BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

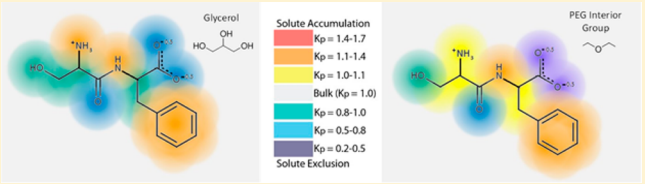
Chemical Interactions of Polyethylene Glycols (PEGs) and Glycerol with Protein Functional Groups: Applications to Effects of PEG and Glycerol on Protein Processes

D. B. Knowles,[†] Irina A. Shkel,[†,‡] Noel M. Phan,[†] Matt Sternke,[†] Emily Lingeman,[†] Xian Cheng,[§] Lixue Cheng,[†,‡] Kevin O'Connor,[†] and M. Thomas Record*[,†,‡,§]

[†]Department of Biochemistry, [‡]Department of Chemistry, and [§]Program in Biophysics, University of Wisconsin—Madison, 433 Babcock Drive, Madison, Wisconsin 53706, United States

Supporting Information

ABSTRACT: In this work, we obtain the data needed to predict chemical interactions of polyethylene glycols (PEGs) and glycerol with proteins and related organic compounds and thereby interpret or predict chemical effects of PEGs on protein processes. To accomplish this, we determine interactions of glycerol and tetraEG with >30 model compounds displaying the major C, N, and O functional groups of proteins. Analysis of these data yields coefficients (α values) that quantify interactions of glycerol, tetraEG, and PEG end ($-CH_2OH$) and interior ($-CH_2OCH_2-$) groups with these groups, relative to interactions with water. TetraEG (strongly) and glycerol (weakly) interact favorably with aromatic C, amide N, and cationic N, but unfavorably with amide O, carboxylate O, and salt ions. Strongly unfavorable O and salt anion interactions help make both small and large PEGs effective protein precipitants. Interactions of tetraEG and PEG interior groups with aliphatic C are quite favorable, while interactions of glycerol and PEG end groups with aliphatic C are not. Hence, tetraEG and PEG300 favor unfolding of the DNA-binding domain of lac repressor (lacDBD), while glycerol and di- and monoethylene glycol are stabilizers. Favorable interactions with aromatic and aliphatic C explain why PEG400 greatly increases the solubility of aromatic hydrocarbons and steroids. PEG400−steroid interactions are unusually favorable, presumably because of simultaneous interactions of multiple PEG interior groups with the fused ring system of the steroid. Using α values reported here, chemical contributions to PEG m-values can be predicted or interpreted in terms of changes in water-accessible surface area (ΔASA) and separated from excluded volume effects.



Polyethylene glycols (PEGs) are widely used in biochemistry, structural biology, and medicine, as well as in the pharmaceutical and chemical industries. Applications include protein crystallization[1] and stabilization, protein separation,[2] and boosting osmotic pressure,[3,4] as well as use as an excipient in medications,[5] an anticancer agent,[6] and a stabilizer of organs for transplants.[7,8] Other recent and significant applications include pressure control in biomedical devices,[9] drug delivery,[10,11] and use as an intelligent biomaterial for treating injures in membranes[12] and an antifouling coating in biomedical devices.[13] Properties of PEGs relevant for these applications include their high solubility or complete miscibility with water, their chemical interactions with molecular and macroscopic biosurfaces, and excluded volume effects. Here, by quantifying the interactions of glycerol and tetraEG with model compounds displaying functional groups of proteins, we deduce the chemical interactions of PEG with protein and other functional groups, relative to interactions with water. From this information, we predict and/or interpret the chemical component of PEG−protein interactions and PEG m-values for protein processes in terms of structural information and thereby separate chemical from excluded volume effects. We also address the question of how the different interactions of PEG with different types of molecular surfaces make both small and large PEGs effective protein precipitants while solubilizing drugs and other aromatic compounds.

Glycerol and PEG are routinely used to perturb biopolymer processes. Glycerol modestly stabilizes globular proteins,[14−17] β-hairpins,[18,19] protein assemblies,[20−23] and protein−DNA complexes[24] but modestly destabilizes nucleic acid duplexes[25−27] and can stabilize or destabilize RNA tertiary structure.[27] Tetraethylene glycol (tetraEG) is a weak destabilizer of nucleic acid duplexes,[26,28] and some PEGs destabilize proteins.[29−32] Oligo ethylene glycols and polymeric PEGs are common protein crystallization agents,[1] stabilize protein−DNA complexes,[24,33,34] and perturb their kinetics.[24,35] Effects of PEGs on the stability of these assemblies have been variously interpreted as osmotic effects (effects primarily on water activity), excluded volume effects, and combinations of preferential (chemical) interactions (including osmotic effects) and excluded volume effects.

Previously, we quantified the chemical noncovalent interactions of urea,[36,37] glycine betaine (GB), proline,[37,38] and the

Received: March 6, 2015
Revised: May 5, 2015
Published: May 11, 2015

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

Determination of chemical potential derivative $\mu_{23} = \left(\frac{\partial \mu_2}{\partial m_3}\right)_{P,T}$		
Determination of μ_{23} by Osmolality Difference	$\Delta\mathrm{Osm} = \mathrm{Osm}(\mathrm{m}_2, \mathrm{m}_3) - \mathrm{Osm}(\mathrm{m}_2, 0) - \mathrm{Osm}(0, \mathrm{m}_3) = \frac{\mu_{23}}{RT} m_2 m_3$	Eq.1
Determination of μ_{23} by Solubility	$\text{solubility } m\text{-}value \equiv -RT\left(\frac{d\ln m_2^{ss}}{dm_3}\right) = -RT\left(\frac{\partial \ln m_2}{\partial m_3}\right)_{\mu_2} \cong \mu_{23}$	Eq.2
Analysis and interpretation of μ_{23}		
Group ASA-based Interpretation of μ_{23}	$\mu_{23} = \left(\frac{\partial \mu_2}{\partial m_3}\right)_{P,T} = \sum_i \alpha_{solute,i} ASA_i^{model\ compound} + \sum_k \beta_{solute,k} \nu_k^{ion}$	Eq.3
SPM Interpretation of $\alpha_{solute,i}$	$\alpha_{solute,i} = -RT\left(K_{p,i} - 1\right) b_{1,i}(1 + \epsilon_3)/m_1^\bullet$ where $K_{p,i} = m_{3i,loc}/m_{3i,bulk}$ and $\epsilon_i = \frac{d\ln\gamma_i}{d\ln m_i}$	Eq.4
m-Value, Relation to μ_{23}	$\mathrm{m-value} \equiv \frac{\partial \Delta G_{obs}}{\partial m_3} = -RT\frac{\partial \ln K_{obs}}{\partial m_3} = \Delta\mu_{23}$	Eq.5
Prediction of μ_{23} and m-values		
Predicting *m*-value	$m\text{-}value = \Delta\mu_{23} = \sum_i \alpha_{solute,i} \Delta ASA_i^{biopolymer} + \sum_k \beta_{solute,k} \Delta\nu_k^{ion}$	Eq.6
Predicting PEG μ_{23} values	$\mu_{23} = \sum_i [\alpha_{2E,i} + (N_3 - 1)\alpha_{I,i}] ASA_i + \sum_k [\beta_{2E,k} + (N_3 - 1)\beta_{I,k}] \nu_k^{ion}$	Eq.7
Predicting PEG *m*-values	$m\text{-}value = \sum_i [\alpha_{2E,i} + (N_3 - 1)\alpha_{I,i}] \Delta ASA_i^{biopolymer} + \sum_k [\beta_{2E,k} + (N_3 - 1)\beta_{I,k}] \Delta\nu_k^{ion}$	Eq.8

Figure 1. Quantifying interactions of solutes (glycerol, tetraEG, and PEG) with model compounds and effects on processes (chemical potential derivatives μ_{23}).

spectrum of protein-stabilizing and -destabilizing Hofmeister salts[39,40] with sets of model compounds selected to display various combinations of functional groups of proteins and (for urea) nucleic acids. These chemical interactions are preferential interactions, i.e., interactions relative to interactions with water; an unfavorable interaction between solute and model compound means that both species prefer to interact with water rather than with each other. Analysis of these data reveals that solute−model compound interactions can be dissected into independent, additive interactions of the solute with functional groups of the model compound. Dissection by functional groups provides a more chemical interpretation of solute interaction data than the backbone−side chain interpretation of amino acid and dipeptide transfer free energies,[17,41] involves many fewer parameters (7 vs 20 for applications to protein processes at pH 7), and eliminates ambiguities in interpretation or prediction when only part of a side chain or backbone unit is buried or exposed in a process. The two approaches have been compared recently.[37]

Large areas of biopolymer surface are buried in interfaces and coupled conformational changes in the steps of biopolymer self-assembly and function. Chemical effects of solutes and Hofmeister salts on biopolymer processes (as distinct from excluded volume effects of large solutes and/or Coulombic effects of salts) are well interpreted or predicted to be additive interactions of the solute or salt ions with the functional groups on the biopolymer that are buried (or exposed) in the process. For example, proline and GB stabilize proteins against unfolding because of the unfavorable chemical interactions of these solutes with the aliphatic hydrocarbon surface and amide oxygens buried during folding, while urea destabilizes in large part because of its favorable interactions with these groups.[36−38] Many of the chemical interactions between two solutes or between the solute and biopolymer in water, including hydrogen bonding and hydrophobic interactions, are the same as those of biopolymer self-assembly. Quantitative information about solute−solute interactions is needed to develop the use of solutes as probes of changes in biopolymer surface area in the key intermediates and transition states in the mechanisms of self-assembly and function, as well as to quantify the contributions of noncovalent interactions to biopolymer stability.

Here we determine the interactions of tetraEG and glycerol with a large set of model compounds chosen to display various

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

combinations of functional groups of proteins and interpret these measurements in terms of chemical interactions (α values) of these solutes and of the terminal -CH_2OH and interior -CH_2OCH_2- groups of PEG with chemically distinct types of protein functional groups [designated aliphatic (sp^3) C, aromatic (sp^2) C, amide N, cationic N, amide O, carboxylate O, carboxylic acid O, and hydroxyl O] and with four inorganic ions (Na^+, K^+, Cl^-, and HPO_4^{2-}). In principle, an excluded volume (EV) contribution to these interactions of two small solutes may also be present. Because the solutes are similar in size and are probably also similar in size to the average cluster size of water molecules, we expect that EV interactions are not significant. SPT calculations including water find that the EV contribution to small solute effects on protein folding is negligible when the solute is the same size as water.[42] Both chemical and EV effects are generally significant for interactions of two polymeric solutes. Quantitative analyses of polymeric PEG−protein interactions[43] and effects of polymeric PEG on DNA helix formation[28] (see the discussion below) have recently been presented.

We apply α value analysis to predict and interpret the chemical interactions of glycerol, tetraEG, and other PEG oligomers with proteins and effects of these solutes on protein solubility, folding, and subunit assembly, including new results quantifying effects of the series from diEG to PEG300 and the polyols EG and glycerol on folding of the lac repressor DNA-binding domain (lacDBD). Our analysis explains why polyols like EG and glycerol are protein stabilizers while tetraEG and larger PEG oligomers are destabilizers, and (in ref 43) why both small and large PEGs can be effective protein precipitants. For these larger PEGs (flexible coils), both excluded volume and shielding of chemical interactions must be considered. Good agreement with literature data quantifying favorable interactions of PEG400 with aromatic compounds and unfavorable interactions of small and large PEGs with native proteins[43] is obtained.

MATERIALS AND METHODS

Details about measurements of μ_{23} by vapor-pressure osmometry and solubility and of m-values of lacDBD unfolding are provided in the Supporting Information. Calculations of the ASA of model compounds (Table S1 of the Supporting Information), PEG oligomers (Table S3 of the Supporting Information), and aromatic compounds (Table S7 of the Supporting Information) and the change in the ASA of proteins processes (Table S5 of the Supporting Information) are performed as described in the Supporting Information. Analyses of experimental data, including uncertainty estimations, with equations listed in Figure 1 were performed as described in the Supporting Information.

Background and Proposed Analysis of PEG Interactions. *Chemical Potential Derivatives μ_{23} Quantifying Chemical (preferential) Interactions.* Our strategy, as in previous studies, is to determine chemical interactions of tetraEG and glycerol (each designated component 3) with large sets of model compounds (designated component 2) displaying the most common functional groups of proteins. These interactions are quantified by chemical potential derivatives $(\partial\mu_2/\partial m_3)_{m_2} = \mu_{23}$, model-independent thermodynamic quantities that characterize solution nonideality arising from interactions of two solutes. When it is independent of solute concentration between 0 and 1 molal, μ_{23} is numerically equal to the transfer free energy for the process of transferring the model compound from a dilute solution in water to a solution with the same low concentration of model compound that is 1 molal in solute.[17,37,41] Values of μ_{23} can be interpreted using Kirkwood−Buff (KB) statistical mechanics[44] or molecular thermodynamic analysis of the solute partitioning model (SPM)[36−38,45] to obtain the local concentration of one solute in the vicinity of another, and by simulations.[46]

A negative μ_{23} indicates a favorable solute−model compound interaction relative to interactions with water. A positive μ_{23} indicates an unfavorable interaction, i.e., that components 2 and 3 prefer to interact with water rather than with one another. Higher-molecular weight PEGs as well as other polymers (e.g., dextran and Ficoll) also exert physical (excluded volume) effects[43,47−52] (see refs 53 and 54 for reviews). One goal of determining chemical interactions of tetraEG and glycerol with protein groups is to predict interactions of PEG end and interior residues with these groups and thereby predict chemical contributions to any PEG−protein interaction. In this way, chemical and excluded volume contributions to effects of large PEGs on protein processes can be dissected.[28,43]

Measurement of μ_{23} for Interactions of Glycerol and TetraEG with Soluble Model Compounds by Vapor-Pressure Osmometry. To quantify preferential interactions of tetraEG and glycerol with protein functional groups, vapor-pressure osmometry (VPO) is used to measure the excess osmolality ΔOsm of each of a series of three-component solutions with varying concentrations of component 3 (either tetraEG or glycerol) and component 2 [one of a set of more than 30 small, nonvolatile model compounds displaying functional groups of proteins (or inorganic ions) to the solution]. Ethylene glycol and the smallest PEG oligomers (diEG and triEG) are too volatile to investigate by VPO. The excess osmolality of a three-component solution, ΔOsm, is defined in eq 1 (Figure 1) as the difference between the osmolality of the three-component solution Osm(m_2,m_3) and the sum of the osmolalities of the corresponding two-component solutions [Osm(m_2,0) + Osm(0,m_3)].

Excess osmolality (ΔOsm) data are analyzed to obtain values of μ_{23}, the chemical potential derivative quantifying solute−model compound interactions. Multicomponent solution theory[55] shows that analysis of ΔOsm as a function of m_2m_3, the product of the molal concentrations of the model compound (component 2) and solute (component 3), yields μ_{23} (see eq 1 in Figure 1). The m_2m_3 product quantifies the probability of an interaction between solutes 2 and 3, and μ_{23} quantifies the strength of this interaction.

Measurement of μ_{23} for Interactions of Glycerol and TetraEG with Low-Solubility Model Compounds by Solubility Assays. Chemical interactions (μ_{23}) of glycerol and tetraEG with sparingly soluble model compounds are quantified from the effect of these solutes on the solubility of the model compound. If the concentration of the model compound (component 2) in a saturated solution is low enough that self-interactions of the model compound are negligible (i.e., provided that the saturated solution is an ideal dilute solution), then the solubility m-value ($-RT$ d ln m_2^{ss}/dm_3, where m_2^{ss} is the molal concentration of the model compound in the saturated solution) is equal to μ_{23} (see eq 2 in Figure 1).

If self-interactions of the model compound in the saturated solution are non-negligible, another method such as osmometry (eq 1) should be used at lower concentrations to determine μ_{23}.

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

Quantifying Interactions of Solutes Like Glycerol and TetraEG with Protein Functional Groups and Salt Ions. Values of μ_{23} for the interactions of the small solutes investigated to date (urea, GB, proline, and the series of Hofmeister salts[36−38]) with model compounds and biopolymers are found to be well-described as a sum of independent short-range interactions of these solutes with the individual functional groups on the model compounds or biopolymers and the inorganic salt ion(s) of the electroneutral component (see eq 3 in Figure 1). In eq 3, chemical interaction coefficient $\alpha_{solute,i}$ (in calories per mole per molal per square angstrom) quantifies the interaction of the solute (here tetraEG or glycerol) with 1 Å^2 of protein functional group type i, relative to interactions with hydration water, and $\beta_{solute,i}$ (in calories per mole per molal) quantifies the interaction of these solutes with inorganic ions. In eq 3, ASA_i is the amount of water-accessible surface (in square angstroms) of functional group type i. To interpret a value of $\alpha_{solute,i}$, one must consider not only the interaction between the solute and this functional group but also the solute−water and functional group−water interactions that are disrupted when this solute−group interaction occurs.

The SPM[36−38,45] provides a molecular interpretation of preferential interactions of a small solute with a chemically homogeneous surface, like the macroscopic air−water interface or the molecular surface of an aromatic or aliphatic hydrocarbon. For interactions of a small solute with a molecular surface with a hydration b_1 (number of molecules of water per unit area of surface), the SPM predicts the observed proportionality of μ_{23} to surface area and interprets the proportionality constant (α_{solute}) in terms of the accumulation or exclusion of the solute in the vicinity of the surface, as quantified by a microscopic local bulk partition coefficient K_p. This relationship is given as eq 4 (Figure 1). In eq 4, $K_p = m_3^{loc}/m_3^{bulk}$, where m_3^{loc} and m_3^{bulk} are local and bulk solute molalities, respectively, and K_p is a microscopic analog of an equilibrium constant for the process of transferring the solute from bulk solution to the vicinity of the surface and as such is relatively independent of bulk solute concentration. Equation 4 has been extended to interpret interactions of a small solute with individual functional groups (i) on the surface of a model compound.[36−38] A negative (favorable) $\alpha_{solute,i}$ means that $K_{p,i} > 1$ and that the solute accumulates in the vicinity of a functional group of type i to a concentration that is higher than its bulk concentration (by factor $K_{p,i}$). Conversely, a positive (unfavorable) $\alpha_{solute,i}$ means that $K_{p,i} < 1$ and the solute is excluded from the water of hydration of the functional group, so the local solute concentration is less than its bulk concentration (by factor $K_{p,i}$). Also in eq 4, $\varepsilon_3 = (\partial \ln \gamma_3/\partial \ln m_3)_{m_2}$ is a solute−solute self-nonideality correction, typically small in magnitude and not very concentration-dependent below 1 molal, and $m_1^\bullet$ is the molality of pure water (55.5 mol/kg).

Predicting and Interpreting Effects of Solutes Like Glycerol and TetraEG on Biopolymer Processes. Experimentally, the effect of a solute on folding, precipitation, and other biopolymer processes is quantified by an m-value, the derivative of the observed standard free energy change for the process (ΔG°_{obs}) with respect to solute concentration (here expressed on the molal scale). Thermodynamically, the m-value is the difference in chemical potential derivatives μ_{23} determined by quantifying the preferential interaction of the solute with the final (product) and initial (reactant) species in the process[56] (see eq 5 in Figure 1). Using eq 3 and the hypothesis of additivity, the quantity $\Delta\mu_{23}$ in eq 5 is interpreted as the interaction of the solute with the ASA exposed or buried in the process (eq 6). In eq 6 (Figure 1), $\Delta ASA_i^{biopolymer}$ is the difference in the ASA of functional group i between product and reactant species in the process, and $\Delta\nu_k^{ion}$ is the corresponding difference in the number of free ions. For precipitation and crystallization, the $\Delta\nu_k^{ion}$ terms are expected to be significant. For biopolymer processes in solution, including folding, subunit assembly, and ligand binding, these salt ion terms generally are negligible and $\Delta\nu_k^{ion} = 0$. We test these predictions by comparisons with available data at or near 25 °C, including glycerol m-values for conversion of ConA dimers to tetramers[21] and glycerol and tetraEG m-values for unfolding lacDBD.

Predicting and Interpreting Effects of PEG Oligomers on Folding, Assembly, and Other Protein Processes. By extension of eq 3, we propose that the preferential interaction (μ_{23} value) of any PEG oligomer (degree of polymerization N_3) with any model compound or biopolymer can be dissected into interactions of the end groups and interior repeats with the functional groups on the biopolymer and with the ions of the biopolymer component. Assigning interaction potentials α_{2E} and α_I for the interactions of the two end groups (E, -CH_2OH) and the interior group (I, -CH_2OCH_2-) with 1 Å^2 of ASA of each type of functional group on the biopolymer and β_{2E} and β_I values for the interactions of end and interior groups of PEG with inorganic ions of the biopolymer component, we obtain eq 7 (Figure 1). Likewise, m-values that equal $\Delta\mu_{23}$, quantifying chemical effects of any PEG oligomer on these processes, are predicted or interpreted using eq 8 (Figure 1). For processes in solution where inorganic ions are not stoichiometric participants, $\Delta\nu_k^{ion} = 0$.

When α and β values for interactions of other PEG oligomers with protein functional groups are determined, analysis of these α and β values as a function of the number of interior -CH_2OCH_2- repeats will yield α and β values for interactions of end and interior repeat groups of PEG with protein groups, from which the chemical interactions of any PEG with any protein surface can be predicted. Such predictions are needed to separate chemical (soft) and excluded volume (hard) contributions to the interaction of any PEG with proteins, as well as PEG effects on protein solubility and other protein processes (see below). However, PEG oligomers shorter than tetraEG are found to be unsuitable for osmometric studies by VPO, and chemically homogeneous longer PEG oligomers are not readily available.

Here, as a practical alternative, we obtain the necessary interaction potentials α_{2E} (also β_{2E}) and α_I (also β_I) for interactions of end and interior groups of any PEG (from diEG to PEG20000 or larger) with protein groups (also salt ions) from analysis of glycerol and tetraEG α and β values for interactions with these groups, assuming additivity. TetraEG α and β values are composed of contributions from interactions with three interior -CH_2OCH_2- groups and two -CH_2OH end groups. Because of the chemical analogy between PEG end groups and glycerol, scaling of glycerol α and β values is used to obtain PEG end group α_{2E} and β_{2E} values. These are approximate because of the difference in composition (aliphatic C:hydroxyl O ASA ratios) of glycerol and PEG end groups. (EG would be a closer, though also approximate, model for PEG end groups but is found to be unsuitable for VPO studies.) As an empirical correction, we use the experimentally determined contribution of PEG end groups to the m-value

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

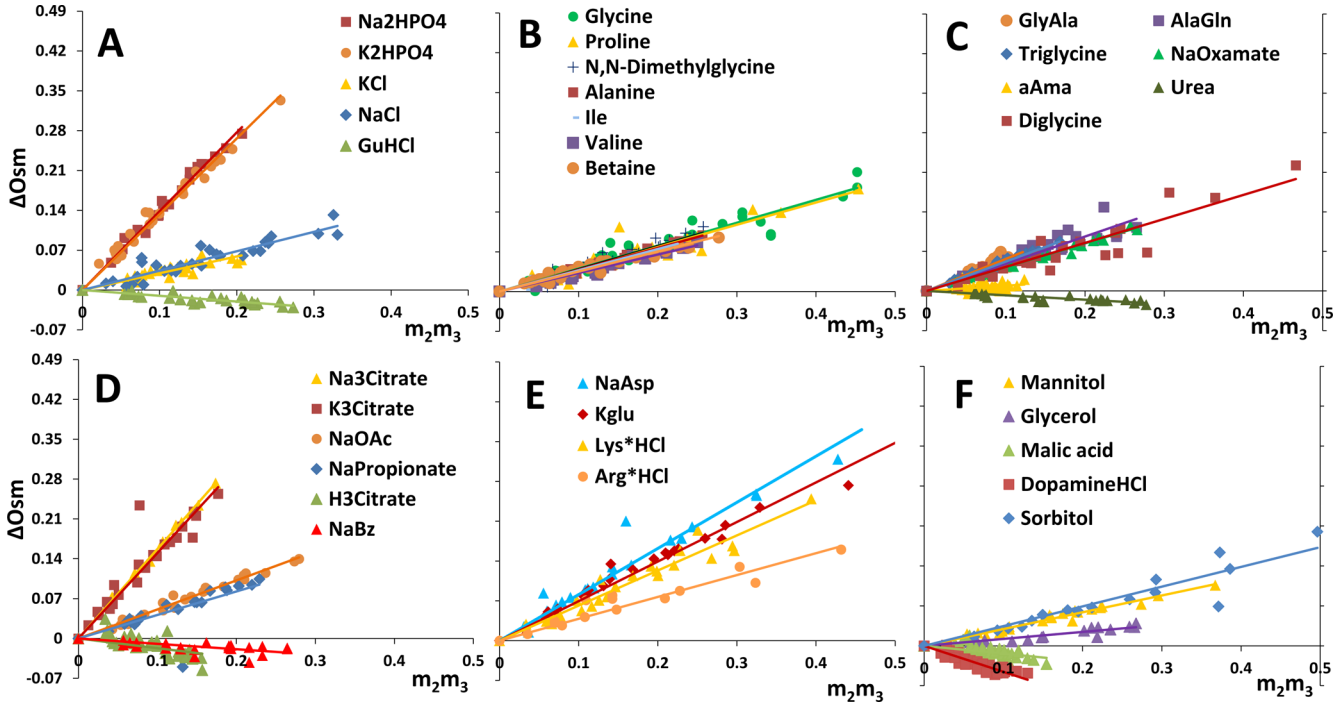


Figure 2. Interactions of tetraEG with salts and compounds displaying protein groups. Determination of μ_{23} quantifying tetraEG−model compound preferential interactions by vapor-pressure osmometry. Excess osmolality ΔOsm plotted vs m_2m_3, the product of the molal concentrations of the two solutes (tetraEG, component 3; model compound, component 2). From eq 1 (Figure 1), slopes are μ_{23}/RT: (A) inorganic salts, (B) amino acids, zwitterions, and derivatives, (C) amides, (D) carboxylate salts and carboxylic acids, (E) amino acid salts, and (F) other model compounds. Abbreviations for model compounds are defined in the footnote of Table S1 of the Supporting Information.

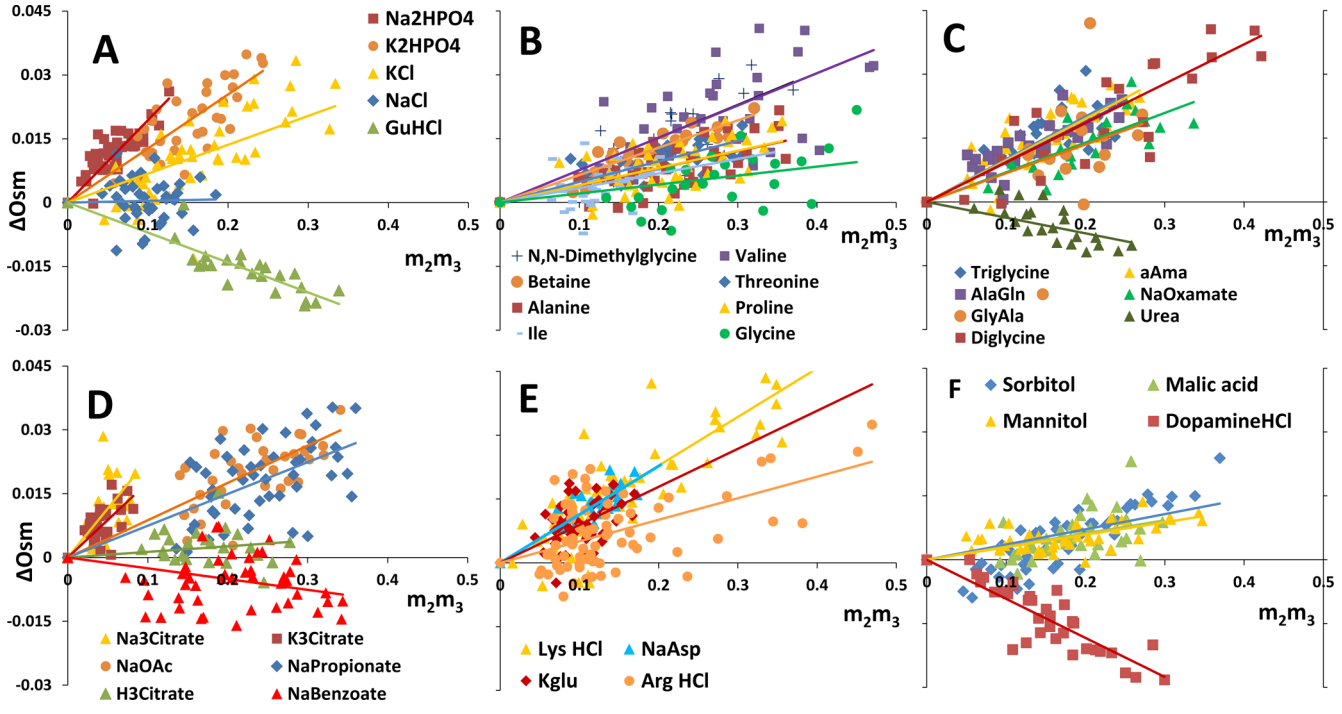


Figure 3. Interactions of glycerol with salts and compounds displaying protein groups. Determination of μ_{23} quantifying glycerol−model compound preferential interactions by vapor-pressure osmometry. Excess osmolality ΔOsm plotted vs m_2m_3, the product of the molal concentrations of the two solutes (glycerol, component 3; model compound, component 2). From eq 1 (Figure 1), slopes are μ_{23}/RT: (A) inorganic salts, (B) amino acids, zwitterions, and derivatives, (C) amides, (D) carboxylate salts and carboxylic acids, (E) amino acid salts, and (F) other model compounds. Abbreviations for model compounds are defined in the footnote of Table S1 of the Supporting Information.

data for unfolding of lacDBD (for which 65% of the ΔASA is aliphatic C) to set the α value for the interaction of the PEG end group with aliphatic C. We test these PEG α and β values by predicting the incremental effect of additional interior

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

Table 1. Interactions of TetraEG, Glycerol, and End and Interior Groups of PEG with Protein Groups and Salt Ions

(A) Group Interaction Potentials ($\alpha_{solute,i}$, cal mol^{-1} molal^{-1} Å^{-2})				
group (i)	$\alpha_{tetraEG,i}$	$\alpha_{glycerol,i}$	$\alpha_{2E,i}$	$\alpha_{I,i}$
aliphatic C	−0.349 ± 0.029	0.0548 ± 0.011	−0.00268 ± 0.020	−0.115 ± 0.012
aromatic C	−2.66 ± 0.028	−0.431 ± 0.011	−0.364 ± 0.0094	−0.765 ± 0.01
hydroxyl O	0.843 ± 0.025	0.0305 ± 0.0079	0.0258 ± 0.0067	0.272 ± 0.0085
amide O	2.94 ± 0.11	0.826 ± 0.059	0.698 ± 0.050	0.747 ± 0.041
carboxylate O	3.59 ± 0.06	0.467 ± 0.024	0.395 ± 0.021	1.07 ± 0.021
carboxylic acid O	−0.41 ± 0.039	0.0446 ± 0.0086	0.0377 ± 0.0072	−0.149 ± 0.013
amide N	−1.67 ± 0.056	−0.491 ± 0.034	−0.415 ± 0.029	−0.418 ± 0.021
cationic N	−0.805 ± 0.041	−0.245 ± 0.017	−0.207 ± 0.015	−0.199 ± 0.014
(B) Ion Interactions ($\beta_{solute,k}$, cal mol^{-1} molal^{-1} ion^{-1})				
ion (k)	$\beta_{tetraEG,k}$	$\beta_{glycerol,k}$	$\beta_{2E,k}$	$\beta_{I,k}$
HPO_4^{2-}	780 ± 16	93.9 ± 6.3	79.3 ± 5.3	234 ± 5.5
Cl^-	132 ± 6.8	22.8 ± 2.9	19.3 ± 2.5	37.6 ± 2.4
K^+	6.19 ± 10	−4.11 ± 3.1	−3.47 ± 2.6	3.22 ± 3.4
Na^+	18.5 ± 5.2	5.34 ± 3.2	4.51 ± 2.7	4.66 ± 2

residues in the PEG series from diEG to PEG300 on the lacDBD stability *m*-value, and by comparisons with literature data quantifying effects of PEG400 on the solubility of 28 aromatic compounds with functional groups that are the same as or similar to those investigated here.

RESULTS AND INTERPRETATIONS

Preferential Interactions of TetraEG with Model Compounds Displaying Protein Groups. VPO determinations of μ_{23} values quantifying preferential interactions of tetraEG with 34 nonvolatile, soluble model compounds, each displaying one or more protein functional groups and/or inorganic ions, are shown in Figure 2. Excess osmolalities ΔOsm (eq 1) are plotted as a function of the concentration product m_2m_3, where m_2 is the molal concentration of the model compound and m_3 is the molal concentration of tetraEG. These plots are linear within uncertainty over the concentration ranges studied here (<0.6 molal in each component); values of μ_{23} for each tetraEG−model compound preferential interaction, obtained from the slope, are reported in Table S1 of the Supporting Information.

For solutions of tetraEG and an amino acid, polyol, or carboxylate salt, Figure 2 shows that ΔOsm increases with an increasing concentration of each component (plotted as the m_2m_3 product). Hence, preferential interactions of tetraEG with these model compounds are unfavorable, with positive values of μ_{23}. TetraEG and these model compounds prefer to interact with water than with one another, causing the local concentration of tetraEG in the vicinity of these model compounds to be lower than its bulk concentration. By contrast, Figure 2 shows that interactions of tetraEG with all aromatic compounds and carboxylic acids studied, as well as with compounds with primarily amide N or cationic N functional groups (urea and GuHCl), are favorable. For these solutes, ΔOsm decreases with an increasing m_2m_3, indicating negative μ_{23} values and local accumulation of tetraEG in the vicinity of these model compounds.

Solubility data characterizing the favorable preferential interactions of tetraEG with two low-solubility model compounds (butylhydroxytoluene and phenylacetic acid) are shown in Figure S1 of the Supporting Information; μ_{23} values obtained from these data (eq 2) are negative (cf. Table S1 of the Supporting Information), indicating accumulation of tetraEG in the vicinity of these aromatic solutes.

Though the chemical compositions of tetraEG and glycine betaine (GB) are very different, signs and in some cases magnitudes of μ_{23} for interactions with the same model compound often agree. For example, values of μ_{23} for interactions of GB and tetraEG with amino acids, carboxylate salts, and polyols are positive (unfavorable) and similar in magnitude, while μ_{23} values for interactions of GB and tetraEG with urea and Na benzoate are both negative (favorable).

Preferential Interactions of Glycerol with Model Compounds Displaying Protein Groups. Preferential interactions of glycerol with 34 soluble model compounds (in most cases, the same as those investigated with tetraEG) are quantified by VPO; a solubility assay is used with an additional model compound. Figure 3 (osmometry) and Figure S1 of the Supporting Information (solubility) summarize the results of these experiments. As observed for tetraEG (Figure 2), values of ΔOsm of solutions of glycerol and an amino acid, polyol, or carboxylate salt increase with an increasing concentration of each solute (corresponding to unfavorable preferential interactions; $\mu_{23} > 0$) while values of ΔOsm of solutions of glycerol and aromatics or compounds with primarily amide N or cationic N functional groups decrease with an increasing concentration of each solute (corresponding to favorable preferential interactions; $\mu_{23} < 0$).

Results for glycerol−model compound interactions (Figure 3) in general parallel those obtained for tetraEG. For all compounds investigated except citric acid, values of μ_{23} have the same sign but are approximately one-third to one-tenth as large in magnitude for glycerol as for tetraEG. (The glycerol−citric acid interaction is unfavorable, while the tetraEG−citric acid interaction is favorable.) In many cases, values of μ_{23} for glycerol−model compound interactions are of opposite in sign but similar in magnitude to those reported previously for urea.[36] Exceptions are those of GuHCl and carboxylate salts (sodium propionate, trisodium and tripotassium citrate, potassium glutamate, and sodium benzoate), for which μ_{23} values for both glycerol and urea have the same sign.

Chemical Interaction Coefficients (α values) for Interactions of TetraEG and Glycerol with Functional Groups of Proteins. Model compounds investigated here were selected to display one or more of the eight (coarse-grained) principal functional groups of proteins (aliphatic and

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

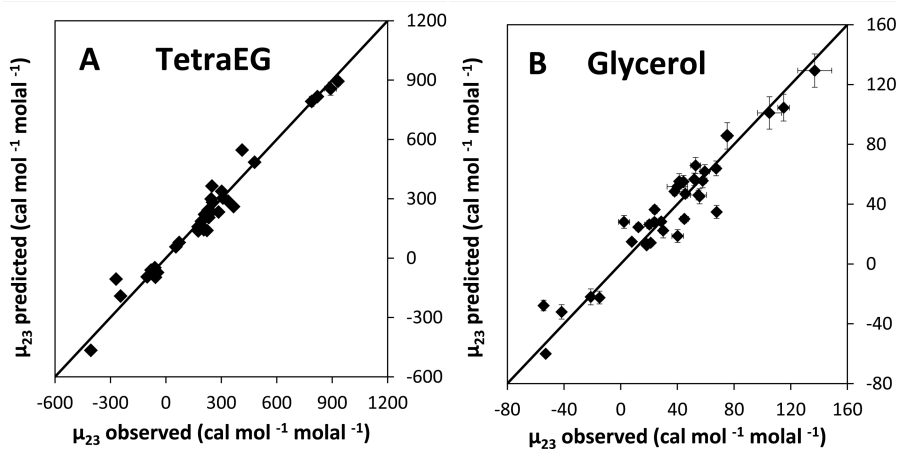


Figure 4. Comparison of predicted and observed preferential interactions μ_{23}. Predicted μ_{23} [calories per mole per molal (Table S1 of the Supporting Information)] for interactions of (A) tetraEG and (B) glycerol with model compounds plotted vs the observed (experimental) value. Predictions are generated using eq 3, ASA information from Table S2 of the Supporting Information, and interaction potentials α_i and β_k (Table 1) obtained from analysis of these model compound data sets. The line indicates the equality of predicted and observed μ_{23}. Error bars are one standard deviation of fitting uncertainty for observed values and determined as described in the Supporting Information for predicted values.

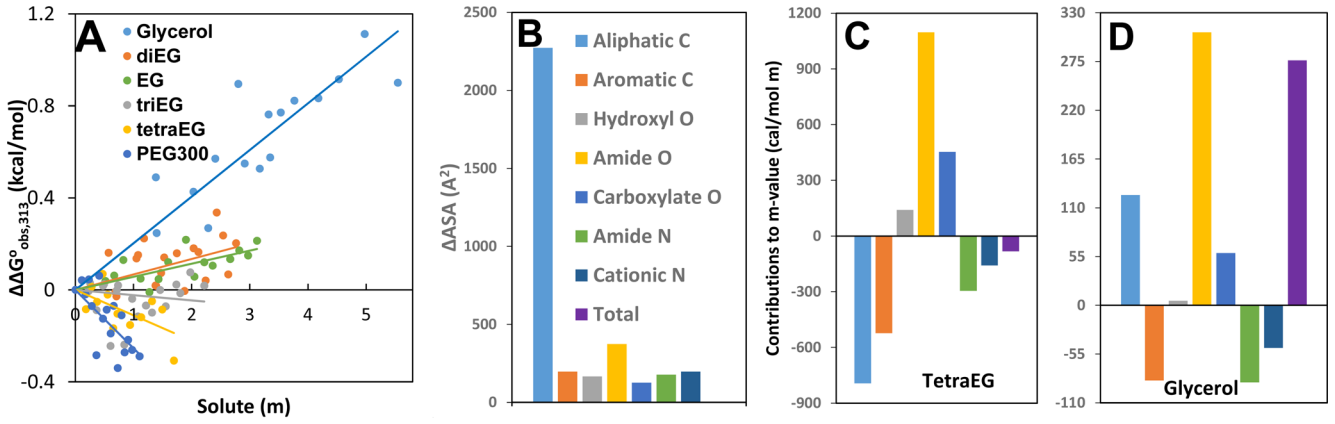


Figure 5. Effects of glycerol and small PEGs on unfolding lacDBD. (A) Standard free energy differences $\Delta\Delta G^{\circ}_{obs,313}$ of unfolding of the lac repressor DNA-binding domain (lacDBD) at 313 K, expressed relative to unfolding in the absence of the solute, plotted as a function of the molality of glycerol, EG, diEG, triEG, tetraEG, and PEG300. Slopes of lines are *m*-values [$\Delta\mu_{23}$ (eq 5 in Figure 1)] for these solutes. (B) Contribution from each surface type to the ΔASA of unfolding lacDBD to an extended polypeptide chain (ref 36). Predicted contribution from solute interactions with each surface type [from α values (Table 1)] to the (C) tetraEG and (D) glycerol *m*-values. Color coding is as in panel B. The sum of the bars is equal to the *m*-value. Purple bars represent the experimental tetraEG and glycerol *m*-values in panel A.

aromatic C; amide, hydroxyl, carboxylate, and carboxylic acid O; and amide and cationic N). All these functional groups are defined to include their covalently bonded proton or protons (one proton on carboxylic acid and hydroxyl O, and on secondary amide N; two protons on primary amide N; one, two, or three protons on cationic N of guanidinium and ammonium groups). This definition is consistent with the unified atom analysis of the surface areas of these model compounds.[57]

To interpret the interactions (μ_{23} values) of tetraEG and glycerol with these model compounds in terms of interactions with individual functional groups, numerical values of μ_{23} (Table S1 of the Supporting Information) and water-accessible surface areas (ASA) of the model compounds (Table S2 of the Supporting Information) are used as inputs to eq 3, and the resulting overdetermined linear systems of 36 (tetraEG) and 35 (glycerol) equations are solved to obtain eight chemical interaction coefficients (α values) quantifying the preferential interactions of tetraEG and glycerol with the unit surface area of these functional groups, as well as four β values quantifying interactions of tetraEG and glycerol with the inorganic ions present in the model compound set. This analysis assumes that chemical interactions of functional groups determine these small solute−small solute μ_{23} values and that excluded volume contributions are negligible. In these analyses, β values for all inorganic ions are allowed to float individually. β values for interactions of glycerol and tetraEG with the anions investigated (Cl^- and HPO_4^{2-}) are large and positive, with small relative uncertainties, while β values for the salt cations (Na^+ and K^+) are smaller in magnitude with larger relative uncertainties. Because β values for interactions of Na^+ with glycerol and of K^+ with tetraEG are approximately zero given these uncertainties, we also tested fits assigning these values to zero. These fits gave α values (and β values for the other ions) that are not significantly different from those listed in Table 1.

Consistent with comparisons of μ_{23} values, α values for interactions of tetraEG and glycerol with individual functional groups are found to differ greatly in magnitude but to have the same sign for most groups; important exceptions include those of aliphatic C and carboxylic acid O. Table 1 shows that, except

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

Table 2. *m*-Values Quantifying Effects of Polyols, Small PEGs, and Other Solutes on LacDBD Unfolding and Dissociation of the ConA Tetramer

	lacDBD unfolding *m*-value (cal mol^{-1} molal^{-1})		concanavalin A tetramer−dimer *m*-value (cal mol^{-1} molal^{-1})	
solute	experiment[a]	predicted[b]	experiment[c,d]	predicted[b]
EG	56.9 ± 11	–		
glycerol	203 ± 18	276 ± 34	170 ± 260	227 ± 37
diEG	67.0 ± 34	53.8 ± 54		239 ± 53
triEG	−21.0 ± 29	−14.4 ± 73		411 ± 72
tetraEG	−111 ± 41	−82.6 ± 98[e]		583 ± 88[f]
PEG300	−253 ± 64	−253 ± 170		1010 ± 180
urea	−440 ± 10[d,g]	−512 ± 84	−855 ± 69	−547 ± 120
proline		773 ± 210	598 ± 41	694 ± 310
GB	820 ± 110[d,h]	631 ± 470	2160 ± 76	724 ± 520
GuHCl	−747 ± 60[i]	−895 ± 170		

[a]Determined at 40 °C. [b]Predicted at 25 °C using α values from Table 1 for glycerol and for PEG end and interior groups, α values from ref 37 for urea, proline, and GB, and α values from ref 39 for GuHCl. In all cases, ASA information is from Table S5 of the Supporting Information; for lacDBD unfolding, an extended chain model of unfolded lacDBD is used.[36] [c]Determined at 20 °C.[21] [d]Experimental *m*-values are reported in calories per mole per molar. Values in calories per mole per molal will be somewhat smaller in magnitude (except for initial slopes at low solute concentration); differences do not exceed the uncertainties listed. [e]The prediction from 25 °C tetraEG α values in Table 1 is −82.2 ± 79 cal mol^{-1} molal^{-1}. [f]The prediction from 25 °C tetraEG α values in Table 1 is 581 ± 91 cal mol^{-1} molal^{-1}. [g]The urea *m*-value at 25 °C is −449 ± 11 cal mol^{-1} M^{-1}.[59] [h]The GB *m*-value at 25 °C is 780 ± 110 cal mol^{-1} M^{-1}.[58] [i]Experimental data and α values are from ref 39.

for hydroxyl O where the magnitude of the glycerol α value is very small, α values for tetraEG are 3−7 times larger than the corresponding glycerol α values. Many glycerol α values are comparable in magnitude (and in some important cases opposite in sign) to those of urea.[36] For tetraEG, many α values are comparable in magnitude and of the same sign as those for GB.[38]

To test additivity (eq 3) and the ability of these eight α values and four β values to describe the preferential interactions of tetraEG and glycerol with the 3-fold larger number of model compounds, Figure 4 compares experimental values of μ_{23} for tetraEG (panel A) and glycerol (panel B) to values of μ_{23} predicted from the interaction potentials of Table 1. The good agreement for the large majority of the model compounds studied, especially in the case of tetraEG, supports the assumption of additivity, the neglect of excluded volume, and the use of ASA as the weighting factor for the solute−group interaction potentials ($\alpha_{solute,i}$) in eq 3. The observed additivity and context independence of these solute−model compound preferential interactions indicate that interactions of water with solute and model compound functional groups are also well-approximated as additive and context-independent. Relative (percent) deviations between predicted and observed μ_{23} values for glycerol interactions are larger than for tetraEG because glycerol interactions are weaker and therefore more difficult to determine by VPO, as comparison of Figures 2 and 3 reveals. Uncertainties in μ_{23} values are estimated as described in the Supporting Information. Applications of these glycerol and tetraEG α values (Table 1) are provided below.

Determination of LacDBD Unfolding *m*-Values for Glycerol and Small PEGs (EG to PEG300). For comparison with α value predictions, equilibrium constants (K_{obs}) for two-state, reversible unfolding of marginally stable lacDBD, the operator-binding domain of lac repressor,[58,59] were determined as a function of the concentration of glycerol, ethylene glycol, and a series of PEG oligomers, including tetraEG, to concentrations of at least 3 monomolal. Values of $\Delta\Delta G^{\circ}_{obs} = -RT \ln(K_{obs}/K_{obs}^{\circ})$ for unfolding lacDBD at 40 °C are plotted versus solute molality in Figure 5A; slopes obtained from these plots are unfolding *m*-values, listed in Table 2. A temperature of 40 °C was chosen because it is within the transition region and does not require extrapolation of the thermodynamic data for unfolding. Glycerol is stabilizing, with a *m*-value approximately half as large in magnitude and opposite in sign to that of urea,[59] while ethylene glycol and diEG are marginally stabilizing, triEG is neutral within the uncertainty, and tetraEG and PEG300 are destabilizing.

Di-, tri-, and tetraEG and PEG300 form a series with the same end group ASA and composition (two -CH_2OH groups) and an increasing number of -CH_2OCH_2- interior repeats. Figure 7A reveals that PEG−lacDBD unfolding *m*-values decrease linearly with the number of -CH_2OCH_2- interior repeats. The intercept (122 ± 40 cal mol^{-1} molal^{-1}) and slope (−70.2 ± 16 cal mol^{-1} molal^{-1}) are the experimentally determined contributions of the two -CH_2OH end groups and each -CH_2OCH_2- PEG repeat to these lacDBD *m*-values.

The contribution of two PEG end groups to the lacDBD *m*-value obtained from this analysis (122 ± 40 cal mol^{-1} molal^{-1}) is twice as large as the EG *m*-value (56.9 ± 11 cal mol^{-1} molal^{-1}) even though the functional groups are the same and the ASA of EG (212 Å^2) is almost the same as that of two PEG end groups [214 Å^2 (Table 2 and Table S3 of the Supporting Information)]. What differs is the composition of the ASA: EG is 47.5% hydroxyl O, while PEG end groups are 49.7% hydroxyl O. Glycerol [55.6% hydroxyl O, an ASA (253 Å^2) only 20% larger than that of EG (212 Å^2)] exhibits a lacDBD *m*-value (203 ± 18 cal mol^{-1} molal^{-1}) that is almost 4 times as large as the EG *m*-value. From these results, we propose that the percent hydroxyl O ASA of a polyol determines whether it is stabilizing or destabilizing and, that, for a given % hydroxyl O, the total amount of hydroxyl O ASA determines the magnitude of the protein stabilizing effect. Tests of this proposal are in progress. These effects indicate a delicate balance between the favorable interaction of polyol aliphatic C and unfavorable interaction of polyol hydroxyl O with the largely (65%) aliphatic C surface exposed in unfolding lacDBD.

Chemical Interaction Coefficients (α values) for Interactions of PEG End (-CH_2OH) and Interior (-CH_2OCH_2-) Groups with Functional Groups of Proteins. Because no direct experimental route is available to determine

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

α and β values for interactions of end and interior repeat groups of PEG with protein groups, we propose to use α and β values for interactions of tetraEG and glycerol with protein functional groups and ions to obtain the corresponding α and β values for interior and end groups of PEG. α and β values for glycerol, scaled by the ratio of the ASA of two PEG end groups to that of glycerol (214/253 = 0.845), provide approximate α and β values for two PEG end groups. These are approximate because of the small but significant difference in the aliphatic C:hydroxyl O ASA ratio between glycerol and PEG end groups discussed above. Therefore, because the surface exposed in unfolding lacDBD is 65% aliphatic C, we correct the α value for interaction of two PEG end groups with aliphatic C to obtain the experiment-based (extrapolated) lacDBD m-value (122 cal mol^{-1} $molal^{-1}$). Given the dominance of interactions of the interior groups of PEG over end group interactions (see below), any inaccuracy introduced by this approximation is not expected to have a significant effect on predictions of chemical interactions of PEGs with protein surfaces. Resulting α and β values for interactions of two PEG end groups with protein functional groups ($\alpha_{2E,i}$) and inorganic ions ($\beta_{2E,i}$) are listed in Table 1.

To obtain α and β values for the interactions of interior -CH_2OCH_2- groups of tetraEG with functional groups of proteins ($\alpha_{I,i}$) and inorganic ions ($\beta_{I,i}$), we interpret tetraEG α values (Table 1) as additive contributions from end group and interior group interactions ($\alpha_{tetraEG,i} = \alpha_{2E,i} + 3\alpha_{I,i}$). TetraEG β values are analyzed similarly to yield $\beta_{I,i}$; values of $\alpha_{I,i}$ and $\beta_{I,i}$ are listed in Table 1. Tests of the predictive ability of these PEG end group and interior group α values are provided in the Discussion.

Table 1 reveals that α and β values quantifying chemical interactions of the interior -CH_2OCH_2- group of PEG with protein functional groups and inorganic ions are as large as or larger than those of two end -CH_2OH groups. Indeed, interactions of the interior group are at least as important as those of the two end groups even for diEG and become progressively more dominant with an increasing number of interior groups in larger PEGs. Although the two -CH_2OH end groups contribute more than half the ASA of tetraEG, they contribute less than a third of the interaction of tetraEG with most protein functional groups and inorganic ions. Interactions of the three interior -CH_2OCH_2- repeats make the dominant contribution to tetraEG α and β values for all groups and ions except cationic N.

DISCUSSION

Comparisons of Chemical Interactions (α values) of Protein Groups with Glycerol (also PEG end groups) and with TetraEG (also PEG interior groups). *Key Differences in Signs of Interactions with Protein Groups.* Because aliphatic C makes by far the largest contribution to the ASA of both folded and unfolded proteins and to the ΔASA of unfolding, interactions of glycerol, tetraEG, and PEG groups with aliphatic C are very important determinants of effects of these solutes on protein processes even though α values (Table 1) and K_p values (Table S4 of the Supporting Information) for interactions with aliphatic C are relatively small in magnitude. TetraEG and PEG interior residues interact moderately favorably with aliphatic C, while glycerol and PEG end groups exhibit small unfavorable interactions, relative to interactions with water. This difference is a major reason why glycerol is a protein stabilizer while tetraEG and larger PEGs are destabilizers.

At acidic pH (<3), carboxylates of proteins are protonated, and hence, the relevant interactions of solutes are with carboxylic acid oxygens (i.e., =O and -OH) instead of with carboxylate oxygens (both -$O^{0.5-}$) present at neutral or alkaline pH. Table 1 shows that tetraEG and PEG interior residues interact moderately favorably with carboxylic acid O, while glycerol and PEG end groups exhibit a small unfavorable interaction with carboxylic acid O. The favorable interaction of PEG interior residues with carboxylic acid O indicates that the carboxylic acid -OH group is a good hydrogen bond donor to the ether O of PEG, relative to interactions with water.

Similarities in Signs of Interactions with Most Protein Groups. α and β values (Table 1) reveal that preferential interactions of tetraEG and PEG interior -CH_2OCH_2- groups with amide O and anionic carboxylate O, as well as with dianionic phosphate, are strongly unfavorable. Interactions of glycerol and PEG -CH_2OH end groups with these functional groups are weakly unfavorable. Strongly unfavorable interactions of PEG -CH_2OCH_2- moieties with these oxygens are expected, as discussed above, because the -CH_2OCH_2- group lacks the hydrogen bond donor capability necessary to compete effectively with water to interact with these oxygens.

Preferential interactions of tetraEG, PEG end and interior residues, and glycerol with hydroxyl O are also unfavorable, though for glycerol this preferential interaction is very weak. The relatively small magnitude (compared to those of tetraEG and PEG interior residues) of glycerol and PEG end group α values for interactions with both hydrogen bond donor and acceptor groups means that these hydrogen bonding interactions are nearly equivalent in free energy to the corresponding hydrogen bonding interactions with water.

α values (Table 1) also reveal that interactions of glycerol and PEG end groups with amide and cationic N are weakly favorable while interactions of PEG -CH_2OCH_2- repeats and therefore of tetraEG with amide and cationic N are strongly favorable. Amide and cationic N groups are good hydrogen bond donors, allowing them to compete with water to interact with the ether oxygen of PEG interior groups.

Interactions of glycerol, tetraEG, and PEG end and interior groups with aromatic C are all very favorable. Studies of the interactions of alcohols, polyols, and small PEGs with aromatic hydrocarbons and nucleobases are in progress to determine the contributions of individual -HO, -O-, and hydrocarbon functional groups to the interaction of these solutes with aromatic C, and the differences in interactions with aromatic and aliphatic C.

Comparisons of Interactions of PEG and Glycerol with Other Osmolytes. Interactions of PEG interior groups and of tetraEG with protein functional groups, quantified by α values (Table 1), are a mixture of strong favorable and unfavorable interactions, resulting in predictions of PEG−protein μ_{23} values and of m-values for protein processes that are small differences between large offsetting contributions and as such are quite sensitive to the details of the composition of the protein surface with which PEG is interacting. This situation parallels what is observed for glycine betaine and proline,[37,38] and indeed, PEG interior group α- values in many cases are of the same sign as and in some cases similar in magnitude to those of the interactions of glycine betaine (GB) and proline with these groups. For interactions with amide and carboxylate O, this parallel has a straightforward explanation. PEG interior groups

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

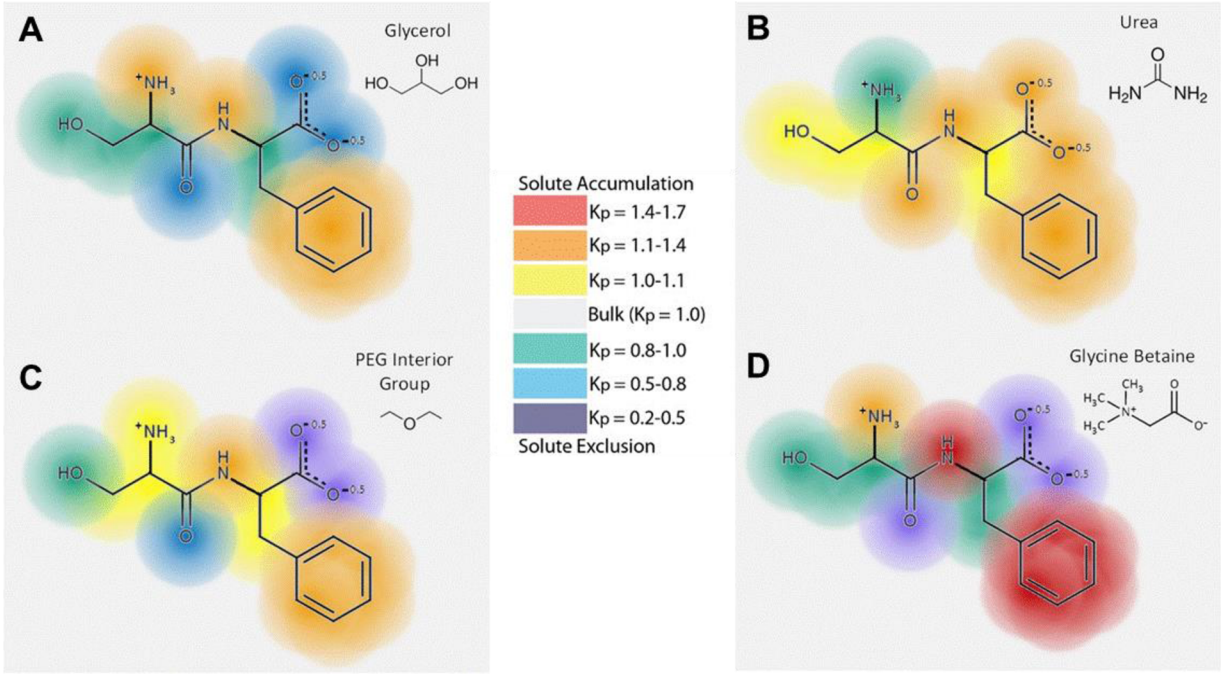


Figure 6. Predicted interactions of glycerol, PEG interior groups, urea, and glycine betaine with the Ser-Phe model dipeptide. Predicted extents of accumulation and exclusion of solutes from functional groups of the arbitrarily chosen model peptide Ser-Phe from K_p values (Table S4 of the Supporting Information) using a color "heat map".[37] Comparison of (A) glycerol and (C) the PEG interior group with each other and with (B) urea[36,37] and (D) glycine betaine.[37,38] Analogous predictions can be made for any other peptide or other compound with these functional groups using this set of K_p values (Table S4 of the Supporting Information).

and GB (also, but to a lesser extent, proline and tetraEG) lack the hydrogen bond donor capability necessary to compete with water to interact with amide and carboxylate O of proteins. Therefore, these solutes interact unfavorably with these groups and are strongly excluded from them. Figure 6C uses local bulk partitioning coefficient K_p (Table S4 of the Supporting Information), derived from α values using eq 4 (Figure 1), to predict the extent of accumulation or exclusion of the PEG interior -CH_2OCH_2- repeat at each functional group of a FS (Phe-Ser) dipeptide and compares these predictions with those for GB. These predictions can be made for any amino acid, peptide, or protein. The FS dipeptide is chosen for the purpose of illustration because it displays all but one of the major protein functional groups. Colors represent the gradation from strong accumulation [red, corresponding to $K_p > 1.4$ (see Table S4 of the Supporting Information)] to strong exclusion (violet, corresponding to $K_p < 0.5$).

Interactions of glycerol and PEG -CH_2OH end groups with the functional groups of proteins are also a mixture of favorable and unfavorable interactions, all of which are relatively weak by comparison with the analogous interactions of PEG interior groups or tetraEG. Comparison with results for other solutes investigated to date shows that interactions of glycerol and PEG end groups with functional groups of proteins are most similar in magnitude, though not in sign, to those of urea. Interactions of glycerol and urea with two of the most significant functional groups for protein unfolding (aliphatic C and amide O) are opposite in sign, making glycerol a protein stabilizer (unfavorable interactions with the functional groups exposed during unfolding) while urea is a destabilizer (favorable interactions). Glycerol and urea exhibit the same sign of interaction with aromatic C and amide N (both favorable), but opposite signs of interaction with hydroxyl and carboxylate O (favorable for urea, unfavorable for glycerol) and with cationic N (unfavorable for urea, favorable for glycerol). Figure 6A,B compares the extent of accumulation or exclusion of glycerol and urea at each functional group of the FS (phe-ser) dipeptide, showing the extent to which glycerol is an "anti-urea". This comparison is based on K_p values for glycerol (Table S4 of the Supporting Information) and urea.[36,37]

Predicting LacDBD Unfolding *m*-Values for Glycerol, TetraEG, and Other PEGs and Comparison with Experiment and with Results for Other Solutes. From α values quantifying interactions of glycerol, tetraEG, and PEG end and interior groups with the unit area of each major functional group of proteins (Table 1) and the amount and composition of the ΔASA of unfolding lacDBD to an extended chain (Figure 5B and Table S5 of the Supporting Information), unfolding *m*-values are readily predicted. Aliphatic C makes the dominant contribution (65%) to the ΔASA of unfolding, with amide O (11%) being the next largest contributor. Figure 5C,D shows the predicted contribution of each type of surface to the glycerol and tetraEG *m*-values. Table 2 compares predicted (25 °C) and observed (40 °C) *m*-values: for both glycerol and tetraEG, these agree within uncertainty. For tetraEG, analysis of the contributions of different surface types to the predicted *m*-value shows substantial compensation between moderately large stabilizing and destabilizing interactions, resulting in the small magnitude of the *m*-value. The uncertainty in the predicted *m*-value is relatively large because propagation of error does not exhibit this compensation. Any effect of the temperature difference between observed and predicted *m*-values is likely to be much smaller than this uncertainty. LacDBD unfolding *m*-values for urea[59] and GB[58] are not significantly dependent on temperature in this range (Table 2).

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

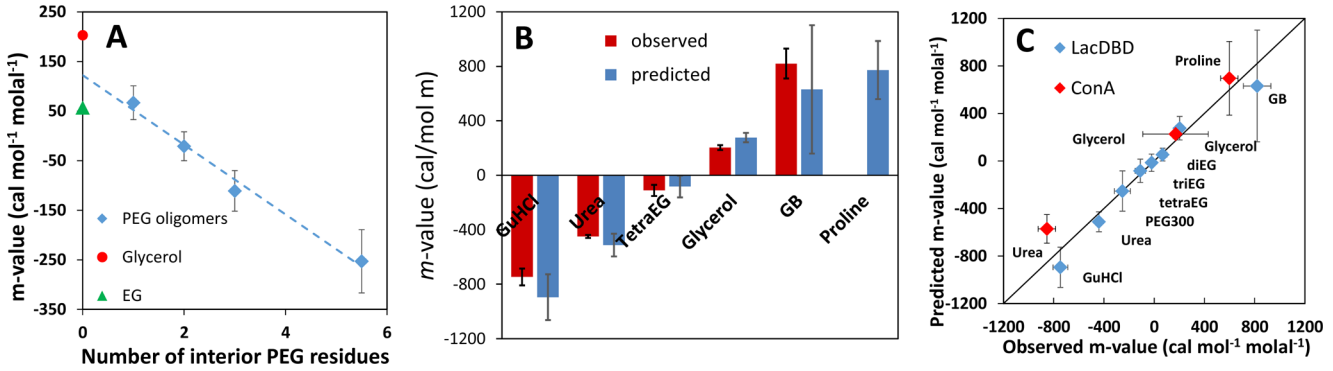


Figure 7. Effects of glycerol, small PEGs, and other solutes on lacDBD unfolding and comparison of predicted and observed *m*-values. (A) Experimental lacDBD unfolding *m*-values (Figure 5A) for diEG−PEG300 oligomers (diamonds) are plotted vs the number of interior -CH_2OCH_2- PEG units. Glycerol and EG *m*-values are shown for comparison; these were not used in obtaining the best-fit line. (B) Comparison of tetraEG and glycerol *m*-values for lacDBD unfolding with previously determined lacDBD unfolding *m*-values for destabilizers (GuHCl[39] and urea[59]) and a stabilizer (GB[58]). Predicted *m*-values (Table 2) for these solutes and for the stabilizer proline are also plotted. (C) Comparison of predicted and observed *m*-values for lacDBD unfolding (blue symbols) and dissociation of ConA tetramers into dimers (red symbols) (see Table 2). Predictions were made at 298 K, while observed lacDBD *m*-values were seen at 313 K and ConA *m*-values at 293 K.[21]

From the PEG end and interior group α values of Table 1 and the ΔASA of unfolding lacDBD, *m*-values for the chemical effect of any small PEG on unfolding of lacDBD are readily predicted. For larger PEG, shielding of chemical interactions (χ effect) and excluded volume effects also need to be considered.[43] The contribution of the PEG end group to these predictions is fixed as described in Results, using the intercept of the plot (Figure 7A) of the diEG−PEG300 *m*-values versus the number of interior residues. Here we examine whether the PEG interior residue α values in Table 1 can predict the slope of this plot (i.e., the increment in PEG *m*-value per interior residue). From Figure 7A, the experimental *m*-value increment is −70.2 ± 16 cal mol^{-1} $molal^{-1}$ at 40 °C, which agrees well with the prediction from the α values of Table 1 (−68.2 ± 31 cal mol^{-1} $molal^{-1}$ at 25 °C). Predictions of di-, tri-, and tetraEG and PEG300 *m*-values from these α values are also in agreement with the experimental values within uncertainty, as shown in Table 2. The EG *m*-value cannot be predicted at this time because α values for interactions of EG with all the different functional groups exposed in unfolding lacDBD are not yet available. For some predictions, the percent uncertainty is large because the propagation of the uncertainty does not exhibit the compensation between large opposing contributions that *m*-values do.

The effect of the temperature difference between observed (40 °C) and predicted (25 °C) lacDBD *m*-values cannot at present be assessed. For urea and glycine betaine, *m*-values are similar at 40 and 25 °C.[58,59] Figure 7B compares experimentally determined lacDBD unfolding *m*-values for glycerol and tetraEG with previously determined results for denaturants (GuHCl and urea) and osmolyte stabilizers (GB and proline). These *m*-values give the destabilization (negative) or stabilization (positive) $\Delta\Delta G^{\circ}_{obs}$ at a solute concentration of 1 molal. TetraEG and glycerol are less perturbing than these other solutes.

The destabilizing effect of tetraEG arises from favorable preferential interactions with the aliphatic C, aromatic C, and amide and cationic N surface exposed in unfolding, partially compensated by unfavorable interactions with amide and carboxylate O. As noted above, parallels exist with GB, which also has favorable interactions with aromatic C and amide and cationic N surface exposed in unfolding, partially compensated by unfavorable interactions with amide and carboxylate O. What makes GB a stabilizer while tetraEG is a destabilizer is the sign of the small-magnitude interaction with aliphatic C surface, which is favorable for tetraEG but unfavorable for GB. For tetraEG, aliphatic C makes a large (destabilizing) contribution despite a relatively small aliphatic C α value, because aliphatic C accounts for the majority of the ΔASA of unfolding (65%). In contrast, though aromatic C accounts for only ∼6% of the total ΔASA, it makes a large (destabilizing) contribution because preferential interactions of tetraEG with aromatic C are highly favorable. This parallels the situation for urea, where weak favorable interactions of urea with aliphatic C and strong favorable interactions of urea with aromatic C are deduced to make similar significant contributions to the *m*-value for unfolding of a typical globular protein.[36] The protein stabilizing effect of glycerol arises from its unfavorable preferential interactions with amide O and aliphatic C; favorable interactions with amide N and aromatic C only partially offset these stabilizing contributions.

Predicting Solute *m*-Values for Association of ConA Dimers into Tetramers and Comparison with Experiment. Effects of solutes and Hofmeister salts on the association of two concanavalin A dimers to form a tetramer have been determined at room temperature.[21,22] Glycerol modestly favors tetramerization [*m*-value = 170 ± 260 cal mol^{-1} $molal^{-1}$ (from ref 21) (Table 2)]. Structural analysis reveals a large dimer−tetramer interface [ΔASA = −4900 $Å^2$ (Table S5 of the Supporting Information)], differing somewhat in composition from that buried in protein folding, with smaller proportions of aliphatic C and amide O and N and larger proportions of both cationic N and anionic O. Using the ASA information in conjunction with the glycerol α values in Table 1, we predict a small stabilizing glycerol *m*-value (Table 2) that is completely consistent with the experimental value (see Figure 7C). *m*-Values were also determined for urea, proline, GB, and other solutes.[21,22] Predicted *m*-values for urea and proline (Table 2) are comparable to the experimental values. For all other cases, Figure 7C summarizes the generally good agreement of observed solute *m*-values for lacDBD unfolding and ConA tetramer−dimer dissociation with those predicted from solute α values and ASA information (Table 2).

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

Predicting Effects of PEG400 on the Solubility of Aromatic Compounds and Comparison with Experiment. Selection of Aromatics for Analysis. Rytting et al.[60] determined solubilities of more than 100 aromatic compounds at PEG400 volume fractions of 0, 0.25, 0.5, 0.75, and 1. Functional groups on 28 of these compounds (see Table S6 of the Supporting Information for names and structures) are sufficiently similar to the protein groups investigated here to warrant comparison of predicted and observed μ_{23} values. These include biphenyl and naphthalene as well as derivatives of benzene, naphthalene, biphenyl, and other aromatic or fused aromatic−aliphatic (steroid) ring systems. Nucleobases and other heterocyclic aromatics, as well as those with halogens and/or ether oxygens, are omitted from this analysis. In addition to aromatic (sp^2) C, these 28 model compounds also display aliphatic (sp^3) C, amide and carbonyl O, amide and amino N, and hydroxyl and carboxylic acid O groups.

Solubilities of these 28 uncharged aromatic compounds in the absence of PEG400 are in the range of 10^{-5} to 10^{-1} M. Panels A and B of Figure S3 of the Supporting Information show the dramatic increases in solubility with increasing PEG400 concentration, by factors of up to 10^4 in 75% PEG400. Qualitatively, these large solubility increases result from the very favorable interactions (large negative α values) of the interior (and end) groups of PEG400 with the aromatic C surface of these model compounds (as quantified by the α values in Table 1). Interactions with aliphatic C, amide (and amino) N, and carboxylic acid O are also predicted to be favorable and contribute to these solubility increases, while interactions with amide (and carbonyl) O are predicted to be unfavorable. Table 1 lists α values for the interactions of end and interior groups of any PEG with these groups in a nonaromatic context. The effect on interactions with PEG400 of placing these groups in an aromatic context can be assessed from analysis of these data. The two other groups present on these aromatic model compounds (carbonyl O and amino N) should have similar hydrogen bonding interactions to amide O and N. Indeed, interactions of urea with carbonyl O and amino N functional groups on the heterocyclic aromatic rings of nucleobases are found to be similar to interactions with amide O and N.[61] On the other hand, interactions of urea with nucleobase methyl groups are found to be more favorable than those with other aliphatic C.[61]

Plots of the experimental data (the logarithm of the molar solubility of the aromatic as a function of PEG400 volume fraction) for these 28 compounds are shown in Figure S3 of the Supporting Information, together with fits of these data to quadratic functions of PEG400 concentration. These data are converted to molal scale solubilities (m_2^{ss}) as a function of PEG400 molality using the equations listed in Table S8 of the Supporting Information, plotted as in Figure S4 of the Supporting Information, and fitted as either a quadratic or linear (first two or three points only) function of PEG400 molality. Values of μ_{23} obtained from the initial slopes are reported in Table S10 of the Supporting Information. The uncertainties in the initial slopes are relatively large, especially where curvature is pronounced, because of the small data sets and the lack of data at low PEG400 concentrations. In addition, for the higher-solubility compounds ($S > 0.1$ M), μ_{23} values estimated from these initial slopes have an additional uncertainty resulting from the neglect of self-interactions in eq 2.

The α values for the interaction of PEG400 with the aromatic C of these compounds are predicted by additivity to be a linear combination of α values from Table 1 for aromatic interactions of its two end groups and 7.7 interior groups ($\alpha_{PEG400,aroC} = \alpha_{2E,aroC} + 7.7\alpha_{I,aroC}$). α values for interactions of PEG400 with the unit area of the other functional groups and types of surface of these aromatic compounds are predicted analogously. Many of these functional groups are the same as those present on the model compound studied here [alkyl, amide, carboxylic acid, and hydroxyl groups (Table S1 of the Supporting Information)], though the context of these groups on an aromatic ring may in principle alter their interaction with PEG400. Other functional groups are different (carbonyl O of ketone or quinone groups, amino N).

Predicted μ_{23} values obtained from these PEG400 α values (Table S9 of the Supporting Information) and ASA information (Table S7 of the Supporting Information) are listed in Table S10 of the Supporting Information. For these predictions, carbonyl O, phenolic hydroxyl O, and amino N of the aromatic compounds are treated as amide O, hydroxyl O, and amide N, respectively. α values for interactions of urea with carbonyl O and amino N of nucleobases are similar though not identical to α values for interactions of urea with amide O and amide N of protein model compounds.[61] For the 16 nonsteroid compounds (derivatives of benzene, naphthalene, and biphenyl), agreement between predicted and observed μ_{23} values is generally within ±25% and is not significantly different for the 10 compounds with carbonyl O, phenolic hydroxyl O, and/or amino N than for the six compounds lacking these groups. These 16 comparisons largely validate the approach developed here for predicting interactions and *m*-values for any PEG from the data for glycerol− and tetraEG−model compound interactions developed here and also show that context (aromatic vs nonaromatic) in general is not of primary significance as a determinant of PEG−functional group α values.

For the 12 steroids with a fused ring system, experimental values of μ_{23} are systematically more favorable than predicted values by approximately −1.5 kcal mol^{-1} $molal^{-1}$, giving rise to the offset in the predicted versus observed comparison shown in Figure 8. This offset is most simply interpreted as a favorable concerted (context-dependent) interaction involving multiple residues of the PEG400 chain molecule and the fused ring system of the steroids.

CONCLUSIONS

We develop the ability to predict or interpret effects of glycerol and any low-molecular weight PEG on the protein processes and solubility of organic compounds in terms of the change in water-accessible surface area (ASA) of the protein or organic compound in these processes. Relatively small differences between interactions of a protein aliphatic carbon surface with -CH_2OH groups (unfavorable) and -CH_2OCH_2- groups (favorable), relative to interactions with water, explain the opposite effects of glycerol and tetraEG on protein stability. We demonstrate the ability to predict the concentration-dependent effects of any PEG oligomer on the solubility of any organic compound with proteinlike functional groups; this should be of practical significance. Current research is focused on expanding the groups of solutes and model compounds characterized to include interactions of nucleobases and sugars with glycerol and other polyols of different ratios of aliphatic C to hydroxyl O ASA and small PEGs, dissecting contributions of unfavorable

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

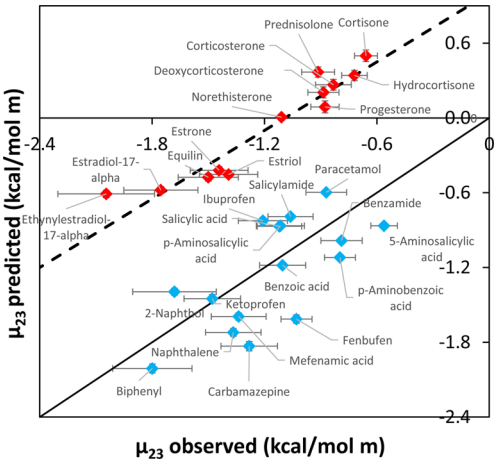


Figure 8. Comparison of predicted and observed interactions of PEG400 with aromatic compounds and steroids. Predicted vs observed μ_{23} values determined by solubility[60] for interactions of PEG400 with 28 aromatic compounds (Table S6 of the Supporting Information) with functional groups that are the same or similar to those of the model compound set investigated here. Predictions are obtained from α values for interactions of PEG400 with functional groups of proteins listed in Table S9 of the Supporting Information, using ASA information from Table S7 of the Supporting Information. Experimental uncertainty is calculated as described in the Supporting Information. Blue symbols are μ_{23} values for 16 compounds without a steroid fused ring system; red symbols are μ_{23} values for 12 steroid compounds with fused ring systems. For all 12 steroids, observed μ_{23} values are systematically more favorable (more negative) than predicted μ_{23} values by ~1.5 kcal mol^{-1} molal^{-1}.

chemical interactions and unfavorable excluded volume interactions to the interactions of low- and high-molecular weight PEGs with native proteins, relevant for understanding the effects of PEG on protein solubility, protein folding, and other protein processes.

■ ASSOCIATED CONTENT

Supporting Information

Details about methods (vapor-pressure osmometry, solubility measurements, and the lacDBD unfolding procedure), data analysis with uncertainty estimations, tabulated experimental data and ASA values, and prediction and comparison of the effect of PEG400 on the solubility of aromatic compounds. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.5b00246.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: mtrecord@wisc.edu. Phone: (608) 262-5332.

Author Contributions

D.B.K. and I.A.S. contributed equally to this research.

Author Contributions

D.B.K., N.M.P., M.S., E.L., X.C., and K.O. collected the data. I.A.S., D.B.K., L.C., and M.T.R. analyzed the data, and M.T.R., D.B.K., and I.A.S. wrote the paper.

Funding

This work was supported by National Institutes of Health Grant GM47022.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Christine Garmoe for help in searching PEG literature. We thank the reviewers for their thorough critiques and many useful recommendations, incorporated in the revision.

■ REFERENCES

(1) Finet, S., Vivares, D., Bonnete, F., and Tardieu, A. (2003) Controlling biomolecular crystallization by understanding the distinct effects of PEGs and salts on solubility. *Methods Enzymol. 368*, 105−129.

(2) Asenjo, J. A., and Andrews, B. A. (2011) Aqueous two-phase systems for protein separation: A perspective. *J. Chromatogr. A 1218*, 8826−8835.

(3) Parsegian, V. A., Rand, R. P., and Rau, D. C. (2000) Osmotic stress, crowding, preferential hydration, and binding: A comparison of perspectives. *Proc. Natl. Acad. Sci. U.S.A. 97*, 3987−3992.

(4) Timasheff, S. N. (2002) Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proc. Natl. Acad. Sci. U.S.A. 99*, 9721−9726.

(5) Strickley, R. G. (2004) Solubilizing excipients in oral and injectable formulations. *Pharm. Res. 21*, 201−230.

(6) Duncan, R. (2011) Polymer therapeutics as nanomedicines: New perspectives. *Curr. Opin. Biotechnol. 22*, 492−501.

(7) Eugene, M. (2004) Polyethyleneglycols and immunocamouflage of the cells tissues and organs for transplantation. *Cell. Mol. Biol. 50*, 209−215.

(8) Thuillier, R., Renard, C., Rogel-Gaillard, C., Demars, J., Milan, D., Forestier, L., Ouldmoulene, A., Goujon, J. M., Badet, L., and Hauet, T. (2011) Effect of polyethylene glycol-based preservation solutions on graft injury in experimental kidney transplantation. *Br. J. Surg. 98*, 368−378.

(9) Bruhn, B. R., Schroeder, T. B. H., Li, S. Y., Billeh, Y. N., Wang, K. W., and Mayer, M. (2014) Osmosis-Based Pressure Generation: Dynamics and Application. *PLoS One 9*, e91350.

(10) Sanna, V., Pala, N., and Sechi, M. (2014) Targeted therapy using nanotechnology: Focus on cancer. *Int. J. Nanomed. 9*, 467−483.

(11) Talelli, M., Rijcken, C. J. F., Hennink, W. E., and Lammers, T. (2012) Polymeric micelles for cancer therapy: 3 C's to enhance efficacy. *Curr. Opin. Solid State Mater. Sci. 16*, 302−309.

(12) Rad, I., Khodayari, K., Alijanvand, S. H., and Mobasheri, H. (2015) Interaction of polyethylene glycol (PEG) with the membrane-binding domains following spinal cord injury (SCI): Introduction of a mechanism for SCI repair. *J. Drug Targeting 23*, 79−88.

(13) Lowe, S., O'Brien-Simpson, N. M., and Connal, L. A. (2015) Antibiofouling polymer interfaces: Poly(ethylene glycol) and other promising candidates. *Polym. Chem. 6*, 198−212.

(14) Gekko, K., and Timasheff, S. N. (1981) Thermodynamic and kinetic investigation of protein stabilization by glycerol. *Biochemistry 20*, 4677−4686.

(15) Haque, I., Islam, A., Singh, R., Moosavi-Movahedi, A., and Ahmad, F. (2006) Stability of proteins in the presence of polyols estimated from their guanidinium chloride-induced transition curves at different pH values and 25 °C. *Biophys. Chem. 119*, 224−233.

(16) Back, J. F., Oakenfull, D., and Smith, M. B. (1979) Increased thermal stability of proteins in the presence of sugars and polyols. *Biochemistry 18*, 5191−5196.

(17) Auton, M., Rösgen, J., Sinev, M., Holthauzen, L. M. F., and Bolen, D. W. (2011) Osmolyte effects on protein stability and solubility: A balancing act between backbone and side-chains. *Biophys. Chem. 159*, 90−99.

(18) Politi, R., and Harries, D. (2010) Enthalpically driven peptide stabilization by protective osmolytes. *Chem. Commun. 46*, 6449−6451.

(19) Sukenik, S., Sapir, L., Gilman-Politi, R., and Harries, D. (2013) Diversity in the mechanisms of cosolute action on biomolecular processes. *Faraday Discuss. 160*, 225−237.

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

(20) Munishkina, L. A., Cooper, E. M., Uversky, V. N., and Fink, A. L. (2004) The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J. Mol. Recognit. 17*, 456−464.
(21) Silvers, T. R., and Myers, J. K. (2013) Osmolyte effects on the self-association of concanavalin A: Testing theoretical models. *Biochemistry 52*, 9367−9374.
(22) Herskovits, T. T., Jacobs, R., and Nag, K. (1983) The effects of salts and ureas on the subunit dissociation of concanavalin-A. *Biochim. Biophys. Acta 742*, 142−154.
(23) Hirota, N., Mizuno, K., and Goto, Y. (1998) Group additive contributions to the alcohol-induced α-helix formation of melittin: Implication for the mechanism of the alcohol effects on proteins. *J. Mol. Biol. 275*, 365−378.
(24) Cao, D. (2002) Cosolute effects as probes of the role of water in DNA binding and catalysis by three restriction endonucleases. Ph.D. Thesis, University of Pittsburgh, Pittsburgh, PA.
(25) Levine, L., Gordon, J. A., and Jencks, W. P. (1963) The relationship of structure to the effectiveness of denaturing agents for deoxyribonucleic acid. *Biochemistry 2*, 168−175.
(26) Spink, C. H., Garbett, N., and Chaires, J. B. (2007) Enthalpies of DNA melting in the presence of osmolytes. *Biophys. Chem. 126*, 176−185.
(27) Lambert, D., and Draper, D. E. (2007) Effects of osmolytes on RNA secondary and tertiary structure stabilities and RNA-Mg^{2+} interactions. *J. Mol. Biol. 370*, 993−1005.
(28) Knowles, D. B., LaCroix, A. S., Deines, N. F., Shkel, I., and Record, M. T. (2011) Separation of preferential interaction and excluded volume effects on DNA duplex and hairpin stability. *Proc. Natl. Acad. Sci. U.S.A. 108*, 12699−12704.
(29) Lee, L. L., and Lee, J. C. (1987) Thermal stability of proteins in the presence of poly(ethylene glycols). *Biochemistry 26*, 7813−7819.
(30) Farruggia, B., Garcia, G., Dangelo, C., and Pico, G. (1997) Destabilization of human serum albumin by polyethylene glycols studied by thermodynamical equilibrium and kinetic approaches. *Int. J. Biol. Macromol. 20*, 43−51.
(31) Kumar, V., Sharma, V. K., and Kalonia, D. S. (2009) Effect of polyols on polyethylene glycol (PEG)-induced precipitation of proteins: Impact on solubility, stability and conformation. *Int. J. Pharm. 366*, 38−43.
(32) Senske, M., Tork, L., Born, B., Havenith, M., Herrmann, C., and Ebbinghaus, S. (2014) Protein Stabilization by Macromolecular Crowding through Enthalpy Rather Than Entropy. *J. Am. Chem. Soc. 136*, 9036−9041.
(33) Jarvis, T. C., Ring, D. M., Daube, S. S., and von Hippel, P. H. (1990) Macromolecular crowding: Thermodynamic consequence for protein-protein interactions within the T4 DNA-replication complex. *J. Biol. Chem. 265*, 15160−15167.
(34) Zimmerman, S. B., and Harrison, B. (1985) Macromolecular crowding accelerates the cohesion of DNA fragments with complementary termini. *Nucleic Acids Res. 13*, 2241−2249.
(35) Dupuis, N. F., Holmstrom, E. D., and Nesbitt, D. J. (2014) Molecular-crowding effects on single-molecule RNA folding/unfolding thermodynamics and kinetics. *Proc. Natl. Acad. Sci. U.S.A. 111*, 8464−8469.
(36) Guinn, E. J., Pegram, L. M., Capp, M. W., Pollock, M. N., and Record, M. T. (2011) Quantifying why urea is a protein denaturant, whereas glycine betaine is a protein stabilizer. *Proc. Natl. Acad. Sci. U.S.A. 108*, 16932−16937.
(37) Diehl, R. C., Guinn, E. J., Capp, M. W., Tsodikov, O. V., and Record, M. T., Jr. (2013) Quantifying Additive Interactions of the Osmolyte Proline with Individual Functional Groups of Proteins: Comparisons with Urea and Glycine Betaine, Interpretation of m-Values. *Biochemistry 52*, 5997−6010.
(38) Capp, M. W., Pegram, L. M., Saecker, R. M., Kratz, M., Riccardi, D., Wendorff, T., Cannon, J. G., and Record, M. T., Jr. (2009) Interactions of the osmolyte glycine betaine with molecular surfaces in water: Thermodynamics, structural interpretation, and prediction of m-values. *Biochemistry 48*, 10372−10379.
(39) Pegram, L. M., Wendorff, T., Erdmann, R., Shkel, I., Bellissimo, D., Felitsky, D. J., and Record, M. T. (2010) Why Hofmeister effects of many salts favor protein folding but not DNA helix formation. *Proc. Natl. Acad. Sci. U.S.A. 107*, 7716−7721.
(40) Pegram, L. M., and Record, M. T. (2008) Thermodynamic origin of Hofmeister ion effects. *J. Phys. Chem. B 112*, 9428−9436.
(41) Nozaki, Y., and Tanford, C. (1963) The solubility of amino acids and related compounds in aqueous urea solutions. *J. Biol. Chem. 238*, 4074−4081.
(42) Davis-Searles, P. R., Saunders, A. J., Erie, D. A., Winzor, D. J., and Pielak, G. J. (2001) Interpreting the effects of small uncharged solutes on protein-folding equilibria. *Annu. Rev. Biophys. Biomol. Struct. 30*, 271−306.
(43) Shkel, I. A., Knowles, B., and Record, M. T. (2015) Separating chemical and excluded volume interactions of polyethylene glycols with native proteins: Comparison with PEG effects on DNA helix formation. *Biopolymers*, DOI: 10.1002/bip.22662.
(44) Smith, P. E. (2006) Chemical potential derivatives and preferential interaction parameters in biological systems from Kirkwood-Buff theory. *Biophys. J. 91*, 849−856.
(45) Record, M. T., Jr., 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*, 786−794.
(46) Holehouse, A. S., Garai, K., Lyle, N., Vitalis, A., and Pappu, R. V. (2015) Quantitative assessments of the distinct contributions of polypeptide backbone amides versus sidechain groups to chain expansion via chemical denaturation. *J. Am. Chem. Soc. 137*, 2984−2995.
(47) Miklos, A. C., Li, C., Sharaf, N. G., and Pielak, G. J. (2010) Volume exclusion and soft interaction effects on protein stability under crowded conditions. *Biochemistry 49*, 6984−6991.
(48) Minton, A. P. (2013) Quantitative assessment of the relative contributions of steric repulsion and chemical interactions to macromolecular crowding. *Biopolymers 99*, 239−244.
(49) Hermans, J. (1982) Excluded-Volume Theory of Polymer Protein Interactions Based on Polymer-Chain Statistics. *J. Chem. Phys. 77*, 2193−2203.
(50) Benton, L. A., Smith, A. E., Young, G. B., and Pielak, G. J. (2012) Unexpected effects of macromolecular crowding on protein stability. *Biochemistry 51*, 9773−9775.
(51) Wang, Y., Sarkar, M., Smith, A. E., Krois, A. S., and Pielak, G. J. (2012) Macromolecular crowding and protein stability. *J. Am. Chem. Soc. 134*, 16614−16618.
(52) Johansen, D., Jeffries, C. M., Hammouda, B., Trewhella, J., and Goldenberg, D. P. (2011) Effects of macromolecular crowding on an intrinsically disordered protein characterized by small-angle neutron scattering with contrast matching. *Biophys. J. 100*, 1120−1128.
(53) Elcock, A. H. (2010) Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Curr. Opin. Struct. Biol. 20*, 196−206.
(54) Zhou, H. X., Rivas, G., and Minton, A. P. (2008) Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys. 37*, 375−397.
(55) Robinson, R. A., and Stokes, R. H. (1961) Activity Coefficients in Aqueous Solutions of Sucrose, Mannitol and Their Mixtures at 25 °C. *J. Phys. Chem. 65*, 1954−1958.
(56) Cannon, J. G., Anderson, C. F., and Record, M. T. (2007) Urea-amide preferential interactions in water: Quantitative comparison of model compound data with biopolymer results using water accessible surface areas. *J. Phys. Chem. B 111*, 9675−9685.
(57) Tsodikov, O. V., Record, M. T., Jr., and Sergeev, Y. V. (2002) Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. *J. Comput. Chem. 23*, 600−609.
(58) Felitsky, D. J., Cannon, J. G., Capp, M. W., Hong, J., Van Wynsberghe, A. W., Anderson, C. F., and Record, M. T. (2004) The exclusion of glycine betaine from anionic biopolymer surface: Why glycine betaine is an effective osmoprotectant but also a compatible solute. *Biochemistry 43*, 14732−14743.

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542

(59) 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*, 2202−2217.
(60) Rytting, E., Lentz, K. A., Chen, X. Q., Qian, F., and Venkatesh, S. (2005) Aqueous and cosolvent solubility data for drug-like organic compounds. *AAPS J. 7*, E78−E105.
(61) Guinn, E. J., Schwinefus, J. J., Cha, H. K., McDevitt, J. L., Merker, W. E., Ritzer, R., Muth, G. W., Engelsgjerd, S. W., Mangold, K. E., Thompson, P. J., Kerins, M. J., and Record, M. T., Jr. (2013) Quantifying Functional Group Interactions That Determine Urea Effects on Nucleic Acid Helix Formation. *J. Am. Chem. Soc. 135*, 5828−5838.

DOI: 10.1021/acs.biochem.5b00246
Biochemistry 2015, 54, 3528−3542