Answers to Problem Set #2:

(1) (a) The unfolding curves are equilibrium measurements that monitor the amounts of various protein species present at each denaturent concentration. The two–state assumption says that only "native" and "unfolded" species are present at observable concentrations during the unfolding process. Then at any specific denaturant concentration, the y–axis value will be just $y_N f_N + y_U f_U$, where y_N and y_U are the y–values for pure native and pure unfolded, f_N and f_U are the fractions of native and unfolded present at that concentration. Clearly f_N and f_U are a function of the denaturant concentration and are independent of the y–axis property being monitored. If the y–axis values for different properties are normalized to the same scale ($y_N = 100$ and $y_U = 0$ in this problem), then the unfolding curves must be *identical* regardless of the property being monitored. In this case, the CD–based curve is clearly very different from the other two. Thus, the two–state assumption cannot be valid for penicillinase. At least some third species, an "intermediate", must be present to cause non–coincidence of the normalized curves.

(b) The CD–based unfolding curve is monitoring the presence of secondary structure as a function of denaturant. The viscosity curve is largely measuring the compactness of the structure. We have proved above that intermediate(s) must be present in large enough quantity to have caused non–coincidence of the the curves. Then at a guanidinium chloride (GuHCl) concentration of about 1M we appear to have a collection of conformational states that retains significant secondary structure while being almost as extended as the "unfolded" state. This reasoning would tend to support folding models in which intermediates acquire some or most secondary structure before finally folding to a fully compact structure; ie, the diffusion–collision–adhesion or framework models.

(2) (a) If ΔH and ΔS do not depend on temperature, then they can be treated as "constants" with respect to temperature variations. Then ΔG will vary linearly with temperature change, and the multiplicative constant for this change will just be $-\Delta S$. In terms of simple differential calculus, the derivative of $f(x) = a - bx$, with a and b constant, is just $df/dx = -b$. This corresponds directly to $d(\Delta G)/dT = -\Delta S$.

(b) At the melting temperature, which is the midpoint of the transition between the folded and unfolded states, we have [folded] = [unfolded] and K = 1. Thus, $\Delta G_m = -RT \ln(1) = 0$ kcal/mole. From part (a), the slope at T_m of the ΔG vs. T plot is just $-\Delta S_m$; ie, $\Delta S_m = +300.3$ cal/mole/degree. Finally, we have $\Delta G_m = 0 = \Delta H_m - T_m \Delta S_m$, so that $\Delta H_m = T_m \Delta S_m = (48.3 + 273.2)(300.3) = 96.55$ kcal/mole. Note that the entropy of protein unfolding is favorable, while the enthalpy of unfolding is unfavorable.

(c) We have $\Delta G = -RT \ln(K) = \Delta H - T\Delta S$. Dividing both sides of the middle equality by $-RT$ and simplifying gives:

$$\ln(K) = (-\Delta H/R) \ (1/T) + (\Delta S/R)$$

Thus, a plot of ln(K) vs. 1/T has a slope of $-\Delta H/R$ and an intercept of $\Delta S/R$. In other words, if the van't Hoff plot is linear, we can determine ΔH and ΔS from the slope and intercept, respectively. As a rough "rule of thumb" the heat capacity change for unfolding, $\Delta C_p = C_p(\text{unfolded}) - C_p(\text{folded}) = d(\Delta H)/dT$, has a value near 12 cal/mole/degree/residue for most proteins. If ΔC_p is nonzero, then ΔH and ΔS vary with temperature and the plot will be nonlinear.

(3) This seemingly unusual data is explained by an irreversible aggregration of the protein near 4M urea. Several proteins have been shown to exhibit such an effect for some "critical" denaturant concentration. β–galctosidase is largely unfolded in 4M urea, and many "unfolded" proteins tend to

aggregrate to minimize the exposure of hydrophobic surface area that occurs upon unfolding. The aggregration does not occur upon standing at very high concentrations of urea, perhaps due to changes in the thermodynamics of nonpolar solvation by the large amounts of urea. The aggregation is aggrevated by high protein concentration, since it is by definition a bimolecular or higher order process. The original data given in this problem is discussed by M. E. Goldberg in "Dynamic Aspects of Conformation Changes in Biological Macromolecules", (C. Sadron, editor). Similar data for tryptophanase is reported by London, *et. al.* in *Eur. J. Biochem.*, 47 409–415 (1979).

(4) An approximate solution for this problem can be found by using $\Delta G = -RTln(K)$ to compute the transfer free energy for each molecule at each of the three temperatures. For 3–MI this gives ΔG = –1.67, –1.75 and –1.82 kcal/mole with increasing temperature, while NMI gives –3.27, –3.36 and –3.44 kcal/mole. The value of ΔS is found for each molecule as the negative of the slope of a plot of ΔG vs. T; and ΔH then found via $\Delta G = \Delta H - T\Delta S$. These fits give the values ΔS = 7.5 cal/mole/deg and ΔH = +0.49 kcal/mole for 3–MI. The NMI data gives ΔS = 8.5 cal/mole/deg and ΔH = –0.82 kcal/mole. The relatively large enthalpic difference is presumably due to the formation of a hydrogen bond between the indole –NH of 3–MI and water. Two factors complicate the simple calculations described above: the units for the partition coefficient (choice of mole–fraction, volume–fraction, corrected volume–fraction, etc.), and the presence of water in the nonpolar phase. Both of these issues are discussed at great length in the original paper by Wimley and White, *Biochemistry*, 31, 12813–12818 (1992). Also see their correction in *Biochemistry*, 32, 9262 (1993).

(5) The GB/SA solvation model consists of a combination cavity creation/vdW term modeled as linear to the solvent accessible surface area, and a polarization term derived from the sum of Coulomb's law and the Born equation. In the

model discussed in *JACS*, 112, 6127 (1990), the cavity creation term is assumed to be independent of the type of atom lining the cavity. In addition, the authors note that the model is very sensitive to the set of molecular mechanics atomic partial charges used. This is of particular concern when trying to model solvation energy differences for conformational isomers since most molecular mechanics models do not allow the charges to change with conformation. In addition, there is some concern about the validity of the model around active site clefts and narrow solvent exposed channels, since it contains a discontinuity at the point where two accessible surfaces are just separated. Nonaqueous solvents have been simulated by changing the dielectric constant in the appropriate equations, but to take the surface area term to be independent of solvent is not well justified.