10372 *Biochemistry* **2009,** *48,* 10372–10379
DOI: 10.1021/bi901273r

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Interactions of the Osmolyte Glycine Betaine with Molecular Surfaces in Water: Thermodynamics, Structural Interpretation, and Prediction of *m*-Values[†]

Michael W. Capp,[⊥] Laurel M. Pegram,[∥] Ruth M. Saecker,[⊥] Megan Kratz,[∥,‡] Demian Riccardi,[∥] Timothy Wendorff,[∥] Jonathan G. Cannon,[¶,§] and M. Thomas Record, Jr.*[,∥,⊥,¶]

Departments of [∥] Biochemistry and [⊥] Chemistry and [¶] Biophysics Program, University of Wisconsin, Madison, Wisconsin 53706.
[‡] Current address: Curriculum in Neurobiology, University of North Carolina, Chapel Hill, NC 27599.
[§] Current address: Department of Physiology and Biophysics, Case Western University, Cleveland, OH 44106.

Received July 24, 2009; Revised Manuscript Received September 16, 2009

ABSTRACT: Noncovalent self-assembly of biopolymers is driven by molecular interactions between functional groups on complementary biopolymer surfaces, replacing interactions with water. Since individually these interactions are comparable in strength to interactions with water, they have been difficult to quantify. Solutes (osmolytes, denaturants) exert often large effects on these self-assembly interactions, determined in sign and magnitude by how well the solute competes with water to interact with the relevant biopolymer surfaces. Here, an osmometric method and a water-accessible surface area (*ASA*) analysis are developed to quantify and interpret the interactions of the remarkable osmolyte glycine betaine (GB) with molecular surfaces in water. We find that GB, lacking hydrogen bond donors, is unable to compete with water to interact with anionic and amide oxygens; this explains its effectiveness as an osmolyte in the *Escherichia coli* cytoplasm. GB competes effectively with water to interact with amide and cationic nitrogens (hydrogen bonding) and especially with aromatic hydrocarbon (cation−π). The large stabilizing effect of GB on lac repressor−lac operator binding is predicted quantitatively from ASA information and shown to result largely from dehydration of anionic DNA phosphate oxygens in the protein−DNA interface. The incorporation of these results into theoretical and computational analyses will likely improve the ability to accurately model intra- and interprotein interactions. Additionally, these results pave the way for development of solutes as kinetic/mechanistic and thermodynamic probes of conformational changes and formation/disruption of molecular interfaces that occur in the steps of biomolecular self-assembly processes.

Biopolymer self-assembly (folding, binding) in vivo and in vitro involves the replacement of interactions with water by more favorable interactions between biopolymer functional groups (*1*). The ability (or inability) of solutes and Hofmeister salt ions to compete with water to interact with biopolymer functional groups results in often large destabilizing (or stabilizing) effects on these assembled states (*2*, *3*). To understand the energetics of self-assembly and how solutes modulate these processes, the strength of interactions of functional groups with water relative to the strength of their interactions with one another must be determined (*4*). To accomplish this, new methodologies and analyses are required. Here, we use GB[1] as proof of principle of a method and analysis that allow us to systematically obtain fundamental, previously unavailable, information about hydrogen bonding, ion pairing, and other interactions between biomolecular functional groups.

Glycine betaine (*N*,*N*,*N*-trimethylglycine) is the most effective *Escherichia coli* osmolyte characterized to date, allowing the cell to efficiently retain intracellular water, maintain cytoplasmic volume and dilution of cellular biopolymers, and therefore grow well under dehydrating conditions (*5*). Many bacterial pathogens accumulate cytoplasmic GB to adapt to osmotic stress, increasing their growth rate and thus affecting colonization and infectivity (*5*, *6*). In vitro, GB drives self-assembly, strongly stabilizing site-specific protein−DNA complexes (e.g., refs (*7*−*10*)), moderately stabilizes the globular (folded) conformation of protein (*11*, *12*), and promotes some (but not all) tertiary interactions in folded RNA (*13*). GB has little if any effect on stability of all-AT DNA duplexes but destabilizes GC-containing nucleic acid duplexes (*13*, *14*). To explain various of these effects, GB is proposed to be excluded from anionic oxygens (phosphate, carboxylate) (*15*, *16*), the peptide backbone (*17*), and hydrocarbon groups (*18*).

Effects of solutes like urea and GB (and non-Coulombic Hofmeister effects of salts) on biopolymer processes like folding and binding are quantified by *m*-values, defined as derivatives with respect to solute or salt concentration (m_3) of the observed standard free energy change of the process $\Delta G^\circ_{obs} = -RT \ln K_{obs}$, where K_{obs} is the observed equilibrium concentration quotient (expressed in terms of concentrations and not thermodynamic activities):

$$m\text{-value} = \frac{d\,\Delta G^\circ_{obs}}{d\,m_3} = -RT\frac{d \ln K_{obs}}{d\,m_3} = RT\frac{d \ln K_\gamma}{d\,m_3} = \Delta\mu_{23} \quad (1)$$

In eq 1, K_γ is the quotient of biopolymer activity coefficients corresponding to the concentration quotient K_{obs} and

[†] This work was supported by National Institutes of Health Grants GM47022 and GM23467 (to M.T.R.).
*To whom correspondence should be addressed. E-mail: mtrecord@wisc.edu. Phone: 608-262-5332. Fax: 608-262-3453.
[1] Abbreviations: GB, glycine betaine; VPO, vapor pressure osmometry; ASA, accessible surface area; SPM, solute partitioning model; DBD, DNA binding domain; BSA, bovine serum albumin; HEWL, hen egg white lysozyme.

 Published on Web 09/16/2009

$\mu_{23} = RT\ \mathrm{d}\ln\gamma_2/\mathrm{d}\ m_3$; μ_{23} is closely related to the preferential interaction coefficient $\Gamma_{\mu 3}$ (*2*, *3*), which can be predicted if the solute distribution near the biopolymer is known (*19*, *20*). (Throughout, subscripts 1, 2, and 3 refer to water, biopolymer or model compound, and glycine betaine, respectively.)

Lee and Richards pioneered the calculation of water-accessible surface area (ASA) of proteins and model compounds and initiated the use of ASA in analyses of protein stability (*21*). Subsequently, ASA-based analyses were used to interpret the thermodynamics of protein folding and other protein processes and to predict coupled folding vs rigid body binding (*22*−*24*). The first systematic quantitative application of ASA to analyze solute effects on protein processes was the landmark study of Myers, Pace, and Scholtz (*25*), wherein the magnitudes of urea and GuHCl *m*-values for globular protein unfolding were observed to be proportional to the ΔASA of unfolding, calculated assuming an extended chain model for the unfolded state. Urea and GuH^+ *m*-value/ΔASA ratios for unfolding of α-helical peptides are 3−4-fold larger than for globular proteins, which correlates with the 3−4-fold greater fraction of amide surface in the ΔASA of unfolding α-helical peptides as compared to globular proteins (*26*) and demonstrates that the primary interaction of these denaturants is with amide surface (*8*, *26*).

In the present study, we develop the use of vapor pressure osmometry (VPO) to quantify the thermodynamics of interaction of a solute like GB with model compounds that display the principal functional groups of proteins and nucleic acids. The osmolality Osm of a solution, which may be thought of as an effective (nonideality-corrected) total solute concentration, is directly related to the water activity ($\mathrm{Osm} = -55.5 \ln a_1$). The excess osmolality of a three-component solution (ΔOsm) is a quantitative, rigorous measure of the favorable or unfavorable interaction of the two nonelectrolyte solute components, relative to their interactions with water:

$$\Delta\mathrm{Osm} = \mathrm{Osm}(m_2, m_3) - \mathrm{Osm}(m_2, 0) - \mathrm{Osm}(0, m_3) \qquad (2)$$

Experimental values of ΔOsm yield the chemical potential derivatives μ_{23} which comprise the *m*-value (*27*−*30*). If μ_{23} is independent of m_2 and m_3, then

$$\Delta\mathrm{Osm} \cong \left(\frac{\mu_{23}}{RT}\right) m_2 m_3 \qquad (3)$$

Hence, μ_{23}/RT is the slope of a plot of ΔOsm, obtained from eq 3 for given choices of molal concentrations m_2 and m_3, vs the m_2m_3 product. Here, we experimentally determine interactions of GB with model compounds containing the functional groups of biopolymers, interpret these as interactions of GB with the water-accessible areas of the major different types of hydrocarbon, nitrogen, and oxygen surfaces of biopolymers, and apply these values to predict GB interactions and GB effects on processes, using ASA information.

MATERIALS AND METHODS

Chemicals. Nitrilotriacetic acid trisodium salt monohydrate (>98%), glycine betaine monohydrate (>99%), potassium acetate (>99%), urea, mannitol, lysine and arginine hydrochlorides, potassium oxalate monohydrate, and sodium benzoate (all >99.5%) were from Fluka. Sodium aspartate, sodium oxamate, potassium citrate tribasic monohydrate (all >98%), glycine, diglycine, sodium chloride, and potassium glutamate (all >99%) were from Sigma. Sodium chloride (>99%), dibasic potassium and sodium phosphates (>99%), sodium acetate (>99.5%), and sucrose (>99.9%) were from Thermo Fisher Scientific. Glycerol (>99.5%) was from Aldrich. Acetyl-Ala-methylamide (>99%) was from Bachem. All samples were dissolved in water purified with a Barnstead E-pure system (Thermo Fisher Scientific).

Vapor Pressure Osmometry. Samples were prepared using all-gravimetric methods in plastic, capped microfuge tubes. Typically, ∼30−250 mg of the model compound (e.g., carboxylate salt) was added as the solid to a preweighed tube, and its weight was accurately determined using an analytical balance (Mettler). Approximately 1 mL of a glycine betaine solution of precisely known concentration (gravimetrically prepared on a larger scale) was added, and the tube was weighed again. Samples were mixed until the solid was completely dissolved and then stored at 4 °C until reading, no more than 24 h later. In some cases, separate stock solutions of the model compound and of GB were prepared gravimetrically. Since the purer form of commercially available GB is the monohydrate, these experiments were most conveniently performed as series in which the GB molality was held constant and the molality of the model compound was varied. To hold the concentration of the model compound constant in a series of experiments at different GB concentrations, the model compound solution was added first and its weight determined. Weights of GB solution and of water required to achieve a constant concentration of the model compound were calculated and added. To avoid addition of a fourth component as well as protonation of carboxylates of model compounds, no buffer component was added, and solution pH (ranging from 5 to 11) was not adjusted. In no case was the pH of the model compound solution sufficiently acidic to protonate the carboxylate group of GB.

Osmometry was performed using a Wescor Vapro 5520 vapor pressure osmometer under conditions of controlled humidity. These experiments typically spanned the range of GB concentrations from 0.4 to 1.2 *m* and model compound concentrations from 0.05 to 1.0 *m* for tricarboxylate salts and 0.1 to 1.25 *m* for monocarboxylate salts. Triplicate measurements of osmolality were performed on each sample. The thermocouple was cleaned extensively before (and often during) assays using 2 M ammonium hydroxide and filtered deionized water (Barnstead E-Pure) and then calibrated extensively until stable readings were obtained. The calibration was checked frequently during the assay. Standard solutions of NaCl at osmolalities of 0.100, 0.290, and 1.000 osmolal for use in calibration of the osmometer were prepared gravimetrically at the appropriate molalities (0.05351, 0.1567, and 0.5422 *m* NaCl, respectively), calculated from isopiestic distillation (ID) data (*31*). Use of these standards significantly improved the agreement between our VPO and literature ID data for NaCl, KCl, GB, and urea. Above 1 osmolal, the highest calibration osmolality, small systematic differences between VPO and literature ID data for NaCl were observed. Fitted values of these differences were used to correct experimental VPO data obtained for 1−3 osmolal solutions measured on the same osmometer.

Surface Area Calculations. Various online resources were utilized to generate PDB files of the model compounds: PubChem (http://pubchem.ncbi.nlm.nih.gov), ChemDB (http://cdb.ics.uci.edu/), Biological Magnetic Resonance Data Bank (http://www.bmrb.wisc.edu), and the SMILES (simplified molecular input line entry specification) translator at http://cactus.nci.nih.gov/translate. Water-accessible surface areas (ASA) are calculated using Surface Racer (*32*) with the Richards' set

of van der Waals radii (*33*) and a 1.4 Å probe radius for water. Each ASA is divided into contributions from eight coarse-grained surface types (Supporting Information Tables S1 and S2): aliphatic carbon, aromatic carbon, hydroxyl oxygen, amide oxygen, anionic carboxylate oxygen, anionic phosphate oxygen, amide nitrogen, and cationic nitrogen. A unified atom model is used, wherein hydrogens which are covalently bonded to these atoms are treated as part of the atom in calculating its van der Waals radius. For the lac DBD, the 20 conformers of PDB (*34*) entry 1OSL (*35*) were used as the model of the folded state; averages for total ASA and composition were calculated (as above) using the first 51 residues of each headpiece monomer (40 conformers). The web application ProtSA (*36*, *37*) was used to generate an unfolded ensemble for these 51 residues, and water-accessible surface areas were calculated for 1919 conformations. The resulting average ΔASA composition, as well as details for other biopolymer ASA and ΔASA calculations, is presented in Supporting Information Table S2.

Data Analysis. To obtain values of μ_{23}/RT for interaction of GB with model compounds, values of ΔOsm (eq 2) were plotted as a function of the m_2m_3 product. No evidence for significant concentration dependence of μ_{23}/RT or deviation from proportionality of ΔOsm to m_2m_3 is observed for interactions of GB with the nonelectrolytes, zwitterions, and salts studied. For all model compounds investigated, these data are well fit by a line with fixed zero intercept, as predicted for nonelectrolytes by eq 3. If the intercept is floated, small nonzero intercepts are observed, but these deviations are not systematic and do not have a significant effect on the slopes. Values of μ_{23}/RT and corresponding uncertainties were obtained from the slopes of plots of ΔOsm vs m_2m_3 with the intercept fixed at zero. Values of μ_{23}/RT for interactions of GB with model compounds (with ASA and surface composition calculated as above) were analyzed by Igor 5.04B (multiple linear regression) to obtain values of GB interaction potentials $(\mu_{23}/RT\text{ASA})_i$ for different surface types.

Repressor−Operator Binding Studies. Effects of GB on binding lac repressor to SymL operator (40 bp) DNA were determined at 25 °C and 0.40 M K^+ by nitrocellulose filter binding of equilibrium mixtures of repressor (0.5 nM) and operator (0.1 nM) at GB concentrations from 0 to 1 *m* (*8*, *38*). Binding constants are averages of four independent experiments, each with duplicate samples. The predicted contribution from burial of DNA phosphate O ASA in the repressor−operator interface (Figure 3) is reduced by 0.25 m^{-1} to correct for the observed increase in KCl activity with increasing GB concentration (Table 1).

RESULTS

Osmometric Quantification of Interactions of GB with Model Compounds. We used osmometry to quantify the interactions of GB with 23 model compounds containing carboxylate, phosphate, amide, hydroxyl, ammonium, guanidinium, and aliphatic and aromatic hydrocarbon moieties. (Literature solubility data for four cyclic dipeptides as a function of GB concentration (0−4 M) (*39*) were analyzed as well.) Values of ΔOsm quantifying interactions of GB with model compounds investigated to date are plotted as a function of the m_2m_3 product in Figure 1. In all cases, these plots are linear over the concentration ranges examined, demonstrating that μ_{23}/RT for each GB interaction is independent of concentration. Values of μ_{23}/RT

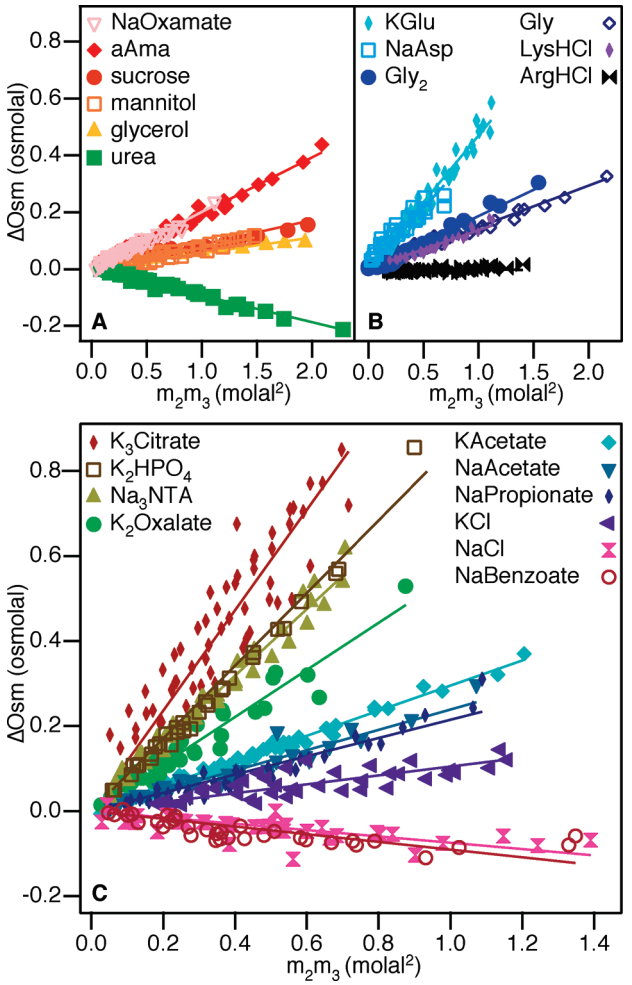

FIGURE 1: Excess osmolality ΔOsm from VPO studies of GB−model compound interactions plotted against m_2m_3, the product of molal concentrations of model compound and GB; the slope is μ_{23}/RT (eq 3).

obtained from the slopes (or solubility data (*40*)) are listed in Table 1. From the plots of Figure 1, we observe the following:

(1) The interactions of GB with the primary amide urea and the secondary diamide acetyl-Ala-methylamide (aAma) are of opposite sign (Figure 1A). GB interacts favorably with urea, for which the N/O ASA ratio is 2.7, but unfavorably with aAma, for which the N/O ASA ratio is 0.3. The simplest interpretation of these data is that GB must interact favorably with (i.e., accumulate in the vicinity of) the amide nitrogen but interact unfavorably with (i.e., be excluded from) the amide oxygen.

(2) The observed small, positive values of ΔOsm and of μ_{23}/RT (Figure 1A) for the interactions of GB with the polyols glycerol (C_3), mannitol (C_6), and sucrose (C_{12}) are most simply explained as weak exclusion of GB from hydroxyl oxygen and aliphatic carbon surface.

(3) Highly positive values of ΔOsm and of μ_{23}/RT are obtained for the interactions of GB with phosphate and carboxylate salts (Figure 1C). In the series potassium acetate, dipotassium oxalate, tripotassium, as the number of carboxylate groups (and K^+ ions) increases from one to three, μ_{23}/RT increases. The simplest interpretation of these data is that GB is strongly excluded from anionic oxygens.

(4) The less unfavorable interactions of GB with the amino acids Gly, LysHCl, and especially ArgHCl (Figure 1B) than with acetate salts (Figure 1C) indicate that interactions of GB with

Table 1: VPO Values of μ_{23}/RT for Interactions of GB with Model Compounds

solute	exptl μ_{23}/RT (m^{-1}) ± SD[a]	calcd μ_{23}/RT (m^{-1}) ± SD
glycerol	0.056 ± 0.003	0.052 ± 0.024
mannitol	0.075 ± 0.004	0.065 ± 0.034
sucrose	0.086 ± 0.004	0.095 ± 0.046
urea	−0.093 ± 0.005	−0.094 ± 0.046
cGG[b]	0.19 ± 0.01	0.20 ± 0.05
cAG[b]	0.20 ± 0.03	0.21 ± 0.05
aAma	0.20 ± 0.01	0.23 ± 0.05
cAA[b]	0.26 ± 0.01	0.22 ± 0.05
cAL[b]	0.26 ± 0.02	0.24 ± 0.06
diglycine	0.19 ± 0.02	0.26 ± 0.04
sodium oxamate	0.20 ± 0.01	0.27 ± 0.03
arginine hydrochloride	−0.0040 ± 0.0020	−0.0038 ± 0.0871
lysine hydrochloride	0.14 ± 0.01	0.062 ± 0.064
glycine	0.15 ± 0.01	0.18 ± 0.04
sodium aspartate	0.43 ± 0.02	0.39 ± 0.04
potassium glutamate	0.47 ± 0.01	0.52 ± 0.04
sodium benzoate	−0.091 ± 0.007	−0.10 ± 0.08
sodium propionate	0.22 ± 0.01	0.30 ± 0.03
sodium acetate	0.24 ± 0.01	0.29 ± 0.02
potassium acetate	0.30 ± 0.02	0.38 ± 0.02
dipotassium oxalate	0.55 ± 0.03	0.71 ± 0.04
Na_3NTA	0.78 ± 0.04	0.71 ± 0.05
Na_2HPO_4	0.82 ± 0.04	0.74 ± 0.06
K_2HPO_4	0.85 ± 0.04	0.92 ± 0.06
tripotassium citrate	1.2 ± 0.1	0.97 ± 0.05
KCl	0.048 ± 0.002	0.060
NaCl	−0.074 ± 0.006	−0.030

[a]Error estimates provided are the larger of 5% or the estimated standard deviation determined from the fit residuals by Igor Pro. [b]Values of μ_{23}/RT for cyclic dipeptides (cycloGlyGly, cycloAlaGly, cycloAlaAla, and cycloAlaLeu) were determined from literature solubility data (*39*), as performed in ref *40*.

cationic ammonium nitrogens (one on Gly and ArgHCl, two on LysHCl) and guanidinium nitrogens (three on ArgHCl) are favorable.

(5) The modest net favorable interaction of GB with sodium benzoate (Figure 1C) indicates a strongly favorable interaction of GB with the benzyl ring, which compensates for the unfavorable interaction of GB with the carboxylate moiety.

Interpretation of GB Interactions (μ_{23}) using Accessible Surface Area (ASA) and the Solute Partitioning Model (SPM). Eight coarse-grained classes of surface were considered in the analysis of the osmometric data: aliphatic carbon, aromatic carbon, hydroxyl oxygen, amide oxygen, anionic carboxylate, phosphate oxygen, amide nitrogen, and cationic nitrogen. For salts, contributions of the interactions of GB with K^+, Na^+, or Cl^- ions were also included. (Supporting Information Table S1 lists amounts of each type of surface for all model compounds investigated here, calculated as described in Materials and Methods.) Therefore, as a first level of interpretation, we dissect experimental values of $\Delta\mu_{23}/RT$ or μ_{23}/RT (see eqs 1 and 3) into additive contributions from chemically distinct, coarse-grained surface types (*15*, *30*, *40*). This is analogous to the approach of Tanford (*41*) and Bolen (*42*), which assumes that a solute *m*-value for protein unfolding can be decomposed into additive contributions from the 20 side chains and the peptide backbone units exposed in unfolding. We propose that the contribution of each type of surface (i) to μ_{23}/RT is the product of a solute interaction potential (contribution per unit of ASA; $(\mu_{23}/RT\text{ASA})_i$) and the ASA of that surface. The experimental value of the chemical potential derivative μ_{23}/RT is therefore represented as the sum of terms:

$$\frac{\mu_{23}}{RT} = \sum_i \left(\frac{\mu_{23}}{RT\text{ASA}}\right)_i (\text{ASA})_i + \nu_j \left(\frac{\mu_{23}}{RT}\right)_j \qquad (4)$$

where the interaction potential $(\mu_{23}/RT\text{ASA})_i$ quantifies the interaction of the solute of interest with 1 Å^2 of surface of type i on any compound or biopolymer, $(\text{ASA})_i$ is the water-accessible area in Å^2 of surface type i on the model compound being analyzed, and $\nu_j(\mu_{23}/RT)_j$ is the product of the number of salt ions (ν_j) per formula unit of a salt and the assigned contribution $(\mu_{23}/RT)_j$ of that type of ion to μ_{23}/RT. The observed μ_{23}/RT are model-independent thermodynamic quantities; the solute potentials $(\mu_{23}/RT\text{ASA})_i$, which quantify the effect of the solute per unit area of a particular type of water-accessible surface on the biomolecule or model compound, require a structural model.

We use an ASA-based analysis of coarse-grained surface types instead of a functional group or atom-by-atom analysis for two reasons. First, the ASA of a particular type of surface takes account of variations in the accessibility of different functional groups or atoms resulting from the global conformation of the molecule and/or local steric effects of neighboring atoms or groups. Second, ASA is a fundamental variable in the solute partitioning model (SPM) of preferential interactions, which proposes that the hydration of a particular type of surface is proportional to its ASA. An SPM-based molecular thermodynamic analysis, using ASA, has been successfully applied to experimental data characterizing the effects of the spectrum of Hofmeister salts on the process of forming a nonpolar air−water surface (*43*) and on processes which expose hydrocarbon and amide molecular surface to water (*40*).

Each solute interaction potential is readily interpreted using the SPM (*40*); for a nonelectrolyte solute at low concentration:

$$\left(\frac{\mu_{23}}{RT\text{ASA}}\right)_i = -\frac{(K_{p,i} - 1)b_1}{55.5} \qquad (5)$$

In eq 5, $K_p = m_3^{\text{local}}/m_3^{\text{bulk}}$ is the microscopic analogue of a macroscopic thermodynamic partition coefficient (equilibrium concentration quotient) characterizing the distribution of a solute like GB or urea between the local water of hydration of a given type of biopolymer or model compound surface and bulk water, and b_1 is the surface density of the local water.

Initially, values of μ_{23}/RT for interactions of GB with nine uncharged solutes (six amides, three polyols; cf. Figure 1A and Table 1) were analyzed using eq 4 and ASA compositions to obtain values of reduced GB interaction potentials (μ_{23}/RTASA) for amide O, amide N, hydroxyl O, and hydrocarbon C surface types (Table 2). Interactions of GB with these solutes were analyzed first because they are uncharged and so may interact more simply with the GB zwitterion than other model compounds in the data set, which are zwitterions or salts. Confirming the qualitative conclusions presented above, the results show that GB interacts unfavorably with amide O but favorably with amide N. Unfavorable interactions of GB with aliphatic C and hydroxyl O are smaller in magnitude.

These GB interaction potentials were held constant in the application of eq 4 to analyze interactions of GB with the zwitterions and salts in Figure 1B,C and Table 1; using the ASA information of Supporting Information Table S1, we obtained reduced GB interaction potentials μ_{23}/RTASA for carboxylate O,

Table 2: GB−Surface Interaction Potentials and Corresponding SPM b_1 and K_p Values

		SPM (eq 5)	
surface type	$\mu_{23}/(RT\text{ASA})$ (m^{-1} Å^{-2})	b_1 (H_2O Å^{-2})	K_p
phosphate O	$(4.9 \pm 0.4) \times 10^{-3}$	0.27 ± 0.02	0^a
carboxylate O	$(2.8 \pm 0.2) \times 10^{-3}$	0.18 ± 0.01^b	0.14 ± 0.08
amide O	$(2.6 \pm 0.5) \times 10^{-3}$	0.18 ± 0.01^b	0.20 ± 0.16
aliphatic C	$(3.5 \pm 1.6) \times 10^{-4}$	0.18 ± 0.01^c	0.89 ± 0.05
hydroxyl O	$(0.9 \pm 1.2) \times 10^{-4}$	0.18 ± 0.01^b	0.97 ± 0.04
cationic N	$(-1.1 \pm 0.4) \times 10^{-3}$	0.18 ± 0.01^b	1.3 ± 0.1
amide N	$(-1.7 \pm 0.3) \times 10^{-3}$	0.18 ± 0.01^b	1.5 ± 0.1
aromatic C	$(-2.2 \pm 0.5) \times 10^{-3}$	0.18 ± 0.01^c	1.7 ± 0.2

[a]Assumed value; used to calculate lower bound b_1 for phosphate O surface. [b]Assumed lower bound value based on results for hydrocarbon and carboxylate O surfaces (see text). [c]Lower bound value determined from application of SPM to analyze effects of salts on hydrocarbon solubility (*40*).

phosphate O, aromatic C, and cationic (ammonium, guanidinium) N surface types (Table 2). In this analysis, we assigned a μ_{23}/RT value of 0 m^{-1} to Na^+. This value, which is qualitatively consistent with a small net interaction due to strong exclusion of Na^+ from hydrocarbon surface (*40*) and a presumed strong favorable interaction of Na^+ with the carboxylate of GB (*44*), results in values of μ_{23}/RT for K^+ and Cl^- of 0.09 m^{-1} and −0.03 m^{-1}. The most unfavorable potentials are for anionic oxygens, and the interaction with phosphate O is more unfavorable than that with carboxylate O. These results provide a higher resolution separation of the previously reported composite GB−anionic oxygen interaction potential, obtained from an analysis of GB−DNA and GB−protein interactions and GB effects on protein folding (*8*). Interpretation of the GB−phosphate O and GB−carboxylate O interaction potentials (Table 2) using the SPM (eq 5) indicates that the extents of hydration of anionic carboxylate and phosphate oxygens are at least 0.16 ± 0.01 H_2O Å^{-2} (~two layers) and 0.27 ± 0.02 H_2O Å^{-2} (~three layers), respectively. We deduce that GB is highly excluded from this water because of its inability to compete with water as a hydrogen bond donor to interact with these oxygens.

The favorable interaction of GB with cationic nitrogens (Table 2) is presumably due to a hydrogen-bonded ion pair or salt bridge between the ammonium or guanidinium group (donor) and the anionic carboxylate of GB (acceptor). If the hydration of these cationic nitrogens (in the absence of GB) is two layers of water (as determined for hydrocarbon (*40*) and anionic carboxylate oxygen surfaces), then the partition coefficient characterizing this interaction is $K_p = 1.3 \pm 0.1$. The GB−amide O interaction potential is similar in sign and magnitude to that for carboxylate oxygen, indicating that GB is quite highly excluded from the water of hydration of amide oxygens (K_p ~0.1). The chemical basis of this exclusion, likewise, must be the inability of GB to compete with water to hydrogen bond to the amide oxygen. The favorable preferential interaction of GB with amide nitrogen (similar to that with cationic nitrogen) likely indicates that the amide NH is a better hydrogen bond donor than water for the GB carboxylate oxygen acceptor. Interpretation of this favorable interaction using the SPM yields a GB−amide N partition coefficient $K_p = 1.5 \pm 0.1$. Auton and Bolen concluded that GB is excluded from the peptide backbone (*17*); our potentials quantify this moderate (net) exclusion and decompose it into contributions from the oxygen, carbon, and nitrogen surfaces.

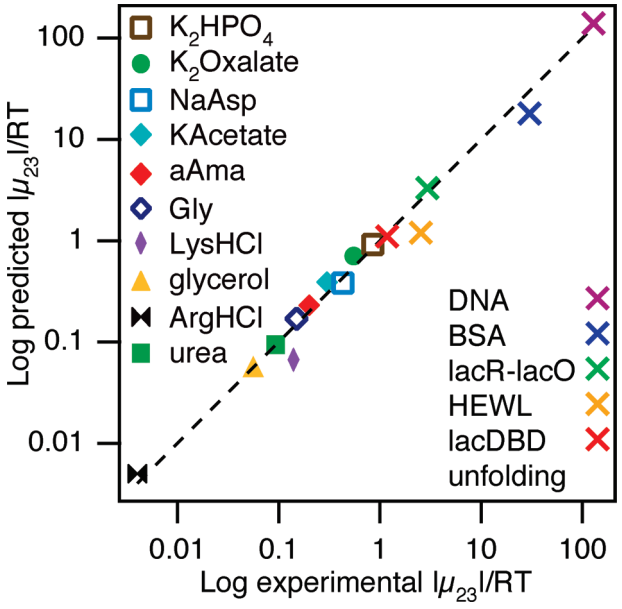

FIGURE 2: Log−log plot of predicted vs observed absolute values of $|\mu_{23}|/RT$ for interactions of GB with model compounds and biopolymers and of $|\Delta\mu_{23}|/RT$ for effects of GB on biopolymer processes.

The most favorable interaction determined here is that between GB and the aromatic carbon surface of the benzyl group of benzoate. Since the hydration of benzyl surface is 0.18 H_2O Å^{-2} (*40*), the GB−benzyl partition coefficient is 1.7 ± 0.2. This is almost certainly a cation−π interaction between the $(CH_3)_3N^+$ group of GB and aromatic surface, like those observed in the crystal structures of complexes of GB with two proteins involved in its transport across the periplasmic membrane. The periplasmic binding protein ProX (*45*, *46*) and the transporter protein BetP (*47*) both feature GB trapped in a "box" formed by aromatic side chains, and the pore of BetP is lined with aromatic amino acids. The strength of the cation−π interaction between GB and the benzyl group supports the structural proposals of the key role of cation−π interactions in both the binding and transport of GB.

Values of μ_{23}/RT for all model compounds studied were calculated using the results of the global fit (Table 2) and are compared to the experimentally observed values in Table 1 (see also Supporting Information Figure S1). Agreement is generally very good, showing the merit of the approach and of the assignment of a single-ion μ_{23}/RT value of 0 for Na^+.

DISCUSSION

Predicting GB−Biopolymer Interactions and GB Effects on Biopolymer Processes. Using GB interaction potentials from Table 2 and ASA analyses from structural data (Supporting Information Table S2) as inputs to eq 4, we predict the strength (μ_{23}) of GB−biopolymer interactions and of GB *m*-values ($\Delta\mu_{23}$) for protein folding and protein−DNA binding. Figure 2 is a log−log plot comparing predicted and observed magnitudes of μ_{23}/RT (or $\Delta\mu_{23}/RT$) for both model compounds and biopolymers. Agreement is quantitative (within ~15%) for interaction of GB with double-stranded DNA and effects of GB on folding of the lac repressor DNA binding domain (DBD) (*15*) and binding of lac repressor to lac operator DNA. Predictions for interactions of GB with native BSA and HEWL are only semiquantitative; both are 40−50% less than the observed values (Supporting Information Figure S2).

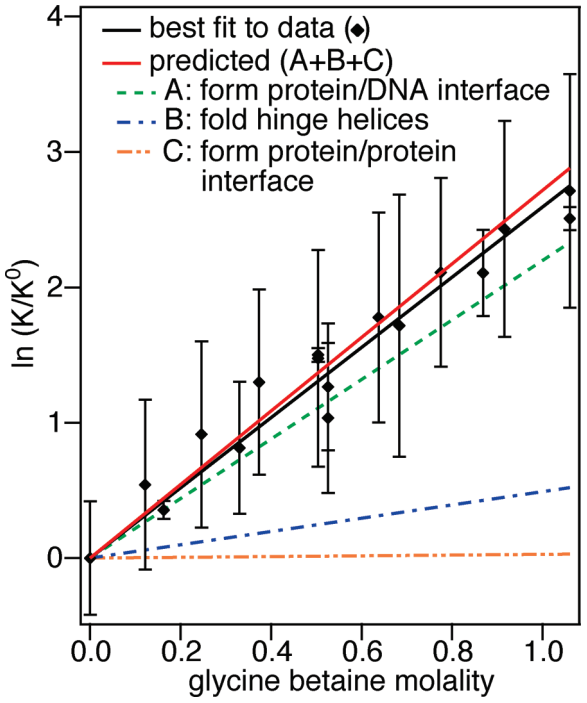

FIGURE 3: Comparison of predicted and observed effects of GB on lac repressor−operator binding at constant salt molality. The predicted contributions to the dependence of ln K_{obs} on GB molality from the three interfaces formed during this assembly process are also shown.

GB effects on processes involving DNA can be attributed in large part to the extraordinary exclusion of GB from anionic phosphate oxygens. Exclusion of GB from the 23700 Å^{-2} of anionic phosphate oxygen ASA of 160 bp DNA is predicted to contribute 120 m^{-1} to the observed value of μ_{23}/RT (130 ± 13; see Supporting Information Figure S2), and the 320 K^+ counterions are predicted to contribute an additional 30 m^{-1}. Since contributions of other surface types are predicted to be small and somewhat compensating, these two contributions determine the observed exclusion of GB from duplex DNA.

Site-specific binding of proteins to helical DNA partially or fully dehydrates and buries significant amounts of anionic DNA phosphate oxygen surface. Since GB is completely excluded from the water of hydration of this surface on free DNA, the reduction in water activity caused by addition of GB drives formation of these complexes. Figure 3 is a semilog plot of binding constants K_{obs} as a function of GB concentration (0−1 *m*), determined from GB titrations of mixtures of lac repressor tetramer and 40 bp lac operator DNA at constant salt molality. From the slope, the GB *m*-value/*RT* (i.e., $\Delta\mu_{23}/RT$) is 2.6 ± 0.4 m^{-1}, ~25% larger than that previously estimated from experiments at higher GB concentration (*8*). The burial of 6900 Å^2 of C, N, and O surface in the protein−DNA and protein−protein interfaces formed in complexation (Supporting Information Table S2) gives rise to a net predicted GB *m*-value/*RT* = 2.7 m^{-1} at constant salt molality, in good agreement with experiment (Figure 3). Formation of the repressor−operator interface, which buries 630 Å^2 of anionic DNA phosphate O surface, is predicted to contribute 2.2 m^{-1} (81%) to the GB slope; folding the hinge helices, which buries 520 Å^2 of amide surface, is predicted to contribute 0.5 m^{-1} (19%); and formation of the core repressor−DBD interface is predicted to make no significant contribution due to small compensating effects (Figure 3).

The GB *m*-value/*RT* for unfolding of the globular lac repressor DNA binding domain is 1.2 ± 0.1 m^{-1} at low GB concentration (*11*). From the potentials of Table 2 and ΔASA of Supporting Information Table S2, weak exclusion of GB from aliphatic hydrocarbon surface (68% of ΔASA) is predicted to contribute 0.6 m^{-1} to the GB *m*-value. Exclusion from amide oxygen and accumulation at amide nitrogen contributes a net 0.8 m^{-1}, and exposure of 150 Å^2 of tyrosine aromatic ASA in unfolding contributes −0.3 m^{-1}. Other contributions are smaller and largely compensate each other. The predicted *m*-value/*RT* for the stabilizing effect of GB on unfolding is 1.0 m^{-1}, within 20% of the experimental value. We conclude that GB modestly stabilizes globular proteins like the lac DBD because of roughly equal contributions from exclusion of GB from amide O and aliphatic C surface, the effects of which are somewhat mitigated by accumulation at aromatic C and amide N surface.

CONCLUSIONS

Our results explain why GB is such an effective osmolyte: it is unable to compete with the water of hydration of anionic (phosphate, carboxylate) and amide oxygens of cellular biopolymers, lipids, metabolites, and GB itself. Cytoplasmic GB is therefore concentrated in the remaining intracellular water, thus increasing cytoplasmic osmolality more than a uniformly distributed or accumulated solute could. Our results also predict that steps in biopolymer assembly mechanisms and other processes which bury (or expose) anionic or amide oxygen surface, as in forming a protein−nucleic acid interface, will be especially sensitive to changes in GB concentration (*8*, *48*), making GB a valuable quantitative probe in mechanistic studies. How cells maintain regulation of protein−nucleic acid interactions of gene expression, replication, and other nucleic acid processes while accumulating strongly perturbing osmolytes to high concentrations during osmotic stress remains an in vivo−in vitro paradox (*5*).

ACKNOWLEDGMENT

The authors thank the reviewers for their thoughtful comments and suggestions and Jorge Estrada for providing the PDB files for the lacDBD unfolded ensemble.

NOTE ADDED AFTER ASAP PUBLICATION

After this paper was published ASAP October 8, 2009, a correction was made to Table 2; the corrected version was reposted October 27, 2009.

SUPPORTING INFORMATION AVAILABLE

Tables containing water-accessible surface areas (ASA) or ΔASA for model compounds, biopolymers, and biopolymer processes, one figure comparing experimental values of μ_{23}/RT for model compounds with those calculated from the global fit of the experimental data, and one figure illustrating the determination of ΔOsm using VPO data on 0.005 *m* BSA solutions as a function of GB concentration and comparing ΔOsm values for BSA, DNA, and HEWL. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Ball, P. (2008) Water as a biomolecule. *ChemPhysChem 9*, 2677–2685.
2. Timasheff, S. N. (1998) Control of protein stability and reactions by weakly interacting cosolvents: The simplicity of the complicated. *Adv. Protein Chem. 51*, 355–432.

3. Record, M. T., Zhang, W., and Anderson, C. F. (1998) Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: A practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. *Adv. Protein Chem. 51*, 281–353.
4. Pegram, L. M., and Record, M. T. (2008) Quantifying accumulation or exclusion of H^+, HO^-, and Hofmeister salt ions near interfaces. *Chem. Phys. Lett. 467*, 1–8.
5. Cayley, S., and Record, M. T. (2003) Roles of cytoplasmic osmolytes, water, and crowding in the response of *Escherichia coli* to osmotic stress: Biophysical basis of osmoprotection by glycine betaine. *Biochemistry 42*, 12596–12609.
6. Sleator, R. D., Wouters, J., Gahan, C. G., Abee, T., and Hill, C. (2001) Analysis of the role of OpuC, an osmolyte transport system, in salt tolerance and virulence potential of *Listeria monocytogenes*. *Appl. Environ. Microbiol. 67*, 2692–2698.
7. Fried, M. G., Stickle, D. F., Smirnakis, K. V., Adams, C., MacDonald, D., and Lu, P. (2002) Role of hydration in the binding of lac repressor to DNA. *J. Biol. Chem. 277*, 50676–50682.
8. Hong, J., Capp, M. W., Saecker, R. M., and Record, M. T. (2005) Use of urea and glycine betaine to quantify coupled folding and probe the burial of DNA phosphates in lac repressor-lac operator binding. *Biochemistry 44*, 16896–16911.
9. Vander Meulen, K. A., Saecker, R. M., and Record, M. T. (2008) Formation of a wrapped DNA-protein interface: Experimental characterization and analysis of the large contributions of ions and water to the thermodynamics of binding IHF to H' DNA. *J. Mol. Biol. 377*, 9–27.
10. Jen-Jacobson, L., and Jacobson, L. (2008) Role of water and effects of small ions in site-specific protein-DNA interactions, in *Structural Biology of Protein-Nucleic Acid Interactions* (Rice, P. A. and Correll, C. Eds.), pp 13−46, Royal Society of Chemistry, Cambridge.
11. Felitsky, D. J., and Record, M. T. (2004) Application of the local-bulk partitioning and competitive binding models to interpret preferential interactions of glycine betaine and urea with protein surface. *Biochemistry 43*, 9276–9288.
12. Santoro, M. M., Liu, Y., Khan, S. M., Hou, L. X., and Bolen, D. W. (1992) Increased thermal stability of proteins in the presence of naturally occurring osmolytes. *Biochemistry 31*, 5278–5283.
13. Lambert, D., and Draper, D. E. (2007) Effects of osmolytes on RNA secondary and tertiary structure stabilities and RNA-Mg^{2+} interactions. *J. Mol. Biol. 370*, 993–1005.
14. Rees, W. A., Yager, T. D., Korte, J., and von Hippel, P. H. (1993) Betaine can eliminate the base pair composition dependence of DNA melting. *Biochemistry 32*, 137–144.
15. Felitsky, D. J., Cannon, J. G., Capp, M. W., Hong, J., Van Wynsberghe, A. W., Anderson, C. F., and Record, M. T. (2004) The exclusion of glycine betaine from anionic biopolymer surface: Why glycine betaine is an effective osmoprotectant but also a compatible solute. *Biochemistry 43*, 14732–14743.
16. Hong, J., Capp, M. W., Anderson, C. F., Saecker, R. M., Felitsky, D. J., Anderson, M. W., and Record, M. T. (2004) Preferential interactions of glycine betaine and of urea with DNA: Implications for DNA hydration and for effects of these solutes on DNA stability. *Biochemistry 43*, 14744–14758.
17. Auton, M., and Bolen, D. W. (2005) Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proc. Natl. Acad. Sci. U.S.A. 102*, 15065–15068.
18. Stanley, C., and Rau, D. C. (2008) Assessing the interaction of urea and protein-stabilizing osmolytes with the nonpolar surface of hydroxypropylcellulose. *Biochemistry 47*, 6711–6718.
19. Pierce, V., Kang, M., Aburi, M., Weerasinghe, S., and Smith, P. E. (2008) Recent applications of Kirkwood-Buff theory to biological systems. *Cell Biochem. Biophys. 50*, 1–22.
20. Ni, H. H., Anderson, C. F., and Record, M. T. (1999) Quantifying the thermodynamic consequences of cation (M^{2+}, M^+) accumulation and anion (X^-) exclusion in mixed salt solutions of polyanionic DNA using Monte Carlo and Poisson-Boltzmann calculations of ion-polyion preferential interaction coefficients. *J. Phys. Chem. B 103*, 3489–3504.
21. Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol. 55*, 379–400.
22. Privalov, P. L., and Gill, S. J. (1988) Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem. 39*, 191–234.
23. Murphy, K. P., and Freire, E. (1992) Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv. Protein Chem. 43*, 313–361.
24. Spolar, R. S., and Record, M. T. (1994) Coupling of local folding to site-specific binding of proteins to DNA. *Science 263*, 777–784.
25. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant *m*-values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci. 4*, 2138–2148.
26. Courtenay, E. S., Capp, M. W., Saecker, R. M., and Record, M. T. (2000) Thermodynamic analysis of interactions between denaturants and protein surface exposed on unfolding: Interpretation of urea and guanidinium chloride m-values and their correlation with changes in accessible surface area (ASA) using preferential interaction coefficients and the local-bulk domain model. *Proteins: Struct., Funct., Genet. 41*, 72–85.
27. Robinson, R. A., and Stokes, R. H. (1961) Activity coefficients in aqueous solutions of sucrose, mannitol, and their mixtures at 25°C. *J. Phys. Chem. 65*, 1954–1958.
28. Schonert, H., and Stroth, L. (1981) Thermodynamic interaction between urea and the peptide group in aqueous solutions at 25°C. *Biopolymers 20*, 817–831.
29. Anderson, C. F., and Record, M. T. (2004) Gibbs-Duhem-based relationships among derivatives expressing the concentration dependences of selected chemical potentials for a multicomponent system. *Biophys. Chem. 112*, 165–175.
30. Cannon, J. G., Anderson, C. F., and Record, M. T. (2007) Urea-amide preferential interactions in water: Quantitative comparison of model compound data with biopolymer results using water accessible surface areas. *J. Phys. Chem. B 111*, 9675–9685.
31. Archer, D. G. (1992) Thermodynamic properties of the NaCl + H_2O system. II. Thermodynamic properties of NaCl(aq), $NaCl{\cdot}2H_2O$(cr), and phase equilibria. *J. Phys. Chem. Ref. Data 21*, 793–829.
32. Tsodikov, O. V., Record, M. T., and Sergeev, Y. V. (2002) Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. *J. Comput. Chem. 23*, 600–609.
33. Livingstone, J. R., Spolar, R. S., and Record, M. T. (1991) Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area. *Biochemistry 30*, 4237–4244.
34. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235–242.
35. Kalodimos, C. G., Biris, N., Bonvin, A. M. J. J., Levandoski, M. M., Guennuegues, M., Boelens, R., and Kaptein, R. (2004) Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science 305*, 386–389.
36. Bernado, P., Blackledge, M., and Sancho, J. (2006) Sequence-specific solvent accessibilities of protein residues in unfolded protein ensembles. *Biophys. J. 91*, 4536–4543.
37. Estrada, J., Bernado, P., Blackledge, M., and Sancho, J. (2009) Protsa: A web application for calculating sequence specific protein solvent accessibilities in the unfolded ensemble. *BMC Bioinf. 10*, 104.
38. Frank, D. E., Saecker, R. M., Bond, J. P., Capp, M. W., Tsodikov, O. V., Melcher, S. E., Levandoski, M. M., and Record, M. T. (1997) Thermodynamics of the interactions of lac repressor with variants of the symmetric lac operator: Effects of converting a consensus site to a non-specific site. *J. Mol. Biol. 267*, 1186–1206.
39. Venkatesu, P., Lee, M.-J., and Lin, H.-M. (2007) Thermodynamic characterization of the osmolyte effect on protein stability and the effect of GdnHCl on the protein denatured state. *J. Phys. Chem. B 111*, 9045–9056.
40. Pegram, L. M., and Record, M. T. (2008) Thermodynamic origin of Hofmeister ion effects. *J. Phys. Chem. B 112*, 9428–9436.
41. Nozaki, Y., and Tanford, C. (1963) The solubility of amino acids and related compounds in aqueous urea solutions. *J. Biol. Chem. 238*, 4074–4081.
42. Auton, M., and Bolen, D. W. (2007) Application of the transfer model to understand how naturally occurring osmolytes affect protein stability. *Methods Enzymol. 428*, 397–418.
43. Pegram, L. M., and Record, M. T. (2007) Hofmeister salt effects on surface tension arise from partitioning of anions and cations between bulk water and the air-water interface. *J. Phys. Chem. B 111*, 5411–5417.
44. Vrbka, L., Vondrasek, J., Jagoda-Cwiklik, B., Vacha, R., and Jungwirth, P. (2006) Quantification and rationalization of the higher affinity of sodium over potassium to protein surfaces. *Proc. Natl. Acad. Sci. U.S.A. 103*, 15440–15444.
45. Schiefner, A., Breed, J., Bosser, L., Kneip, S., Gade, J., Holtmann, G., Diederichs, K., Welte, W., and Bremer, E. (2004) Cation-pi

interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligand-binding protein ProX from *Escherichia coli*. *J. Biol. Chem. 279*, 5588–5596.

46. Horn, C., Sohn-Bosser, L., Breed, J., Welte, W., Schmitt, L., and Bremer, E. (2006) Molecular determinants for substrate specificity of the ligand-binding protein OpuAC from *Bacillus subtilis* for the compatible solutes glycine betaine and proline betaine. *J. Mol. Biol. 357*, 592–606.
47. Ressl, S., Terwisscha van Scheltinga, A. C., Vonrhein, C., Ott, V., and Ziegler, C. (2009) Molecular basis of transport and regulation in the Na^+/betaine symporter BetP. *Nature 458*, 47–52.
48. Kontur, W. S., Saecker, R. M., Davis, C. A., Capp, M. W., and Record, M. T. (2006) Solute probes of conformational changes in open complex (RP_o) formation by *Escherichia coli* RNA polymerase at the λ P_R promoter: Evidence for unmasking of the active site in the isomerization step and for large-scale coupled folding in the subsequent conversion to RP_o. *Biochemistry 45*, 2161–2177.