Electronic Supplementary Material (ESI) for RSC Advances.

Supporting Information: Solvent accessible surface area-assessed molecular basis of osmolyte-induced protein stability

Shampa Raghunathan*

École Centrale School of Engineering, Mahindra University, Hyderabad 500043, India

E-mail: shampa.raghunathan@mahindrauniversity.edu.in

1 Methods and models

3-site model. Given that the three peptide interaction sites (i, j, k) are independent (i.e., interaction energies of each site are additive), the total degeneracy, Ω_{ijk}, of a specific microstate can be written as product of their SAs at particular peptide-interaction site (as proposed by Street et al., Proc. Natl. Acad. Sci. USA 2006, 103, 13997) :

$$\Omega_{ijk} = SA_i.SA_j.SA_k \tag{1}$$

In the case of aqueous-osmolyte solutions, the total SAs will have contributions from both the solvents. In Y molar (Y M) osmolyte solution:

$$SA_i = 55.5 \cdot SA_{w,i} + Y \cdot SA_{o,i} \tag{2}$$

$$SA_j = 55.5 \cdot SA_{w,j} + Y \cdot SA_{o,j} \tag{3}$$

$$SA_k = 55.5 \cdot SA_{w,k} + Y \cdot SA_{o,k} \tag{4}$$

Here, SA_w and SA_o refer surface areas coming from water and osmolyte, respectively. The probability of any specific microstate is given by

$$p_{ijk} = \frac{\Omega_{ijk} \quad e^{-E_{ijk}/k_B T}}{\sum_i \sum_j \sum_k \Omega_{ijk} \quad e^{-E_{ijk}/k_B T}} \quad , \tag{5}$$

where k_B is Boltzmann's constant and T is 298.15 K. These probabilities can be used to get the average energy of the system

$$\langle E \rangle = \sum_i \sum_j \sum_k E_{ijk}.p_{ijk} \quad , \tag{6}$$

and ΔG_{tr} is calculated by the difference between the average system energy at 0 and n M osmolyte concentrations

$$\Delta G_{tr} = \langle E \rangle_{water} - \langle E \rangle_{osmolyte} \quad . \tag{7}$$

The average osmolyte occupancy on an interaction site can be computed as

$$\langle O_{pref} \rangle = \sum_i \sum_j \sum_k p_{ijk} \left(\frac{SA_{o,i}}{SA_{o,i} + SA_{w,i}} + \frac{SA_{o,j}}{SA_{o,j} + SA_{w,j}} + \frac{SA_{o,k}}{SA_{o,k} + SA_{w,k}} \right) \tag{8}$$

whereas, in the bulk it is

$$\langle O_{bulk} \rangle = \left(\frac{SA_o}{SA_o + SA_w} \right) \quad . \tag{9}$$

The difference between $< O_{pref} >$ and $< O_{bulk} >$ will give the local osmolyte concentration.

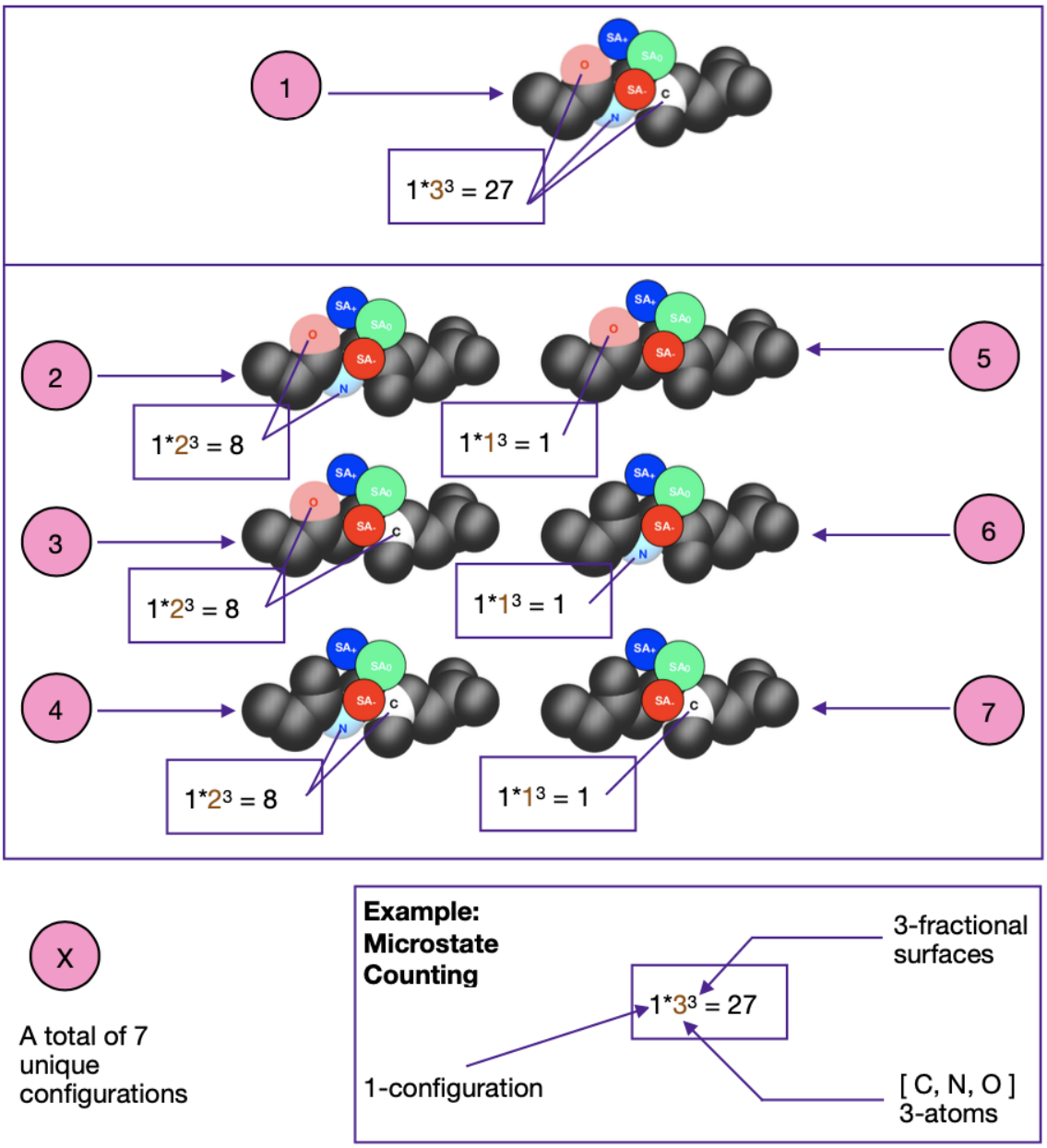
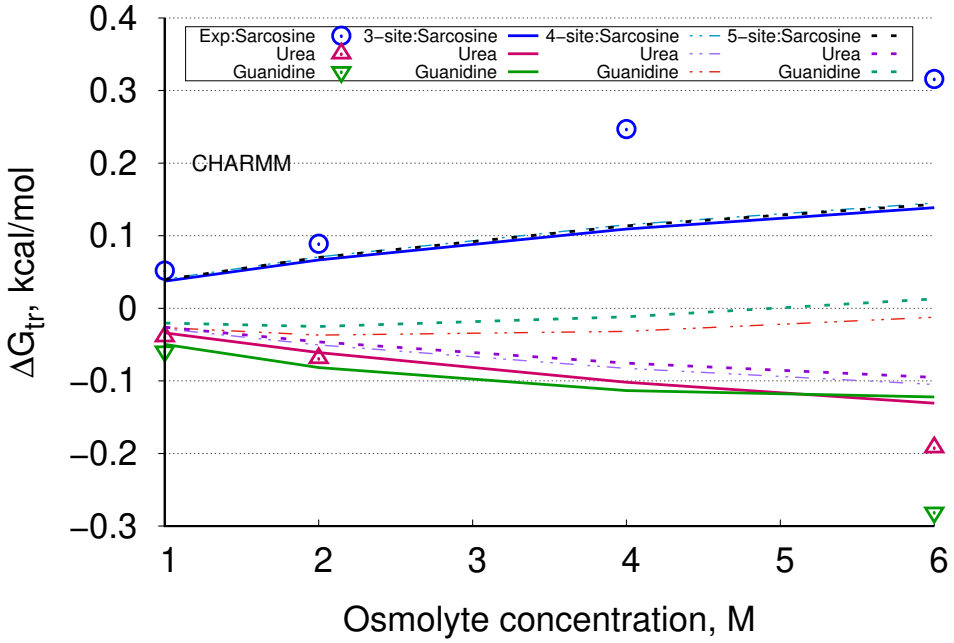


Figure S1: A 3-site model with solute atoms C, N, and O and solvent accessible surface areas with partially positive (SA_+), negative (SA_-), and neutral (SA_0) charges as proposed in the current work. Unique configurations (sets/sub-sets), and an example of microstate counting shows contribution from each term. Here, the 3-site model has a total of $(27 + 3\times8 + 3\times1) = 54$ microstates.

(a)



(b)

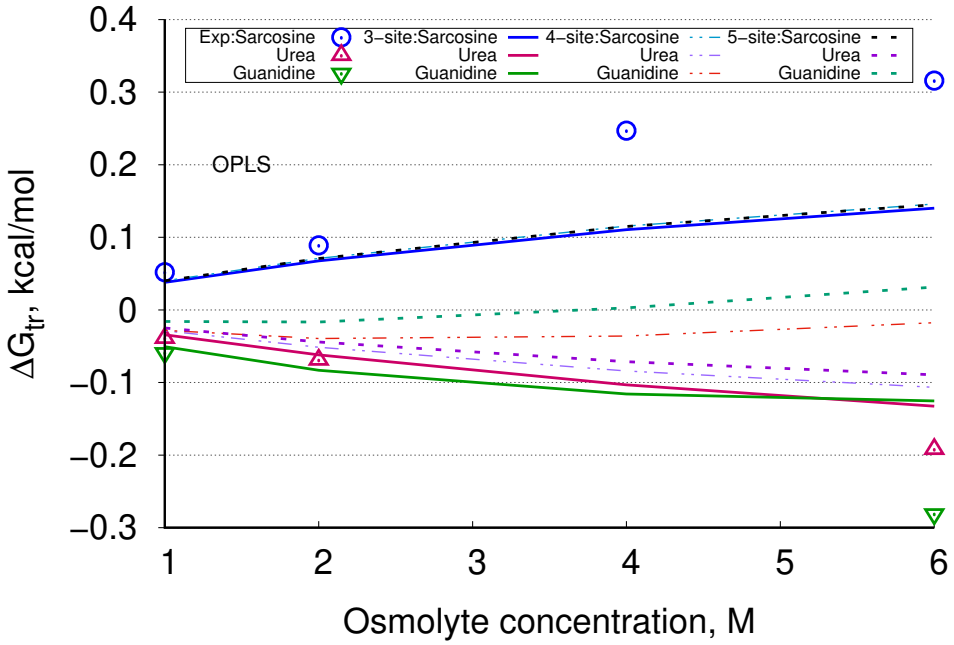


Figure S2: Calculated ΔG_{tr} values for the backbone model as osmolyte concentration is increased from > 1 M. Available experimental data are plotted using blue circles (sarcosine), magenta triangles (urea), and green inverted triangles (guanidine). Lines represent values calculated in this study using (a)CHARMM and (b)OPLS partial charges for various 3-site, 4-site and 5-site models.

2 Partial charges

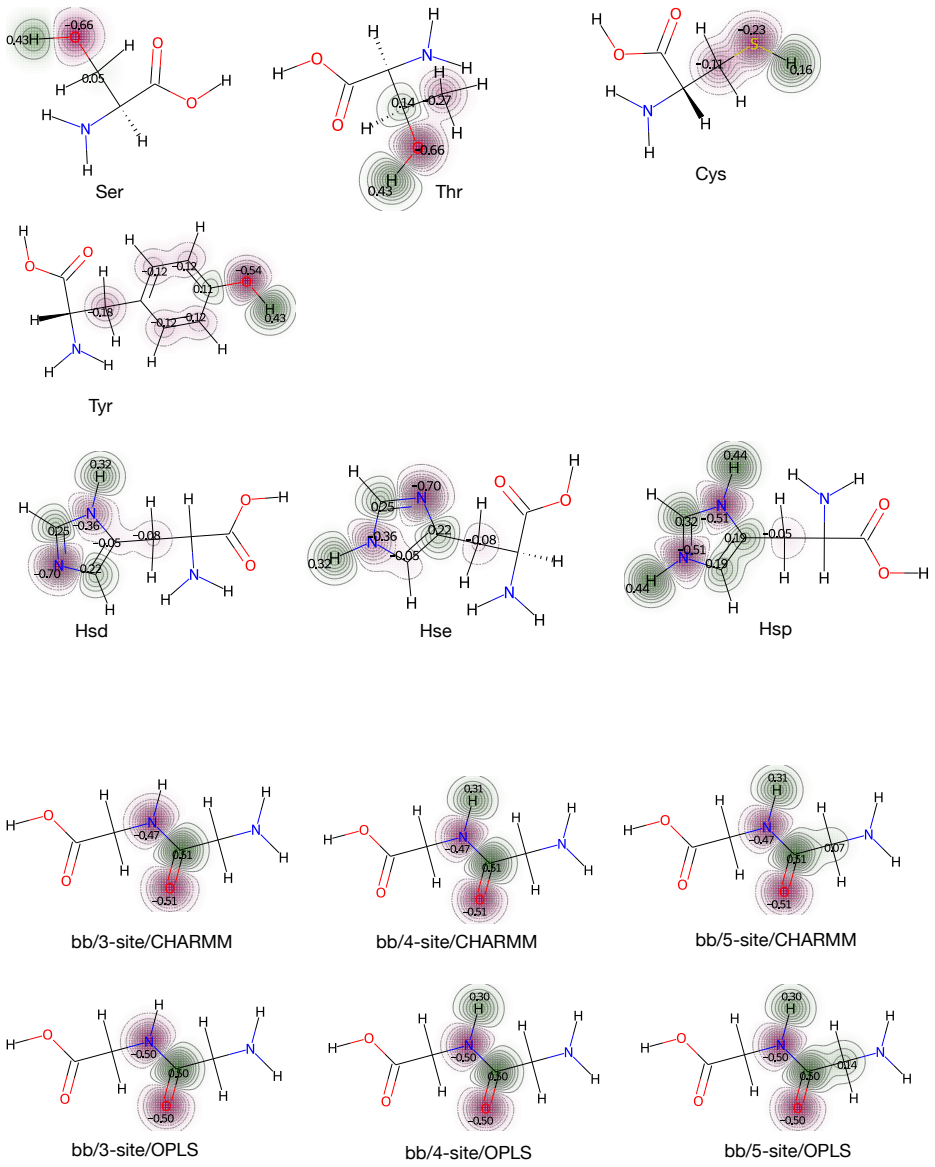


Figure S3: **Molecular electrostatics of amino acid side chains and backbone.** Atomic partial charges used in the CHARMM force field for Ser, Thr, Cys, Tyr, Hsd, Hse, and Hsp with an additional charge site on hydrogen atom. Backbone as 3-site, 4-site, and 5-site models with CHARMM/OPLS partial charges.

3 ΔG_{tr} values for n-site models including hydrogen as an additional charge site

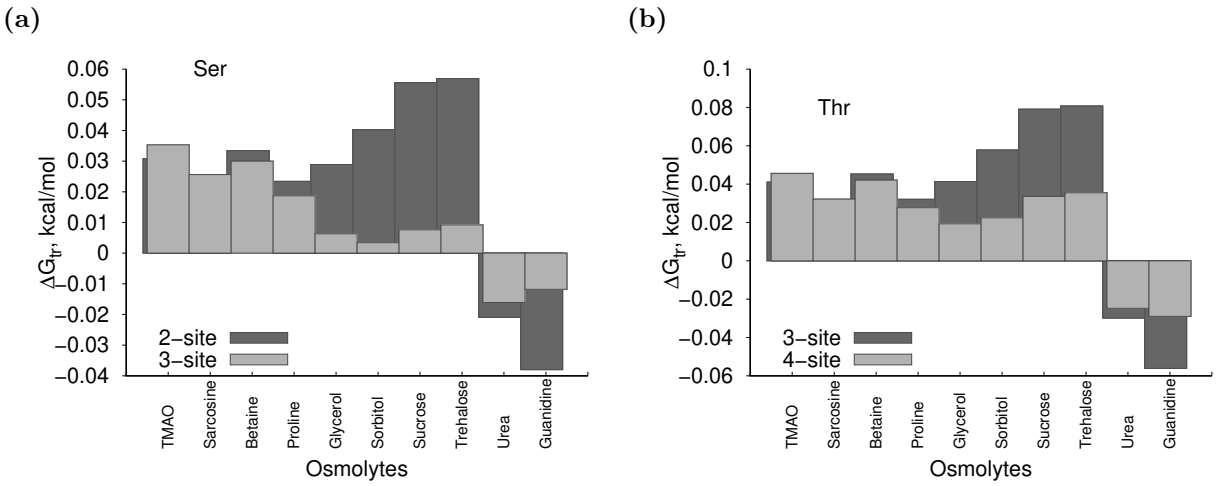


Figure S4: Computed ΔG_{tr} values for varying n-site model of (a) Ser and (b) Thr.

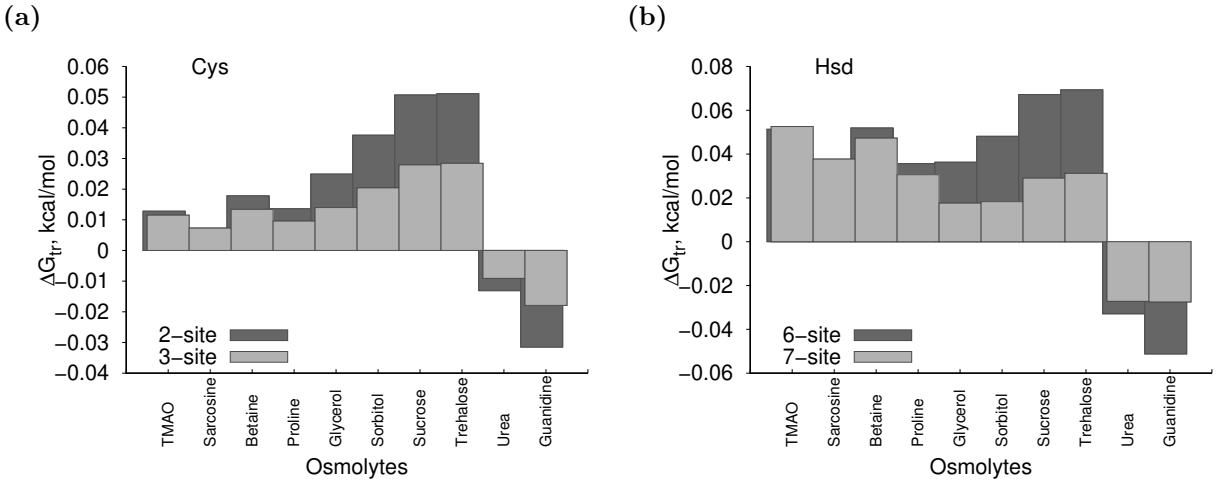


Figure S5: Computed ΔG_{tr} values for varying n-site model of (a) Cys and (b) Hsd.

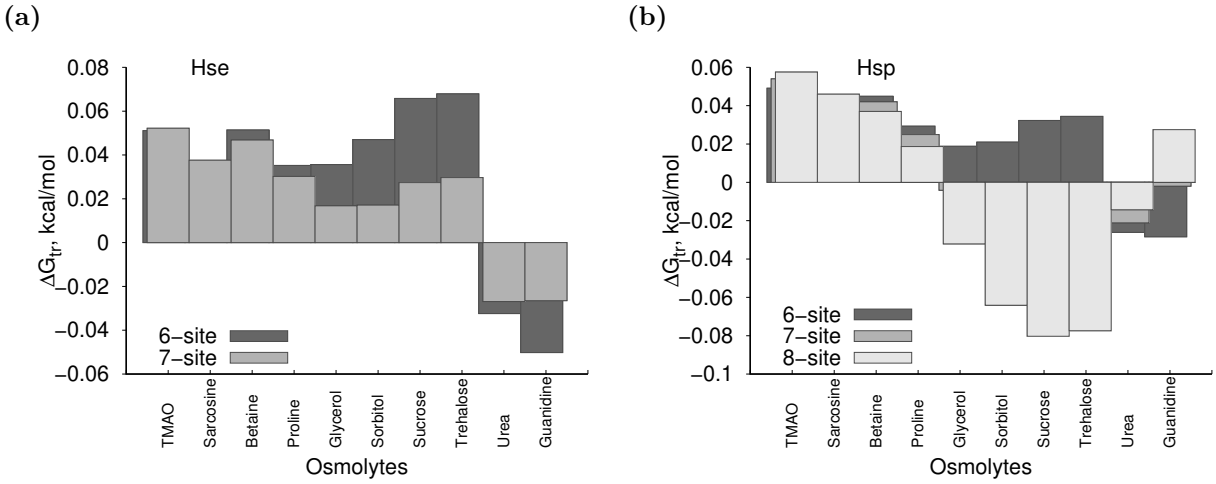


Figure S6: Computed ΔG_{tr} values for varying n-site model of (a) Hse and (b) Hsp.