The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

SMR/1845-1

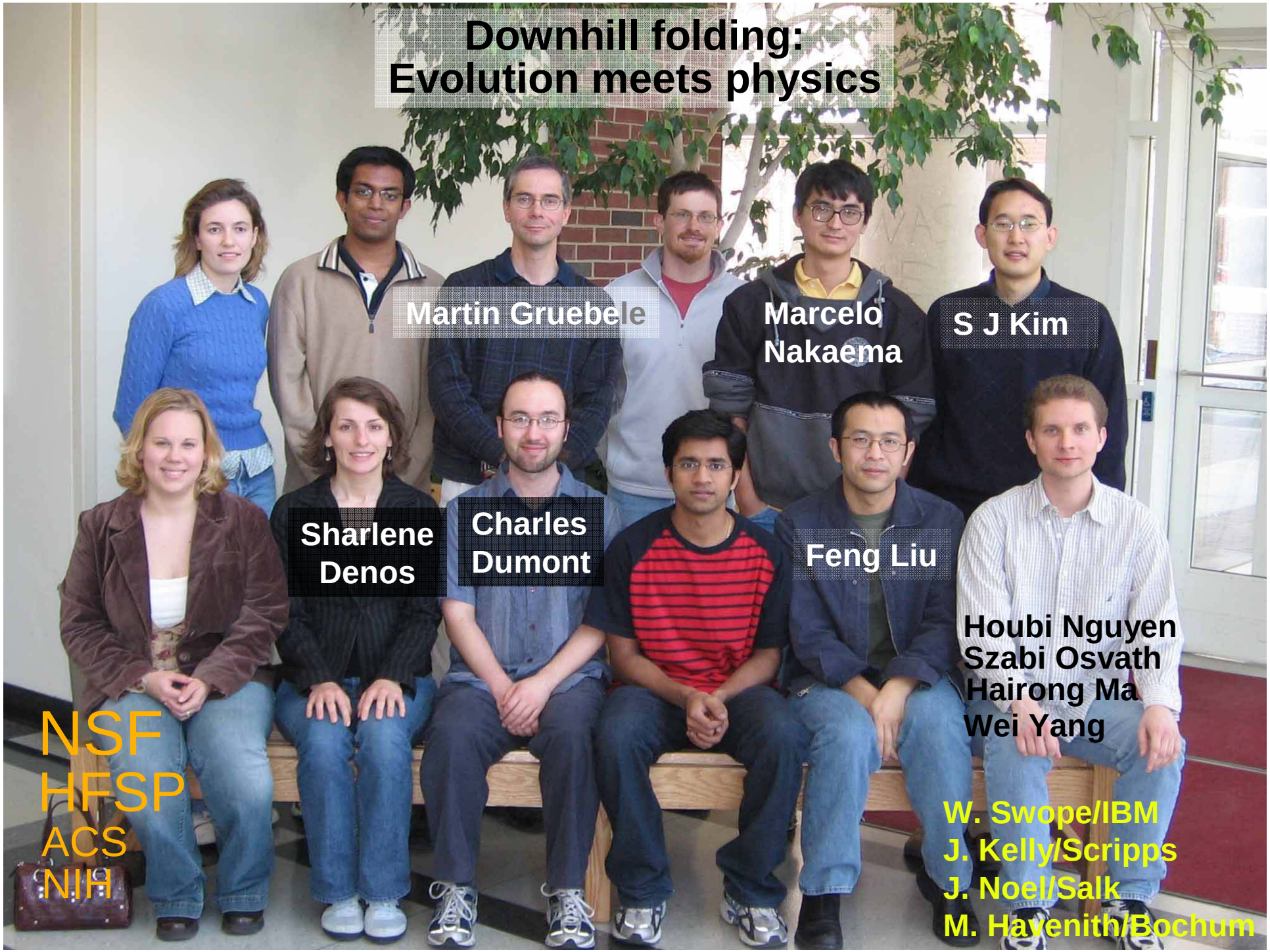
Conference on Structure and Dynamics in Soft Matter and Biomolecules: From Single Molecules to Ensembles

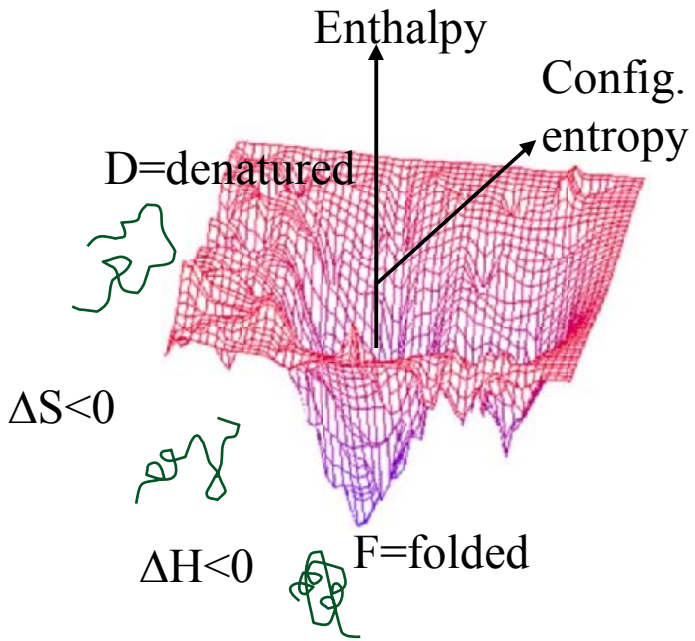
4 - 8 June 2007

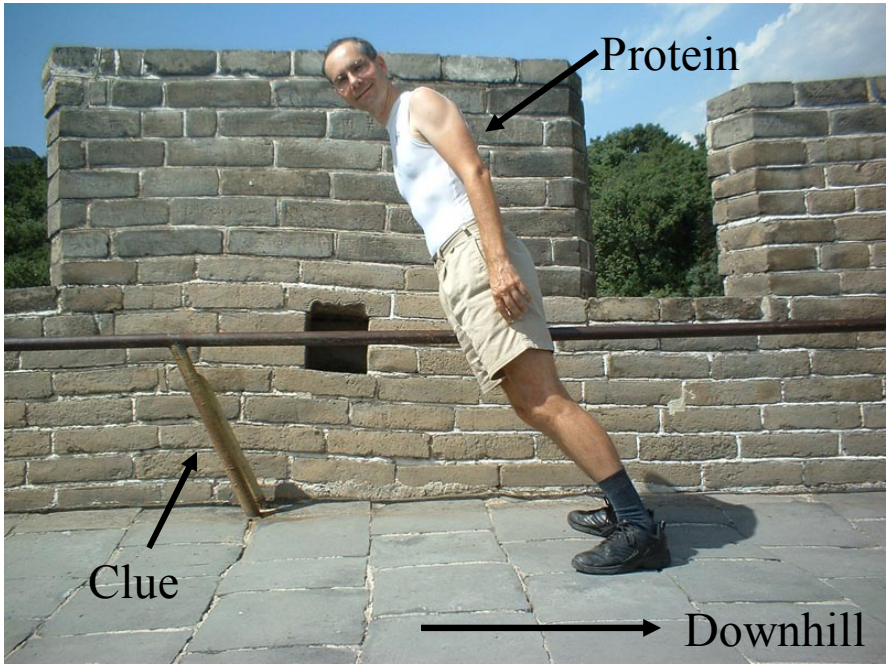
Downhill folding: Evolution meets physics

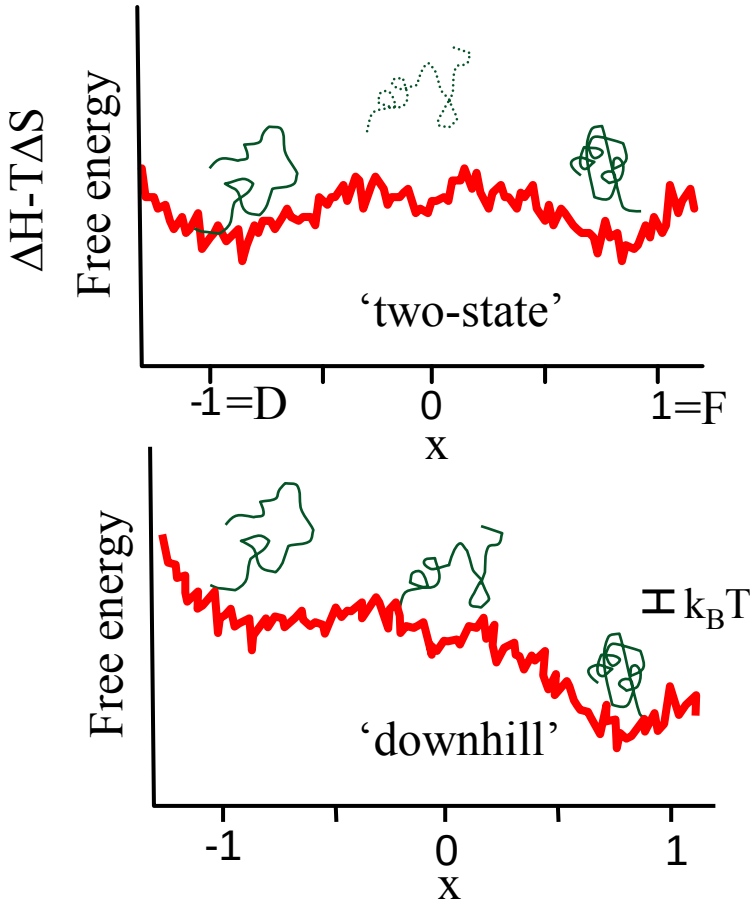
Martin GRUEBELE

University of Illinois at Urbana-Champaign
Department of Chemistry
600 South Mathews Ave.,
P.O. Box 5-6
IL 61801 Urbana


Downhill folding:
Evolution meets physics
Martin Gruebele
Marcelo Nakaema
S J Kim
Sharlene Denos
Charles Dumont
Feng Liu
Houbi Nguyen
Szabi Osvath
Hairong Ma
Wei Yang
NSF
HFSP
ACS
NIH
W. Swope/IBM
J. Kelly/Scripps
J. Noel/Salk
M. Havenith/Bochum







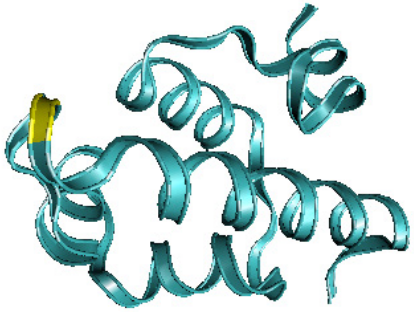
Interest:

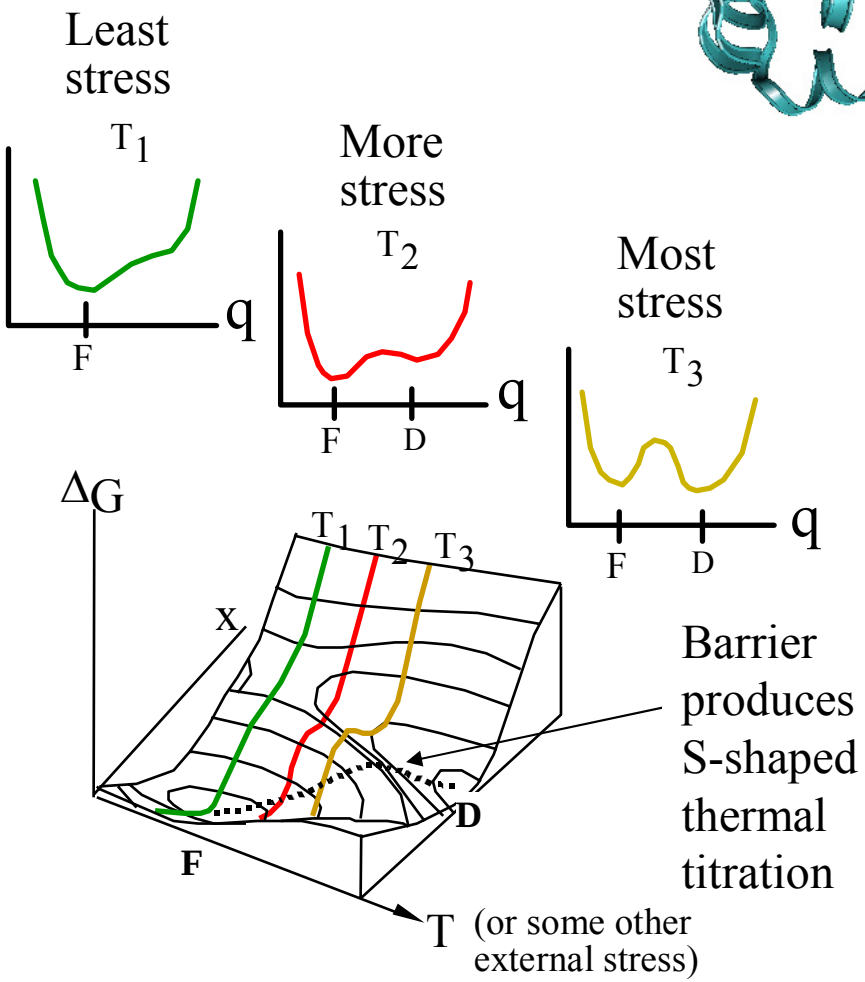
-ability to simulate folding by molecular dynamics (speed limited by diffusion)

-if physics is barrier-free, what causes barriers?

Thermodynamic measurements may not be sufficient: Downhill folding usually disappears under stress (denaturation)

ex: λ_{6-85}





Tuning the kinetics can provide the necessary proof:

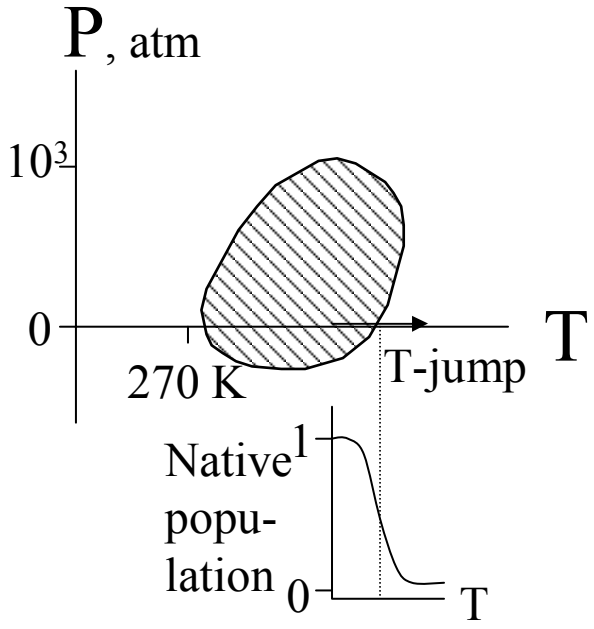
At T_3, kinetics interconvert 2 states: single exponential

At T_1, kinetics are diffusive (could be single exponential for simple diffusion, or stretched for more complex diffusion

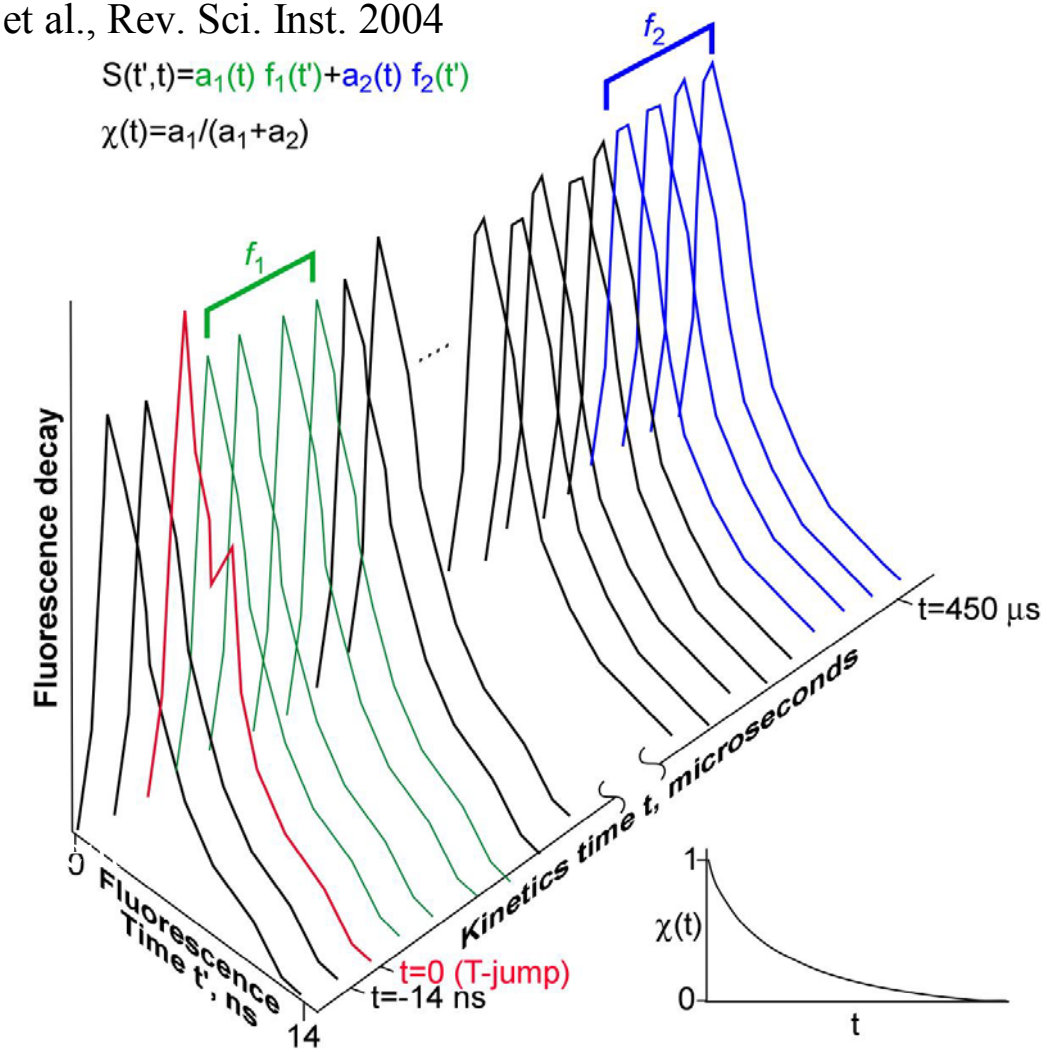
At T_2, see a mixture of the two and get the most complex kinetics.

Of course, folding speeds up from T_3 to T_1

How do you measure such fast kinetics: relaxation à la Manfred Eigen



Ballew et al., PNAS and Rev. Sci. Inst. 1996
Ervin et al., J. PCPB A, 2000
Ma et al., Rev. Sci. Inst. 2004



Fluorescence lifetime analysis

(14 ns dead time, 250 ps time resolution)

Fluorescence intensity analysis

(14 ns dead time)

Fluorescence spectral analysis

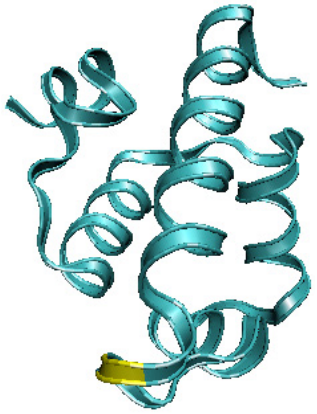
(SVD; 300 ns dead time)

Infrared spectral analysis

(SVD; 300 ns dead time)

Time course analyzed by singular value decomposition or nonorthogonal linear decomposition

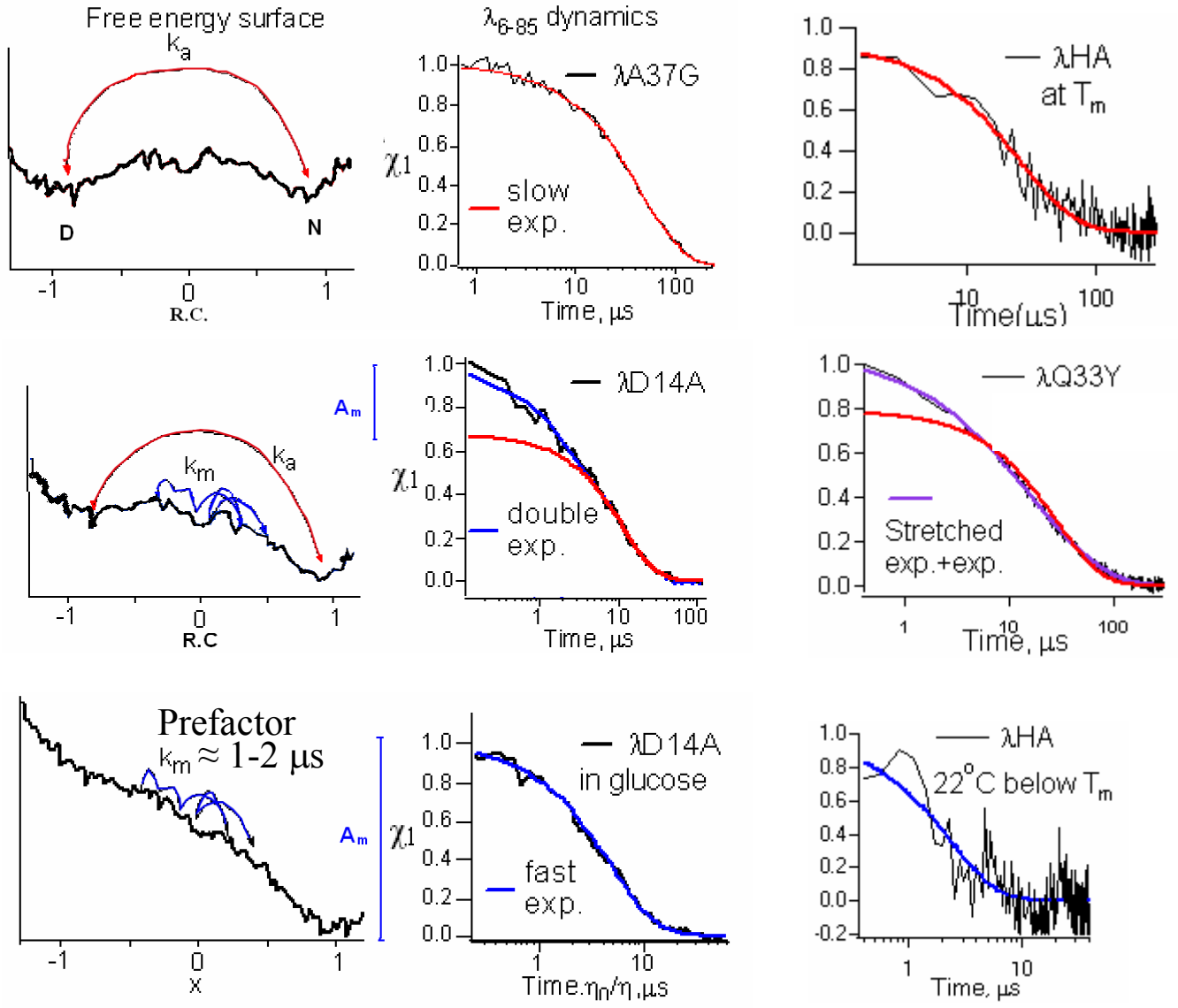
The full range: **exponential to complex to simple** as thermodynamic stress is relieved by T, mutation or solvent :



λ_{6-85}

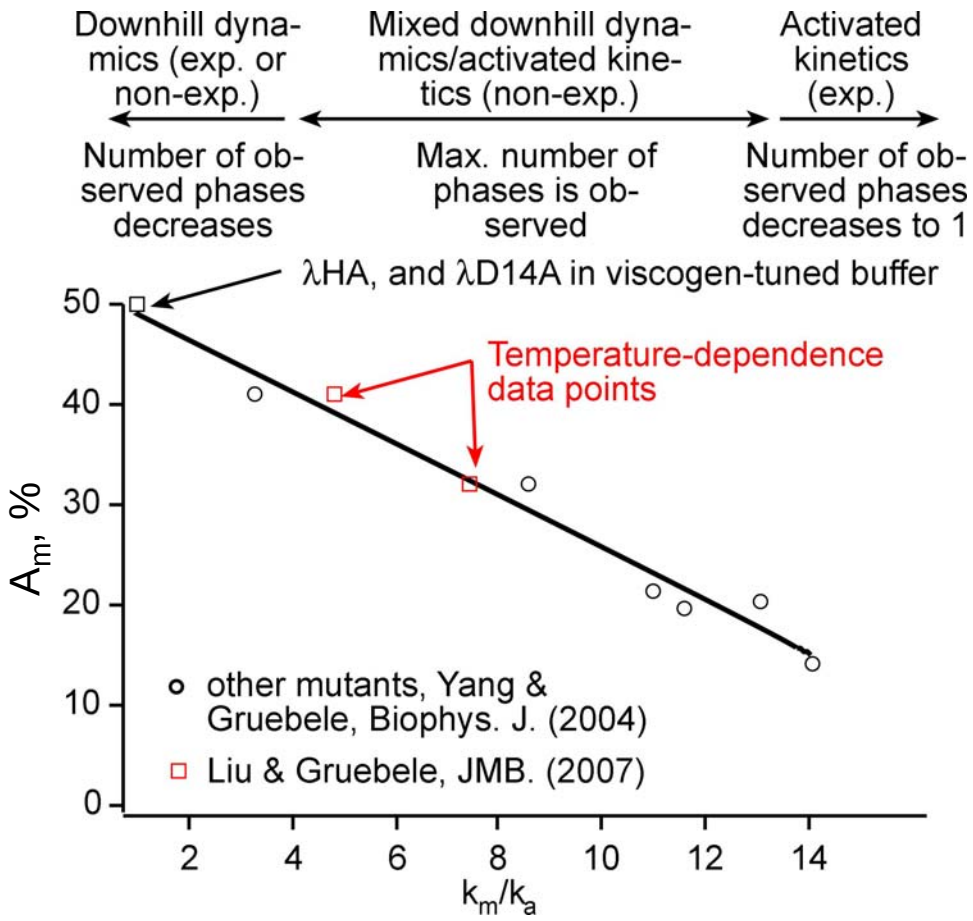
See also:
Oas & Co.

Native Bias



Distinguishing intermediates from downhill folders: **smooth correlation** of the molecular Phase k_m and the activated phase k_a:

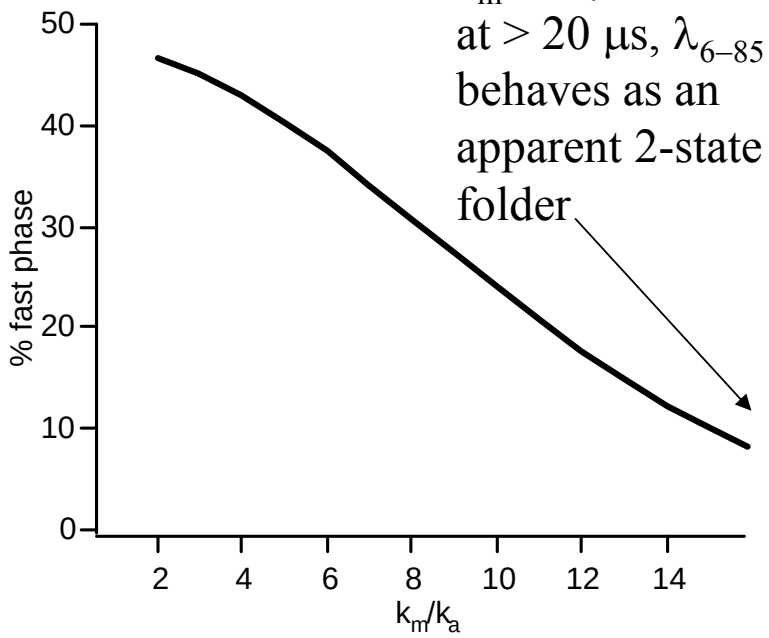
Experiment:



Liu & Gruebele, JMB 2007

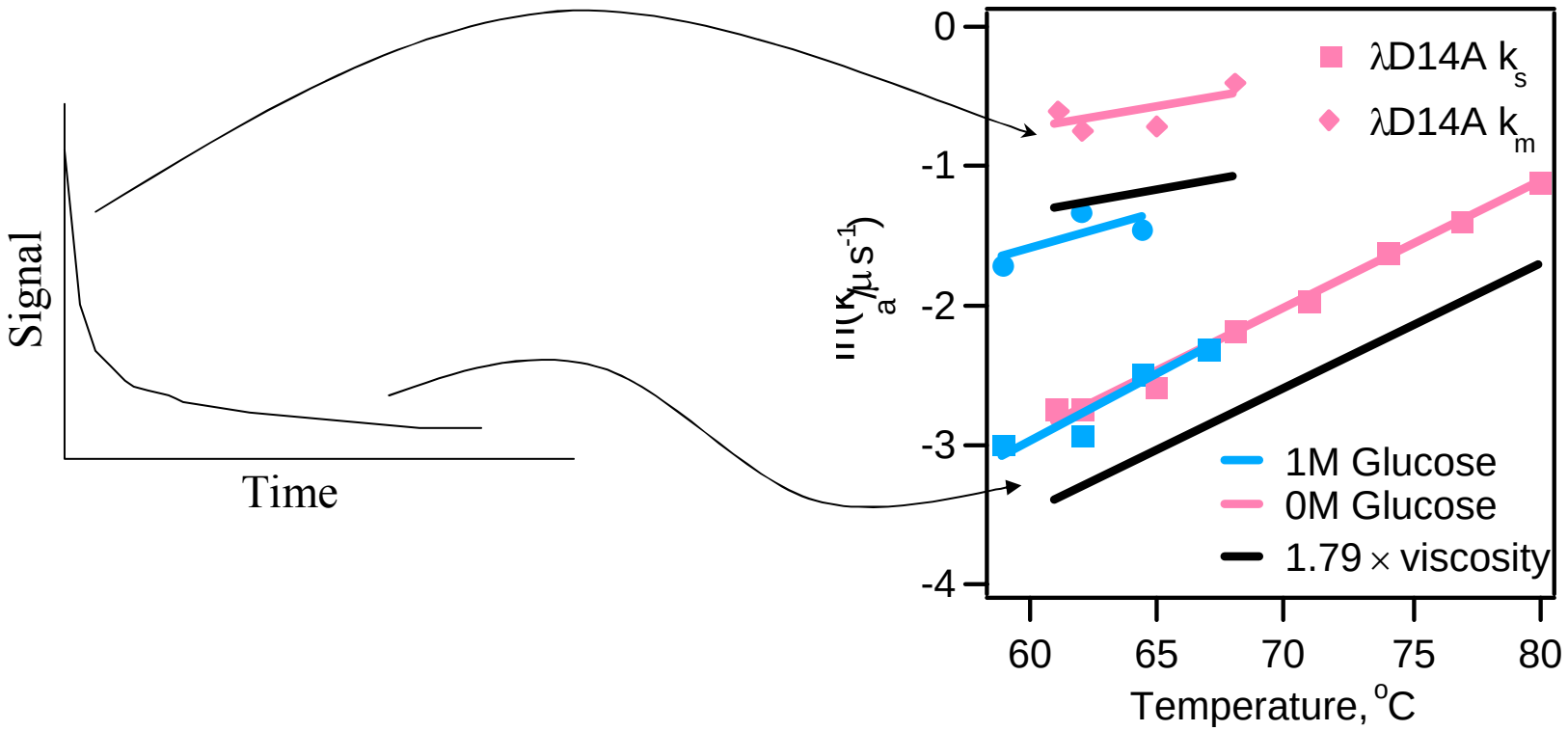
Theory:

Langevin dynamics on a double-well surface that becomes downhill when native bias (stability) increases yields this result:



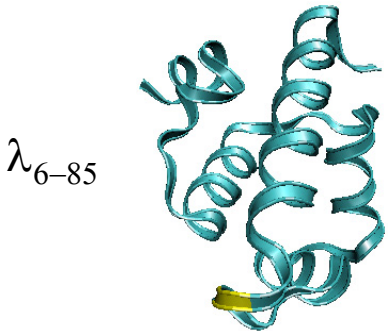
Ma & Gruebele, JCC 2006

Another key behavior: the faster the dynamics becomes,
the more it shows only a viscosity-dependence:

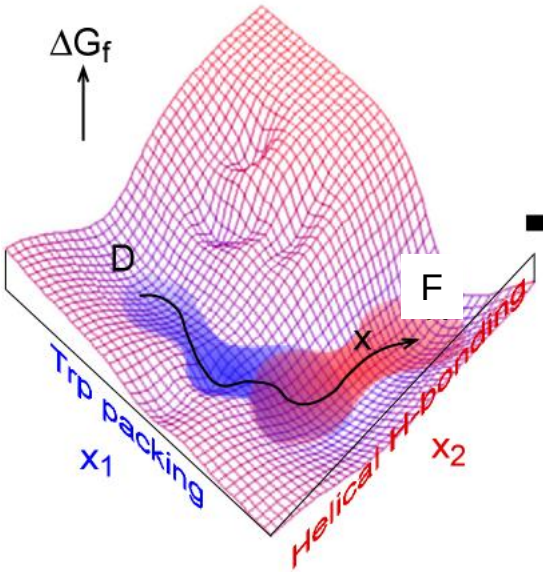


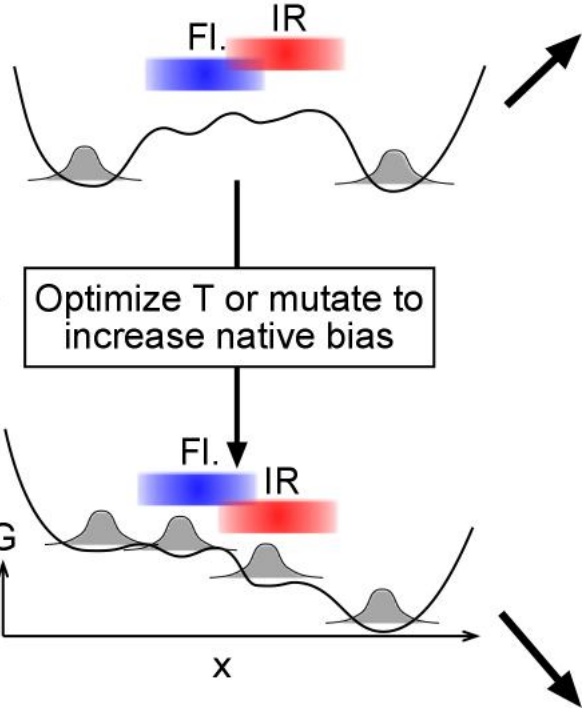
Similarly,

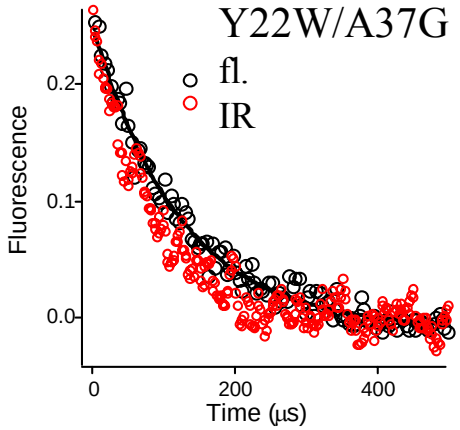
The difference between probes is also tuned by stress:

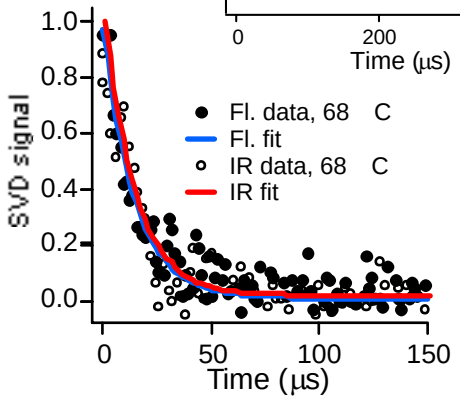


λ_{6-85}

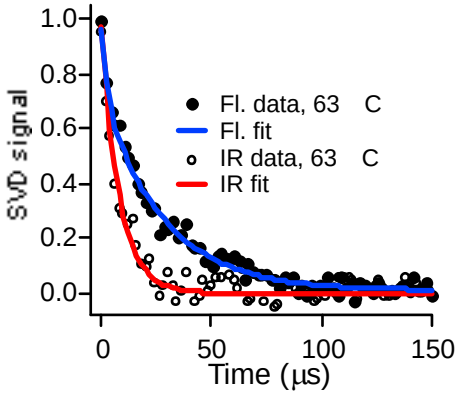








Y22W/Q33Y/
G46A/G48A



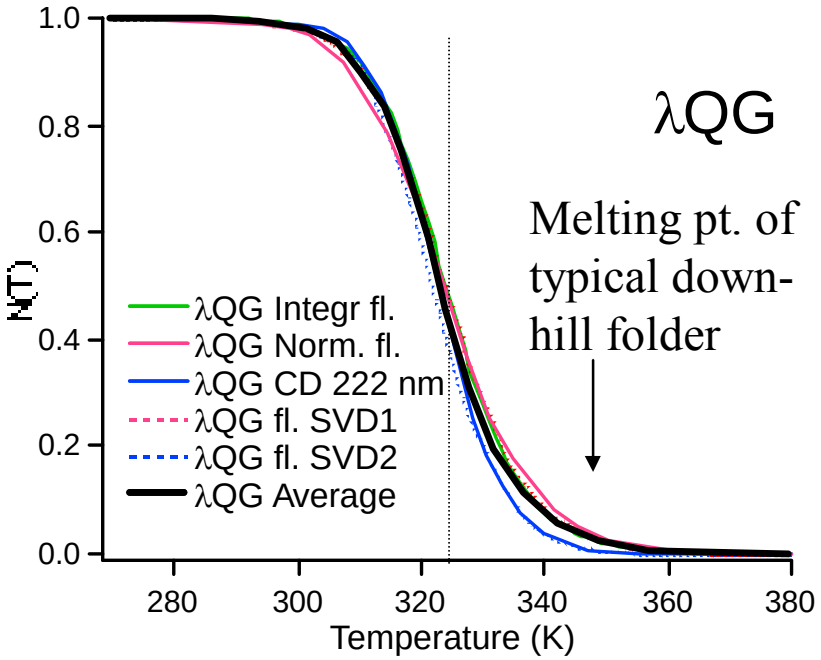
Ma & Gruebele, PNAS 2005, JCC 2005

Thus, in the BO&W scenario, **downhill folding turns to 2-state folding under stress**.

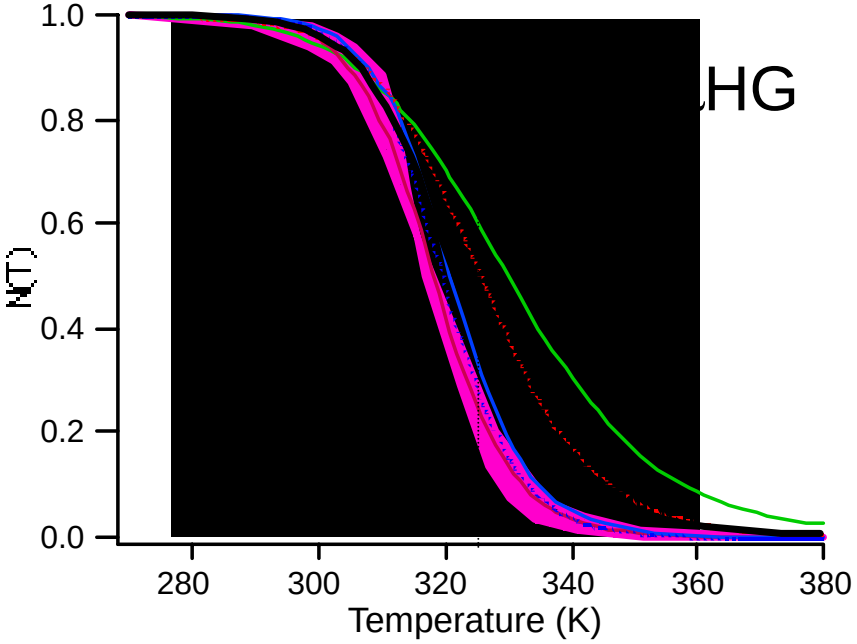
Possible reasons: energy required to expel a sheet of water upon hydrophobic core formation, or nonadditive forces? **[see posters by Badasyan & Ferguson from Chan group!]**

Possible 'cure': weakening the hydrophobic core/packing while maintaining very fast folding:

$\lambda_{6\text{-}85}$QG: 30 μs 2-state folder:

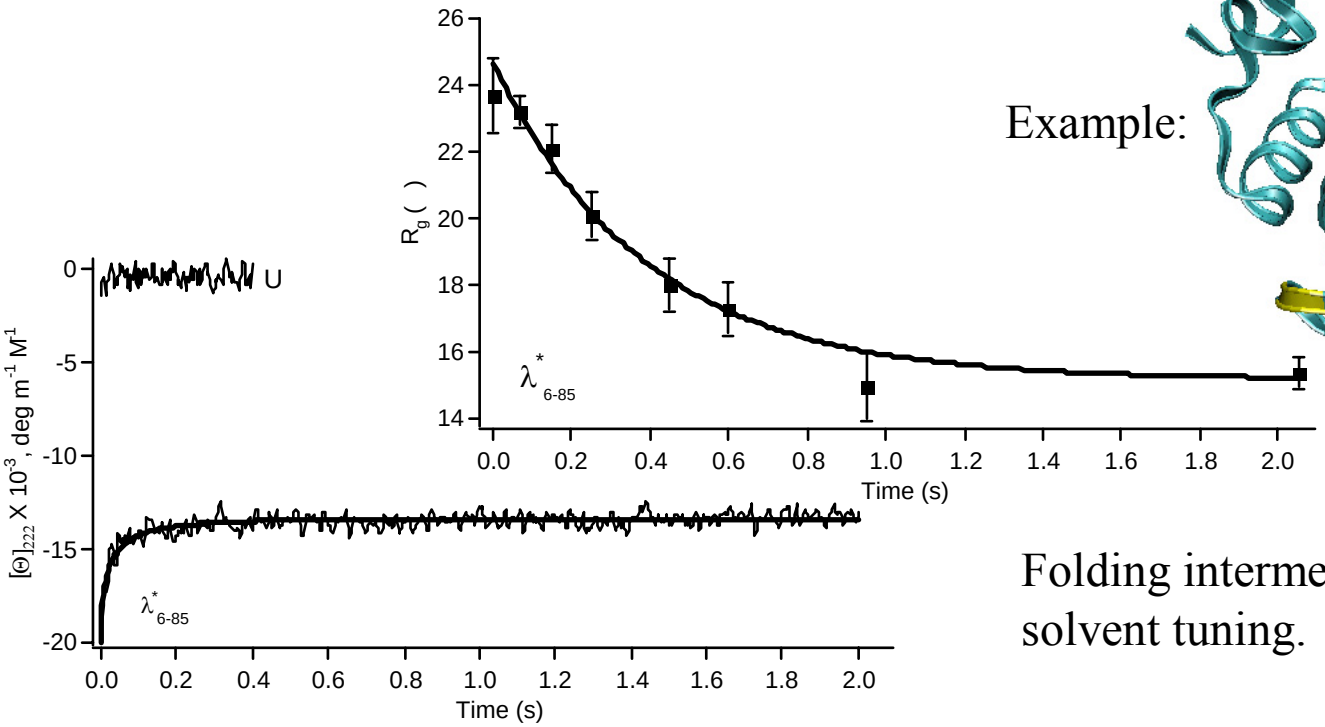


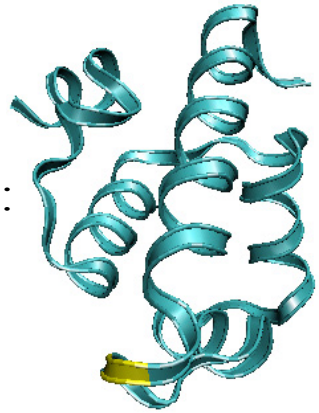
His33 $\lambda_{6\text{-}85}$: 10 μs folder at T_m, which is About the same as for λQG



Whether protein folding barriers are absent, or merely small, the consequence is the same:

We can tune proteins to fold by any mechanism!



Example:  again.

Folding intermediate stabilized by solvent tuning.

We have also tuned through all possible scenarios with a beta sheet protein, WW domain. (Liu & Gruebele, tbs 2007)

$\exp[-G(x)/k_BT]$ is a sensitive function of G(x) !

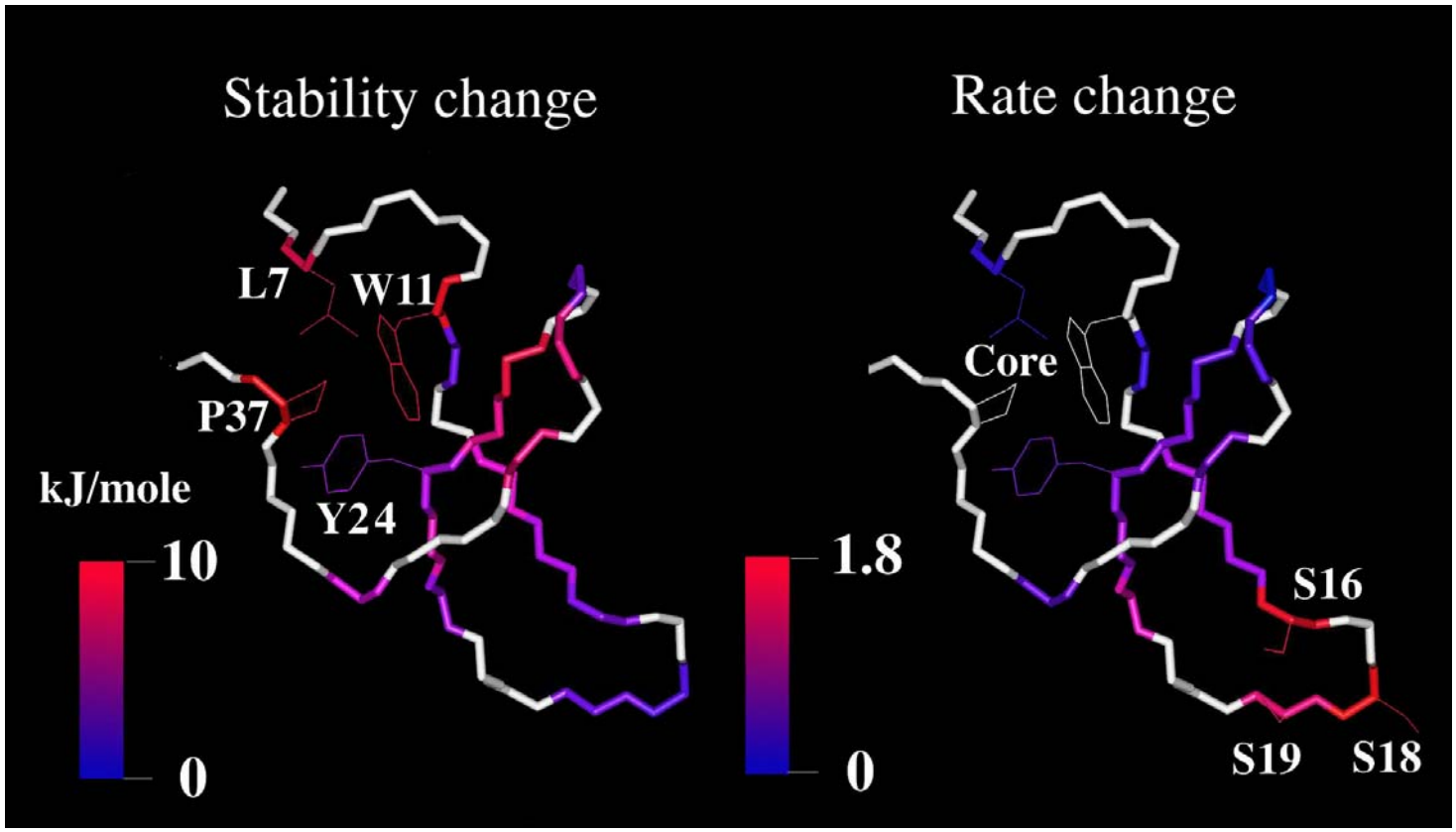
Makes sense from a perspective of evolutionary robustness

We have seen this smooth tuning behavior and speedup of folding for several proteins, α-helical, α /β and β-sheet.

Physical principles allow the same fold to form witout a TS, with 1 TS, or with many TS…

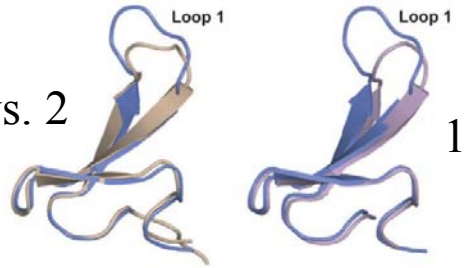
..how come natural globular protein domains fold over a barrier?
Answer: evolution for function and against aggregation

Reason # 1: Energetic frustration caused by evolution for function (Gruebele, COSB 2002)



Jäger et al. JMB 2001; collaboration with Kelly/Noel groups at Scripps

Crystal structure

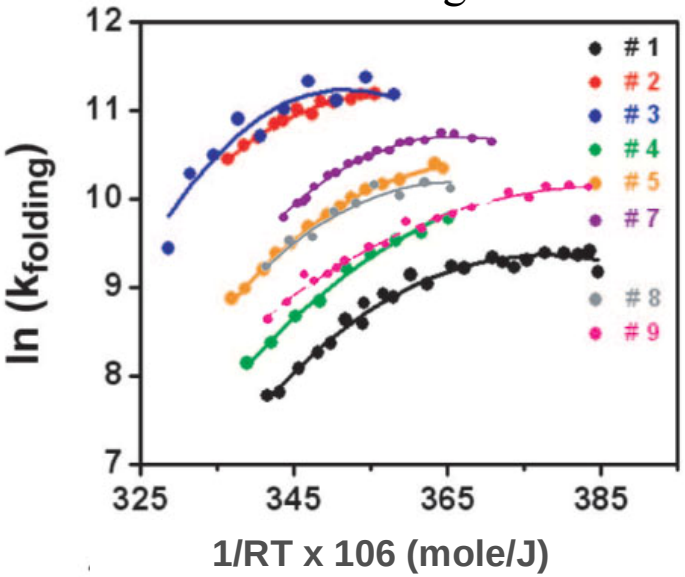


1 vs. 2

1 vs. 7

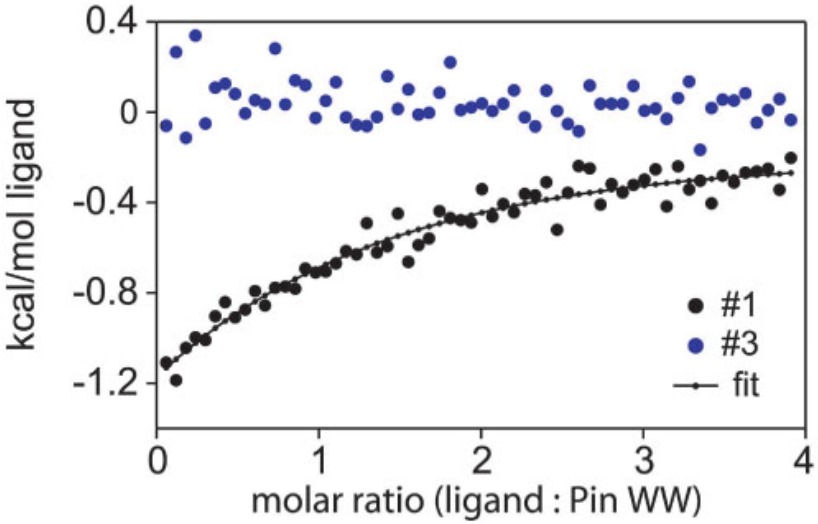
```
Variant          β1          β2          β3
  1  KLPPGWEKRMSRSSGRVYYFNHITNASQWERPSG  (4:6 loop)
  2  KLPPGWEKRMS-ADGRVYYFNHITNASQWERPSG  (3:5 loop)
  3  KLPPGWEKRMS-RDGRVYYFNHITNASQWERPSG  (3:5 loop)
  4  KLPPGWEKRMS-ARGRVYYFNHITNASQWERPSG  (3:5 loop)
  5  KLPPGWEKRMS-SSGRVYYFNHITNASQWERPSG  (3:5 loop)
  6  KLPPGWEKRMSRSS-RVYYFNHITNASQWERPSG  (3:5 loop)
  7  KLPPGWEKRMS--NGRVYYFNHITNASQWERPSG  (2:2 loop)
  8  KLPPGWEKRMS--RGRVYYFNHITNASQWERPSG  (2:2 loop)
  9  KLPPGWEKRMS--SGRVYYFNHITNASQWERPSG  (2:2 loop)
 10     SEWTERKT-ADGKTYYYNNRTLESTFEKP    (3:5 loop)
 11     SEWTERKSRSSGRTYYYNNRTLESTFEKP    (4:6 loop)
```

Folding



Biology is back in the picture: function causes energetic frustration

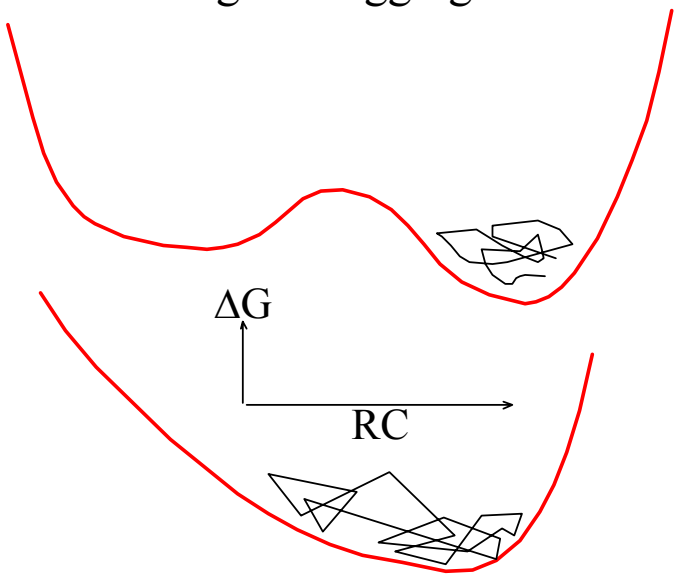
Function



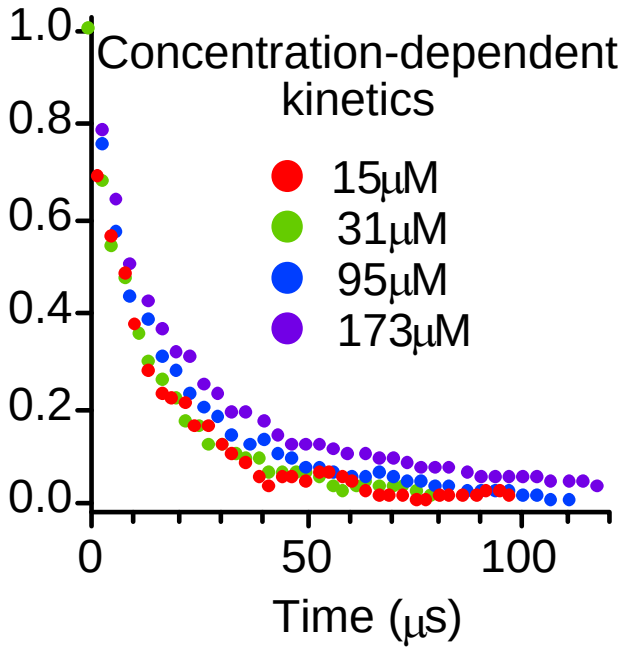
Jäger et al., PNAS 2006
(with Kelly and Noel groups)

Reason # 2 (F X Schmid, M Oliveberg, C Dobson,…):

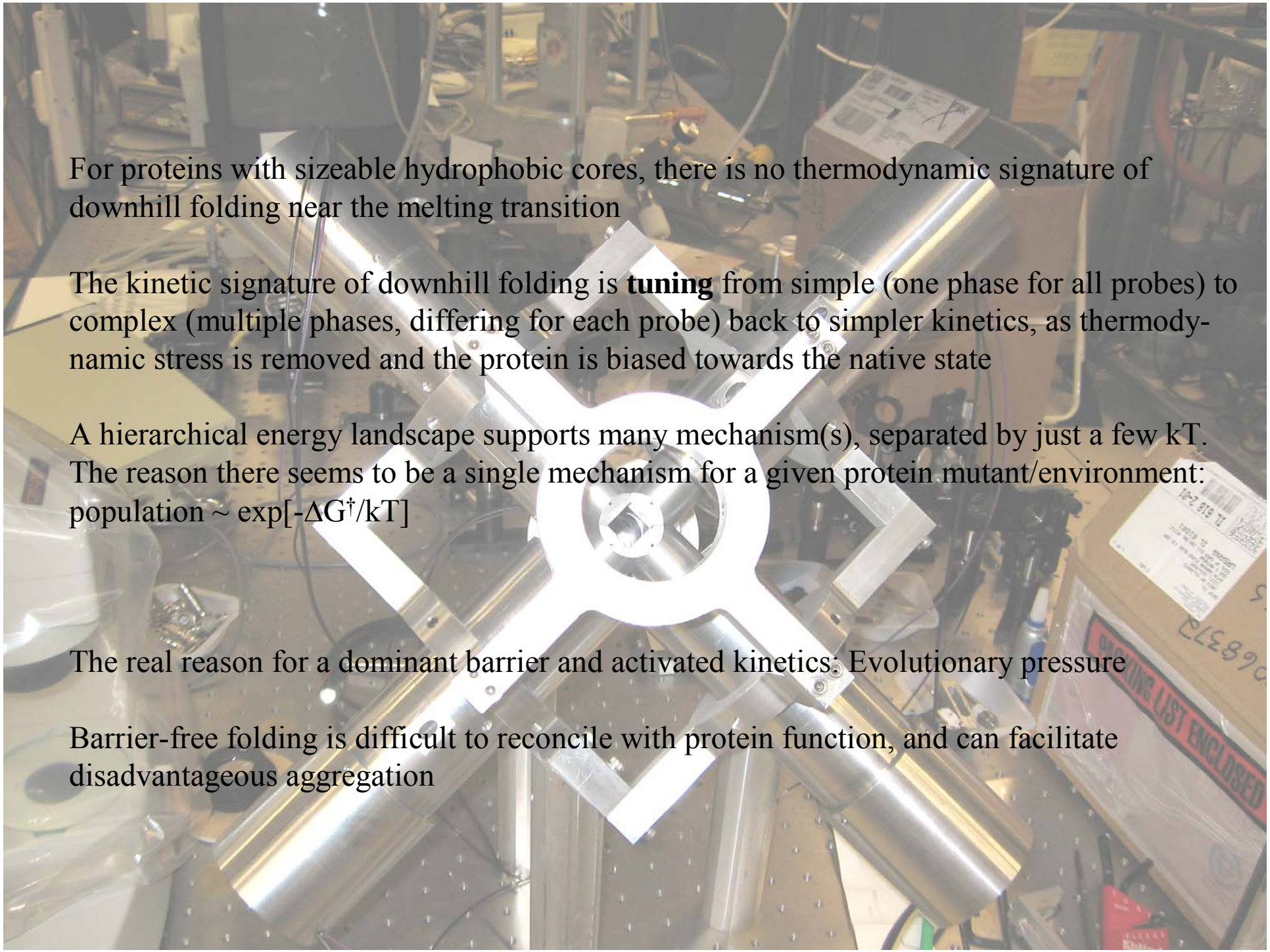
Evolution against aggregation



And we find? The more stable the 'downhill' folder, the more it aggregates



Biology is back in the picture: Proteins that fold faster also unfold faster and are more prone to aggregation.



For proteins with sizeable hydrophobic cores, there is no thermodynamic signature of downhill folding near the melting transition

The kinetic signature of downhill folding is **tuning** from simple (one phase for all probes) to complex (multiple phases, differing for each probe) back to simpler kinetics, as thermodynamic stress is removed and the protein is biased towards the native state

A hierarchical energy landscape supports many mechanism(s), separated by just a few kT. The reason there seems to be a single mechanism for a given protein mutant/environment: population ~ $\exp[-\Delta G^{\dagger}/kT]$

The real reason for a dominant barrier and activated kinetics: Evolutionary pressure

Barrier-free folding is difficult to reconcile with protein function, and can facilitate disadvantageous aggregation