Defining the transition state for folding – how much heterogeneity is there?

Michael Baxa and Tobin Sosnick, Department of Biochemistry and Molecular Biology, The University of Chicago

To determine the structure(s) of the transition state ensemble (TSE), we have applied our metal binding ψ analysis to a variety of small kinetically two-state proteins.

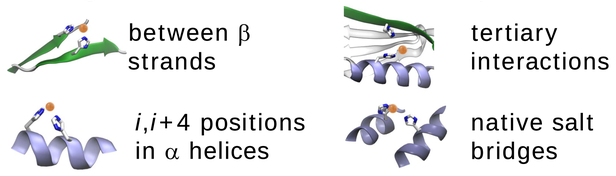
ψ analysis directly identifies residue-residue contacts in the TSE through the utilization of bi-Histidine metal binding sites.

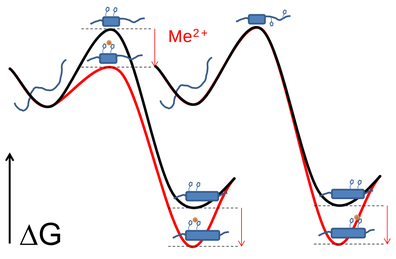
Because one can extrapolate to zero metal ion concentration, ψ analysis identifies the degree of structure formation *prior to the perturbation*.

The TSEs so determined have much more structure than those derived from traditional mutational ϕ analysis, in part due to the latter's reliance of inferring structure directly from energy perturbations, which can be challenging especially as the more flexible TSE can relax to accommodate a mutation relative to the native state.

What is ψ analysis?

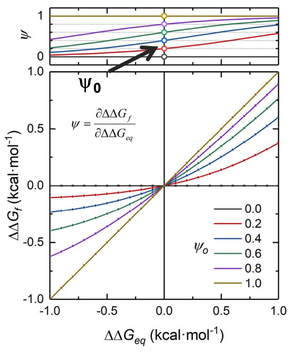
Introduce individual bi-Histidine (biHis) metal binding sites on the surface of the protein.





Metal ions stabilize binding competent confirmations.

Produces changes in k_f (ΔG_f) and stability, ΔG_{eq}.

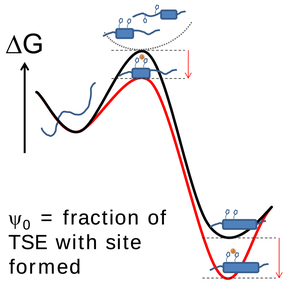


Varying [Me^{2+}] generates continuous Leffler plot

$$\Delta\Delta G_f = RT \ln\left((1-\psi_o)+\psi_o e^{\Delta\Delta G_{eq}/RT}\right)$$

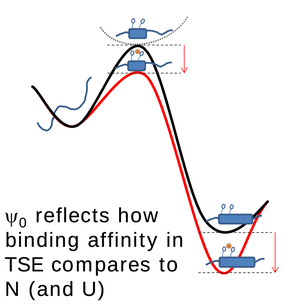
But what is ψ_0?
Reflects intrinsic degree of structure formation in *zero perturbation* – i.e. [Me^{2+}] = 0

Heterogeneous TSE



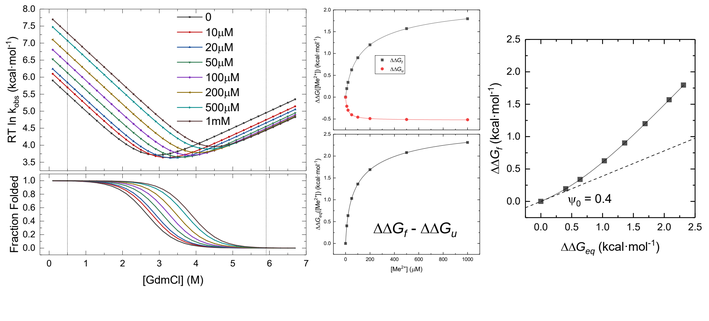
ψ_0 = fraction of TSE with site formed

Singular TSE



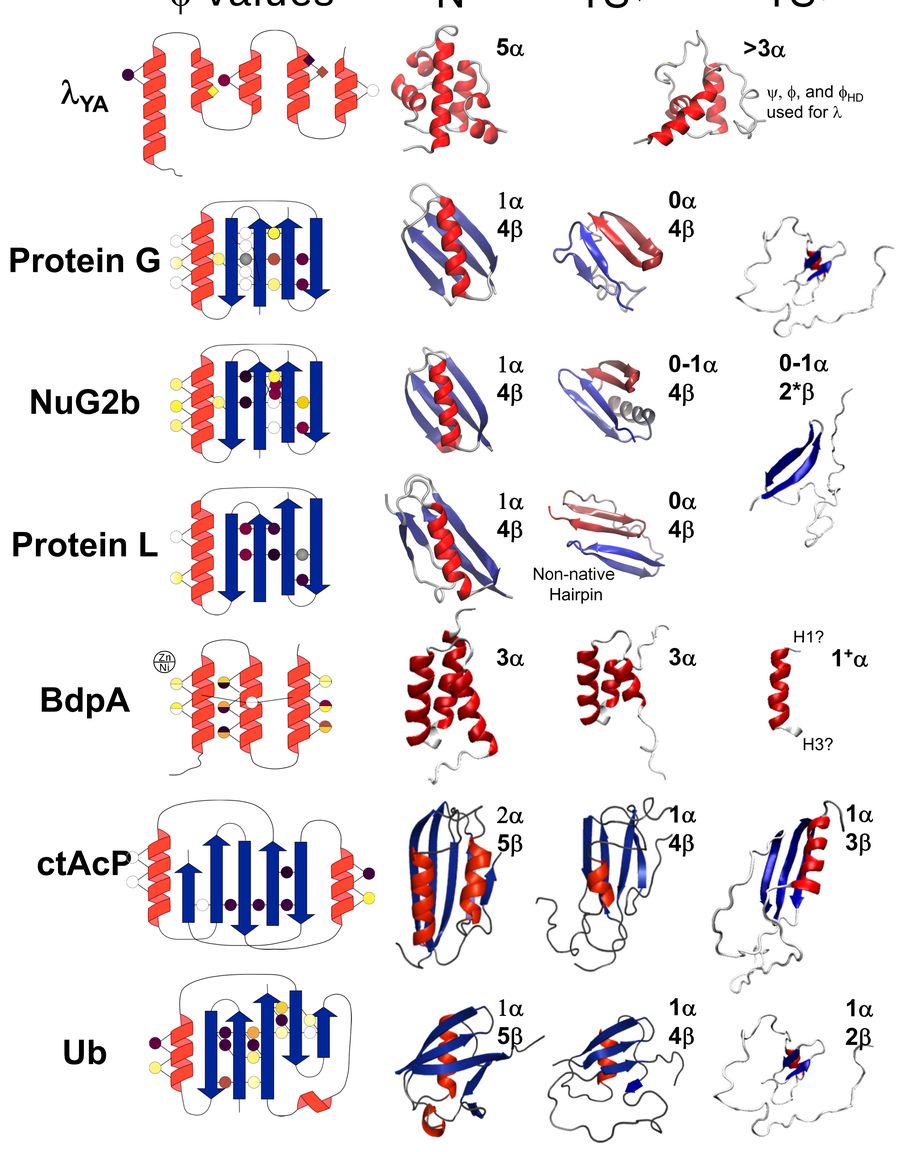
ψ_0 reflects how binding affinity in TSE compares to N (and U)

Ideal chevrons and equilibrium curves for ψ_0 = 0.4



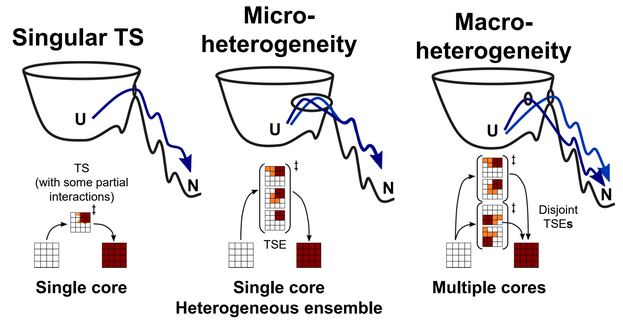
ψ values of several small proteins

TSEs are much more structured than what is deduced from traditional ϕ analysis. Further, every protein has at least 1 unity ψ value – limits extent of structurally disjoint TSEs.



How much heterogeneity is there?

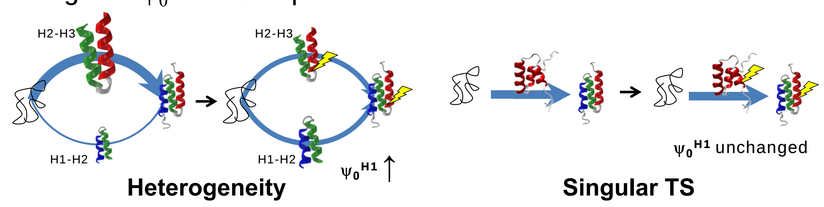
Can use ψ with other methods to distinguish between different levels of heterogeneity.



But every protein we have studied has at least 1 unity ψ value – at least one region is present in all TSEs.

Combine ψ with other methods to probe heterogeneity, e.g. (de)stabilize one region using mutation or crosslinking and re-measure ψ at another region.

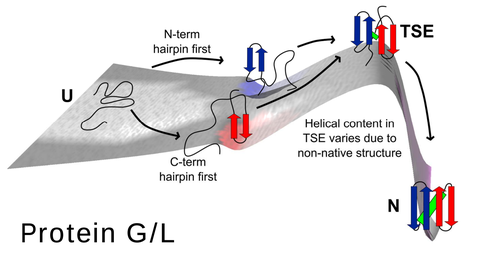
In a macro-heterogeneous model, destabilizing H2-H3 via mutation would increase the flux along the H1-H2 route, and hence lead to a higher ψ_0 in H1. In a single TS model, no change in ψ_0 in H1 is predicted.



Our results are consistent with other "pump-probe" approaches which provided no evidence of TSE heterogeneity.
(cp. Serrano et al., Fersht et al., Baker et al.)

Based on these data and the high RCO^{TSE}/RCO^{N} values, we believe that the major pathways have converged by the TS – i.e. minimal macro-heterogeneity.

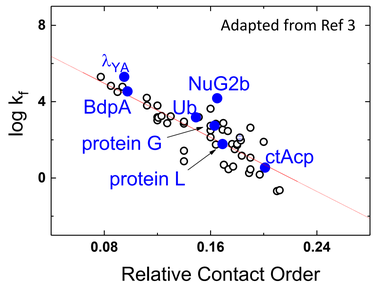
BUT there may be pathway diversity on the way up to the TSE.



Protein G/L

What have we learned from ψ?

Rationalizing the ln k_f - RCO^{native} correlation.



The TSEs of Ub, ctAcP, BdpA, prot G/L, NuG2b, and λ_{YA} are extensively structured and adopt a common high degree of the native state topology, i.e. **~70% of the native states' contact order** (RCO).

Fairly robust TSEs

We do not observe significant structural heterogeneity in the TSE at the level of secondary structural elements (4), in accord with other studies which similarly point to a robust TSE.

Rich dataset for simulations

Vast amount of experimental data provide opportunity for computational comparisons.

Protein	Reference
Lambda	Yu W, et al. (2016) Cooperative folding near the downhill limit determined with amino acid resolution by hydrogen exchange. Proc Natl Acad Sci U S A 113(17):4747-4752.
Protein G, NuG2b	Baxa MC, et al. (2015) Even with nonnative interactions, the updated folding transition states of the homologs Proteins G & L are extensive and similar. Proc Natl Acad Sci U S A 112(27):8302-8307.
Protein L	Yoo TY, et al. (2012) The folding transition state of protein L is extensive with nonnative interactions (and not small and polarized). J. Mol. Biol. 420(3):220-234
BdpA	Baxa M, Freed KF, & Sosnick TR (2008) Quantifying the Structural Requirements of the Folding Transition State of Protein A and Other Systems. J. Mol. Biol. 381:1362-1381.
ctAcP	Pandit AD, Jha A, Freed KF, & Sosnick TR (2006) Small proteins fold through transition states with native-like topologies. J. Mol. Biol. 361(4):755-770.
Ub	Krantz BA, Dothager RS, & Sosnick TR (2004) Discerning the structure and energy of multiple transition states in protein folding using psi-analysis. J. Mol. Biol. 337(2):463-475.

References

1. Yu W, Baxa MC, Gagnon I, Freed KF, & Sosnick TR (2016) Cooperative folding near the downhill limit determined with amino acid resolution by hydrogen exchange. *Proc Natl Acad Sci U S A* 113(17):4747-4752.
2. Baxa MC, *et al.* (2015) Even with nonnative interactions, the updated folding transition states of the homologs Proteins G & L are extensive and similar. *Proc Natl Acad Sci U S A* 112(27):8302-8307.
3. Plaxco KW, Simons KT, & Baker D (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* 277(4):985-994.
4. Shandiz AT, Baxa MC, & Sosnick TR (2012) A "Link-Psi" strategy using crosslinking indicates that the folding transition state of ubiquitin is not very malleable. *Protein Sci* 21(6):819-827.