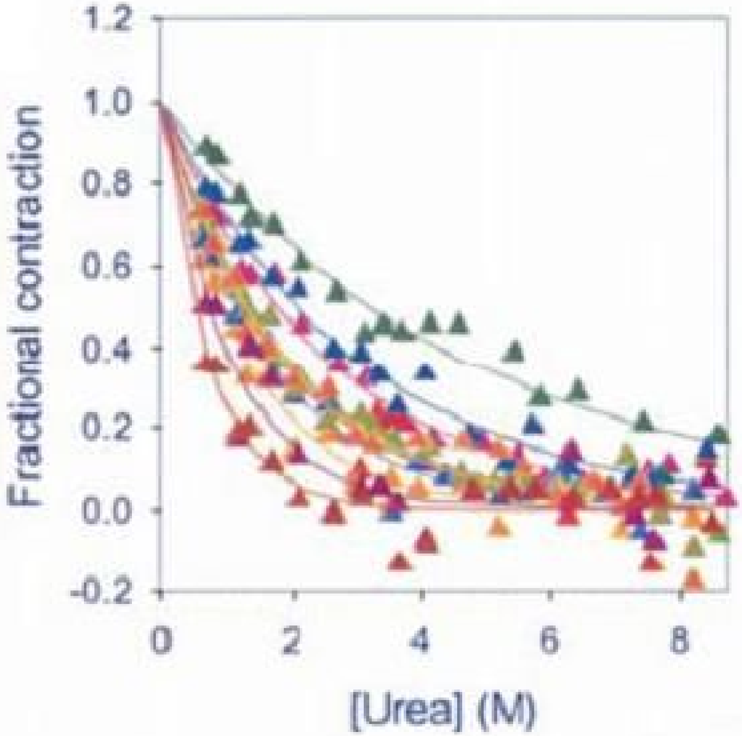
The energy barrier for unfolding occurs between the N-like and U-like forms

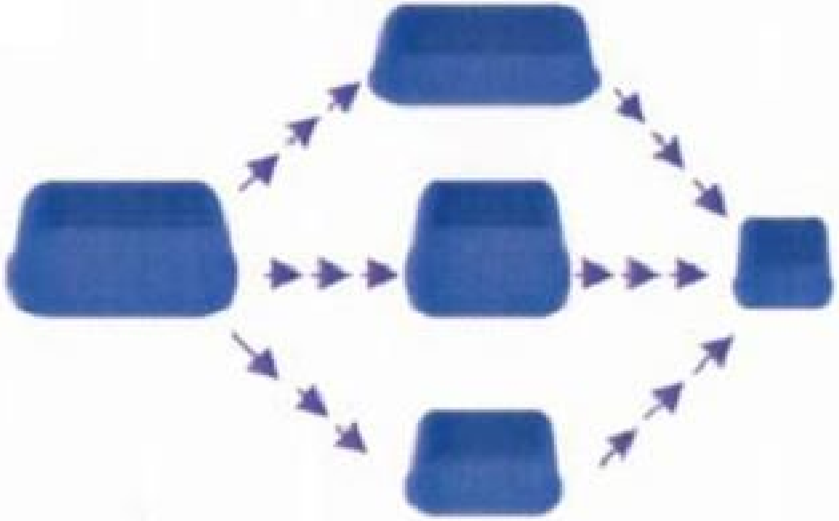
[N-like forms] ↔ [U-like forms]

Lakshmikanth et al (2001) Nat. Structural Biol. 8 , 799-804

Are multiple pathways operative during initial folding events?

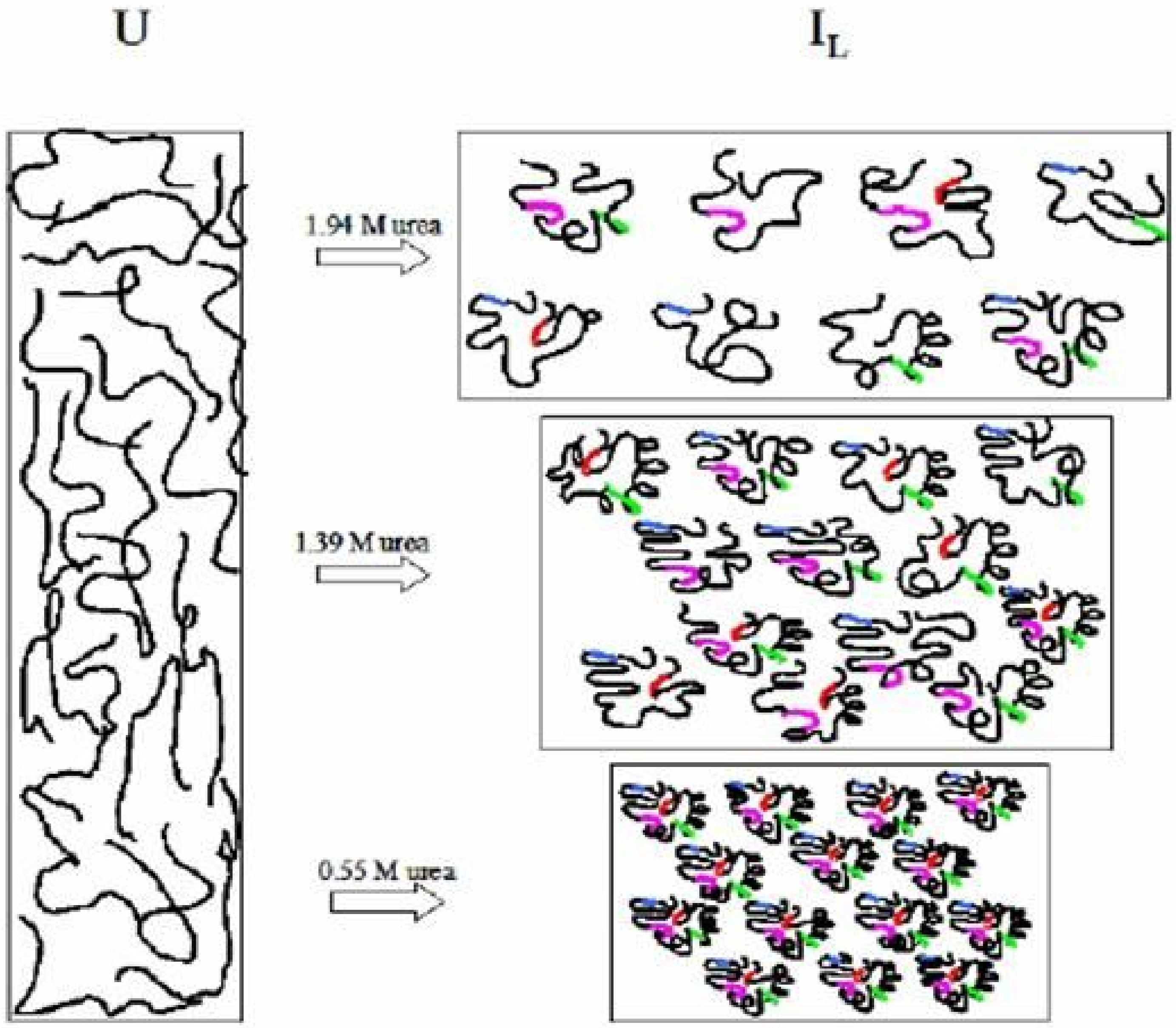






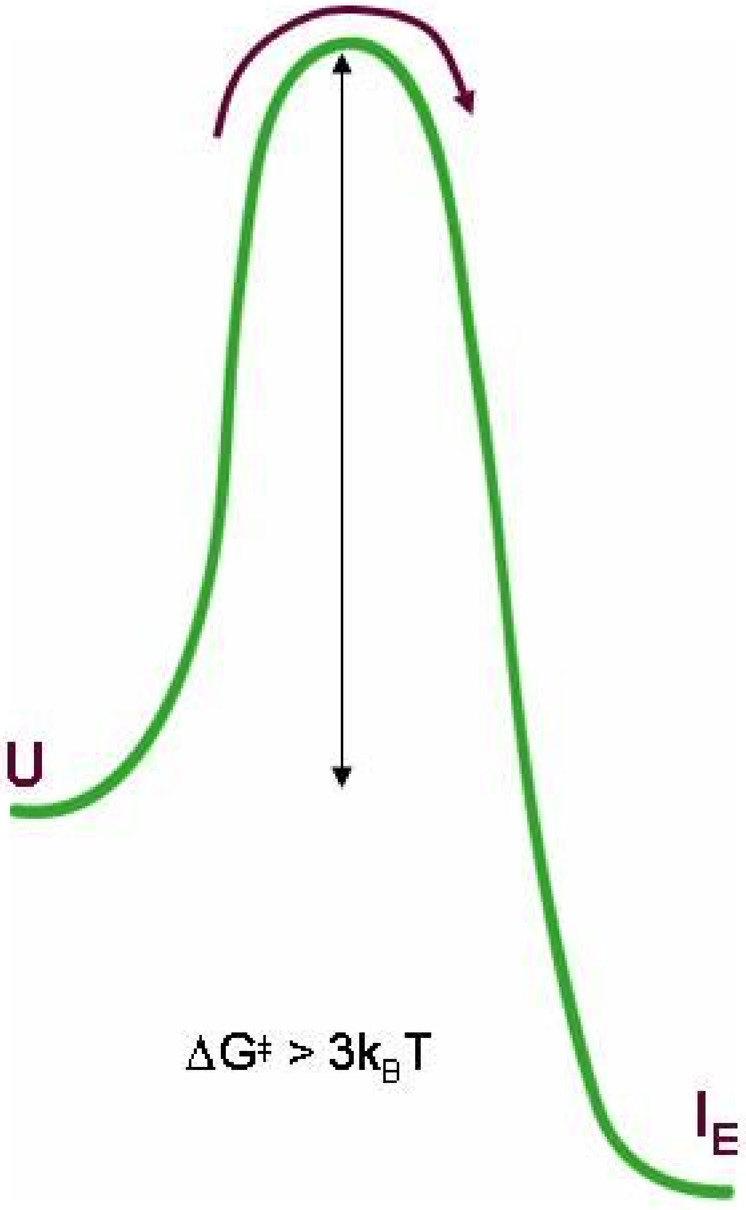
Udgaonkar (2008) Annu. Rev. Biophys.

Increasing stability decreases conformational heterogeneity in a protein folding intermediate ensemble

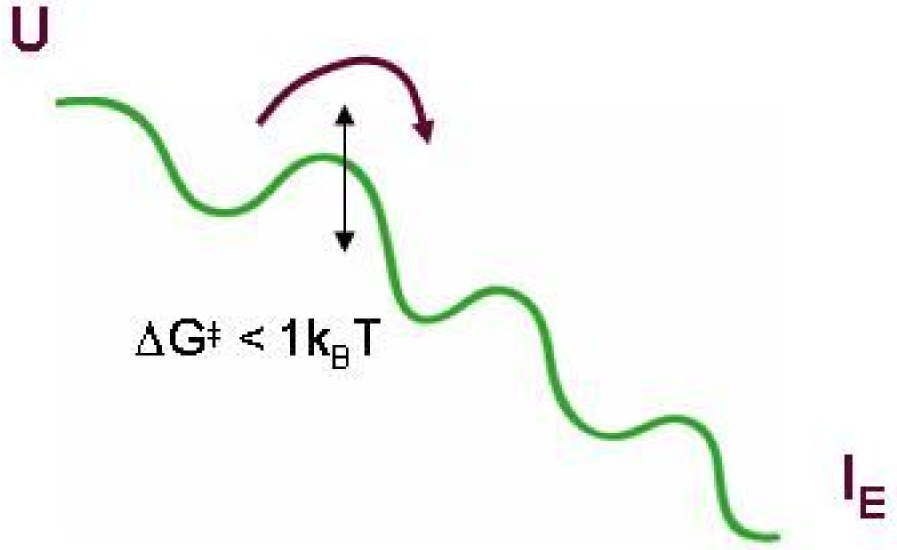


Sridevi et al. (2004) J. Mol. Biol 337, 699-711

Is there one dominant free energy barrier (or many barriers)?

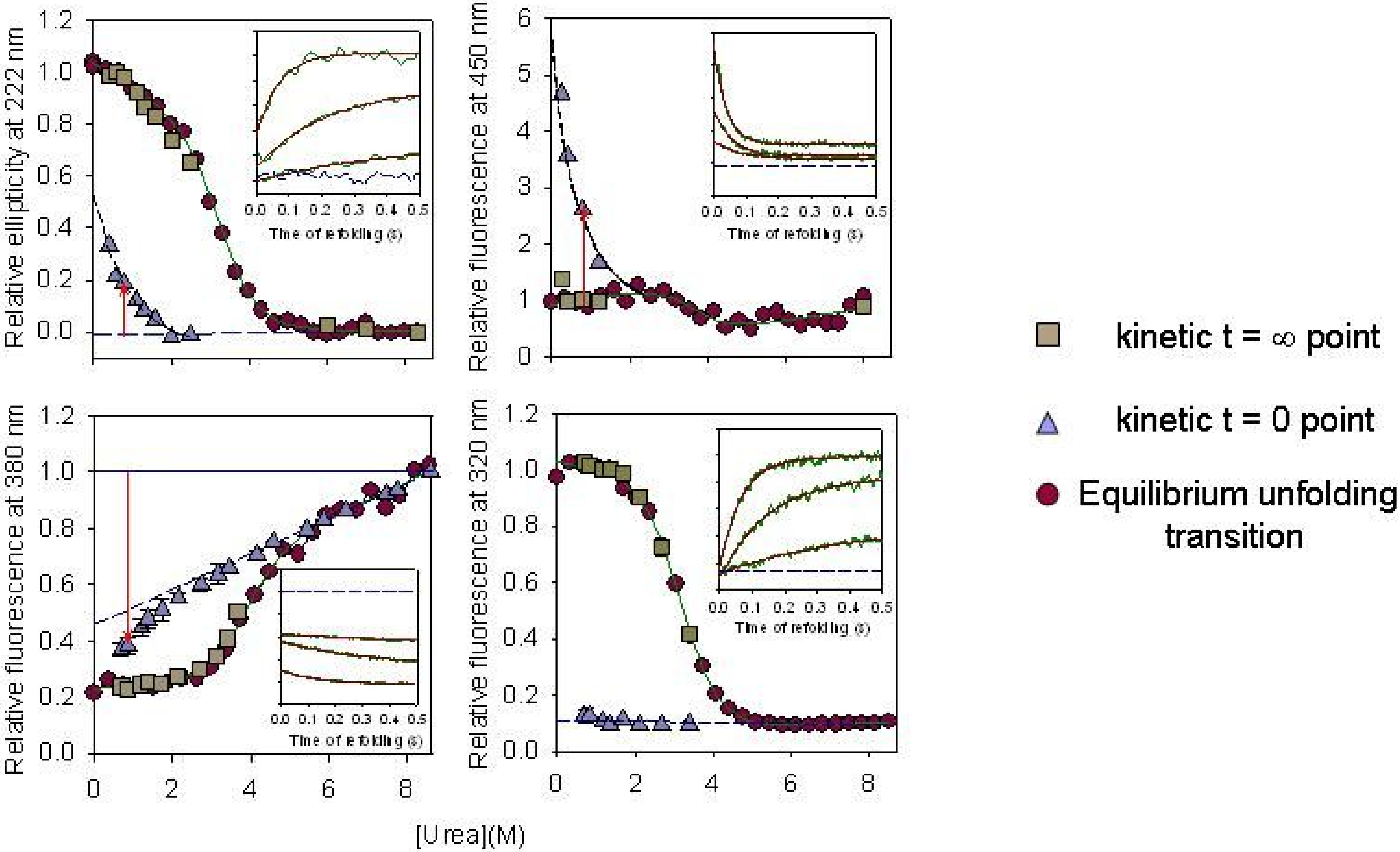


Barrier crossing
one dominant energy barrier



Barrierless: several small steps

The product of the sub-ms folding reaction possesses specific structure

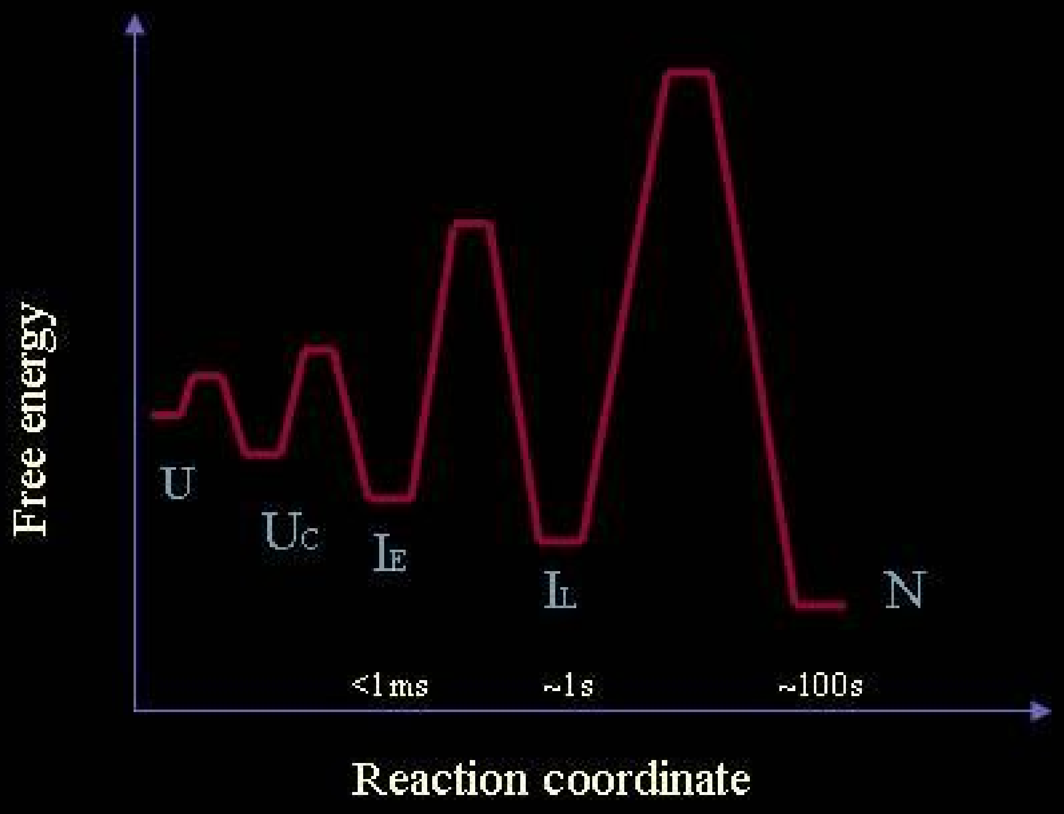


Sinha and Udgaonkar (2008) PNAS

Detection of specific structure during the initial folding reaction

Folding in strongly stabilizing conditions

The product of sub-millisecond folding reaction possesses specific structure.

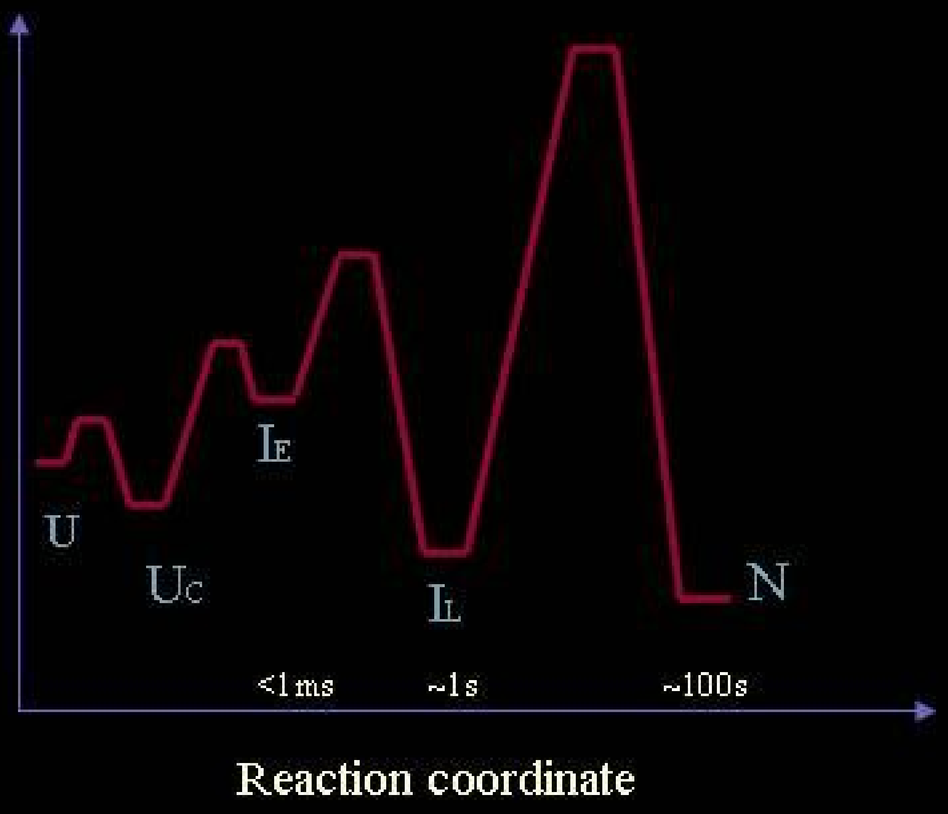


The specific intermediate I_E is populated at ~1ms of folding.

The collapsed globule U_C is not detected.

Folding in marginally stabilizing conditions

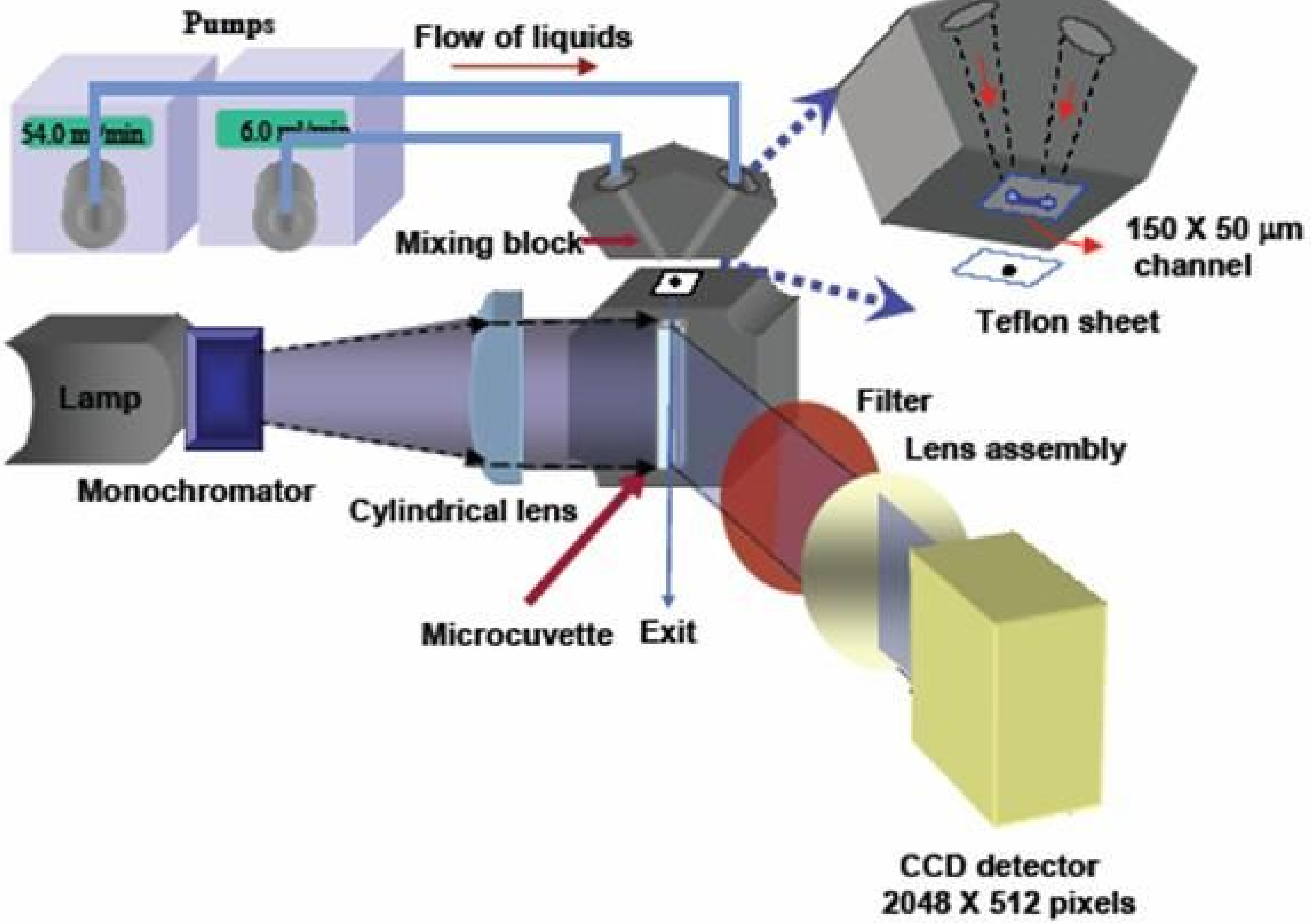
The product of sub-millisecond folding reaction does not possess specific structure.

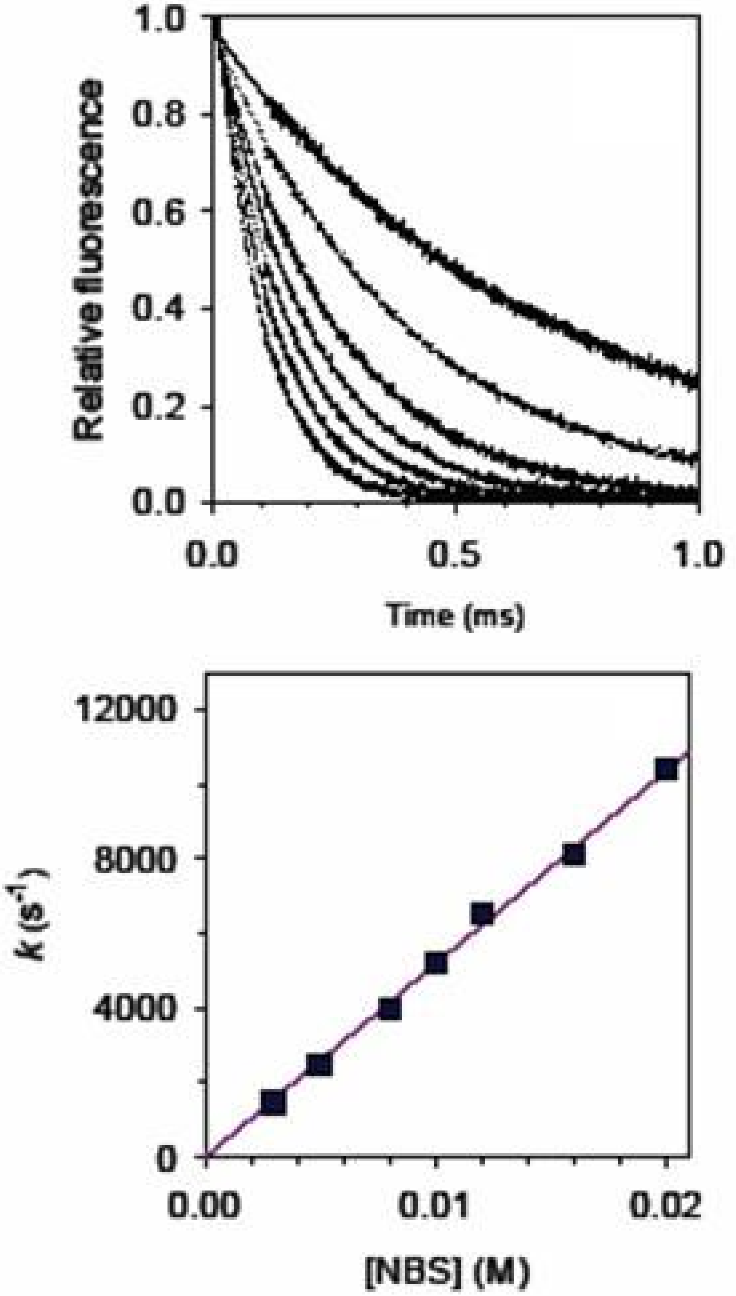


The collapsed globule U_C is populated at ~1ms of folding.

The specific intermediate I_E is not detected.

Measurement of sub-millisecond folding kinetics using a continuous flow mixer

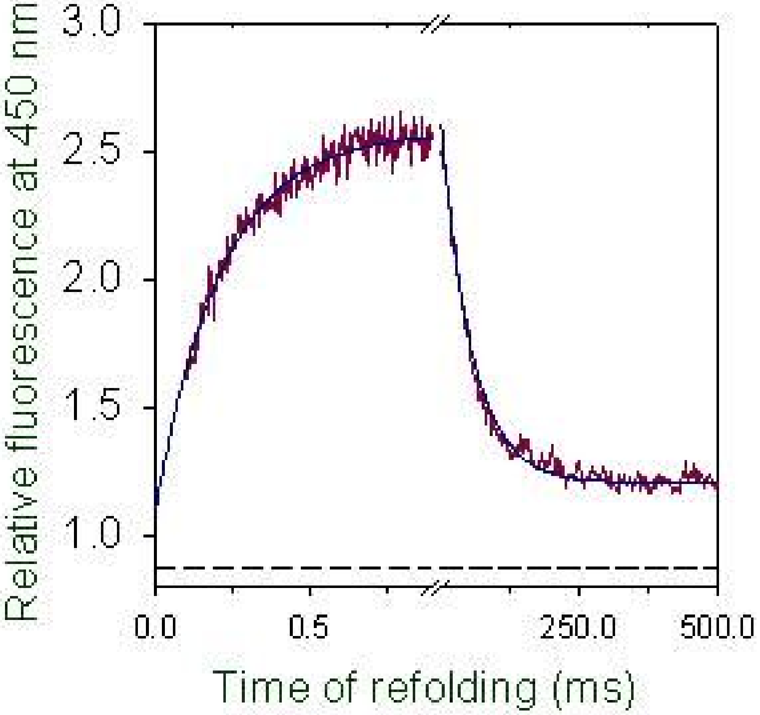




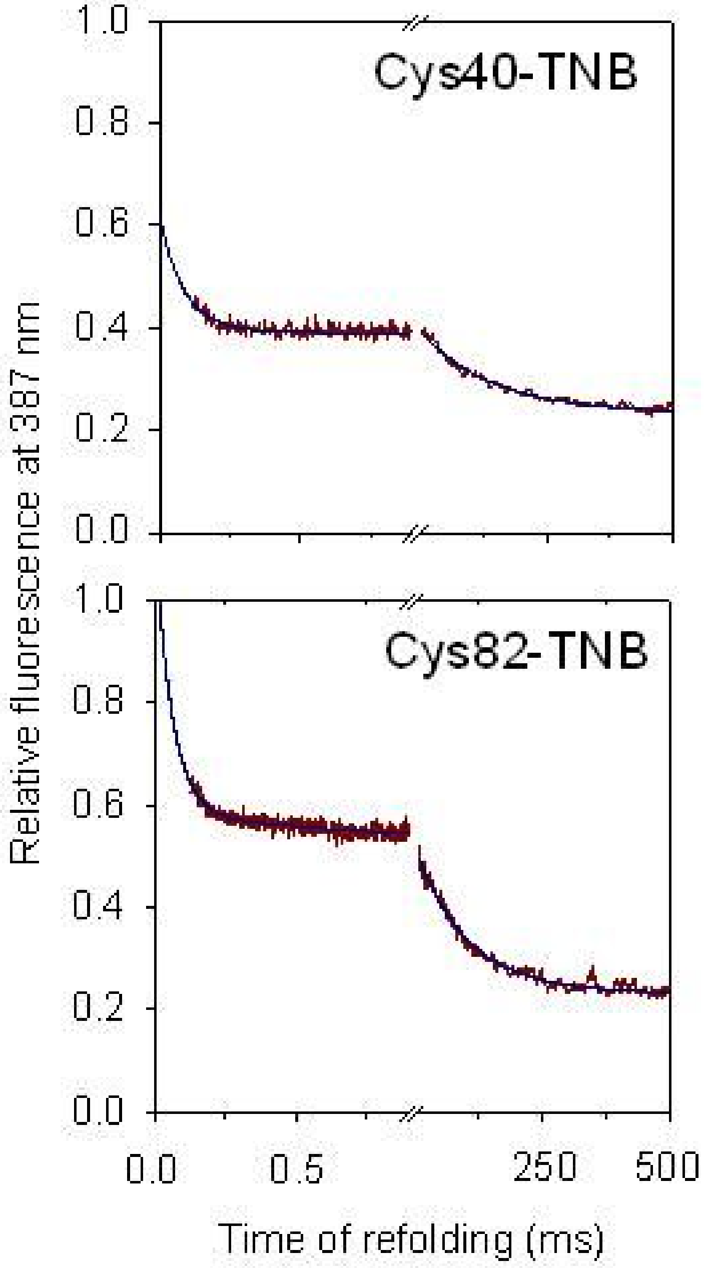
Sinha and Udgaonkar (2008) PNAS

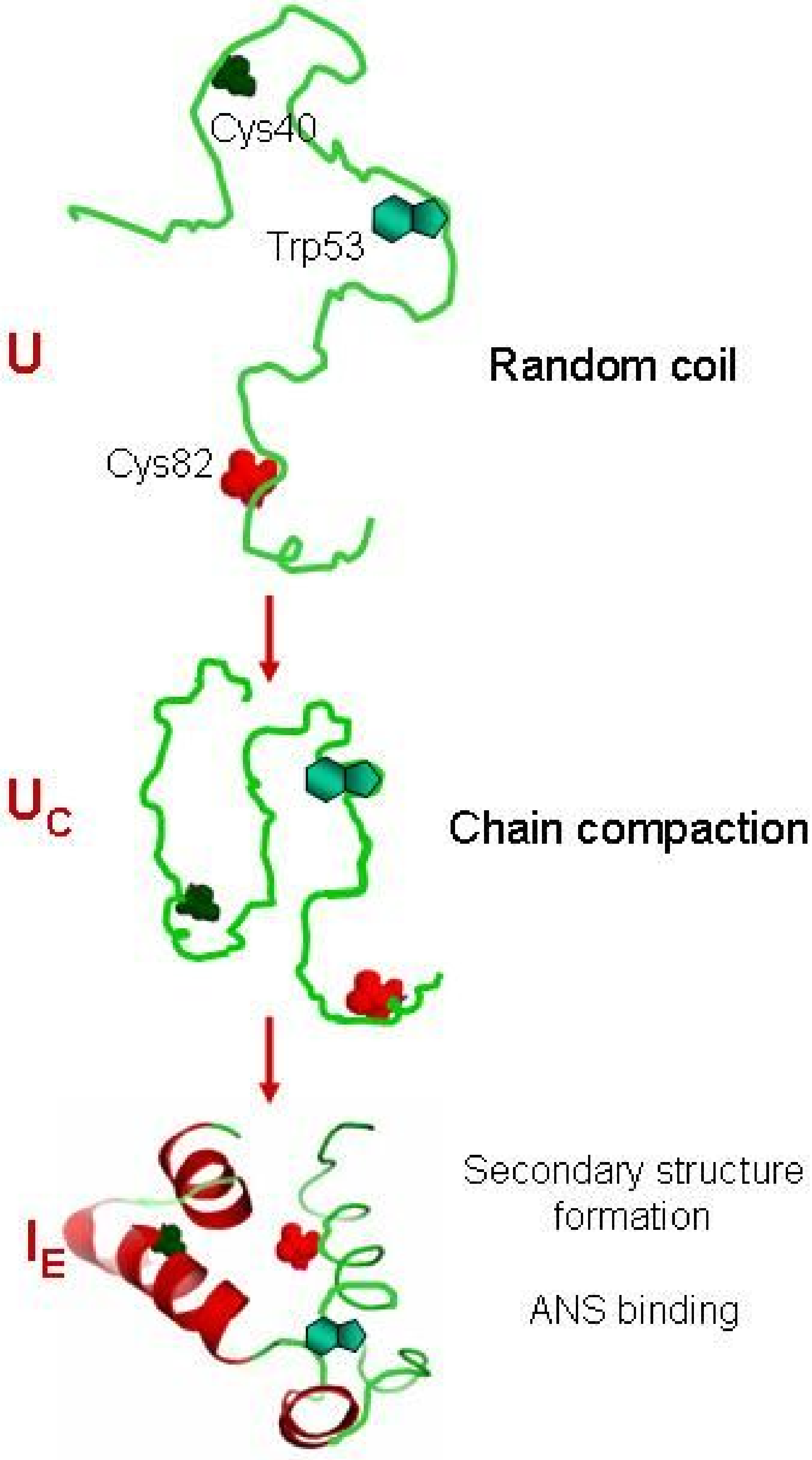
Sub-ms contraction of the polypeptide chain occurs faster than hydrophobic clustering

ANS-fluorescence



FRET





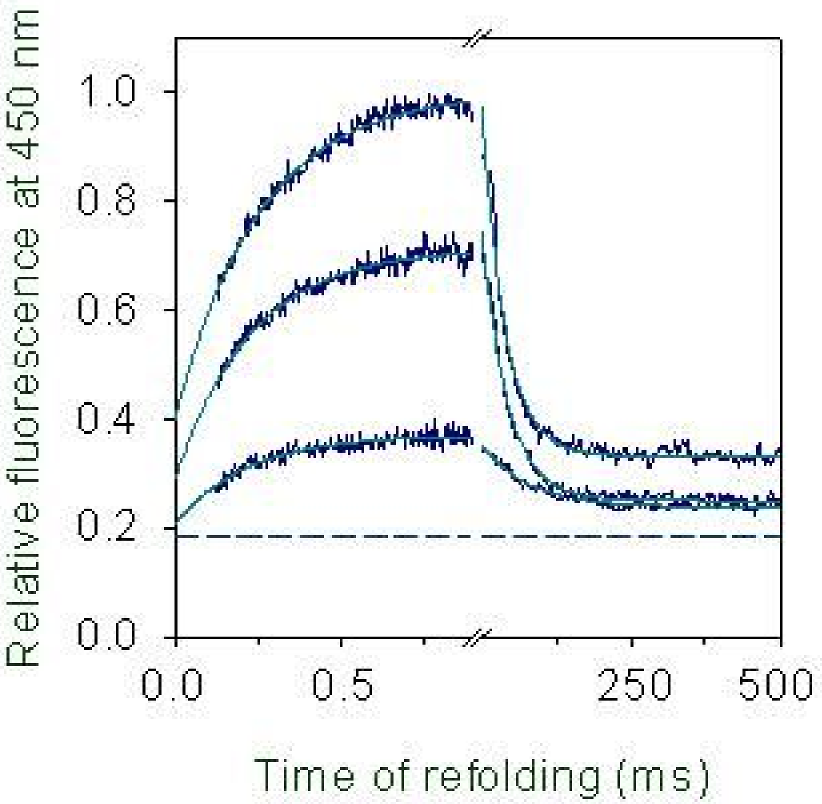
Random coil

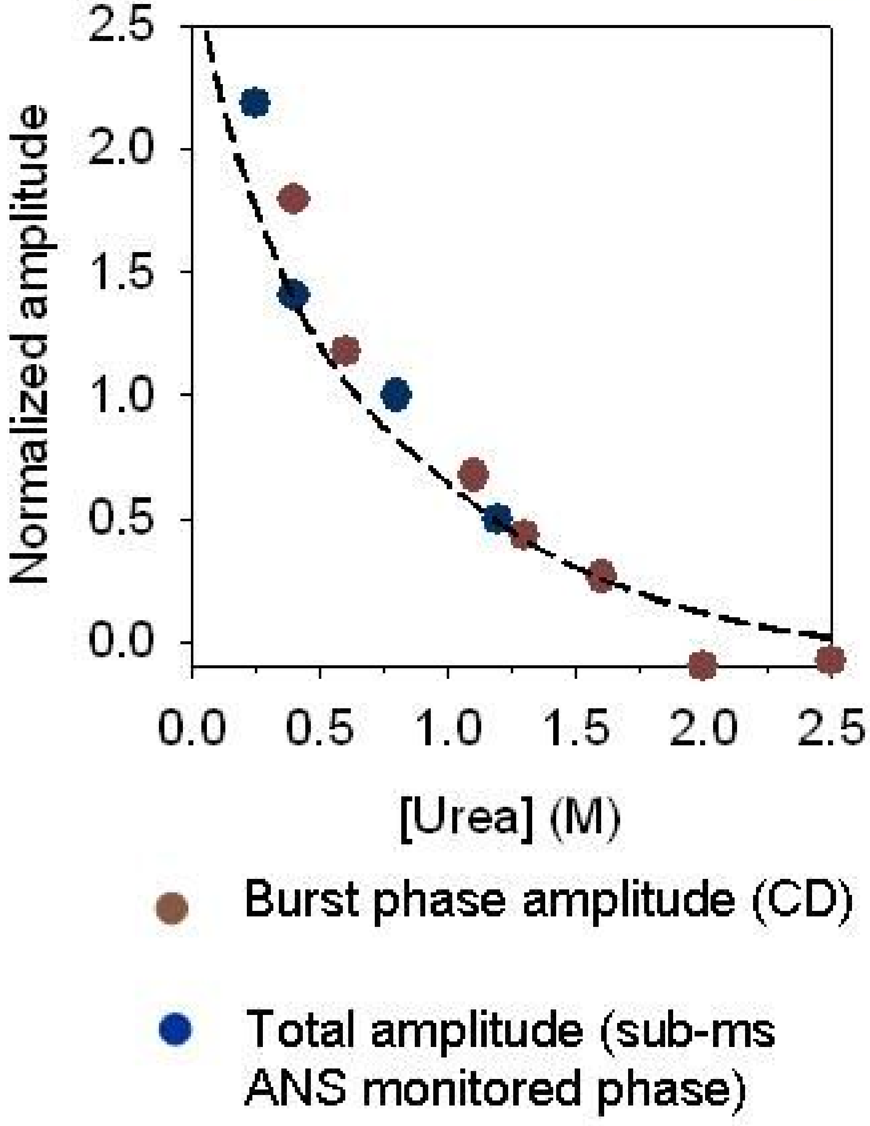
Chain compaction

Secondary structure formation

ANS binding

ANS binding reports on the formation of I_E



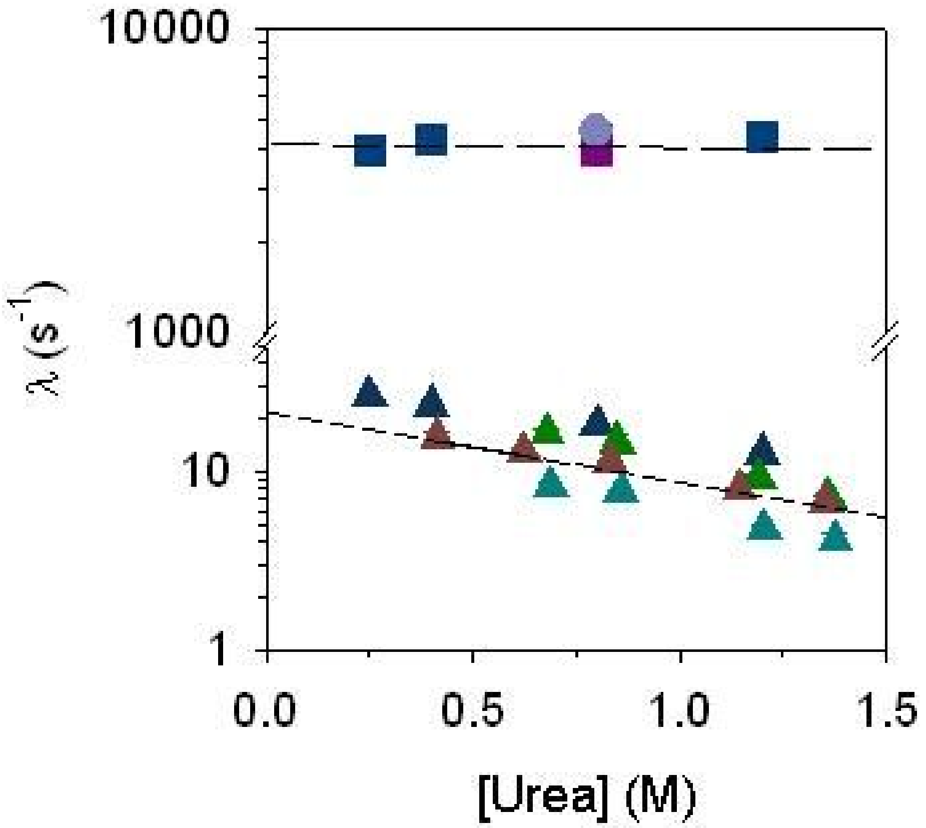


Burst phase amplitude of the CD-monitored kinetics changes in parallel with the amplitude of the ANS fluorescence monitored sub-ms kinetics.

The dependences of amplitudes on [urea] are non-sigmoidal and gradual.

Sinha and Udgaonkar (2008) PNAS

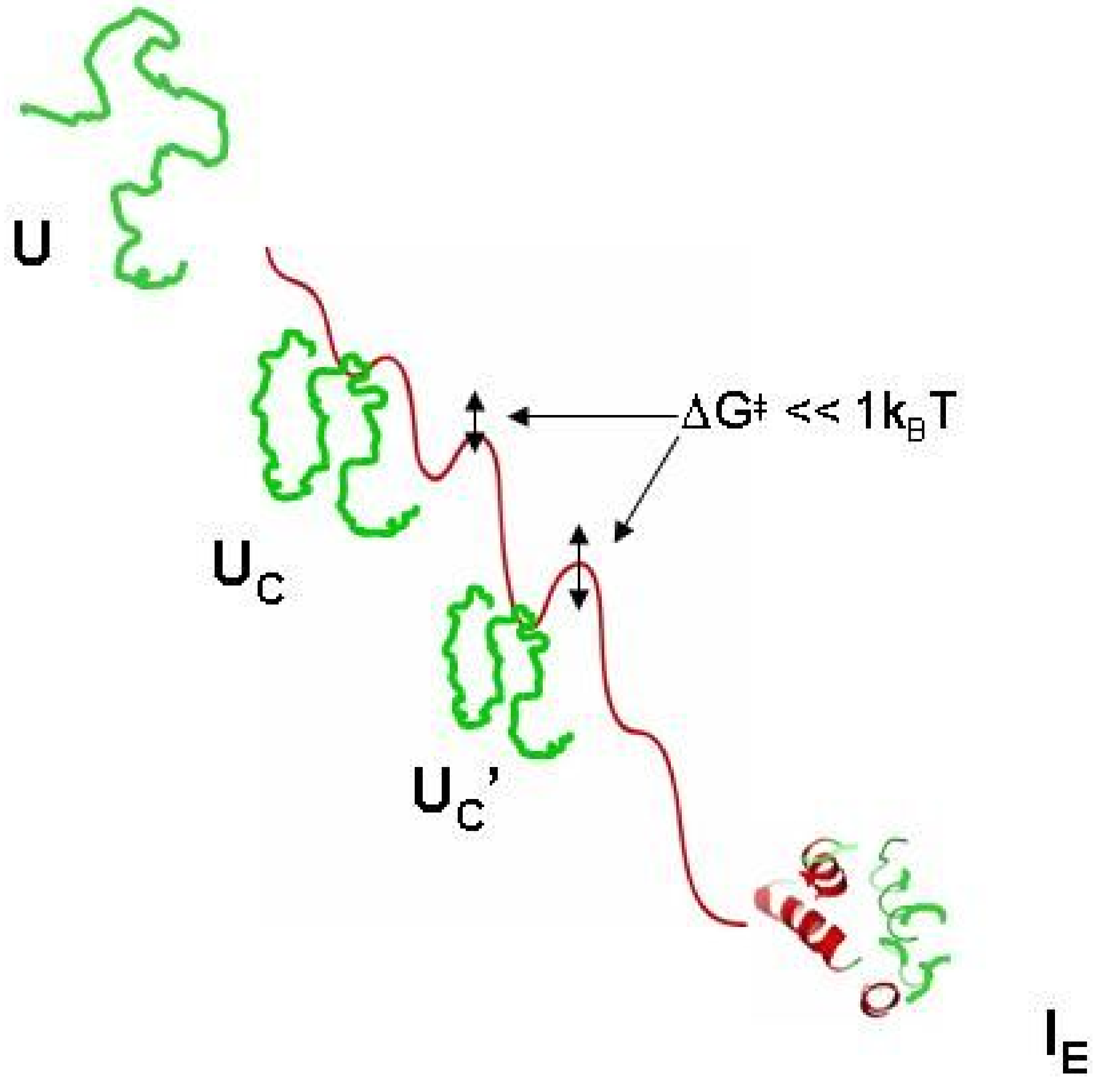
Sub-ms folding reaction of barstar is a barrierless transition



The very fast rate constant does not change with [urea]

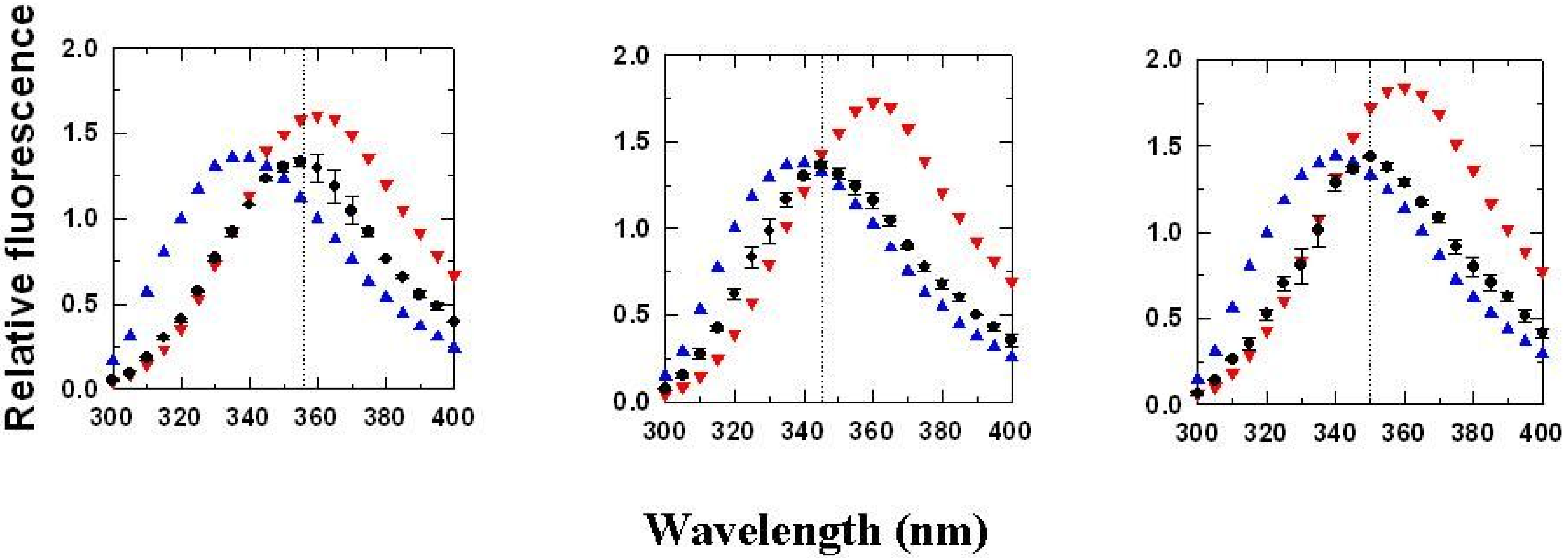
No change in the solvent accessible surface area : SASA is the same in TS and U

Barrierless evolution of specific structure during sub-ms folding reaction



The structure of the initial folding intermediate appears to depend upon the folding conditions

Folding in 0.9 M Urea, pH 8


no added salt

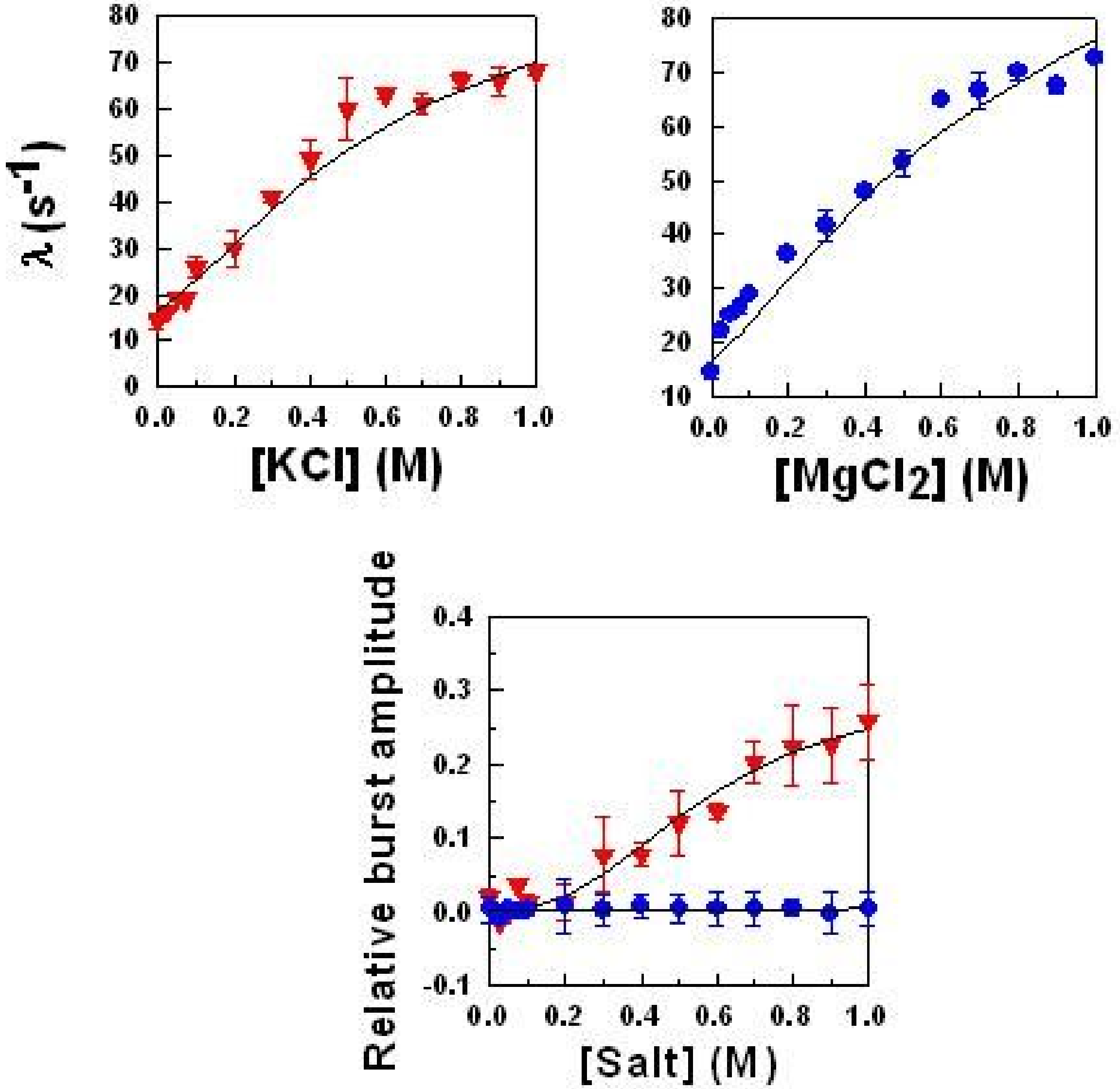
+ 1 M KCl

+1 M $MgCl_2$

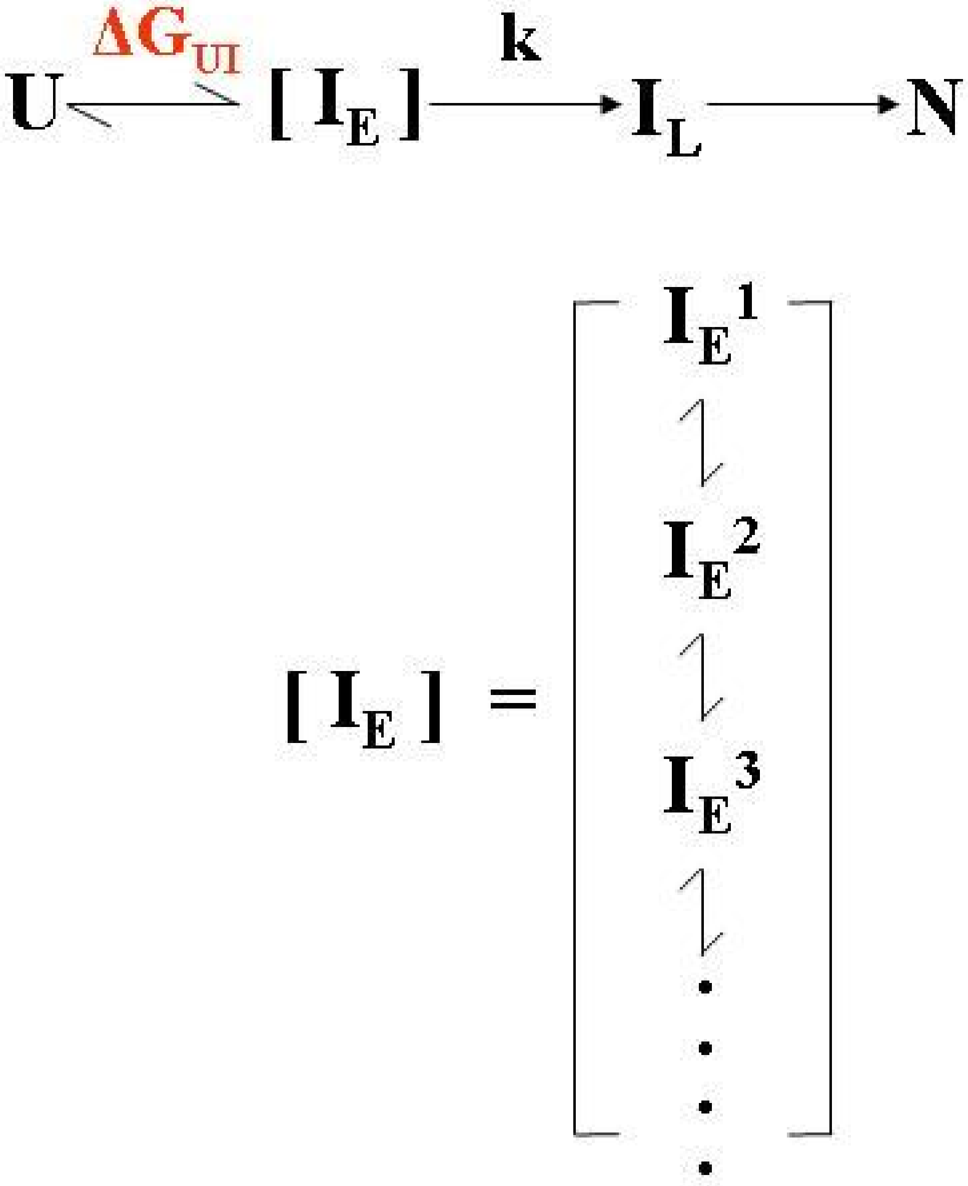
Pradeep and Udgaonkar (2002) J. Mol. Biol. 324, 331-347

Differential salt-induced stabilization of structure in the initial folding intermediate ensemble, I_E

Salt dependence of folding kinetics



Folding mechanism of barstar

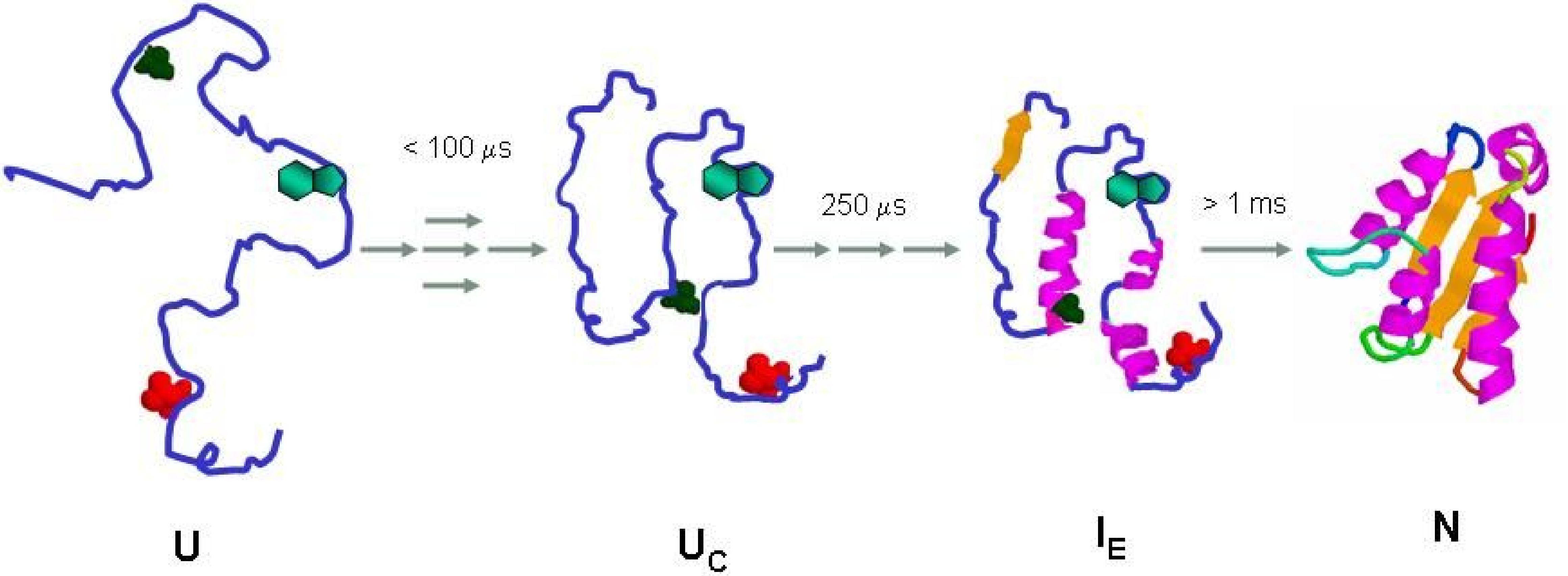


$$U \overset{\Delta G_{UI}}{\rightleftharpoons} [\ I_E\ ] \xrightarrow{k} I_L \longrightarrow N$$

$$[\ I_E\ ] = \left[\begin{array}{c} I_E^{\,1} \\ \upharpoonleft\!\downharpoonright \\ I_E^{\,2} \\ \upharpoonleft\!\downharpoonright \\ I_E^{\,3} \\ \upharpoonleft\!\downharpoonright \\ \vdots \end{array}\right]$$

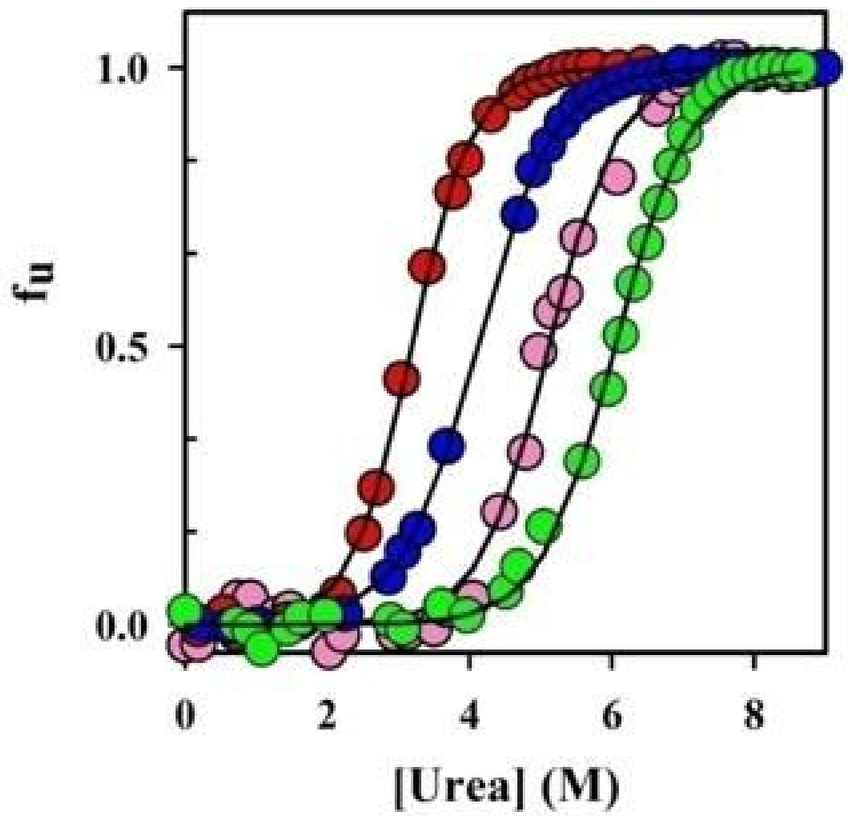
Pradeep and Udgaonkar (2002) J. Mol. Biol. 324, 331-347

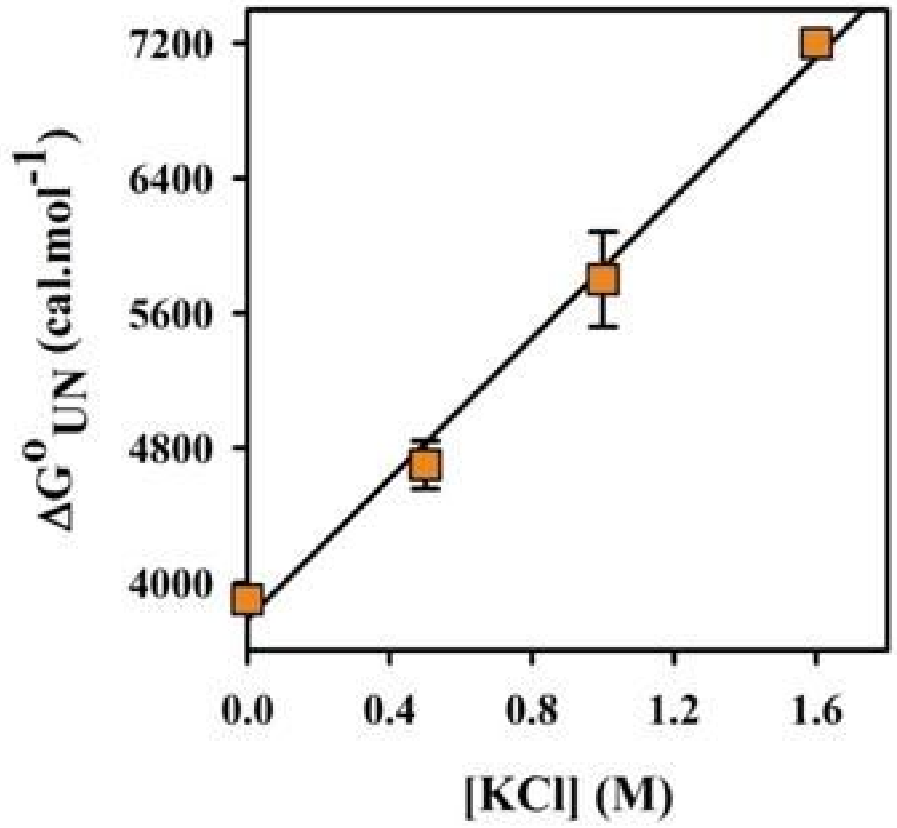
Folding reaction of barstar – the first ms



Sub-ms folding reaction of barstar involves more than one step – non-cooperative

Salt Induced Stabilization of Native Barstar

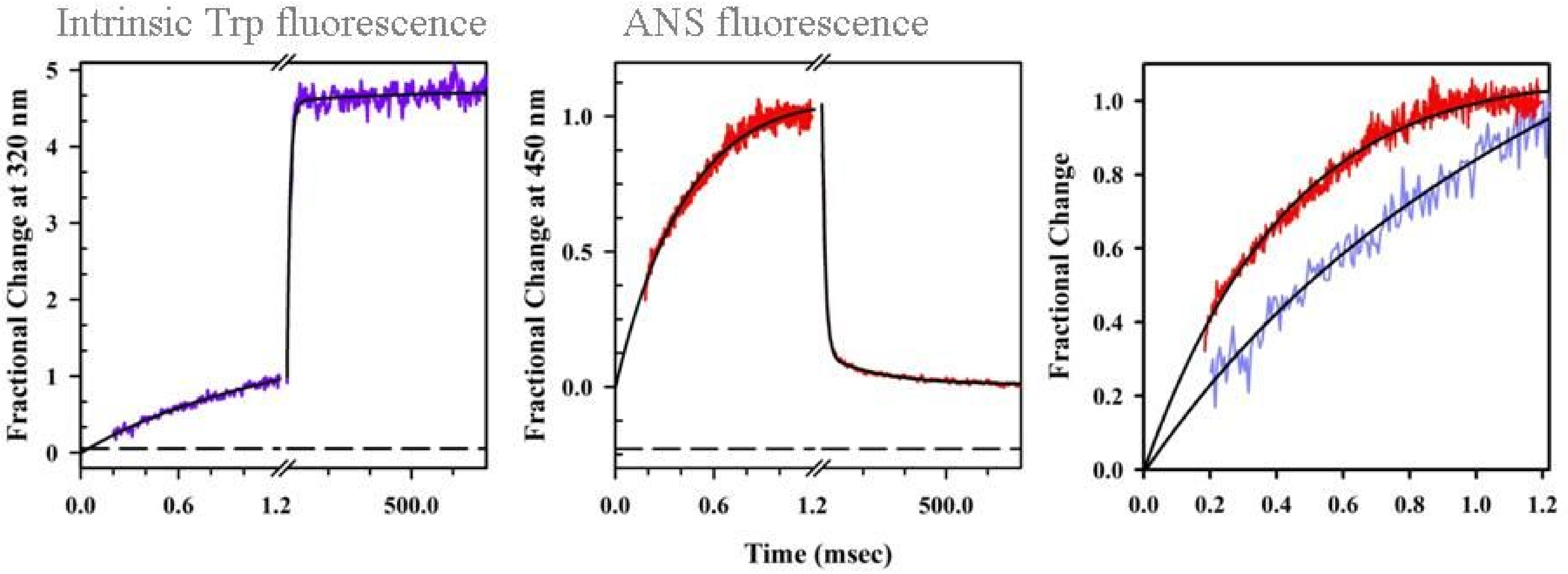




Dasgupta and Udgaonkar (2010)

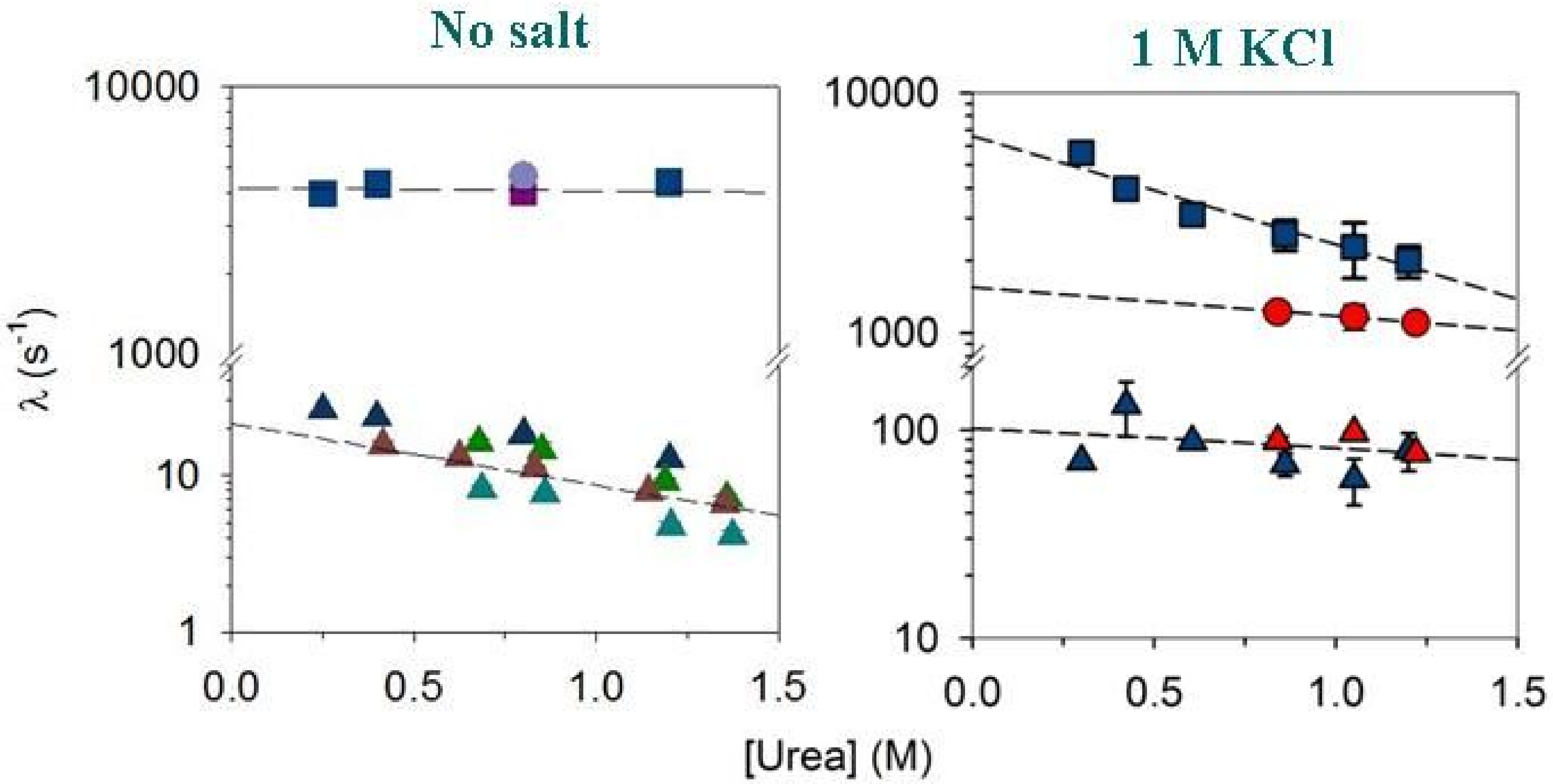
Sub-millisecond folding occurs in multiple stages

Two stages can be detected in the transformation of the structure-less globule U_C to the specifically structured intermediate I_E, for folding in 1 M KCl



Dasgupta and Udgaonkar (2010)

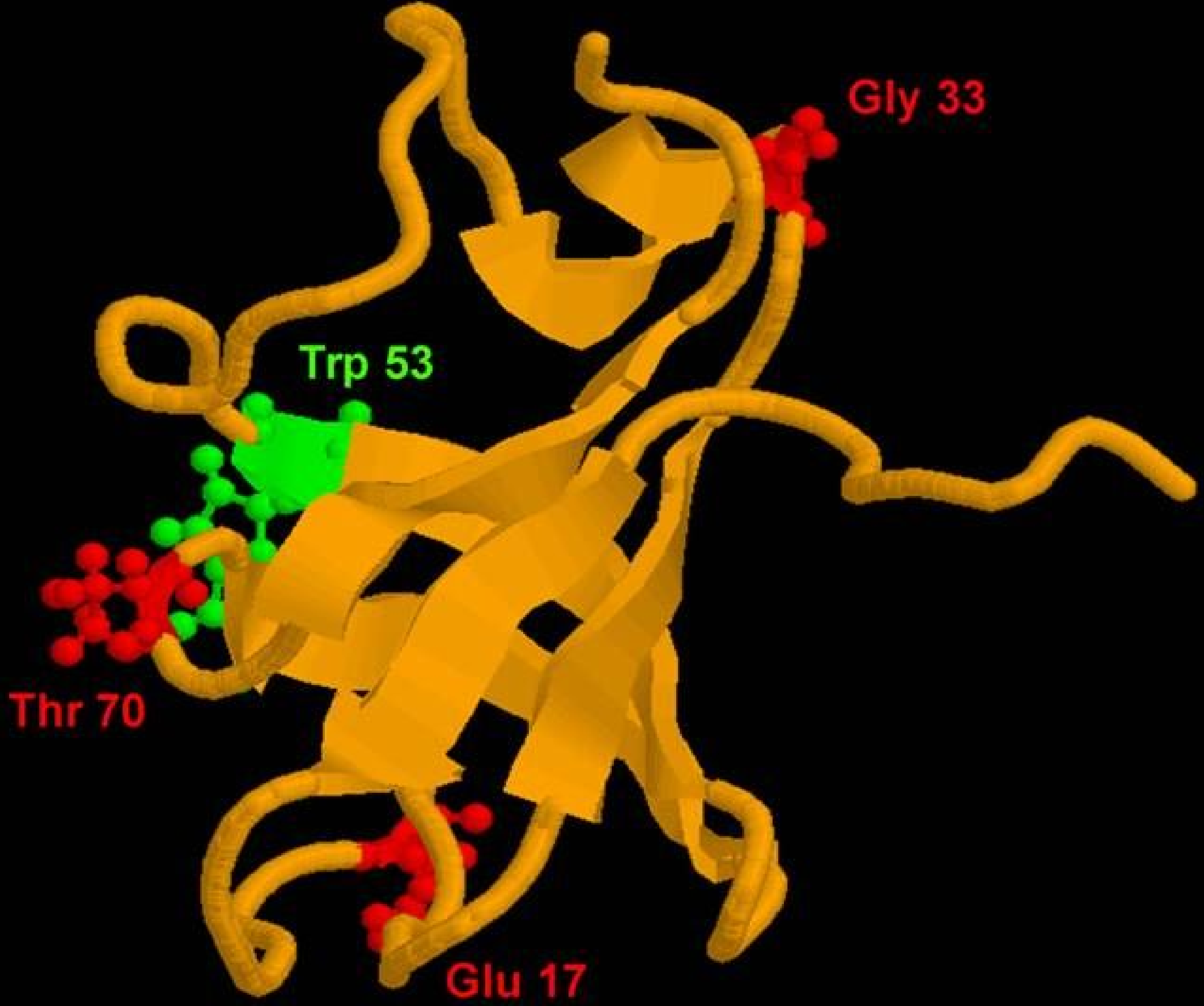
Salt-induced modulation of sub-ms folding



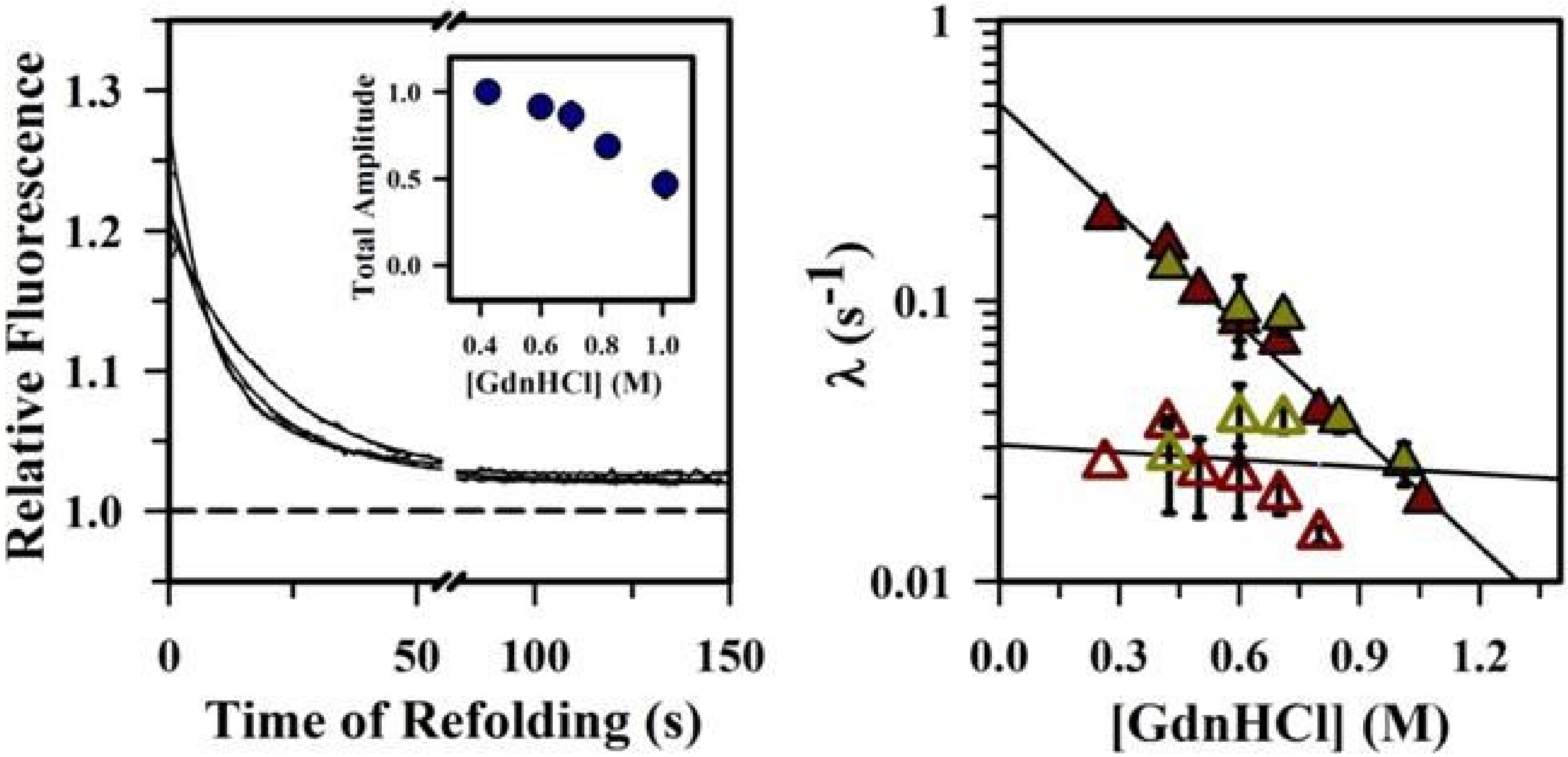
$$U \overset{<\ 100\ \mu s}{\longrightarrow\longrightarrow\longrightarrow} U_C \overset{\sim\ 100\ \mu s}{\longrightarrow\longrightarrow\longrightarrow} U_C' \overset{\sim\ 0.5\ ms}{\longrightarrow\longrightarrow} I_E$$

Dasgupta and Udgaonkar (2010)

SH3 domain of PI3 Kinase

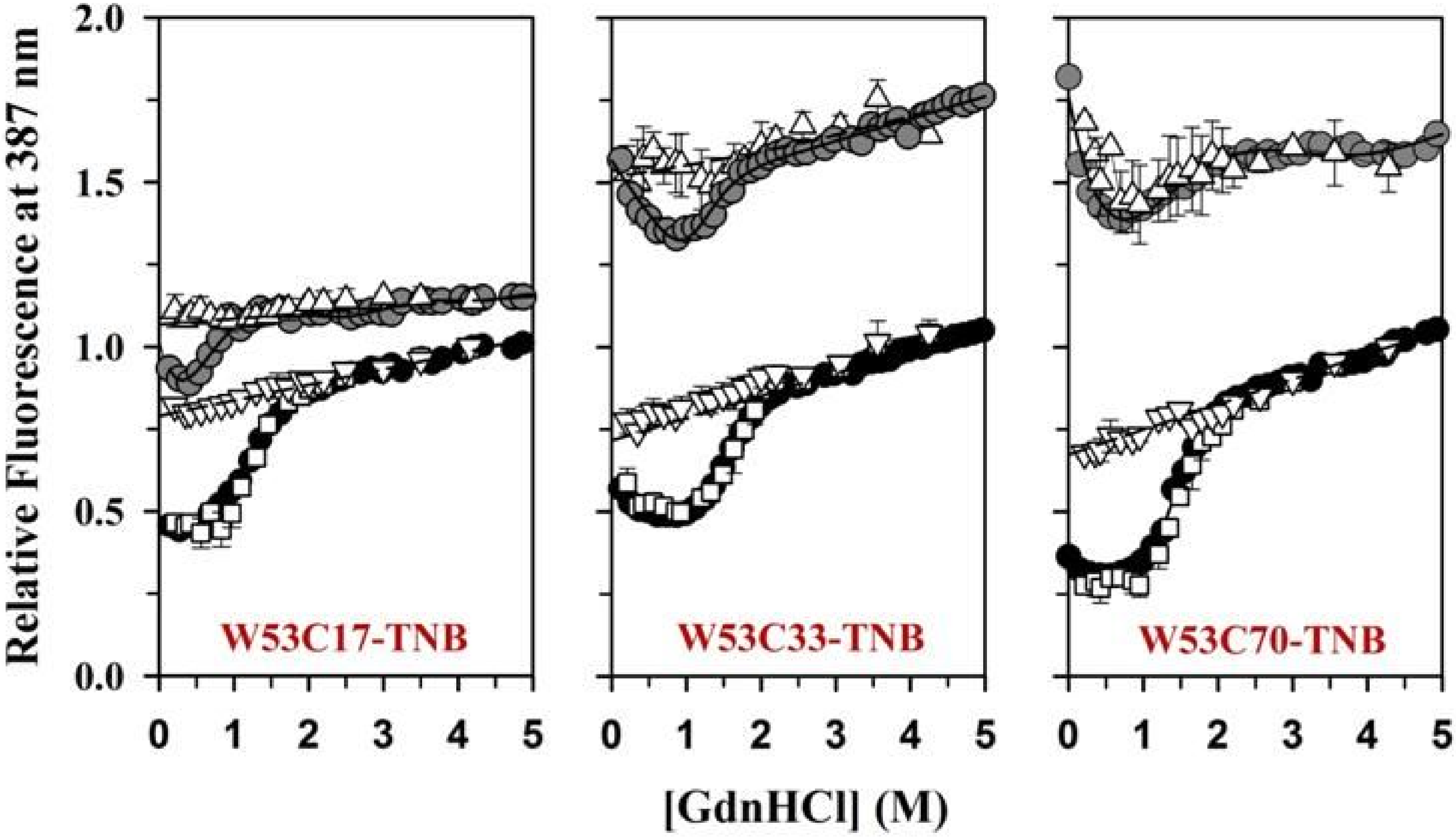


Evidence for an initial hydrophobic collapse during the refolding of the SH3 domain of PI3K



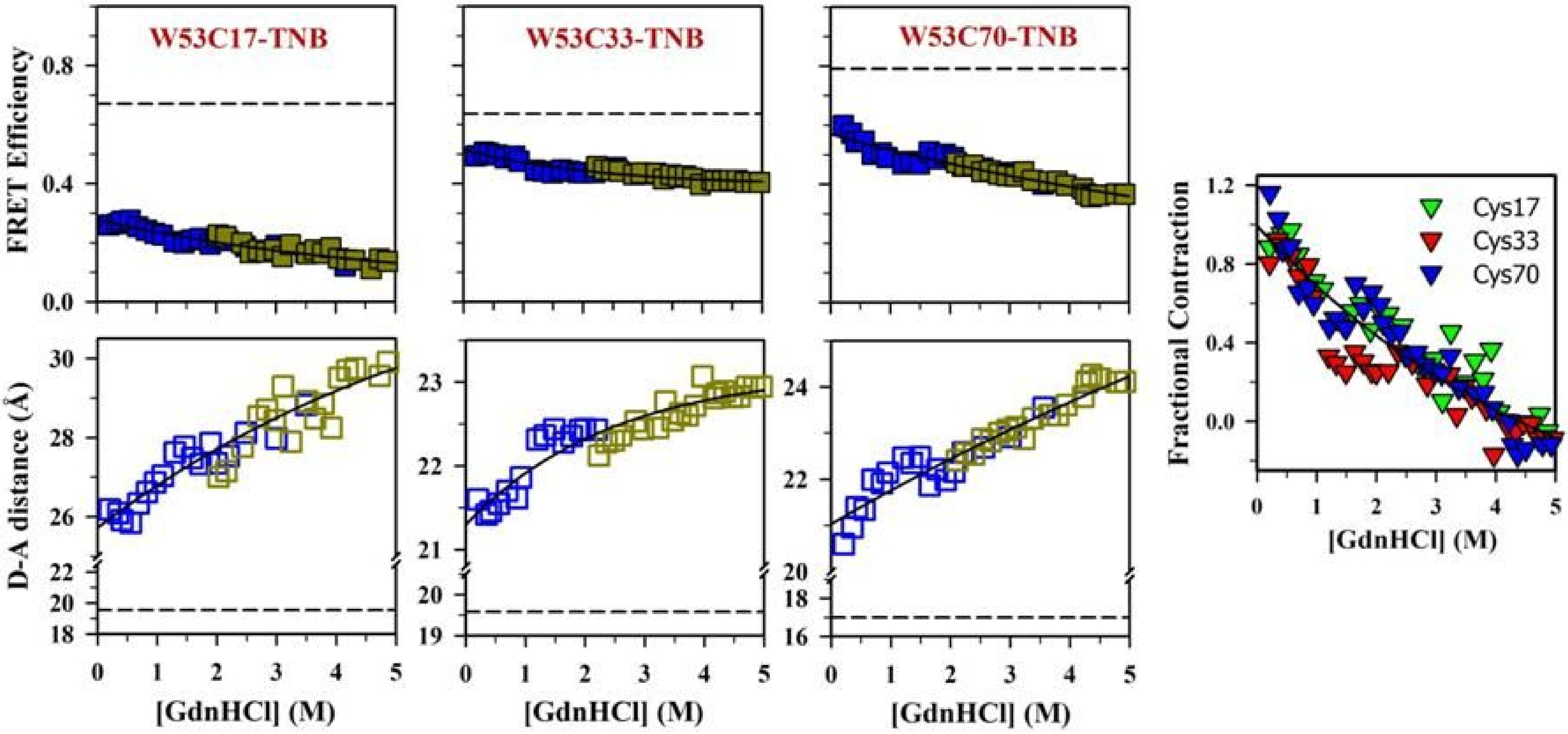
Dasgupta and Udgaonkar (2010)

The initial collapse during the refolding of the SH3 domain of PI3K is non-specific

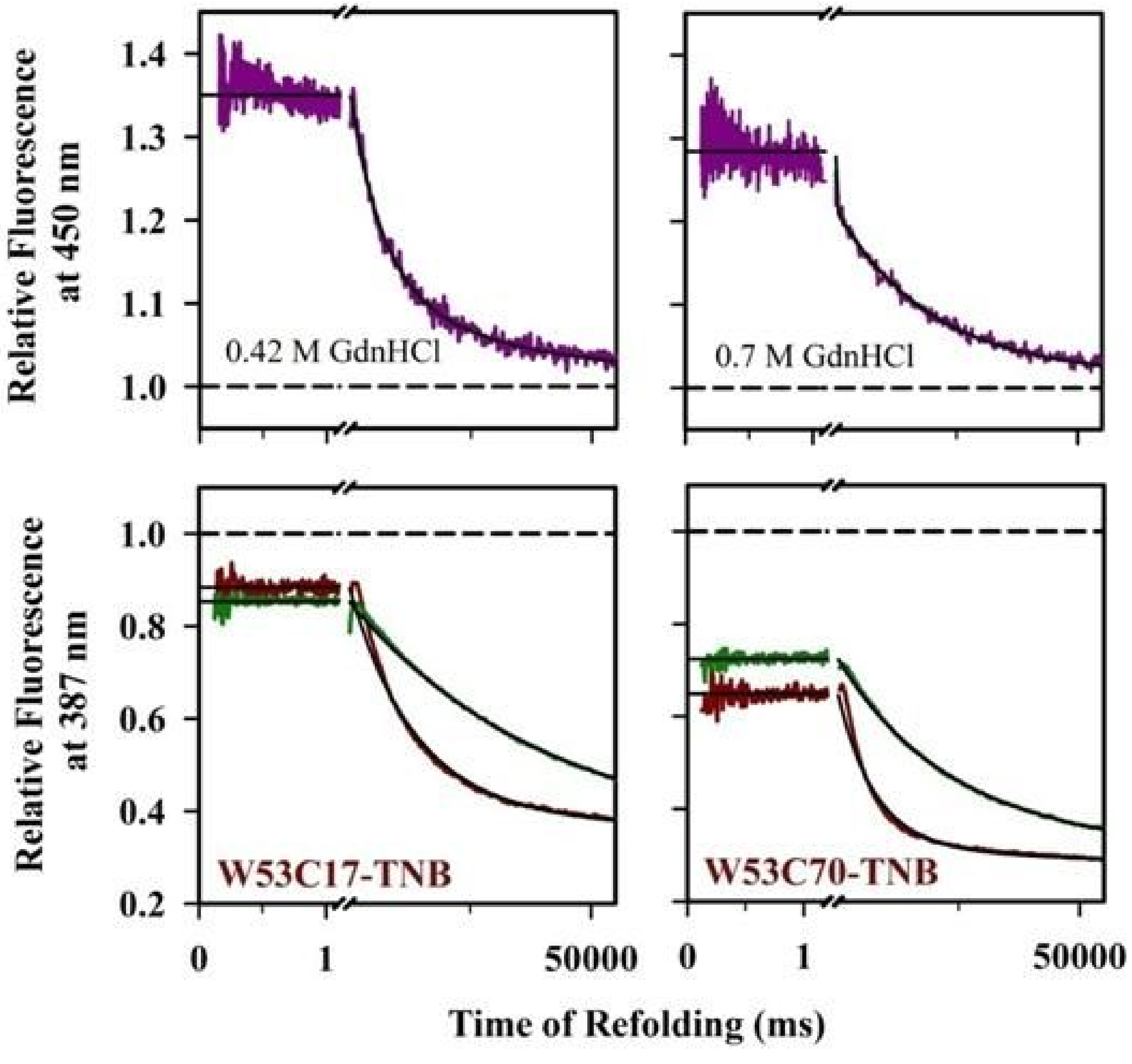


Dasgupta and Udgaonkar, (2010)

Different segments of the structure appear to collapse in a synchronized manner



Sub-millisecond measurements using a continuous flow mixer indicate that chain collapse occurs faster than 20, 000 s^{-1}



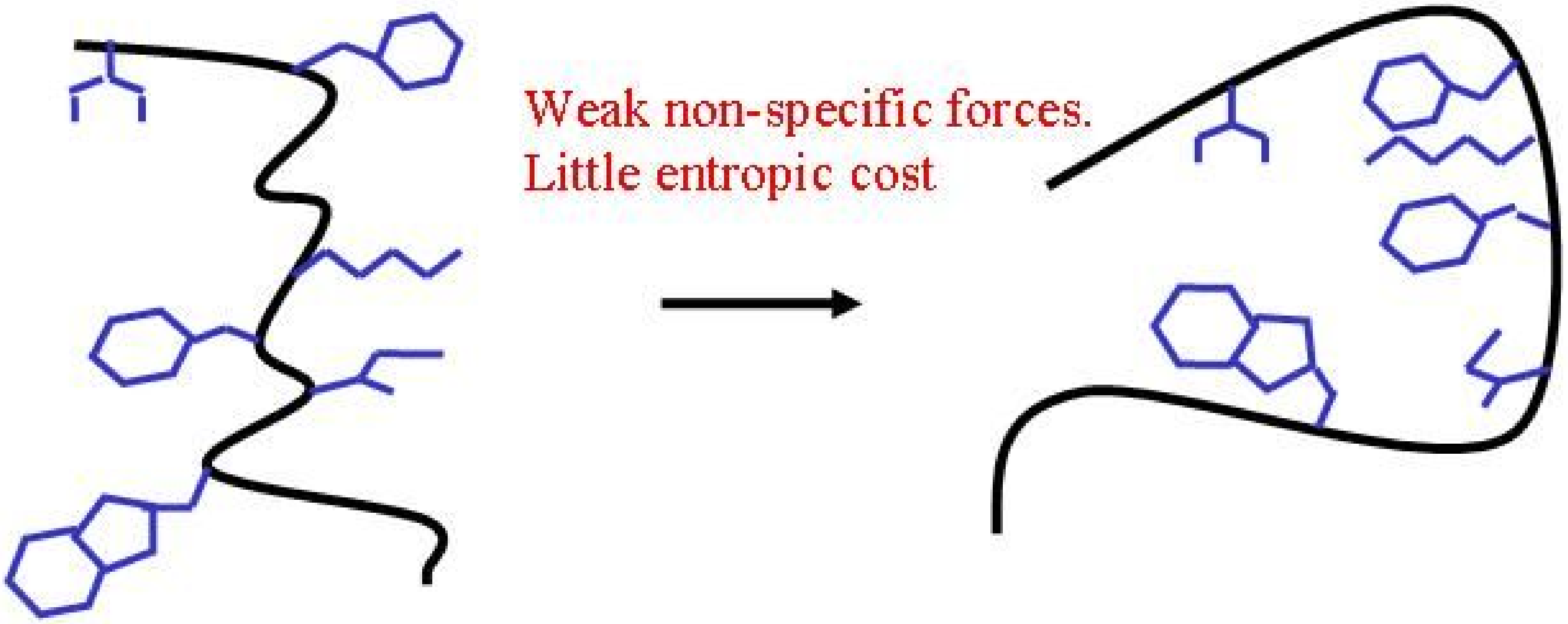
Hydrophobic Collapse

Hydrophobic interactions.

1) Intuitive picture: Water drives non-polar molecules together. Thought to be the dominant interaction in proteins but very little evidence.

➡ Hydrophobic (non-polar) residues are generally found in the interior of globular proteins.

2) Second component from Dispersion forces.
van der Waals interactions are not very specific.
All =CH, >CH2, -CH3 groups attract each other



Unfolded chain

Weak non-specific forces.
Little entropic cost

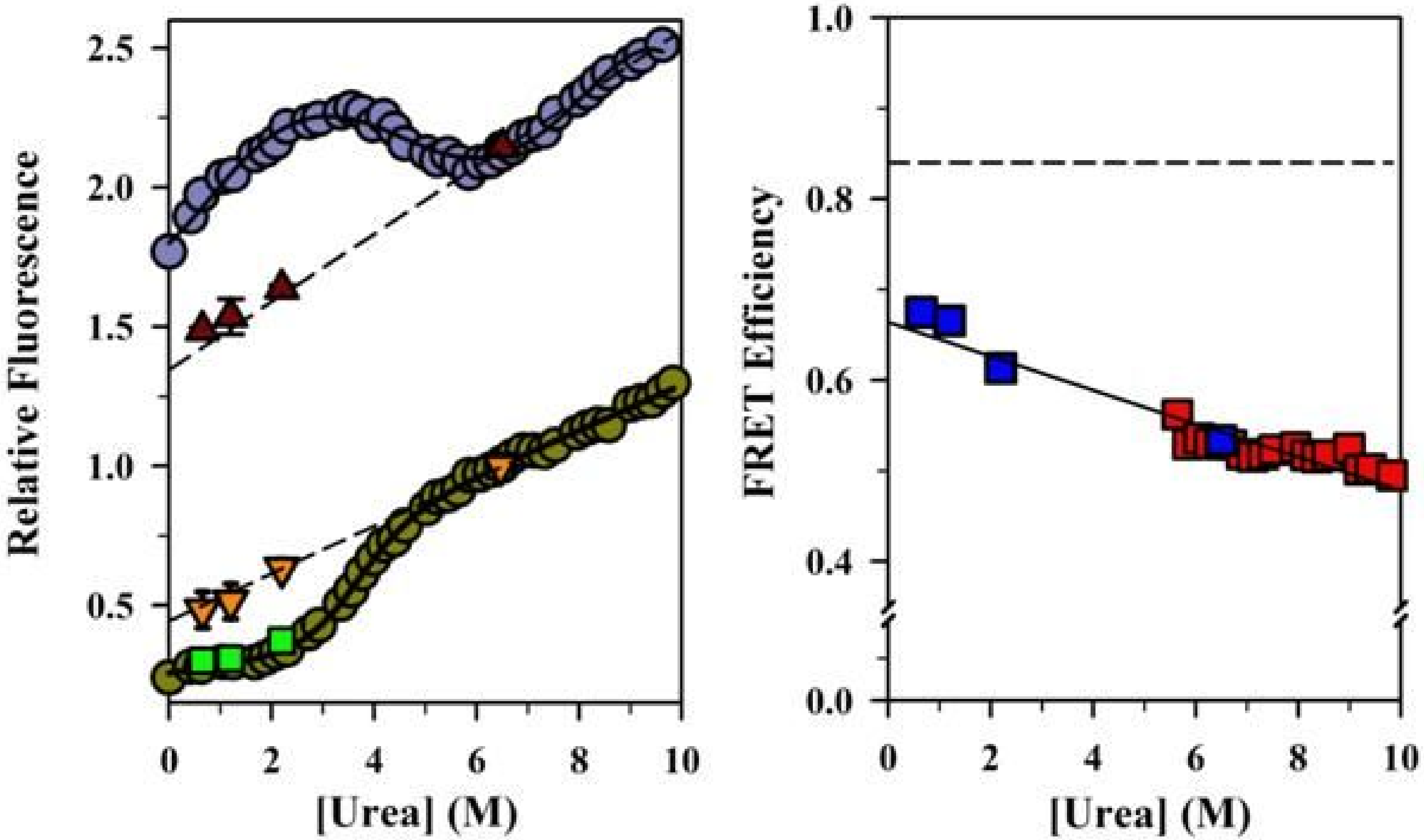
Loose compact structure

Exposed non-polar side chains

Non-polar side chains buried

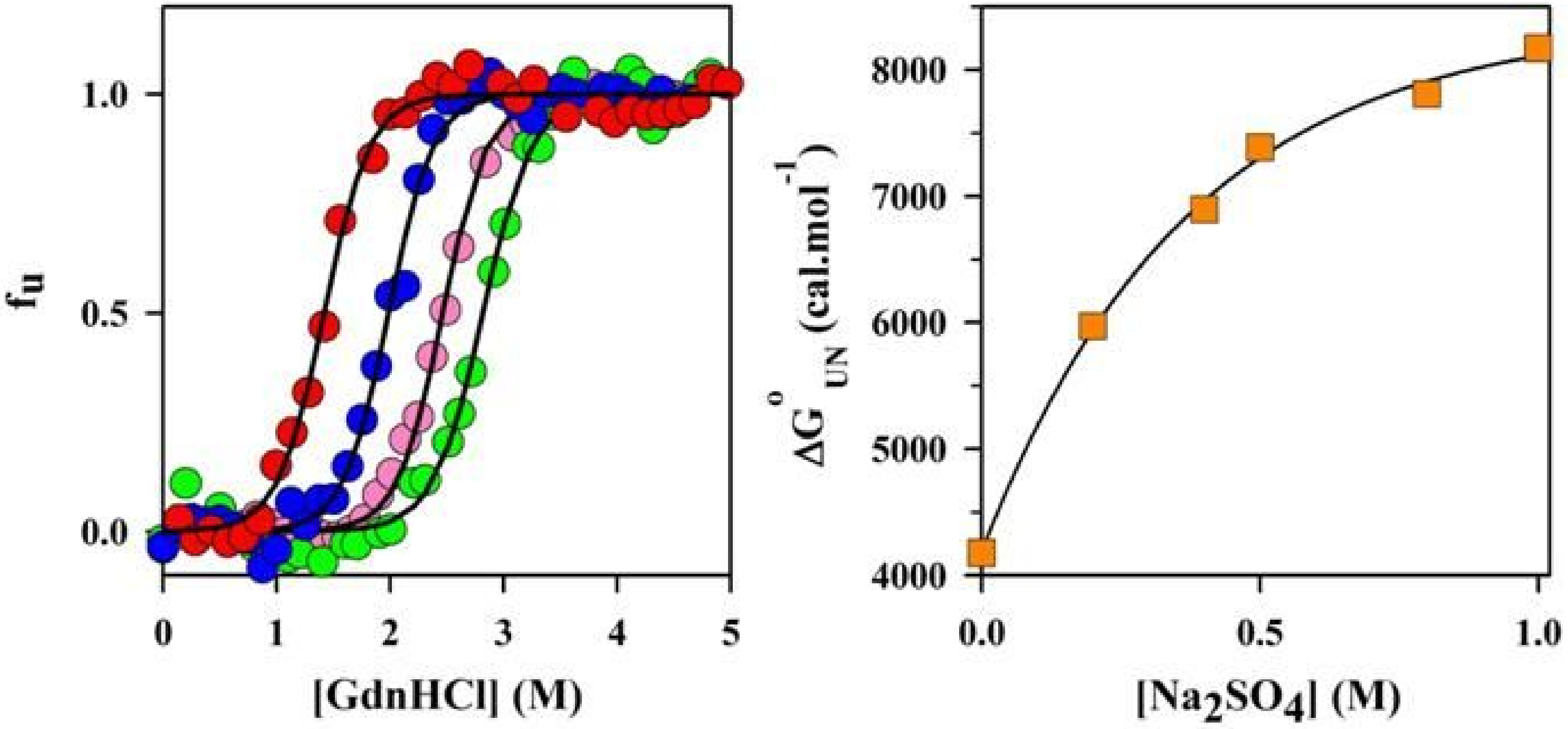
Hydrogen Bonding may play a role in driving initial chain collapse during the folding of the SH3 domain of PI3K

Upon dilution of urea hydrogen bonding interactions becomes stronger while hydrophobic interactions remain relatively unaffected (Bolen and coworkers, 2010).



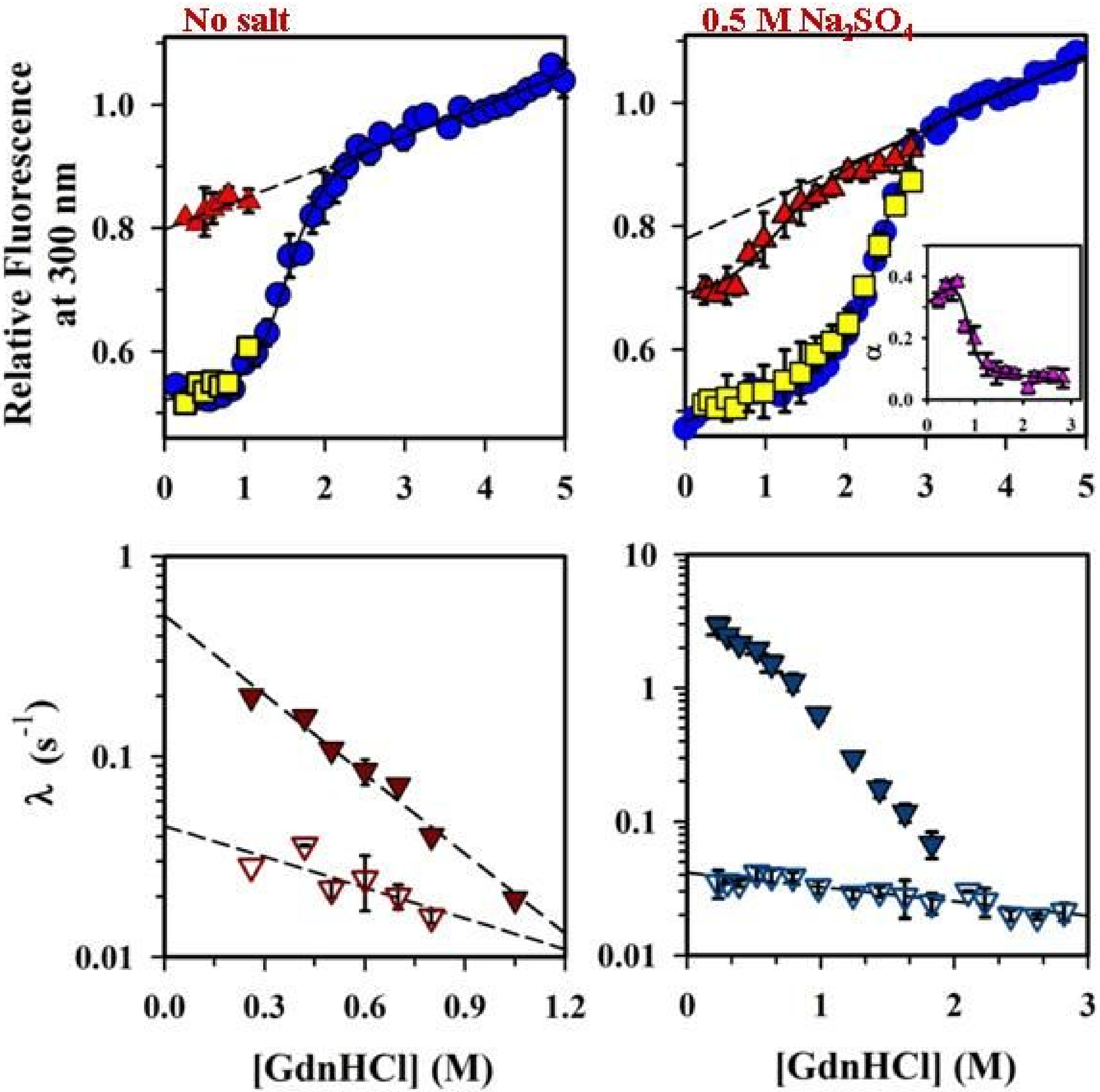
Dasgupta and Udgaonkar, (2010)

Salt Induced Stabilization of Native SH3 domain of PI3K



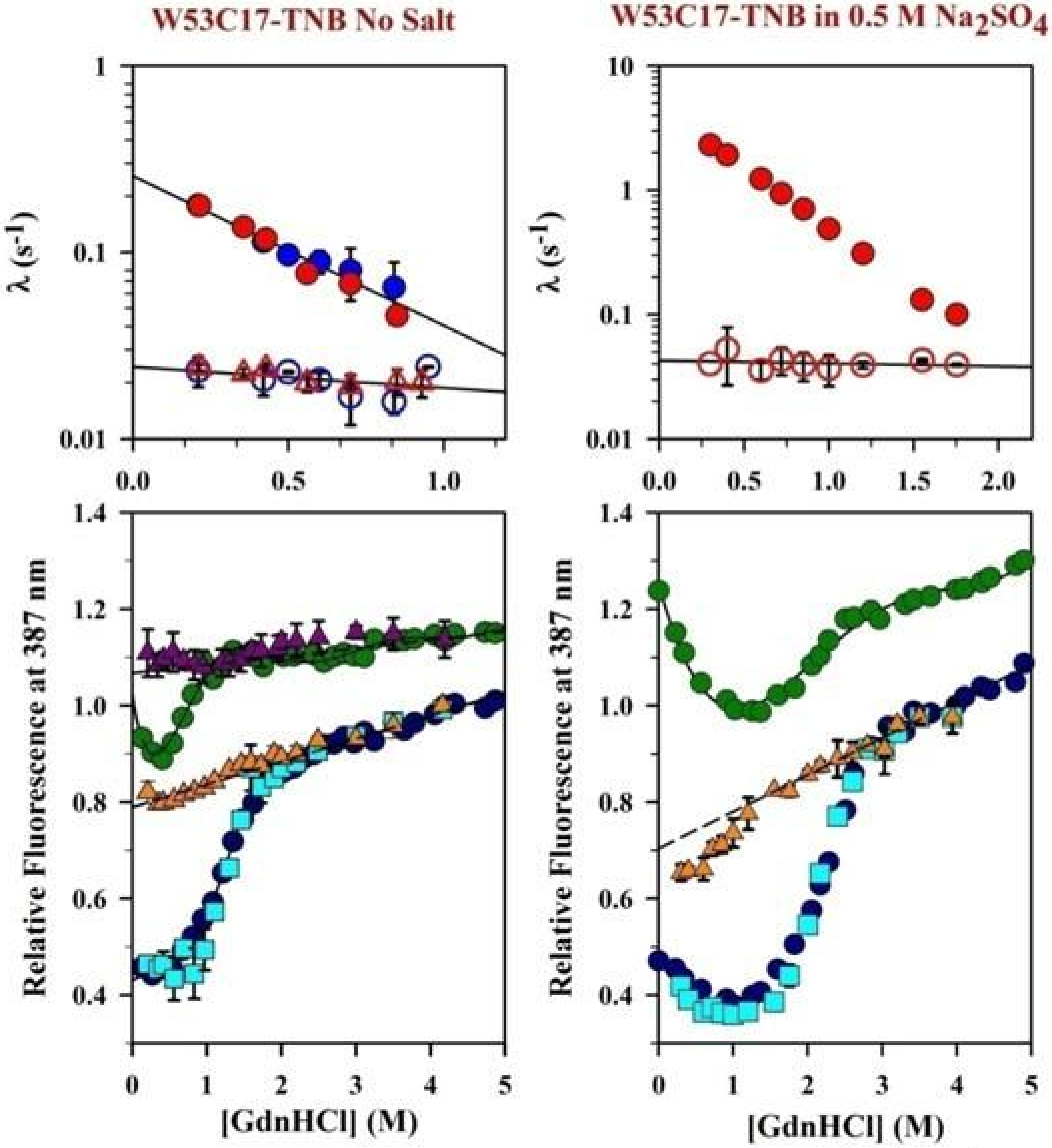
Dasgupta and Udgaonkar, (2010)

A specific intermediate is formed during sub-ms chain collapse during folding in strongly stabilizing conditions



Dasgupta and Udgaonkar, (2010)

Site-specific structure formation during initial chain collapse can be detected by FRET

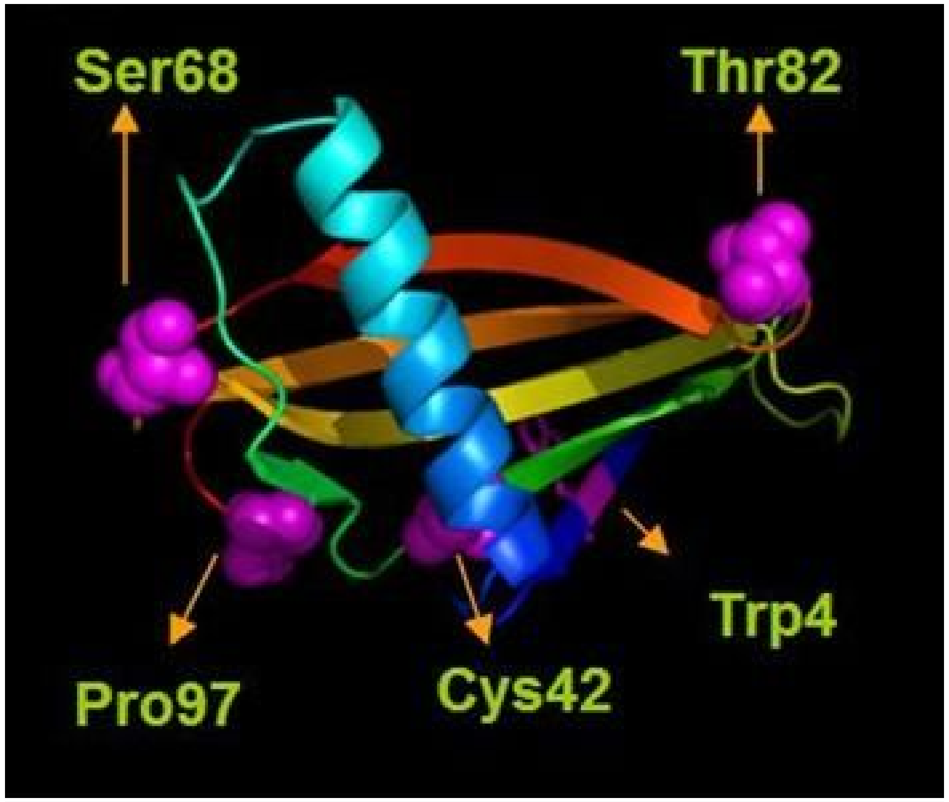


Dasgupta and Udgaonkar, (2010)

Continuous dissolution of structure during the unfolding of a small protein

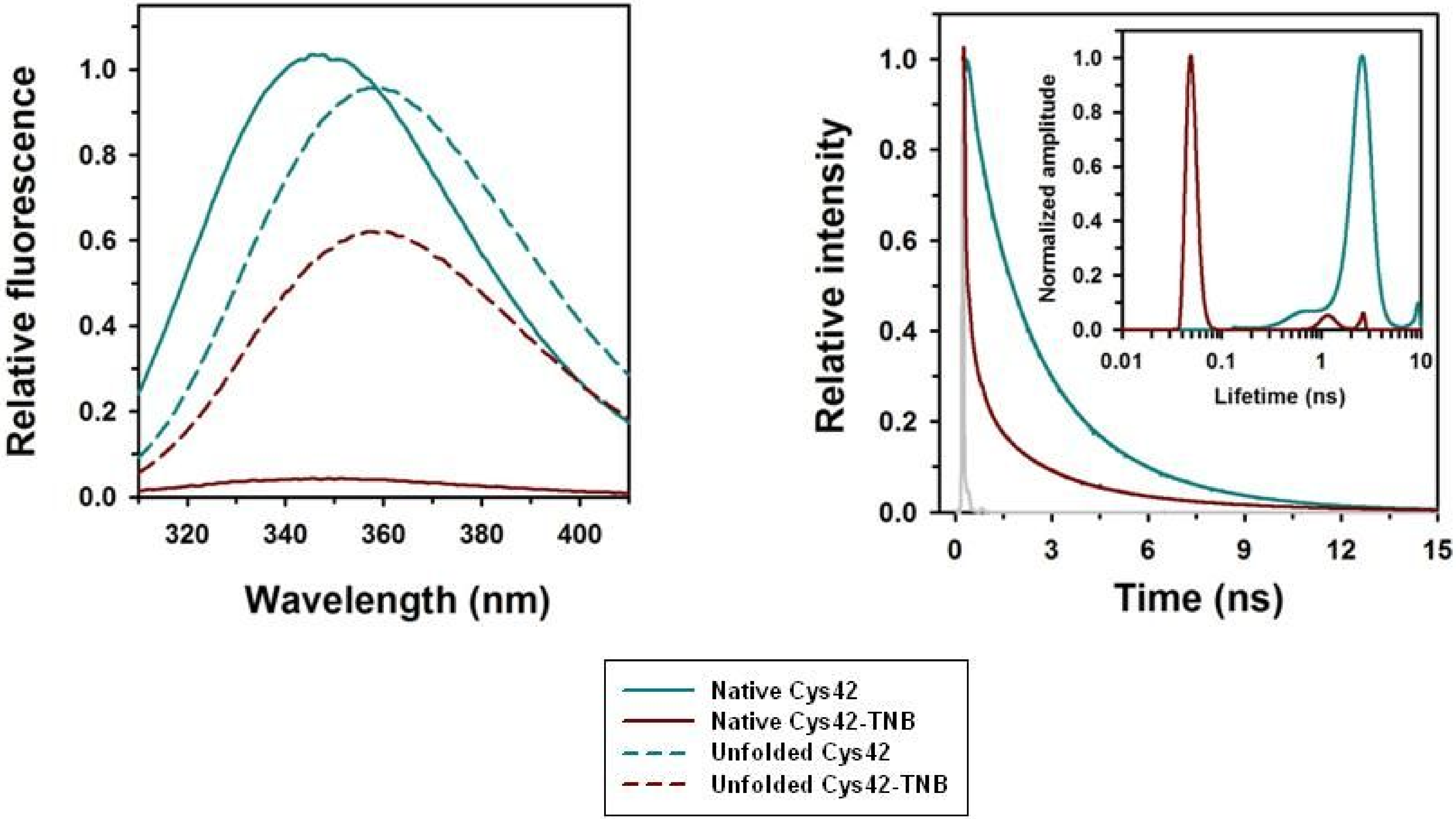
A set of four single Trp and single Cys-containing mutant forms of monellin was used, in each of which a thionitrobenzoate (TNB) adduct was attached to the cysteine thiol.

Locations of amino acid residues which were mutated to cysteine and their distances from W4 in the structure of monellin

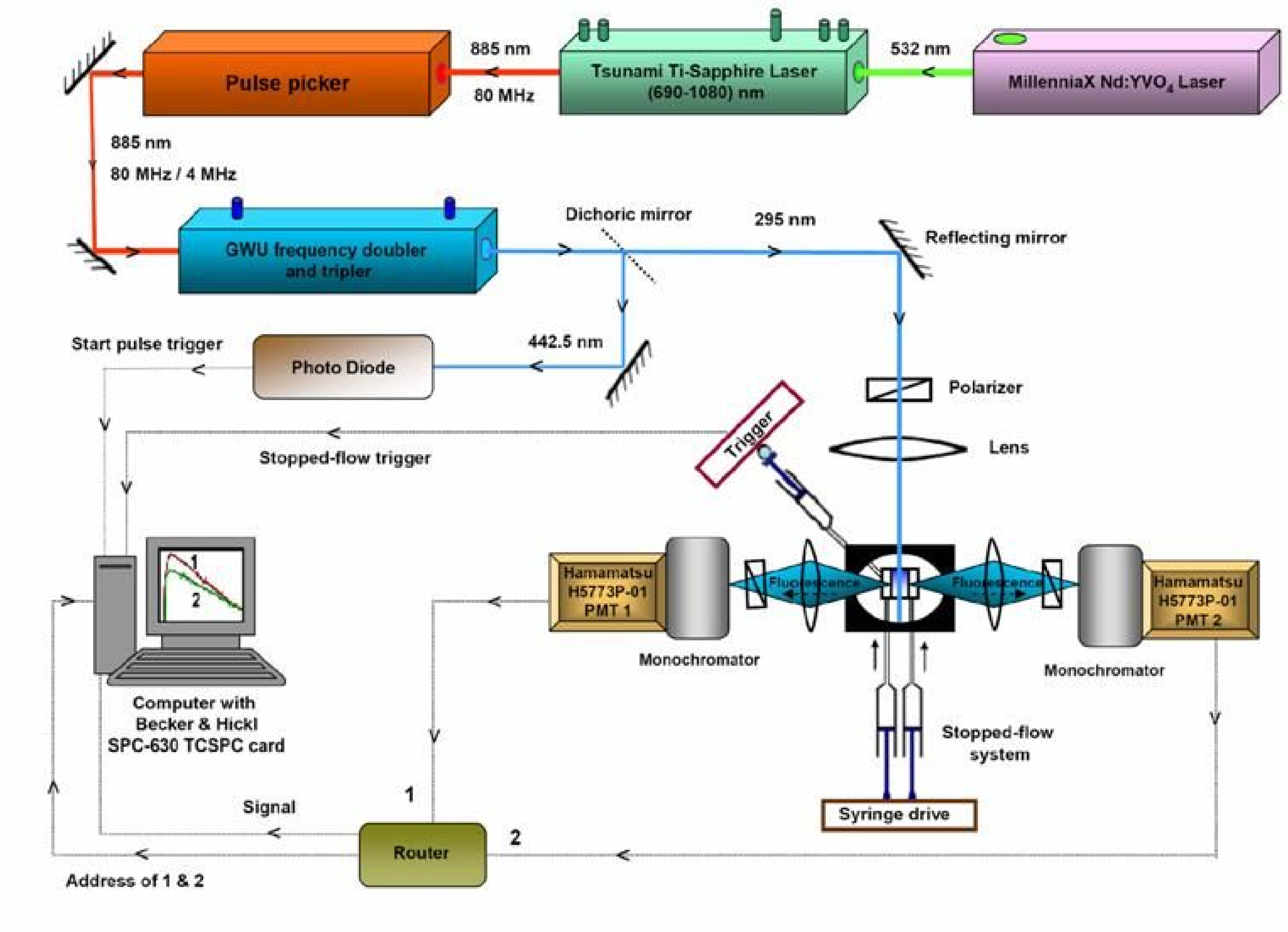


*An average co-ordinate was calculated for each side-chain from the X-ray structure of monellin. The size of the TNB adduct is not included in this calculation.

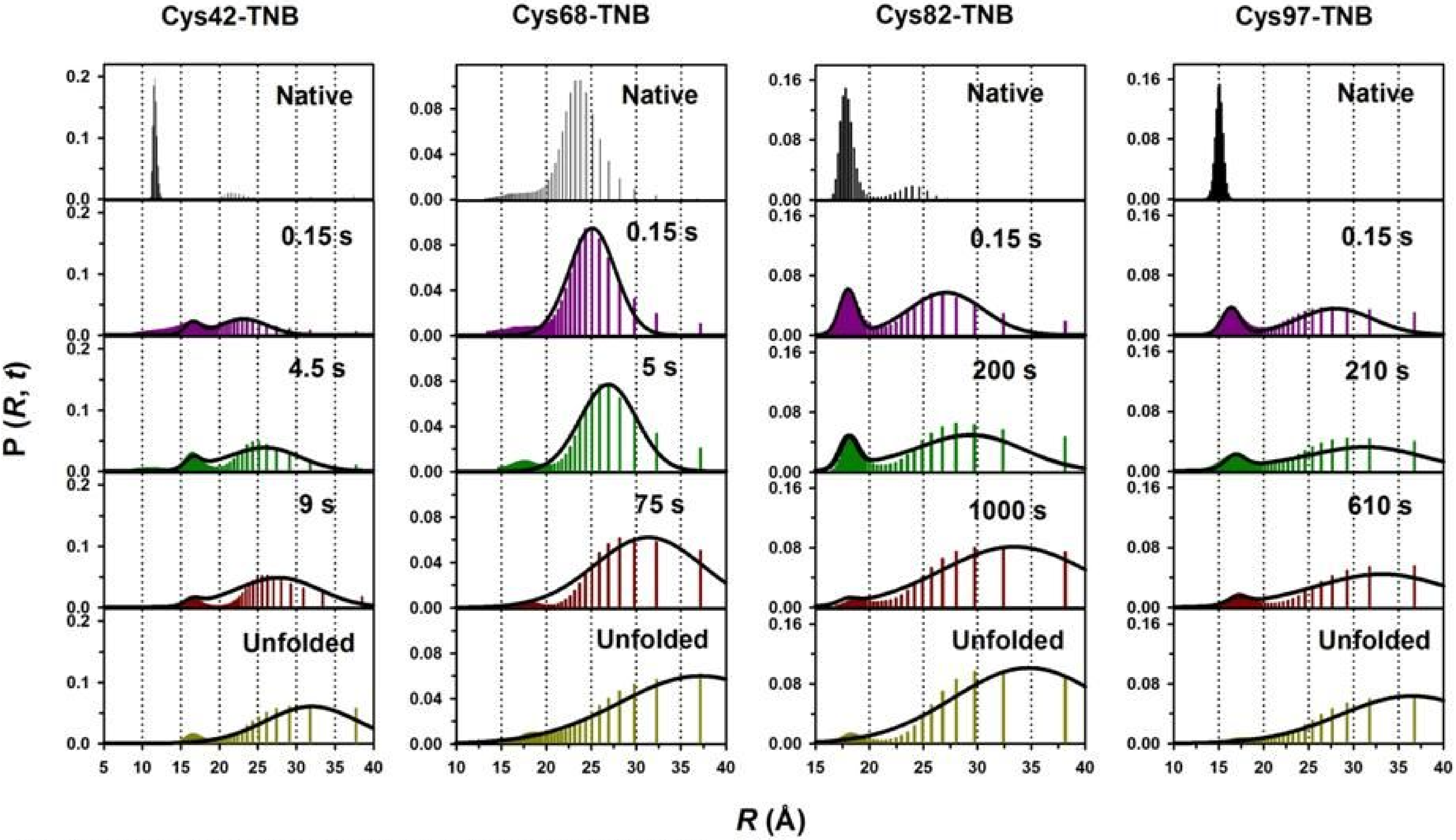
Time-resolved fluorescence resonance energy transfer (TR-FRET) coupled to maximum entropy method analysis (MEM) can reveal the populations of species with different fluorescence lifetimes



Schematic of dual-channel time-correlated single photon counting (TCSPC) set-up coupled to a stopped-flow mixing module 'Double-kinetic' experiments



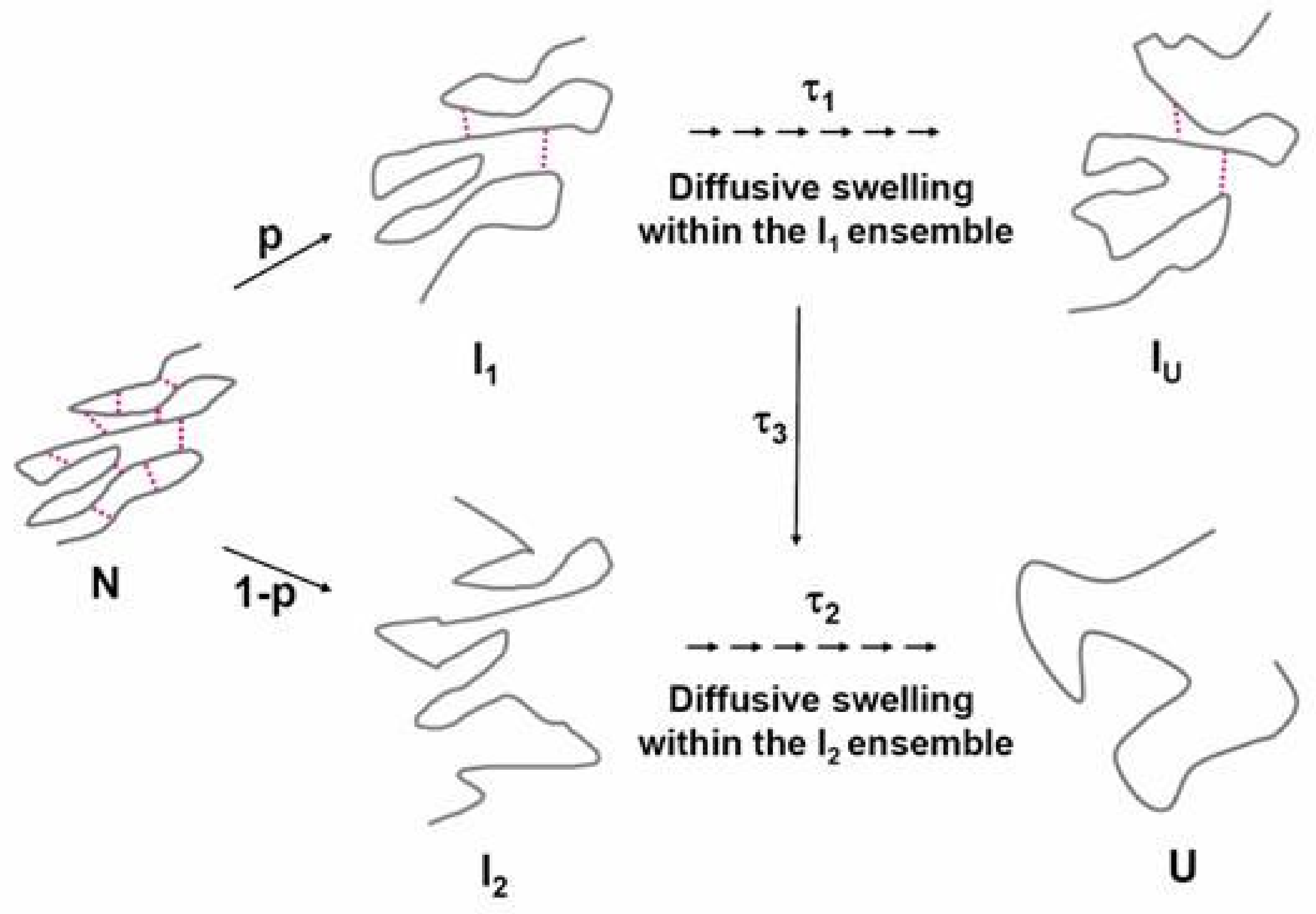
D-A distance distributions, at different times of unfolding, are broad and change continuously with the time of unfolding



Vertical bars: Experimental distance distributions.

Continuous black lines: Theoretical distance distributions generated from model.

A simple phenomenological model based upon the slow diffusive swelling of a Rouse-like chain with some cross-links can explain the experimental data

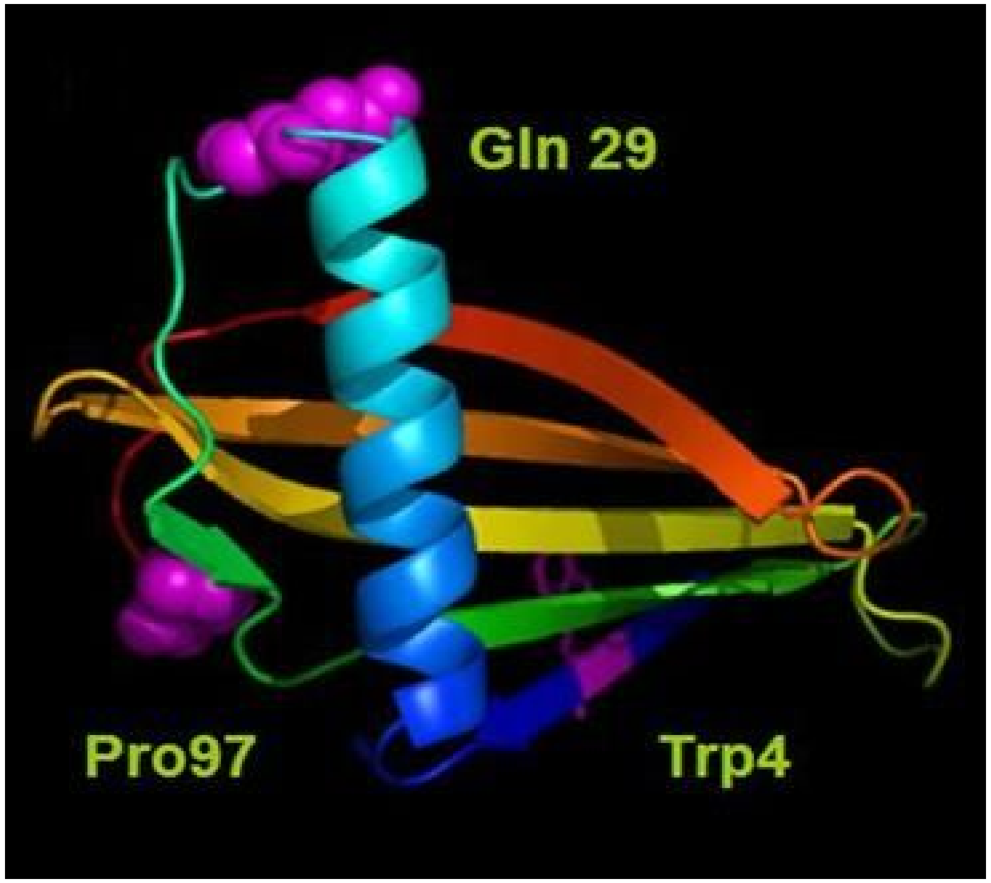


$$P\,(R,t) = \frac{1}{\sqrt{2\pi}\sigma_i(t)} e^{-\frac{(R-R_i(t))^2}{2\sigma_i^2(t)}}$$

where, $R_i(t)$ and $\sigma_i^2\,(t)$ are the *time-dependent* mean and variance, respectively, of the distance distribution in the ensemble I_i given by $R_i(t) = R_i\,(\infty) - [R_i\,(\infty) - R_i\,(0)]\, e^{-\frac{t}{\tau_i}}$ and $\sigma_i^2\,(t) = \sigma_i^2\,(\infty) - \left[\,\sigma_i^2\,(\infty) - \sigma_i^2(0)\right] e^{-\frac{t}{\tau_i}}$

Jha, Dhar, Krishnamoorthy and Udgaonkar (2009) Proc. Natl. Acad. Sci. USA. 106, 11113-11118.

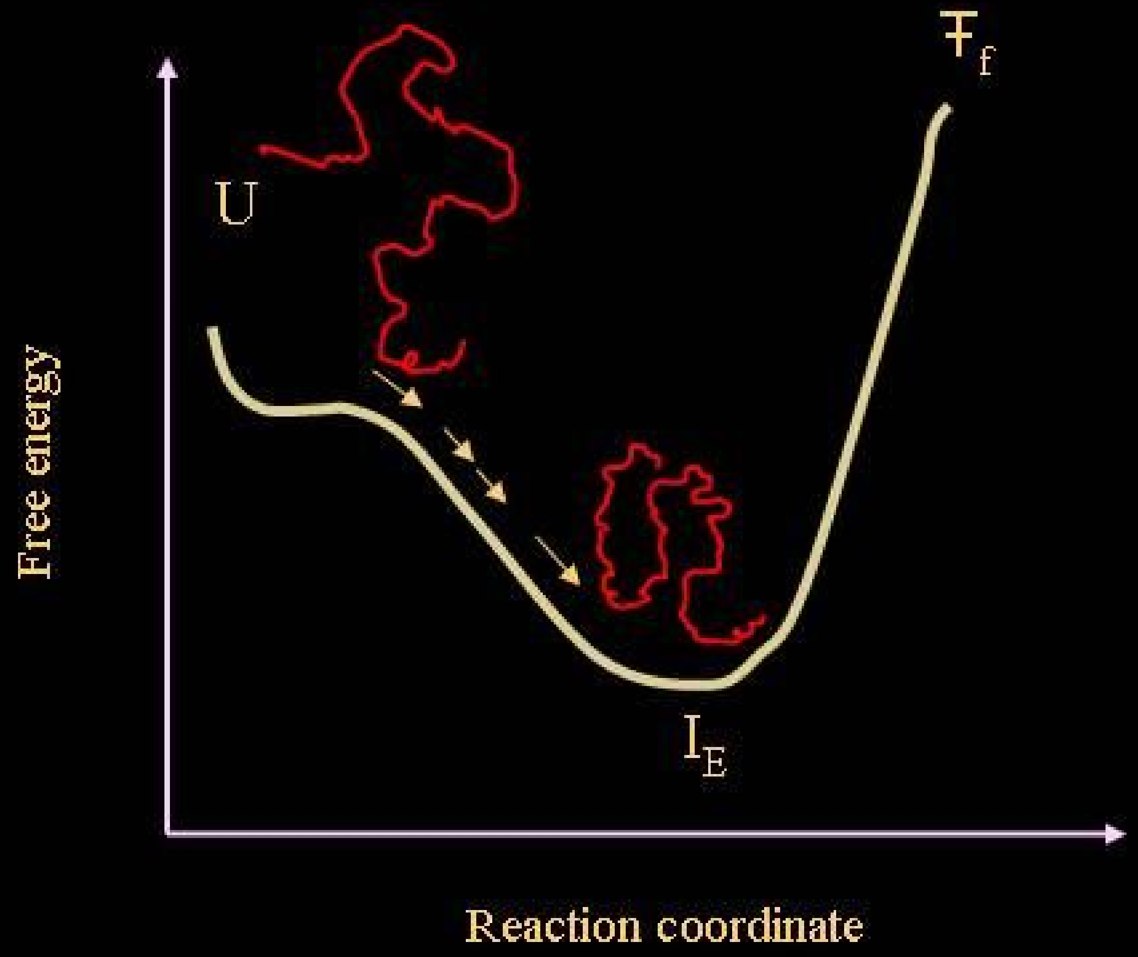
Direct observation of a dry molten globule intermediate during the unfolding of a small protein



The helix moves away from the rest of the protein during the initial stage of unfolding

Jha and Udgaonkar (2009) Proc. Natl. Acad. Sci. USA. 106, 12289-12294

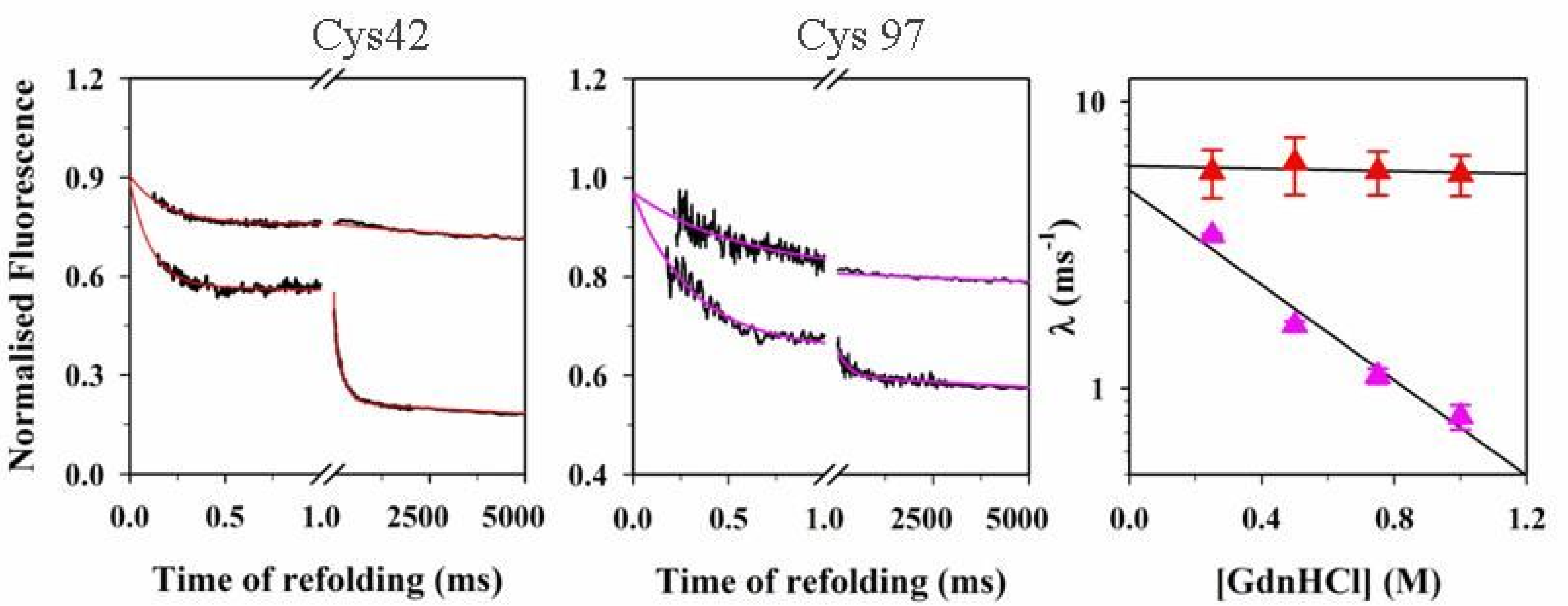
The unfolded polypeptide chain collapses in a continuous barrierless manner



Sinha, K.K. & Udgaonkar, J.B. (2005) *J Mol Biol* **353**, 704-718
Sinha, K.K. & Udgaonkar, J.B. (2007) *J Mol Biol* **370**, 385-405
Sinha, K.K. & Udgaonkar, J.B. (2008) *Proc. Natl. Acad. Sci. USA* 105, 7998-8003
Dasgupta. A. & Udgaonkar, J.B. (2010) submitted

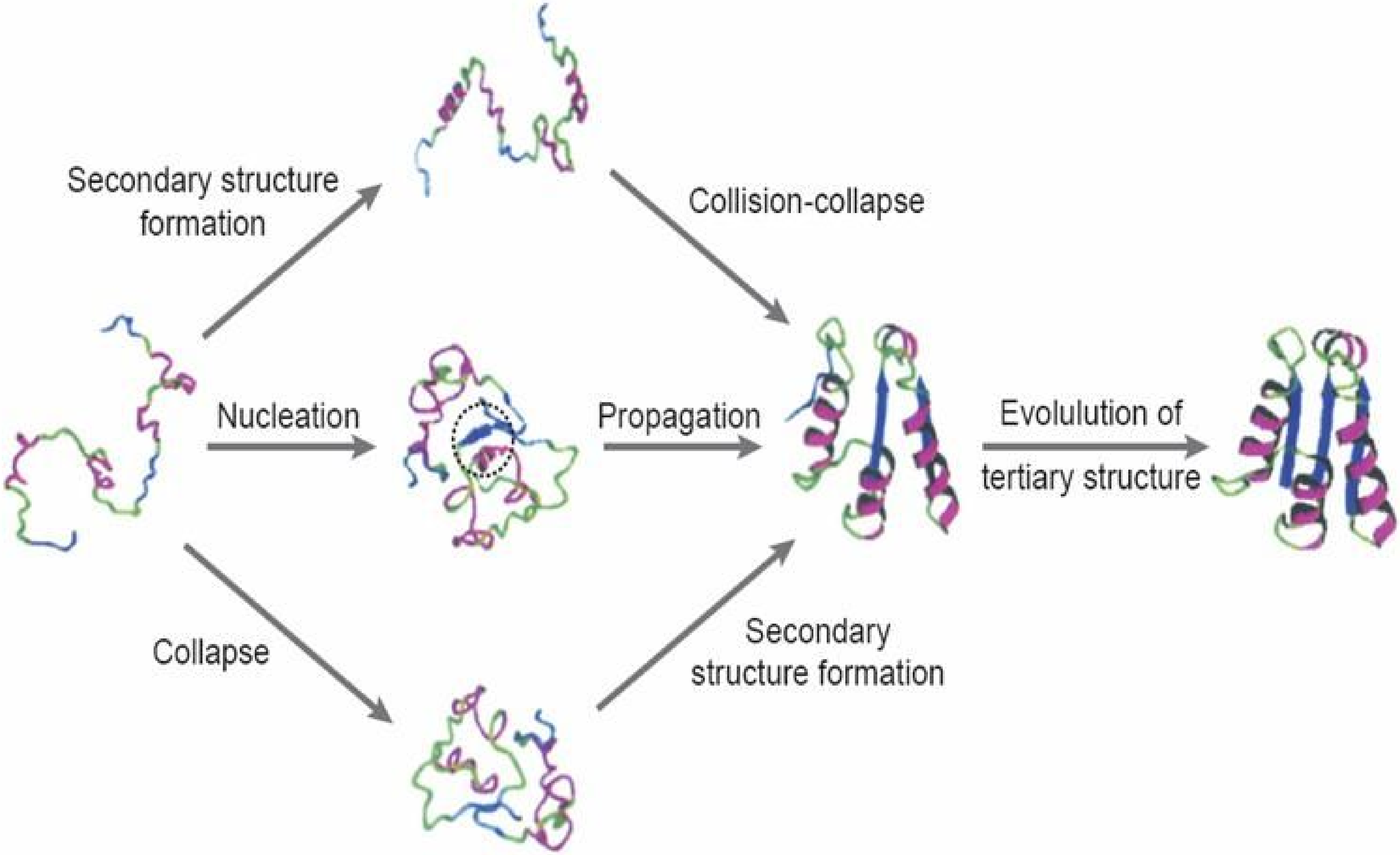
Initial chain collapse during the folding of monellin

Barrier-limited and barrier-less contraction of intramolecular distances



Patra, A.K. and Udgaonkar, J.B (2010)

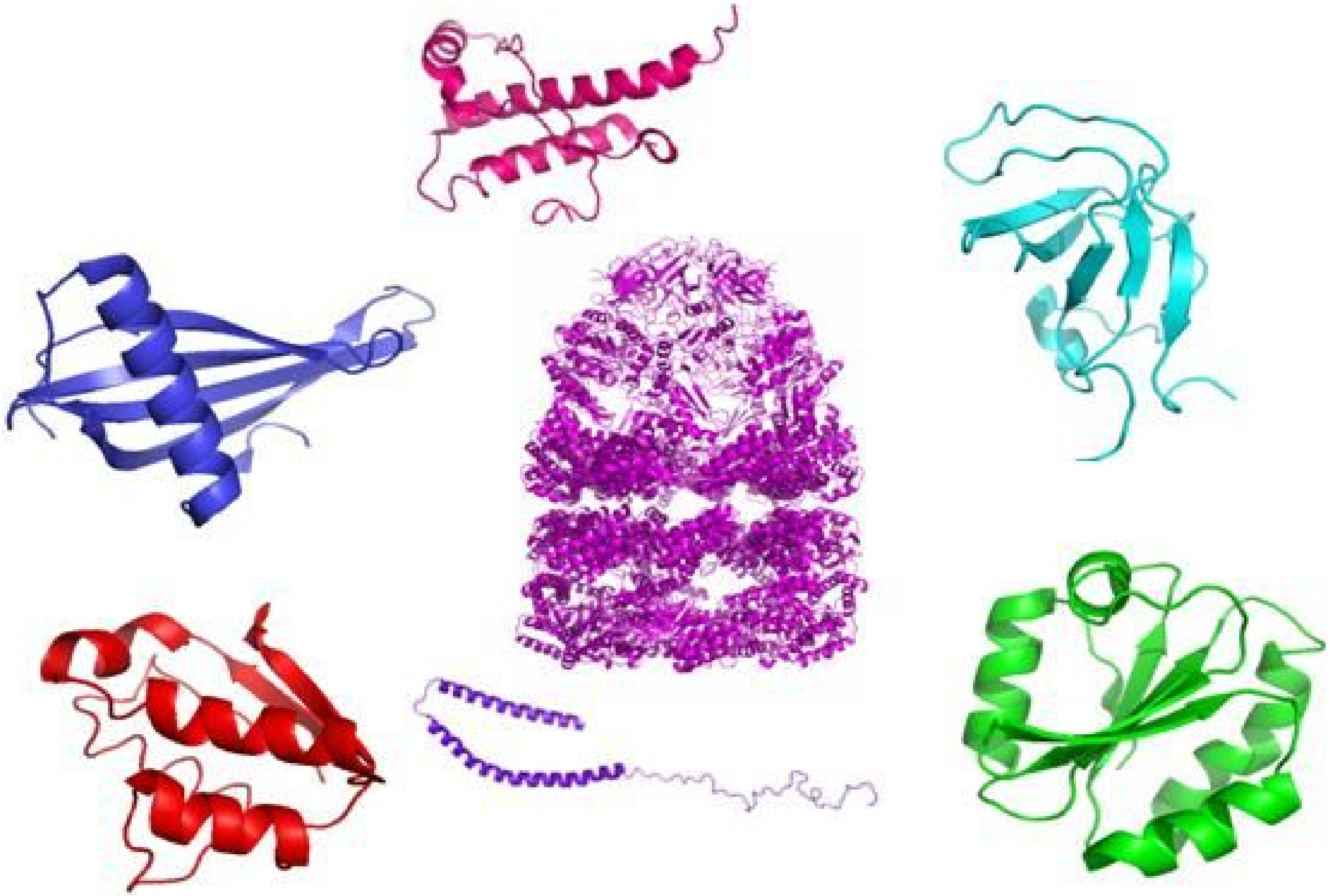
Pathways of Protein Folding



Udgaonkar, J.B. (2008) Annual Review of Biophysics

How do proteins fold, unfold and misfold?

JAYANT B. UDGAONKAR
NATIONAL CENTRE FOR BIOLOGICAL SCIENCES
TATA INSTITUTE OF FUNDAMENTAL RESEARCH



Nilesh Aghera
Amrita Dasgupta
Shweta Jain
Megha Kishore
Pooja Malhotra
Gayathri Ramachandran
Roumita Moulick
Amrita Sekhar
Vishal