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] *

Program in Biophysics[1] and Departments of Biochemistry[2] and Chemistry[3]
University of Wisconsin – Madison, Madison WI 53706
[#]Current address: Department of Immunology, Washington University, St. Louis, MO.

Materials and Methods:

Chemicals

Methyl urea (mu), ethyl urea (eu), 1,1-diethlyurea (1,1-deu), malonamide (mad; purity >97% for all), 1,1-dimethylurea (1,1-dmu), 1,3-dimethylurea (1,3-dmu), N-methylacetamide (nma), N-methylformamide (mfad; all >99%) and naphthalene (>99%) were from Aldrich. Urea and acetamide (aa; both >99%), formamide (fad; >99.5%) and anthracene (>99%) were from Sigma. 1,3-diethylurea (1,3-deu; >98%) was from TCI. Propionamide (ppa; >98%) was from Alfa Aesar. Acetyl-L-ala-methylamide (aama, >99%) was from Bachem. All the samples were obtained in anhydrous form and were dissolved in deionized water purified with a Barnstead E-pure system (Thermo-Fischer Scientific). Structures of these compounds with color-coded surface types are shown in Figure S14.

Vapor Pressure Osmometry (VPO) Method to Quantify Interactions of Amide Compounds

Sample Preparation: For each pair of amides investigated, four or more series of ten three-component solutions were independently prepared in which the molality of one amide (designated component 2, molality m_2) was held constant at 0.35 molal or 0.6 molal and the molality of the other amide (designated component 3, molality m_3) was varied from 0 to 0.95m. In addition, four or more series of ten to fifteen two component solutions of each amide were independently prepared, spanning the concentration range 0.05 molal to 1.5 molal. Thirty KCl standard solutions in the concentration range of 0.02-1.2 molal were also prepared and used as bracketing standards.

Sample measurement: All samples were measured on a Wescor Vapro 5600 Vapor Pressure Osmometer (VPO) using published methods.[1] Osmometers were calibrated using 0.10, 0.29, and 1.00 Osm Wescor NaCl calibration standards from Wescor before reading each series of samples. Osmolalities of all samples at room temperature (23 ± 1℃) were determined by

averaging at least three osmometer readings. Bracketing KCl standards were read with each sample and used to correct its osmolality using literature isopiestic distillation (ID) data for KCl.[2] Concentrations were chosen so solution osmolalities did not exceed 1.5 Osm. At higher concentrations, osmolalities of samples and KCl bracketing standards were less reproducible.

Two-Component Solutions: Determination of the Excess Chemical Potential Derivative $d\mu_2^{ex}/dm_2 = \mu_{22}^{ex}$ from the Excess Osmolality $\Delta Osm(m_2)$

The osmolality of a two-component nonelectrolyte solution is represented as a power series in solute molality m_2^{3-4}). For situations considered here, the power series can be truncated at or before the m_2^3 term as shown in Eq. S1.

$$Osm(m_2) = \beta_2 m_2 + \beta_{22} m_2^2 + \beta_{222} m_2^3 \ldots \qquad Eq.\,S1$$

For nonelectrolytes, $\beta_2 = 1$. From the two-component Gibbs-Duhem equation, at constant T and P,

$$\frac{dOsm}{dm_2} = 1 + m_2 \frac{dln\gamma_2}{dm_2} = 1 + m_2 \frac{\mu_{22}^{ex}}{RT} \qquad Eq.\,S2$$

In Eq S2, $\mu_2^{ex} = RTln\gamma_2$ where γ_2 is the molal-scale solute activity coefficient relative to the ideal dilute solution, and $\mu_{22}^{ex} = \left(\frac{\partial \mu_2^{ex}}{\partial m_2}\right)_{P,T}$. From Eqs. S1-2,

$$\frac{dOsm}{dm_2} = 1 + 2\beta_{22} m_2 + 3\beta_{222} m_2^2 + \cdots = 1 + m_2 \frac{\mu_{22}^{ex}}{RT} \qquad Eq.\,S3$$

and $\qquad \frac{\mu_{22}^{ex}}{RT} \equiv \frac{\frac{dOsm}{dm_2} - 1}{m_2} = 2\beta_{22} + 3\beta_{222} m_2 + \cdots \qquad Eq.\,S4$

In sufficiently dilute solutions ($m_2 < 1$ molal) where the contribution of the m_2^3 term in Eq. S1 is typically insignificant,

$$\frac{\mu_{22}^{ex}}{RT} \cong 2\beta_{22} \qquad\qquad \text{Eq. S5}$$

The excess osmolality of a two component solution is given by

$$\Delta Osm(m_2) = Osm(m_2) - Osm^{ideal}(m_2) = Osm(m_2) - m_2 = \beta_{22}m_2^2 + \beta_{222}m_2^3 + \cdots \qquad \text{Eq. S6}$$

For sufficiently small m_2

$$\Delta Osm(m_2) \cong \beta_{22}m_2^2 = \frac{\mu_{22}^{ex}}{2RT}m_2^2 \qquad\qquad \text{Eq. S7}$$

The excess osmolality of the two-component solution is interpreted as the product of the probability of a pairwise solute-solute interaction (proportional to m_2^2) and the strength of that interaction relative to interactions with water ($\mu_{22}^{ex}/2$), relative to thermal energy (RT). Experimentally-determined excess osmolalities $\Delta Osm(m_2) = Osm(m_2) - m_2$ of two component solutions of urea and alkyl ureas are plotted in Figure S4 as a function of m_2^2 for the concentration range 0.1 molal $< m_2 <$ 0.8 molal. Linear fits with zero intercept yield self-interaction coefficients $\mu_{22}^{ex}/2$, values of which are listed together with μ_{23} values in Table S2.

Linear fitting of two-component osmolality data to Equation S7 to obtain $\mu_{22}^{ex}/2$ is analogous to the approach used to obtain μ_{23} from three- and two-component osmolality data (see Equation 1). Alternatively, values of β_{22} and $\mu_{22}^{ex}/2$ are obtained from nonlinear fitting of two-component osmolality data (0.1-1.5 molal) to Eq. S1 (see Figure S1). These two component VPO data for urea and alkyl ureas are consistent with published results from isopiestic distillation studies where available, as demonstrated by the comparison of osmotic coefficients $\varphi = Osm(m_2)/m_2$ in Supplemental Table S13. Table S14 shows that values of $RT\beta_{22}(\cong \mu_{22}^{ex}/2)$ agree with each other and with predicted values obtained from α-values (Table 1) within $\pm$ 30% on average.

In general, two- component osmolality data for alkyl ureas in the entire concentration range examined (<1.5 molal) were fit to Eq. S1 setting $\beta_2 = 1$ and floating β_{22} and β_{222} (Figure S1A). Fitted values of β_{22} and β_{222} are listed in Table S1. Table S13 shows that osmotic coefficients of ureas predicted at different concentrations using these β_{22} and β_{222} values agree well with experimentally determined osmotic coefficients from isopiestic distillation.[5-6]

To obtain good fits to two component data for N-methylacetamide, formamide and N-methylformamide (all liquids) and propionamide (Figure S1B), it was necessary to float β_2 as well as β_{22}, though the cubic term was unnecessary ($\beta_{222} = 0$). Values of β_2 and β_{22} for these four solutes are shown in the lower part of Table S1. Possibly this is the result of a trace impurity in these samples with sufficient volatility to interfere with the VPO assay. This artifact appears to largely subtract out in calculating ΔOsm, therefore having no significant effect on α-values.

Three-Component Solutions: Obtaining the Chemical Potential Derivative μ_{23} from the Excess Three-Component Osmolality $\Delta Osm(m_2, m_3)$

In two component (2c) solutions of water (component 1) and one solute (component 2 or 3), the Gibbs-Duhem equations at constant T and P are

$$dOsm\,(m_2) \equiv -m_1^{\bullet}d\ln a_1^{(2c)} = (RT)^{-1}m_2 d\mu_2^{(2c)} \qquad\qquad Eq.\,S8$$

$$dOsm\,(m_3) \equiv -m_1^{\bullet}d\ln a_1^{(2c)} = (RT)^{-1}m_3 d\mu_3^{(2c)} \qquad\qquad Eq.\,S9$$

while for a three-component (3c) solution at constant T and P

$$dOsm\,(m_2, m_3) \equiv -m_1^{\bullet}d\ln a_1^{(3c)} = (RT)^{-1}m_2 d\mu_2^{(3c)} + (RT)^{-1}m_3 d\mu_3^{(3c)} \qquad\qquad Eq.\,S10$$

The osmolality difference ΔOsm(m_2,m_3) of interest here is defined as

$$\Delta Osm(m_2, m_3) = Osm(m_2, m_3) - Osm(m_2) - Osm(m_3) \qquad\qquad Eq.\,1$$

Subtracting Equations S8 and S9 from Eq.S10 and rearranging terms we obtain:

$$d\Delta Osm\ (m_2, m_3) = (RT)^{-1}m_2\ \left(d\mu_2^{(3c)} - d\mu_2^{(2c)}\right) + (RT)^{-1}m_3\ \left(d\mu_3^{(3c)} - d\mu_3^{(2c)}\right) \qquad Eq.\,S11$$

Representing the differences in chemical potential in Equation S11 by the first terms in Taylor expansions, one obtains

$$d\ \Delta Osm(m_2, m_3) = (RT)^{-1}\left(m_2\ \left(\frac{\partial\mu_2}{\partial m_3}\right)_{P,T,m_2;\ m_3=0} dm_3 + m_3\ \left(\frac{\partial\mu_3}{\partial m_2}\right)_{P,T,m_3;\ m_2=0} dm_2\right)$$

$$\approx (RT)^{-1}\ \mu_{23}\ d(m_2 m_3) \qquad\qquad Eq.\,S12$$

In Equation S12, the derivative $\left(\frac{\partial\mu_2}{\partial m_3}\right)_{P,T,m_2;\ m_3=0}$ is taken at constant m_2 and then evaluated in the limiting situation where m_3 approaches zero, and the derivative $\left(\frac{\partial\mu_3}{\partial m_2}\right)_{P,T,m_3;\ m_2=0}$ is taken at constant m_3 and then evaluated in the limiting situation where m_2 approaches zero. The approximation in Eq. S12 is that μ_{23} is sufficiently concentration-independent in the range of our assays (0 – 0.8 m in each solute) so that

$$\left(\frac{\partial\mu_2}{\partial m_3}\right)_{P,T,m_2;\ m_3=0} \cong \left(\frac{\partial\mu_3}{\partial m_2}\right)_{P,T,m_3;\ m_2=0} \cong \mu_{32} = \mu_{23} \qquad Eq.\,S13$$

Equations analogous to the integrated form of Equation S12 have been obtained previously[4, 7-8] using other assumptions. The derivation of Equation S12 presented here is straightforward and provides direct insight into the approximations (Equations S12-S13) used to obtain μ_{23}.

Cannon et al[4] analyzed ID and solubility data of urea-amino acid 3-component solutions as functions of both solute molalities using Equation 1, and concluded that μ_{23} to a good approximation is independent of solute concentration up to approximately 1 molal.

Solubility Method to Quantify Interactions of Amides with Aromatic Compounds

Solubility assays at 25°C are used to determine interactions of urea, alkyl ureas and other amides with sparingly-soluble naphthalene and/or anthracene in aqueous solution. Solubility data quantifying interactions of other solutes with aromatics, nucleobases, and other sparingly-soluble model compounds have been reported previously.[9-11] For more soluble model compounds, interactions with a solute are best quantified by an osmometric assay.[1, 12-16]

At least two series of 12 solutions of urea or the alkyl urea or amide at concentrations in the range 0 – 4 molal (limited to 0 – 1.5 molal for 1,1- and 1,3-diethyl urea) are prepared gravimetrically in 10 mM K_2HPO_4 (adjusted to pH = 7.3 with HCl). in 15 ml conical centrifuge tubes. An excess of naphthalene or anthracene is added and solutions are heated at 40 ℃ for 1-2 hours before placing them in a thermostatted shaking water bath at 25°C for 10 – 14 days. This is sufficient time to obtain solubility equilibrium, as shown previously for naphthalene[11] and checked for naphthalene and anthracene by the spectrophotometric assay described below.

Supernatants were clarified by centrifugation and absorbance scans (350 -200 nm) vs buffer were performed with an Agilent Cary 300 double beam spectrophotometer against a reference cell containing buffer. A buffer vs. buffer scan with the same cells was subtracted to correct for cell to cell differences. Anthracene solutions were read without dilution; naphthalene solutions were diluted 1:4 gravimetrically with buffer before reading.

Calculations of Water-Accessible Surface Area (ASA)

Water-accessible surface areas (ASA) of the different types of unified atoms present on the alkyl ureas, other amides, and aromatic compounds investigated in this research were calculated as in previous research [1, 12-14, 16-21] using the Surface Racer program[22] with a conventional set of van der Waals radii[20] and a 1.4 Å probe radius for water. For each compound, ASA values for amide sp^{2}C, aliphatic sp^{3}C, amide sp^2 N, amide sp^2 O or aromatic sp^2 C are

reported in Table S4. Structures of all compounds investigated were obtained from Cactus as described previously.[14] PubChem structures are also available for all compounds investigated, and BMRB structures are available for some. In a subsequent SI section and Table SI 15 we compare ASA values obtained from these three sources.

Alternative programs that calculate and dissect ASA from structural information include GETAREA[23] and VMD.[24] Surface Racer was used because it was most suitable for the calculations of this type. Other alternatives, like GROMACS,[25] are primarily designed for other applications and are less user-accessible. Surface Racer and GETAREA, both based on analytical theory, give identical results when used to calculate the same structure with the same choices of atomic radii. VMD, which uses a numerical Monte Carlo algorithm, predicts atomic ASA which deviate by small to moderate extents (±0.7% to ±14%) from Surface Racer and GETAREA for large ASA values. For atomic ASA values less than 4 $Å^2$ (as measured by Surface Racer and GETAREA), the absolute VMD - Surface Racer or GETAREA difference is less than 0.8 $Å^2$. VMD predictions of total ASA of amide molecules are in better agreement with Surface Racer and GETAREA, exhibiting deviations of ±0.2 to ±3.6%.

To test if the differences in ASA values obtained with VMD as compared to Surface Racer or GETAREA affect the analysis reported in Figure 4 and Table 3, we recalculated α-values for urea and alkylureas using ASA values from VMD (Table S21). As shown in Table S22, no significant differences in α-values are obtained. We consider the ASA predictions from Surface Racer and GETAREA to be more exact and robust, especially for small molecules like those which are the subject of this study, and use these ASA in this research.

Analysis of μ23 Values Using a Dissection Based on Numbers of Atoms instead of ASA

By analogy with the ASA-based analysis of μ23 values for interactions of an alkyl urea or urea with the unified atoms of amide compounds (Equation 3) and aromatic compounds (Equation 4), the number-based analysis represents μ23 for these interactions as

$$\mu_{23} = \alpha_{n,amide\ sp^2O}\, n_{amide\ sp^2O} + \alpha_{n,amide\ sp^2N} \sum_{j=0}^{2} w_{sp^2N,j}\, n_{amide\ sp^2N,j}$$

$$+ \alpha_{n,amide\ sp^2C} \sum_{i=0}^{1} w_{sp^2C,i}\, n_{amide\ sp^2C,i} + \alpha_{n,sp^3C} \sum_{k=1}^{3} w_{sp^3C,k}\, n_{sp^3C,k} \qquad Eq.\,S14$$

and

$$\mu_{23} = \alpha_{n,aromatic\ sp^2C}\, n_{aromatic\ sp^2C} \qquad\qquad Eq.\,S15$$

In Equations S14 and S15, α_n-values are strengths of interaction of the alkyl urea (or urea) component 3 with the specified type of atom or unified atom (O, N, C) on the amide or aromatic compound. These α_n-values are per atom, expressed in cal mol^{-1} molal^{-1}, while α-values (in cal mol^{-1} molal^{-1} A^{-2}) are per unit area of that type of atom. There is only one type of amide sp^{2}O and one type of aromatic sp^{2}C unified atom in our data set, but multiple types of amide sp^{2}N, amide sp^{2}C, and aliphatic sp^{3}C unified atoms. Hence Equation S14 includes sums over the various numbers of H atoms integrated in the C and N unified atoms; $n_{x,y}$ is the number of atoms in the structure of type x with y H atoms and $w_{x,y}$ is a weighting factor for this unified atom relative to that with the maximum number of H. For sp^{3}C, the reference unified atom is–CH$_3$ ($w_{sp^3C,3}$ =1); for amide sp^{2}C, the reference is the unified atom with one H ($w_{amide\ sp^2C,1}$ =1); for sp^{2}N, the reference unified atom is –NH$_2$ ($w_{sp^2N,2}$ =1).

Values of μ23 for the interaction of each alkyl urea with naphthalene and anthracene are interpreted by Equation S15 to obtain the parameter $\alpha_{n,aromatic\ sp^2C}$. These number-based α_n-

values differ from the corresponding ASA-based α-values (Table 1) by the average ASA per CH_2 united atom of naphthalene and anthracene (25 Å^2).

Global fitting of the set of μ_{23} values for interactions of seven ureas with 12 amide compounds to Equation S14 was used to obtain three or four α_n-values for each urea and zero to five global weighting factors for the set of 7 ureas. (Figure S12; Table S9) Early research[26] used a number-based analysis analogous to Equation S14 to interpret primarily two-component thermodynamic data for polyols, amides and other solutes in water and estimate strengths of interactions between individual atom types or groups. In their analysis, the amide group was not subdivided, and CH, CH_2 and CH_3 sp^{3}C groups were assigned relative weights of 0.5, 1 and 1.5. Here global fitting is used to obtain global relative weights for the different types of C and N unified atoms to improve the quality of the number-based fits.

If no global weighting factors are introduced in fitting Equation S14 to the set of alkyl urea-amide compound μ_{23} values (i.e. all $w_i = 1$), the fit yields three α_n-values for the interactions of each urea with amide sp^{2}N, aliphatic sp^{3}C, and combined amide sp^{2}C-amide sp^{2}O. Individual α_n-values for amide sp^{2}O and amide sp^{2}C cannot be separated because the two atoms occur together in a 1:1 ratio in all amide compounds. Figure S12 panel A compares predicted and observed μ_{23} values for this number-based fit with 3 parameters for each urea. Agreement is relatively poor by comparison to the ASA-based fit (Figure 5A), which has 4 parameters for each urea. The quality of fit (the sum of squares of residuals) is about 10 times larger for this number-based fit than for the ASA-based fit. Table S9 lists α_n-values for this and other fits (below) for comparison with ASA-based α-values (Table 1).

Panel B floats the relative weight of amide sp^{2}C with no integrated H atoms. This is equivalent to treating the α_n-values for amide sp^{2}O and amide sp^{2}C as separate fitting parameters. Panel C fit floats the relative weights of the various unified sp^{2}N atoms (i.e. N, NH relative to NH_2)

and sp^{3}C (i.e. CH, CH$_2$ relative to CH$_3$) atoms but does not distinguish between types of amide sp^{2}C atoms and hence (like the panel A fit) cannot separate contributions of amide O and amide C. The fit in Panel D floats all five relative weights.

Tests Demonstrating that Cactus ASA (Table S4) and α-Values (Table 1) are Robust

a) Comparison of Cactus ASA Information with PubChem and BMRB

As described in Methods, structures of all solutes investigated here were predicted from the NIH Cactus website (https://cactus.nci.nih.gov/translate/) and used to obtain ASA values for the different types of C, N and O unified atoms. The Cactus prediction algorithm[27] has been tuned using the Cambridge crystal structure data base.[28]

The alternative prediction method is PubChem,[29] which is based on quantum calculations with the MMFF94s force field.[30] For 20 amide and aromatic compounds investigated here, Table S15 compares ASA values of unified atoms calculated from PubChem structures with Cactus ASA values and with those calculated from Biological Magnetic Resonance Bank (BMRB) structures.[31] For 14 of these 20 solutes, PubChem predictions agree with Cactus predictions within $\pm 10\%$ (or ± 1 Å^2 for amide C), which is comparable to the experimental % uncertainty. Four of these 14 solutes also have a BMRB structure (BMRB IDs given in Table S15). Agreement of ASA values from the BMRB structure and Cactus and PubChem structures is also within $\pm 10\%$ (or ± 1 Å^2 for amide C).

For three amides (1,3-diethylurea, 1,3-dimethylurea and malonamide) PubChem ASA predictions deviate much more significantly from Cactus predictions, especially for amide O (17%-65%). For three other amides (aama, nma, tmu), deviations of PubChem from Cactus are in the 10-15% range for amide O or amide N. Should PubChem or Cactus ASA values be selected for this research? For 1,3-dimethyl urea, where a BMRB structure is available, BMRB and Cactus-predicted ASA values agree within $\pm 1\%$ (or ± 0.2 Å for amide C), but the PubChem amide O

and amide N ASA values differ from the corresponding BMRB values by 16% and 12%, respectively. For 1,3-diethylurea, the Cactus amide O ASA value is much more consistent with the trend observed for the other alkyl ureas than is the case for the corresponding PubChem amide O ASA value. For these reasons, Cactus ASA values were used in the analysis presented in the main text. The analysis in the section below further supports the use of Cactus ASA in the present context.

b) Comparison of α-Values and Uncertainties Predicted Using ASA Values from PubChem and Cactus Structures

To test how differences in ASA values obtained from different structures affect α-values, the α-value analysis in the main text (based on Cactus structures) was repeated using PubChem structures. If malonamide and 1,3-diethylurea are included in the data set analyzed using PubChem ASA values, the magnitude of the amide O α-values for 1,3-diethylurea is reduced by 26% and magnitudes of amide O α-values for the other five alkyl ureas are reduced by 38-52%. These discrepancies are caused by the large discrepancy in amide O ASA for 1,3-diethylurea (65%) and the smaller but significant discrepancies for amide O ASA of malonamide (21%) and 1,3-dimethyl urea (17%). Table S16 shows that if malonamide and 1,3-diethylurea are removed from the data set, the analysis based on Cactus ASA values is relatively unaffected, while the analysis based on PubChem ASA values is improved. Also, PubChem α-values obtained from analysis of this reduced data set predict μ_{23} values for malonamide and 1,3-diethylurea which are in better agreement with the experimental values (Table S17).

c) Tests for Possible Inter-relationships of α-Values and their Consequences

In a linear regression analysis of the type performed here to determine α-values from μ_{23} values using ASA information and Eq. 3, correlations between inputs (ASA values for different

types of unified atoms on the molecules of the data set) can lead to ambiguity and/or large uncertainties in outputs (α-values). An extreme example of this is present in some versions of the alternative analysis of the experimental μ_{23} values in SI, using Eq. S14 and the numbers of atoms of each type in each molecule studied instead of the ASA contribution from each atom type. As discussed in the text, in this alternative analysis amide C and amide O, always present in a 1:1 ratio, are completely correlated in scenarios A and C in Table S9. This correlation is resolved in the more highly parameterized fits (B, D) where the amide CH unified atom is assigned a different α-value than amide C. The ASA-based analysis is better in this regard, but constraints imposed by amide chemical bonding patterns and experimental considerations (restriction to nonvolatile amides) result in some correlation between the different ASA types for the majority of the amides studied.

Two approaches were used to test whether correlations in ASA contributions in the alkyl urea and amide data sets have any significant effect (in comparison to the experimental uncertainty) on the determination of α-values. One approach is to calculate a "condition number"[32] which characterizes the extent of correlations, and interpret this using ranges given by the authors. Condition numbers for urea and alkylureas are found to be approximately 20, indicating modest correlations (<10 indicates weak correlations; >30 indicates strong correlations).

To clarify the meaning of this statistical result and provide another test of whether the analysis in the main text is robust, we examined whether significantly different combinations of α-values give significantly worse fits to the μ_{23} data. Table S18 shows the situation for methyl urea, the alkyl urea with the largest uncertainties in α-values and therefore the largest possibility of having different combinations of α-values giving equally significant fits. In Table S18, one α-value is fixed at each of the listed range of different values bracketing the best fit value in the center row of the table (bold, italic). The values immediately next to the best fit value (italic) represent the

range given by the uncertainty in this α-value in Table 1 of the main text. In each case, use of either of these α-values causes compensating changes in other α-values that are also within their uncertainties, and does not significantly degrade the fit, as judged by the sum of squares of residuals in the last column. But changing the fixed α-value by more than its uncertainty degrades the fit in each case. We therefore conclude that the uncertainties in α-values listed are appropriate, and that these uncertainties, based on propagation of the experimental error, are sufficient to include any effects of correlations between the fitting parameters.

d) Tests of Effect of Size of Dataset on α-Values

As additional tests of robustness of the α-values in Table 1, effects of adding to or removing compounds from the data set of μ_{23}-values were examined. Table S19 shows that α-values for interactions of urea with amide and hydrocarbon atoms agree well with those obtained previously[14] (Table S11) from analysis of a larger data base of atom types containing some different amide compounds than those studied here. Inclusion of seven amides from that data base in the set analyzed here did not significantly affect any of the urea α-values (Table S19). We also extended the data sets for all the alkyl ureas (and urea) by including $\mu_{22}{}^{ex}$ values (Table S2). Inclusion of $\mu_{22}{}^{ex}$ does not affect any α-values significantly (Table S20). As noted previously in SI (Table S16), deletion of 1,3-diethylurea and malonamide from the data set also does not significantly affect the α-values reported in Table 1.

Table S1. Fitting Coefficients Describing Concentration Dependence of Amide Osmolality[a]

Solute[b]	β_2	β_{22}	β_{222}
urea	1	-0.05 ± 0.003	0 ± 0
mu	1	-0.089 ± 0.007	0.034 ± 0.008
eu	1	-0.16 ± 0.046	0.026 ± 0
1,1-dmu	1	-0.047 ± 0.014	0.001 ± 0
1,3-dmu	1	-0.129 ± 0.006	0.028 ± 0.011
1,1-deu	1	-0.251 ± 0.011	0.084 ± 0.013
1,3-deu	1	-0.275 ± 0.01	0.076 ± 0.01
mad	1	-0.139 ± 0.014	0.065 ± 0.014
aama	1	0.01 ± 0.001	0 ± 0
Solute[c]	β_2	β_{22}	
ppa	0.89 ± 0.006	-0.042 ± 0.008	
nma	0.848 ± 0.007	-0.069 ± 0.008	
fad	0.805 ± 0.005	-0.039 ± 0.004	
mfad	0.701 ± 0.004	-0.068 ± 0.005	

[a]Solute abbreviations: mu: methyl urea; eu: ethyl urea; dmu: dimethyl urea; deu: diethyl urea; mad: malonamide; aama: N-Acetylalanine N-methylamide; ppa: propionamide; nma: N-methyl acetamide; fad: formamide; mfad: N-methyl formamide.

[b]Osmolalities in the range 0 – 1.5 molal were fit to the equation $Osm(m_2) = m_2 + \beta_{22}m_2^2 + \beta_{222}m_2^3$ with β_{22} and β_{222} floated. See SI text Figure S1A.

[c]Osmolalities in the range 0 – 1.5 molal were fit to the equation $Osm(m_2) = \beta_2 m_2 + \beta_{22}m_2^2$ with β_2 and β_{22} floated. See SI text and Figure S1B.

Table S2. Quantifying Amide-amide Interactions by Osmometry: Values of μ_{23} (or $\mu_{22}^{ex}/2$) in cal mol^{-1} molal^{-1} at 23 °C[a]

	urea		Methylurea		Ethylurea		1,1-dimethylurea	
	Observed[b]	Predicted[c]	Observed[b]	Predicted[c]	Observed[b]	Predicted[c]	Observed[b]	Predicted[c]
urea	(-42.4 ± 4.3)[d]	-41.9 ± 3.5	-37.8 ± 1.9	-39.6 ± 5.9	-43.8 ± 2.3	-42.3 ± 4.2	-35.7 ± 2.1	-37.9 ± 4.5
mu	-37.8 ± 1.9	-38.2 ± 2.5	(-38.6 ± 3.9)[d]	-52.9 ± 4.4	-59.4 ± 2.3	-63.2 ± 3.2	-46.8 ± 2.1	-52.4 ± 3.4
eu	-43.8 ± 2.3	-40 ± 2.4	-59.4 ± 2.3	-61.8 ± 4.4	(-75.6 ± 7.6)[d]	-75.6 ± 3.2	-69.2 ± 2.7	-61.7 ± 3.4
1,1-dmu	-35.7 ± 2.1	-38.8 ± 2	-46.8 ± 2.1	-56.9 ± 4	-69.2 ± 2.7	-70.3 ± 2.9	(-28.8 ± 2.9)[d]	-52.8 ± 3.2
1,3-dmu	-30.2 ± 1.1	-34.5 ± 1.6	-59.1 ± 2.8	-66.1 ± 3.4	-83.1 ± 2.9	-84.1 ± 2.6	-61.6 ± 1.9	-66.7 ± 2.7
1,1-deu	-39.7 ± 1.2	-39.3 ± 1.9	-78.3 ± 1.9	-72.8 ± 4	-96.2 ± 2.7	-92.6 ± 3	-77.7 ± 2.7	-70.3 ± 3.3
1,3-deu	-40.3 ± 2.5	-38.2 ± 1.6	-87.8 ± 3.3	-84 ± 3.7	-103 ± 1.9	-109 ± 2.9	-79.8 ± 2.8	-85.4 ± 3.1

	1,3-dimethylurea		1,1-diethylurea		1,3-diethylurea	
	Observed[b]	Predicted[c]	Observed[b]	Predicted[c]	Observed[b]	Predicted[c]
urea	-30.2 ± 1.1	-33.2 ± 5.8	-39.7 ± 1.2	-40.4 ± 6.3	-40.3 ± 2.5	-34.8 ± 6.4
mu	-59.1 ± 2.8	-64.6 ± 4.4	-78.3 ± 1.9	-76.2 ± 4.8	-87.8 ± 3.3	-80.1 ± 4.9
eu	-83.1 ± 2.9	-80.6 ± 4.5	-96.2 ± 2.7	-94.6 ± 4.9	-91.7 ± 2.7	-102 ± 5
1,1-dmu	-61.6 ± 1.9	-72.2 ± 4.2	-77.7 ± 2.7	-86.1 ± 4.6	-79.8 ± 2.8	-91.8 ± 4.6
1,3-dmu	(-66.6 ± 6.7)[d]	-96 ± 3.6	-114 ± 3.6	-112 ± 4	-125 ± 3.5	-125 ± 4.1
1,1-deu	-114 ± 3.6	-102 ± 4.4	(-115 ± 12)[d]	-121 ± 4.8	-146 ± 4.5	-134 ± 4.9
1,3-deu	-125 ± 3.5	-128 ± 4.2	-146 ± 4.5	-149 ± 4.6	(-133 ± 13)[d]	-169 ± 4.7

	malonamide		propionamide		N-methylacetamide		aama	
	Observed[b]	Predicted[c]	Observed[b]	Predicted[c]	Observed[b]	Predicted[c]	Observed[b]	Predicted[c]
urea	-54.9 ± 2.5	-54.4 ± 3.8	-36.8 ± 2	-35.8 ± 2	-36.3 ± 1.7	-34 ± 1.5	-52.4 ± 5.9	-54 ± 2.5
mu	-43.8 ± 1.4	-43.8 ± 6.8	-62.3 ± 1.7	-51.3 ± 3.9	-49.7 ± 1.8	-49.4 ± 3.5	-61.3 ± 3.3	-61.9 ± 5.9
eu	-48.5 ± 2.4	-46.6 ± 4.9	-67.6 ± 1.7	-62.8 ± 2.8	-72.2 ± 1.5	-62.5 ± 2.7	-70.6 ± 1.7	-76.9 ± 4.4
1,1-dmu	-30.6 ± 1.8	-33.2 ± 5.3	-53.5 ± 2.3	-47.3 ± 3.1	-42.2 ± 1.7	-40.7 ± 2.9	-39.8 ± 2.3	-41.3 ± 4.7
1,3-dmu	-27.3 ± 2.4	-27.4 ± 6.9	-68.5 ± 1.9	-63 ± 4	-66.2 ± 2.8	-61.4 ± 3.8	-60.9 ± 2.6	-63.2 ± 6.3
1,1-deu	-34.6 ± 1.3	-37.1 ± 7.5	-78.2 ± 1.5	-75.6 ± 4.4	-89 ± 1.9	-75.1 ± 4.2	-78.3 ± 3.1	-82.1 ± 7
1,3-deu	-21.7 ± 1.2	-25.8 ± 7.6	-72.7 ± 2.5	-79.7 ± 4.5	-87.1 ± 2.2	-79.4 ± 4.3	-80.3 ± 2.9	-79.9 ± 7

[a]Abbreviations defined in Table S1. [b]Experimental values of μ_{23} and $\mu_{22}^{ex}/2$ are obtained from VPO assays at 23°C (Figure 1; S2, S3 and S4) using Equation 1. For μ_{23}, error estimates are standard deviation determined from the linear fit of the data from Figures 1 by Igor Pro. For $\mu_{22}^{ex}/2$, error estimates are the larger of 10% or the estimated standard deviation determined from the linear fit of the data from Figure S4. [c]Predicted μ_{23} values are calculated from Equation 3 using α_i values from Table 1 and ASA information from Supplemental Table S4. [d]$\mu_{22}^{ex}/2$ obtained from Figure S4 using Equation S7.

Table S3. Amide Interactions with Compounds Displaying Amide sp^2 C or Aromatic sp^2 C: Predicted and Observed μ$_{23}$-Values (cal mol^{-1} molal^{-1})

Solutes	Interactions with Formamide		Interactions with N-methylformamide		Interactions with Anthracene	
	Observed μ$_{23}$[a]	Predicted μ$_{23}$[b]	Observed μ$_{23}$[a]	Predicted μ$_{23}$[b]	Observed μ$_{23}$[c]	Predicted μ$_{23}$[d]
urea	-62.9 ± 2.2	-60.6 ± 3.5	-56.1 ± 2.9	-57 ± 2.9	-194 ± 5.9[h]	-197 ± 4.6
mu	-48.8 ± 3	-44.2 ± 5.6	-53.7 ± 1.7	-57.2 ± 4.6	-472 ± 43	-459 ± 27
eu	-39.2 ± 1.1	-46.6 ± 4	-74.9 ± 1.9	-67.3 ± 3.3	-584 ± 40	-540 ± 25
1,1-dmu	-51.6 ± 1.7	-48.4 ± 4.3	-60.2 ± 1.7	-62.5 ± 3.6	-693 ± 33	-685 ± 25
1,3-dmu	-52.6 ± 1.6	-43.5 ± 5.5	-66.7 ± 1.9	-74.5 ± 4.6	-635 ± 24	-657 ± 15
1,1-deu	-51.4 ± 1.7	-46.6 ± 6.1	-77.5 ± 2.3	-81.9 ± 5	-907 ± 56	-908 ± 35
1,3-deu	-48.9 ± 1.3	-42.4 ± 6.1	-81.6 ± 2.5	-87.1 ± 5	-1064 ± 74	-999 ± 46

				Interactions with Naphthalene				
Solutes	Observed μ$_{23}$[c]	Predicted μ$_{23}$[b]	Predicted μ$_{23}$[d]	Literature μ$_{23}$[e]	Solutes[g]	Observed μ$_{23}$[c]	Predicted μ$_{23}$[d]	Literature μ$_{23}$[e]
urea	-165 ± 5.9[h]	-161 ± 3.7	-122 ± 54	-153	aa	-380 ± 12	-402 ± 41	-275
mu	-359 ± 15	-375 ± 22	-353 ± 41	-337	mfad	-538 ± 25	-515 ± 48	-377
eu	-387 ± 15	-441 ± 20	-456 ± 40	---	dfad	-678 ± 20	-700 ± 43	-638
1,1-dmu	-551 ± 32	-560 ± 21	-544 ± 36	---	tmu	-918 ± 25	-862 ± 31	-828
1,3-dmu	-565 ± 9.4	-537 ± 12	-584 ± 30	-514	eeu	---	-496 ± 44	-471[f]
1,1-deu	-691 ± 9.1	-742 ± 28	-692 ± 35	---	ndma	---	-762 ± 33	-761[f]
1,3-deu	-738 ± 16	-817 ± 37	-789 ± 31	---				

[a]Determined by VPO at 23 °C (Figure S2 and S3). Error estimates are standard deviation determined from the linear Igor Pro fit. [b]Calculated from Equation 3 using α-values from Table 1 and ASA information from Supplemental Table S4. Uncertainties propagated from uncertainties in α-values (Table 1). [c]Determined by solubility assays at 25 °C (Figure 2 and S5). Error estimates are standard deviation determined from quadratic or linear Igor Pro fit. [d]Predicted from naphthalene α-values (Table 1). [e]Calculated from literature data,[11] correcting molar μ$_{23}$ derivative to molal as in reference.[16] [f]Literature[11] data used in naphthalene α-value analysis (Table 1). [g]Abbreviations: **aa**: acetamide; **mfad**: N-methyl formamide; **dfad**: N, N-dimethyl formamide; **tmu**: tetramethyl urea; **eeu**: ethyleneurea; **ndma**: N,N-dimethyl acetamide. [h]Determined in literature[14]

Table S4. Contributions to the ASA of Solutes and Amide Model Compounds[a]

Model compound	ASA contribution, Å^2				
	Amide sp^2 O	Amide sp^2 N	Aliphatic sp^3 C	Amide/Aromatic sp^2 C	Total
urea	47.9	130.2	0	7.2	185.3
methylurea	38.3	87.5	88.4	6.5	220.7
ethylurea	38.3	82	124.6	6.5	251.4
1,1-dimethylurea	38.3	54.5	148.7	6.2	247.7
1,3-dimethylurea	28.7	44.9	176.7	5.8	256
1,1-diethylurea	35.7	50	208.9	3.7	298.3
1,3-diethylurea	28.7	33.8	249.2	5.8	317.5
malonamide	65.7	123.2	48.5	8.5	245.9
propionamide	36.8	61.6	124.5	4.3	227.2
N-methylacetamide	35.3	19	178	3.5	235.9
aama	62.5	21.1	257.7	4.3	345.5
formamide	51.3	70.8	0	40.2	162.2
N-methylformamide	41.7	27.7	88.4	39.5	197.3
naphthalene	0	0	0	273	273
anthracene	0	0	0	334	334
acetamide	44.9	61.6	89.7	4.3	200.5
N,N-dimethylforamide	41.7	0.8	157.7	30.7	230.9
tetramethylurea	29.9	0.8	266.4	3.3	300.3
ethyleneurea	49	57.1	113.6	7	226.6
N,N-dimethylacetamide	35.3	0.8	220.1	3.6	259.7

[a]Surface areas calculated as in SI text. All structures generated by the Cactus website as described elsewhere in SI and reference[16]

Table S5. Pairs of Amides with ASA Difference (ΔASA) Primarily in One Type of Unified Atom

Unified atom: pair of amides	ΔASA			
	Amide sp^2 O	Amide sp^2 N	Aliphatic sp^3 C	Total ΔASA
Amide O: (1,3-deu, aama; see Fig 3B)	33.8	-12.7	8.4	29.6
Amide N: (1,3-dmu, nma; see Fig 3C)	-6.6	25.9	-1.3	17.9
sp^3 C: (1,1-deu; 1,1-dmu; see Fig 3D)	-2.7	-4.5	60.2	53.0
Amide N: (eu; proprionamide; see Fig S7)	1.5	20.4	0.1	22.0

Table S6. Interaction Potentials (α – Values) for Urea and Alkyl Ureas from Different Treatments of sp^{2}C (See Table 1)

(Fit) Atom Type	α_i (cal mol^{-1} molal^{-1} Å^{-2})						
	urea	mu	eu	1,1-dmu	1,3-dmu	1,1-deu	1,3-deu
(C) Amide sp^{2}O[a]	-0.55 ± 0.03	0.78 ± 0.09	1.16 ± 0.14	1.43 ± 0.07	1.65 ± 0.08	2.41 ± 0.13	2.78 ± 0.18
(C) Amide sp^{2}N[a]	-0.06 ± 0.02	-0.5 ± 0.04	-0.67 ± 0.08	-0.68 ± 0.03	-0.74 ± 0.03	-1.02 ± 0.05	-1.11 ± 0.06
(C) sp^{3}C[a]	-0.07 ± 0.02	-0.34 ± 0.01	-0.47 ± 0.02	-0.38 ± 0.01	-0.54 ± 0.02	-0.73 ± 0.03	-0.8 ± 0.04
(C) Aro/Amide sp^{2}C[a]	-0.59 ± 0.01	-1.36 ± 0.08	-1.6 ± 0.07	-2.03 ± 0.07	-1.95 ± 0.04	-2.69 ± 0.1	-2.97 ± 0.13
(I) Aro sp^{2}C[b]	-0.59 ± 0.01	-1.36 ± 0.08	-1.61 ± 0.07	-2.04 ± 0.08	-1.95 ± 0.04	-2.7 ± 0.1	-2.97 ± 0.13
(I) Amide sp^{2}C[b]	-0.69 ± 0.06	-0.99 ± 0.07	-1.2 ± 0.05	-1.53 ± 0.05	-1.87 ± 0.06	-1.88 ± 0.07	-2.28 ± 0.07
(P) C$_{10}$H$_8$ sp^{2}C[c]	-0.45 ± 0.2	-1.29 ± 0.15	-1.67 ± 0.15	-1.99 ± 0.13	-2.14 ± 0.11	-2.53 ± 0.13	-2.89 ± 0.11

[a] Combined (C) sp^{2}C analysis: aromatic sp^2 C is assumed to be the same as amide sp^2 C in fitting μ_{23} values (Tables S2, S3) to Equation 3.
[b] Individual (I) sp^{2}C analysis: individual α-values for aromatic sp^2 C and amide sp^2 C are reproduced from Table 1.
[c] Predicted (P) C$_{10}$H$_8$ sp^{2}C analysis: aromatic sp^{2}C α-values are predicted from naphthalene α-values in Table 1 (see text).

Table S7. Predicted μ_{23}-Values (cal mol^{-1} molal^{-1}) for Amide Interactions with Compounds Displaying Amide sp^{2}C or Aromatic sp^{2}C using Combined α-Values for Amide and Aromatic sp^{2}C

Solute	Formamide C-Predicted μ_{23}	N-methylformamide C-Predicted μ_{23}	Naphthalene C-Predicted μ_{23}	Anthracene C-Predicted μ_{23}
urea	-57.9 ± 2.5	-53.7 ± 1.7	-160 ± 3.7	-198 ± 4.5
mu	-50.3 ± 5.6	-65.2 ± 4.8[a]	-372 ± 21	-460 ± 27
eu	-53.1 ± 8.1	-75.8 ± 7.1	-438 ± 20	-542 ± 25
1,1-dmu	-56.7 ± 8.3	-73.4 ± 7.3	-555 ± 20	-687 ± 25
1,3-dmu	-44.8 ± 6.9	-76.2 ± 6.1	-533 ± 12	-660 ± 15
1,1-deu	-59.6 ± 6.7	-99.2 ± 5.8[a]	-736 ± 28	-911 ± 34
1,3-deu	-51.8 ± 14.4	-101 ± 13[a]	-810 ± 37	-1003± 45

[a]These underlined C-predicted μ_{23} values are the only cases where use of the combined amide/aromatic sp^{2}C α-value provides significantly worse agreement than individual amide and aromatic sp^{2}C α-values.

Table S8. Comparison of μ_{23} and μ_{32} Calculated from α-Values for Each of the Two Solutes

Solute 2	Solute 3	Observed μ_{23} (cal mol^{-1} molal^{-1})	Predicted μ_{32} or μ_{23} (cal mol^{-1} molal^{-1})	
			α-values of solute 2[a]	α-value of solute 3[b]
urea	mu	-37.8 ± 1.9	-38.2 ± 2.5	-39.6 ± 5.9
urea	eu	-43.8 ± 2.3	-40 ± 2.4	-42.3 ± 4.2
urea	1,1-dmu	-35.7 ± 2.1	-38.8 ± 2	-37.9 ± 4.5
urea	1,3-dmu	-30.2 ± 1.1	-34.5 ± 1.6	-33.2 ± 5.8
urea	1,1-deu	-39.7 ± 1.2	-39.3 ± 1.9	-40.4 ± 6.3
urea	1,3-deu	-40.3 ± 2.5	-38.2 ± 1.6	-34.8 ± 6.4
mu	eu	-59.4 ± 2.3	-61.8 ± 4.4	-63.2 ± 3.2
mu	1,1-dmu	-46.8 ± 2.1	-56.9 ± 4	-52.4 ± 3.4
mu	1,3-dmu	-59.1 ± 2.8	-66.1 ± 3.4	-64.6 ± 4.4
mu	1,1-deu	-78.3 ± 1.9	-72.8 ± 4	-76.2 ± 4.8
mu	1,3-deu	-87.8 ± 3.3	-84 ± 3.7	-80.1 ± 4.9
eu	1,1-dmu	-69.2 ± 2.7	-70.3 ± 2.9	-61.7 ± 3.4
eu	1,3-dmu	-83.1 ± 2.9	-84.1 ± 2.6	-80.6 ± 4.5
eu	1,1-deu	-96.2 ± 2.7	-92.6 ± 3	-94.6 ± 4.9
eu	1,3-deu	-103 ± 1.9	-109 ± 2.9	-102 ± 5
1,1-dmu	1,3-dmu	-61.6 ± 1.9	-66.7 ± 2.7	-72.2 ± 4.2
1,1-dmu	1,1-deu	-77.7 ± 2.7	-70.3 ± 3.3	-86.1 ± 4.6
1,1-dmu	1,3-deu	-79.8 ± 2.8	-85.4 ± 3.1	-91.8 ± 4.6
1,3-dmu	1,1-deu	-114 ± 3.6	-102 ± 4.4	-112 ± 4
1,3-dmu	1,3-deu	-125 ± 3.5	-128 ± 4.2	-125 ± 4.1
1,1-deu	1,3-deu	-146 ± 4.5	-149 ± 4.6	-134 ± 4.9

[a]Predicted μ_{32} from interaction potentials of α-values (**Table 1**) of solute 2 at 23°C.
[b]Predicted μ_{23} from interaction potentials of α-values (**Table 1**) of solute 3 at 23°C.

Table S9. Number-based Fits: α_n-Values (cal mol^{-1} molal^{-1} Å^{-2}; Eq. S14) for Interactions of Alkyl Ureas with Atoms of Amide Compounds

	3 α_n-Values, 0 global weights (Figure S10 Panel A)			**4 α_n-Values, 1 global weight (Figure S10 Panel B)**			
	Amide N	Amide C,O	Aliphatic C	Amide N	Amide O	Aliphatic C	Amide C
urea	-3.7 ± 1.2	-25.7 ± 1.8	-1.7 ± 0.2	-6.1 ± 1.2	0 ± 2.7	-2.1 ± 0.2	-52.4 ± 3.4
mu	-16.3 ± 1.4	-7.7 ± 1.7	-9.1 ± 0.6	-18.6 ± 1.4	21.4 ± 2.4	-11.5 ± 0.5	-48.4 ± 3.1
eu	-23.5 ± 1.2	-2.2 ± 1.5	-12.2 ± 0.5	-26.7 ± 1.2	30.4 ± 2.3	-14.6 ± 0.5	-53.4 ± 2.3
1,1-dmu	-26.3 ± 1.2	6.1 ± 1.5	-7 ± 0.5	-30.4 ± 1.2	47.7 ± 2.2	-10.1 ± 0.5	-68.2 ± 2.4
1,3-dmu	-27 ± 1.3	12.7 ± 1.8	-17.3 ± 0.6	-31.6 ± 1.3	60.3 ± 2.7	-20.8 ± 0.6	-77.9 ± 2.6
1,1-deu	-36.3 ± 1.2	18.7 ± 1.5	-20.7 ± 0.5	-41.7 ± 1.2	68.5 ± 2.3	-25.2 ± 0.6	-78.7 ± 2.4
1,3-deu	-41.3 ± 1.5	26.9 ± 1.9	-21.4 ± 0.8	-47.2 ± 1.5	81.7 ± 2.8	-26.3 ± 0.8	-86.6 ± 2.8

	3 α_n-Values, 4 global weights (Figure S10 Panel C)			**4 α_n-Values, 5 global weights (Figure S10 Panel D)**			
	Amide N	Amide C,O	Aliphatic C	Amide N	Amide O	Aliphatic C	Amide C
urea	0.5 ± 0.6	-34 ± 1.3	-2.2 ± 0.3	-4.3 ± 1.7	-2.6 ± 2.6	-3.6 ± 0.3	-51.9 ± 3.6
mu	-4.1 ± 0.8	-22.1 ± 1.3	-14.5 ± 0.6	-23.1 ± 1.7	30.6 ± 3.3	-24.3 ± 0.8	-52.5 ± 3.5
eu	-6.6 ± 0.8	-21.2 ± 1.3	-20 ± 0.6	-25.1 ± 1.4	34.7 ± 2.5	-32.5 ± 0.7	-57 ± 2.4
1,1-dmu	-4.7 ± 0.7	-15.6 ± 1.2	-16.4 ± 0.7	-31.1 ± 1.4	58 ± 2.6	-27.5 ± 0.7	-77.2 ± 2.6
1,3-dmu	-10.6 ± 1	-7.4 ± 1.6	-25.4 ± 0.8	-33.5 ± 1.5	67.9 ± 3.1	-42.5 ± 1	-81.6 ± 2.8
1,1-deu	-15.5 ± 0.8	-6.8 ± 1.3	-26.7 ± 0.9	-38.6 ± 1.6	69 ± 3.3	-50 ± 1.1	-79.9 ± 3.1
1,3-deu	-18.4 ± 1	1.9 ± 1.4	-32.5 ± 1	-43 ± 1.8	85.7 ± 3.2	-57.8 ± 1.1	-90.2 ± 3

Relative Weight of Group ($w_{x,y}$; Equation S14)

	N	NH	NH2	amide O	CH	CH2	CH3	amide C	amide CH	ssresid[a]
Panel A	1	1	1	1	1	1	1	1	1	32761
Panel B	1	1	1	1	1	1	1	0.39	1	14161
Panel C	0.52[b]	1.86	1	1	-2.73[b]	0.57	1	1	1	15129
Panel D	-0.59[b]	0.48	1	1	-0.83[b]	0.47	1	0.45	1	2500
ASA	--	--	--	--	--	--	--	--	--	3025

[a]Ssresid: Sum of squares of residuals for predicted - observed differences in the set of μ_{23} values $\sum(\mu_{23}^{predicted} - \mu_{23}^{observed})^2$

[b]Determined from very small sample sizes (1 compound with CH, 2 compounds with N).

Table S10. Composite Amide, Hydrocarbon and Amide-Hydrocarbon α-Values to Predict Solute Effects on Protein Folding

Solute	α-value (cal mol^{-1} molal^{-1} Å^{-2})		
	Amide[a]	**Hydrocarbon**[b]	**Combined**[c]
urea	-0.39 ± 0.03	-0.12 ± 0.01	-0.18 ± 0.01
methylurea	0.24 ± 0.06	-0.42 ± 0.01	-0.27 ± 0.02
ethylurea	0.4 ± 0.04	-0.54 ± 0.01	-0.33 ± 0.01
butyl urea[d]	0.8 ± 0.14	-0.82 ± 0.01	-0.46 ± 0.03
1,1-dimethylurea	0.59 ± 0.05	-0.51 ± 0.01	-0.27 ± 0.01
1,3-dimethylurea	0.96 ± 0.06	-0.69 ± 0.01	-0.33 ± 0.02
1,1-diethylurea	0.97 ± 0.07	-0.84 ± 0.02	-0.44 ± 0.02
1,3-diethylurea	1.36 ± 0.07	-0.97 ± 0.02	-0.46 ± 0.02

[a]Composite amide (O, N) α value for the amide surface exposed in protein unfolding (ΔASA_{amide}) calculated for amide O : amide N = 2.4 : 1[15]

[b]Composite hydrocarbon (sp^{3}C, sp^{2}C) α value for the hydrocarbon surface exposed in protein unfolding ($\Delta ASA_{hydrocarbon}$) calculated for aliphatic sp^{3}C : aromatic sp^{2}C = 9.3 : 1[15]

[c]Composite overall α value for the entire surface exposed in protein unfolding (ΔASA_{total}) calculated for hydrocarbon (sp^3, sp^2) : amide(O,N) = 3.5 : 1

[d]butyl urea α-values estimated by extrapolation of methyl and ethyl urea α-values (Table 1).

Table S11. Comparison of urea α-values (cal mol^{-1} molal^{-1} Å^{-2})

	Amide sp^{2}O	Amide sp^{2}N	Aliphatic sp^{3}C	Amide sp^{2}C	Aromatic sp^{2}C
This work	-0.52 ± 0.04	-0.09 ± 0.02	-0.07 ± 0.01	-0.69 ± 0.06	-0.59 ± 0.01
Guinn et al[14]	-0.5 ± 0.11	-0.22 ± 0.09	-0.07 ± 0.03	--[a]	-0.53 ± 0.03

[a]Not determined.

Table S12. Thermodynamics of Amide Interactions from Experiment and Molecular Dynamics (MD) Simulations

Solute	Solute	Interaction (μ_{23} or $\mu_{22}^{ex}/2$, cal mol^{-1} molal^{-1})	
		Simulation	Prediction
urea	benzene	μ_{23} = -109 ± 16 (at 4 - 6 molal)[a]	μ_{23} = -124 ± 3 (<1 molal)[d]
nma	nma	$\mu_{22}^{ex}/2 \sim$ -40[b] or 16[c] (at 5.5 molal)	$\mu_{22}^{ex}/2 \sim$ (-40[e] to -80[f]) (<0.8 molal)

[a]Average μ_{23} calculated using tabulated Kirkwood – Buff integrals from 15 MD simulations of interactions of urea (4 – 6 molal) with benzene in water, using various force fields.[33]

[b, c]Estimates of $\mu_{22}^{ex}/2$ obtained from graphs of Kirkwood – Buff integrals for self-interaction of aqueous nma (5.5 molal) from MD simulations using CHARMM[b] or Drude[c] force field.[34]

[d]Predicted μ_{23} value of urea with benzene using urea-aromatic sp^{2}C α-value (Table 1).

[e]methyl urea $\mu_{22}^{ex}/2$

[f]ethyl urea $\mu_{22}^{ex}/2$

Table S13. Comparison of Amide Osmotic Coefficients φ from Isopiestic Distillation (ID) and VPO Measurements

			Urea		
molal	φ_{ID}[a]	Φ_{VPO}[b]	molal	φ_{ID}[a]	Φ_{VPO}[b]
0.2054	0.991	0.99	0.6192	0.976	0.969
0.2101	0.99	0.989	0.8094	0.968	0.96
0.4071	0.984	0.98	0.8365	0.966	0.958
0.4137	0.984	0.979	1.028	0.961	0.949
0.61	0.977	0.969	1.0372	0.961	0.948

	methylurea			ethylurea	
molal	φ_{ID}[a]	Φ_{VPO}[b]	molal	φ_{ID}[a]	Φ_{VPO}[b]
0.535	0.973	0.962	0.525	0.951	0.923
0.675	0.966	0.955	0.701	0.931	0.905
0.93	0.955	0.947	1.034	0.906	0.862

	1,1-dimethylurea			1,3-dimethylurea	
molal	φ_{ID}[a]	Φ_{VPO}[b]	molal	φ_{ID}[a]	Φ_{VPO}[b]
0.522	0.957	0.976	0.55	0.947	0.937
0.69	0.946	0.968	0.698	0.935	0.924
1.011	0.927	0.954	0.971	0.915	0.901

	1,1-diethylurea			1,3-diethylurea	
molal	φ_{ID}[a]	Φ_{VPO}[b]	molal	φ_{ID}[a]	Φ_{VPO}[b]
0.423	0.889	0.909	0.432	0.866	0.895
0.592	0.88	0.881	0.586	0.852	0.865
0.693	0.854	0.866	0.871	0.818	0.818
0.96	0.827	0.836	0.987	0.805	0.803
1.047	0.811	0.829	1.124	0.79	0.787

[a]Literature values of φ_{ID} from isopiestic distillation results at 25°C.[5-6]
[b]Values of Φ_{VPO} calculated from fits to VPO data using the fitting coefficients β_{22} and β_{222} in Table S1.

Table S14. Two Component Values of $RT\beta_{22}$ from Nonlinear Fitting

Solute	$RT\beta_{22} \approx \mu_{22}^{excess}/2$ (cal mol^{-1} molal^{-1})	
	Observed[a]	Predicted[b]
urea	-29.4 ± 2.9	-41.9 ± 3.5
methylurea	-52.6 ± 5.3	-52.9 ± 4.4
ethylurea	-94.2 ± 27	-75.6 ± 3.2
1,1-dimethylurea	-27.5 ± 8	-52.8 ± 3.2
1,3-dimethylurea	-76 ± 7.6	-96 ± 3.6
1,1-diethylurea	-148 ± 15	-121 ± 4.8
1,3-diethylurea	-162 ± 16	-169 ± 4.7

[a]Observed $RT\beta_{22}$ from fitted β_{22} values from Table S1 and RT = 588 cal mol^{-1}. Error estimates are the larger of 10% or the estimated standard deviation determined from the fit of the data from Figure S1A.
[b] Predicted values of $RT\beta_{22}$ from α values in Table 1.

Table S15. Comparison of ASA Values Calculated from Cactus[a], PubChem[b] and BMRB[c] Structures

		ASA (Å^2)				
		Amide sp^{2}O	Amide sp^{2}N	Aliphatic sp^{3}C	Amide or aromatic sp^{2}C	Total
urea	Cactus	47.9	130.2	0	7.2	185.3
	PubChem	49.5	129.3	0	7.7	186.5
	% Difference	3.3	0.7	---	6.9	0.6
methylurea	Cactus	38.3	87.5	88.4	6.5	220.7
	PubChem	41.8	84.2	89.9	7	222.9
	% Difference	9.1	3.8	1.7	7.7	1
	BMRB	40.3	85.1	89.8	6.5	221.7
ethylurea	Cactus	38.3	82	124.6	6.5	251.4
	PubChem	41.8	84.2	121.3	6.7	254
	% Difference	9.1	2.7	2.6	3.1	1
1,1-dimethylurea	Cactus	38.3	54.5	148.7	6.2	247.7
	PubChem	40.8	54.4	147.5	6.7	249.4
	% Difference	6.5	0.2	0.8	8.1	0.7
1,3-dimethylurea	Cactus	28.7	44.9	176.7	5.8	256.1
	PubChem	33.7	39.7	179.4	6.3	259.1
	% Difference	17.4	11.6	1.5	8.6	1.2
	Cactus	29	45.2	175.6	5.6	255.4
1,1-diethylurea	Cactus	35.7	50	208.9	3.7	298.3
	PubChem	38.6	49.8	207.2	4.4	300
	% Difference	8.1	0.4	0.8	18.9	0.6
1,3-diethylurea	Cactus	28.7	33.8	249.2	5.8	317.5
	PubChem	10.1	39.7	242.5	5.8	298.1
	% Difference	64.8	17.5	2.7	0	6.1
malonamide	Cactus	65.7	123.2	48.5	8.5	245.9
	PubChem	79.8	121.8	42	7.4	251
	% Difference	21.5	1.1	13.4	12.9	2.1
propionamide	Cactus	36.8	61.6	124.5	4.3	227.2
	PubChem	39.9	60.7	122.9	4.7	228.2
	% Difference	8.4	1.5	1.3	9.3	0.4
N-methylacetamide	Cactus	35.3	19	178	3.5	235.8
	PubChem	38.9	16.1	178.2	3.9	237.1
	% Difference	10.2	15.3	0.1	11.4	0.6
aama	Cactus	62.5	21.1	257.7	4.3	345.6
	PubChem	69.9	19.4	256.9	4.5	350.7
	% Difference	11.8	8.1	0.3	4.7	1.5

formamide	Cactus	51.3	70.8	0	40.2	162.3
	PubChem	52.7	71.6	0	38.7	163
	% Difference	2.7	1.1	---	3.7	0.4
	BMRB	53.5	69.5	0	39.8	162.8
N-methylformamide	Cactus	41.7	27.7	88.4	39.5	197.3
	PubChem	45.4	26.6	90.2	37.5	199.7
	% Difference	8.9	4	2	5.1	1.2
naphthalene	Cactus	0	0	0	273	273
	PubChem	0	0	0	275	275
	% Difference	---	---	---	0.7	0.7
anthracene	Cactus	0	0	0	334	334
	PubChem	0	0	0	338	338
	% Difference	---	---	---	1.2	1.2
acetamide	Cactus	44.9	61.6	89.7	4.3	200.5
	PubChem	46.5	61.1	88.6	4.6	200.8
	% Difference	3.6	0.8	1.2	7	0.1
	BMRB	46	60.6	89.5	4.4	200.5
N-N-dimethylformamide	Cactus	41.7	0.8	157.7	30.7	230.9
	PubChem	45	0.8	157.6	29.2	232.6
	% Difference	7.9	0	0.1	4.9	0.7
	BMRB	46.7	0.2	160.4	26.8	234.1
tetramethylurea	Cactus	29.9	0.8	266.4	3.3	300.4
	PubChem	33.8	0.8	267.6	2.8	305
	% Difference	13	0	0.5	15.2	1.5
ethyleneurea	Cactus	49	57.1	113.6	7	226.7
	PubChem	50.4	56.9	112.6	7.8	227.7
	% Difference	2.9	0.4	0.9	11.4	0.4
N-N-dimethylacetamide	Cactus	35.3	0.8	220.1	3.6	259.7
	PubChem	37.7	0.9	218.5	3.8	260.8
	% Difference	6.8	11.4	0.7	8.1	0.4

[a]Reference[27];[b]Reference[29];[c]Reference[31]; BMRB ID for solutes: methylurea: bmse000738; 1,3-dimethylurea: bmse000248; formamide: bmse000267; acetamide: bmse000825; N,N-dimethylformamide: bmse000709

Table S16. Comparison of α-Values Calculated Using Cactus and PubChem Structures for Different Amide Data Sets

Solute	Structure Source	α-Values (cal mol^{-1} molal^{-1}Å^{-2})				
		Amide sp^{2}O	Amide sp^{2}N	Aliphatic sp^{3}C	Amide sp^{2}C	Aromatic sp^{2}C
urea	Cactus	-0.52 ± 0.04	-0.09 ± 0.02	-0.07 ± 0.01	-0.69 ± 0.06	-0.59 ± 0.01
	Cactus (w/o 1,3-deu, mad)[a]	-0.52 ± 0.04	-0.09 ± 0.03	-0.06 ± 0.01	-0.69 ± 0.06	-0.59 ± 0.01
	PubChem	-0.49 ± 0.03	-0.08 ± 0.02	-0.07 ± 0.01	-0.74 ± 0.06	-0.58 ± 0.01
	Pubchem (w/o 1,3-deu, mad)[a]	-0.53 ± 0.04	-0.07 ± 0.03	-0.06 ± 0.01	-0.7 ± 0.07	-0.58 ± 0.01
methylurea	Cactus	0.52 ± 0.08	-0.44 ± 0.03	-0.31 ± 0.01	-0.99 ± 0.07	-1.37 ± 0.08
	Cactus (w/o 1,3-deu, mad)[a]	0.42 ± 0.11	-0.42 ± 0.04	-0.3 ± 0.01	-0.92 ± 0.08	-1.37 ± 0.08
	PubChem	0.32 ± 0.05	-0.4 ± 0.02	-0.29 ± 0.01	-0.89 ± 0.06	-1.36 ± 0.08
	Pubchem (w/o 1,3-deu, mad)[a]	0.4 ± 0.1	-0.41 ± 0.03	-0.3 ± 0.02	-0.96 ± 0.09	-1.36 ± 0.08
ethylurea	Cactus	0.79 ± 0.06	-0.55 ± 0.02	-0.43 ± 0.01	-1.2 ± 0.05	-1.62 ± 0.07
	Cactus (w/o 1,3-deu, mad)[a]	1.02 ± 0.08	-0.6 ± 0.03	-0.47 ± 0.01	-1.36 ± 0.07	-1.62 ± 0.07
	PubChem	0.38 ± 0.04	-0.43 ± 0.02	-0.38 ± 0.01	-0.98 ± 0.04	-1.6 ± 0.07
	Pubchem (w/o 1,3-deu, mad)[a]	0.95 ± 0.08	-0.58 ± 0.03	-0.48 ± 0.02	-1.4 ± 0.07	-1.6 ± 0.07
1,1-dimethylurea	Cactus	1.09 ± 0.07	-0.61 ± 0.03	-0.35 ± 0.01	-1.53 ± 0.05	-2.05 ± 0.08
	Cactus (w/o 1,3-deu, mad)[a]	1.17 ± 0.08	-0.63 ± 0.03	-0.37 ± 0.01	-1.57 ± 0.07	-2.05 ± 0.08
	PubChem	0.6 ± 0.04	-0.48 ± 0.02	-0.29 ± 0.01	-1.27 ± 0.05	-2.03 ± 0.08
	Pubchem (w/o 1,3-deu, mad)[a]	1.14 ± 0.08	-0.63 ± 0.03	-0.39 ± 0.01	-1.65 ± 0.07	-2.03 ± 0.07
1,3-dimethylurea	Cactus	1.68 ± 0.09	-0.77 ± 0.03	-0.56 ± 0.01	-1.87 ± 0.06	-1.97 ± 0.04
	Cactus (w/o 1,3-deu, mad)[a]	1.81 ± 0.12	-0.8 ± 0.04	-0.58 ± 0.02	-1.96 ± 0.08	-1.97 ± 0.04
	PubChem	0.94 ± 0.05	-0.58 ± 0.02	-0.46 ± 0.01	-1.47 ± 0.05	-1.95 ± 0.04
	Pubchem (w/o 1,3-deu, mad)[a]	1.66 ± 0.11	-0.76 ± 0.04	-0.6 ± 0.02	-2 ± 0.08	-1.95 ± 0.04
1,1-diethylurea	Cactus	1.73 ± 0.09	-0.84 ± 0.03	-0.64 ± 0.01	-1.88 ± 0.07	-2.72 ± 0.1
	Cactus (w/o 1,3-deu, mad)[a]	1.76 ± 0.12	-0.86 ± 0.04	-0.65 ± 0.02	-1.89 ± 0.09	-2.72 ± 0.1
	PubChem	0.97 ± 0.06	-0.64 ± 0.03	-0.54 ± 0.01	-1.48 ± 0.06	-2.69 ± 0.1
	Pubchem (w/o 1,3-deu, mad)[a]	1.7 ± 0.12	-0.84 ± 0.04	-0.68 ± 0.02	-1.98 ± 0.1	-2.69 ± 0.1
1,3-diethylurea	Cactus	2.35 ± 0.09	-1.01 ± 0.04	-0.76 ± 0.01	-2.28 ± 0.07	-2.99 ± 0.13
	Cactus (w/o 1,3-deu, mad)[a]	2.24 ± 0.12	-0.99 ± 0.04	-0.74 ± 0.02	-2.19 ± 0.09	-2.99 ± 0.13
	PubChem	1.74 ± 0.07	-0.89 ± 0.03	-0.71 ± 0.01	-1.96 ± 0.06	-2.96 ± 0.13
	Pubchem (w/o 1,3-deu, mad)[a]	2.09 ± 0.11	-0.95 ± 0.04	-0.77 ± 0.02	-2.26 ± 0.1	-2.96 ± 0.13

[a]α-analysis of μ_{23} values for all amides in Tables S2-3 except malonamide and 1,3-diethyl urea, using structures from Cactus or PubChem as indicated.

Table S17. Comparison of Predicted μ_{23} Values For Interactions of Alkyl Ureas with 1,3 - Diethyl Urea and Malonamide Using Different α-Values from Table S16

Solute	Observed μ_{23}	malonamide			
		Predicted μ_{23} (cal mol^{-1} molal^{-1})			
		Cactus[a]	Cactus (w/o 1,3-deu, mad)	PubChem	PubChem (w/o 1,3-deu, mad)
urea	-54.9 ± 2.5	-54.4 ± 3.8	-54.1 ± 4	-57 ± 4.1	-58 ± 4.4
methylurea	-43.8 ± 1.4	-43.8 ± 6.8	-45.9 ± 8.3	-41.6 ± 5.2	-37.6 ± 9
ethylurea	-48.5 ± 2.4	-46.6 ± 4.9	-41.2 ± 6.6	-45.6 ± 3.6	-25.5 ± 7.1
1,1-dimethylurea	-30.6 ± 1.8	-33.2 ± 5.3	-32.7 ± 6.5	-32.2 ± 4.3	-14.6 ± 7.3
1.3-dimethylurea	-27.3 ± 2.4	-27.4 ± 6.9	-24.8 ± 8.9	-25.8 ± 5.2	-0.6 ± 9.6
1,1-diethylurea	-34.6 ± 1.3	-37.1 ± 7.5	-37.4 ± 9.4	-34.8 ± 6	-10.4 ± 10.6
1,3-diethylurea	-21.7 ± 1.2	-25.8 ± 7.6	-29.8 ± 9.1	-13.6 ± 6.6	1.6 ± 10.1

Solute	Observed μ_{23}	1,3-diethylurea			
		Predicted μ_{23} (cal mol^{-1} molal^{-1})			
		Cactus[a]	Cactus (w/o 1,3-deu, mad)	PubChem	PubChem (w/o 1,3-deu, mad)
urea	-40.3 ± 2.5	-38.2 ± 1.6	-37.9 ± 1.7	-29.4 ± 1.4	-26.8 ± 1.5
methylurea	-87.8 ± 3.3	-84 ± 3.7	-81 ± 4.9	-87.7 ± 2.3	-91.5 ± 4.1
ethylurea	-103 ± 1.9	-109 ± 2.9	-115 ± 4.5	-111 ± 1.8	-138 ± 4
1,1-dimethylurea	-79.8 ± 2.8	-85.4 ± 3.1	-88.4 ± 4.1	-90.8 ± 2	-118 ± 3.7
1.3-dimethylurea	-125 ± 3.5	-128 ± 4.2	-132 ± 5.8	-134 ± 2.6	-170 ± 5
1,1-diethylurea	-146 ± 4.5	-149 ± 4.6	-150 ± 6	-156 ± 3	-193 ± 5.4
1,3-diethylurea[b]	(-133 ± 13)	(-169 ± 4.7)	(-167 ± 6)	(-202 ± 3.4)	(-217 ± 5.4)

[a]From Tables S2

[b]Values of $\mu_{22}^{ex}/2$

Table S18. Methylurea: Effects of Changes in One α-Value on Other α-Values

(A) Fix α-value of amide O					(B) Fix α-value of amide N				
Amide O	Amide N[a]	Aliphatic C[a]	Amide C[a]	Ssresid[b]	Amide O[a]	Amide N[a]	Aliphatic C[a]	Amide C[a]	Ssresid[b]
0.02	-0.27 ± 0.03	-0.26 ± 0.01	-0.68 ± 0.09	664	0.08 ± 0.05	-0.24	-0.27 ± 0.01	-0.79 ± 0.08	670
0.27	-0.35 ± 0.03	-0.29 ± 0.01	-0.83 ± 0.09	438	0.3 ± 0.05	-0.34	-0.29 ± 0.01	-0.89 ± 0.08	440
0.44[d]	*-0.41 ± 0.03*	*-0.31 ± 0.01*	*-0.94 ± 0.09*	*370*	*0.45 ± 0.05*	*-0.41[d]*	*-0.31 ± 0.01*	*-0.96 ± 0.08*	*370*
0.52[c]	**-0.44 ± 0.03**	**-0.31 ± 0.01**	**-0.99 ± 0.09**	**362**	**0.52 ± 0.05**	**-0.44[c]**	**-0.31 ± 0.01**	**-0.99 ± 0.08**	**362**
0.6[d]	*-0.47 ± 0.03*	*-0.32 ± 0.01*	*-1.04 ± 0.09*	*370*	*0.58 ± 0.05*	*-0.47[d]*	*-0.32 ± 0.01*	*-1.01 ± 0.08*	*369*
0.77	-0.53 ± 0.03	-0.34 ± 0.01	-1.14 ± 0.09	438	0.74 ± 0.05	-0.54	-0.33 ± 0.01	-1.08 ± 0.08	438
1.02	-0.61 ± 0.03	-0.37 ± 0.01	-1.29 ± 0.09	664	0.96 ± 0.05	-0.64	-0.35 ± 0.01	-1.18 ± 0.08	665
(C) Fix α-value of aliphatic sp^{3}C					(D) Fix α-value of amide sp^{2}C				
Amide O[a]	Amide N[a]	Aliphatic C[a]	Amide C[a]	Ssresid[b]	Amide O[a]	Amide N[a]	Aliphatic C[a]	Amide C[a]	Ssresid[b]
0.12 ± 0.06	-0.33 ± 0.03	-0.25	-0.76 ± 0.06	641	0.2 ± 0.07	-0.36 ± 0.03	-0.28 ± 0.01	-0.53	650
0.31 ± 0.06	-0.38 ± 0.03	-0.28	-0.86 ± 0.06	441	0.36 ± 0.07	-0.4 ± 0.03	-0.3 ± 0.01	-0.76	433
0.43 ± 0.06	*-0.42 ± 0.03*	*-0.3[d]*	*-0.94 ± 0.06*	*376*	*0.47 ± 0.07*	*-0.43 ± 0.03*	*-0.31 ± 0.01*	*-0.92[d]*	*368*
0.49 ± 0.06	**-0.43 ± 0.03**	**-0.31[c]**	**-0.97 ± 0.06**	**364**	**0.52 ± 0.07**	**-0.44 ± 0.03**	**-0.31 ± 0.01**	**-0.99[c]**	**362**
0.56 ± 0.06	*-0.45 ± 0.03*	*-0.32[d]*	*-1.01 ± 0.06*	*365*	*0.57 ± 0.07*	*-0.45 ± 0.03*	*-0.32 ± 0.01*	*-1.06[d]*	*370*
0.68 ± 0.06	-0.49 ± 0.03	-0.34	-1.08 ± 0.06	408	0.68 ± 0.07	-0.48 ± 0.03	-0.33 ± 0.01	-1.22	438
0.87 ± 0.06	-0.54 ± 0.03	-0.37	-1.19 ± 0.06	575	0.85 ± 0.07	-0.52 ± 0.03	-0.35 ± 0.01	-1.45	659

[a] Propagated uncertainties in α-values are calculated as described previously [16]

[b] Ssresid: Sum of squares of residuals for predicted - observed differences in the set of μ_{23} values $\sum(\mu_{23}^{predicted} - \mu_{23}^{observed})^2$

[c] Best-fit value from Table 1.

[d] ±1 SD from the best-fit value from Table 1.

Table S19. Urea α-Value: Effects of Dataset Size on α-Values

	Amide sp^{2}O	Amide sp^{2}N	Aliphatic sp^3 C	Amide sp^{2}C	Aromatic sp^{2}C
14 μ_{23}[a]	-0.5 ± 0.12	-0.11 ± 0.05	-0.06 ± 0.01	-0.71 ± 0.09	-0.59 ± 0.01
21 μ_{23}[b]	-0.52 ± 0.04	-0.09 ± 0.02	-0.07 ± 0.01	-0.69 ± 0.06	-0.59 ± 0.01

[a]From Tables S2-3 above.

[b]Including μ_{23} values for seven amides from reference[14] in addition to those in Tables S2-3 above.

Table S20. Comparison of α-values Calculated with/without Inclusion of Self-Interaction μ_{22}^{ex}

		α-values (cal mol^{-1} molal^{-1}Å^{-2})				
		Amide sp^{2}O	Amide sp^{2}N	Aliphatic sp^{3}C	Amide sp^{2}C	Aromatic sp^{2}C
urea	without μ_{22}^{ex}	-0.52 ± 0.04	-0.09 ± 0.02	-0.07 ± 0.01	-0.69 ± 0.06	-0.59 ± 0.01
	with μ_{22}^{ex}	-0.52 ± 0.04	-0.1 ± 0.02	-0.06 ± 0.01	-0.68 ± 0.06	-0.59 ± 0.01
mu	without μ_{22}^{ex}	0.52 ± 0.08	-0.44 ± 0.03	-0.31 ± 0.01	-0.99 ± 0.07	-1.36 ± 0.08
	with μ_{22}^{ex}	0.46 ± 0.08	-0.4 ± 0.03	-0.31 ± 0.01	-0.98 ± 0.07	-1.36 ± 0.08
eu	without μ_{22}^{ex}	0.79 ± 0.06	-0.55 ± 0.02	-0.43 ± 0.01	-1.2 ± 0.05	-1.61 ± 0.07
	with μ_{22}^{ex}	0.79 ± 0.07	-0.55 ± 0.03	-0.43 ± 0.01	-1.2 ± 0.05	-1.61 ± 0.07
1,1-dmu	without μ_{22}^{ex}	1.09 ± 0.07	-0.61 ± 0.03	-0.35 ± 0.01	-1.53 ± 0.05	-2.04 ± 0.08
	with μ_{22}^{ex}	1.09 ± 0.06	-0.6 ± 0.03	-0.34 ± 0.01	-1.54 ± 0.05	-2.04 ± 0.08
1,3-dmu	without μ_{22}^{ex}	1.68 ± 0.09	-0.77 ± 0.03	-0.56 ± 0.01	-1.87 ± 0.06	-1.95 ± 0.04
	with μ_{22}^{ex}	1.45 ± 0.09	-0.69 ± 0.03	-0.52 ± 0.01	-1.74 ± 0.07	-1.95 ± 0.04
1,1-deu	without μ_{22}^{ex}	1.73 ± 0.09	-0.84 ± 0.03	-0.64 ± 0.01	-1.88 ± 0.07	-2.7 ± 0.1
	with μ_{22}^{ex}	1.7 ± 0.11	-0.83 ± 0.04	-0.63 ± 0.02	-1.87 ± 0.07	-2.7 ± 0.1
1,3-deu	without μ_{22}^{ex}	2.35 ± 0.09	-1.01 ± 0.04	-0.76 ± 0.01	-2.28 ± 0.07	-2.97 ± 0.13
	with μ_{22}^{ex}	1.95 ± 0.17	-0.89 ± 0.05	-0.68 ± 0.03	-2.03 ± 0.11	-2.97 ± 0.13

Table S21. Comparison of ASA Values (Å^2) Calculated from Surface Racer[a] (or equivalently GETAREA[b]) and VMD[c] Using Cactus Structures[d] (% Differences expressed as (100)*(1 – VMD ASA/Surface Racer ASA)

		ASA (Å^2)				
		Amide sp^{2}O	Amide sp^{2}N	Aliphatic sp^{3}C	Amide or aromatic sp^{2}C	Total
urea	Surface Racer/GETAREA	47.9	130.2	0	7.2	185.3
	VMD	49.2	131.1	0	6.3	186.7
	% Difference	-2.8	-0.7	---	+12.8	-0.7
methylurea	Surface Racer/GETAREA	38.3	87.5	88.4	6.5	220.7
	VMD	39.5	95.2	83.4	5.8	223.9
	% Difference	-3.2	-8.8	+5.7	+10.8	-1.4
ethylurea	Surface Racer/GETAREA	38.3	82	124.6	6.5	251.4
	VMD	42.9	83.3	122.9	6.3	255.4
	% Difference	-12	-1.6	+1.4	+3.4	-1.6
1,1-dimethylurea	Surface Racer/GETAREA	38.3	54.5	148.7	6.2	247.7
	VMD	41	56.5	141.5	6	245.1
	% Difference	-7.1	-3.7	+4.8	+2.6	+1.1
1,3-dimethylurea	Surface Racer/GETAREA	28.7	44.9	176.7	5.8	256.1
	VMD	27.5	46.1	173.2	4.8	251.6
	% Difference	+4.3	-2.7	+2	+16.7	+1.8
1,1-diethylurea	Surface Racer/GETAREA	35.7	50	208.9	3.7	298.3
	VMD	37	50	202.2	4.1	293.3
	% Difference	-3.6	0	+3.2	-11	+1.7
1,3-diethylurea	Surface Racer/GETAREA	28.7	33.8	249.2	5.8	317.5
	VMD	30	30.7	248.1	5.1	313.9
	% Difference	-4.6	+9.2	+0.4	+12.6	+1.1
malonamide	Surface Racer/GETAREA	65.7	123.2	48.5	8.5	245.9
	VMD	67.2	119.1	47.6	8.5	242.4
	% Difference	-2.3	+3.4	+1.8	0	+1.4
propionamide	Surface Racer/GETAREA	36.8	61.6	124.5	4.3	227.2
	VMD	37.6	66.9	120.9	4.3	229.7
	% Difference	-2.2	-8.6	+2.9	0	-1.1
N-methyl acetamide	Surface Racer/GETAREA	35.3	19	178	3.5	235.8
	VMD	36.6	16.4	178.1	4.3	227.2
	% Difference	-3.7	+13.6	0	-22.9	+3.6
aama	Surface Racer/GETAREA	62.5	21.1	257.7	4.3	345.6
	VMD	59.8	18.1	252.8	3.9	334.6
	% Difference	+4.3	+14.1	+1.9	+10.1	+3.2
formamide	Surface Racer/GETAREA	51.3	70.8	0	40.2	162.3
	VMD	55.8	66.9	0	39.6	162.3
	% Difference	-8.8	+5.5	---	+1.5	0

N-methyl formamide	Surface Racer/GETAREA	41.7	27.7	88.4	39.5	197.3
	VMD	44.2	26.8	85.1	41.5	197.7
	% Difference	-5.9	+3.2	+3.7	-5.2	-0.2
naphthalene	Surface Racer/GETAREA	0	0	0	273	273
	VMD	0	0	0	270	270
	% Difference	---	---	---	+1.1	+1.1
anthracene	Surface Racer/GETAREA	0	0	0	334	334
	VMD	0	0	0	332.3	332.3
	% Difference	---	---	---	+0.5	+0.5
acetamide	Surface Racer/GETAREA	44.9	61.6	89.7	4.3	200.5
	VMD	48	63.5	84.5	4.6	200.6
	% Difference	-6.9	-3.1	+5.7	-6.7	-0.1
N,N-dimethyl formamide	Surface Racer/GETAREA	41.7	0.8	157.7	30.7	230.9
	VMD	45.7	1	154.6	30.7	231.9
	% Difference	-9.5	-25	+2	0	-0.4
tetramethyl urea	Surface Racer/GETAREA	29.9	0.8	266.4	3.3	300.4
	VMD	27.7	0	260.3	2.7	290.7
	% Difference	+7.4	+100	+2.3	+19.5	3.2
ethyleneurea	Surface Racer/GETAREA	49	57.1	113.6	7	226.7
	VMD	44.2	52.7	118.2	6.5	221.6
	% Difference	+9.8	+7.8	-4.1	+6.8	+2.3
N,N-dimethyl acetamide	Surface Racer/GETAREA	35.3	0.8	220.1	3.6	259.7
	VMD	32.6	0	220.8	2.9	256.3
	% Difference	+7.9	+100	-0.3	+18.4	+1.3

[a]Reference[22]; [b]Reference[23]; [c]Reference[24]; [d]Reference[27]

Table S22. Comparison of α-values Calculated using ASA from Surface Racer[a] and VMD[b]

		α-values (cal mol^{-1} molal^{-1}Å^{-2})				
		Amide sp^{2}O	Amide sp^{2}N	Aliphatic sp^{3}C	Amide sp^{2}C	Aromatic sp^{2}C
urea	Surface Racer	-0.52 ± 0.04	-0.09 ± 0.02	-0.07 ± 0.01	-0.69 ± 0.06	-0.59 ± 0.01
	VMD	-0.53 ± 0.04	-0.08 ± 0.02	-0.07 ± 0.01	-0.63 ± 0.06	-0.59 ± 0.01
mu	Surface Racer	0.52 ± 0.08	-0.44 ± 0.03	-0.31 ± 0.01	-0.99 ± 0.07	-1.37 ± 0.08
	VMD	0.5 ± 0.08	-0.45 ± 0.03	-0.32 ± 0.01	-1.02 ± 0.07	-1.38 ± 0.08
eu	Surface Racer	0.79 ± 0.06	-0.55 ± 0.02	-0.43 ± 0.01	-1.2 ± 0.05	-1.62 ± 0.07
	VMD	0.77 ± 0.06	-0.55 ± 0.02	-0.43 ± 0.01	-1.27 ± 0.05	-1.63 ± 0.07
1,1-dmu	Surface Racer	1.09 ± 0.07	-0.61 ± 0.03	-0.35 ± 0.01	-1.53 ± 0.05	-2.05 ± 0.08
	VMD	1.03 ± 0.06	-0.61 ± 0.02	-0.35 ± 0.01	-1.58 ± 0.05	-2.07 ± 0.08
1,3-dmu	Surface Racer	1.68 ± 0.09	-0.77 ± 0.03	-0.56 ± 0.01	-1.87 ± 0.06	-1.97 ± 0.04
	VMD	1.66 ± 0.08	-0.79 ± 0.03	-0.57 ± 0.01	-1.99 ± 0.07	-1.98 ± 0.05
1,1-deu	Surface Racer	1.73 ± 0.09	-0.84 ± 0.03	-0.64 ± 0.01	-1.88 ± 0.07	-2.72 ± 0.1
	VMD	1.68 ± 0.08	-0.86 ± 0.03	-0.64 ± 0.01	-2 ± 0.07	-2.74 ± 0.1
1,3-deu	Surface Racer	2.35 ± 0.09	-1.01 ± 0.04	-0.76 ± 0.01	-2.28 ± 0.07	-2.99 ± 0.13
	VMD	2.28 ± 0.09	-1.03 ± 0.03	-0.76 ± 0.01	-2.42 ± 0.07	-3.01 ± 0.14

[a]Reference[22]; [b]Reference[24]
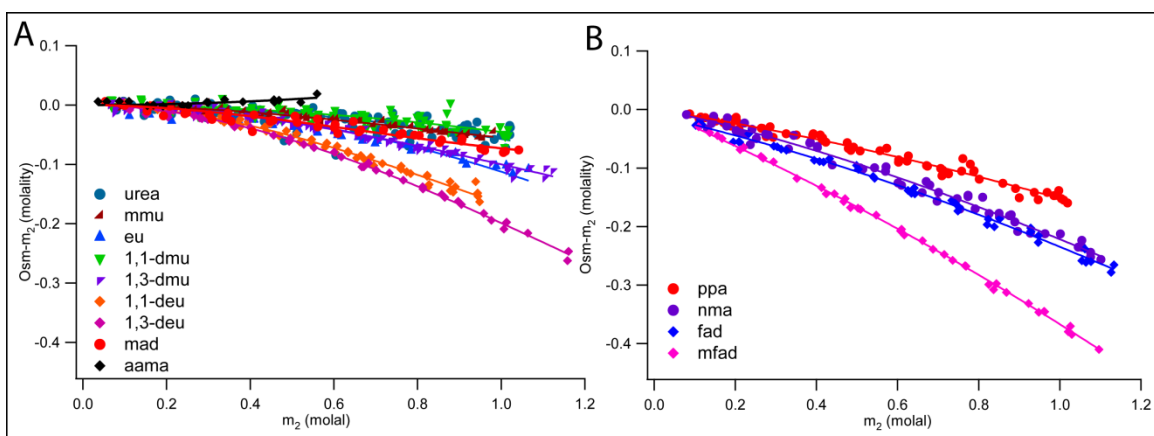
Figure S1. Excess osmolalities $\Delta Osm = Osm(m_2) - m_2$ of two component amide solutions determined by VPO at 23 ^{o}C are plotted as a function of amide molality m_2. Curves are two-parameter polynomial fits to Equation S1. Fitting coefficients β_{22} and either β_{222} (Panel A, where $\beta_2 = 1$) or β_2 (Panel B) and residuals are reported in Table S1. Abbreviations for solutes are in Table S1. (Also in Materials and Methods)
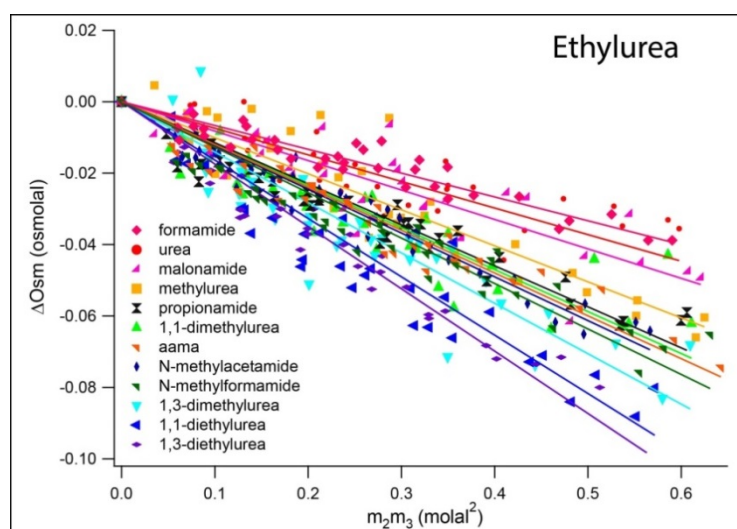
Figure S2. Interactions of Ethyl Urea with Twelve Other Amides Determined by VPO at 23 ^{o}C. In all cases the osmolality difference $\Delta Osm(m_2, m_3) = Osm\,(m_2, m_3) - Osm\,(m_2) - Osm\,(m_3)$ between a three-component solution and the corresponding two-component solutions is plotted according to Equation 1 vs. the product of molal concentrations ($m_2 m_3$) of the two amides. Slopes of linear fits with zero intercept yield chemical potential derivatives $(\partial\mu_2/\partial m_3)_{P,T,m_2} = \mu_{23}$ quantifying preferential interactions between the two amides (Equation 1).
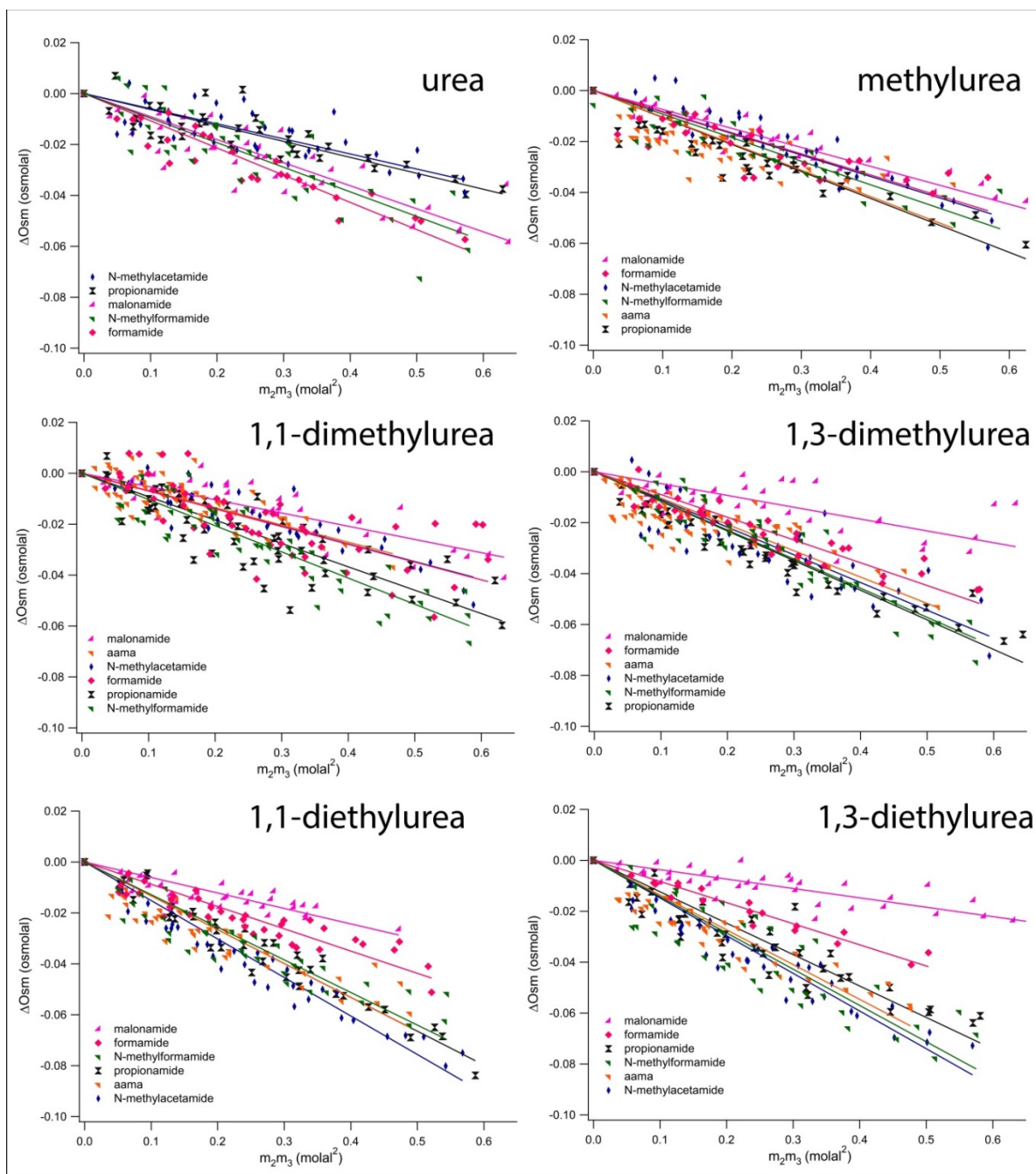
Figure S3. Interactions of Urea and Five Alkyl Ureas with other Amides Determined by VPO at 23°C. Additional VPO data for interactions of these six ureas and ethyl urea with alkyl ureas and other amides are shown in Figures 1 and S2, respectively. The osmolality difference $\Delta Osm(m_2, m_3) = Osm\,(m_2, m_3) - Osm\,(m_2) - Osm\,(m_3)$ between a three-component solution and the corresponding two-component solutions is plotted according to Equation 1 vs. the product of molal concentrations ($m_2 m_3$) of the two amides. Slopes of linear fits with zero intercept yield chemical potential derivatives $(\partial \mu_2 / \partial m_3)_{P,T,m_2} = \mu_{23}$ quantifying preferential interactions between the two amides (Equation 1).
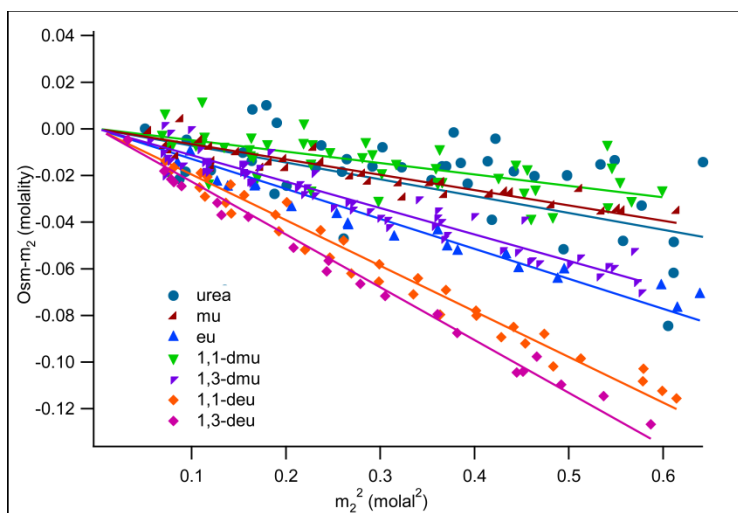
Figure S4. Amide-amide Interactions: Self-Interactions of Alkyl Ureas Determined by VPO at 23°C. Differences $\Delta Osm(m_2) = Osm\,(m_2) - m_2$ between the observed osmolality and the concentration of the alkyl urea in the two-component solution (see Figure S1A) are plotted according to Equation S7 vs. the square of its molal concentration (m_2^2). Slopes of linear fits with zero intercept yield excess chemical potential derivatives $(\partial\mu_2^{ex}/\partial m_2)_{P,T} = \mu_{22}^{ex}$ quantifying self-interactions of that alkyl urea (Equations S2, S7).
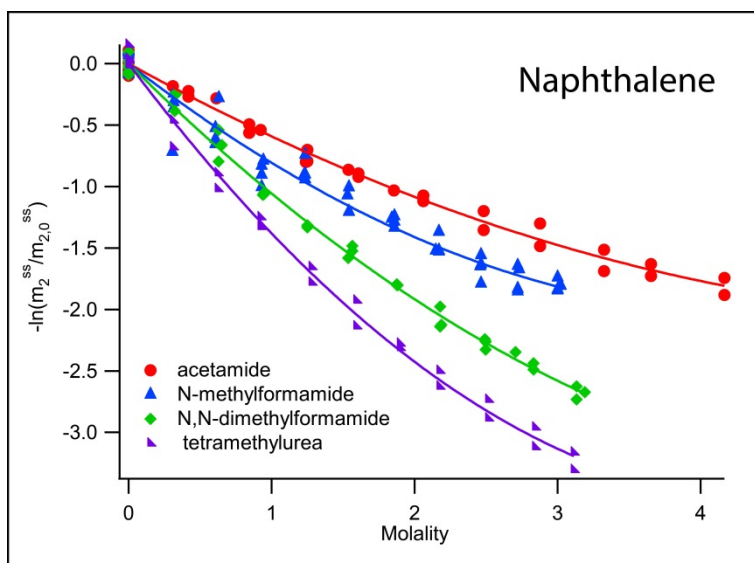
Figure S5. Interactions of Four Additional Amides with Naphthalene by Solubility Assays at 25 °C. (See Figure 3 for results for urea and alkyl ureas.) The negative logarithm of naphthalene solubility (m_2^{ss}), normalized by the extrapolated molal solubility in the absence of amide ($m_{2,0}^{ss}$), is plotted versus amide molality m_3 and fit to a quadratic equation as shown to obtain μ_{23} from the initial slope (see Equation 2).
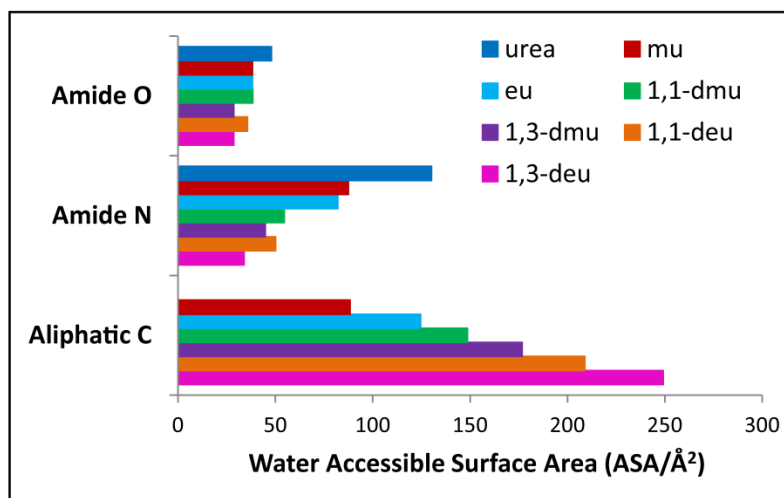
Figure S6. Water Accessible Surface Areas (ASA) of Amide sp^2O, Amide sp^2N and Aliphatic sp^3C Unified Atoms of Urea and Alkylated Ureas. (See Table S4)
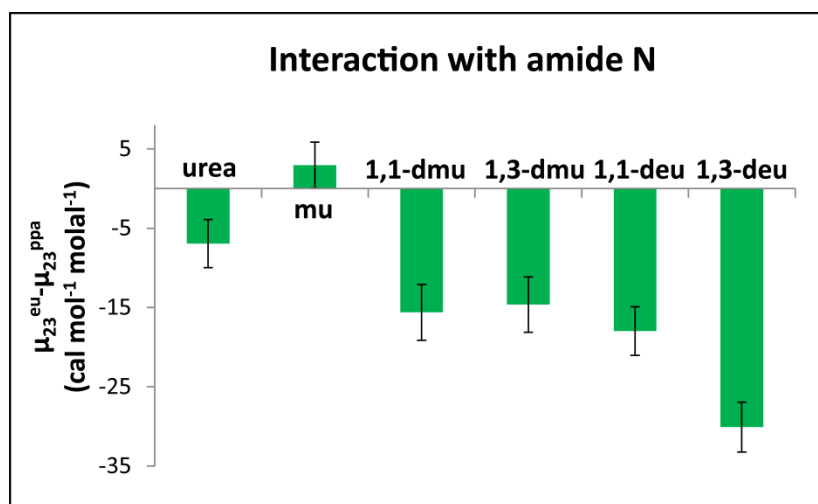
Figure S7. Additional Semi-quantitative Dissection of Interactions of Urea and the Alkyl Urea Series with Amide sp^2N. (See Figure 3C and Table S5.) Differences in μ_{23} values for interactions of the alkyl urea series with ethyl urea (eu) and proprionamide (ppa), differing primarily in amide sp^2N ASA, are shown.
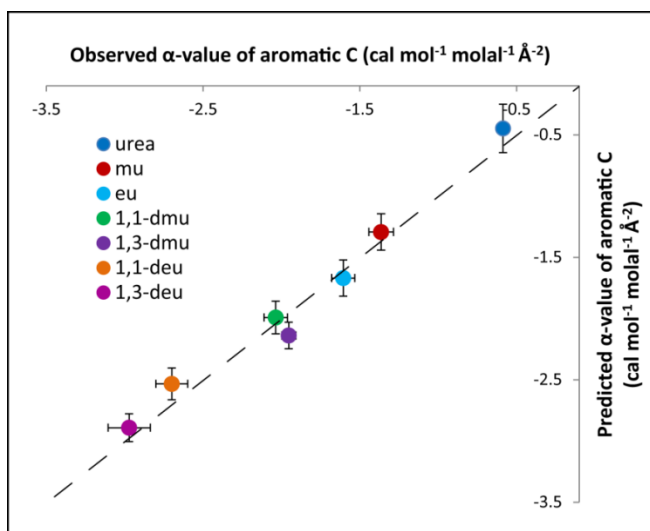
Figure S8. Observed α-Values for Interactions of Alkyl Ureas with Aromatic sp^{2}C versus Predicted α-Values using Naphthalene α-Values in Table 1.
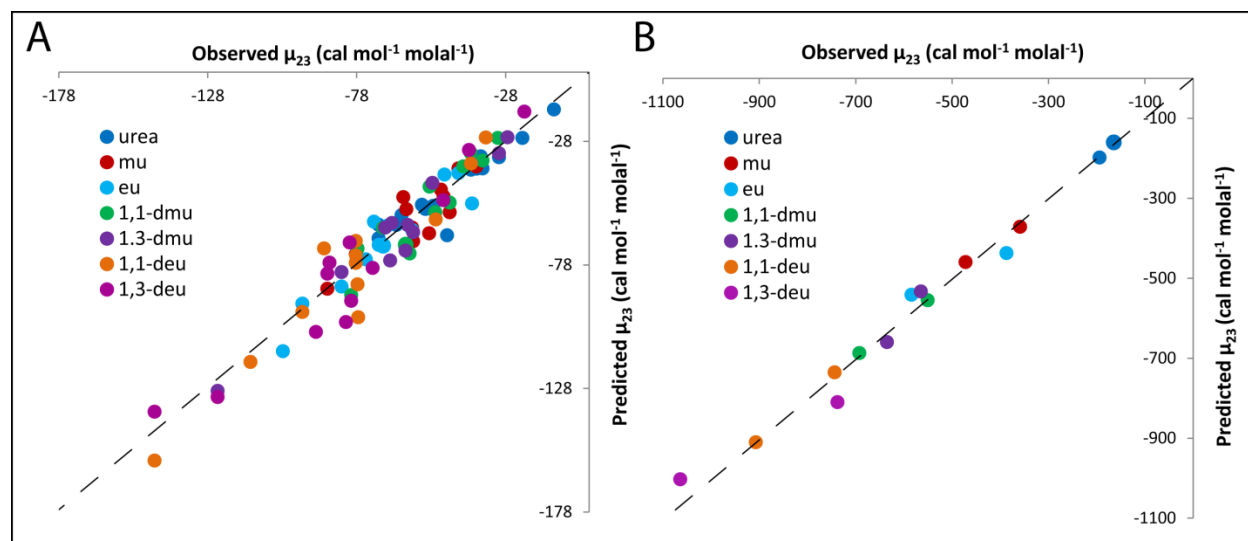
Figure S9. Predicted vs Experimental (Observed) μ_{23} Values for: A) Alkyl Urea – Alkyl Urea and Alkyl Urea- Amide Interactions at 23 ℃. B) Alkyl Urea – Aromatic Interactions at 25 ℃. Predictions use α-values for combined (amide, aromatic) sp^{2}C, aliphatic sp^3 C, amide sp^{2}N and amide sp^{2}O. (Compare with Figure 5, obtained using individual α-values for amide sp^{2}C and aromatic sp^{2}C.) All α-values used to predict μ_{23} values are listed in Table S6.
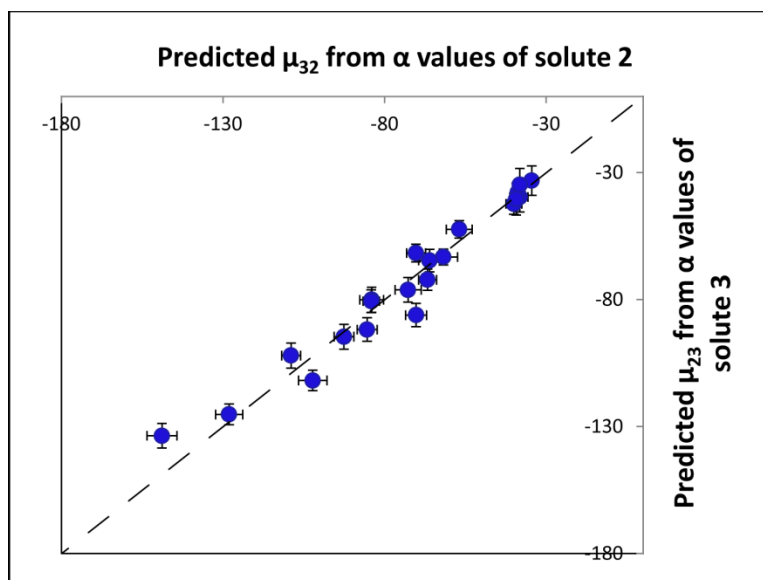
Figure S10. Comparison of Predicted μ_{23} Values for Interactions of Two Alkyl Ureas, Obtained from α-Values for Each of the Alkyl Ureas in Table 1 (see Table S8).
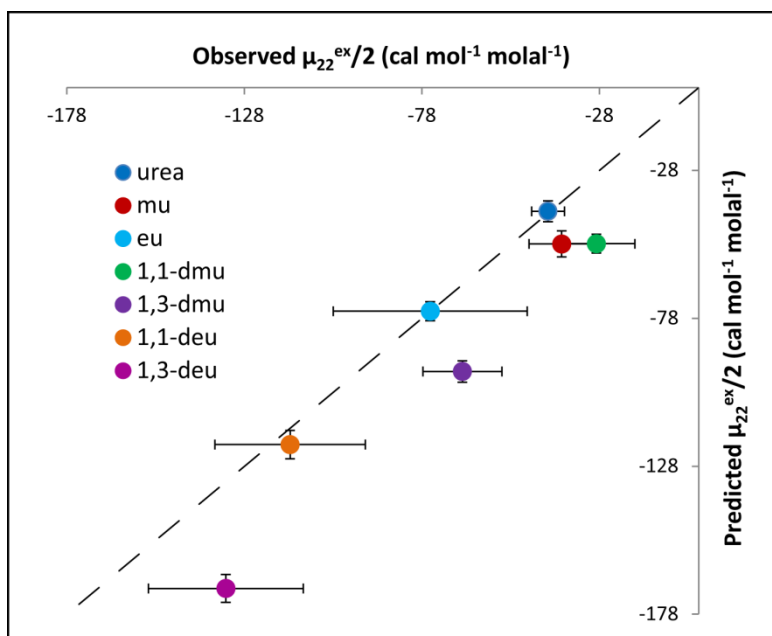
Figure S11. A) Predicted vs Observed $\mu_{22}^{ex}/2$ Values for Urea and Alkyl Ureas Determined at 23°C. Values of $\mu_{22}^{ex}/2$ were determined using linear slopes of Figure S4.

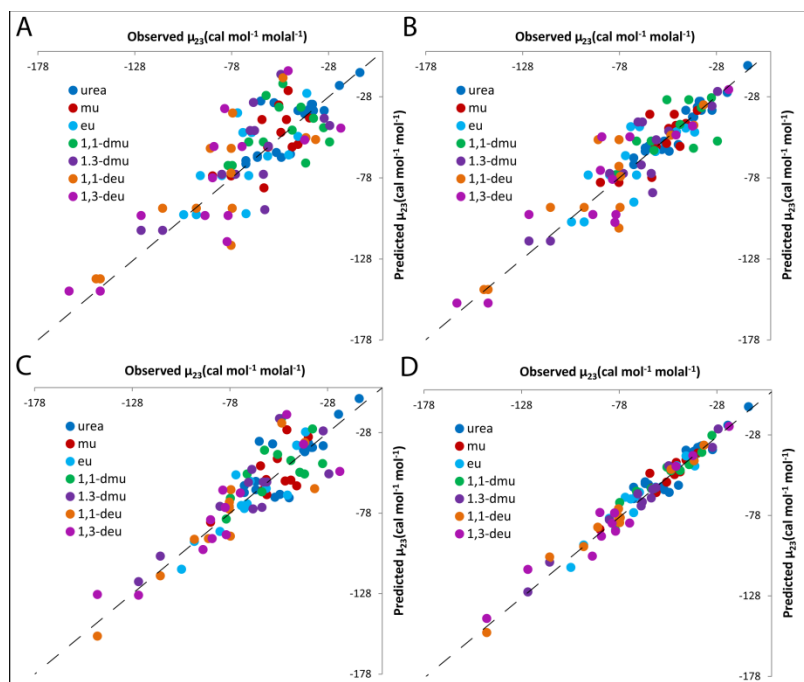**Figure S12.** Predicted vs Observed μ_{23} Values for Alkyl Urea Interactions for Different Number-based Analyses (Equation S14; Table S9) A) 3 α_n-values; 0 global weights, B) 4 α_n-values; 1 global weight, c) 3 α_n-values; 4 global weights, and d) 4 α-values; 5 global weights. See SI text for explanation of the α_n-values and global weights involved in each fit.
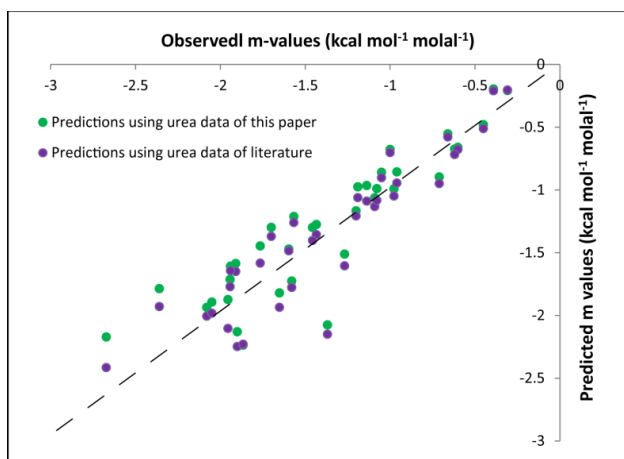
Figure S13. Comparison of predicted and experimental urea *m*-values for a series of globular proteins[14] using previously reported ΔASA values and urea α-values obtained previously or in this work (See Table 1 and Table S11). **Green**: Predicted *m*-values obtained from only four urea **α**-values (Table 1: aromatic sp^{2}C, aliphatic sp^{3}C, amide sp^{2}O and amide sp^{2}N). **Purple**: Previously-reported predictions of urea *m*-values using seven urea **α**-values (including hydroxyl O, carboxylate O and cationic N in addition to above amide and hydrocarbon unified atoms).   Amide sp^{2}C represents less than 1% of the ΔASA of unfolding and was not accounted for in these comparisons.
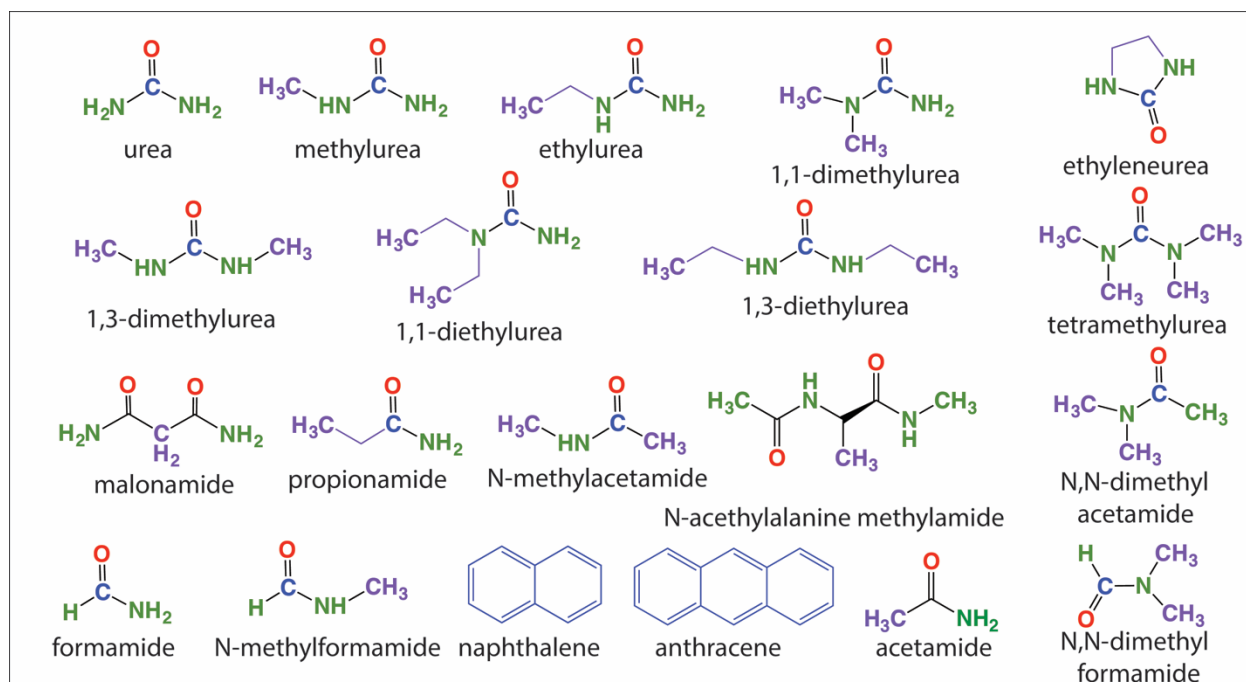
Figure S14. Color-coded unified atoms of solutes (Table S4). Red: amide O, Green: amide N; Purple: Aliphatic sp^{3}C; Blue: amide and aromatic sp^{2}C.

References:

(1) Cheng, X.; Guinn, E. J.; Buechel, E.; Wong, R.; Sengupta, R.; Shkel, I. A.; Record, M. T., *Biophys. J.* **2016,** *111*, 1854-1865.

(2) Archer, D. G., *J. Phys. Chem. Ref. Data* **1992,** *21*, 793-829.

(3) Schonert, H.; Stroth, L., *Biopolymers* **1981,** *20*, 817-831.

(4) Cannon, J. G.; Anderson, C. F.; Record, M. T., Jr., *J. Phys. Chem. B* **2007,** *111*, 9675-9685.

(5) Barone, G.; Rizzo, E.; Volpe, V., *J. Chem. Eng. Data* **1976,** *21*, 59-61.

(6) Ellerton, H. D., Dunlop, P. J. , *J. Phys. Chem.* **1966,** *70*, 7.

(7) Anderson, C. F.; Record, M. T., Jr., *Biophys. Chem.* **2004,** *112*, 165-175.

(8) Robinson, R. A. S., R. H., *J. Phys. Chem.* **1961,** *65*, 1954-1958.

(9) Herskovits, T. T.; Bowen, J. J., *Biochemistry* **1974,** *13*, 5474-5483.

(10) Nandi, P. K.; Robinson, D. R., *J. Am. Chem. Soc.* **1972,** *94*, 1308-1315.

(11) Roseman, M.; Jencks, W. P., *J. Am. Chem. Soc.* **1975,** *97*, 631-640.

(12) Capp, M. W.; Pegram, L. M.; Saecker, R. M.; Kratz, M.; Riccardi, D.; Wendorff, T.; Cannon, J. G.; Record, M. T., Jr., *Biochemistry* **2009,** *48*, 10372-10379.

(13) Diehl, R. C.; Guinn, E. J.; Capp, M. W.; Tsodikov, O. V.; Record, M. T., Jr., *Biochemistry* **2013,** *52*, 5997-6010.

(14) Guinn, E. J.; Pegram, L. M.; Capp, M. W.; Pollock, M. N.; Record, M. T., Jr., *Proc. Natl. Acad. Sci. U.S.A.* **2011,** *108*, 16932-16937.

(15) Guinn, E. J.; Kontur, W. S.; Tsodikov, O. V.; Shkel, I.; Record, M. T., Jr., *Proc. Natl. Acad. Sci. U.S.A.* **2013,** *110*, 16784-16789.

(16) Knowles, D. B.; Shkel, I. A.; Phan, N. M.; Sternke, M.; Lingeman, E.; Cheng, X.; Cheng, L. X.; O'Connor, K.; Record, M. T., *Biochemistry* **2015,** *54*, 3528-3542.

(17) Guinn, E. J.; Schwinefus, J. J.; Cha, H. K.; McDevitt, J. L.; Merker, W. E.; Ritzer, R.; Muth, G. W.; Engelsgjerd, S. W.; Mangold, K. E.; Thompson, P. J.; Kerins, M. J.; Record, M. T., Jr., *J. Am. Chem. Soc.* **2013,** *135*, 5828-5838.

(18) Pegram, L. M.; Record, M. T., Jr., *J. Phys. Chem. B* **2008,** *112*, 9428-9436.

(19) Sengupta, R.; Pantel, A.; Cheng, X.; Shkel, I.; Peran, I.; Stenzoski, N.; Raleigh, D. P.; Record, M. T., Jr., *Biochemistry* **2016,** *55*, 2251-2259.

(20) Livingstone, J. R.; Spolar, R. S.; Record, M. T., Jr., *Biochemistry* **1991,** *30*, 4237-4244.

(21) Spolar, R. S.; Livingstone, J. R.; Record, M. T., Jr., *Biochemistry* **1992,** *31*, 3947-3955.

(22) Tsodikov, O. V.; Record, M. T., Jr.; Sergeev, Y. V., *J. Comput. Chem.* **2002,** *23*, 600-609.

(23) Fraczkiewicz, R.; Braun, W., *J. Comput. Chem.* **1998,** *19*, 319-333.

(24) Humphrey, W.; Dalke, A.; Schulten, K., *J. Mol. Graph.* **1996,** *14*, 33-38, 27-38.

(25) Pall, S.; Abraham, M. J.; Kutzner, C.; Hess, B.; Lindahl, E., *Lect. Notes Comput. Sc.* **2015,** *8759*, 3-27.

(26) Okamoto, B. Y.; Wood, R. H.; Thompson, P. T., *J. Chem. Soci. Faraday Trans. I* **1978,** *74*, 1990-2007.

(27) Sadowski, J.; Gasteiger, J.; Klebe, G., *J. Chem. Inf. Comput. Sci.* **1994,** *34*, 1000-1008.

(28) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C., *Acta Crystallogr. Sect. B* **2016,** *72*, 171-179.

(29) Bolton, E. E.; Chen, J.; Kim, S.; Han, L. Y.; He, S. Q.; Shi, W. Y.; Simonyan, V.; Sun, Y.; Thiessen, P. A.; Wang, J. Y.; Yu, B.; Zhang, J.; Bryant, S. H., *J. Cheminform.* **2011,** *3*.

(30) Halgren, T. A., *J. Comput. Chem.* **1996,** *17*, 520-552.

(31) Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Wenger, R. K.; Yao, H. Y.; Markley, J. L., *Nucleic Acids Res.* **2008,** *36*, D402-D408.

(32) Belsley, D. A.; Kuh, E., *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* Wiley: New York, 1980.

(33) Ganguly, P.; van der Vegt, N. F., *J. Chem. Theory Comput.* **2013,** *9*, 5247-5256.

(34) Lin, B.; Lopes, P. E.; Roux, B.; MacKerell, A. D., Jr., *J. Chem. Phys.* **2013,** *139*, 084509.