Urea Interactions with Protein Groups: A Volumetric Study

Soyoung Lee, Yuen Lai Shek, and Tigran V. Chalikian*

Department of Pharmaceutical Sciences, Leslie Dan Faculty of Pharmacy, University of Toronto, 144 College Street, Toronto, Ontario M5S 3M2, Canada

Running Title: Urea Binding to Protein Groups

Keywords: urea; protein groups; thermodynamics; volume; compressibility

Tel: (416)946-3715

Fax: (416)978-8511

E-mail: chalikan@phm.utoronto.ca

* Author to whom correspondence should be addressed

Abstract

We determined the partial molar volumes and adiabatic compressibilities of N-acetyl amino acid amides, N-acetyl amino acid methylamides, N-acetyl amino acids, and short oligoglycines as a function of urea concentration. We analyze these data within the framework of a statistical thermodynamic formalism to determine the association constants for the reaction in which urea binds to the glycyl unit and each of the naturally occurring amino acid side chains replacing two waters of hydration. Our determined association constants, k, range from 0.04 to 0.39 M. We derive a general equation that links k with changes in free energy, ΔG_{tr}, accompanying the transfer of functional groups from water to urea. In this equation, ΔG_{tr} is the sum of a change in the free energy of cavity formation, $\Delta\Delta G_C$, and the differential free energy of solute-solvent interactions, $\Delta\Delta G_I$, in urea and water. The observed range of affinity coefficients, k, corresponds to the values of $\Delta\Delta G_I$ ranging from highly favourable to slightly unfavourable. Taken together, our data support a direct interaction model in which urea denatures a protein by concerted action *via* favourable interactions with a wide range of protein groups. Our derived equation linking k to ΔG_{tr} suggests that $\Delta\Delta G_I$ and, hence, the net transfer free energy, ΔG_{tr}, are both strongly influenced by the concentration of a solute employed in the experiment. We emphasize the need to exercise caution when two solutes differing in solubility are compared to determine the ΔG_{tr} contribution of a particular functional group.

INTRODUCTION

There are not many areas in biophysical chemistry which have attracted as much attention as the problem of elucidation of the molecular mechanisms underlying the action of urea and other water-soluble cosolvents on protein stability and function.[1-16] Two major proposals concerning the mode of urea action have been put forward.[15, 17] The direct mechanism implies the existence of direct van der Waals or hydrogen bonding or other electrostatic interactions between urea and protein groups.[8, 9, 18] In the indirect mechanism, urea exerts its effect *via* perturbation of the structure of water and the related modulation of protein-water interactions.[15, 19, 20] The issue of determination of the specific mode of urea action remains controversial, although the direct mechanism appears to be increasingly favoured by the researchers.[17, 21, 22] The quest for understanding the relative importance of urea interactions with polar *versus* nonpolar groups is yet another recurrent theme in current investigations.[6, 8-10, 14, 15, 23-26]

There are two experimental approaches that have been employed in studying the effect of cosolvents on protein folding and binding. In one approach, the free energy of the transfer of a solute from water to a water-cosolvent mixture, ΔG_{tr}, is determined from a change in solubility.[1, 5, 25, 27-31] In an alternative approach, equilibrium dialysis measurements are performed in macromolecular solutions containing increasing concentrations of cosolvent to determine the preferential interaction parameter, $(\partial\mu_3/\partial m_2)_{m_3} = (\partial\mu_2/\partial m_3)_{m_2}$, or the preferential binding parameter $\Gamma_{23} = (\partial m_3/\partial m_2)_{\mu_3} = -[(\partial\mu_3/\partial m_2)_{m_3}/(\partial\mu_3/\partial m_3)_{m_2}]$, where μ and m denote chemical potential and molal concentration, respectively; while subscripts 2 and 3

refer to solute (macromolecule) and cosolvent, respectively.[3, 32, 33] The preferential binding parameter, Γ_{23}, is related to the effective numbers of the principal solvent, n_1, and cosolvent, n_3, molecules in the vicinity of a solute *via* $\Gamma_{23} = n_3 - (m_3/m_1)n_1$.[3, 32, 34, 35] Similarly, the preferential hydration parameter, $\Gamma_{21} = (\partial m_1/\partial m_2)_{\mu_1}$, is related to n_1 and n_3 *via* $\Gamma_{21} = n_1 - (m_1/m_3)n_3$. Although the two techniques are complementary and linked *via* $\Delta G_{tr} = \int_o^{m3} (\partial\mu_2/\partial m_3)_{m_2} dm_3$, their results cannot be compared directly. The transfer free energy method is restricted to low-molecular weight solutes, while the equilibrium dialysis method is applicable to macromolecular solutes. The former can potentially provide insights into the interactions of cosolvents with individual functional groups, while the latter provides collective information about the net accumulation or depletion of cosolvent in the vicinity of a solute.

Schellman has pioneered the use of statistical thermodynamics to rationalize experimental data on solute-cosolvent interactions by treating them as a stoichiometric binding.[36-39] He has proposed a model in which the binding of cosolvent to a solute is presented as an exchange reaction in which cosolvent replaces waters of hydration. The model and the related formalism have resolved a paradox first posed by von Hippel *et al.*[40] In this paradox, negative cosolvent binding constants are required to account for the preferential exclusion of cosolvent from a solute (expressed as negative values of Γ_{23}). In a later development, the free energy of the transfer of a solute from water to a water-cosolvent mixture, ΔG_{tr}, has been presented as the sum of the differential free

energy of cavity formation, $\Delta\Delta G_C$, and the free energy of solute-cosolvent interactions, $\Delta\Delta G_I$.[39, 41]

Despite the wealth of information provided by these and other studies, we still lack the thermodynamic knowledge of the interactions of urea with individual protein groups. Information about the energetics of urea interactions with a specific functional group has been predominantly obtained by comparing ΔG_{tr} of two solutes differing in chemical structure by that group.[30, 42] However, solubility-based studies are conducted, by definition, at the solubility limit of the solute under study which may range from mM to several M. As discussed below, direct comparison between the data for solutes exhibiting significantly different solubilities may be fraught with error.

Decomposition of ΔG_{tr} into the cavity, $\Delta\Delta G_C$, and interaction, $\Delta\Delta G_I$, terms is not a simple matter. A standard way to evaluate the differential free energy of cavity formation, $\Delta\Delta G_C$, is based on scaled particle theory (SPT) calculations.[41, 43-45] However, SPT-based calculations may be unreliable due to the critical sensitivity of calculated $\Delta\Delta G_C$ on the assumed diameters of solvent and cosolvent molecules.[45]

We describe here a novel way of probing solute-cosolvent interactions which is based on high precision volumetric measurements. The method is not restricted with respect to either the concentration or the molecular weight of a solute. We measure the partial molar volume and adiabatic compressibility of oligoglycines and a set of amino acid derivatives with blocked termini as a function of urea concentration. We use these data in conjunction with a

statistical thermodynamic formalism to determine the equilibrium constants, k, for water/urea exchange reactions in the solvation shell of the glycyl unit (-CH_2CONH-) and amino acid side chains.[46] We further calculate the differential free energy of solute-solvent interactions in a concentrated urea solution and water for all naturally occurring amino acid side chains and the glycyl unit.

Our results support the direct interaction model of urea action. More specifically, our data are consistent with the picture in which urea denatures a protein by concerted action *via* favourable solute-cosolvent interactions with a wide range of protein groups, including the peptide backbone and most of the amino acid side chains. This conclusion is in agreement with the results of recent molecular dynamics simulations.[8, 9, 14] Hua *et al.* have found that urea interacts with the backbone of a polypeptide chain and its amino acid side chains by stronger dispersion interactions than water.[8] One manifestation of enhanced dispersion interactions of urea with solutes is the apparent weakening of the hydrophobic effect in concentrated urea solutions.[9] Enhanced dispersion interactions are augmented by preferential hydrogen bonds formed between urea and peptide groups.[8] Our data will find further use in modeling the volumetric properties of unfolded polypeptide chains in concentrated urea solutions which are required for analyzing the post-denaturational baselines in volumetric investigations of urea-induced protein unfolding transitions.

MATERIALS AND METHODS

Materials

Urea, glycine, diglycine, triglycine, tetraglycine, pentaglycine, N-methyl acetamide, N-acetyl glycine amide, N-acetyl tyrosine amide, N-acetyl glycine, N-acetyl alanine, N-acetyl phenylalanine, N-acetyl tryptophan, N-acetyl cysteine, N-acetyl serine, and N-acetyl threonine, as well as sodium acetate were purchased from Sigma-Aldrich Canada, Ltd. (Oakville, Ontario, Canada). N-acetyl alanine amide, N-acetyl valine amide, N-acetyl leucine amide, N-acetyl isoleucine amide, N-acetyl proline amide, N-acetyl phenylalanine amide, N-acetyl tryptophan amide, N-acetyl methionine amide, N-acetyl glutamine amide, N-acetyl aspartic acid amide, N-acetyl glutamic acid amide, N-acetyl lysine amide chloride, N-acetyl arginine amide acetate, N-acetyl glycine methylamide, and N-acetyl histidine methylamide were purchased from Bachem Bioscience, Inc (King of Prussia, PA, USA). N-acetyl asparagine was obtained from Fluka (Buchs, Switzerland). All amino acid derivatives except N-acetyl serine and N-acetyl threonine were in L-stereoisomeric form. N-acetyl serine and N-acetyl threonine were a mixture of D- and L-stereoisomeric forms. All the reagents used in the studies reported here were of the highest purity commercially available and used without further purification.

Solution Preparation

Aqueous solutions of urea with concentrations of 2, 4, 6, and 8 M were prepared by weighing 10 to 50 g of urea and adding pre-estimated amounts of water to achieve the desired molalities, m. The molar concentration, C, of a urea solution

was computed from the molal value, m, using $C = [1/(m\rho_W) + \phi V/1000]^{-1}$, where ρ_W is the density of water and ϕV is the apparent molar volume of urea. The concentrated urea solutions were used as solvents for respective oligoglycines and amino acid derivatives. The concentrations of the samples were determined by weighing 10 to 20 mg of a solute material with a precision of ±0.02 mg and dissolving the sample in a known amount of solvent (urea solution). All chemicals were dried under vacuum in the presence of phosphorus pentoxide for 72 hours prior to weighing.

Methods

All densities were measured at 25 °C with a precision of $\pm 1.5 \times 10^{-4}$ % using a vibrating tube densimeter (DMA-5000, Anton Paar, Gratz, Austria). The apparentmolar volumes, ϕV, of the solutes were calculated from the relationship $\phi V = M/\rho - (\rho - \rho_0)/(\rho\rho_0 m)$, where M is the molecular weight of the solute; m is the molal concentration of the solute; ρ and ρ_0 are the densities of the solution and the solvent (urea solution), respectively.

Solution sound velocities, U, and absorptions per wavelength, $\alpha\lambda$, were measured at 25 °C at a frequency of 7.2 MHz using the resonator method and a previously described differential technique.[47-50] The analysis of the frequency characteristics of the ultrasonic resonator cells required for sound velocity measurements was performed by a Hewlett Packard model E5100A network/spectrum analyzer (Mississauga, ON, Canada). For the type of ultrasonic resonators used in this work, the accuracies of the sound velocity and

absorption measurements are about $\pm 1 \times 10^{-4}$ % and ±1 %, respectively.[48, 51, 52] The acoustic characteristics of a solute which can be derived directly from ultrasonic measurements are the relative molar sound velocity increment, [U], and the molar increment of ultrasonic absorption per wavelength, [$\alpha\lambda$]. The relative molar sound velocity increment, [U], of a solute is equal to $(U - U_0)/(U_0C)$, where C is the molar concentration of a solute; and U and U_0 are the sound velocities in the solution and the solvent, respectively. The molar increment of ultrasonic absorption per wavelength, [$\alpha\lambda$], is equal to $\Delta(\alpha\lambda)/C$, where α is the coefficient of sound absorption; λ is the sound wavelength; $\Delta(\alpha\lambda)$ is the difference in the ultrasonic absorption per wavelength between the solution and the solvent.

The values of [U] were used in conjunction with the ϕV values derived from densimetric measurements to calculate the apparent molar adiabatic compressibility, ϕK_S, using the relationship $\phi K_S = \beta_{S0} (2\phi V - 2[U] - M/\rho_0)$, where $\beta_{S0} = \rho_0^{-1} U_0^{-2}$ is the coefficient of adiabatic compressibility of the solvent. The values of ρ_0, U_0, and β_{S0} were directly determined for each urea solution from our densimetric and acoustic measurements. For each evaluation of ϕV or ϕK_S, three to five independent measurements were carried out within a concentration range of 2 - 3 mg/ml. Our reported values of ϕV or ϕK_S represent the averages of these measurements, while the errors were calculated as standard deviations.

Measurements in the Solutions of N-acetyl Amino Acids

To minimize the influence of the ionizable carboxyl terminus of the N-acetyl amino acids (N-acetyl glycine, N-acetyl alanine, N-acetyl phenylalanine, N-acetyl tryptophan, N-acetyl cysteine, N-acetyl serine, N-acetyl threonine, and N-acetyl asparagine) on the side chains, their partial molar volume and adiabatic compressibilities were determined in each urea solution at low ~pH 2 where the carboxyl terminus can be considered to be fully neutralized. The initial values of pH of the N-acetyl amino acid solutions were within the range of 2.3 to 2.8. The pH of the N-acetyl amino acid solutions were lowered by HCl; equal aliquots of HCl were incrementally added to both the solution and the solvent. Not to alter the initial urea concentration of the amino acid samples in the course of titration, the HCl solution used for the titrations was adjusted to the same urea concentration as that in the amino acid sample being titrated. The relative molar sound velocity increment, [U], and apparent molar volume, ϕV, of the solute were determined from the differential solution-*versus*-solvent measurements at each pH point. To ensure full neutralization of the carboxyl terminus, the pH-dependent measurements of [U] and ϕV were performed until these volumetric parameters level off (at ~pH 2). The plateau values of [U] and ϕV were used to calculate apparent molar adiabatic compressibilities, ϕK_S, at each experimental temperature.

Volume and Compressibility Changes Accompanying Neutralization of Ionizable Side Chains

To determine changes in relative molar sound velocity increment, volume, and compressibility accompanying neutralization of the aspartic acid, glutamic acid, histidine, lysine, and arginine side chains, we performed pH-dependent densimetric and acoustic measurements at each urea concentration studied in this work. The pH-dependent density and ultrasonic velocity and absorption measurements were performed following the previously described experimental protocol.[53]

RESULTS

Previous studies have revealed that the apparent molar volumes and adiabatic compressibilities of oligopeptides and N-acetyl amino acid amides in water do not strongly depend on concentration.[54-57] By extension, we assume that the concentration dependences of the volumetric properties of these solutes should be insignificant in concentrated solutions of urea. Consequently, we do not discriminate below between the apparent molar and partial molar characteristics of the amino acid derivatives. Tables S1, S2, and S3 of Supplementary Information present the relative molar sound velocity increments, [U], partial molar volumes, $V°$, and partial molar adiabatic compressibility, $K°_S$, of the solutes investigated in this study at 0, 2, 4, 6, and 8 M urea. To the best of our knowledge, no data of this kind have been reported. Therefore, our results cannot be compared with the literature.

Ionization/neutralization reactions of aspartic and glutamic acids (reaction 1), histidine (reaction 2), and lysine and arginine (reaction 3) are described by the equilibria:

$-COO^- + H^+ \leftrightarrow -COOH$ (Reaction 1)

$\equiv N + H^+ \leftrightarrow \equiv NH^+$ (Reaction 2)

$-NH_3^+ + OH^- \leftrightarrow -NH_2 + H_2O$ (Reaction 3)

We measured the relative molar sound velocity increments, [U], molar increments of ultrasonic absorption per wavelength, [$\alpha\lambda$], and partial molar volumes, V°, of the amino acid derivatives with ionizable side chains as a function of pH in the acidic and alkaline range (data not shown). Our measured pH-dependences of [U], [$\alpha\lambda$], and V° exhibit profiles typical of such dependences in water.[58, 59] We analyzed these pH-dependences as described previously to determine changes in volume and compressibility accompanying protonation of each ionizable group we study in this work.[53]

Tables S4a, b, c, d, and e of Supplementary Information present the values of pK_a, ΔV, and ΔK_S for protonation of the aspartic acid, glutamic acid, histidine, lysine, and arginine side chains at 0, 2, 4, 6, and 8 M urea, respectively. For lysine and arginine, changes in volume or compressibility accompanying the protonation of their side chains ($-NH_2 + H^+ \leftrightarrow -NH_3^+$) were calculated by subtracting the volume or compressibility of water ionization from the changes in volume or compressibility associated with Reaction 3. The volumetric

characteristics of water ionization have been determined at each urea concentration from ΔX_{ion} = X°(HCl)+ X°(NaOH) -X°(H_2O) - X°(NaCl), where X°(H_2O), X°(NaOH), X°(HCl), and X°(NaCl) are our measured partial molar volumes or adiabatic compressibilities of water, NaOH, HCl, and NaCl, respectively (data not shown). Our determined changes in volume accompanying protonation of water, ΔV_{ion}, are -21.7±0.2, -21.2±0.7, -20.5±0.8, -20.0±0.6, and -18.6±0.7 cm^3mol^{-1} at 0, 2,4, 6, and 8 M urea, respectively, while the changes in adiabatic compressibility are $(-49.1\pm0.2)\times10^{-4}$, $(-47.1\pm0.7)\times10^{-4}$, $(-45.1\pm0.8)\times10^{-4}$, $(-41.9\pm0.7)\times10^{-4}$, and $(-38.1\pm1.5)\times10^{-4}$ $cm^3mol^{-1}bar^{-1}$ at 0, 2, 4, 6, and 8 M urea, respectively.

Alkaline titration measurements in arginine solutions were conducted up to ~pH 13. Densimetric and acoustic measurements at higher pH values are problematic and may result in large error because of the high concentrations of the added NaOH and its respective contributions to the solution density and sound velocity. The protonation-related changes in volume, ΔV, and adiabatic compressibility, ΔK_S, for arginine with pK_a of ~12.5 have been determined based on incomplete titration plots by fitting the available experimental points with well-known equations as described previously.[53] At high urea concentrations, the incomplete nature of the measured pH-dependences of the volumetric characteristics of arginine coupled with the noisiness of the data did not permit us to reliably determine the values of pK_a. Therefore, we assume, as a first approximation, that the pK_a of arginine changes with an increase in urea concentration in parallel to that of lysine.

DISCUSSION

Volumetric Contributions of Amino Acid Side Chains

The volume or compressibility contribution of a specific amino acid side chain can be obtained as the difference in the partial molar volume, V°, or adiabatic compressibility, $K°_S$, between the corresponding amino acid and glycine derivatives studied in this work. Tables 1 and 2 list, respectively, the volume and adiabatic compressibility contributions for the 19 amino acid side chains as a function of urea concentration. For amino acids containing titrable groups (aspartic and glutamic acids, histidine, lysine, and arginine), the data presented in Tables 1 and 2 refer to the unionized state of the side chain. For aspartic and glutamic acids, the group contributions of uncharged species, X(-R), were calculated from $X(-R) = X°(pH) - X°(Gly) + \Delta X/(1 + 10^{pKa - pH})$, where X°(pH) is the partial molar volume or adiabatic compressibility of N-acetyl aspartic or glutamic acid amide at the experimental pH; X°(Gly) is the partial molar volume or adiabatic compressibility of N-acetyl glycine amide; ΔX and pK_a are, respectively, the protonation volume or adiabatic compressibility and the dissociation constant of the side chain (see Tables S4a-e of Supplementary Information). The group contribution of the histidine side chain was calculated from $X(-R) = X°(pH) - X°(Gly) - \Delta X/(1 + 10^{pH - pKa})$, where X°(pH) is the partial molar volume or adiabatic compressibility of N-acetyl histidine methylamide at the experimental pH; and X°(Gly) is the partial molar volume or adiabatic compressibility of N-acetyl glycine methylamide. The group contribution of lysine side chain was calculated from $X(-R) = X°(pH) - X°(Gly) - X°(HCl) - \Delta X/(1 + 10^{pH - pKa})$, where X°(pH) is the partial

molar volume or adiabatic compressibility of N-acetyl lysine amide hydrochloride at the experimental pH; and X°(HCl) refers to our measured urea-dependent values of the partial molar volume or adiabatic compressibility of HCl (data not shown). The group contribution of arginine side chain was calculated from X(-R) = X°(pH) - X°(Gly) - X°(HCl) + X°(NaCl) - X°(CH_3COONa) - $\Delta X/(1 + 10^{pH - pKa})$, where X°(pH) is the partial molar volume or adiabatic compressibility of N-acetyl arginine amide acetate at the experimental pH; X°(NaCl) and X°(CH_3COONa) are our measured partial molar volumes or adiabatic compressibilities of sodium chloride and sodium acetate, respectively (data not shown).

Volumetric Contributions of Glycyl Residue

The contribution of the glycyl residue (-CH_2CONH-) can be obtained as the difference between the values corresponding to N-acetyl glycine methylamide (CH_3-CO-NH-CH_2-CO-NH-CH_3) and N-methyl acetamide (CH_3-NH-CO-CH_3) or as the incremental change in the dependence of the volumetric properties of triglycine, tetraglycine, and pentaglycine on the number of glycyl units in the molecule.[60] The volume and compressibility contributions of the glycyl residue determined in both ways are listed in Tables 1 and 2, respectively.

Inspection of Tables 1 and 2 reveals that the volume and compressibility contributions of the glycyl residue in N-acetyl glycine methylamide are significantly smaller than the respective contributions of the glycyl residue in oligoglycines. These disparities parallel the results of enthalpy measurements which suggest that hydration of a peptide group is strongly influenced by its

microenvironment.[61-63] As the urea concentration increases the volumetric differences become smaller, practically disappearing in 8 M urea. This observation suggests that, while the hydration properties of the glycyl residue in the two solutes are significantly distinct in water, in concentrated urea solutions the solvation differences subside.

Urea Affinity for Various Functional Groups

We analyze our measured urea-dependences of the volumetric properties of solutes under the following assumptions. Each bound urea replaces *r* water molecules from the binding site. All binding sites are identical and independent. There are *n* binding sites for the principal solvent (water) and, hence, *n/r* binding sites for cosolvent (urea) in each analyzed solute. The elementary solvation reactions involving a cosolvent-binding site can be presented as follows:

$S_0 + rW \leftrightarrows SW_r$ (Reaction 4)

$S_0 + U \leftrightarrows SU$ (Reaction 5)

where S_0 denotes the dry (unsolvated) binding site.

Based on the combinatorial approach, the total concentration of a solute with *n/r* identical and independent cosolvent-binding sites in water, $[S_1]$, and a concentrated urea solution, $[S_3]$, are given by $[S_1] = [S_{01}](1 + k_1 a_{10}^{\,r})^{n/r}$ and $[S_3] = [S_{03}](1 + k_1 a_1^{\,r} + k_3 a_3)^{n/r}$, respectively; where $[S_{01}]$ and $[S_{03}]$ are the concentrations of unsolvated solute in water and urea solution, respectively; a_{10} and a_1 are the

activities of water in the absence and presence of urea, respectively; a_3 is the activity of urea; and k_1 and k_3 are the elementary binding constants for Reactions 4 and 5, respectively. The Gibbs free energies of solvation of a solute site in water, ΔG^*_1, and a concentrated urea solution, ΔG^*_3, can be expressed as follows:[64-66]

$$\Delta G^*_1 = -RT\ln([S_1]/[S_g]) = \Delta G_{C1} - (n/r)RT\ln(1 + k_1 a_{10}{}^r) \quad (1)$$

$$\Delta G^*_3 = -RT\ln([S_3]/[S_g]) = \Delta G_{C3} - (n/r)RT\ln(1 + k_1 a_1{}^r + k_3 a_3) \quad (2)$$

where $[S_g]$ is the molar concentrations of a solute in the ideal gas phase that exists in equilibrium with the solute in the liquid phase; and $\Delta G_{C1} = -RT\ln([S_{01}]/[S_g])$ and $\Delta G_{C3} = -RT\ln([S_{03}]/[S_g])$ are the free energies of formation of a cavity that accommodates a solute in water and urea solution, respectively.

The partial molar volume of a solute in water is described by the expression:[46,67]

$$V^\circ_1 = (\partial \Delta G^*_1/\partial P)_T + RT(\partial \ln[S_1]/\partial P)_T = V_{C1} + (n/r)k_1 a_{10}{}^r \Delta V_{10}/(1 + k_1 a_{10}{}^r) + \beta_{T1} RT \quad (3)$$

where $V_{C1} = (\partial \Delta G_{C1}/\partial P)_T$ is the cavity volume in water; $\Delta V_{10} = -RT(\partial \ln k_1/\partial P)_T$ is the change in volume accompanying the binding of *r* water molecules to the dry

binding site in pure water; and β_{T1} is the coefficient of isothermal compressibility of water.

The ratio ($1 / k_1a_{10}^r$) represents the fraction of non-interacting (unsolvated) solute species which is negligibly small. Therefore, to a good approximation $1 << k_1a_{10}^r$, with Eq. (3) simplifying to the form:

$$V°_1 = V_{C1} + (n/r)\Delta V_{10} + \beta_{T1}RT \quad (4)$$

Similarly, the partial molar volume of a solute in a concentrated urea solution is described by the expression:

$$V°_3 = V_{C3} + (n/r)(k_1a_1^r\Delta V_1 + k_3a_3\Delta V_3)/(1 + k_1a_1^r + k_3a_3) + \beta_{T3}RT \quad (5)$$

where $V_{C3} = (\partial\Delta G_{C3}/\partial P)_T$ is the cavity volume in a concentrated urea solution; ΔV_1 is the change in volume accompanying the binding of *r* water molecules to the dry binding site in a concentrated urea solution; $\Delta V_3 = -RT(\partial \ln k_3/\partial P)_T$ is the change in volume accompanying the binding of a urea molecule to the dry binding site; and β_{T3} is the coefficient of isothermal compressibility of the urea solution.

With $1 << (k_1a_1^r + k_3a_3)$, Eq. (5) simplifies to the relationship:

$$V°_3 = V_{C3} + (n/r)(k_1a_1^r\Delta V_1 + k_3a_3\Delta V_3)/(k_1a_1^r + k_3a_3) + \beta_{T3}RT = V_{C3} + (n/r)[\Delta V_1 + k(a_3/a_1^r)\Delta V_3]/[1 + k(a_3/a_1^r)] + \beta_{T3}RT \quad (6)$$

where $k = k_3/k_1$ is the equilibrium constant for the reaction in which a urea molecule replaces *r* water molecules at the binding site.

A change in the partial molar volume of a solute arising from the presence of urea can be found by subtracting Eq. (4) from Eq. (6):

$$\Delta V^\circ = \Delta V_C + (n/r)(\Delta V_1 - \Delta V_{10})/[1 + k(a_3/a_1^r)] + (n/r)k(a_3/a_1^r)(\Delta V_1 - \Delta V_{10})/[1 + k(a_3/a_1^r)] + (\beta_{T3} - \beta_{T1})RT \quad (7)$$

where $\Delta V_C = (V_{C3} - V_{C1})$ is the differential cavity volume in a concentrated urea solution and water.

By employing SPT-based calculations, we have shown that, for a solute with the diameter between 4 and 10 Å, the cavity volume is essentially the same between 0 and 8 M urea.[46] Consequently, we neglect the $\Delta\Delta V_C$ term in Eq. (7). One can also neglect the differential term $(\beta_{T3} - \beta_{T1})RT$. Firstly, $\beta_T RT$ is only ~1 cm^3mol^{-1} and, secondly, the difference between β_{TU} and β_{TW} is small not exceeding ~25 % even at 8 M urea solution.

As previously discussed, $\Delta V_1 = \Delta V_{10} - \gamma_1 r \Delta V^\circ_1$ and $\Delta V_3 = \Delta V_{30} - \gamma_3 \Delta V^\circ_3$, where ΔV°_1 and ΔV°_3 are the excess partial molar volumes of water and cosolvent in a concentrated solution; ΔV_{30} is the change in volume accompanying the binding of urea to the binding site in an ideal solution; and γ_1 and γ_3 are the correction factors reflecting the influence of the bulk solvent on the properties of solvating water and cosolvent, respectively.[46] The values of γ_1 and γ_3 may change from 0

(the properties of the solvation shell change in parallel with those of the bulk) to 1 (the properties of the solvation shell are independent of those of the bulk).

One finally arrives at the following expression:

$$\Delta V^\circ = - \gamma_1 n \Delta V^\circ_1 + \Delta V (n/r)(a_3/a_1^r)k / [1 + (a_1/a_3^r)k] \quad (8)$$

where $\Delta V = \Delta V_3 - \Delta V_1 = \Delta V_0 + \gamma_1 r \Delta V^\circ_1 - \gamma_3 \Delta V^\circ_3$ is the change in volume associated with replacement of water with urea in the binding site in a concentrated urea solution; and $\Delta V_0 = \Delta V_{30} - \Delta V_{10}$ is the exchange volume in an ideal solution.

Differentiating Eq. (8) with respect to pressure, one obtains a relationship for a urea-dependent change in the partial molar isothermal compressibility of a solute:

$$\Delta K^\circ_T = -\gamma_1 n \Delta K^\circ_{T1} + \gamma_1 \Delta V^\circ_1 (\partial n/\partial P)_T + \Delta K_T (n/r)(a_3/a_1^r)k / [1 + (a_3/a_1^r)k] + \Delta V^2 (n/r)(a_3/a_1^r)k / RT[1 + (a_3/a_1^r)k]^2 \quad (9)$$

where $\Delta K_T = \Delta K_{T3} - \Delta K_{T1} = \Delta K_{T0} + \gamma_1 r \Delta K^\circ_{T1} - \gamma_3 \Delta K^\circ_{T3}$ is the change in compressibility associated with the replacement of water with urea at the binding site in a concentrated urea solution; ΔK°_{T1} and ΔK°_{T3} are the excess partial molar isothermal compressibilities of water and cosolvent in a concentrated solution; $\Delta K_{T0} = \Delta K_{T30} - \Delta K_{T10}$ is the change in compressibility associated with the replacement of water with urea at the binding site in an ideal solution; and ΔK_{T10}

and ΔK_{T30} are, respectively, the changes in compressibility accompanying the binding of water and urea to the dry binding site in an ideal solution.

Note that Eqs. (8) and (9) are identical to analogous expressions we have previously reported following a conceptually different derivation pathway.[46] Given $(\partial n/\partial P)_T \approx 0$, Eq. (9) reduces to the relationship:

$$\Delta K°_T = -\gamma_1 n\Delta K°_{T1} + \Delta K_T(n/r)(a_3/a_1{}^r)k / [1 + (a_3/a_1{}^r)k] + \Delta V^2(n/r)(a_3/a_1{}^r)k / RT[1 + (a_3/a_1{}^r)k]^2 \quad (10)$$

When analysing below our experimental partial molar adiabatic compressibility data, $K°_S$, we use Eq. (10), although it was derived for isothermal compressibility, $K°_T$. Note that $K°_T$ relates to $K°_S$ *via* $K°_T = K°_S + (T\alpha_0{}^2/\rho_0 c_{P0})(2E°/\alpha_0 - C°_P/\rho_0 c_{P0})$, where ρ_0 is the density of the solvent; α_0 is the coefficient of thermal expansion of the solvent; c_{P0} is the specific heat capacity at constant pressure of the solvent; E° is the partial molar expansibility of a solute; and $C°_P$ is the partial molar heat capacity of a solute.[68] Due to a small value of α_0 and a large value of c_{P0} of water-based solvents, the difference between $K°_T$ and $K°_S$ in aqueous solutions is not large. Therefore, we employ Eq. (10) obtained for partial molar isothermal compressibility to treat our partial molar adiabatic compressibility data.

Data Analysis

Eqs. (8) and (10) have been derived under the assumption of a solute with identical and non-interacting binding sites. This assumption restricts the use of the model to chemically homogeneous functional entities to ensure homogeneity of the determined values of k, ΔV_0, and ΔK_{T0}. When applied to solutes with heterogeneous binding sites, Eqs. (8) and (10) will produce the apparent values of k, ΔV_0, and ΔK_{T0}. The feasibility of application of a "homogeneous" solvent exchange model to treating solutes with heterogeneous binding sites with subsequent determination of apparent thermodynamic parameters has been discussed and analyzed in detail by Schellman.[69]

We used Eq. (8) to treat the volume data shown in Table 1, while Eq. (10) was used for analyzing the compressibility data presented in Table 2. The activity of water was taken equal to its mole fraction $a_1 = 1 - 0.0173\ [\text{urea}] - 7.1\times10^{-4}\ [\text{urea}]^2$.[46] This is an appropriate assumption since the activity coefficient of water remains close to 1 at urea concentrations of up to 8 M.[70] The activity of urea a_3 was taken equal to the product of its molar concentration and the activity coefficient, γ, which is approximated by the polynomial $\gamma = 0.99877 - 0.0878\ [\text{urea}] + 0.00868\ [\text{urea}]^2 - 6.74\times10^{-4}\ [\text{urea}]^3 + 2.29\times10^{-5}\ [\text{urea}]^4$. This polynomial was calculated from the reported values of the coefficient of activity of urea as a function of its concentration.[71]

The excess volumetric parameters of water and urea required for calculations with Eqs. (8) and (10) have been measured in our recent work.[46] Our choice of the values of γ_1 and γ_3 in Eqs. (8) and (10) for a specific solute stems from the following considerations. Owing to their small size, waters of hydration, if

strongly influenced by a solute, can be considered to be relatively insensitive to the properties of water the bulk. At low to moderate temperatures, waters solvating nonpolar groups, in an attempt to form hydrogen bonds with each other within a restricted configurational space, become highly oriented. It is, therefore, reasonable to assume that the structural and thermodynamic properties of such waters are relatively insensitive to changes in the properties of bulk water and, hence, they should exhibit a γ_1 close to 1. On a similar note, we have proposed that, for charged groups interacting with their solvating waters *via* strong charge-dipole interactions, γ_1 can be approximated by 1.[46] In contrast, waters hydrating polar (but uncharged) groups form continuous networks of hydrogen bonds extending from solute to water in the bulk and, therefore, should be significantly influenced by the latter. Consequently, we assume that, for polar groups, $\gamma_1 \approx 0$. For a solute with a mixture of polar and nonpolar atomic groups, γ_1 were calculated as the nonpolar fraction of the solvent accessible surface area [taken from ref.[72]].

Urea is bulkier than water and can potentially form up to eight hydrogen bonds with its neighbours. Consequently, despite its being engaged in solute-solvent interactions, urea can still develop numerous interactions with solvent in the bulk. Therefore, it is reasonable to assume that the thermodynamic properties of solvating urea molecules should be, in general, influenced to a significant degree by the bulk solvent. We use in our analysis an approximation of $\gamma_3 \approx 0$.

Figure 1 shows representative urea dependences of the volume (panel A) and compressibility (panel B) contributions of the leucine side chain in N-acetyl leucine amide which are fitted by Eqs. (8) and (10), respectively. The number of binding sites for water, n, in Eqs. (8) and (10) was calculated for each functionality as the ratio of its solvent accessible surface area to 9 $Å^2$, the effective cross-section of a water molecule. The solvent-accessible surface areas of the amino acid side chains and the glycyl residue have been taken from ref.[73]. The number of water molecules replaced by urea, r, is taken equal to 2 as the ratio of cross-sectional area of urea to that of a water molecule.[46] Table 3 presents our calculated binding constants, k, and changes in volume, ΔV_0, and compressibility, ΔK_{T0}, for an elementary reaction in which urea replaces two water molecules in the vicinity of the glycyl residue and the 19 naturally occurring amino acid side chain. It should be noted that the binding constant, k, does not appear to be strongly dependent on the value of r. If we set r in our analysis to 3 instead of 2, our calculated values of k change by less than 10 %. On the other hand, the values of ΔV_0 and ΔK_{T0} increase by ~50 %, proportionally to 3/2, the ratio of the assumed numbers of water molecules replaced by urea.

Table 3 also shows the binding parameters for the zwitterionic amino acid glycine. Note that the urea-binding parameters for glycine are somewhat different compared to the same parameters that have been previously evaluated without taking into account the coefficient of activity of urea.[46] As a general observation, the compressibility-based determined binding constants, k, are

characterized by lower error than the volume-based estimates due to larger relative changes in ΔK_T relative to ΔV.

Side Chains and Glycyl Unit

Inspection of data presented in Table 3 reveals that the binding constants, k, for urea association with amino acid side chains range from 0.04 to 0.39 M with the average of 0.16±0.09 M. There is no apparent correlation between the values of k and the ratio of polar to nonpolar solvent accessible surface areas. The affinity of a specific side chain for urea appears to be governed by a fine balance of structural and chemical determinants rather than by the trivial ratio of polar-to-nonpolar solvent accessible surface areas.

Changes in volume, ΔV_0, accompanying replacement of two waters of hydration with a urea molecule around amino acid side chains range from -0.45 to 1.02 cm^3mol^{-1} with the average of 0.18±0.30 cm^3mol^{-1}. Compressibility changes, ΔK_{S0}, vary from -2.26 to 3.55 cm^3mol^{-1} with the average of 1.76±1.27 cm^3mol^{-1}. The values of ΔV_0 and ΔK_{S0}, respectively, reflect changes in volume and compressibility accompanying dehydration of the binding site and urea molecule and formation of the solute-urea complex. As such, they represent a fine quantitative measure of changes in hydration accompanying solvent exchange in the vicinity of a solute. Our correlation analysis revealed that ΔV_0 shows a tendency to slightly increase with an increase in the polar fraction of the solvent accessible surface area (data not shown). On the other hand, ΔK_{S0} does not exhibit any significant correlation with the polar-to-nonpolar surface ratio of a

solute (data not shown). Further studies are needed to investigate the molecular origins of the differential sensitivity of the volume and compressibility observables to the ratio of polar-to-nonpolar atomic groups of a solute.

Table 3 presents the urea-binding parameters for the glycyl unit derived in two ways - from the results on oligoglycines and as the difference between the parameters of N-acetyl glycyl methylamide and N-methyl acetamide. The two sets of data differ significantly. The affinity for urea of the glycyl unit in oligoglycines appears to be weaker than that of the glycyl unit in N-acetyl glycine methylamide (0.08 *versus* 0.23 M). In addition, the volume change, ΔV_0, accompanying urea-water exchange is negative for oligoglycines (-0.83 cm^3mol^{-1}) while being positive for N-acetyl glycine methylamide (0.43 cm^3mol^{-1}). Although of the same sign, the value of ΔK_{S0} for the glycyl unit in N-acetyl glycine methylamide (4.68×10^{-4} $cm^3mol^{-1}bar^{-1}$) is higher than that for the glycyl unit in oligoglycines (3.16×10^{-4} $cm^3mol^{-1}bar^{-1}$). These disparities reflect the differential interaction of the glycyl unit in N-acetyl glycine methylamide and oligoglycines with water rather than with urea. This conclusion is based on the observation that the volumetric properties of the glycyl unit in N-acetyl glycine methylamide and oligoglycines, which are significantly different in water, nearly converge at elevated concentrations of urea (see Tables 1 and 2).

Linking Binding Thermodynamics with Transfer Thermodynamics

Interactions of a solute with cosolvent in the presence of water as the principal solvent have been traditionally described in terms of a change in standard

chemical potential accompanying the transfer of a solute from water to a water-cosolvent mixture, ΔG_{tr}.[30, 42] Below, we derive relationships linking our derived solute-cosolvent binding constants, k, with transfer free energies, ΔG_{tr}.

The chemical potential of a solute in the liquid phase in equilibrium with a solute in the vapor phase is described by the equation:[64-66]

$$\mu_l = \mu^\circ_g + \Delta G^* + RT \ln[S_l] = \mu^\circ_l + RT \ln[S_l] \quad (11)$$

where μ°_g is the standard chemical potential of a solute in the ideal gas phase; $\mu^\circ_l = \mu^\circ_g + \Delta G^*$ is the standard chemical potential of a solute in the liquid phase; $[S_l]$ is the molar concentrations of a solute in the liquid phase; and ΔG^* is the Gibbs free energy of solvation.

A change in free energy accompanying transfer of a solute from the principal solvent (water) to a solvent-cosolvent (water-urea) mixture is given by $\Delta G_{tr} = \mu^\circ_3 - \mu^\circ_1 = \Delta G^*_3 - \Delta G^*_1$. By combining Eq. (11) with Eqs. (1) and (2), one derives the following expression:

$$\Delta G_{tr} = \Delta\Delta G^* = (\Delta G_{C3} - \Delta G_{C1}) - (n/r)RT \ln[(1 + k_1 a_1{}^r + k_3 a_3) / (1 + k_1 a_{10}{}^r)] = \Delta\Delta G_C + \Delta\Delta G_I \quad (12)$$

where $\Delta\Delta G_C = (\Delta G_{C3} - \Delta G_{C1})$ is the differential free energy of cavity formation in a concentrated urea solution and water; and $\Delta\Delta G_I = -(n/r)RT \ln[(a_1/a_{10})^r + k(a_3/a_{10}{}^r)]$ is the interaction contribution to the transfer free energy. In fact, $\Delta\Delta G_I$ represents

the differential free energy of solute-solvent interactions in a concentrated urea solution and water.

In the absence of urea, the activity of water can be approximated by unity ($a_{10} \approx 1$). Note that, with this approximation and the assumption of $\Delta\Delta G_C = 0$, Eq. (12) simplifies for a one-to-one binding stoichiometry (r = 1) to the relationship derived by Schellman $\Delta G_{tr} = -nRT \ln(a_1 + ka_3)$.[37, 38, 74] However, in contrast to differential cavity volume $\Delta\Delta V_C$, differential free energy of cavity formation, $\Delta\Delta G_C$, is not insignificant and cannot be ignored.[45] It is the general consensus that the cavity formation term, $\Delta\Delta G_C$, contributes unfavourably to the water-to-urea transfer free energy, ΔG_{tr}, while the contribution of the interaction term, $\Delta\Delta G_I$, is favourable.[39, 45] The differential free energy of cavity formation, $\Delta\Delta G_C$, can be calculated based on scaled particle theory (SPT).[43-45] However, such calculations may be quite unreliable due to their critical dependence on the assumed hard sphere diameter of the cosolvent molecule.[45] The latter is not easy to determine given the necessity to approximate a non-spherical molecule by a sphere.[45]

The interaction contribution, $\Delta\Delta G_I = -(n/r)RT \ln[(a_1/a_{10})^r + k(a_3/a_{10}^r)]$, on the other hand, can be readily calculated for a solute or a functional group from its equilibrium constant, k. Figure 2 presents the simulated urea-dependences of $\Delta\Delta G_I$ for the 19 naturally occurring amino acid side chains and the glycyl unit. Inspection of Figure 2 reveals that our determined binding constants for the amino acid side chains correspond to changes in the interaction free energies ranging from highly favourable to slightly unfavourable. To stress this point quantitatively, the histogram in Figure 3 shows changes in the interaction free

energy, $\Delta\Delta G_I$, accompanying the transfer of the 19 amino acid side chains and a glycyl unit from water to 2 M urea. Inspection of Figure 3 reveals that, with the exception of serine and aspartic acid, the transfer of all the amino acid side chains and the glycyl unit from water to 2 M urea is accompanied by favourable changes in interaction free energy. These results support the direct mechanism of urea action. More specifically, our data are consistent with the picture in which urea denatures a protein by concerted action *via* favourable solute-cosolvent interactions with a wide range of protein groups, including the peptide backbone and most amino acid side chains.

Effect of Solute Concentration on Transfer Free Energy

In principle, by comparing the experimentally determined data on ΔG_{tr} with the volumetrically determined values of $\Delta\Delta G_I$, one can evaluate the contribution of the differential free energy of cavity formation, $\Delta\Delta G_C$. This is a promising development that may allow one to test the results of SPT-based calculations of $\Delta\Delta G_C$ and fine-tune the adjustable parameters (including the diameter of cosolvent molecules) used in the calculations. However, direct comparison of the data shown in Figure 3 with transfer free energy, ΔG_{tr}, data determined from solubility measurements may not be justified. [7, 30, 42] The values of k and $\Delta\Delta G_I$ reported in this study were all determined at solute concentrations of ~0.01 M. On the other hand, determination of transfer free energies from the ratio of solubilities is performed, by definition, at the limit of solute solubility that may be quite high. Note that, in Eq. (12), a_3 is related to the concentration of unbound

urea in solution. If the solubility of a solute is large (e.g., for glycine, it is on the order of 3 M), the activity of cosolvent interacting with the solute will be greatly reduced depending on the binding constant, k. For example, the concentration of free urea with a total concentration of 1 M in the presence of 3 M glycine is on the order of ~0.35 M. This estimate can be made based on the independent binding site model (see below) with the solvent exchange constant, k, of 0.08 M under the assumption of ~7 binding sites per glycine molecule. In this case, using Eq. (12), one calculates an unfavourable (positive) value of $\Delta\Delta G_I$ of 8 cal mol^{-1} at 1 M urea [in qualitative agreement with experimental data of ΔG_{tr}[30, 42]]. On the other hand, for the millimolar range of glycine concentrations used in this study, the concentration of free urea remains close to its initial value of 1 M with $\Delta\Delta G_I$ being favourable and equal to -21 cal mol^{-1}.

These considerations can be put on a more quantitative footing by incorporating into Eq. (12) the independent site binding model with one-to-one urea-to-binding site stoichiometry. The concentration of free urea can be calculated as the difference between its total concentration and the concentration of occupied binding sites of the solute. Hence, the activity of urea in Eq. (12) is given by:

$$a_3 = \gamma([\text{urea}] - \alpha(n/r)[S]) \qquad (13)$$

where [S] is the total concentration of a solute; and α is the fraction of occupied binding sites. The value of α can be calculated from the one-to-one

stoichiometric binding model from the relationship, $\alpha = (2k[S])^{-1} + 0.5(1 + [urea]/[S]) + [(4k^2[S]^2)^{-1} + (1 + [urea]/[S])/(2k[S]) + 0.25([urea]/[S] - 1)^2]^{0.5}$.

Substituting Eq. (13) into Eq. (12), one obtains the equation for the interaction free energy for a solute at an arbitrary concentration:

$$\Delta\Delta G_I = - (n/r)RT\ln[(a_1/a_{10})^r + k(\gamma([urea] - \alpha(n/r)[S])) / a_{10}^r] \quad (14)$$

Figure 4 shows our simulated urea-dependences of the differential interaction free energy, $\Delta\Delta G_I$, for a solute with five binding sites with k = 0.15 M for concentrations of 0.1 M, 1 M, and 3 M. Inspection of Figure 4 reveals that $\Delta\Delta G_I$ strongly depends on the solute concentration decreasing in magnitude and even changing the sign from negative to positive upon an increase in concentration from 0.1 to 3 M.

It is instructive to scrutinize the urea-dependent solubility data on amino acids reported in literature against the concepts outlined by Eq. (14) and the related simulations presented in Figure 4.[30, 42] The scrutiny reveals an intriguing regularity. Well-soluble amino acids (with aqueous solubility greater than ~6 g/100 g) all exhibit a decrease in solubility with an increase in urea concentration, while the amino acids with lower solubility all exhibit an increase in solubility.[30, 42] The observed trend is consistent with the simulations presented in Figure 4 and raises the possibility that the solubility-based values of the transfer free energies of amino acids are strongly influenced by the effect of high solute concentration and the related decrease in the concentration of free urea. This consideration

should be taken into account when two closely-related solutes differing in solubility are compared to determine the ΔG_{tr} contribution of a functional group (e.g., comparing an amino acid with glycine to determine the contribution of the side chain).

An alternative approach is to evaluate the transfer free energy, ΔG_{tr}, for individual protein groups from $\Delta\Delta G_I$ data coupled with SPT calculations of the differential free energy of cavity formation, $\Delta\Delta G_C$. Data on ΔG_{tr} evaluated in this way can be subsequently used in conjunction with additive schemes to predict protein stability and respective m-values for urea-induced protein denaturation.[5, 6, 25] However, as mentioned above, SPT calculations of $\Delta\Delta G_C$ critically depend on the assumed diameters of solvent and cosolvent molecules which are difficult to estimate. This limits the applicability of SPT-based analysis for quantitative determination of ΔG_{tr} and, further, for protein stability prediction. Alternative ways of evaluating $\Delta\Delta G_C$, perhaps, using molecular dynamics simulations, may allow one to circumvent the problem.

As a final note, recent all-atom Replica exchange MD simulations have revealed a strong pressure-dependence of m-values.[75] In this respect, it should be pointed out that our evaluated changes in volume, ΔV_0, and compressibility, ΔK_{S0}, associated with the water-urea exchange at the binding sites of various protein groups (see Table 3) collectively represent the molecular basis for the pressure-dependence of protein m-values. Further studies along these lines are required to quantify the interplay between the protein stability (m-values) and its pressure dependence, the thermodynamic and volumetric characteristics of

specific protein-urea interactions, and the differential solvation properties of individual protein groups in water and water-urea mixtures.

CONCLUSIONS

We measured the partial molar volumes and adiabatic compressibilities of N-acetyl amino acid amides, N-acetyl amino acid methylamides, N-acetyl amino acids, and oligoglycines at urea concentrations ranging from 0 to 8 M. We used the resulting data to evaluate the volumetric contributions of the 19 naturally occurring amino acid side chains and the glycyl unit (-CH_2CONH-) as a function of urea concentration. We analyzed these data in terms of a statistical thermodynamic formalism to evaluate the equilibrium constant for the reaction in which a urea molecule binds each of the functionalities under study replacing two water molecules. We derived an equation linking the equilibrium constants with changes in free energy, ΔG_{tr}, accompanying the transfer of functional groups from water to concentrated urea solutions. In this equation, ΔG_{tr} is the sum of a change in the free energy of cavity formation, $\Delta\Delta G_C$, and the differential free energy of solute-solvent interactions, $\Delta\Delta G_I$, in a concentrated urea solution and water. With the exception of serine and aspartic acid, the transfer of all amino acid side chains and the glycyl unit from water to 2 M urea is accompanied by favourable changes in $\Delta\Delta G_I$. These results support a direct interaction model in which urea denatures a protein by concerted action *via* favourable solute-cosolvent interactions with a wide range of protein groups, including the peptide backbone and most of the amino acid side chains. We emphasize and

analytically substantiate the need to take into consideration the concentration of a solute when the transfer free energy is determined based on differential water-*versus*-urea solubility measurements. This notion has important ramifications in experimental studies when the transfer free energies of solutes greatly differing in solubility are compared in an attempt to determine the ΔG_{tr} contribution of a particular functional group.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to TVC. SL and YLS acknowledge their graduate support from the CIHR Protein Folding Training Program.

REFERENCES

1. Whitney, P. L.; Tanford, C. J Biol Chem 1962, 237, 1735-1737.
2. Timasheff, S. N. Annu Rev Biophys Biomol Struct 1993, 22, 67-97.
3. Timasheff, S. N. Adv Protein Chem 1998, 51, 355-432.
4. Timasheff, S. N. Biochemistry 2002, 41, 13473-13482.
5. Auton, M.; Bolen, D. W. Proc Natl Acad Sci USA 2005, 102, 15065-15068.
6. Street, T. O.; Bolen, D. W.; Rose, G. D. Proc Natl Acad Sci USA 2006, 103, 13997-14002.
7. Rosgen, J.; Pettitt, B. M.; Bolen, D. W. Protein Sci 2007, 16, 733-743.
8. Hua, L.; Zhou, R.; Thirumalai, D.; Berne, B. J. Proc Natl Acad Sci USA 2008, 105, 16928-16933.
9. Zangi, R.; Zhou, R.; Berne, B. J. J Am Chem Soc 2009, 131, 1535-1541.
10. Lim, W. K.; Rosgen, J.; Englander, S. W. Proc Natl Acad Sci USA 2009, 106, 2595-2600.
11. Makhatadze, G. I.; Privalov, P. L. J Mol Biol 1992, 226, 491-505.
12. Makhatadze, G. I. J Phys Chem B 1999, 103, 4781-4785.

13. Zou, Q.; Habermann-Rottinghaus, S. M.; Murphy, K. P. Proteins 1998, 31, 107-115.

14. Stumpe, M. C.; Grubmuller, H. J Am Chem Soc 2007,129, 16126-16131.

15. Bennion, B. J.; Daggett, V. Proc Natl Acad Sci USA 2003, 100, 5142-5147.

16. Shimizu, S. Proc Natl Acad Sci USA 2004, 101, 1195-1199.

17. Rossky, P. J. Proc Natl Acad Sci USA 2008, 105, 16825-16826.

18. Roseman, M.; Jencks, W. P. J Am Chem Soc 1975, 97, 631-640.

19. Frank, H. S.; Franks, F. J Chem Phys 1968, 48, 4746-4757.

20. Sagle, L. B.; Zhang, Y. J.; Litosh, V. A.; Chen, X.; Cho, Y.; Cremer, P. S. J Am Chem Soc 2009, 131, 9304-9310.

21. O'Brien, E. P.; Dima, R. I.; Brooks, B.; Thirumalai, D. J Am Chem Soc 2007,129, 7346-7353.

22. Almarza, J.; Rincon, L.; Bahsas, A.; Brito, F. Biochemistry 2009, 48, 7608-7613.

23. Stumpe, M. C.; Grubmuller, H. PLoS Comput Biol 2008, 4, e1000221.

24. Lee, M. E.; Van der Vegt, N. F. A. J Am Chem Soc 2006, 128, 4948-4949.

25. Auton, M.; Holthauzen, L. M.; Bolen, D. W. Proc Natl Acad Sci USA 2007, 104, 15317-15322.

26. Bolen, D. W.; Rose, G. D. Annu Rev Biochem 2008, 77, 339-362.

27. Lapanje, S. Biopolymers 1984, 23, 1943-1950.

28. Bolen, D. W. Methods 2004, 34, 312-322.

29. Qu, Y.; Bolen, C. L.; Bolen, D. W. Proc Natl Acad Sci USA 1998, 95, 9268-9273.

30. Nozaki, Y.; Tanford, C. J Biol Chem 1963, 238, 4074-4081.

31. Ahmad, F.; Bigelow, C. C. Biopolymers 1990, 29, 1593-1598.

32. Timasheff, S. N. Biochemistry 1992, 31, 9857-9864.

33. Record, M. T., Jr.; Anderson, C. F. Biophys J 1995, 68, 786-794.

34. Timasheff, S. N. Proc Natl Acad Sci USA 2002, 99, 9721-9726.

35. Tanford, C. J Mol Biol 1969, 39, 539-544.

36. Schellman, J. A. Biopolymers 1990, 29, 215-224.

37. Schellman, J. A. Biophys Chem 1990, 37, 121-140.

38. Schellman, J. A. Biopolymers 1994, 34, 1015-1026.

39. Schellman, J. A. Biophys. J. 2003, 85, 108-125.

40. von Hippel, P. H.; Peticolas, V.; Schack, L.; Karlson, L. Biochemistry 1973, 12, 1256-1264.

41. Davis-Searles, P. R.; Saunders, A. J.; Erie, D. A.; Winzor, D. J.; Pielak, G. J. Ann Rev Biophys Biomol Struct 2001, 30, 271-306.

42. Wang, A. J.; Bolen, D. W. Biochemistry 1997, 36, 9101-9108.

43. Pierotti, R. A. Chem Rev 1976, 76, 717-726.

44. Desrosiers, N.; Desnoyers, J. E. Can J Chem 1976, 54, 3800-3808.

45. Tang, K. E. S.; Bloomfield, V. A. Biophys J 2000, 79, 2222-2234.

46. Lee, S.; Chalikian, T. V. J Phys Chem B 2009, 113, 2443-2450.

47. Eggers, F.; Funck, T. Rev Sci Instrum 1973, 44, 969-977.

48. Sarvazyan, A. P. Ultrasonics 1982, 20, 151-154.

49. Eggers, F. Acustica 1992, 76, 231-240.

50. Eggers, F.; Kaatze, U. Meas Sci Technol 1996, 7, 1-19.

51. Sarvazyan, A. P.; Selkov, E. E.; Chalikian, T. V. Sov Phys Acoust-USSR 1988, 34, 631-634.

52. Sarvazyan, A. P.; Chalikian, T. V. Ultrasonics 1991, 29, 119-124.

53. Lee, S.; Tikhomirova, A.; Shalvardjian, N.; Chalikian, T. V. Biophys Chem 2008, 134, 185-199.

54. Hedwig, G. R.; Reading, J. F.; Lilley, T. H. J Chem Soc Faraday Trans 1991, 87, 1751-1758.

55. Hakin, A. W.; Hedwig, G. R. Phys Chem Chem Phys 2000, 2, 1795-1802.

56. Liu, J. L.; Hakin, A. W.; Hedwig, G. R. J Solution Chem 2001, 30, 861-883.

57. Hedwig, G. R.; Hoiland, H. Phys Chem Chem Phys 2004, 6, 2440-2445.

58. Chalikian, T. V.; Kharakoz, D. P.; Sarvazyan, A. P.; Cain, C. A.; Mcgough, R. J.; Pogosova, I. V.; Gareginian, T. N. J Phys Chem 1992, 96, 876-883.

59. Taulier, N.; Chalikian, T. V. Biophys Chem 2003, 104, 21-36.

60. Chalikian, T. V.; Sarvazyan, A. P.; Funck, T.; Breslauer, K. J. Biopolymers 1994, 34, 541-553.

61. Avbelj, F.; Baldwin, R. L. Proteins: Struct Funct Bioinf 2006, 63, 283-289.

62. Baldwin, R. L. J Biol Chem 2003, 278, 17581-17588.

63. Avbelj, F.; Baldwin, R. L. Proc Natl Acad Sci USA 2009, 106, 3137-3141.

64. Ben-naim, A. J Phys Chem 1978, 82, 792-803.

65. Ben-Naim, A. Statistical Thermodynamics for Chemists and Biochemists; Plenum Press: New York, London, 2002.

66. Ben-Naim, A. Molecular Theory of Solutions; Oxford University Press; Oxford, 2006.

67. Chalikian, T. V. J Phys Chem B 2008, 112, 911-917.

68. Blandamer, M. J.; Davis, M. I.; Douheret, G.; Reis, J. C. R. Chem Soc Rev 2001, 30, 8-15.

69. Schellman, J. A. Biopolymers 1994, 34, 1015-1026.

70. Miyawaki, O.; Saito, A.; Matsuo, T.; Nakamura, K. Biosci Biotechnol Biochem 1997, 61, 466-469.

71. Bower, V. E.; Robinson, R. A. J Phys Chem 1963, 67, 1524-1527.

72. Makhatadze, G. I.; Privalov, P. L. Adv Protein Chem 1995, 47, 307-425.

73. Gong, H.; Rose, G. D. Proc Natl Acad Sci USA 2008, 105, 3321-3326.

74. Schellman, J. A. Annu Rev Biophys Biophys Chem 1987, 16, 115-137.

75. Canchi, D. R.; Paschek, D.; Garcia, A. E. J Am Chem Soc 2010, 132, 2338-2344.

Table 1

Partial molar volume contributions of amino acid side chains, V (-R) (cm^3mol^{-1}), as a function of urea concentration.

SC	0 M	2 M	4 M	6 M	8 M
Ala	16.8±0.3	17.0±0.1	17.0±0.1	17.0±0.3	16.9±0.4
Ala[a]	17.1±0.2	17.1±0.1	17.4±0.1	17.4±0.3	17.4±0.4
Val	47.7±0.1	47.9±0.1	48.0±0.2	48.0±0.1	48.2±0.4
Leu	65.4±0.2	65.4±0.1	65.5±0.1	65.6±0.1	65.8±0.4
Ile	63.2±0.3	63.0±0.2	63.0±0.3	63.2±0.3	63.3±0.4
Pro	35.6±0.1	35.4±0.3	35.4±0.1	35.3±0.1	35.3±0.3
Phe	79.9±0.3	80.0±0.1	80.3±0.3	80.6±0.3	80.7±0.3
Phe[a]	79.9±0.1	80.4±0.1	80.6±0.1	80.9±0.2	80.9±0.4
Trp	102.1±0.1	102.1±0.2	102.4±0.1	102.4±0.2	102.5±0.6

Trp[a]	100.5±0.1	102.4±0.8	102.6±0.3	102.9±0.4	103.1±0.4
Met	62.8±0.1	63.2±0.1	63.4±0.4	63.6±0.1	63.5±0.4
Cys[a]	29.7±0.1	30.5±0.1	30.9±0.2	31.0±0.3	31.2±0.2
Tyr	82.2±0.2	82.7±0.1	83.0±0.1	83.1±0.1	83.1±0.4
Ser[a]	17.1±0.1	17.4±0.1	17.8±0.1	17.9±0.3	17.9±0.4
Thr[a]	33.1±0.1	33.5±0.5	34.0±0.5	34.3±0.3	34.5±0.4
Asn[a]	34.0±0.2	34.8±0.3	35.3±0.2	35.5±0.2	35.7±0.3
Gln	50.8±0.1	51.1±0.2	51.3±0.1	51.5±0.7	51.5±0.4
Asp	31.7±0.1	32.1±0.5	32.3±0.3	32.8±0.2	33.1±0.4
Glu	47.7±0.7	47.9±0.3	48.3±0.2	48.7±0.2	48.7±0.5
His[b]	57.0±0.6	57.5±0.6	58.0±0.3	58.2±0.4	58.5±0.6

Lys	70.1±0.4	70.6±0.5	70.8±1.0	70.8±0.6	71.1±0.6
Arg	67.4±0.3	68.7±0.6	69.8±0.9	70.7±0.6	71.9±0.7
-CH_2CONH-[c]	37.5±0.2	37.0±0.6	37.0±0.7	36.9±0.6	36.8±0.5
-CH_2CONH-[d]	34.9±0.4	35.1±0.3	35.6±0.1	35.7±0.4	35.8±0.2

[a] calculated from N-acetyl amino acid data

[b] calculated from N-acetyl amino acid methylamide data

[c] calculated from the data on oligoglycines

[d] calculated as the difference between the data on N-acetyl glycine methylamide and N-methyl acetamide

Table 2

Partial molar adiabatic compressibility contributions of amino acid side chains, K_S (-R) (10^{-4} $cm^3mol^{-1}bar^{-1}$), as a function of urea concentration.

SC	0 M	2 M	4 M	6 M	8 M
Ala	1.1±0.5	2.7±0.9	3.2±1.1	3.6±0.7	4.0±0.5
Ala[a]	2.8±0.6	3.7±0.7	5.0±1.1	5.2±1.5	5.4±1.2
Val	0.6±0.4	6.6±0.6	10.0±1.0	11.8±0.7	13.4±0.7
Leu	0.4±0.6	8.7±0.5	13.9±0.6	16.6±1.7	18.4±0.6
Ile	-0.9±0.4	5.9±0.7	10.8±0.8	15.6±0.9	18.5±0.6
Pro	-4.0±0.6	1.2±1.2	2.9±0.8	4.5±1.2	4.9±1.0
Phe	1.1±0.5	9.1±0.4	13.3±0.7	16.2±1.1	18.3±0.5
Phe[a]	3.6±0.6	10.0±1.1	13.9±1.0	16.8±1.5	18.6±1.1
Trp	4.0±0.6	11.0±1.1	14.3±0.8	17.0±0.7	19.1±1.5

Trp[a]	3.6±0.6	11.1±1.1	15.5±1.1	19.8±2.0	21.7±1.0
Met	-1.4±0.4	6.6±0.8	10.3±0.7	11.4±0.8	12.2±0.7
Cys[a]	-4.9±0.6	-0.2±0.7	1.8±1.6	2.4±1.5	3.2±1.1
Tyr	7.8±0.5	13.0±0.6	15.2±0.8	16.5±1.1	16.4±0.7
Ser[a]	-2.3±0.6	-2.0±0.4	-1.8±1.1	-1.9±1.5	-1.8±1.0
Thr[a]	-2.2±0.6	1.0±0.8	2.6±1.6	4.2±1.9	4.3±1.1
Asn[a]	-3.3±0.6	-1.1±0.9	0.1±1.1	0.8±1.7	0.8±1.0
Gln	-1.0±0.4	1.7±0.6	3.0±0.6	3.9±1.1	4.2±1.0
Asp	0±0.4	2.8±1.2	3.9±0.8	5.4±0.9	6.4±1.0
Glu	1.9±0.8	5.5±1.2	7.7±2.2	9.9±2.1	10.9±1.5
His[b]	1.8±0.7	5.0±1.3	7.8±0.9	8.2±0.8	8.7±1.4

Lys	-2.4±0.6	5.5±1.2	8.5±2.0	10.7±1.6	12.0±2.4
Arg	-5.1±0.6	1.6±0.8	4.7±1.1	7.2±0.9	7.2±1.1
-CH_2CONH-[c]	-1.8±0.7	-0.1±0.2	1.1±0.2	1.9±0.3	2.3±0.1
-CH_2CONH-[d]	-7.5±0.6	-3.1±0.6	-1.0±0.8	-0.4±0.8	0.3±0.4

[a] calculated from N-acetyl amino acid data

[b] calculated from N-acetyl amino acid methylamide data

[c] calculated from the data on oligoglycines

[d] calculated as the difference between the data on N-acetyl glycine methylamide and N-methyl acetamide

Table 3

The correction factor, γ_1, the number of binding sites for water, n, equilibrium constants, k, and changes in volume, ΔV_0, and adiabatic compressibility, ΔK_{S0}, accompanying the binding of urea to amino acid side chains and the glycyl unit in an ideal solution.

SC	γ_1	n	ΔV_0, cm^3mol^{-1}	$\Delta K_{S0}\times10^4$, $cm^3mol^{-1}bar^{-1}$	k^a, M	k^b, M
Ala	1.0	7	0.05±0.03	0.41±0.08	N/A	0.31±0.07
Ala*	1.0	7	0.08±0.02	-0.37±0.70	0.53±0.67	0.11±0.08
Val	1.0	13	0.03±0.01	2.13±0.03	1.06±0.46	0.22±0.01
Leu	1.0	16	-0.04±0.01	2.60±0.08	0.39±0.06	0.22±0.02
Ile	1.0	16	-0.22±0.07	3.63±0.20	0.15±0.05	0.08±0.01
Pro	1.0	12	-0.14±0.01	1.40±0.03	N/A	0.39±0.06
Phe	1.0	19	-0.03±0.02	1.86±0.01	0.20±0.05	0.21±0.04

Phe*	1.0	19	0.08±0.01	1.47±0.01	0.55±0.11	0.16±0.01
Trp	0.9	23	0±0.02	1.28±0.03	0.62±0.51	0.16±0.02
Trp*	0.9	23	0.22±0.01	1.87±0.05	0.92±0.18	0.12±0.01
Met	0.7	17	0.08±0.01	2.22±0.04	0.51±0.16	0.31±0.02
Cys*	0.3	11	0.34±0.01	3.24±0.09	0.39±0.04	0.18±0.01
Tyr	0.8	20	0.08±0.01	1.02±0.01	0.61±0.13	0.26±0.02
Ser*	0.6	8	0.37±0.17	-2.26±0.57	0.15±0.12	0.04±0.01
Thr*	0.7	11	0.35±0.03	1.56±0.05	0.19±0.04	0.14±0.02
Asn*	0.4	10	0.45±0.03	2.05±0.03	0.32±0.05	0.11±0.01
Gln	0.4	13	0.12±0.01	2.12±0.05	0.29±0.06	0.09±0.01
Asp	0.5	12	0.69±0.28	1.97±0.21	0.04±0.02	0.04±0.01
Glu	0.4	15	0.27±0.17	3.20±0.13	0.07±0.07	0.06±0.01

His	0.4	16	0.24±0.03	2.19±0.17	0.22±0.06	0.11±0.02
Lys	0.7	19	0.08±0.01	2.14±0.05	0.81±0.25	0.23±0.02
Arg	0.5	22	1.02±0.12	2.33±0.09	0.06±0.01	0.12±0.01
-CH_2CONH-[c]	0.4	6	-0.83±0.06	3.16±0.07	N/A	0.08±0.01
-CH_2CONH-[d]	0.4	6	0.43±0.12	4.68±0.13	0.24±0.16	0.23±0.02
Glycine	1.0	15	0.75±0.06	3.79±0.21	0.12±0.02	0.08±0.01

[a] calculated from volume data with Eq. (8)

[b] calculated from compressibility data with Eq. (10)

[c] calculated from the data on oligoglycines

[d] calculated as the difference between the data on N-acetyl glycine methylamide and N-methyl acetamide

Figure Legends

Figure 1

The volume (panel A) and compressibility (panel B) contributions of the leucine side chain as a function of urea. The fitting of the experimental data (continuous lines) was accomplished using Eq. (8) (panel A) and Eq. (10) (panel B) as explained in the text.

Figure 2

The differential free energy of solute-solvent interactions, $\Delta\Delta G_I$, in a urea solution and water calculated as a function of urea concentration with Eq. (12); plot 1 - k=0.04 M (Ser, Asp); plot 2 - k=0.06 M (Glu); plot 3 - k=0.08 M (Gly, Ile); plot 4 - k=0.09 M (Gln); plot 5 - k=0.11 M (Asn, His); plot 6 - k=0.12 M (Arg); plot 7 - k=0.14 M (Thr, Trp); plot 8 - k=0.16 M (glycyl backbone); plot 9 - k=0.18 M (Cys); plot 10 - k=0.19 M (Phe); plot 11 - k=0.21 M (Ala); plot 12 - k=0.22 M (Val, Leu); plot 13 – k=0.23 M (Lys); plot 14 - k=0.26 M (Tyr); plot 15 - k=0.31 M (Met); plot 16 - k=0.39 M (Pro). For the alanine, phenylalanine, tryptophan side chains and the glycyl unit, the average of the two binding constants, k, presented in Table 3 was used in the calculations.

Figure 3

The differential free energy of solute-solvent interactions, $\Delta\Delta G_I$, in a 2 M urea solution and water calculated for the amino acid side chains and the glycyl unit (BB) from water to 2 M urea. For the alanine, phenylalanine, tryptophan side

chains and the glycyl unit, the average of the two binding constants, k, presented in Table 3 was used in the calculations.

Figure 4

The differential free energy of solute-solvent interactions, $\Delta\Delta G_I$, in a urea solution and water for a solute with five binding sites calculated as a function of urea concentration with Eq. (14). The urea binding constant, k, used in calculations is 0.15 M. The concentrations of a solute are 0.1 M (red), 1 M (blue), and 3 M (green).

Figure 1a

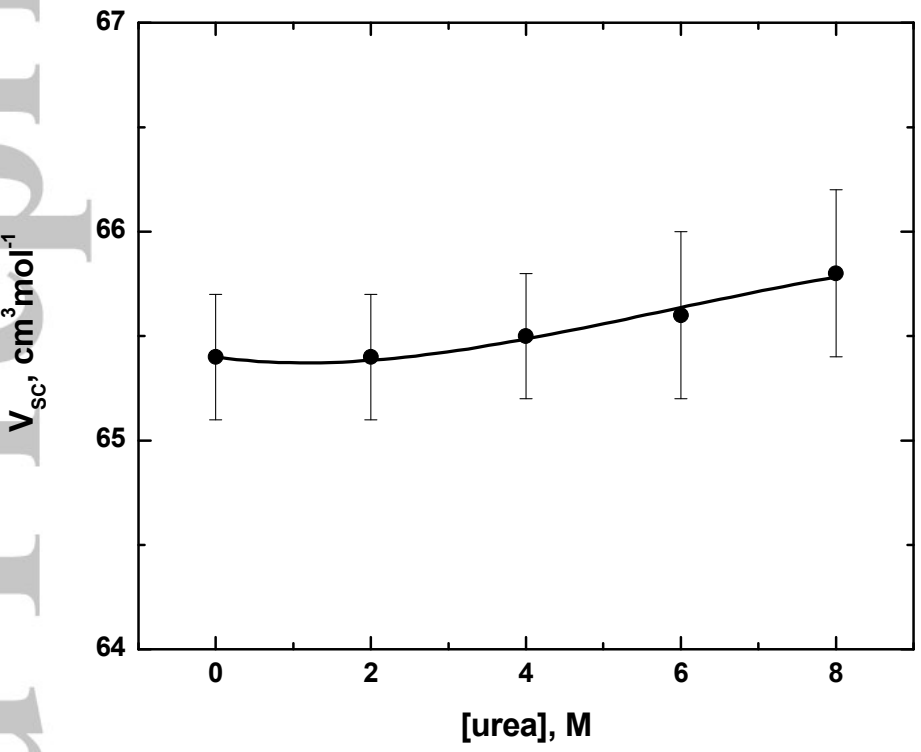

Figure 1b

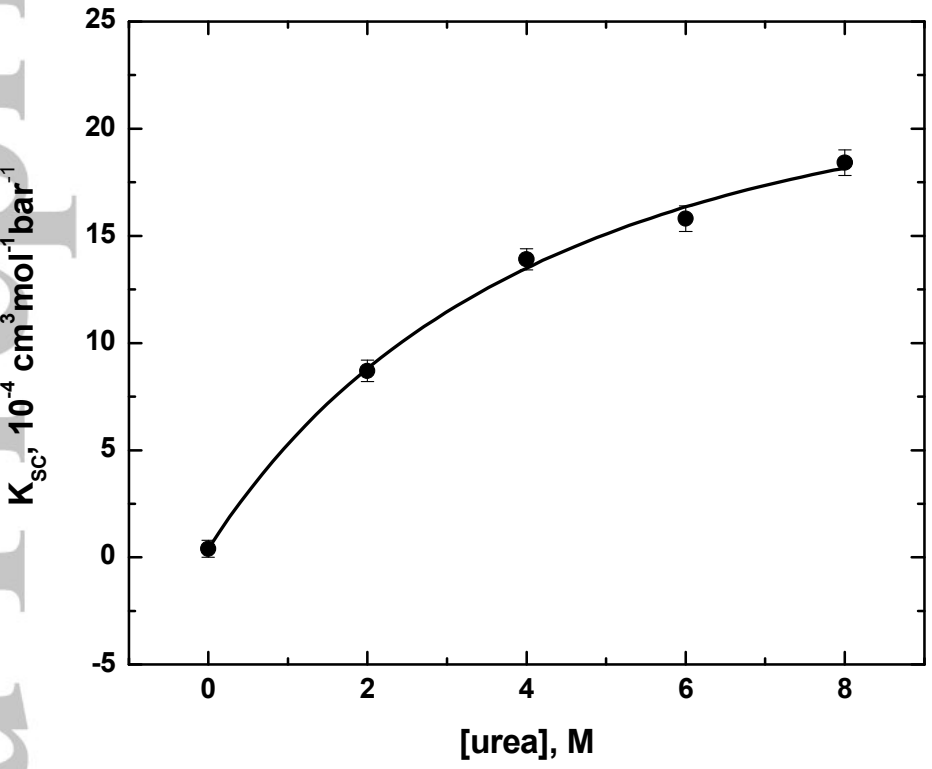

Figure 2

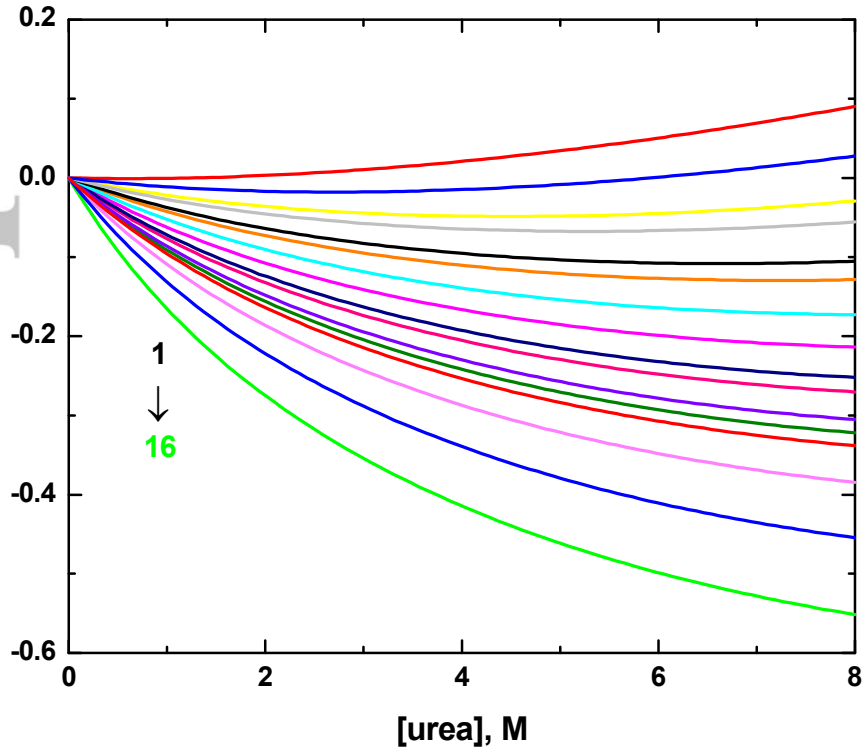

Figure 3

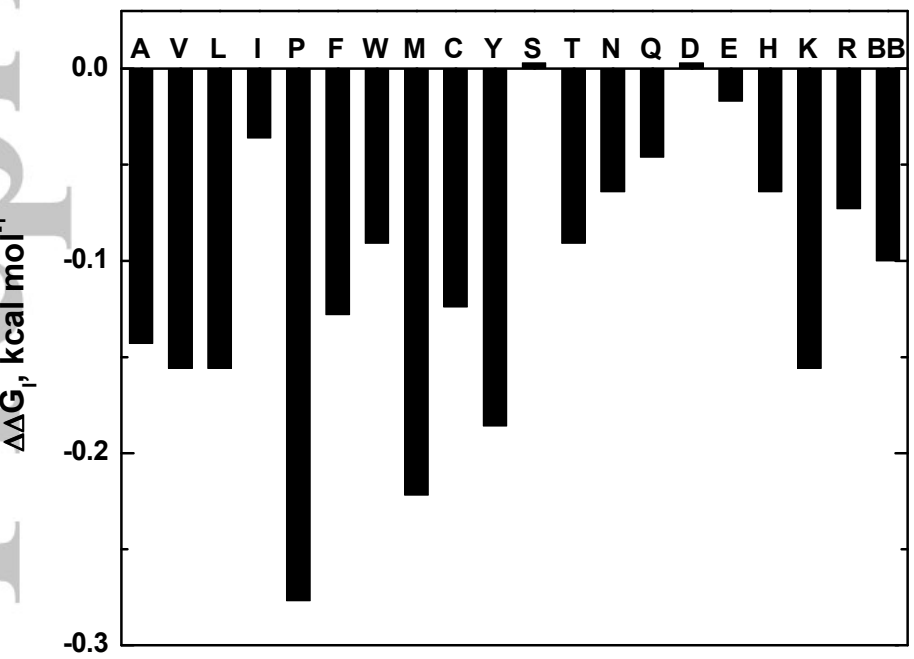

Figure 4