Experimental Atom-by-Atom Dissection of Amide-Amide and Amide-Hydrocarbon Interactions in H₂O

Xian Cheng[1], **Irina A. Shkel**[2,3], **Kevin O' Connor**[2,#], **John Henrich**[2], **Cristen Molzahn**[2], **David Lambert**[2], and **M. Thomas Record Jr**[1,2,3]

[1]Program in Biophysics, University of Wisconsin – Madison, Madison WI 53706

[2]Department of Biochemistry, University of Wisconsin – Madison, Madison WI 53706

[3]Department of Chemistry, University of Wisconsin – Madison, Madison WI 53706

Abstract

Quantitative information about amide interactions in water is needed to understand their contributions to protein folding and amide effects on aqueous processes, and to compare with computer simulations. Here we quantify interactions of urea, alkylated ureas and other amides by osmometry and amide-aromatic hydrocarbon interactions by solubility. Analysis of these data yields strengths of interaction of ureas and naphthalene with amide sp^2O, amide sp^2N, aliphatic sp^3C, and amide and aromatic sp^2C unified atoms in water. Interactions of amide sp^2O with urea and naphthalene are favorable, while amide sp^2O-alkyl urea interactions are unfavorable, becoming more unfavorable with increasing alkylation. Hence amide sp^2O-amide sp^2N interactions (proposed $n-\sigma^*$ hydrogen bond) and amide sp^2O-aromatic sp^2C (proposed $n-\pi^*$) interactions are favorable in water, while amide sp^2O-sp^3C interactions are unfavorable. Interactions of all ureas with sp^3C and amide sp^2N are favorable and increase in strength with increasing alkylation, indicating favorable sp^3C-amide sp^2N and sp^3C-sp^3C interactions. Naphthalene results show that aromatic sp^2C-amide sp^2N interactions in water are unfavorable while sp^2C-sp^3C interactions are favorable. These results allow interactions of amide and hydrocarbon moieties and effects of urea and alkyl ureas on aqueous processes to be predicted or interpreted in terms of structural information. We predict strengths of favorable urea-benzene and N-methylacetamide interactions from experimental information to compare with simulations, and indicate how amounts of hydrocarbon and amide surface buried in protein folding and other biopolymer processes and transition states can be determined from analysis of urea and diethyl urea effects on equilibrium and rate constants.

Corresponding Author mtrecord@wisc.edu.

#Current address: Department of Immunology, Washington University, St. Louis, MO.

ASSOCIATED CONTENT

Supporting Information

Details of experimental methods (vapor pressure osmometry (VPO) and solubility assay). Derivation of equations for, and figures and tables of values of, excess chemical potential derivatives (μ_{22}^{ex}) for two component solutions and chemical potential derivatives (μ_{23}) for three component solutions. Tables and figures of water accessible surface areas (ASA) of solutes and comparison of ASA-based and number-based analyses of μ_{23}. ASA comparison using different structure sources and programs, tests of robustness of α-values.

The authors declare no competing financial interest.

Introduction

Molecular interactions of alkyl ureas with amide or aromatic compounds in water include the N–H ⋯ O=C hydrogen bond (a n – σ^* interaction;[1–4]) between the unified amide sp^2N atom (i.e. –NH- or -NH$_2$) and amide sp^2O, the n – π^* interaction of amide sp^2O with amide sp^2C and aromatic sp^2C,[5–8] interactions between sp^3C and/or sp^2C unified atoms (e.g. the hydrophobic effect,[9–12] CH-π interactions[13–14]), as well as less-characterized interactions between sp^3C and amide sp^2N or amide sp^2O. While of central importance in many areas of the chemical sciences including ligand-protein interactions, protein folding and other self-assembly processes, as well as in applications of amides as solubilizers or destabilizers, little quantitative information has been available to predict or interpret the strengths of these interactions in aqueous solution or the effects of amide solutes on aqueous processes. Recent thermodynamic studies of interactions of a variety of solutes with compounds displaying subsets of protein and nucleic acid functional groups[15–20] indicate the feasibility of obtaining quantitative information about amideamide and amide-hydrocarbon molecular interactions from systematic measurements of interactions of a series of related amide solutes in water. This is the goal of our research.

Interactions of the prototype amide urea with selected model compounds displaying subsets of protein and nucleic acid functional groups in water were quantified by osmometry (for relatively soluble model compounds[16]) and solubility (for sparingly soluble model compounds[18]). These experiments determined the change in chemical potential of the model compound with increasing urea concentration, a quantity which is related to but more fundamental than the free energy of transfer of the model compound from water to 1 molal urea. These experimentally-determined chemical potential increments may be directly and quantitatively compared with predictions from radial distributions of urea in the vicinity of the model compound obtained from molecular dynamics simulations using a Kirkwood-Buff analysis.[21–23]

In aqueous solution, urea reduces the chemical potential of nonelectrolytes displaying protein and nucleic acid functional groups, indicating favorable interactions. Analysis of these data explained quantitatively why urea is an effective solubilizer of many biomolecules and biopolymers and why urea destabilizes aggregates, folded proteins and protein assemblies. Urea interacts favorably with almost all of the major (coarse-grained) functional groups and unified atoms of proteins and nucleic acids and hence favors the direction of processes that expose these groups, like dissolving or unfolding biopolymers.

Interactions of urea with unit surface area (ASA) of amide, carboxylate, and hydroxyl O, amide and cationic N, aromatic sp^2C and aliphatic sp^3C unified atoms of proteins were previously quantified assuming additivity to obtain interaction potentials (α-values[16]). The use of ASA in this analysis is justified by statistical thermodynamic analysis of the solute partitioning model.[15–20] Since the number of model compounds analyzed far exceeded the number of different functional groups, the analysis tested and verified the assumption of additivity. These α-values allow effects of urea on protein processes to be predicted or interpreted in terms of structural information on the amount and composition of the biopolymer surface (ASA) exposed (or buried) in the process. Using these α-values, effects

of urea on the stability of folded proteins (thermodynamic *m*-values) have been interpreted in terms of the ASA of unfolding.[16] Effects of urea or GuHCl on the rate constants of folding and unfolding thirteen globular proteins (kinetic *m*-values) have been interpreted together with activation heat capacities of folding and unfolding to characterize the high free energy folding-unfolding transition state (TS) and obtain compelling evidence for the framework or hierarchic mechanism of folding.[24]

This general approach to the analysis and interpretation of solute interactions and solute effects has also been applied to osmolytes (including glycine betaine,[15, 25] TMAO,[26] proline,[17] trehalose,[27] KGlutamate,[20] glycerol[19]), as well as to urea-nucleobase and -nucleic acid backbone interactions.[18, 28] Chemical (non-coulombic) interactions of the series of Hofmeister salts[29] and chemical (non-excluded volume) interactions of the series of oligo- and polyethylene glycols[19] with amide and hydrocarbon moieties of proteins have also been determined. An alternative approach is based on the analysis of amino acid solubility in the presence and absence of these solutes assuming additivity of solute-backbone and solute-side chain interactions to transfer free energies.[30–34] These approaches have been recently compared.[17] Peptide backbone transfer free energies were correlated with the fraction of polar surface on the solute. This correlation was interpreted using an ASA-based analysis at the level of combined polar (N, O) and nonpolar (C) moieties of the solute.[35]

Here we extend previous studies of urea interactions with the different (coarse-grained) unified atoms or groups of proteins and nucleic acids[16–18, 28] to a series of alkylated ureas and other amides. Osmometry and solubility assays are used to quantify interactions of these amides with each other and with aromatic hydrocarbons. Analysis of these thermodynamic results provides quantitative free energy information about the interactions of these solutes with unified atoms of hydrocarbon and amide groups, sufficient to predict or interpret effects of these ureas on protein folding and other processes. In addition, detailed chemical information is obtained about the various amide-amide and amide-hydrocarbon interactions responsible for protein folding and assembly. These studies also provide the thermodynamic information needed to compare with Kirkwood-Buff analyses of molecular dynamics simulations of amides, peptides and proteins, as has been done for urea.[21–23]

EXPERIMENTAL SECTION

VPO has been shown to be an appropriate assay to quantify interactions of soluble, nonvolatile solutes in water.[15–17, 19–20]

Analysis of Osmolality Data to Quantify Amide Interactions

Differences in osmolality between a three-component amide solution and the corresponding two-component amide solutions ($\Delta Osm(m_2, m_3) = Osm(m_2,m_3) - (Osm(m_2) + Osm(m_3))$) quantify the free energy consequences of interactions between the two amides in water. Analysis based on the Gibbs-Duhem equations for three- and two- component solutions and the relationship between osmolality and solvent activity ($Osm = -m_1 ln a_1$) relates the derivative of $\Delta Osm(m_2,m_3)$ with respect to the concentration product $m_2 m_3$ to the chemical potential derivative ($\partial \mu_2 / \partial m_3)_{P, T,m_2} = \mu_{23} = RT(\partial ln \gamma_2 / \partial m_3)_{P, T,m_2}$ (see Supplemental Eqs. 12–13) with P, T and molality of component 2 constant, which quantifies the interaction

between the two solutes, relative to their interactions with water. [36–38] Integration over the low concentration range where μ_{23} is independent of m_2 and m_3 yields

$$\Delta\mathrm{Osm}(m_2, m_3) = \frac{\mu_{23}}{\mathrm{RT}} m_2 m_3 \quad (1)$$

For the common situation where μ_{23} is independent of m_3 up to 1 molal, μ_{23} can be interpreted as the free energy of transfer of species 2 from water to a 1 molal solution of species 3. Equation 1 is applicable at solute concentrations (m_2, m_3) where only pairwise interactions of solutes are important (typically < 1 molal).

From Equation 1, the slope of a plot of $Osm(m_2,m_3)$ vs. $m_2 m_3$ is μ_{23}/RT. The excess osmolality $Osm(m_2,m_3)$ of the three-component solution is interpreted as the product of the probability of a pairwise interaction of the two solutes (proportional to (m_2,m_3) and the intrinsic strength of that interaction relative to interactions with water (μ_{23}), relative to thermal energy (RT). Two-component osmolalities used to calculate $Osm(m_2, m_3)$ from Eq. 1 are obtained by VPO and analyzed as described in SI. The osmolality $Osm(m_2)$ of solute 2 at concentration m_2 is determined directly. The osmolality $Osm(m_3)$ of solute 3 is determined by interpolation of two-component osmolality data (see Figure S1) using Equation S1 with the fitting coefficients in Table S1. Values of μ_{23}/RT and the corresponding uncertainty are obtained from slopes of plots of $Osm(m_2, m_3)$ vs. $m_2 m_3$ using the statistics program Igor Pro 5.05A.

Solubility Method to Quantify Interaction of Amide and Aromatic Compound

Effects of urea, alkyl ureas and other amides on the solubility m_2^{ss} of naphthalene and/or anthracene are analyzed to obtain μ_{23} using Eq.2, as previously described.[19]

$$\mu_{23} \cong -\mathrm{RT}\left(\frac{\partial \ln(m_2^{ss})}{\partial m_3}\right)_{P,T,\mu_2} \quad (2)$$

Values of $m_{2,0}^{ss}$, the extrapolated molal solubility in the absence of solute, are used to normalize the solubility data. For situations where component 2 (here the aromatic compound) is only sparingly soluble and component 3 (urea, alkyl urea or amide) is in great excess, as is the case here, Eq.2 provides an accurate determination of μ_{23}.

Results

Quantifying Interactions of Ureas with Amide Compounds by VPO

Experimentally-determined excess osmolalities $Osm(m_2, m_3)$ of series of three component solutions of a specified urea and six other ureas as a function of the product of concentrations of the two ureas ($m_2 m_3$) are plotted according to Equation 1 in Figure 1. The upper left panel displays interactions of urea with methyl and ethyl urea and with 1,1-and 1,3-dimethyl and -diethyl ureas. Interactions of these six alkyl ureas with urea and the other alkyl ureas are shown in the remaining panels of Figure 1 and in Figure S2. Excess

osmolality data analogous to Figure 1 for three-component solutions of alkyl ureas and other amides (including formamide and N-methyl formamide, which display a much larger water accessible surface area (ASA) of the amide sp^2C unified atom than other amides studied) are plotted in Figure S3.

For all pairs of amides investigated, excess osmolalities are negative (Osm (m$_2$, m$_3$) < 0). Hence interactions between these amide compounds are favorable, relative to interactions with water. Because the plots are linear, $\mu_2 3$ values determined from the slopes are independent of m$_2$ and m$_3$ in the range investigated. Values of μ_{23} quantifying preferential interactions between pairs of alkyl ureas, determined from Figure 1, are reported in Table S2. Values of μ_{23} for interactions of ureas with other amides, obtained from Figures S2 and S3, are reported in Table S2 and Table S3. Although preferential interactions between smaller members of the urea-alkyl urea series are quite weak (i.e. solute-solute interactions are of similar strength to interactions with water), their μ_{23} values are nevertheless quite well-determined.

Several general trends emerge from these tabulated data.

    i)   Urea (Table S2) and formamide (Table S3), amides that are devoid of aliphatic (sp^3C) hydrocarbon groups, exhibit similar moderately favorable preferential interactions with all alkyl ureas investigated. Values of μ_{23} are −30 to −44 cal mol^{-1} molal^{-1} for urea-alkyl urea interactions and −39 to −53 cal mol^{-1} molal^{-1} for formamide-alkyl urea interactions. The urea-formamide interaction is more favorable (−63 cal mol^{-1} molal^{-1}).

    ii)   Interactions of di-amides malonamide and N-acetylalanine N-methylamide (aama) with urea (μ_{23} values of −53 and −55 cal mol^{-1} molal^{-1}) are significantly more favorable than interactions of mono-amides with urea (μ_{23} values in the range −30 to −44 cal mol^{-1} molal^{-1}).

    iii)   Interactions of the di-amide malonamide with alkyl ureas are less favorable than its interaction with urea, while interactions of the di-amide aama with most alkyl ureas are more favorable than its interaction with urea.

    iv)   Interactions between two alkyl ureas are more favorable than the above interactions of urea or formamide with alkyl ureas, and become increasingly favorable as the extend of alkylation increases. Values of μ_{23} vary from −59 to −146 cal mol^{-1} molal^{-1}

These correlations of strength of interaction with the characteristics of the alkylated urea and/or amide result from differences in the amount and composition (amide O, N, C; aliphatic C) of their water accessible surface area (ASA) and in the intrinsic strengths of the interactions of each of these unified atoms. These are dissected in the Analysis and Discussion section below.

Self-interactions of urea and alkyl ureas are quantified from excess two component osmolalities, as described in Supplemental Text and Figure S4. Values of the chemical potential derivative analogous to μ_{23} (i.e. $\mu_{22}^{\text{ex}}/2$; see Supplemental Eq S7 and Figure S4),

obtained in an analogous manner to the determination of μ_{23} values (Figures 1,S2 and S3) are listed in Table S2. Uncertainties in $\mu_{22}^{ex}/2$ are typically larger than those in μ_{23}.

Quantifying Interactions of Ureas with Aromatic Hydrocarbons by Solubility

Preferential interactions of urea and alkyl ureas with naphthalene and anthracene were determined by solubility assays. The negative logarithm of the solubility, normalized by the fitted solubility in the absence of each urea series solute (intercept value), is plotted as a function of solute molality m_3 in Figure 2. Effects of other amides on the solubility of naphthalene are shown in Figure S5. For the smaller alkylated ureas, where aromatic solubility data were obtained over a wide range of solute concentrations (0 – 4 molal), the initial slope of a fit of the data of Figure 2 obtain μ_{23} at 25°C using Eq 2. Effects of both diethyl ureas on aromatic solubility were determined up to 1.5 molal urea, and linear fits with zero intercept to the data of Figure 2 were used to determine μ_{23}.

Large negative μ_{23} values are observed for the interactions of urea and alkyl ureas with both naphthalene and anthracene (Table S3), indicating favorable interactions of these ureas with aromatic sp^2C unified atoms which become more favorable as the extent of alkylation of the urea increases. Significantly, Table S3 shows that, for every alkyl urea, μ_{23} for the interaction with anthracene exceeds μ_{23} for the interaction with naphthalene by a factor of 1.30 ± 0.16. The ratio of water accessible surface areas (ASA) of these compounds is 334 $Å^2$/273 $Å^2$ = 1.22, while the ratio of sp^2C groups is 14/10 =1.40, so clearly the strengths of these favorable interactions are proportional to the size of the aromatic compound.

Analysis and Discussion

ASA Trends for Ureas and Other Amides—As in previous analyses of interactions of two small solutes in water [15–20] we assume that chemical interactions dominate μ_{23} and that excluded volume effects on these interactions are small. Situations involving interactions of polymeric solutes where excluded volume effects on μ_{23} are significant have been analyzed previously. [39–40] We assume that contributions to μ23 from chemical interactions of one solute with each of the unified atoms of the other solute are additive and that these contributions are proportional to ASA of those unified atoms. Tests of these assumptions are provided here and in the references cited.

The major contributions to water accessible surface areas (ASA) from aliphatic sp^3C, amide sp^2N and amide sp^2O unified atoms for the series of ureas investigated are summarized in Figure S6 and Table S4. For these ureas, the amount of amide sp^2C ASA (3.7 – 7.2 $Å^2$) is always only a small percentage (1 – 4%) of total ASA. All ASA values were obtained from Surface Racer,[41] a robust and user-friendly program. As discussed in SI, the analytic program GETAREA[42] yields identical ASA values, while the numerical program VMD[43] agrees within 4% for the overall ASA, while exhibiting larger deviations in individual atom ASA values (see SI text and Table S21).

Ranked by total ASA, the alkyl ureas form a series marked by large increases in sp^3C ASA and smaller-magnitude, mostly correlated reductions in amide sp^2N ASA, with only minor

reductions in amide sp^2O ASA. Changes in ASA of all atoms in progressing from urea to methyl urea (mu) are larger than those between neighboring members of the series.

ASA information for other amides is also listed in Table S4. Formamide and N-methyl formamide display relatively large amounts of water-accessible amide sp^2C (40.2 $Å^2$ and 39.5 $Å^2$ respectively; 5–10 times that of urea and alkyl ureas). Naphthalene and anthracene display large amounts of aromatic sp^2C ASA (273 $Å^2$, 334 $Å^2$ respectively) for comparison with amide sp^2C and aliphatic sp^3C. Figure 2 reveals favorable interactions of urea and alkyl ureas with these aromatics, indicating favorable interactions with aromatic sp^2C.

Qualitative Dissection of Amide-Amide and Amide-Hydrocarbon Interactions
—Several pairs of amides from the set investigated here differ primarily in ASA of one type of group or unified atom. Figure 3A and Table S5 show that the di-amide aama (N-acetylalanine N-methylamide) differs from 1,3-diethylurea (1,3-deu) primarily in amide sp^2O ASA. Hence differences in μ_{23} values ($\Delta\mu_{23}$) for interactions of aama and 1,3-deu with the alkyl urea series provides information about the interaction of the amide sp^2O group with this series, relative to interactions with water. The bar graph of $\Delta\mu_{23}$ values in Figure 3B shows that the interaction of urea with amide sp^2O is moderately favorable, as found previously.[16–17] Interactions of the series of alkylated ureas in Figure 3B with amide sp^2O are unfavorable and become increasing unfavorable as the extent of alkylation increases.

Figure 3A and Table S5 also show that 1,3-dimethylurea (1,3-dmu) differs from N-methyl acetamide (nma) primarily in amide sp^2N ASA. Hence differences in μ_{23} values ($\Delta\mu_{23}$) for interactions of 1,3-dimethylurea (1,3-dmu) and N-methyl acetamide (nma) with the alkyl urea series provides information about the interaction of the amide sp^2N unified atom with this series. The bar graph of $\Delta\mu_{23}$ values in Figure 3C shows that interactions of alkylated ureas with amide sp^2N are favorable and become increasingly favorable with increasing alkylation of the urea. The behavior of urea, which exhibits a small unfavorable interaction with the predominantly-amide sp^2N ASA in Figure 3C, is unexpected because of the previous finding,[16–17] reinforced below, that the interaction of urea with amide sp^2N is modestly favorable. This is explained by the observations (Figure 3A, Table S5) that the amount of amide O ASA on nma is about 20% greater than for 1,3-dmu and that the interaction of urea with amide sp^2O is very favorable, as deduced here and previously.[16] A second comparison for two amides differing primarily in amide sp^2N ASA (ethyl urea, proprionamide) is shown in Figure S7. In this case, where the difference in amide sp^2O ASA is only 3% (and in the opposite direction; Table S5) the expected favorable interaction of urea with the predominantly amide sp^2N ASA is observed.

Figure 3A shows that 1,1-diethylurea (1,1-deu) differs from 1,1-dimethylurea (1,1-dmu) primarily in aliphatic sp^3C ASA. Hence differences in μ_{23} values ($\Delta\mu_{23}$) for interactions of 1,1-diethylurea (1,1-deu) and 1,1-dimethylurea (1,1-dmu) with the alkyl urea series provide information about the interaction of sp^3C with this series. The bar graph of $\Delta\mu_{23}$ values in Figure 3D shows that the interaction of urea with sp^3C is modestly favorable, as observed previously.[16] Interactions of the series of alkylated ureas in Figure 3D with sp^3C become increasingly favorable as the amount of sp^3C on the alkyl urea increases, indicating that the sp^3C-sp^3C interaction is favorable relative to sp^3C-water interactions, as expected.

Atom-By-Atom Dissection of Interactions of Ureas with Amide Compounds—
The alkylated ureas and other amides investigated here consist entirely of aliphatic sp^3C and amide sp^2O, sp^2N and sp^2C unified atoms. Together these make up a majority of the water-accessible surface (ASA) of most globular protein, and almost all of the ASA exposed in protein unfolding.

New quantitative thermodynamic information of general applicability and significance about how each of these atoms interacts with amide and aromatic solutes and with one other in water (relative to interactions with water) is obtained by analysis of μ_{23} values for these interactions. As in previous analyses of solute-solute interaction, [15–20] we assume that contributions of different interactions to μ_{23} are additive and that strengths of these contributions, relative to interactions with water, are proportional to ASA. The hypothesis of additivity and the use of an ASA-based dissection are tested as part of this analysis of μ_{23} values.

For each amide-amide interaction, μ_{23} is represented as:

$$\mu_{23}=\alpha_{\text{amide } sp^2O}\text{ASA}_{\text{amide } sp^2O}+\alpha_{\text{amide } sp^2N}\text{ASA}_{\text{amide } sp^2N}+\alpha_{\text{amide } sp^2C}\text{ASA}_{\text{amide } sp^2C}+\alpha_{sp^3C}\text{ASA}_{sp^3C}$$

(3)

In Equation 3, the four α-values are intrinsic strengths of interaction of each urea with amide sp^2O, sp^2N and sp^2C and with aliphatic sp^3C, all expressed per unit ASA of the unified atom. For each alkyl urea, a total of twelve μ_{23} values were fitted together to Equation 3 to determine the four α-values. For urea, an additional seven μ_{23} values for urea interactions with five other amides and two hydrocarbons[16] are included in the analysis.

Table 1 lists all α-values with uncertainties obtained from these fits for six alkyl ureas and urea. These are also plotted on the bar graphs in Figure 4. In general, uncertainties are small (<10%). Because these four α-values are in each case highly over-determined (twelve or more equations like Equation 3 in four unknowns), the quality of the fits test and validate the hypothesis of additivity. The ability of these α-values to predict μ_{23} values of the training set and effects of urea on protein processes (see below) also validate additivity.

The left and center panels of Figure 4 show α-values for interactions of urea and the series of alkyl ureas with a unit area of amide sp^2O, amide sp^2N and aliphatic sp^3C groups. The series of ureas is arranged in order of increasing ASA of aliphatic C and concomitant decreasing ASA of amide N (see Figure S6). The order of α-values for interaction of urea and the alkyl urea series with each type of unified atom closely parallels that deduced from the pairwise comparisons of interactions of the alkyl urea series with two amides differing primarily in ASA of one atom type (see Figure 3 and Figure S7). This comparison further validates the hypotheses of additivity and dissection by ASA in Equation 3.

Interactions of Ureas with Amide sp^2O: Figure 4 (left-most panel) compares α-values for interaction of the urea series with amide sp^2O. As deduced previously, the interaction of urea

with amide sp^2O is favorable, which indicates that the NH … O=C hydrogen bond between urea amide sp^2N (which makes up 70% of urea ASA) and amide sp^2O is favorable in water, as expected. Alkylation of urea with one methyl group shifts the interaction with amide sp^2O to unfavorable, and the amide sp^2O-alkyl urea interaction becomes more unfavorable as the extent of alkylation increases. The trend in amide sp^2O α-values for the urea-alkyl urea series in Table 1 (increasingly unfavorable amide sp^2O α-value with increasing ASA of the alkyl urea) correlates with the trends in ASA of the three constituent atom types (Table S4). ASA of amide sp^2N decreases, ASA of aliphatic sp^3C increases and ASA of amide sp^2O decreases as overall ASA increases in this series. The contribution of the amide sp^2O - aliphatic sp^3C interaction to these amide sp^2O α-values can be deduced by comparing 1,1-deu with 1,1-dmu. Figure 3A (also Table S4) shows that the ASA difference between these two alkyl ureas is almost entirely aliphatic sp^3C. Because the amide sp^2O-1.1-dmu α-values is much more unfavorable than for 1,1-dmu, we conclude that the amide sp^2O-aliphatic sp^3C interaction must be unfavorable. Comparison of amide sp^2O α-values for eu and mu, where again the difference in ASA is primarily from aliphatic sp^3C, reinforces this conclusion. The trend in amide sp^2O α-values in Table 1 therefore reflects both the reduction in the favorable amide sp^2N-amide sp^2O, and the increase in the unfavorable aliphatic sp^3C-amide sp^2O interaction with increasing alkylation. (There also is a reduction in the amide sp^2O-amide sp^2O in this series. This interaction is almost certainly unfavorable, relative to interactions with water and therefore must not contribute as much to the trend in amide sp^2O α-values as do amide sp^2N and aliphatic sp^3C.) The amide sp^2O series in Figure 4 closely parallels that in Figure 3B obtained from the difference in μ_{23} values for aama and 1,3-deu which differ primarily in amide sp^2O ASA.

Interactions of Ureas with Amide sp^2N: The left-center panel of Figure 4 ranks α-values quantifying the favorable interactions of the entire urea series with amide sp^2N. Interaction strengths increase from marginally favorable for urea to moderately favorable for the most alkylated ureas. Therefore, the interaction of amide sp^2N with sp^3C must be favorable, relative to interactions of these groups with water. The marginally favorable interaction of urea with amide sp^2N indicates that the interaction of the dominant amide sp^2N groups of urea with amide sp^2N (–NH … N- hydrogen bonds) must be somewhat unfavorable in water, and that this unfavorable interaction is counterbalanced by the favorable interaction of the amide sp^2O of urea with amide sp^2N groups. This amide N series closely parallels that in Figure 3C and Figure S7 obtained from the difference in μ_{23} values for 1,3-dmu and nma which differ primarily in amide sp^2N ASA.

Interactions of Ureas with Aliphatic sp^3C: The center panel of Figure 4 ranks α-values quantifying strengths of interaction of urea and the alkyl urea series with aliphatic sp^3C on the amides investigated. Urea, with no aliphatic sp^3C, exhibits a slightly favorable interaction with aliphatic sp^3C, as previously observed.[16] Generally, as the amount of aliphatic sp^3C ASA on the alkyl urea increases, its interaction with aliphatic sp^3C become increasingly favorable. (The explanation for the deviation of 1,1-dmu from this series order is that the increase in aliphatic sp^3C ASA, relative to ethyl urea, is insufficient to compensate for the loss of amide sp^2N ASA. (See Table S4) Interestingly, for urea and all alkyl ureas investigated, α-values for aliphatic sp^3C are smaller in magnitude than those of

other types of unified atoms on these amides. The aliphatic sp^3C series closely parallels that in Figure 3D, obtained from the difference in μ_{23} values for 1,1-deu and 1,1-dmu which differ primarily in aliphatic sp^3C ASA.

Quantifying Interactions of Ureas with Amide sp^2C and Aromatic sp^2C

Amide sp^2C: Figure 4 (near right panel) shows that α-values for interactions of urea and alkyl ureas with amide sp^2C are favorable, and become increasingly favorable as the extent of alkylation of the urea increases. For all alkyl ureas and urea, α-values for interactions with amide sp^2C are significantly (3 to 10 times) more favorable than α-values for interactions with aliphatic sp^3C (Table 1, Figure 4). The α-value for interaction of urea with amide sp^2C is more favorable than any other urea-amide atom interaction. We propose that this is a very favorable $n - \pi^*$ interaction of n electrons of urea amide sp^2O and the π^* molecular orbital of amide sp^2C[5]. Ongoing dissection of the amide sp^2C - urea α-value into contributions from the amide sp^2O, amide sp^2N and amide sp^2C atoms of urea will test this proposal and quantify the strength of this proposed $n - \pi^*$ interaction.

Aromatic sp^2C: Are strengths of interactions of ureas with aromatic sp^2C similar to those determined above for amide sp^2C? To answer this question, solubility-determined μ_{23} values for interactions of ureas with naphthalene and anthracene were analyzed using Equation 4 to obtain a-values for interactions of each urea with aromatic sp^2C:

$$\mu_{23} = \alpha_{\text{aromatic } sp^2C} \text{ASA}_{\text{aromatic } sp^2C} \quad (4)$$

These aromatic sp^2C α-values are reported in Table 1 and on the right side of the bar graph in Figure 4. Figure 4 shows that aromatic sp^2C α-values for ureas exhibit the same pattern of behavior as observed for amide sp^2C α-values. α-Values for interactions of all ureas with aromatic sp^2C are favorable, becoming increasing favorable with increasing alkylation of the urea. Table 1 shows that aromatic sp^2C α-values are on average 1.3 ± 0.1 times larger than amide sp^2C α-values for these ureas. These sp^2C α-values are so similar that analysis of the combined amide-aromatic μ_{23} data set for each urea using a single sp^2C α-value (floated but found to be the same as that for aromatic sp^2C) provides almost as good agreement between predicted and observed μ_{23} values as that shown in Tables S2 and S3 (see below and Table S6).

Strengths of Interaction of Naphthalene with Amide and Hydrocarbon Groups

—In the previous section, μ_{23} values for interactions of ureas with naphthalene and anthracene (Figure 2) were interpreted by Equation 4 to obtain α-values for interactions of each alkyl urea with aromatic sp^2C, reported in Table 1. These same data can also be analyzed by Equation 3 to determine α-values for interactions of naphthalene with amide sp^2O, amide sp^2N and amide sp^2C and with aliphatic sp^3C. These are also reported in Table 1. Notably, the most favorable interaction of naphthalene is with amide sp^2O. We propose that this is a very favorable $n - \pi^*$ interaction of n electrons of amide sp^2O of these alkyl ureas and the π^* molecular orbital of naphthalene, as identified previously.[6] The intrinsic strength (α-value) of this proposed $n - \pi^*$ interaction is one and one-half times as favorable

as interactions of naphthalene with sp^3C and amide sp^2C, which may be examples of the hydrophobic effect (or CH-π interactions). These C – C interactions in water are favorable because they reduce C ASA and therefore reduce unfavorable interactions between C and water. The naphthalene-amide sp^2N interaction is the only unfavorable component of the interaction of naphthalene with the amides investigated here. Naphthalene and amide sp^2N prefer to interact with water than with one another.

If the α-values for the interactions of naphthalene (entirely aromatic sp^2C) with each type of amide unified atom are assumed to apply to sp^2C generally, they can be used to predict α-values for interactions of the urea series with amide sp^2C. These predictions (Table S6), which as expected are in excellent agreement with the aromatic sp^2C α-values reported in Table 1, are approximately 30% larger in magnitude than the amide sp^2C α-values reported in Table 1 for the alkyl ureas, and about 30% smaller in magnitude for urea. (See Figure S8 for a comparison.) This reinforces the conclusions in previous sections that interactions of amide sp^2C and aromatic sp^2C with ureas are similar, and that an ASA-based analysis of these interactions is appropriate.

Comparison of Observed μ_{23} Values with Those Predicted from α-Values and ASA—Predicted μ_{23} values for alkyl urea – amide and alkyl urea – aromatic interactions, obtained from a-values (Table 1) and ASA information using Equations 3 and 4 are compared with observed μ_{23} values from VPO and solubility experiments in Figure 5A and B. These predicted values of μ_{23} are also listed in S2 and Table S3. Predicted and observed values of μ_{23} agree within the combined uncertainties (+/− 1 SD) for about two-thirds of the interactions investigated. Additivity is verified within the accuracy of the experimental data. No evidence is obtained for significant context-dependence or non-additivity of contributions to μ_{23} or μ_{22}^{ex}. Figure S9 shows an analogous comparison to that of Figure 5 using combined aromatic/amide sp^2C α-values (Table S6) which numerically are found to be the same within uncertainty as aromatic sp^2C α-values (see above). Table S7 displays predicted values for interactions of ureas with amides (formamide and N-methylformamide) containing relative large amounts of amide sp^2C ASA, and also with aromatic hydrocarbons sp^2C; (naphthalene and anthracene) using these combined α values. These predictions are in good agreement with observed μ_{23} values (Table S3). The combined sp^2C fit is almost as good as that obtained using individual sp^2C.

On the molal scale, the chemical potential derivatives μ_{23} and μ_{32} are of necessity equal. Table 1 α-values provide predictions of both μ_{23} and μ_{32} for interactions all pairs of ureas investigated. The comparisons in Figure S10 and Table S8 (see also Table S2) reveal that predicted values of μ_{23} and μ_{32} are in excellent agreement with one another and with the observed values of μ_{23} (Table S2), indicating that α-values of these ureas are well-determined and not context-dependent.

Predicted and observed values of the analogous quantity to μ_{23} that quantifies self-interactions of these ureas (designated $\mu_{22}^{ex}/2$; see SI Eqs S1–7) are compared in Figure S11 and Table S2. Though the experimental uncertainties are larger than for μ_{23}, the agreement again is good. Values of $\mu_{22}^{ex}/2$ for self-interactions of the urea series are listed together with μ_{23} values for mutual interactions in Table S2.

Comparison of Analyses of Amide Interactions Based on Atom ASA and Atom Number—Figure 5 shows good agreement between observed μ_{23} values and those predicted by the SPM-motivated, ASA-based analysis of individual solute-atom interactions in Eq. 3, and Figure 4 shows the clear chemical significance of the alpha values obtained from this analysis. Equally good agreement between predicted and observed μ_{23} values for other solutes has been obtained previously, [15–20] and the chemical significance of the alpha values obtain in these studies was also clear. An alternative method of analysis of similar thermodynamic data was based on the number of atoms of each type[44] and not their ASA. Here we compare predictions of μ_{23} from ASA-based (Equation 3) and number-based (Equation S14) analyses with experimental μ_{23} values for the training set of interactions of alkyl ureas with amide compounds.

Figure S12 plots predicted and observed μ_{23} values for four possible number-based analyses of the 12-member alkyl urea –amide μ_{23} data sets, using Equation S14 with different numbers of fitting parameters (3 or 4 interaction potentials (α_n-values), 0 to 5 global relative weights ($w_{x,y}$) for unified C or N atoms with different numbers of H). For comparison, the ASA-based analysis (Equation 3) used here to interpret these μ_{23} values has 4 α-values as parameters. Details of the different number-based analyses are presented in Supplemental and fitting parameters for each case are summarized in Table S9. Versions of the number-based analyses with a similar number of fitting parameters to the ASA-based analysis are statistically less satisfactory, as seen by comparing Figure S12A and B with Figure 5A. The sum of squares of residuals for these number-based fits is 5 – 10 times larger than for the ASA-based fit. (See Table S9)

Increasing the number of parameters in the number-based analysis by including 5 relative weighting factors (Table S9) in a global fit to Equation S14 provides a fit of comparable statistical quality to the ASA-based fit, as seen by comparing Figure S12D with Figure 5A. In this number-based analysis, NH and CH_2 unified atoms have relative weights of approximately 0.5, so their α_n-values are half as large as those of NH_2 and CH_3 unified atoms, making this analysis more like the ASA-based analysis. A similar pattern in relative weighting factors was observed in a number-based analysis of logarithms of activity coefficients.[45] A concern about the physical significance of the number-based fit of Figure S12D is that α_n-values quantifying interactions of urea with amide sp^2C and amide sp^2O are unusual. In the ASA-based urea analysis (Table 1), α-values for these interactions are both very favorable, relative to interactions of urea with other atom types, and urea α-values for interactions of urea with amide sp^2C and aromatic sp^2C are approximately equal. The strength and significance of the urea-amide sp^2O interaction was quantified previously.[16, 24, 38, 46] But in the number-based analysis, the interaction of urea with sp^2O is implausibly weak, with a near-zero α_n-value, and the urea-amide sp^2C interaction is implausibly strong, with an α_n-value that is 3- to 4-fold more favorable than that for aromatic sp^2C. We therefore conclude that the ASA-based analysis is both statistically and chemically more appropriate to analyze interactions involving both polar and nonpolar portions of these solutes.

Alkyl Ureas as Solubilizing and Unfolding Agents; Use as Solute Probes of Coupled Folding—Urea has been widely used to destabilize (unfold) globular proteins

and α-helices, to disassemble protein quaternary structures, and to solubilize inclusion bodies and protein aggregates. Only limited applications of alkyl ureas have been made in these areas, but the results of this paper allow us to predict the effectiveness of alkyl ureas. Composite α-values useful to predict effects of alkyl ureas on protein stability are given in Table S10. For example, the two diethyl ureas studied are predicted to be as effective as protein unfolding agents as GuHCl and twice as effective as urea. While the reduced solubility of the diethyl ureas compared to urea makes them less suitable than urea and GuHCl for routine applications, the prediction that diethyl ureas unfold proteins entirely because of their favorable interactions with hydrocarbon (sp^3C, sp^2C) ASA that is exposed in unfolding, while GuHCl and urea unfold proteins primarily because of their favorable interactions with the amide sp^2O ASA exposed in unfolding, should make diethyl ureas very useful reagents (Table S10). Salts of the protein-destabilizing tetrapropyl ammonium cation[47] would be an alternative to the diethyl ureas for this purpose, once quantitative information (β-values)[16] about the interactions of this cation with amide and hydrocarbon atoms is obtained.

Urea has been used in conjunction with other solutes as a quantitative probe of the amount and composition of protein surface exposed or buried in global or local folding or other large conformational changes/interface formation in the intermediates and high free energy transition state (TS) of protein processes. In particular, urea kinetic m-values for protein folding, together with activation heat capacity changes for folding, have been interpreted to characterize TS for folding of 13 single domain globular proteins.[24] Urea interacts most favorably with amide sp^2O and aromatic sp^2C groups, and less favorably with amide sp^2N and aliphatic sp^3C groups. The urea kinetic m-value is therefore primarily a measure changes in amide ASA in folding to TS, while the activation heat capacity change is primarily a measure of changes in hydrocarbon ASA in folding to TS. Results of the current analysis show that kinetic m-values for 1,3- or 1,1-diethylurea (or, by extrapolation, butyl urea) should be effective replacements for (and much more straightforward to measure than) activation heat capacity changes as measures of hydrocarbon ASA buried in folding to TS. Determination of kinetic or thermodynamic m-values for this combination of solutes (urea or GuHCl; 1,3- or 1,1-diethylurea) should be effective to detect and quantify coupled folding in any protein process, including protein-nucleic acid binding. Comparison of predicted thermodynamic folding m-values using only hydrocarbon and amide α-values (this paper) with predictions including minor contributors to the ASA (Supp. Figure S13;[16–17]) reveals that agreement with experiment is equally good for both sets of α-values. Urea α-values (Table 1) are in good agreement with previously-published results[16] (Table S11).

In a model protein process investigated both experimentally and by simulations, addition of urea modestly favors hydrophobic collapse of the polyamide PNIPAM in water, while methyl, dimethyl and tetramethyl urea oppose hydrophobic collapse.[48–49] Hydrophobic collapse buries hydrocarbon moieties of PNIPAM, leaving side chain amides exposed.[49] Based on the α-values of Table 1, alkyl ureas oppose hydrophobic collapse because they interact favorably with aliphatic sp^3C and unfavorably with amide O. Urea may favor hydrophobic collapse by being integrated into the collapsed form by favorable hydrogen bonding interactions with amide O,[48] and/or because the burial of side chain isopropyl groups may preferentially expose amide O on the surface of the collapsed form. Both these

interpretations are consistent with the very favorable urea-amide O α-value (Table 1 and reference[16]).

Comparison with MD Simulations—Molecular dynamics (MD) simulations with various force fields (FF) have been used to characterize the interactions of urea with hydrocarbons, amino acids, peptides, and proteins, and the interactions of alkyl ureas with the low-solubility aromatic compound nifedipine. In general, MD results can be compared with experimental thermodynamic data and analysis in two ways. The MD simulation yields the radial distribution of urea in the vicinity of the compound or individual side chain or backbone atoms, indicating the extent of global or local accumulation or exclusion of urea for qualitative comparison with the sign and magnitude of μ_{23} or α values. At a quantitative level the integrated radial distribution function can be compared with μ_{23} values using Kirkwood Buff theory.[22–23, 50–51]

Here and previously, the experimental finding is that urea and alkyl ureas interact favorably with and are therefore accumulated near most biochemical compounds.[16–17] α-Values obtained by dissection of these data predict that urea accumulates in the vicinity of all the common unified atoms of proteins except cationic N. In particular urea is predicted to interact favorably with the peptide backbone, the C-terminal carboxylate, and most side chains (except cationic nitrogens of lysine and arginine) of peptides and proteins. Simulations of urea-hydrocarbon interactions in water predict favorable interactions of urea with aliphatic sp^3C[52–53] and aromatic sp^2C,[23] consistent with the negative α-values of these interactions (Table 1, also[16]). Simulation of urea-peptide interactions (i.e.[54–60]) predict favorable interactions of urea with amide O and amide N, also in agreement with the observed negative urea α-values (Table 1, also[16]). Simulations reveal that the extent of accumulation of alkylated ureas in the vicinity of nifedipine increases with increasing length of the alkyl side chain of the urea,[61] consistent with the experimentally determined increase in solubility with increasing alkyl chain length and with expectation based on α-values for interaction of these alkyl ureas with sp^3C reported here (Table 1).

MD simulations of two component solutions of N-methyl acetamide (nma; mole fraction $X_{nma} = 0.1$; about 6 molal) in water at 313 K were performed with additive (CHARMM) and polarizable (Drude) force fields to obtain radial distributions predicting the local accumulation or exclusion of amide sp^2O, amide sp^2N, methyl sp^3C and water sp^2O unified atoms in the vicinity of one another.[58] CHARMM radial distributions for the two sp^3C unified atoms of nma, one bonded to the amide C and the other to the amide N are very similar, revealing little if any context-dependence of the interactions of these sp^3C atoms of one nma with sp^3C, sp^2O and sp^2N atoms of another nma. Some context dependence could be present for sp^3C – water sp^2O interactions. For the Drude force field, sp^3C-sp^3C and sp^3C-water interactions could exhibit some context dependence while sp^3C- sp^2O and sp^3C-sp^2N interactions appear not to. Most of these MD predictions agree with the finding, in the current study and previous research,[16] that experimental μ_{23} values for interactions of a solute with model compounds displaying different combinations of the same unified atoms can be dissected using additivity (Equation 3) to obtain context-independent α-values quantifying the strength of interaction of each coarse-grained type of unified atom (Table 1).

Kirkwood-Buff integrals of MD simulation results for nma-water mixtures provide the information needed to evaluate $\mu_{22}^{ex}/2$ for the self-interaction of nma at 6 molal in water.[58] From the published graphical results, we obtain estimates of $\mu_{22}^{ex}/2$ for these force fields (Table S12). For CHARMM, $\mu_{22}^{ex}/2 = -40$ cal mol^{-1} molal^{-1}, while for Drude $\mu_{22}^{ex}/2 = 16$ cal mol^{-1} molal^{-1}. The favorable $\mu_{22}^{ex}/2$ obtained from the CHARMM force field is quite consistent with our observation of moderately favorable $\mu_{22}^{ex}/2$ or μ_{23} for interactions of alkyl ureas of similar size or composition to nma (-40 to -80 cal mol^{-1} molal^{-1} at amide concentrations below 1 molal; see Table S2). Previous thermodynamic studies of two component nma solutions at higher temperature yield Kirkwood-Buff estimates of $\mu_{22}^{ex}/2$ (in one case extrapolated to infinite dilution and 25 °C) that are favorable, though at the small-magnitude end of the range of our estimates.[62–63] Hence the CHARMM predictions of nma interactions appear more reliable than those obtained from the Drude force field as currently parameterized. Simulations of solutions of the alkyl ureas investigated in the present study should allow optimization of both force fields for amide simulations.

Two-component thermodynamic information for urea solutions was used to develop a Kirkwood-Buff force field (KBFF) for simulating urea interactions.[21] Using this KBFF, Ganguly et al simulated interactions of urea with benzene in water.[23] The molal scale μ_{23}-value for the urea – benzene interaction is readily predicted from their results using Kirkwood – Buff analysis.[22] A range of μ_{23} values is obtained depending on concentrations of urea and benzene and the details of how the coarse-grained potential of mean force is normalized. The average μ_{23} value at the lower concentrations of urea simulated (4 molal, 6 molal) is -109 ± 16 cal mol^{-1} molal^{-1}, the same within uncertainty as the experiment-based prediction of the urea – benzene μ_{23} value of -124 ± 3 cal mol^{-1} molal^{-1} at low urea concentration (< 1 molal), obtained using the urea-aromatic C α-value (-0.59 ± 0.01 cal mol^{-1} molal^{-1} Å^{-2}) and the ASA of benzene (212 Å^2).

Comparison of the extensive set of μ_{23} and μ_{22}^{ex} values for interactions of urea and alkyl ureas with one another, with other amides and with aromatic hydrocarbons reported here with results of MD simulations may be useful to obtain a more detailed accounting of amide-amide and amide-hydrocarbon interactions, relative to interactions with water. MD force fields that reproduce these μ_{23} values should be useful to simulate interactions of amide compounds and amide effects on protein processes, and to simulate protein folding where the surface buried is almost entirely amide and hydrocarbon.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Alabugin IV, Manoharan M, Peabody S, Weinhold F. J. Am. Chem. Soc. 2003; 125:5973–5987. [PubMed: 12733938]

2. Weinhold F. Adv. Prot. Chem. 2005; 72:121–155.

3. Khaliullin RZ, Cobar EA, Lochan RC, Bell AT, Head-Gordon M. J. Phys. Chem. A. 2007; 111:8753–8765. [PubMed: 17655284]

4. Newberry RW, Raines RT. Nat. Chem. Biol. 2016; 12:1084–1088. [PubMed: 27748749]

5. Bartlett GJ, Choudhary A, Raines RT, Woolfson DN. Nat. Chem. Biol. 2010; 6:615–620. [PubMed: 20622857]

6. Gorske BC, Bastian BL, Geske GD, Blackwell HE. J. Am. Chem. Soc. 2007; 129:8928–8929. [PubMed: 17608423]

7. Bartlett GJ, Newberry RW, VanVeller B, Raines RT, Woolfson DN. J. Am. Chem. Soc. 2013; 135:18682–18688. [PubMed: 24256417]

8. Newberry RW, VanVeller B, Guzei IA, Raines RT. J. Am. Chem. Soc. 2013; 135:7843–7846. [PubMed: 23663100]

9. Baldwin RL. Proc. Natl. Acad. Sci. U.S.A. 2013; 110:1670–1673. [PubMed: 23319615]

10. Tanford C. Science. 1978; 200:1012–1018. [PubMed: 653353]

11. Baldwin RL. Proc. Natl. Acad. Sci. U.S.A. 2014; 111:13052–13056. [PubMed: 25157156]

12. Spolar RS, Record MT Jr. Science. 1994; 263:777–784. [PubMed: 8303294]

13. Dey RC, Seal P, Chakrabarti S. J. Phys. Chem. A. 2009; 113:10113–10118. [PubMed: 19678694]

14. Plevin MJ, Bryce DL, Boisbouvier J. Nat. Chem. 2010; 2:466–471. [PubMed: 20489715]

15. Capp MW, Pegram LM, Saecker RM, Kratz M, Riccardi D, Wendorff T, Cannon JG, Record MT Jr. Biochemistry. 2009; 48:10372–10379. [PubMed: 19757837]

16. Guinn EJ, Pegram LM, Capp MW, Pollock MN, Record MT Jr. Proc. Natl. Acad. Sci. U.S.A. 2011; 108:16932–16937. [PubMed: 21930943]

17. Diehl RC, Guinn EJ, Capp MW, Tsodikov OV, Record MT Jr. Biochemistry. 2013; 52:5997–6010. [PubMed: 23909383]

18. Guinn EJ, Schwinefus JJ, Cha HK, McDevitt JL, Merker WE, Ritzer R, Muth GW, Engelsgjerd SW, Mangold KE, Thompson PJ, Kerins MJ, Record MT Jr. J. Am. Chem. Soc. 2013; 135:5828–5838. [PubMed: 23510511]

19. Knowles DB, Shkel IA, Phan NM, Sternke M, Lingeman E, Cheng X, Cheng LX, O'Connor K, Record MT. Biochemistry. 2015; 54:3528–3542. [PubMed: 25962980]

20. Cheng X, Guinn EJ, Buechel E, Wong R, Sengupta R, Shkel IA, Record MT. Biophys. J. 2016; 111:1854–1865. [PubMed: 27806267]

21. Weerasinghe S, Smith PE. J. Phys. Chem. B. 2003; 107:3891–3898.

22. Smith PE. Biophys. J. 2006; 91:849–856. [PubMed: 16679363]

23. Ganguly P, van der Vegt NF. J. Chem. Theory Comput. 2013; 9:5247–5256. [PubMed: 26592264]

24. Guinn EJ, Kontur WS, Tsodikov OV, Shkel I, Record MT Jr. Proc. Natl. Acad. Sci. U.S.A. 2013; 110:16784–16789. [PubMed: 24043778]

25. Bhojane PP, Duff MR, Bafna K, Rimmer GP, Agarwal PK, Howell EE. Biochemistry. 2016; 55:6282–6294. [PubMed: 27768285]

26. Hong J, Xiong SQ. Biophys. J. 2016; 111:1866–1875. [PubMed: 27806268]

27. Hong J, Gierasch LM, Liu Z. Biophys. J. 2015; 109:144–153. [PubMed: 26153711]

28. Lambert D, Draper DE. Biochemistry. 2012; 51:9014–9026. [PubMed: 23088364]

29. Pegram LM, Record MT Jr. J. Phys. Chem. B. 2008; 112:9428–9436. [PubMed: 18630860]

30. Nozaki Y, Tanford C. J. Biol. Chem. 1963; 238:4074–4081. [PubMed: 14086747]

31. Auton M, Rosgen J, Sinev M, Holthauzen LM, Bolen DW. Biophys. Chem. 2011; 159:90–99. [PubMed: 21683504]

32. Auton M, Bolen DW. Biochemistry. 2004; 43:1329–1342. [PubMed: 14756570]

33. Auton M, Bolen DW. Proc. Natl. Acad. Sci. U.S.A. 2005; 102:15065–15068. [PubMed: 16214887]

34. Auton M, Holthauzen LM, Bolen DW. Proc. Natl. Acad. Sci. U.S.A. 2007; 104:15317–15322. [PubMed: 17878304]

35. Street TO, Bolen DW, Rose GD. Proc. Natl. Acad. Sci. U.S.A. 2006; 103:13997–14002. [PubMed: 16968772]

36. Timasheff SN. Adv. Protein Chem. 1998; 51:355–432. [PubMed: 9615174]

37. Ben-Naim A. Pure & Appl. Chem. 1990; 62:10.

38. Cannon JG, Anderson CF, Record MT Jr. J. Phys. Chem. B. 2007; 111:9675–9685. [PubMed: 17658791]

39. Shkel IA, Knowles DB, Record MT Jr. Biopolymers. 2015; 103:517–527. [PubMed: 25924886]

40. Knowles DB, LaCroix AS, Deines NF, Shkel I, Record MT Jr. Proc. Natl. Acad. Sci. U.S.A. 2011; 108:12699–12704. [PubMed: 21742980]

41. Tsodikov OV, Record MT Jr, Sergeev YV. J. Comput. Chem. 2002; 23:600–609. [PubMed: 11939594]

42. Fraczkiewicz R, Braun W. J. Comput. Chem. 1998; 19:319–333.

43. Humphrey W, Dalke A, Schulten K. J. Mol. Graph. 1996; 14:33–38. 27–38. [PubMed: 8744570]

44. Okamoto BY, Wood RH, Thompson PT. J. Chem. Soc. Faraday Trans. I. 1978; 74:1990–2007.

45. Admire B, Lian B, Yalkowsky SH. Chemosphere. 2015; 119:1441–1446. [PubMed: 25454206]

46. Felitsky DJ, Record MT Jr. Biochemistry. 2003; 42:2202–2217. [PubMed: 12590610]

47. Dempsey CE, Mason PE, Jungwirth P. J. Am. Chem. Soc. 2011; 133:7300–7303. [PubMed: 21520945]

48. Sagle LB, Zhang Y, Litosh VA, Chen X, Cho Y, Cremer PS. J. Am. Chem. Soc. 2009; 131:9304–9310. [PubMed: 19527028]

49. Rodriguez-Ropero F, van der Vegt NF. J. Phys. Chem. B. 2014; 118:7327–7334. [PubMed: 24927256]

50. Ben-Naim A. J. Chem. Phys. 1977; 67:4884–4890.

51. Chitra R, Smith PE. J. Phys. Chem. B. 2001; 105:11513–11522.

52. Trzesniak D, van der Vegt NFA, van Gunsteren WF. Phys. Chem. Chem. Phys. 2004; 6:697–702.

53. Trzesniak D, van der Vegt NFA, Van Gunsteren WF. Mol. Phys. 2007; 105:7.

54. Hua L, Zhou R, Thirumalai D, Berne BJ. Proc. Natl. Acad. Sci. U.S.A. 2008; 105:16928–16933. [PubMed: 18957546]

55. Ma L, Pegram L, Record MT Jr, Cui Q. Biochemistry. 2010; 49:1954–1962. [PubMed: 20121154]

56. Canchi DR, Garcia AE. Biophys. J. 2011; 100:1526–1533. [PubMed: 21402035]

57. Canchi DR, Garcia AE. Annu. Rev. Phys. Chem. 2013; 64:273–293. [PubMed: 23298246]

58. Lin B, Lopes PE, Roux B, MacKerell AD Jr. J. Chem. Phys. 2013; 139:084509. [PubMed: 24007020]

59. Holehouse AS, Garai K, Lyle N, Vitalis A, Pappu RV. J. Am. Chem. Soc. 2015; 137:2984–2995. [PubMed: 25664638]

60. Steinke N, Gillams RJ, Pardo LC, Lorenz CD, McLain SE. Phys. Chem. Chem. Phys. 2016; 18:3862–3870. [PubMed: 26764567]

61. Cui Y. J Pharm (Cairo). 2013; 2013:791370. [PubMed: 26555993]

62. Zielkiewicz J. J. Chem. Thermodyn. 1999; 31:819–825.

63. Marcus Y. J. Chem. Soc. Faraday Trans. 1990; 86:2215–2224.
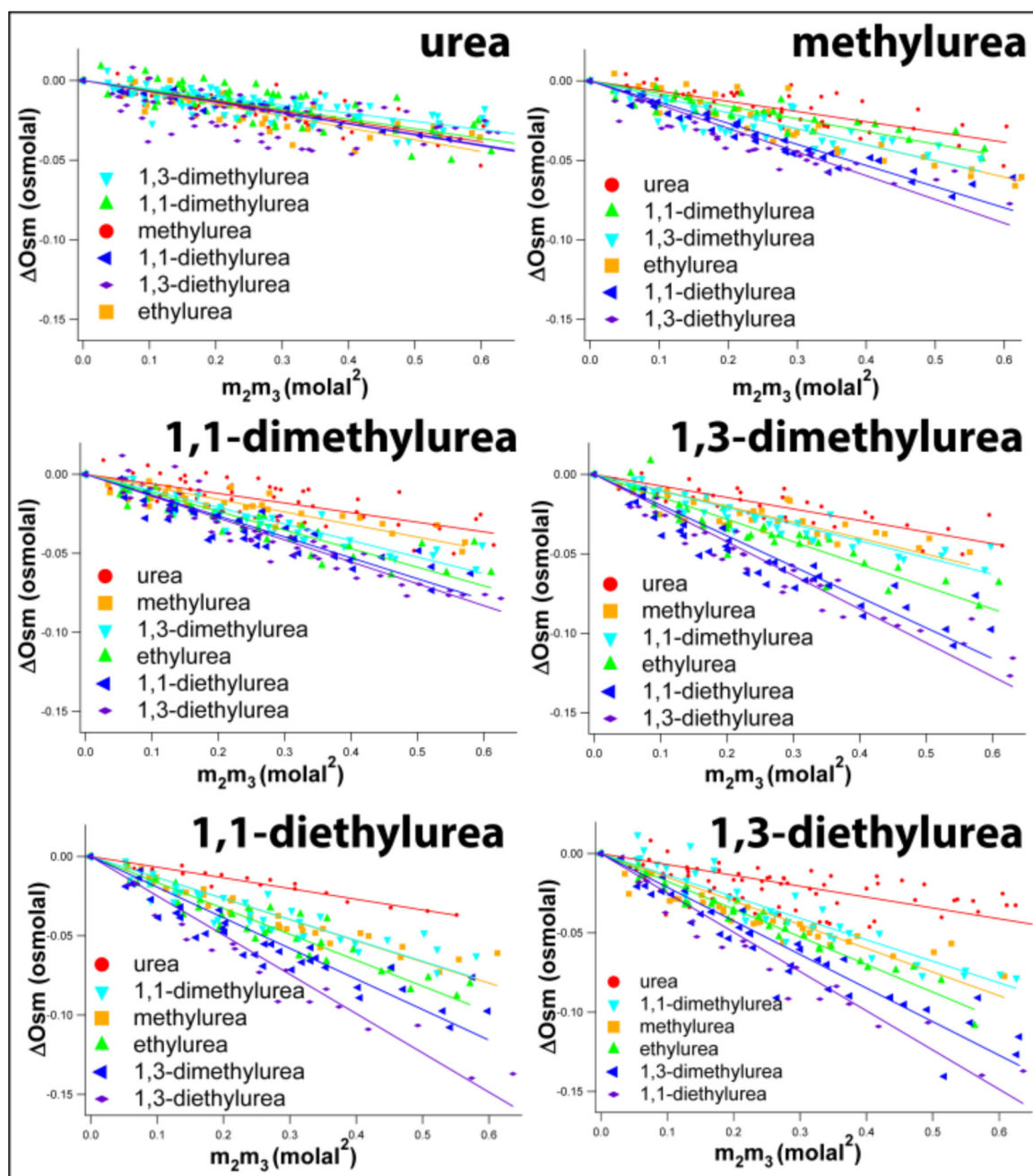
Figure 1.
Amide-Amide Interactions: Interactions of Urea and Five Alkyl Ureas with other Ureas Determined by Vapor Pressure Osmometry (VPO) at 23°C. In all cases the osmolality difference $\Delta Osm(m_2, m_3) = Osm\,(m_2,m_3) - Osm\,\mathbf{(m_2)} - Osm\,\mathbf{(m_3)}$ between a three-component solution and the corresponding two-component solutions is plotted according to Equation 1 vs. the product of molal concentrations (m_2m_3) of the two ureas. Slopes of linear fits with zero intercept yield chemical potential derivatives ($\partial\mu_2/\partial m_3)_{P,T,m2}$ quantifying preferential interactions between the two amides in water (Equation 1). Analogous VPO data for interactions of ethyl urea and of these six ureas with amides are shown in Supplemental
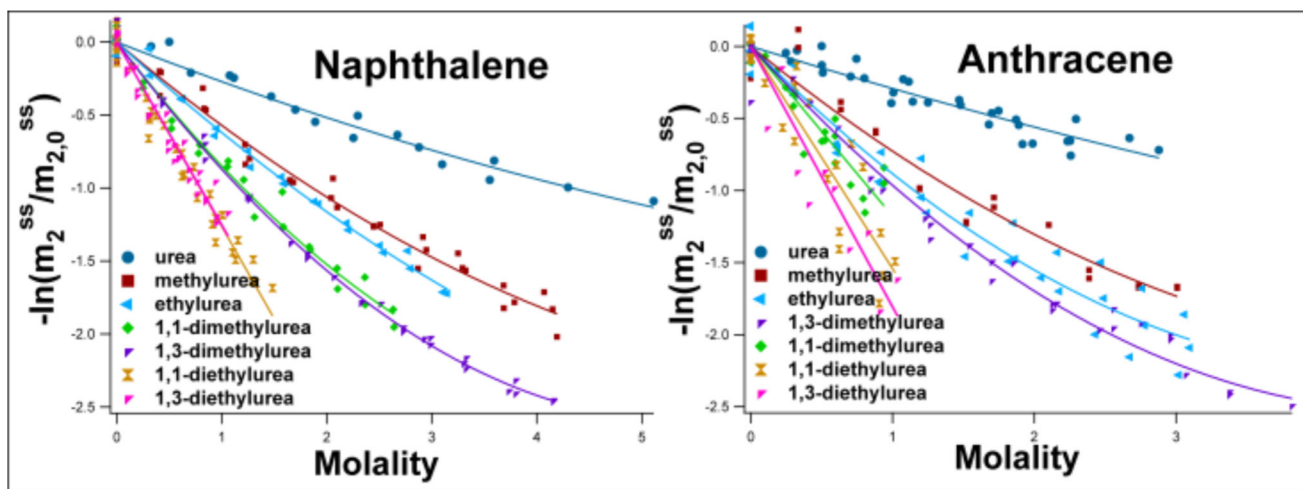
Figures S2 and S3, and VPO data for self-interactions of ureas and amides are shown in Supplemental Figures S1 and S4.
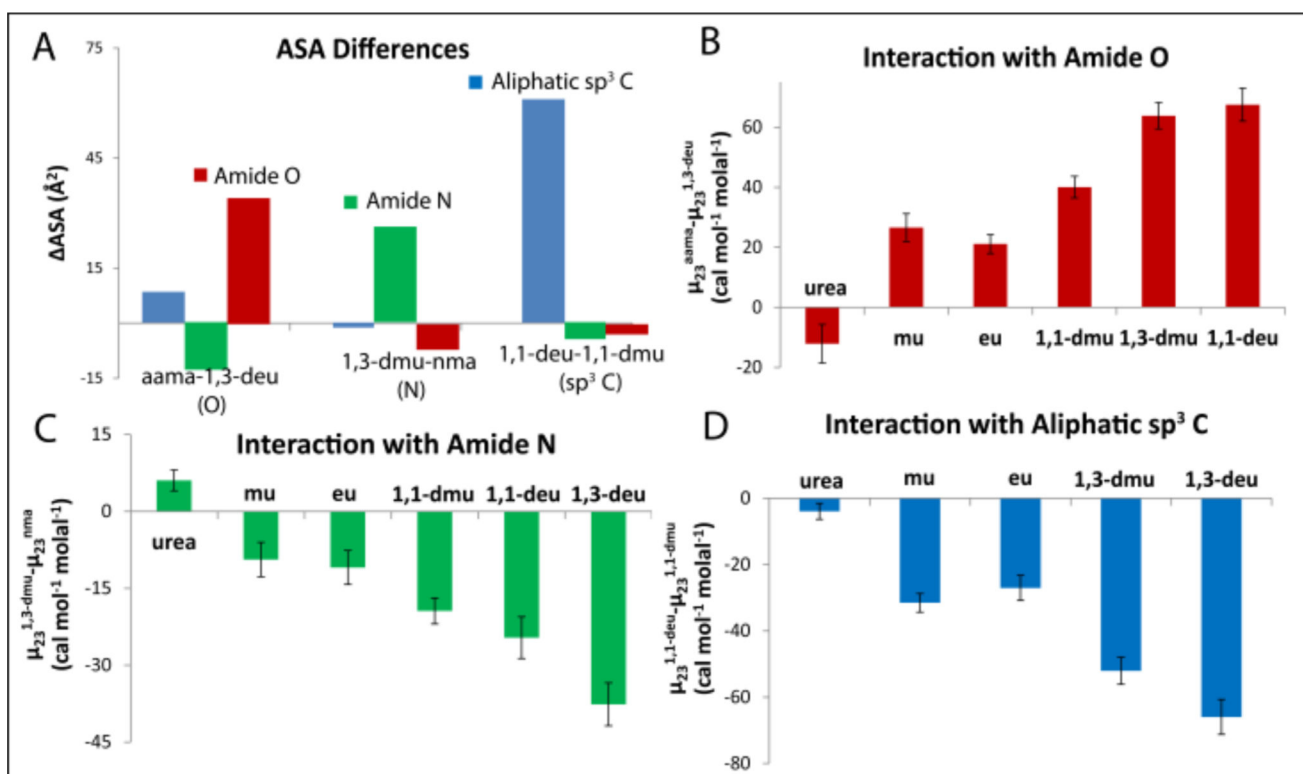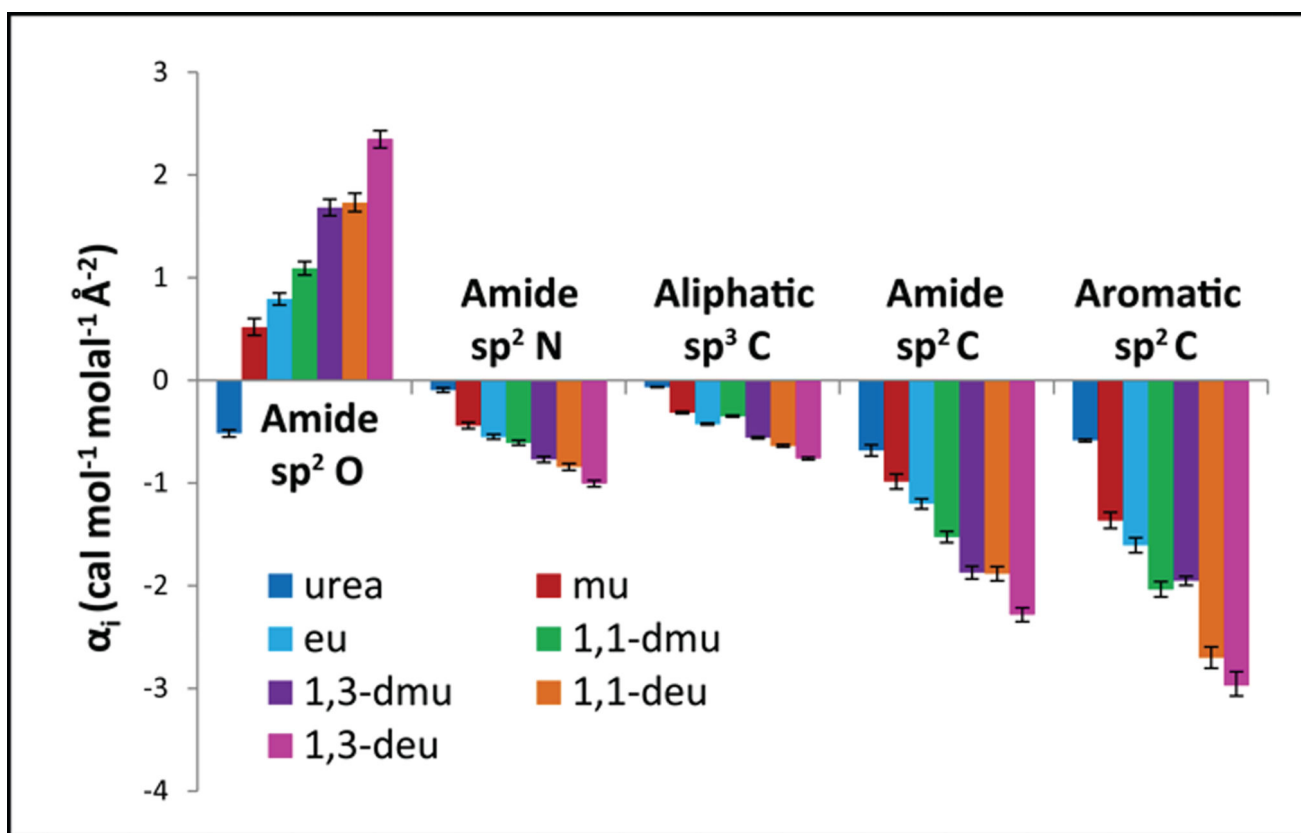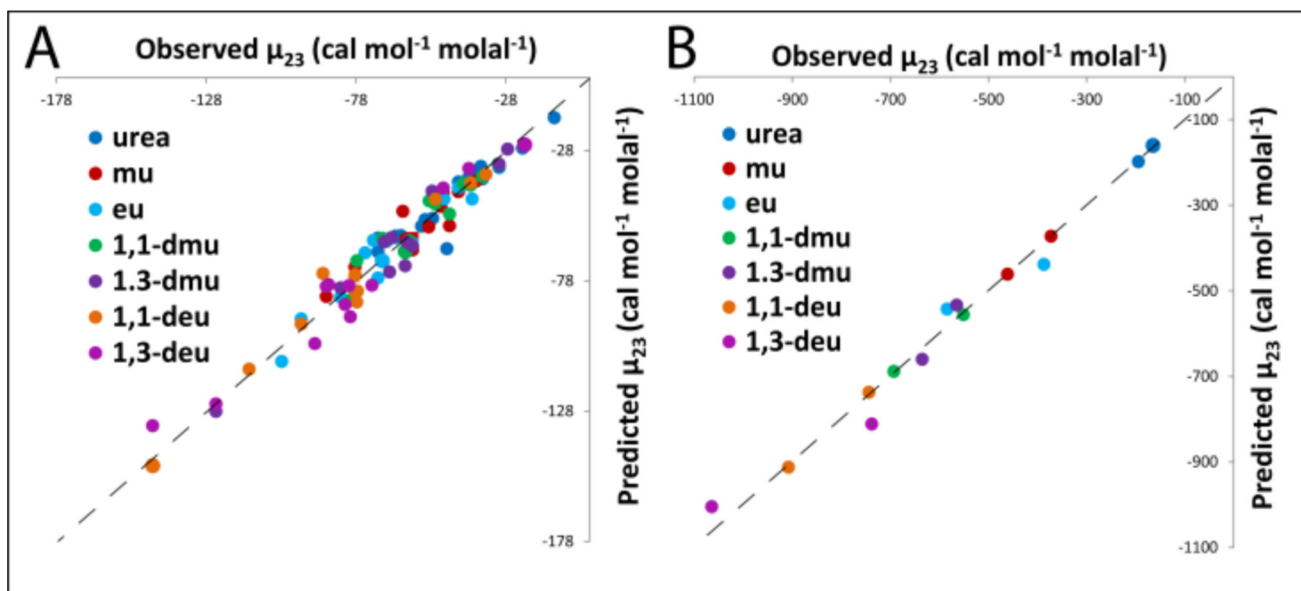
Figure 2.
Amide-aromatic Interactions: Interactions of Alkyl Ureas with Naphthalene and Anthracene by Solubility Assays at 25°C. The negative logarithm of the solubility of the aromatic hydrocarbon (m_2^{ss}) , normalized by its extrapolated molal solubility in the absence of solute ($m_{2,0}^{ss}$) is plotted versus alkyl urea molality m_3 and fit to a linear or quadratic equation as shown to obtain μ_{23} from the initial slope (see Equation 2). Solubility data quantifying interactions of naphthalene with four additional amides are shown in Supplemental Figure S5.

Figure 3.
Visual Dissection of Amide-amide Interactions from Pairwise Comparisons. Panel A shows ASA differences (DASA) between selected pairs of amides where the ASA difference (ΔASA) is mostly from one type of unified atom (amide sp^2O, amide sp^2N, or aliphatic sp^3C). See also Supplemental Figure S6. Panels B, C, and D show the corresponding differences in μ_{23} values for interactions with the urea series, ranked in order of increasing amount of aliphatic C and decreasing amount of amide N ASA. See also Supplemental Figure S7. These Δμ_{23} values show the trends in strength of interaction of this amide series with B) amide sp^2N ASA, C) amide sp^2O ASA, and D) aliphatic sp^3C ASA, which are supported by the quantitative analysis summarized in Figure 4 below.

Figure 4.
Trends in Strengths of Interaction of the Alkyl Urea Series with Individual Amide and Hydrocarbon Functional Groups. Bar graphs compare interaction potentials (α-values; Table 1)) quantifying interactions of alkyl ureas with a unit area of amide sp^2O, amide sp^2N, aliphatic sp^3C, and with amide sp^2C and aromatic sp^2C at 23 °C. Favorable interactions have negative α-values while unfavorable interactions have positive α-values.

Figure 5.
Comparison of Predicted and Observed μ_{23} Values. A) Interactions of Urea and Alkyl Ureas with other Ureas and Amides at 23°C. B) Interactions of Urea and Alkyl Ureas with Naphthalene and Anthracene at 25°C. Predictions of μ_{23} use α-values for amide sp^{2}O, amide sp^{2}N and amide sp^{2}C, aliphatic sp3C, and aromatic sp^{2}C in Table 1. Observed μ_{23} values are from Tables S2 and S3. Dashed lines represent equality of predicted and observed values.

Table 1

Strengths of Interaction (α-values) of Alkyl Ureas with Amide and Hydrocarbon Unified Atoms[a]

| Unified atom | α (cal mol⁻¹ molal⁻¹ Å⁻²)[b] | | | | |
Solute	Amide O	Amide N	Aliphatic sp³C	Amide sp²C	Aromatic sp²C
urea	−0.52 ±0.04	−0.09 ±0.02	−0.07 ±0.01	−0.69 ±0.06	−0.59 ±0.01
mu	0.52 ±0.08	−0.44 ±0.03	−0.31 ±0.01	−0.99 ±0.07	−1.37 ±0.08
eu	0.79 ±0.06	−0.55 ±0.02	−0.43 ±0.01	−1.2 ±0.05	−1.62 ±0.07
1,1-dmu	1.09 ±0.07	−0.61 ±0.03	−0.35 ±0.01	−1.53 ±0.05	−2.05 ±0.08
1,3-dmu	1.68 ±0.09	−0.77 ±0.03	−0.56 ±0.01	−1.87 ±0.06	−1.97 ±0.04
1,1-deu	1.73 ±0.09	−0.84 ±0.03	−0.64 ±0.01	−1.88 ±0.07	−2.72 ±0.1
1,3-deu	2.35 ±0.09	−1.01 ±0.04	−0.76 ±0.01	−2.28 ±0.07	−2.97 ±0.13
naphthalene	−4.44 ± 0.8	0.85 ±0.29	−2.71 ±0.07	−2.88 ±0.83	---[c]

[a]Solute abbreviations: mu: methylurea; eu: ethylurea; dmu: dimethylurea; deu: diethylurea;

[b]
α-values are obtained by fitting experimental μ23 values (Tables S2 and S3) for amide-amide interactions to Equations 3, and for amide-aromatic interactions to Equation 4. Propagated uncertainties in α-values are calculated as described previously.[19]

[c]Not determined.