Use of Urea and Glycine Betaine To Quantify Coupled Folding and Probe the Burial of DNA Phosphates in Lac Repressor−Lac Operator Binding†

Jiang Hong,‡ Mike W. Capp,§ Ruth M. Saecker,*,§ and M. Thomas Record, Jr.*,‡,§

Departments of Biochemistry and Chemistry, University of Wisconsin, Madison, Wisconsin 53706

Received July 29, 2005; Revised Manuscript Received October 7, 2005

ABSTRACT: Thermodynamic analysis of urea−biopolymer interactions and effects of urea on folding of proteins and α-helical peptides shows that urea interacts primarily with polar amide surface. Urea is therefore predicted to be a quantitative probe of coupled folding, remodeling, and other large-scale changes in the amount of water-accessible polar amide surface in protein processes. A parallel analysis indicates that glycine betaine [*N*,*N*,*N*-trimethylglycine (GB)] can be used to detect burial or exposure of anionic (carboxylate, phosphate) biopolymer surface. To test these predictions, we have investigated the effects of these solutes (0−3 *m*) on the formation of 1:1 complexes between lac repressor (LacI) and its symmetric operator site (SymL) at a constant KCl molality. Urea reduces the binding constant K_{TO} [initial slope $d\ln K_{TO}/dm_{urea} = -1.7 \pm 0.2$], and GB increases K_{TO} [initial slope $d\ln K_{TO}/dm_{GB} = 2.1 \pm 0.2$]. For both solutes, this derivative decreases with an increase in solute concentration. Analysis of these initial slopes predicts that $(1.5 \pm 0.3) \times 10^3$ Å^2 of polar amide surface and $(4.5 \pm 1.0) \times 10^2$ Å^2 of anionic surface are buried in the association process. Analysis of published structural data, together with modeling of unfolded regions of free LacI as extended chains, indicates that 1.5×10^3 Å^2 of polar amide surface and 6.3×10^2 Å^2 of anionic surface are buried in complexation. Quantitative agreement between structural and thermodynamic results is obtained for amide surface (urea); for anionic surface (GB), the experimental value is ∼70% of the structural value. For LacI−SymL binding, two-thirds of the structurally predicted change in amide surface (1.0×10^3 Å^2) occurs outside the protein−DNA interface in protein−protein interfaces formed by folding of the hinge helices and interactions of the DNA binding domain (DBD) with the core of the repressor. Since urea interacts principally with amide surface, it is particularly well-suited to detect and quantify the extent of coupled folding and other large-scale remodeling events in the steps of protein−nucleic acid interactions and other protein associations.

One of the most intriguing results from analyses of the ever-expanding genome sequence database has been the discovery of the prevalence of intrinsically disordered regions in proteins (*1*, *2*). These unstructured domains (and sometimes entire protein chains) play key roles in transcription regulation, translation, and cellular signal transduction, where folding is coupled to binding and recognition (see ref *3* for a review and references therein). Thus, the very lack of structures in these systems determines function.

Prior to the explosion in sequence information, analysis of available thermodynamic data for protein−ligand, protein−protein, and site-specific protein−DNA interactions for which structural information was available foreshadowed the generality of this functional design (*4*). Comparisons of the observed heat capacity and entropy changes of protein interactions with those predicted from model compound data for formation of a rigid-body interface indicated that large-scale folding transitions are often coupled to binding and recognition (*4*). However, heat capacity changes can have contributions from other coupled processes (see Discussion). NMR, CD, and other biophysical techniques provide powerful probes of the dynamic states of free proteins, but many systems are not suitable for such studies. Additionally, conformational changes may occur transiently as part of multistep mechanisms such as DNA opening in transcription initiation. Thus, a current challenge is to find alternative methods for detecting and quantifying these remodeling events in solution to relate them to function.

Analysis of the temperature dependence of interactions between lac repressor and various operator sequences (*5*, *6*) led to the proposal that folding was coupled to binding (*4*) and that the extent of folding was a function of operator sequence (*7*). NMR and X-ray crystallographic studies confirmed these proposals and elegantly elucidated that the region connecting the DNA recognition element (helix−turn−helix domain) to the core domain of the protein is fundamentally the origin of the transitions detected by biophysical solution studies (*8*−*13*). Using this system, we test the hypothesis that urea at nondenaturing concentrations can be used as a quantitative probe of ordering of intrinsically disordered regions and other remodeling events in protein processes.

Precedent for this study is provided by the seminal work of Myers, Pace, and Scholtz (*14*), who observed that the

† Support of this research by NIH Grant GM 23467 is gratefully acknowledged.

* To whom correspondence should be addressed: Department of Biochemistry, 433 Babcock Dr., University of Wisconsin, Madison, WI 53706. Telephone: (608) 262-5332. Fax: (608) 262-3453. E-mail: record@biochem.wisc.edu.

‡ Department of Biochemistry.

§ Department of Chemistry.

10.1021/bi0515218 CCC: $30.25
Published on Web 11/30/2005

derivative of the protein unfolding free energy change with respect to urea molarity ($d\Delta G^{\circ}_{unfold}/d[urea]$, defined as the urea *m*-value) is proportional to the calculated (disulfide-corrected) change in water-accessible surface area (ΔASA).[1] Another important cornerstone for this study is provided by Baldwin and co-workers (*15*), who determined the per-residue *m*-value for unfolding of a series of α-helices (AEAAKA repeat unit). Using this per-residue α-helix *m*-value to estimate contributions from α-helical regions to observed *m*-values for unfolding of globular proteins, they concluded that the interaction of urea with peptide groups accounted for a major part of the destabilizing effect of urea on these systems. Significantly, ASA analysis reveals that the *m*-value/ΔASA_{total} ratio for unfolding the AEAAKA α-helices is more than 3 times as large as for unfolding globular proteins (*16*). Our quantitative analysis of these unfolding data for proteins and α-helical peptides, together with data for interactions of urea with native proteins and DNA, demonstrates that the only significant preferential interaction of urea is with polar amide (N, O) surface (*16*−*18*), and is the basis for the determination of its intrinsic interaction coefficient per unit of amide surface (in Analysis).

Typically, as exemplified for lac repressor−lac operator binding (see Results), only a small fraction (5−15%) of the surface buried in the interface between a protein and duplex DNA is composed of polar amide groups. Therefore, the effect of urea concentration on a rigid-body interaction of a protein with duplex DNA or RNA (in which no conformational changes in the protein occur) is predicted to be relatively small. However, polar amide surface constitutes a large fraction (40−60%) of the surface buried in α-helix formation, and a significant fraction (20−30%) of the surface buried in protein folding. Consequently, urea is predicted to be particularly suitable as a quantitative probe of α-helix formation and other large-scale folding/remodeling processes coupled to protein−DNA binding or other protein associations.

In this study, we also test the hypothesis that glycine betaine (GB), which interacts primarily with anionic oxygen surface of proteins and DNA (*18*, *19*), can be used to quantify the amount of this surface buried or exposed in a biopolymer process. For reasons detailed below (see Analysis), GB is not expected to be as unambiguous a probe of changes in ASA of anionic surface as urea appears to be for amide surface.

For urea and GB, we obtain and analyze both initial slopes and curvature at higher solute concentrations characterizing the dependence of the standard binding free energy (equivalently the logarithm of the observed binding constant) on solute molality at a constant KCl molality. Using osmometric data, we correct these initial slopes from a constant KCl molality to a constant KCl activity. We summarize the current status of the thermodynamic characterization of interactions of urea and GB with a spectrum of biopolymer surfaces, and obtain values for the intrinsic interaction coefficients (per unit area of the specified type of biopolymer surface and per unit of solute molality). Using these intrinsic interaction coefficients, we interpret the initial slopes and curvatures of plots of $\ln K_{TO}$ versus molality of urea and of GB to characterize coupled conformational changes and burial of polar amide and anionic surface, respectively, in the repressor−operator interaction. The agreement between predicted and experimental values of amide surface buried on forming the lac repressor−lac operator DNA complex indicates that urea will be a powerful tool for detecting coupled conformational changes in other protein associations.

MATERIALS AND METHODS

Chemicals. Glycine betaine monohydrate (>99% pure, FW 135.2) was obtained from Sigma (St. Louis, MO). Ultrapure urea (>99.5% pure, FW 60.06) was obtained from Life Technologies. Glycerol (>99.5% pure, FW 92.09) was obtained from Aldrich (Milwaukee, WI). All other chemicals and enzymes are reagent or molecular biology grade.

Preparation and Operator Binding Activity of Lac Repressor. Tetrameric lac repressor was purified from *Escherichia coli* strain HB101/lac pIQ by the method of refs *20* and *21* as modified as described previously (*22*, *23*). Aliquots of purified repressor were stored at a total protein concentration of 27.1 μM (tetramer) at −70 °C in 10 mM HEPES (pH 7.5), 0.1 M KCl, 0.1 mM DTT, and 30% (v/v) glycerol. Molar concentrations of repressor tetramer were determined from absorbance at 280 nm using an extinction coefficient ϵ of 9.2×10^4 M^{-1} cm^{-1} (*24*). Working stocks of repressor were prepared by dilution from the −70 °C stock to the micromolar range and stored for up to 10 days at −20 °C in 10 mM KH_2PO_4 (pH 7.5), 0.075 M KCl, 0.1 mM DTT, and 30% (v/v) glycerol. Activities of these samples, defined as the fraction of active repressor tetramers, were determined from filter binding titrations in the CBLR (complete binding of limiting reagent) regime (*25*). Activities ranged from 36 to 44%; repressor activity was not significantly affected by urea or GB [tested up to 1.6 *m* urea and 1.1 *m* GB (data not shown)], or by storage at −20 °C during use for titrations of a labeled DNA sample (up to 10 days). All concentrations specified in the text are total (active and inactive) molar concentrations of repressor tetramer.

Preparation of ^{32}P-Labeled SymL Plasmid DNA. pAS 8-1 (*6*) plasmid DNA (2514 bp), carrying the symmetric 20 bp SymL (formerly designated O^{sym}) operator (*26*) and no other operator sites, was isolated from *E. coli* HB101 using the QIAGEN plasmid purification method and linearized with restriction endonuclease *Pst*I to obtain a centrally located SymL operator. Linear plasmid DNA was replacement labeled (*27*) with [^{32}P]dATP and [^{32}P]dCTP (>6000 Ci/mmol), purified as previously described (*25*), and stored at a concentration of approximately 6 μM at −20 °C in TE buffer [10 mM Tris (pH 8.0) and 1 mM Na_2EDTA].

The concentration of ^{32}P-labeled DNA was determined by a novel quantitative fluorescence assay using Gel-Star (Cambrex BioScience RockLand Inc.) stain to detect DNA; this reagent gives a lower background, a higher sensitivity, and a broader range of linearity than ethidium bromide. In this assay, fluorescence at 545 nm was measured with

[1] Abbreviations: ASA, water-accessible surface area; ASA_{amide}, contribution to ASA from the polar atoms (N and O) of amide groups; GB, glycine betaine (*N*,*N*,*N*-trimethylglycine); LacI, lac repressor; SymL, symmetric lac operator, a palindrome of the left (upstream) half-site of the primary O_1 operator; *m*-value, slope of the plot of ΔG° of protein unfolding versus the molar concentration of the chemical denaturant or osmolyte; AEAAKA, Ala-Glu-Ala-Ala-Lys-Ala; DBD, DNA binding domain (residues 1−62) of LacI; LEM, linear extrapolation method of analysis of urea unfolding data.

excitation at 310 nm in an SLM 8000 spectrofluorimeter. A standard curve was prepared using DNA standards in the concentration range of 0.035−0.18 μg/mL (0.11−0.53 μM in nucleotide monomer). These standards were prepared by accurate dilutions of a stock solution of a known concentration of purified linearized plasmid DNA [3824 bp pKO^c (*28*)], first with TE buffer, followed by a 1:20 dilution with a Gel Star solution prepared by 1:5000 dilution of the proprietary Gel Star stock. The concentration of the DNA stock solution was determined from the absorbance at 260 nm by UV spectroscopy using an ϵ of 0.02 (μg/mL)$^{-1}$ cm^{-1} (*27*). Concentration determinations for one DNA sample using independently prepared standard curves are reproducible to approximately 5%.

Filter Binding Assays. The nitrocellulose filter assay for repressor−operator DNA binding (*25*, *29*−*33*) was used to determine binding isotherms and binding constants (equilibrium concentration quotients) K_{TO} for formation of 1:1 repressor tetramer−operator complexes (TO) as a function of urea or GB molality in binding buffer (BB) [0.010 *m* K_2HPO_4, 0.001 *m* K_2EDTA, 0.001 *m* dithiothreitol (DTT), 0.752 *m* glycerol, 50 μg/mL bovine serum albumin (BSA), and either 0.189 or 0.404 *m* KCl; pH adjusted to 7.3 with concentrated HCl]. The K^+ concentration of BB (0.212 or 0.427 *m*) is the sum of contributions from K_2HPO_4, K_2EDTA, and KCl.

For binding studies in the presence of urea or GB, a series of binding buffers with varying molal concentrations (m_3) of urea or GB were prepared by adding accurately weighed amounts of solid urea or GB to BB. In the case of urea, this procedure kept the molality of all other solutes constant. For GB, added as the monohydrate, this procedure reduced the molality of K^+, glycerol, and other constituents in BB by approximately 5% at 3 *m* GB. Experimentally determined values of K_{TO} at each m_{GB} were corrected for the small and opposing effects of the changes in molality of K^+ and glycerol. To determine the glycerol correction, complete titrations of operator DNA with repressor at 0.212 *m* K^+ and 0 or 1.59 *m* urea were performed at least twice at 0 and 0.752 *m* glycerol. These data provide an estimate of $\partial \ln K_{TO}/\partial m_{glycerol}$ of 1.5 ± 0.5 (*23*), used here to correct K_{TO} to a constant glycerol concentration (0.752 *m*). The log−log [salt]-derivative of the LacI−SymL binding constant in the vicinity of 0.4 M KCl ($SK_{TO} = -6.3$; *30*) was used to correct K_{TO} to a constant K^+ concentration (0.427 *m*) as shown in the tabulated values of K_{TO} in the Supporting Information. The calculated (opposing) effects on K_{TO} of the small reductions in K^+ and glycerol molality with addition of hydrated GB increase the initial slope $d\ln K_{TO}/dm_{GB}$ by only 3%, which is substantially less than the uncertainty in this slope (10%). Repressor and ^{32}P SymL plasmid DNA stocks were serially diluted into BB at the desired solute concentration; dilution factors exceeded 10^3 except in activity assays, where they exceeded 60, so that solutes in the storage buffer make no significant contribution to the composition of the binding buffer.

In urea, binding isotherms were determined at 0.212 and 0.427 *m* K^+ [corresponding to KCl concentrations (m_4) of 0.189 and 0.404 *m*, respectively]; in GB, binding isotherms were determined at only ~0.43 *m* K^+ because K_{TO} is too large to measure at high m_{GB} in 0.2 *m* K^+. Ranges of solute concentration investigated were 0−3.2 *m* urea at 0.212 *m* K^+, 0−1.3 *m* urea at 0.427 *m* K^+, and 0−3.0 *m* GB at 0.43 *m* K^+. All protein DNA binding experiments were performed at 24 °C.

For each forward titration of SymL operator DNA with repressor, 15 sample tubes with the same concentration of ^{32}P labeled SymL operator DNA and increasing concentrations of repressor were equilibrated in binding buffer at a specified concentration of urea or GB for 0.8−2 h at 24 °C. Duplicate aliquots from each sample were then filtered under 250−350 mmHg vacuum through Schleicher & Schuell BA-85 nitrocellulose filters, pretreated as described previously (*25*). After filtration, filters were washed once and dried. Cerenkov radiation of each filter was measured in a Beckman LS 1801 scintillation counter. The total radioactivity filtered in each aliquot was measured by counting multiple aliquots of unfiltered reaction mixture. Experiments were carefully designed, on the basis of preliminary results, to obtain data characterizing the entire binding isotherm. At least two independent determinations of the binding isotherm were performed for each solute and salt concentration investigated, with one exception (one isotherm at 1.06 *m* urea and 0.212 *m* K^+).

The experimentally observed fraction of operator DNA retained on the filter by repressor (θ_{obs}) was calculated by eq 1 from the counts per minute (C) retained on the filter for each sample in a repressor titration experiment:

$$\theta_{obs} = (C - C_b)/(C_t - C_b) \quad (1)$$

where C_b and C_t are the background and total values of C, respectively, for a sample. Values of θ_{obs} are related to θ, the fractional saturation of operator DNA with repressor, by the efficiency E of retention of operator DNA on the filter at saturation with lac repressor:

$$\theta = \theta_{obs}/E \quad (2)$$

The efficiency E is invariably less than 1.0 because of DNA damage and/or incomplete retention of preexisting complexes; values of E in the range 0.37−0.83 were obtained in this study, and no solute effect on E was observed.

Titrations of a fixed operator DNA concentration with increasing repressor concentrations were performed over a range of $[SymL]_{total}$ (0.69−2.5 pM) chosen to be low enough so that even in the initial stages of the titration only 1:1 TO complexes formed (*25*). Lac repressor tetramer is stable to dissociation in the concentration range that was investigated (*25*, *34*). The two-site binding model of the repressor tetramer (*25*, *30*) was used to analyze binding isotherms, with the constraint that the two sites are identical at the salt concentrations that were investigated (equal site binding constants, $ks_1 = ks_2$; *30*). Using the independently determined protein activity, values of θ_{obs} as a function of the total (active and inactive) molar concentration of lac repressor for each isotherm (15 points) were fitted to the two-site binding model to determine binding constants K_{TO} ($=2ks_1$) and filter efficiency E for each protein titration experiment. The activity of repressor tetramer was determined by fitting values of the observed fractional saturation of SymL operator DNA (θ_{obs}) for a protein titration in the CBLR regime (*30*) to the two-site model and fixing $ks_1 \gg [SymL]_{total}^{-1}$.

Data Fitting. The fitting program for determining the binding constant K_{TO} from the binding isotherm (*25*) uses the Newton−Raphson algorithm (*35*) linked to the nonlinear least-squares parameter minimization program NONLIN (*36*). Multiple linear regression (*37*) was used to fit the dependences of $\ln K_{TO}$ on solute concentration (m_3) to second-order polynomials and in analysis of thermodynamic data for urea−biopolymer interactions and urea effects on proteins and α-helix folding (see Analysis).

Modeling and Calculation of Changes in Water-Accessible Surface Areas. All water-accessible surface areas (ASAs) were calculated using ANAREA (*38*) as previously described (*39*).

(a) Lac Repressor−SymL Operator. For the LacI−SymL interaction, entry 1EFA (*40*) from the Protein Data Bank (*41*) was used to calculate rigid-body changes in ASA upon formation of the LacI−SymL operator complex. This file contains the coordinates for the repressor dimer (lacking only the C-terminal helix required for tetramerization) bound to SymL operator DNA and the anti-inducer orthonitrophenyl fucoside (*40*). The resolution of the crystal structure is 2.6 Å. No water molecules are bound in the relevant protein−DNA or protein−protein interfaces in 1EFA. Although chemically both the repressor and this DNA site are 2-fold symmetric, the degree of resolution in each half-site differs. In the protein−DNA interface, side chains of residues 2 and 26−37 in monomer A and residues 2, 31, 36, and 37 in monomer B are missing. In the half-site contacted by the A "headpiece" domain (residues 1−49), the terminal 5′ end is missing density for nucleic acid bases and backbone assigned as D1−D4 and E19−E21. In the "B" half-site, only the terminal base E1 is missing. We chose a model that utilizes the B half-site where more atoms are resolved.

In 1EFA, SymL is bent by ∼40°. Insight II (Accelyrs Software Inc.) was used to build a B-form DNA corresponding to the E2−E11 and D12−D21 sequences as a model of the free DNA half-site. Values of ΔASA for forming the rigid-body LacI−SymL interface were then calculated as the difference in ASA between the bound half-site (B chain, DNA bases D12−D21, E2−E11 in 1EFA) and the free B chain (1EFA) and free DNA model. This difference was then multiplied by 2.

In the free repressor, residues connecting headpiece to the core are unfolded (*12*), but fold upon binding the SymL operator (*8*−*13*). To calculate the ΔASA of this folding process, Insight II was used to put the conformation of residues 50−60 from 1EFA in an extended conformation, a reproducible representation of unfolded regions used to calculate ΔASA of unfolding (*14*, *39*, *42*). Under conditions where secondary structures or tertiary interactions (e.g., hydrophobic collapse) are stable, this model may overestimate the ASA of these regions (see part c below). Since the side chains missing in chain A are not involved in this coupled folding transition, chains A and B were used in this calculation. The difference between ASA in this extended state model of the LacI dimer and in the bound conformation in 1EFA was calculated. Values of ΔASA for this calculation include the folding transition in residues 50−58 as well as the surface changes arising from "docking" the DBD on the core domain. The change in ASA in residues 50−58 was subtracted from the total to obtain the amount buried in docking.

(b) α-Helical and Unfolded States of Alanine-Containing Peptides. Scholtz et al. (*15*) reported and analyzed the urea denaturation at 273 K of a series of α-helical peptides with the sequence Ac-Tyr-(Ala-Glu-Ala-Ala-Lys-Ala)$_k$-Phe-NH_2, with values of k of 2, 3, 4, 5, and 8. To calculate the changes in total and amide surface area in the unfolding of these peptides, a representative α-helical peptide with $k = 3$ was built using the standard parameters from the Insight II Biopolymer library (Accelrys). The denatured state was modeled as an extended chain. The difference in ASA calculated for the central six-residue internal repeat for conversion from the α-helix to the unfolded state was divided by 6 to obtain the ΔASA of unfolding per average residue. Three different conformations of the lysine and glutamate side chains in the α-helix were used, and the results were averaged to obtain the values reported in Table 3. The first was the conformation modeled by Insight II and reported previously (*16*). In the second model, these side chains were rotated to be maximally extended from the helix backbone. In the third model, glutamate and lysine were rotated to make the best possible salt bridge consistent with the α-helical geometry (the side chain rotamer library in Insight II allows the carbonyl oxygens of glutamate to be positioned 3.4 Å from the terminal nitrogen in the side chain of lysine). Although the change in total ASA varies, the change in polar amide ASA upon unfolding is the same (±0.5 Å^2) for all three models of the α-helix.

(c) Folded and Unfolded States of Monomeric and Dimeric Globular Proteins. Criteria for selection of monomeric and dimeric proteins for the analysis of change in surface area and *m*-value were high-resolution native structures containing no disulfide cross-links or heme groups, reversible and two-state unfolding in urea, and spectroscopically determined *m*-values in the ranges pH 5−8 and 0−40 °C. The data set in the Supporting Information includes all proteins analyzed by Myers et al. (*14*) and all more recent systems which meet these criteria. Surface areas of the folded states of proteins were calculated using the appropriate PDB entry, and unfolded states were modeled as extended chains in which side chains were rotated to avoid steric clash using Insight II (see the Supporting Information). Various models have been used to quantify the ASA of the unfolded state. We and others have proposed that, in the absence of evidence for secondary structure or tertiary interactions, the extended polypeptide chain provides the most reasonable estimate of the ASA (*14*, *39*, *42*). Tripeptide models overestimate ASA (*39*, *42*), and models that include secondary structure should be used only where evidence of secondary structure exists. In unfolding studies at high urea concentrations, no stable secondary structure is detected by circular dichroism of the unfolded state.

RESULTS

Surface Area Analysis of LacI−SymL Interactions. The high-resolution structure of a LacI dimer−SymL complex (*40*), together with NMR structural data for the DNA binding domain (*8*), provides the data for predicting the amount and composition of the surface buried in complexation. These results are summarized in Table 1. Of the 6882 Å^2 of total surface of LacI and SymL operator buried in complexation, 1526 Å^2 is polar amide surface and 632 Å^2 is anionic oxygen surface. All anionic oxygen surface buried upon binding is

Table 1: Changes in Accessible Surface Area (ΔASA) in LacI−SymL Complexation[a]

process	$-\Delta ASA_{total}$ (Å^2)	$-\Delta ASA_{amide}$ (Å^2) (%)	$-\Delta ASA_{anionic}$ (Å^2)
formation of LacI−SymL interface (docking two DBDs on SymL site)	3689	491 (13%)	632
folding of two hinge helices (residues 50−58)	1144	515 (45%)	0
formation of two DBD−core interfaces	2049	520 (25%)	0
total	6882	1526 (22%)	632

[a] ΔASA values calculated using ANAREA (*38*), PDB entry 1EFA (*40*), and models of the free states of SymL and LacI as described in Materials and Methods.

from DNA phosphate oxygens buried in the protein−DNA interface. Since amide-like functional groups comprise only 2.5% of the ASA of duplex DNA, almost all of the change in polar amide surface (95%) arises from burial of carbonyl oxygen and nitrogen atoms of amide groups of the peptide backbone and of Q and N side chains of the repressor. Polar amide surface makes up only 13% of the surface buried in the protein−DNA binding interface, but constitutes 25% of the DBD−core interface and 45% of the surface buried in folding the hinge helices. Consequently, less than one-third of the polar amide surface buried in the binding process (491 Å^2 of 1526 Å^2) is in the repressor−operator binding interface; the majority is buried in folding the two hinge helices (515 Å^2) and forming protein−protein interfaces between the folded DNA binding domain and the core of lac repressor (520 Å^2).

Opposite Effects of Urea (Destabilizing) and Glycine Betaine (Stabilizing) on Binding of Lac Repressor to Lac Operator DNA. Since the structural data predict that substantial amounts of amide and anionic oxygen surface are buried in forming the LacI−SymL complex, binding studies as a function of urea and glycine betaine (GB) concentration were performed to test the following hypotheses (*18*, *19*): (1) Urea is a quantitative thermodynamic probe of large-scale changes in the amount of water-accessible polar amide surface, and (2) GB can be used to detect burial or exposure of anionic (carboxylate and phosphate) biopolymer surface in biopolymer processes. Figures 1 and 2 show representative binding isotherms over the ranges of urea and GB concentrations that were investigated. Nitrocellulose filter binding data at 24 °C at each solute concentration are plotted as the fraction (θ) of total SymL operator DNA complexed with LacI as a function of the logarithm of the total molar concentration of repressor tetramer.

For urea, Figure 1A compares binding isotherms at 0.212 *m* K^+ for urea molalities of 0, 0.796, 1.59, and 2.39 *m*. Figure 1B compares binding isotherms at 0.427 *m* K^+ for urea molalities of 0, 0.43, and 1.28 *m*. At both salt concentrations, the binding curve shifts to higher repressor concentrations with an increase in urea concentration, demonstrating that urea shifts the binding equilibrium toward reactants.

For GB, Figure 2 compares isotherms at ∼0.43 *m* K^+ for GB molalities of 0, 0.634, 1.25, and 2.06 *m*. Figure 2 shows that the binding curve shifts to lower protein concentrations with an increase in GB concentration, in accordance with the previous finding that GB and other osmolytes increase the stability of the repressor−O_1 operator complex (*43*). The shape of these isotherms, typical of noncooperative binding, is independent of solute and salt concentration. Fitted curves are nonlinear least-squares fits to a noncooperative isotherm involving only 1:1 (TO) LacI−SymL operator complexes (cf. ref *25*). No significant contribution from 1:2 (TO_2) repressor−operator complexes is expected for the conditions of these experiments (*25*, *30*).

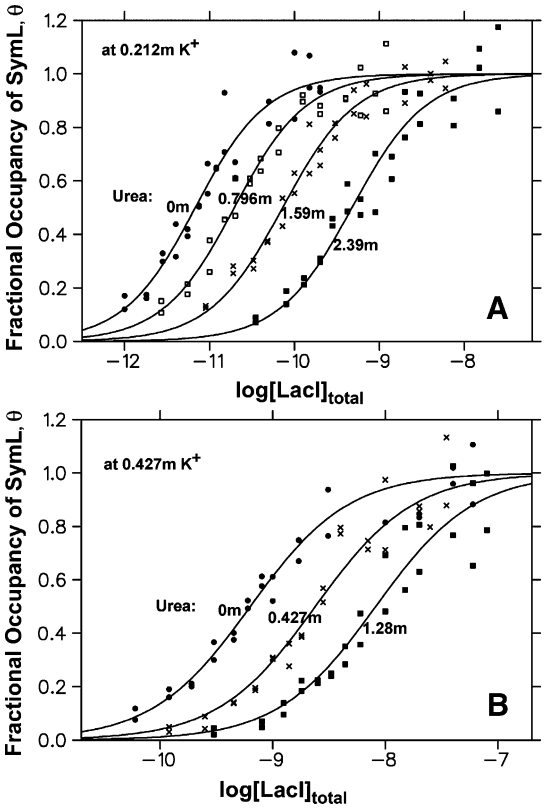

FIGURE 1: Effect of urea concentration on specific binding of lac repressor to lac operator DNA. Representative binding isotherms determined by the nitrocellulose filter assay in binding buffer (BB) at 24 °C in the absence and presence of urea are plotted as the fractional occupancy (θ) of the SymL operator site by LacI protein as a function of the logarithm of the total (bound and free, uncorrected for activity) molar concentration of LacI tetramer. (A) Lower-salt concentration titrations (0.212 *m* K^+). Isotherms at 6.9 pM DNA in the absence of urea (●) and at 1.8 pM DNA in 0.796 *m* urea (□), 1.59 *m* urea (×), or 2.39 *m* urea (■). LacI activity is 0.43. (B) Higher-salt concentration titrations (0.427 *m* K^+). Isotherms at 1.0 pM DNA in the absence of urea (●) and with 1.28 *m* urea (■) and at 1.25 pM DNA with 0.427 *m* urea (×). LacI activity is 0.39. In both panels, the curves are nonlinear least-squares fits to a noncooperative model with two equivalent sites for operator DNA on a repressor tetramer (see Materials and Methods), assuming that only 1:1 LacI−operator complexes (TO) form under the conditions that have been investigated. Table S1 summarizes the fitted values of the binding constants K_{TO}.

Binding constants K_{TO} obtained from these fits are tabulated together with results at these and other urea or GB molalities in Table S1 (in the Supporting Information). K_{TO} decreases with an increase in urea concentration and increases with an increase in GB concentration at a constant K^+ molality. At 0.427 *m* K^+, where K_{TO} is ∼4 × 10^9 M^{-1} in

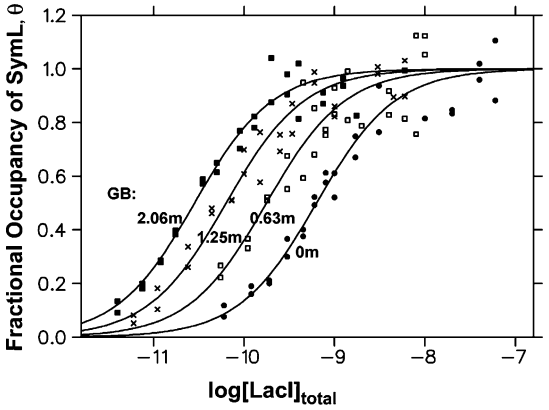

FIGURE 2: Effect of glycine betaine concentration on the specific binding of lac repressor to lac operator DNA. Representative binding isotherms determined by the nitrocellulose filter assay in binding buffer (BB) at 24 °C in the absence and presence of GB are plotted as the fractional occupancy (θ) of the SymL operator site by LacI protein as a function of the logarithm of the total molar concentration of LacI tetramer. Isotherms at 1.0 pM DNA in the absence of GB [0.427 *m* K^+ (●)] and in 2.06 *m* GB [0.411 *m* K^+ (■)], at 1.25 pM DNA in 0.63 *m* GB [0.422 *m* K^+ (□)], and at 1.50 pM DNA in 1.25 *m* GB [0.418 *m* K^+ (×)]. LacI activity is 0.36. The curves are nonlinear least-squares fits to a noncooperative model with two equivalent sites for operator DNA on a repressor tetramer (see Materials and Methods), assuming that only 1:1 LacI operator complexes (TO) form under the conditions that have been investigated. Table S1 summarizes the fitted values of the binding constants K_{TO}.

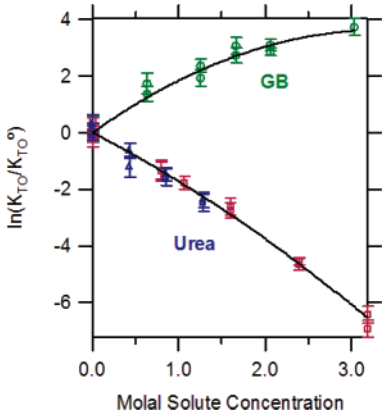

FIGURE 3: Effects of urea and GB molality on the repressor−operator binding constant K_{TO}. The quantity $\ln(K_{TO}/K^o_{TO})$, the natural logarithm of the ratio of the repressor−operator binding constant to that in the absence of added solute, is plotted against solute (GB or urea) molality. Binding constants K_{TO} and K^o_{TO} at the corresponding salt concentration are obtained from Table S1. For GB, all binding constants are at 0.427 *m* K^+ (circles); for urea, binding data at 0.212 *m* K^+ (squares) and 0.427 *m* K^+ (triangles) are plotted. Equations for the fitted curves are as follows: $\ln(K_{TO}/K^o_{TO}) = (-1.7 \pm 0.2)m_3 - (0.12 \pm 0.06)m_3^2$ for urea (combining data at both K^+ molalities) and $\ln(K_{TO}/K^o_{TO}) = (2.1 \pm 0.2)m_3 - (0.31 \pm 0.07)m_3^2$ for GB.

the absence of perturbing solute (cf. Table S1), addition of 1.28 *m* urea reduces K_{TO} by somewhat more than an order of magnitude (to $\sim 3 \times 10^8$ M^{-1}), while addition of 1.25 *m* GB increases K_{TO} by somewhat less than an order of magnitude (to $\sim 3 \times 10^{10}$ M^{-1}). Reduction in the K^+ concentration from 0.427 to 0.212 *m* increases K_{TO} more than two orders of magnitude (to $\sim 6 \times 10^{11}$ M^{-1}) in the absence of perturbing solute (cf. Table S1). This strong dependence of K_{TO} on salt concentration above 0.2 *m* (log−log [salt]-derivative $SK_{obs} = -7 \pm 1$ near 0.3 M salt; *30*) arises primarily from formation of the specific interface with operator. [A smaller, variable contribution to SK_{obs} is attributed to local wrapping of flanking nonoperator DNA on the repressor core (*30*).] Because both urea and GB at concentrations in the molal range significantly reduce the thermodynamic activity of KCl (*18*, *44*) and because K_{TO} depends so strongly on salt activity, it is necessary to correct the observed derivatives of K_{TO} with respect to solute (urea and GB) concentration to constant salt activity, by the method developed in Appendix I and summarized below.

Solute Concentration Dependences of the LacI−SymL Binding Constant K_{TO}. Binding constants K_{TO} for the entire range of urea and GB concentrations investigated are summarized in Figure 3, which plots the natural logarithm of the ratio of K_{TO} to the reference binding constant, K^o_{TO}, as a function of solute molality. [K^o_{TO} is the binding constant in the absence of the solute (urea or GB) at the same KCl concentration.] The effect of urea concentration on $\ln K_{TO}$ is linear within the uncertainty at low urea concentrations (<1.6 *m*) at both 0.212 and 0.427 *m* K^+. Data at higher urea concentrations (obtained at only 0.212 *m* K^+) reveal detectable downward curvature. The fit of all values of $\ln(K_{TO}/K^o_{TO})$ at both 0.212 and 0.427 *m* K^+ to a second-order polynomial in m_{urea} (0−3.2 *m* urea) is shown in Figure 3:

$$\ln(K_{TO}/K^o_{TO}) = (-1.7 \pm 0.2)m_{urea} - (0.12 \pm 0.06)m_{urea}^2 \quad (3)$$

The initial slope $(\partial \ln K_{TO}/\partial m_{urea})_{m_{urea}\to 0} = -1.7 \pm 0.2$ is most directly interpretable in terms of the ASA buried upon binding repressor to operator (see Discussion, and also refs *16* and *45*). The significant downward curvature (quadratic term) at higher molal concentrations of urea is opposite in direction from that expected if urea were only perturbing the intrinsic strength of the interaction and not affecting the structure of the DNA binding domain of the unbound lac repressor. Possible origins of this effect are discussed below (see the Discussion).

For GB, the plot of $\ln(K_{TO}/K^o_{TO})$ versus m_{GB} in Figure 3 is fitted as a quadratic function of m_{GB} (shown below in eq 4). The initial slope is opposite in sign but comparable in magnitude to that obtained with urea. Although the effects of the two solutes on $\ln(K_{TO}/K^o_{TO})$ are in opposite directions, both solute plots exhibit the same direction of curvature at higher solute concentrations. Stronger downward curvature is exhibited in GB than in urea, and indeed K_{TO} approaches a GB-independent plateau value or maximum at high GB concentrations; the predicted maximum from the fit (given below) is at ~3.5 *m* GB. The fit of all values of $\ln(K_{TO}/K^o_{TO})$ at 0.427 *m* K^+ to a second-order polynomial in m_{GB} (0−3.03 *m* GB) is shown in Figure 3:

$$\ln(K_{TO}/K^o_{TO}) = (2.1 \pm 0.2)m_{GB} - (0.31 \pm 0.07)m_{GB}^2 \quad (4)$$

As for urea, the initial slope $(\partial \ln K_{TO}/\partial m_{GB})_{m_{GB}\to 0} = 2.1 \pm 0.2$ is most directly interpretable in terms of the ASA buried upon binding repressor to operator (see Discussion, and also refs *16* and *45*). Possible origins of the strong curvature

(quadratic term) at higher GB concentrations are discussed below (see the Discussion). Both the direction and magnitude of this curvature are consistent with those observed in studies of protein thermal stability as a function of GB concentration (*45*, *46*).

Calculation of the Solute Concentration Dependence of the Repressor−Operator Binding Constant at Constant Salt Activity and in the Limit of Low Solute Concentrations. Effects of denaturants (urea and guanidinium chloride) on the standard free energy change for protein unfolding have traditionally been reported as "*m*-values" (*47*). We (*16*, *17*, *48*) applied a thermodynamic analysis to interpret these *m*-values in terms of differences in preferential interactions of denaturant with denatured and native states of the protein, using the solute partitioning model of preferential interaction coefficient (*16*). For such an analysis, the molal scale of solute concentrations is more fundamental. Hence, we analyze the dependence of $\ln K_{obs} = -\Delta G^{\circ}_{obs}/RT$ for the process of interest (here repressor−operator binding) as a function of molal concentration of solute (urea or GB).

Protein−nucleic acid interactions are very strongly dependent upon the type and thermodynamic activity of salt. In the determinations of the effects of urea and GB on repressor−operator binding constants reported here, the total K^+ molal concentration and individual anion molal concentrations (primarily Cl^-, also phosphate and EDTA) were maintained constant in each series of experiments. Because the osmolalities of three-component aqueous solutions of KCl and urea (or GB) are always lower than the sum of the osmolalities of the individual two-component solutions, the KCl activity in experiments at constant KCl molality decreases with an increase in concentration of urea or GB. The strong dependence of the repressor−operator binding constant on salt activity makes it necessary to correct the experimental data from constant salt molality to constant salt activity to isolate the effect of urea or GB on the LacI−SymL binding constant. This is most conveniently done at the level of the derivative $\partial \ln K_{obs}/\partial m_3$, using experimental thermodynamic data obtained by VPO (vapor pressure osmometry) for the interaction of urea and GB with KCl. Appendix 1 provides the derivation of the relationship between the derivative $(\partial \ln K_{obs}/\partial m_3)_{m_4}$ and the corresponding derivative at constant a_4, where component 4 is the salt and component 3 is the perturbing solute (urea and GB). The result is

$$\left(\frac{\partial \ln K_{obs}}{\partial m_3}\right)_{a_4} = \left(\frac{\partial \ln K_{obs}}{\partial m_3}\right)_{m_4} + \frac{SK_{obs}\Delta Osm_{34}}{2m_3m_4(1+\epsilon_{\pm})_{m_3}} \quad (5)$$

In eq 5, the quantity ΔOsm_{34} is the difference in osmolality between the three-component solute (urea and GB)/salt solution and the sum of its constituent two-component solutions at the same molalities (m_3 and m_4). [Because the quantity $RT\Delta Osm_{34}/m_3m_4$ is approximately equal to the derivative of chemical potential of urea or GB with respect to salt molality, it is expected to be relatively independent of m_3 over the range of molalities of interest here, including the situation $m_3 \rightarrow 0$ (cf. Appendix 1).] Thermodynamic consequences of any interactions of urea or GB with glycerol or other components of the binding buffer are expected to be insignificant and are therefore neglected. For three-component solutions containing KCl and either urea or GB, osmolality fitting functions to evaluate ΔOsm_{34} have been published (*18*).

Table 2: Calculation of $(\partial \ln K_{obs}/\partial m_3)_{m_3\rightarrow 0,a_4}$ from $(\partial \ln K_{obs}/\partial m_3)_{m_4,m_3\rightarrow 0}$

	$(\partial \ln K_{obs}/\partial m_3)_{m_4,m_3\rightarrow 0}$	correction to constant KCl activity	$(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0}$
urea effect	-1.7 ± 0.2	-0.34 ± 0.09[a]	-2.1 ± 0.2
GB effect	2.1 ± 0.2	-0.37 ± 0.10[b]	1.8 ± 0.2

[a] For KCl concentrations in range of 0.212−0.427 *m*. [b] For a KCl concentration of approximately 0.43 *m*.

In applications of eq 5 to the initial value of the derivative $(\partial \ln K_{obs}/\partial m_3)_{m_4,m_3\rightarrow 0}$, values of the electrolyte nonideality term $(1+\epsilon_{\pm})_{m_3=0}$ and of the salt derivative of the protein−DNA binding constant (SK_{obs}) in the absence of perturbing solute were used. For binding of lac repressor to plasmid operator DNA in the vicinity of 0.3 M salt, $-SK_{obs} = 7 \pm 1$ (*30*). For KCl solutions in the range 0.2−0.43 *m*, the nonideality term $(1+\epsilon_{\pm})_{m_3=0} = 0.893 \pm 0.001$ (*49*). Table A1 in Appendix 1 provides numerical values for the second term on the right-hand side of eq 5 for urea and GB at the KCl concentrations used in this study. Solute concentration derivatives $(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0}$ at constant salt activity for both urea and GB are calculated using eq 5 and listed in Table 2. For the salt concentration range of our experiments (0.21−0.43 *m* KCl), the KCl correction terms for urea and GB have the same sign and similar magnitudes. Thus, for urea (GB), the derivative $(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0}$ is larger (smaller) in magnitude than $(\partial \ln K_{obs}/\partial m_3)_{m_4,m_3\rightarrow 0}$.

Analysis of initial slopes $(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0}$ is not only convenient from the point of view of the correction to constant salt activity (eq 5), but also provides the most direct information regarding changes in ASA in the absence of solute. Possible origins of the changes in slope observed at higher concentrations of both urea and GB are discussed below.

ANALYSIS

Intrinsic Interaction Coefficient for Urea and Polar Amide Surface. To obtain intrinsic interaction coefficients for urea and different types of biopolymer surface, we analyze the preferential interactions of urea with a wide range of biopolymer surfaces, differing in composition and in total surface area (*18*). Our approach is based on dissection of urea−biopolymer interactions and urea effects on well-defined biopolymer processes to obtain intrinsic interaction coefficients for the solute and various types of biopolymer surface. This contrasts with analyses that obtain analogous quantities (transfer free energies) from the effect of the solute on the solubility of model compounds (*50*). To date, the two approaches have not been systematically compared. A limited comparison of free energies and enthalpies of transfer of two model peptides (di- and triglycine) from water to urea (*51*, *52*) calculated from isopiestic distillation and solubility data revealed significant systematic differences that were attributed to the treatment of activity coefficients in the analysis of the solubility data.

For our dissection of observed urea−biopolymer interactions, quantitative information about the interactions of urea with the protein surface exposed in unfolding [which is

typically relatively uncharged (<5% charged ASA), largely nonpolar (65−75%), with 15−20% polar amide ASA] is obtained from the characterization of urea effects on the stability of globular proteins. Myers et al. (*14*) compiled urea *m*-values for the set of 45 globular proteins for which structural data were available, and found a linear relationship (with a small positive intercept) between the *m*-value and the ΔASA of unfolding. Correction for the effect of disulfide bonds on the ASA of the unfolded state eliminated the intercept and yielded a proportionality between the *m*-value and the total ΔASA.

In general, the proteins in the Myers et al. (*14*) data set have stable folds, and exhibit midpoint unfolding concentrations of urea in excess of 4 M. Given the relatively high concentrations of urea used to denature proteins, the *m*-value only describes the initial slope if $\Delta G°$ is a linear function of urea concentration. (Initial slopes characterizing solute dependence on molal and molar concentration scales are the same.) Evidence for the validity of the linear extrapolation method (LEM) for urea is provided by detailed studies of the urea-induced helix−coil transitions of a series of marginally stable α-helical peptides at 273 K (*15*) and unfolding studies of marginally stable proteins as functions of temperature and urea concentration (*48*, *53*). For these marginally stable systems, effects of urea on stability are observable over the concentration range from 0 to 4−6 M; in all cases, the *m*-value is independent of urea concentration within experimental uncertainty, justifying the LEM. Theoretical justification for the validity of the LEM for urea is provided by the two-domain, solute partitioning model of the preferential interaction, which should be valid for non-Coulombic interactions of solutes with biopolymers (*16*, *45*). For solutes other than urea, especially in cases where the solute concentration dependence of bulk solute nonideality (i.e., in two-component solution) is larger in magnitude than that observed for urea, the LEM may be less successful (*45*).

We selected from the Myers et al. (*14*) data set the 14 proteins which lack both disulfide cross-links and heme groups, and added to this set an additional 26 proteins (with high-resolution folded structures and lacking disulfides and hemes) for which urea *m*-values have been recently reported at or near neutral pH [pH 5−8; in this range, typically no significant effect of pH on the *m*-value is observed (*54*)]. Determinations at temperatures other than 298 K were corrected to 298 K using the small temperature dependence of the *m*-value determined previously for lac HTH (*48*), weighted by the relative amounts of polar amide surface exposed in the two processes. In Figure 4A, these 40 *m*-values are plotted as a function of the total ΔASA of unfolding. (All protein *m*-values, conditions, PDB files, and ΔASA calculations are reported in Table S2 of the Supporting Information.) An excellent correlation is observed; the intercept is zero within the uncertainty of the best-fit line, indicating proportionality of the *m*-value to ΔASA_{total}. The error-weighted average and standard deviation of the *m*-value/ΔASA_{total} ratio for the 40-protein data set is 0.15 ± 0.03. This result is plotted as the line in Figure 4A. [For the disulfide-corrected, 45-protein set of Myers et al. (*14*), the proportionality constant is 0.14, within error of the current result.] The largest deviation from the proportionality of the urea *m*-value to ΔASA_{total} is observed for the largest protein in the data set, human glutathione transferase A1-1 (hGSTA1-1) (*55*). Unfolding of hGSTA1-1 is two-state (folded dimer → two unfolded monomers) and reversible. The surface exposed in unfolding hGSTA1-1 is only 13% amide (one of the two lowest values in the data set); for this protein,

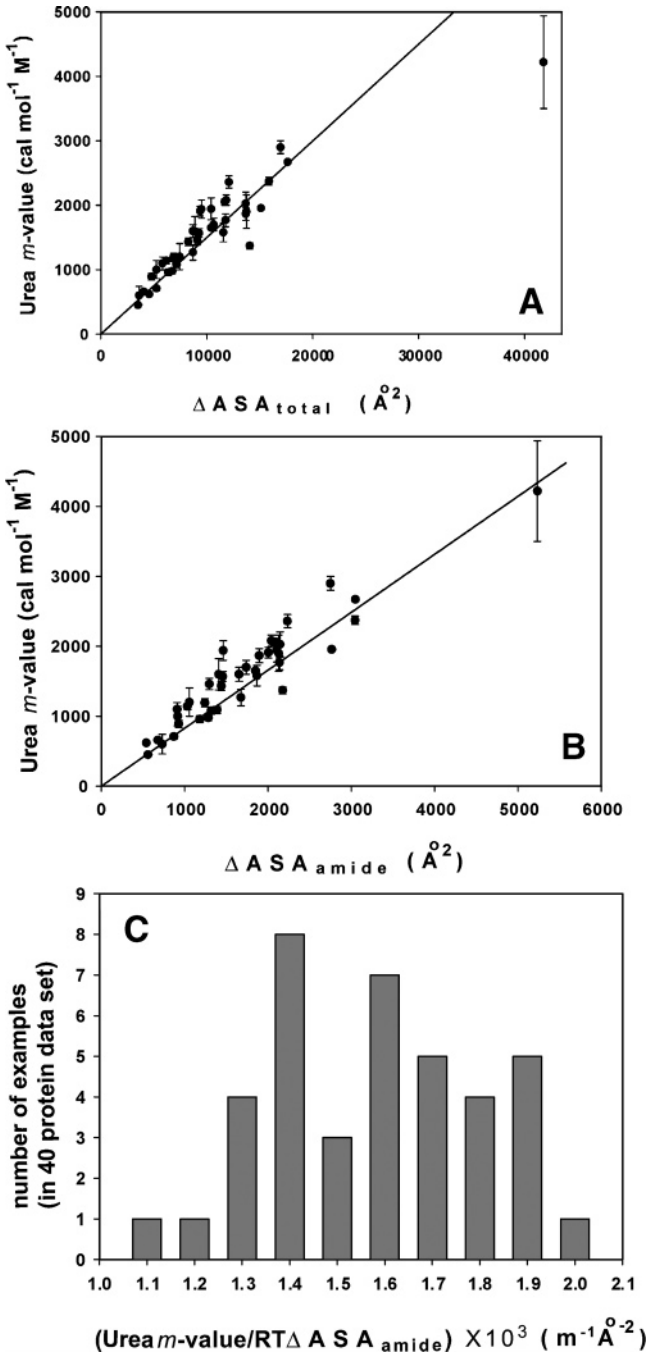

FIGURE 4: ASA analysis of urea *m*-values for unfolding of monomeric and dimeric globular proteins. (A) Literature determinations of urea *m*-values (converted to 25 °C) are plotted vs calculated values of the total ΔASA of unfolding for 40 monomeric and dimeric proteins (cf. Table S2 of the Supporting Information). The line corresponds to the error-weighted average value of the quantity *m*-value/ΔASA_{total} (0.15 ± 0.03). (B) The same set of urea *m*-values at 25 °C as in Figure 4A is plotted vs calculated values of the polar amide ΔASA of unfolding. The line corresponds to the error-weighted average value of the quantity *m*-value/ΔASA_{amide} (0.83 ± 0.18), from which an intrinsic urea−amide ASA interaction coefficient $\Gamma_{\mu_3}/m_3 ASA_{amide} = (1.4 \pm 0.3) \times 10^{-3}$ m^{-1} Å^{-2} is calculated. (C) Distribution of values of the intrinsic urea−amide ASA interaction coefficient (equal to *m*-value/*RT*ΔASA_{amide}) for the set of 40 monomeric and dimeric proteins.

Table 3: Quantifying Urea Accumulation at Polar Amide Surface[a]

process	example	m-value/ΔASA_{total} (cal mol^{-1} M^{-1} Å^{-2})	% polar amide ΔASA	urea accumulation per Å^2 polar amide ΔASA [$\Gamma_{\mu_3}/m_3 ASA_{amide}$ (m^{-1} Å^{-2})]	footnote
protein unfolding	set of 40 globular proteins (without S−S or heme groups)	0.15 ± 0.03	17 ± 2	$(1.4 \pm 0.3) \times 10^{-3}$	*b*
α-helix melting	$AcY(AEAAKA)_kFNH_2$; $\Delta ASA_{total} = 42 \pm 6$ Å^2 per residue	0.52 ± 0.07	49 ± 7	$(1.8 \pm 0.1) \times 10^{-3}$	*c*

system	ASA_{total} (Å^2)	$RT(\Gamma_{\mu_3}/m_3)/ASA_{total}$	% polar amide ASA	urea accumulation per Å^2 polar amide ASA [$\Gamma_{\mu_3}/m_3 ASA_{amide}$ (m^{-1} Å^{-2})]	footnote
BSA	2.78×10^4	0.12 ± 0.03	14.6	$(1.3 \pm 0.4) \times 10^{-3}$	*d*
double-stranded DNA	171 per nucleotide monomer	-0.07 ± 0.07	2.5	−	*e*

[a] Values of ΔASA calculated as described in Materials and Methods. [b] See the Supporting Information. [c] The m-value is from ref *15* and corrected to 298 K using the temperature effect (*48*). [d] The value of Γ_{μ_3}/m_3 is from Cannon et al. (manuscript in preparation) and agrees within uncertainty with ref *17*. [e] ASA analysis assumes that preferential interactions of urea with K^+ counterions are negligible [interaction of urea with KCl is weakly favorable (*44*)] (*18*).

the m-value/ΔASA_{total} ratio is smaller than the average by almost three standard deviations.

The m-value is related by fundamental three-component thermodynamics to the difference in urea preferential interaction coefficients for unfolded and folded states of the protein (*16*, *56*−*58*). Assuming that interactions of urea with the protein surface which is solvent-accessible in both folded and unfolded forms are not affected by unfolding, the m-value is a direct measure of the preferential interaction coefficient Γ_{μ_3} characterizing the net interaction of urea with the entire protein surface exposed in unfolding. The m-value/$RT\Delta ASA_{total}$ ratio is interpreted as the intrinsic preferential interaction coefficient $\Gamma_{\mu_3}/m_3 ASA_{total}$ (per unit of urea concentration and per unit area of protein surfaces exposed in unfolding) in the limit of low urea concentrations (*16*). To determine whether the favorable preferential interaction with urea is specific for some subset of the surface exposed in unfolding (e.g., amide, nonpolar, etc.), we determined thermodynamic coefficients characterizing the interactions of urea with native DNA and protein (BSA) surfaces (refs *17* and *18* and manuscript in preparation by J. Cannon et al.) and analyzed literature data (*15*) for effects of urea on α-helix melting (*16*).

For α-helix melting, Table 3 shows that the m-value/ΔASA_{total} ratio is more than 3 times larger than for protein unfolding at 25 °C, and that the fraction of total surface exposed upon unfolding which is polar amide surface (backbone and side chain ASA attributable to amide N and O atoms) is also more than 3 times larger for α-helix melting than for protein unfolding (0.52 vs 0.17; *16*). The per residue ΔASA_{total} varies from 36 to 47 Å^2 for various plausible models of the conformations of the K and E side chains in the helix (*16*, *18*), which introduces substantial uncertainty in the m-value/ΔASA_{total} ratio. However, the average amount of polar amide ASA exposed per residue of $(AEAAKA)_k$ α-helix unfolded is the same (20.6 ± 0.3 Å^2) for all three structural models assumed for the α-helix.

We proposed (*16*, *18*) that interaction coefficients for interaction of urea with protein and other biopolymer surfaces should be interpreted as primarily due to a favorable preferential interaction with polar amide surface. Since the m-value is independent of urea concentration, the intrinsic interaction coefficient for urea and the polar amide surface exposed in unfolding proteins and α-helices at low [urea] ($\Gamma_{\mu_3}/m_3 ASA_{amide}$) is equal to the m-value/$RT\Delta ASA_{amide}$ ratio. Calculated from the α-helix unfolding m-value, the intrinsic interaction coefficient for urea and polar amide surface is $(1.8 \pm 0.1) \times 10^{-3}$ m^{-1} Å^{-2}. Figure 4B shows the corresponding analysis for the protein unfolding data set of Figure 4A; in Figure 4B, 25 °C m-values are plotted as a function of the amount of polar amide surface exposed in unfolding. Comparison of panels A and B of Figure 4 shows that in general the m-values correlate equally well with either total ΔASA or polar amide ΔASA, which is the expected result because most of the proteins in the data set form a homologous series in which 17 ± 2% of the surface exposed in unfolding is polar amide surface. The line in Figure 4B represents the error-weighted average value of the intrinsic urea−amide interaction coefficient m-value/$RT\Delta ASA_{amide} = (1.4 \pm 0.3) \times 10^{-3}$ m^{-1} Å^{-2} for all 40 proteins (cf. the Supporting Information). The value for the protein hGSTA1-1, which deviates by the largest amount from the average in Figure 4A, agrees with the average in Figure 4B.

The intrinsic interaction coefficient for urea and polar amide surface obtained from the protein unfolding data [$(1.4 \pm 0.3) \times 10^{-3}$] is the same within uncertainty as that obtained from analysis of the α-helix unfolding m-value [$(1.8 \pm 0.1) \times 10^{-3}$], which supports the hypothesis that a favorable interaction of urea with polar amide surface is the dominant factor contributing to urea denaturation of proteins and of α-helices. Figure 4C plots the distribution of the protein folding data set around this error-weighted average. It will be important to ascertain whether the breadth of this distribution indicates the existence of a secondary interaction of urea with other types of protein surface, or whether it arises from experimental factors (e.g., deviations from two-state, reversible transitions or differences in baseline assignments used to calculate m-values) or from the choice of structural models for the folded and unfolded states of these proteins in solution used for calculation of ΔASA.

Native biopolymer surfaces, such as those of BSA and duplex DNA, differ profoundly in composition from the surfaces exposed in the processes of unfolding proteins and melting α-helices. These native surfaces are much more highly charged and much less nonpolar than are the surfaces exposed in unfolding and melting. As shown in Table 3, the intrinsic interaction of urea with BSA surface (2.78×10^4 Å^2), which is 14.6% polar amide but 29% charged and 53%

nonpolar, expressed per square angstrom of polar amide surface, is the same within error as that obtained for the surfaces exposed in unfolding proteins and α-helices. No significant effect of the large percentage of charged surface on this ratio is detected. For duplex DNA, the surface of which is 44% charged (anionic) but only 2.5% polar amide, no significant preferential interaction with urea is observed, indicating that urea is not significantly accumulated or excluded from anionic or other types of DNA surface. (Because we do not detect the predicted slight accumulation of urea at the 2.5% of the surface of duplex DNA which is amide-like, it is possible that there is a compensating slight exclusion from anionic or other DNA surface, which we are unable to dissect.)

Since protein unfolding comprises the most extensive data set in Table 3, we select the intrinsic urea−polar amide surface preferential interaction coefficient and uncertainty specified by these data for use in analysis of the effects of urea on binding of LacI to SymL. All other data in Table 3 are consistent with this result within the uncertainty. We therefore propose that, for any biopolymer process with equilibrium concentration quotient K_{obs}, the derivative $(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0}$ at low urea concentrations is, to a good approximation, proportional to the change in the amount of water-accessible polar amide surface (ΔASA_{amide}; Å^2) in the process.

$$(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0} = (1.4 \pm 0.3) \times 10^{-3} \times \Delta ASA_{amide} \quad (6)$$

When the data of Table 3 are analyzed as a linear combination of the contribution from polar amide, nonpolar, and anionic surface (by the method of Appendix 2), fitted coefficients for nonpolar and anionic surface are an order of magnitude smaller than for polar amide surface, and comparable to their uncertainties; the coefficient for polar amide surface agrees within uncertainty with eq 6 above.

Intrinsic Interaction Coefficient for Glycine Betaine and Anionic Surface. GB is strongly excluded from the surfaces of native BSA and DNA, and modestly excluded from native hen egg white lysozyme (*19*). GB modestly increases the thermal stability of RNase and lysozyme (*46*), and more significantly stabilizes the lac HTH (*19*). One interpretation of the stabilizing effects of GB is that it is modestly excluded from the surface exposed upon unfolding of these proteins (*19*). Unpublished experiments from our laboratory show that GB at concentrations up to 2 M has no significant effect on the stability of an $(AEAAKA)_6$ α-helix and therefore is neither excluded from nor accumulated at the surface exposed in melting this α-helix (J. Cannon, S. Heitkamp, et al., unpublished observations). We also find that GB is significantly excluded from the dicarboxylic acid salt potassium glutamate (J. Cannon and M. Kratz, unpublished observations), although it accumulates slightly in the vicinity of potassium chloride (*18*), indicating that the exclusion is probably attributed to the glutamate anion.

From these data, we are able to refine our previous analysis (*19*, *59*) to conclude that GB is detectably excluded only from anionic surface (phosphate and carboxylate) and that it exhibits no detectable preferential interaction with amide and nonpolar surface. The observed exclusion from anionic surface may be explained by the fact that GB, unlike water, has no hydrogen bond donor groups and so no ability to hydrogen bond to anionic groups of proteins or nucleic acids. Analysis of the published data for interactions of GB with BSA and duplex DNA as an interaction only with anionic surface yields an intrinsic interaction coefficient (per unit molality of GB and per unit anionic surface) of -4.0×10^{-3} m^{-1} Å^{-2}, as previously reported (*19*). In addition to this exclusion from anionic surface, GB has been observed to interact favorably with (i.e., accumulates at) some aromatic surfaces, forming a cation−π complex in the cocrystal structure of its transporter protein (*60*), and to selectively destabilize GC (but not AT) DNA at molar concentrations, presumably by a weak favorable interaction with G or C bases in the single-stranded state (*18*, *61*). A weak cation−π interaction of GB with tyrosines on the surface of the folded HTH which are buried in the protein−DNA interface could explain the greater-than-expected stabilization of the lac HTH against thermal unfolding by GB, as discussed below.

For biopolymer processes in which the interaction of GB with aromatic surface is not significant, the semilogarithmic derivative $(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0}$ at low GB concentrations is expected to be proportional to the change in the amount of water-accessible anionic surface ($\Delta ASA_{anionic}$; Å^2).

$$(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0} = -4.0 \times 10^{-3} \times \Delta ASA_{anionic} \quad (7)$$

We estimated an uncertainty in this value of ~20% based on the standard deviation of the much larger data set used to establish the corresponding proportionality for urea and polar amide surface (cf. eq 6).

DISCUSSION

Use of Urea To Quantify Burial of Polar Amide Surface in Repressor−Operator Binding: Contributions from the Interface and from Coupled Folding. Previously, we used heat capacity changes upon binding to quantify changes in nonpolar and polar water-accessible surface area in protein−ligand and protein−protein interactions (*62*); this analysis has successfully predicted the occurrence and often the amount of coupled folding in these protein associations (*4*). However, other large-scale temperature-dependent coupled processes also can contribute to the heat capacity change in reactions involving biopolymers (*63*−*66*). These coupled equilibria (e.g., protonation, changes in base stacking, and salt bridge disruption) are not generally related to overall structural changes, necessitating the development of more direct probes of coupled folding. Here we propose that the study of the urea concentration dependence of the binding constant provides one such method.

The initial slope describing the urea dependence of the logarithm of the LacI−SymL binding constant at low urea molality and at constant KCl activity is (cf. Table 2)

$$(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0} = -2.1 \pm 0.2 \quad (8)$$

Interpretation of the initial effect of urea on the binding constant of the LacI−SymL interaction (eq 8) using the intrinsic urea−polar amide interaction coefficient (eq 6) predicts that $(1.5 \pm 0.3) \times 10^3$ Å^2 of polar amide surface is buried upon specific binding of a SymL operator to the lac repressor tetramer. This thermodynamics-based structural prediction greatly exceeds the amount of amide surface

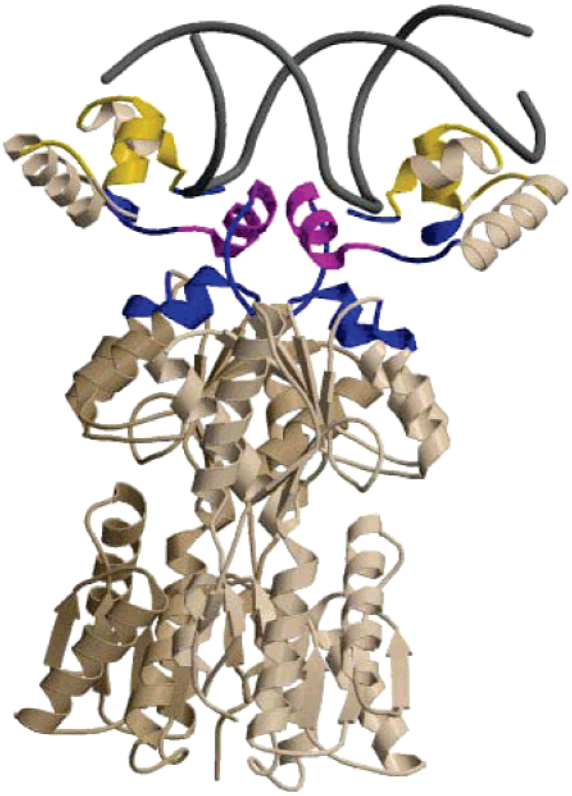

FIGURE 5: Structural model showing lac repressor surfaces buried in binding operator DNA. Binding of LacI to SymL involves the formation of the protein−DNA interface, folding of the hinge helices, and docking of the DNA binding domains on the core domains of the repressor. In this cartoon, residues buried during the formation of the rigid-body interface are colored yellow, residues buried during the folding of the hinge helices are colored red, and residues buried in the core−DBD interface are colored blue. This model is based on PDB entry 1EFA (*40*), and was created using MolScript (*76*) and Raster3D (*77*).

buried in the rigid-body LacI−SymL interface (491 Å^2; see Table 1 and Figure 5). However, this apparent discrepancy [$(1.0 \pm 0.3) \times 10^3$ Å^2] is fully explained by inclusion of the coupled conformational changes inferred from the structure. The DNA binding domains (DBDs), composed of helix−turn−helix (HTH) residues 1−49 and flexible tethers ("hinges") connecting the HTH to the core repressor (residues 50−58), are not observed in the crystal structure of the intact (DBD-containing) free repressor (*12*). Hence, in the free (unbound) form of the repressor, the DBDs are not assembled on the repressor core, whereas they are visualized in the structure of the complex in an ordered and docked state (see Figure 5). Indeed, solution NMR experiments reveal that the hinge residues are unstructured in the free state and fold upon binding to the SymL operator (*8*, *10*).

Coupled folding of the two hinge helices is predicted (see Table 1) to bury 515 Å^2 of polar amide surface (258 Å^2 for folding of each hinge helix). As shown in Figure 5, the folded hinge helices not only participate in the interface with SymL operator DNA but also participate in an interface between the LacI DBD and the LacI core in the operator complex (*40*). Docking of the pairs of DNA-bound hinge helices and HTH domains on the repressor core buries 2×260 Å^2 (520 Å^2) of polar amide surface. Therefore, the total amount of polar amide surface deduced to be buried in coupled folding of the hinge helices and in forming an interface between the folded headpieces/hinge helices and core is 1035 Å^2 of polar amide surface. This value calculated from structural data is the same as that obtained by urea analysis within the limits of uncertainty.

The structural data of Table 1 reveal why urea is a sensitive probe of the existence and amount of folding that is coupled to binding in protein interactions. Only 1144 Å^2 of total surface is buried in folding the two nine-residue hinge helices, but 45% of this is polar amide surface (515 Å^2). Therefore, the contribution of folding 18 residues and burying a total of only 1144 Å^2 of surface to the observed effect of urea on the binding constant is just as large as the contribution from the burial of 2049 Å^2 of surface in the interface between the folded DBD and core repressor (which is only 25% polar amide), and from the burial of 3689 Å^2 of surface in the repressor−operator (protein−DNA) interface (which is only 13% polar amide surface because of the lack of these functional groups on the accessible surface of duplex DNA).

If the hinge helices are completely unfolded in the free state of repressor and completely folded and docked in the complex with SymL, and if no other coupled processes contribute to the thermodynamics, the heat capacity change predicted (*4*) from the ASA data in Table 1 and Figure 5 is approximately −0.80 kcal/K. By contrast, the experimental value is approximately −1.5 kcal/K. Calorimetric studies with a cross-linked dimer of the DNA binding domain (residues 1−62) are in progress in an attempt to determine the other factor(s) contributing to the heat capacity change of LacI−SymL binding.

It is instructive to compare the structural prediction of the polar amide ASA analysis of the urea effect with that which would be obtained using the original expression of Myers et al. (*14*) in terms of total ASA, applicable only to analysis of unfolding data for the homologous series of globular proteins. Analysis of the urea dependence of the repressor−operator binding constant (eq 8), using the original protein folding expression [m-value/$\Delta ASA_{total} = 0.14$ (*14*)], yields a predicted ΔASA_{total} of -8.9×10^3 Å^2, which exceeds the structural value by 74% (cf. Table 1; 5.1×10^3 Å^2 of repressor ASA buried in complexation). Even if the ASA of DNA buried in complexation were included, which cannot be justified given the observation of no significant preferential interaction of urea with duplex DNA (*18*; see Table 3), the protein folding correlation overpredicts the ΔASA_{total} by 29%. The overprediction occurs because the m-value/ΔASA_{total} ratio obtained from analysis of globular protein folding data is only applicable to processes which expose or bury surface with the same polar amide percentage (17%) as that characteristic of folding globular proteins. However, the percentage of total surface buried in LacI−SymL complexation which is polar amide surface is 22% (including DNA surface) or 30% (including only repressor surface).

Interpreting the Urea Concentration Dependence of $d\ln K_{obs}/dm_3$ for Repressor−Operator Binding. Figure 3 and eq 3 indicate that the magnitude of the derivative $d\ln K_{obs}/dm_3$ increases with an increase in urea concentration in the range examined (up to 3.19 m). This urea concentration dependence of $d\ln K_{obs}/dm_3$ is in the direction opposite from that expected from the solute partitioning model, in which three factors make contributions to the concentration dependence of $d\ln K_{obs}/dm_3$: (1) urea nonideality ($1 + \epsilon_3$), (2) attenuation of the urea local−bulk partition coefficient (*45*),

and (3) the effect of replacement of local water by urea when it accumulates at biopolymer surfaces (*16*). These three factors, predicted to be approximately equal in significance, together are expected to reduce the magnitude of $d\ln K_{obs}/dm_3$ by approximately 25% at 1.59 *m* urea and 43% at 3.19 *m* urea. However, we observe that the magnitude of $d\ln K_{obs}/dm_3$ increases by approximately 18% at 1.59 *m* urea and 35% at 3.19 *m* urea.

A plausible origin of these effects is urea-driven unfolding of either the HTH or the core of lac repressor. At 298 K and under salt conditions which differ from those investigated here, the core of lac repressor unfolds in a very narrow transition at approximately 3 M urea (*34*, *67*). At 298 K, the HTH unfolds in a very broad transition starting in the absence of urea at low salt (25 mM K_2HPO_4; *48*) and near 1 M urea at 0.2 M K_2HPO_4 (*67*), which is probably a more stabilizing salt than our binding buffer. Therefore, urea-induced unfolding of the DNA-binding HTH appears to be the most likely explanation of the direction and magnitude of curvature in Figure 3. Calculations of the curvature predicted as a result of a broad unfolding transition for the HTH are consistent with that observed (data not shown).

Use of Glycine Betaine To Quantify Burial of DNA Phosphate Surface in Repressor−Operator Binding. The initial slope describing the GB dependence of the logarithm of the LacI−SymL binding constant at low GB molality and at constant KCl activity is (cf. Table 2)

$$(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0} = 1.8 \pm 0.2 \qquad (9)$$

Comparison of eq 9 with eq 7 for the repressor−operator interaction yields the following structural prediction from the thermodynamic analysis: $\Delta ASA_{anionic} = -(4.5 \pm 1.0) \times 10^2$ Å^2. By this analysis, approximately 450 Å^2 of anionic surface is predicted to be buried upon binding one SymL operator DNA to one pair of headpieces on a LacI tetramer. This prediction is ∼70 ± 16% of the amount of anionic surface which we calculated as buried in the LacI−SymL (TO) interface (632 Å^2, obtained by doubling the amount of anionic ASA buried in the more highly resolved half-site).

One possible explanation of the discrepancy between the structural value and the GB-based prediction for the ΔASA of anionic surface, which, if correct, will limit the utility of GB as a quantitative probe for changes in exposure of anionic surface in biopolymer processes, involves the same proposal discussed above to rationalize the greater-than-expected stabilization of the folded state of the lac HTH by GB. Interaction of the cationic end of GB with the aligned tyrosines in the DNA binding site of the folded HTH would not only stabilize the folded state against unfolding but also compete with DNA binding, and hence reduce the otherwise stabilizing effect of GB on a process which buries anionic surface. A second possible explanation, developed in Appendix 2, is that residual water in the interface of the complex, not detected in the crystal or NMR structures, reduces the magnitude of the GB effect. Even with the uncertainty regarding this effect, it appears that GB can be used as a semiquantitative probe of changes in exposure of anionic surface in processes involving proteins and double-stranded DNA. For example, the rate-determining conformation change in open complex formation by *E. coli* RNA polymerase ($E\sigma^{70}$) at the λP_R promoter is significantly disfavored by GB, consistent with exposure of anionic surface in this step (*68*). Since DNA phosphates presumably are buried in this process, the GB result (*68*) may indicate exposure of carboxylates on RNA polymerase in this step. One possible candidate, shown to be exposed in open complex formation (*69*), is the negatively charged N-terminal domain of σ^{70}, exposure of which would unmask the active site. Since GB is inferred to interact with guanine and/or cytosine bases in single-stranded DNA (*18*), perhaps also by a cation−π interaction involving stacked bases, it will probably not be a suitable probe of processes involving single-stranded DNA. If GB were also excluded from amide surface, as originally proposed (*19*, *50*), the discrepancy between structural and thermodynamic analyses of the change in "GB-affected" surface area would be even greater.

Interpreting the Glycine Betaine Concentration Dependence of $d\ln K_{obs}/dm_3$ for Repressor−Operator Binding. Figure 3 and eq 4 indicate that the magnitude of the derivative $d\ln K_{obs}/dm_3$ decreases with an increase in GB concentration in the range examined (up to 3.03 *m*). This GB concentration dependence of $d\ln K_{obs}/dm_3$ is consistent with an attenuation model (*45*), in which the GB concentration dependence of its nonideality in the local domain is expected to be attenuated relative to that of the bulk domain due to GB−protein surface interaction in the local domain. The same amount of attenuation as that deduced for the GB effect on lac HTH unfolding (*45*) can be used to interpret the curvature of $d\ln K_{obs}/dm_3$ in our LacI binding study. (Details of the calculation are given in the Supporting Information.) Using the GB attenuation model (*45*), $\ln(K_{obs}/K^o_{obs})$ is expressed as a fourth-power polynomial in m_3. Fitting experimental values of $\ln(K_{obs}/K^o_{obs})$ to this polynomial yields $(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0} = 1.6 \pm 0.2$ which is the same within uncertainty as determined above (see eq 9) from analyzing the initial slope $(\partial \ln K_{obs}/\partial m_3)_{m_4,m_3\rightarrow 0}$ of the quadratic fitting of $\ln(K_{obs}/K^o_{obs})$ versus m_3.

CONCLUSION

This study provides the basis for, and an example of the use of, urea as a quantitative probe of coupled folding and other large-scale coupled conformational changes in protein−DNA interactions and other protein associations. Urea selectively detects changes in exposure of polar amide surface, and consequently is a very sensitive probe of α-helix formation (or disruption) and other local folding or unfolding processes. Urea is also a sensitive probe of formation (or disruption) of protein−protein interfaces coupled to DNA binding, and is relatively insensitive to the formation of the protein−duplex DNA interface because of the virtual absence of water-accessible amide-like surface on the DNA double helix. Urea should be very useful in determining the order and the magnitude of coupled conformational changes that occur in a multistep protein−DNA binding process, such as formation of an open promoter complex, by observing urea concentration-dependent effects on the rate and equilibrium constants of the individual steps of the mechanism. Studies of urea effects on open complex formation by RNA polymerase are in progress.

APPENDIX 1

Conversion of $(\partial \ln K_{obs}/\partial m_3)_{m_4}$ to $(\partial \ln K_{obs}/\partial m_3)_{a_4}$. Here we obtain eq 5 in the text, which relates the experimentally determined derivative $(\partial \ln K_{obs}/\partial m_3)_{m_4}$ at constant temperature, pressure, and molality of a 1-1 salt (4) to the derivative $(\partial \ln K_{obs}/\partial m_3)_{a_4}$ at constant salt activity a_4 and thereby eliminates the effect on $\ln K_{obs}$ of changes in salt activity brought about by changing solute (urea or GB) concentration at constant salt molality. This analysis is accomplished using standard manipulations of partial derivatives. The basic mathematical relationship between $(\partial \ln K_{obs}/\partial m_3)_{m_4}$ and $(\partial \ln K_{obs}/\partial m_3)_{a_4}$ in an excess of solute components 3 and 4 [relative to biopolymer participants (*70*)] is

$$(\partial \ln K_{obs}/\partial m_3)_{a_4} = (\partial \ln K_{obs}/\partial m_3)_{m_4} + (\partial \ln K_{obs}/\partial m_4)_{m_3} (\partial m_4/\partial m_3)_{a_4} \quad \text{(A1)}$$

The partial derivatives in the additive correction term in eq A1 are evaluated using experimental data:

$$(\partial \ln K_{obs}/\partial m_4)_{m_3} \cong SK_{obs}/m_4 \quad \text{(A2)}$$

where $SK_{obs} \equiv (\partial \ln K_{obs}/\partial \ln C_4)_{m_3} \cong (\partial \ln K_{obs}/\partial \ln m_4)_{m_3}$

$$(\partial m_4/\partial m_3)_{a_4} = -\frac{\mu_{43}}{\mu_{44}} = -\frac{\mu_{34}}{\mu_{44}} \quad \text{(A3)}$$

$$\text{where } \mu_{34} \equiv (\partial \mu_3/\partial m_4)_{m_3} \cong RT\Delta \mathrm{Osm}_{34}/m_3 m_4 \quad \text{(A4)}$$

$$\text{and } \mu_{44} \equiv (\partial \mu_4/\partial m_4)_{m_3} = 2RT(1+\epsilon_{\pm})_{m_3}/m_4 \quad \text{(A5)}$$

Euler reciprocity is used in the second equality of eq A3. In eq A4, $\Delta \mathrm{Osm}_{34} \equiv \mathrm{Osm}(m_3,m_4) - \mathrm{Osm}(m_3) - \mathrm{Osm}(m_4)$, where $\mathrm{Osm}(m_3)$ and $\mathrm{Osm}(m_4)$ are osmolalities of two-component solutions at the same molality as that of these solutes in the three-component solution with osmolality $\mathrm{Osm}(m_3,m_4)$. The approximate relationship between μ_{34} and $\Delta \mathrm{Osm}_{34}$, sufficiently accurate for this purpose, was originally presented by Robinson and Stokes (*71*) and was recently derived by Anderson and Record (*72*). In eq A5, $\epsilon_{\pm}$ symbolizes the derivative $\partial \ln \gamma_{\pm}/\partial \ln m_4$ of the mean ionic activity coefficient of the 1-1 salt with respect to its concentration.

Combining eqs A2−A5 in eq A1 yields

$$\left(\frac{\partial \ln K_{obs}}{\partial m_3}\right)_{a_4} = \left(\frac{\partial \ln K_{obs}}{\partial m_3}\right)_{m_4} + \frac{SK_{obs}\Delta \mathrm{Osm}_{34}}{2m_3 m_4 (1+\epsilon_{\pm})_{m_3}} \quad \text{(A6)}$$

In the limiting case where m_3 approaches zero, the log−log [salt] derivative SK_{obs} and the activity coefficient derivative $\epsilon_{\pm}$ may be approximated by their values in the absence of the perturbing solute (component 3). The quantity $\Delta \mathrm{Osm}_{34}/m_3 m_4 = \mu_{34}/RT$ does not vanish as the urea concentration is reduced because it represents the intrinsic effect of salt concentration on the chemical potential of urea. For binding of lac repressor to plasmid operator DNA, $SK_{obs} = -7.0 \pm 1.0$ in the vicinity of 0.3 *m* KCl (*30*). For KCl solutions in the range 0.2−0.43 *m*, $(1+\epsilon_{\pm})_{m_3=0} = 0.893 \pm 0.001$ (*49*).

APPENDIX 2

Analysis of the Solute Concentration Dependence of $\ln K_{obs}$ Using the Solute Partitioning Model: Initial Slope Method.

	KCl molality (m_4)	SK_{obs} for LacI−SymL	$\Delta \mathrm{Osm}_{34}/m_3 m_4$ at $m_3 \rightarrow 0$	$2(1+\epsilon_{\pm})_{m_3=0}$	$SK_{obs}\Delta \mathrm{Osm}_{34}/2m_3 m_4 (1+\epsilon_{\pm})_{m_3}$
urea	0.212	−7.0	−0.09 ± 0.02	1.788	−0.35 ± 0.09
	0.427	−6.3	−0.09 ± 0.01	1.785	−0.33 ± 0.09
GB	0.212	−7.0	−0.12 ± 0.01	1.788	−0.43 ± 0.08
	0.404	−6.3	−0.10 ± 0.02	1.785	−0.34 ± 0.09
	0.427	−6.3	−0.12 ± 0.02	1.785	−0.42 ± 0.08

Hong (*23*) generalized the basic thermodynamic relationship that expresses a concentration effect of a nonelectrolyte solute (3) on the observed equilibrium constant K_{obs} (or free energy change $\Delta G^{o}_{obs} = -RT \ln K_{obs}$) for a biopolymer process from a system containing one perturbing solute (*58*, *73*) to the case where two perturbing solutes (e.g., GB and KCl) are present:

$$\left(\frac{\partial \ln K_{obs}}{\partial m_3}\right)_{a_4} \cong \frac{\Delta \Gamma_{\mu_3}}{m_3}(1+\epsilon_3)_{a_4} \quad \text{(A7)}$$

where Γ_{μ_3} is a solute−biopolymer preferential interaction coefficient, $\Delta \Gamma_{\mu_3}$ is the stoichiometrically weighted difference in preferential interaction coefficients between products and reactants, and $\epsilon_3 \equiv (\partial \ln \gamma_3/\partial \ln m_3)_{T,P,a_4}$. Major approximations in eq A7 are discussed by Hong (*23*). In the limit $m_3 \rightarrow 0$, but maintaining the excess of solute over any biopolymer reactant or product (denoted by numerical number 2) ($m_3 \gg m_2$), there is no significant difference between the bulk and the total solute concentration (*74*), and the initial slope of $\ln K_{obs}$ versus m_3 is given by

$$\left(\frac{\partial \ln K_{obs}}{\partial m_3}\right)^{m_3 \rightarrow 0}_{a_4} = \left(\frac{\Delta \Gamma_{\mu_3}}{m_3}\right)^{m_3 \rightarrow 0} \quad \text{(A8)}$$

For the interaction of a nonelectrolyte solute with a biopolymer, in the solute partitioning model (also called the local-bulk or two-domain model) (*16*, *75*), Γ_{μ_3}/m_3 is expressed in terms of molecular parameters:

$$\frac{\Gamma_{\mu_3}}{m_3} = \frac{B_3}{m_3} - \frac{B_1}{m_1^{\bullet}} = \frac{(K_p - 1)b_1^{o}\mathrm{ASA}}{m_1^{\bullet}(1 + K_p S_{1,3} m_3/m_1^{\bullet})} \quad \text{(A9)}$$

In the first part of eq A9, B_3 and B_1 are the numbers of solute and water molecules in the local domain at the surface of a biopolymer, respectively, and $m_1^{\bullet} = 55.5$ mol of H_2O/kg of H_2O. In the second part of eq A9, the local-bulk solute partition coefficient K_p is defined as $K_p \equiv (B_3/B_1)/(m_3^{bulk}/m_1^{\bullet})$, the hydration of the biopolymer per unit area in the absence of solute $b_1^{o} = B_1^{o}/\mathrm{ASA}$, and $S_{1,3}$ is the cumulative exchange stoichiometry (the number of water molecules displaced from the local domain divided by the number of solute molecules accumulated) (*16*, *75*). As $m_3 \rightarrow 0$, using the superscript "o" to denote this limit

$$\left(\frac{\Gamma_{\mu_3}}{m_3}\right)^{m_3 \rightarrow 0} = \frac{B_3}{m_3} - \frac{B_1^{o}}{m_1^{\bullet}} = \frac{(K_p^{o} - 1)b_1^{o}\mathrm{ASA}}{m_1^{\bullet}} \quad \text{(A10)}$$

For the binding of a protein to a nucleic acid where the creation of the binding interface(s) is characterized by reductions in ASA of different types of biopolymer surface and where some water molecules (and, in principle, some

solute molecules) are retained in these interface(s), eqs A8 and A10 yield

$$\left(\frac{\partial \ln K_{obs}}{\partial m_3}\right)_{T,P,a_4}^{m_3\rightarrow 0} = \left[\left(\frac{B_{3,I}}{m_3} - \frac{B_{1,I}}{55.5}\right) + \sum_i f_i \left(\frac{\Gamma_{\mu_3}}{m_3 ASA}\right)_i^{m_3\rightarrow 0} \Delta ASA_i\right] \quad \text{(A11)}$$

In eq A11, $B_{3,I}$ and $B_{1,I}$ are the numbers of solute and water molecules retained in the interface of the complex, respectively. Equation A11 assumes that contributions to $\Delta\Gamma_{\mu_3}$ from interactions of solute with different types of biopolymer surface are additive (*16*−*19*). In eq A11, f_i is the fraction of total surface of the *i*th type and ΔASA_i is the change in the amount of water-accessible biopolymer surface of this type in the process. The sum in eq A11 is over all types of biopolymer surface (*i*) for which ASA_i changes; types of surface considered at the current level of resolution include anionic, polar amide, other polar, cationic, and nonpolar (*18*, *19*). For binding processes, these ΔASA_i values are typically negative (as a result of forming the binding interfaces), although coupled conformational changes may expose some biopolymer surface (e.g., coupled unfolding). Each ΔASA_i in eq A11 should be calculated without regard for any buried water (or solute) within the interface (no such water is present in relevant interfaces of the lac repressor−operator complex in 1EFA). Then the term $(\Gamma_{\mu_3}/m_3 ASA)_i^{m_3\rightarrow 0}\Delta ASA_i$ represents the contribution to $(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0}$ from the *i*th type of surface for the situation where its interactions with water and solute are eliminated during the formation of the complex.

For urea, the quantity $(\Gamma_{\mu_3}/m_3 ASA)_i^{m_3\rightarrow 0}$ is significantly different from zero only for interactions with polar amide surface (*18*). Interactions with cationic, anionic, and nonpolar surface are an order of magnitude weaker and comparable to their uncertainties. The observation that GB has no significant effect on α-helix stability (J. Cannon, S. Heitkamp, et al., unpublished observations), together with the data analyzed previously (*19*), indicates that $(\Gamma_{\mu_3}/m_3 ASA)_i^{m_3\rightarrow 0}$ is most significant for interaction of GB with anionic surface, although interactions with other types of biopolymer surface may also contribute (see the Discussion).

The term $B_{1,I}$ (and $B_{3,I}$ if relevant) in eq A11 accounts for the effect of any water (or solute) retained in the interface. No water is detected in the binding interface in crystal and NMR structures of repressor−operator complexes, though the 2.6 Å resolution of the crystal structure cannot unambiguously allow one to conclude that water is totally absent from the interface. For every 10 waters remaining in the interface of the complex, the predicted value of the derivative $(\partial \ln K_{obs}/\partial m_3)_{a_4,m_3\rightarrow 0}$ (for any solute) is reduced by the amount $10/55.5 = 0.18$ from that calculated by considering the contributions of the different types of dehydrated surface in eq A11. Therefore, if the presence of 40 residual water molecules in the interface of the complex were the explanation for the deviation between predicted and observed effects of GB [$\Delta(\partial \ln K_{obs}/\partial m_3) = 2.5-1.8 = 0.7$], 40 water molecules would also contribute to the effect of urea on K_{obs} according to eq A11. If 40 water molecules were retained in the interfaces of the repressor−operator complex, the contribution to $\partial \ln K_{obs}/\partial m_{urea}$ ascribed to burial of polar amide surface would be changed from −2.1 to −1.4, a 33% reduction in the magnitude of the thermodynamic prediction for polar amide ΔASA, thereby introducing a 33% discrepancy between structural and thermodynamic predictions. Since such a discrepancy is well outside of experimental uncertainty, we interpret it to mean that substantially fewer than 40 water molecules are present in the interfaces of the repressor−operator complex, and that the discrepancy in the GB result is largely due to other factors, as discussed in the text.

ACKNOWLEDGMENT

We thank Dr. Charles Anderson, Jonathan Cannon, and Melissa Anderson for discussions of preferential interactions of these solutes and for their comments on the manuscript and Megan Kratz, Sara Heitkamp, and Jonathan Cannon for the use of their unpublished data on the preferential interactions of GB and urea with biopolymer surfaces. We are grateful to Dr. Linda Jen-Jacobson and the reviewers for their valuable comments and suggestions for revision.

SUPPORTING INFORMATION AVAILABLE

Observed values of K_{TO} at the concentrations of urea, GB, and KCl investigated here and the corrected values of K_{TO} in GB experiments in which KCl and glycerol molality varied slightly with GB molality; individual values of surface area changes and corresponding *m*-values for the set of monomeric and dimeric proteins summarized in Table 3; and the details of the thermodynamic analysis using a solute partitioning model. This analysis is used to interpret the curvature in $\ln K_{obs}$ as a function of GB concentration. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) Intrinsic disorder and protein function, *Biochemistry 41*, 6573−82.
2. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life, *J. Mol. Biol. 337*, 635−45.
3. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions, *Nat. Rev. Mol. Cell Biol. 6*, 197−208.
4. Spolar, R. S., and Record, M. T. (1994) Coupling of Local Folding to Site-Specific Binding of Proteins to DNA, *Science 263*, 777−84.
5. Whitson, P. A., Olson, J. S., and Matthews, K. S. (1986) Thermodynamic analysis of the lactose repressor-operator DNA interaction, *Biochemistry 25*, 3852−8.
6. Ha, J. H., Spolar, R. S., and Record, M. T., Jr. (1989) Role of the hydrophobic effect in stability of site-specific protein-DNA complexes, *J. Mol. Biol. 209*, 801−16.
7. 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*, 1186−206.
8. Spronk, C. A., Bonvin, A. M., Radha, P. K., Melacini, G., Boelens, R., and Kaptein, R. (1999) The solution structure of Lac repressor headpiece 62 complexed to a symmetrical lac operator, *Struct. Folding Des. 7*, 1483−92.

9. Kalodimos, C. G., Folkers, G. E., Boelens, R., and Kaptein, R. (2001) Strong DNA binding by covalently linked dimeric Lac headpiece: Evidence for the crucial role of the hinge helices, *Proc. Natl. Acad. Sci. U.S.A. 98*, 6039−44.
10. Kalodimos, C. G., Bonvin, A. M. J. J., Salinas, R. K., Wechselberger, R., Boelens, R., and Kaptein, R. (2002) Plasticity in protein-DNA recognition: lac repressor interacts with its natural operator O1 through alternative conformations of its DNA-binding domain, *EMBO J. 21*, 2866−76.
11. Kalodimos, C. G., Biris, N., Bonvin, A. M., Levandoski, M. M., Guennuegues, M., Boelens, R., and Kaptein, R. (2004) Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes, *Science 305*, 386−9.
12. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer, *Science 271*, 1247−54.
13. Spronk, C. A., Slijper, M., van Boom, J. H., Kaptein, R., and Boelens, R. (1996) Formation of the hinge helix in the lac repressor is induced upon binding to the lac operator, *Nat. Struct. Biol. 3*, 916−9.
14. 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*, 2138−48.
15. Scholtz, J. M., Barrick, D., York, E. J., Stewart, J. M., and Baldwin, R. L. (1995) Urea Unfolding of Peptide Helices as a Model for Interpreting Protein Unfolding, *Proc. Natl. Acad. Sci. U.S.A. 92*, 185−9.
16. 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* (Suppl. 4), 72−85.
17. Courtenay, E. S., Capp, M. W., and Record, M. T. (2001) Thermodynamics of interactions of urea and guanidinium salts with protein surface: Relationship between solute effects on protein processes and changes in water-accessible surface area, *Protein Sci. 10*, 2485−97.
18. 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*, 14744−58.
19. Felitsky, D. J., Cannon, J. G., Capp, M. W., Hong, J., Van Wynsberghe, A. W., Anderson, C. F., and Record, M. T., Jr. (2004) The exclusion of glycine betaine from anionic biopolymer surface: Why glycine betaine is an effective osmoprotectant but also a compatible solute, *Biochemistry 43*, 14732−43.
20. Laiken, S. L., Gross, C. A., and Von Hippel, P. H. (1972) Equilibrium and kinetic studies of *Escherichia coli* lac repressor-inducer interactions, *J. Mol. Biol. 66*, 143−55.
21. Rosenberg, J. M., Khallai, O. B., Kopka, M. L., Dickerson, R. E., and Riggs, A. D. (1977) Lac repressor purification without inactivation of DNA binding activity, *Nucleic Acids Res. 4*, 567−72.
22. Levandoski, M. M. (1995) Ph.D. Disertation, University of Wisconsin, Madison, WI.
23. Hong, J. (2004) Ph.D. Disertation, University of Wisconsin, Madison, WI.
24. Butler, A. P., Revzin, A., and von Hippel, P. H. (1977) Molecular parameters characterizing the interaction of *Escherichia coli* lac repressor with non-operator DNA and inducer, *Biochemistry 16*, 4757−68.
25. Levandoski, M. M., Tsodikov, O. V., Frank, D. E., Melcher, S. E., Saecker, R. M., and Record, M. T. (1996) Cooperative and anticooperative effects in binding of the first and second plasmid O-sym operators to a LacI tetramer: Evidence for contributions of non-operator DNA binding by wrapping and looping, *J. Mol. Biol. 260*, 697−717.
26. Simons, A. T. D., von Wilcken-Bergmann, B., and Muller-Hill, B. (1984) Possible ideal lac operator: *Escherichia coli* lac operator-like sequences from eukaryotic genomes lack the central G X C pair, *Proc. Natl. Acad. Sci. U.S.A. 81*, 1624−8.
27. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
28. Law, S. M., Bellomy, G. R., Schlax, P. J., and Record, M. T., Jr. (1993) In vivo thermodynamic analysis of repression with and without looping in lac constructs. Estimates of free and local lac repressor concentrations and of physical properties of a region of supercoiled plasmid DNA in vivo, *J. Mol. Biol. 230*, 161−73.
29. Riggs, A. D., Suzuki, H., and Bourgeois, S. (1970) Lac repressor-operator interaction. I. Equilibrium studies, *J. Mol. Biol. 48*, 67−83.
30. Tsodikov, O. V., Saecker, R. M., Melcher, S. E., Levandoski, M. M., Frank, D. E., Capp, M. W., and Record, M. T. (1999) Wrapping of flanking non-operator DNA in lac repressor-operator complexes: Implications for DNA looping, *J. Mol. Biol. 294*, 639−55.
31. Winter, R. B., Berg, O. G., and von Hippel, P. H. (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The *Escherichia coli* lac repressor−operator interaction: Kinetic measurements and conclusions, *Biochemistry 20*, 6961−77.
32. Barkley, M. D. (1981) Salt dependence of the kinetics of the lac repressor-operator interaction: Role of nonoperator deoxyribonucleic acid in the association reaction, *Biochemistry 20*, 3833−42.
33. Barry, J. K., and Matthews, K. S. (1999) Substitutions at histidine 74 and aspartate 278 alter ligand binding and allostery in lactose repressor protein, *Biochemistry 38*, 3579−90.
34. Barry, J. K., and Matthews, K. S. (1999) Thermodynamic analysis of unfolding and dissociation in lactose repressor protein, *Biochemistry 38*, 6520−8.
35. Wong, I., Chao, K. L., Bujalowski, W., and Lohman, T. M. (1992) DNA-induced dimerization of the *Escherichia coli* rep helicase. Allosteric effects of single-stranded and duplex DNA, *J. Biol. Chem. 267*, 7596−610.
36. Johnson, M. L., and Frasier, S. G. (1985) Non-linear least-squares analysis, *Methods Enzymol. 117*, 301−42.
37. Bevington, P. R., and Robinson, D. K. (1992) *Data Reduction and Error Analysis for the Physical Sciences*, 2nd ed., WCB McGraw-Hill, Boston.
38. Richmond, T. J. (1984) Solvent Accessible Surface-Area and Excluded Volume in Proteins: Analytical Equations for Overlapping Spheres and Implications for the Hydrophobic Effect, *J. Mol. Biol. 178*, 63−89.
39. 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*, 4237−44.
40. Bell, C. E., and Lewis, M. (2000) A closer view of the conformation of the Lac repressor bound to operator, *Nat. Struct. Biol. 7*, 209−14.
41. 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*, 235−42.
42. Creamer, T. P., Srinivasan, R., and Rose, G. D. (1997) Modeling unfolded states of proteins and peptides. 2. Backbone solvent accessibility, *Biochemistry 36*, 2832−5.
43. 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*, 50676−82.
44. Hong, J., Capp, M. W., Anderson, C. F., and Record, M. T. (2003) Preferential interactions in aqueous solutions of urea and KCl, *Biophys. Chem. 105*, 517−32.
45. 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*, 9276−88.
46. Santoro, M. M., Liu, Y. F., Khan, S. M. A., Hou, L. X., and Bolen, D. W. (1992) Increased Thermal-Stability of Proteins in the Presence of Naturally-Occurring Osmolytes, *Biochemistry 31*, 5278−83.
47. Greene, R. F., and Pace, C. N. (1974) Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, α-chymotrypsin, and *β*-lactoglobulin, *J. Biol. Chem. 249*, 5388−93.
48. Felitsky, D. J., and Record, M. T., Jr. (2003) Thermal and urea-induced unfolding of the marginally stable lac repressor DNA-binding domain: A model system for analysis of solute effects on protein processes, *Biochemistry 42*, 2202−17.
49. Archer, D. G. (1999) Thermodynamic properties of the KCl+H_2O system, *J. Phys. Chem. Ref. Data 28*, 1−17.
50. Auton, M., and Bolen, D. W. (2004) Additive transfer free energies of the peptide backbone unit that are independent of the model

compound and the choice of concentration scale, *Biochemistry 43*, 1329−42.
51. Rafflenbeul, L., Pang, W. M., Schonert, H., and Haberle, K. (1973) Thermodynamics of hydrophobic interaction in the systems water + glycine + urea and water + alanine + urea at 25 °C, *Z. Naturforsch. C28*, 533−54.
52. Schonert, H., and Stroth, L. (1981) Thermodynamic Interaction between Urea and the Peptide Group in Aqueous Solutions at 25 °C, *Biopolymers 20*, 817−31.
53. Naik, M. T., and Huang, T. H. (2004) Conformational stability and thermodynamic characterization of the lipoic acid bearing domain of human mitochondrial branched chain α-ketoacid dehydrogenase, *Protein Sci. 13*, 2483−92.
54. Khurana, R., Hate, A. T., Nath, U., and Udgaonkar, J. B. (1995) pH dependence of the stability of barstar to chemical and thermal denaturation, *Protein Sci. 4*, 1133−44.
55. Wallace, L. A., Sluis-Cremer, N., and Dirr, H. W. (1998) Equilibrium and kinetic unfolding properties of dimeric human glutathione transferase A1-1, *Biochemistry 37*, 5320−8.
56. Schellman, J. A. (1994) The Thermodynamics of Solvent Exchange, *Biopolymers 34*, 1015−26.
57. 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.
58. Wyman, J., Jr. (1964) Linked Functions and Reciprocal Effects in Hemoglobin: A Second Look, *Adv. Protein Chem. 19*, 223−86.
59. Felitsky, D. J., Cannon, J. G., Capp, M. W., Hong, J., Van Wynsberghe, A. W., Anderson, C. F., and Record, M. T., Jr. (2004) The exclusion of glycine betaine from anionic biopolymer surface: Why glycine betaine is an effective osmoprotectant but also a compatible solute, *Biochemistry 43*, 14732−43.
60. Schiefner, A., Breed, J., Bosser, L., Kneip, S., Gade, J., Holtmann, G., Diederichs, K., Welte, W., and Bremer, E. (2004) Cation-π 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*, 5588−96.
61. 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*, 137−44.
62. Spolar, R. S., Livingstone, J. R., and Record, M. T. (1992) Use of Liquid-Hydrocarbon and Amide Transfer Data to Estimate Contributions to Thermodynamic Functions of Protein Folding from the Removal of Nonpolar and Polar Surface from Water, *Biochemistry 31*, 3947−55.
63. Kozlov, A. G., and Lohman, T. M. (2000) Large contributions of coupled protonation equilibria to the observed enthalpy and heat capacity changes for ssDNA binding to *Escherichia coli* SSB protein, *Proteins Suppl. 4*, 8−22.
64. Eftink, M. R., Anusiem, A. C., and Biltonen, R. L. (1983) Enthalpy−entropy compensation and heat capacity changes for protein−ligand interactions: General thermodynamic models and data for the binding of nucleotides to ribonuclease A, *Biochemistry 22*, 3884−96.
65. Bergqvist, S., Williams, M. A., O'Brien, R., and Ladbury, J. E. (2004) Heat capacity effects of water molecules and ions at a protein-DNA interface, *J. Mol. Biol. 336*, 829−42.
66. Holbrook, J. A., Tsodikov, O. V., Saecker, R. M., and Record, M. T. (2001) Specific and non-specific interactions of integration host factor with DNA: Thermodynamic evidence for disruption of multiple IHF surface salt-bridges coupled to DNA binding, *J. Mol. Biol. 310*, 379−401.
67. Schnarr, M., and Maurizot, J. C. (1981) Unfolding of lac repressor and its proteolytic fragment by urea: Headpieces stabilize the core within lac repressor, *Biochemistry 20*, 6164−9.
68. Kontur, W. S., Saecker, R. M., Davis, C. A., Capp, M. W., and Record, M. T., Jr. (2005) Solute probes of conformational changes in open complex formation by *E. coli* RNA polymerase at the λP_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, submitted to *Biochemistry*.
69. Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A. N., Niu, W., Ebright, Y. W., Levy, R., and Ebright, R. H. (2002) Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex, *Cell 108*, 599−614.
70. Anderson, C. F., and Record, M. T. (1993) Salt Dependence of Oligoion Polyion Binding: A Thermodynamic Description Based on Preferential Interaction Coefficients, *J. Phys. Chem. 97*, 7116−26.
71. 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*, 1954−8.
72. Anderson, C. F., and Record, M. T., Jr. (2004) Gibbs−Duhem-based relationships among derivatives expressing the concentration dependences of selected chemical potentials for a multicomponent system, *Biophys. Chem. 112*, 165−75.
73. Record, M. T., Zhang, W. T., 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.
74. Courtenay, E. S., Capp, M. W., Anderson, C. F., and Record, M. T. (2000) Vapor pressure osmometry studies of osmolyte-protein interactions: Implications for the action of osmoprotectants in vivo and for the interpretation of "osmotic stress" experiments in vitro, *Biochemistry 39*, 4455−71.
75. Record, M. T., and Anderson, C. F. (1995) Interpretation of Preferential Interaction Coefficients of Nonelectrolytes and of Electrolyte Ions in Terms of a Two-Domain Model, *Biophys. J. 68*, 786−94.
76. Kraulis, P. J. (1991) MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures, *J. Appl. Crystallogr. 24*, 946−50.
77. Merritt, E. A., and Bacon, D. J. (1997) Raster3D: Photorealistic Molecular Graphics, *Methods Enzymol. 277*, 505−24.

BI0515218