Fluoroalkyl and Alkyl Chains Have Similar Hydrophobicities in Binding to the "Hydrophobic Wall" of Carbonic Anhydrase

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · AUGUST 2011

Impact Factor: 12.11 · DOI: 10.1021/ja2045293 · Source: PubMed

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

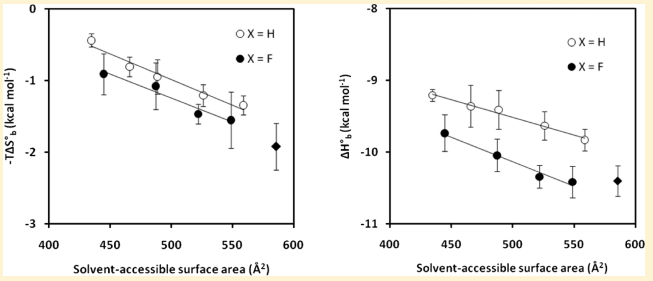
Fluoroalkyl and Alkyl Chains Have Similar Hydrophobicities in Binding to the "Hydrophobic Wall" of Carbonic Anhydrase

Jasmin Mecinović,[†] Phillip W. Snyder,[†] Katherine A. Mirica,[†] Serena Bai,[†] Eric T. Mack,[†] Richard L. Kwant,[†] Demetri T. Moustakas,[†] Annie Héroux,[§] and George M. Whitesides*[,†,‡]

[†]Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, United States
[‡]Wyss Institute for Biologically Inspired Engineering and [§]National Synchrotron Light Source, Brookhaven National Laboratory, 725 Brookhaven Avenue, Upton, New York 11973-5000, United States

Supporting Information

ABSTRACT: The hydrophobic effect, the free-energetically favorable association of nonpolar solutes in water, makes a dominant contribution to binding of many systems of ligands and proteins. The objective of this study was to examine the hydrophobic effect in biomolecular recognition using two chemically different but structurally similar hydrophobic groups, aliphatic hydrocarbons and aliphatic fluorocarbons, and to determine whether the hydrophobicity of the two groups could be distinguished by thermodynamic and biostructural analysis. This paper uses isothermal titration calorimetry (ITC) to examine the thermodynamics of binding of benzenesulfonamides substituted in the para position with alkyl and fluoroalkyl chains ($H_2NSO_2C_6H_4$-$CONHCH_2(CX_2)_nCX_3$, n = 0−4, X = H, F) to human carbonic anhydrase II (HCA II). Both alkyl and fluoroalkyl substituents contribute favorably to the enthalpy *and* the entropy of binding; these contributions increase as the length of chain of the hydrophobic substituent increases. Crystallography of the protein−ligand complexes indicates that the benzenesulfonamide groups of all ligands examined bind with similar geometry, that the tail groups associate with the hydrophobic wall of HCA II (which is made up of the side chains of residues Phe131, Val135, Pro202, and Leu204), and that the structure of the protein is indistinguishable for all but one of the complexes (the longest member of the fluoroalkyl series). Analysis of the thermodynamics of binding as a function of structure is compatible with the hypothesis that hydrophobic binding of both alkyl and fluoroalkyl chains to hydrophobic surface of carbonic anhydrase is due primarily to the release of nonoptimally hydrogen-bonded water molecules that hydrate the binding cavity (including the hydrophobic wall) of HCA II and to the release of water molecules that surround the hydrophobic chain of the ligands. This study defines the balance of enthalpic and entropic contributions to the hydrophobic effect in this representative system of protein and ligand: hydrophobic interactions, here, seem to comprise approximately equal contributions from enthalpy (plausibly from strengthening networks of hydrogen bonds among molecules of water) and entropy (from release of water from configurationally restricted positions).



INTRODUCTION

Hydrophobic Interactions in Protein−Ligand Binding. Hydrophobic interactions, the free energetically favorable aggregation of nonpolar molecules in aqueous media, are centrally important in biology because they dominate the folding of proteins, the formation of lipid bilayers, and the association of proteins and ligands.[1−3] The classical concept of hydrophobic interactions, which we attribute to Kauzmann and Tanford (KT), predicts that (i) water near the surface of hydrophobic groups is more (or, perhaps, just differently) structured than bulk water, and that (ii) entropy dominates the favorable free energy of hydrophobic interactions because association of two nonpolar surfaces causes the release of structured molecules of water near nonpolar surfaces.[1,2] The molecular basis of hydrophobic interactions in protein−ligand association is more complicated than this classical description and still incompletely understood.[4−6] The distinction between hydrophobic effects when different types of groups, aliphatic and aromatic hydrocarbons, or fluorocarbons, interact within a protein−ligand complex is also not clear. This lack of understanding (probably) contributes to the present difficulty in *designing* ligands that associate tightly with proteins.

Aliphatic Hydrocarbons and Fluorocarbons in Biomolecular Recognition. Both aliphatic hydrocarbons (R_H) and aliphatic fluorocarbons (R_F) are hydrophobic, in that they are poorly soluble in water,[7] but the thermodynamic basis of this hydrophobicity, at least in the context of protein−ligand interactions, is poorly characterized. In drug discovery, replacement of hydrocarbon groups by fluorocarbon groups has been used to modify solubility and basicity, to test for hydrogen bonding interactions, and to improve the

Received: May 17, 2011
Published: July 26, 2011

Table 1. Comparison of the Dependence of the ΔG°_b, ΔH°_b, and $-T\Delta S^\circ_b$ on the Chain Length (n), Ligand Surface Area (A), Ligand Solvent-Accessible Surface Area (SASA), and Volume of the Ligand (V) for Alkyl (X = H) and Fluoroalkyl (X = F) Tails of $H_2NSO_2C_6H_4CONHCH_2(CX_2)_nCX_3$ (n = 0−4) to HCA II

	hydrocarbon, R_H	fluorocarbon, R_F
$\Delta\Delta G^\circ_b/\Delta n^a$(cal mol^{-1})	−366 ± 30	−479 ± 35
$\Delta\Delta G^\circ_b/\Delta A^b$(cal mol^{-1} Å^{-2})	−18 ± 1	−18 ± 1
$\Delta\Delta G^\circ_b/\Delta$SASA^c(cal mol^{-1}Å^{-2})	−12 ± 1	−14 ± 1
$\Delta\Delta G^\circ_b/\Delta V^d$(cal mol^{-1} Å^{-3})	−20 ± 2	−20 ± 2
$\Delta\Delta H^\circ_b/\Delta n^a$(cal mol^{-1})	−150 ± 30	−247 ± 37
$\Delta\Delta H^\circ_b/\Delta A^b$(cal mol^{-1} Å^{-2})	−7 ± 1	−9 ± 1
$\Delta\Delta H^\circ_b/\Delta$SASA^c(cal mol^{-1} Å^{-2})	−5 ± 1	−7 ± 1
$\Delta\Delta H^\circ_b/\Delta V^d$(cal mol^{-1} Å^{-3})	−8 ± 2	−10 ± 2
$-T\Delta\Delta S^\circ_b/\Delta n^a$ (cal mol^{-1})	−216 ± 23	−232 ± 48
$-T\Delta\Delta S^\circ_b/\Delta A^b$(cal mol^{-1}Å^{-2})	−11 ± 1	−9 ± 2
$-T\Delta\Delta S^\circ_b/\Delta$SASA^c(cal mol^{-1} Å^{-2})	−7 ± 1	−7 ± 1
$-T\Delta\Delta S^\circ_b/\Delta V^d$(cal mol^{-1}Å^{-3})	−12 ± 1	−10 ± 2

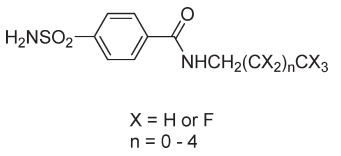
[a] Obtained from the slope of ΔG°_b, ΔH°_b, or $-T\Delta S^\circ_b$ vs the number of carbon atoms of the tail (Figures 2−4). [b] Obtained from the slope of ΔG°_b, ΔH°_b or $-T\Delta S^\circ_b$ vs the surface area (A) of the tail in the fully extended conformation (Figures S1−S3 in Supporting Information). [c] Obtained from the slope of ΔG°_b, ΔH°_b, or $-T\Delta S^\circ_b$ vs the solvent-accessible surface area (SASA) of the tail in the fully extended conformation (Figures 2−4). [d] Obtained from the slope of ΔG°_b, ΔH°_b or $-T\Delta S^\circ_b$ vs the volume (V) of the tail in the fully extended conformation (Figures S1−S3).

metabolic stability, binding affinity, and bioavailability of several compounds.[8] Incorporation of fluorocarbons into proteins and peptides results in the stabilization of folded proteins and promotion of self-assembly of α-helical peptides into coiled coils.[9,10] Resistance of these structures to thermal denaturation suggests that fluorinated analogues are more stable than hydrocarbons, although it remains unclear whether this effect is due to increased hydrophobic surface area alone or to a difference in the character of hydrophobicity.

The results of our own studies involving (i) binding of ligands modified with R_H and R_F tails to bovine carbonic anhydrase (BCA)[11] and (ii) denaturation of BCA modified with a series of R_H and CF_3 substituents in the presence of sodium dodecyl sulfate[12] suggest, however, that the free energy of interaction of R_H and R_F with hydrophobic surfaces can be rationalized entirely or predominantly based on the amount of solvent-accessible surface area (CF_3CONH groups are 0.05−0.07 kcal mol^{-1} more hydrophobic than R_HCONH groups with the same surface area). (A previous paper and relevant reviews summarize the background on the hydrophobic effect due to R_F).[12,13]

The Thermodynamics of Association of Series of Ligands Presenting R_H and R_F Groups with Human Carbonic Anhydrase. The current study uses isothermal titration calorimetry to measure the values of the free energy (ΔG°_b), enthalpy (ΔH°_b), and entropy ($-T\Delta S^\circ_b$) for the binding to human carbonic anhydrase II (HCA II, EC 4.2.1.1) of two series of benzenesulfonamide ligands ($H_2NSO_2C_6H_4CONHCH_2R_{H/F}$): in one series, the substituents in the 4-position of benzenesulfonamide are R_H groups of increasing length, and in the second series, the substituents are R_F groups of increasing length. To compare the two series, we estimate the incremental values of the thermodynamic parameters of binding ($\Delta\Delta G^\circ_b$, $\Delta\Delta H^\circ_b$, and $-T\Delta\Delta S^\circ_b$; defined as a difference of thermodynamics of binding for two ligands in the same series and obtained as a slope of ΔJ°_b (where $J = G, H, S$) vs chain length or surface area) based on measurements of (i) the solvent-accessible surface area, and (ii) the molecular volume of the R_H and R_F groups in the crystal structures of the protein−ligand complexes. We determined that values of $\Delta\Delta G^\circ_b$ are indistinguishable within statistical uncertainty on the basis of ligand solvent-accessible surface area (for R_H $\Delta\Delta G^\circ_b = -12 \pm 1$ cal mol^{-1} Å^{-2}; for R_F $\Delta\Delta G^\circ_b = -14 \pm 1$ cal mol^{-1} Å^{-2}), and on the basis of the volume of the ligand (for R_H $\Delta\Delta G^\circ_b = -20 \pm 2$ cal mol^{-1} Å^{-3}; for R_F $\Delta\Delta G^\circ_b = -20 \pm 2$ cal mol^{-1} Å^{-3}) (Table 1). Both alkyl and fluoroalkyl groups contribute to $\Delta\Delta G^\circ_b$ through favorable values of *both* $\Delta\Delta H^\circ_b$ and $-T\Delta\Delta S^\circ_b$. The magnitude of $\Delta\Delta H^\circ_b$ is indistinguishable (within experimental error) for alkyl and fluoroalkyl groups (for alkyls $\Delta\Delta H^\circ_b = -5 \pm 1$ cal mol^{-1} Å^{-2} and for fluoroalkyls $\Delta\Delta H^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2}). Alkyls and fluoroalkyls also have indistinguishable values of $-T\Delta\Delta S^\circ_b$ (for alkyls $-T\Delta\Delta S^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2}, for fluoroalkyls $-T\Delta\Delta S^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2}). Table 1 summarizes the results of this study as a guide to subsequent details.

We used para-substituted benzenesulfonamides connected to hydrophobic side chains, "greasy tails",[11] via an amide linkage ($H_2NSO_2C_6H_4CONHCH_2(CX_2)_nCX_3$, where $n = 0-4$, X = H, F) as ligands for HCA II.



In the system of HCA II and derivatives of benzenesulfonamide, association of the benzenesulfonamide moiety ($^-HNSO_2C_6H_4$-) is essentially invariant to most changes in the structure of the R group in the $H_2NSO_2C_6H_4R$ group.[14] Binding is determined by association of this $^-HNSO_2C_6H_4$- group to the active site Zn^{2+} ion, and many biostructural data establish that the geometry of the phenyl group in the active site is highly conserved.[15] We have used the extreme simplicity of the system of HCA (and structurally very similar BCA) as the basis for detailed physical−organic studies of binding of ligands to HCA II.[14]

This paper reports the values of free energy of binding (ΔG°_b), enthalpy of binding (ΔH°_b), and entropy of binding ($-T\Delta S^\circ_b$) for the two series of ligands (R = R_H, R_F; $H_2NSO_2C_6H_4CONHR$) measured using isothermal titration calorimetry (ITC). We calculated the incremental changes in free energy, enthalpy, and entropy ($\Delta\Delta G^\circ_b$, $\Delta\Delta H^\circ_b$, and $-T\Delta\Delta S^\circ_b$) of binding by correlating the thermodynamic parameters for binding with (i) the change in solvent-accessible surface area of binding, and (ii) the molecular volume of the ligand. These incremental values represent the contribution to binding per unit area of hydrocarbon−hydrocarbon and hydrocarbon−fluorocarbon interaction, and per unit volume of hydrocarbon or fluorocarbon for each series.

Residues Phe131, Val135, Pro202, and Leu204 comprise the so-called "hydrophobic wall" of HCA II.[14] We guessed, based on the crystallography of structurally similar ligands, and validated by our own structural studies that the hydrophobic tails of para-substituted benzenesulfonamides would form van der Waals contacts with the hydrophobic wall. Examination of many ligands

for HCA II (and BCA II) has demonstrated that hydrophobic groups, and specifically groups of type $H_2NSO_2C_6H_4CONHCH_2R$, with R = *n*-alkyl, *n*-fluoroalkyl, increase the strength of binding as the putative area of contact between the ligand and the protein increases. The system that examines binding of benzenesulfonamide ligands ($H_2NSO_2C_6H_4$-R, where R represents various organic moieties) to carbonic anhydrase is thus an excellent one (the best, we believe, so far developed) for physical−organic studies of the hydrophobic effect in a biologically relevant system comprising protein and ligand.[14] It is particularly interpretable since the rigidity of the tertiary structure of CA II makes contributions to binding from protein plasticity negligible.[14]

The first objective of this study was to explore the relationship between the hydrophobic effect and ligand structure, using two chemically different, but structurally related classes of hydrophobic groups: alkyls (R_H) and fluoroalkyls (R_F). Our hypothesis was that either (i) the hydrophobic effect is due primarily to exclusion of water from the hydrophobic surfaces of the active site and of the ligand, in which case, the magnitude of the effect for homologous R_H and R_F tails interacting with the hydrophobic wall of HCA II would be the same when adjusted for differences in the solvent-accessible surface areas of the tails, or (ii) the hydrophobic effect results from the physical properties of R_H and R_F (i.e., polarizability, van der Waals interactions, etc.), in which case, the magnitude of the effect might be quite different for the two types of tails, since these properties are different for R_H and R_F.

Our second objective in comparing R_H and R_F was to define their relative hydrophobicity in the context of protein−ligand interactions. Incorporation of fluorine into small molecules is an important tactic in designing inhibitors of proteins.[16] This strategy is often used to increase binding affinity, to improve membrane permeability, and to augment metabolic stability of pharmaceuticals. There is a widespread belief, based primarily and qualitatively on the hydrophobicity and oleophobicity of Teflon, that R_H and R_F are fundamentally different in their hydrophobicity.[17] In many (in fact, most) of the systems studied, the hydrophobicities for R_H and R_F are different, but there is no thermodynamic evidence to support this argument.

Improved understanding of hydrophobic interactions involving R_H and R_F in the context of protein−ligand binding will clarify the nature of the hydrophobic effect in biomolecular recognition, assist advances in the rational design of inhibitors for enzymes, and help to understand the basis of interactions of both fluorocarbons and hydrocarbons with proteins.

Additional background information on the hydrophobic effect is summarized in the Supporting Information.

EXPERIMENTAL DESIGN

Choice of Protein−Ligand System. We choose HCA II as a model system for our physical−organic study for five reasons: (i) HCA II is an exceptionally stable and rigid protein. It can be obtained readily by expression in *E. coli* (∼100 mg L^{-1} of growth medium), using techniques with which we are familiar, and obtained in the quantities necessary for ITC (∼0.5 mg per experiment)[18] and X-ray crystallography; (ii) numerous benzenesulfonamide-containing ligands ($H_2NSO_2C_6H_4$-R, with some constraints on R), bind in the active site of HCA II with a conserved geometry;[14] (iii) previous X-ray and neutron crystallographic studies have detailed the structure of sulfonamide ligands bound in the active site of HCA;[14,19] (iv) a single protocol for growing crystals of the protein with different ligands can be used; (v) one face of the active site of HCA II is a hydrophobic "shelf" or "wall", comprising residues

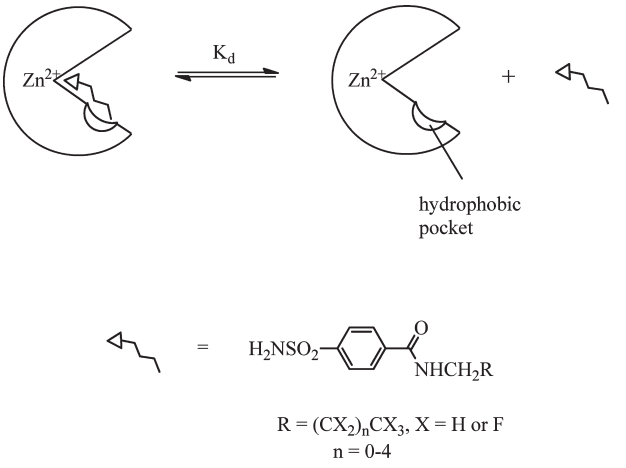


Figure 1. Our approach to increasing the binding affinity of para-substituted benzensulfonamide ligands to HCA II employs hydrophobic interactions between hydrophobic tails of ligands and the hydrophobic wall of the protein.

Phe131, Val135, Pro202, and Leu204.[14] This wall has ∼250 Å^2 of solvent-accessible hydrophobic surface, and substituents in the para-position of benzenesulfonamide are positioned over that part of the active site (Figure 1).

Perturbational Approach for Probing Binding. To probe the interactions of "greasy tails" with the hydrophobic region adjacent to the active site of HCA II, we have followed a perturbational approach: we used the *p*-carboxamido benzenesulfonamide group to anchor the ligand in the active site of the protein in a well-defined, conserved geometry, and we systematically varied the length of R_H and R_F chains ($(CX_2)_n$, where X = H, F and n = 0−4) in the para-position. Previous structural analyses, and data we present here, indicated that this anchor would preserve the geometry of the arylsulfonamide group, which makes the dominant contribution to thermodynamics of binding in this system (∼8 kcal mol^{-1}), regardless of the nature of the greasy tail.[15] This structural rigidity is essential to the perturbational approach that we use because, as we show below, the difference in the contribution to the thermodynamics of the interaction between methylene or fluoromethylene groups and the hydrophobic wall of HCA is less than 0.5 kcal mol^{-1}.

Since this value is roughly the same as the uncertainty in the measurement of $\Delta H°_b$ (or $-T\Delta S°_b$) by ITC for any single ligand, the comparison of any pair of ligands would not be statistically meaningful. The perturbational approach, which in this work includes analyses of five ligands of each series, thus allows us to evaluate the similarities (or differences) between R_H and R_F tails with greater precision than would be possible for pairs of structurally homologous compounds.

One potential limitation of using the *p*-carboxamido benzenesulfonamide anchor, rather than the *N*-methylcarboxamides, for example, could be differences in the values of pK_a of the carboxamide group for R_H and R_F series. Involvement of the amide NH group may contribute favorably to the enthalpy of binding via hydrogen-bonding (NH···H_2O···Thr200, Pro201), a hydrogen bond that we observe by X-ray crystallography (Figure S4 in Supporting Information). In our previous studies, indeed, we demonstrated that the pK_a of the carboxamide group for the R_F series is lower than that of the R_H series, and the value of $\Delta G°_b$ of the R_F tails were more favorable than that of the R_H tails.[11] In that work, however, we also measured the values of $\Delta\Delta G°_b$ for both series of *N*-methylcarboxamides and determined that the difference in the pK_a of the carboxamide group did not influence the values of $\Delta\Delta G°_b$ for either the R_H or the R_F series. We also show here that, although the NMR shifts of the carboxamide protons of R_H and R_F are different, they are the same across each series, and we infer that the values of $\Delta\Delta H°_b$ and $-T\Delta\Delta S°_b$

reflect contributions from the hydrophobic interactions between $R_{H/F}$ and the hydrophobic wall.

From our previous work with these groups, we anticipated negative values of $\Delta\Delta G°_b$ for both R_H and R_F tails.[11] Our objective was to analyze the enthalpic and entropic contributions to this free energy of binding by using ITC and to correlate these contributions with the structures and physical properties of the molecules.

RESULTS

Synthesis of the Ligands. We prepared benzenesulfonamides with alkyl and fluoroalkyl tails following the previously reported procedures.[11]

Purification of the Protein. To purify HCA II (EC 4.2.1.1, >95% pure), we followed the procedures reported by Fierke et al., who also kindly provided the plasmids containing the gene for the protein.[20−22] Details of this procedure appear in the Supporting Information.

Collection of Data by ITC. Because of the low solubility (<50 μM) in aqueous buffer of the ligands that had long ($n > 2$) tails, we expected it to be challenging to conduct ITC experiments, which require that the concentration of molecule in the cell of the calorimeter to be no higher than $10^3 \times K_d$, and that the concentration of molecule in the syringe be ∼10 times the concentration of the molecule in the cell.[18] Placing solutions of ligand in the cell not only set a lower limit on the concentration of ligand needed but also allowed us to minimize the contribution to the uncertainty in $\Delta\Delta H°_b$ from the uncertainty in the concentration of ligand.[23] We titrated aliquots of HCA II (20 μM), taken from a single batch, into solutions of each of the 10 ligands (∼2.0 μM). The concentration of ligands was measured in the presence of maleic acid as an internal standard in DMSO-d_6 by ^{1}H NMR. The observed protein:ligand stoichiometry was typically around 0.9 ± 0.02, possibly due to the adsorption of greasy tail ligands to the plastic tray. On the basis of these, and other ITC experiments in our group using carbonic anhydrase, we exclude a possibility that HCA II adsorbs to the surface of the glass syringe; if HCA II were to adsorb to the syringe, the protein:ligand stoichiometry would be greater than 1. By assuming that the concentration of active protein was the same in each experiment, we were able to adjust the stoichiometry of protein−ligand binding to n = 1 during analysis of the data. ITC experiments with each ligand were repeated 7−9 times. We report the average values of $\Delta G°_b$, $\Delta H°_b$ and $-T\Delta S°_b$ and estimate their uncertainties as standard deviations (for number of repeated experiments $N \geq 7$).

$\Delta G°_b$ Is Proportional to the Solvent-Accessible Surface Area of the Ligand. ITC experiments confirmed our previous observation that extending the length of the "greasy tail" resulted in more negative values of $\Delta G°_b$ and lower values of K_d for dissociation from HCA II (Figure 2). Fluoroalkyls, in general, display higher affinity (∼1 kcal mol^{-1}) for HCA II than alkyls with the same number of carbon atoms in the "greasy tail". We rationalize this effect, at least in part, by the fact that CF_2 groups have larger hydrophobic surface area than CH_2 groups (see below). The electron-withdrawing properties of fluoroalkyl chains, and their absence in alkyl chains, moreover, result in more acidic amidic NH in fluoroalkyl amides ($ArCONHCH_2R_F$) than in alkyl amides ($ArCONHCH_2R_H$). This inference of a difference in acidity is supported by chemical shifts in ^{1}H NMR spectra (Figure S5 in Supporting Information). Table 2 summarizes the thermodynamic values for binding of para-substituted benzenesulfonamides to HCA II. Our data are consistent

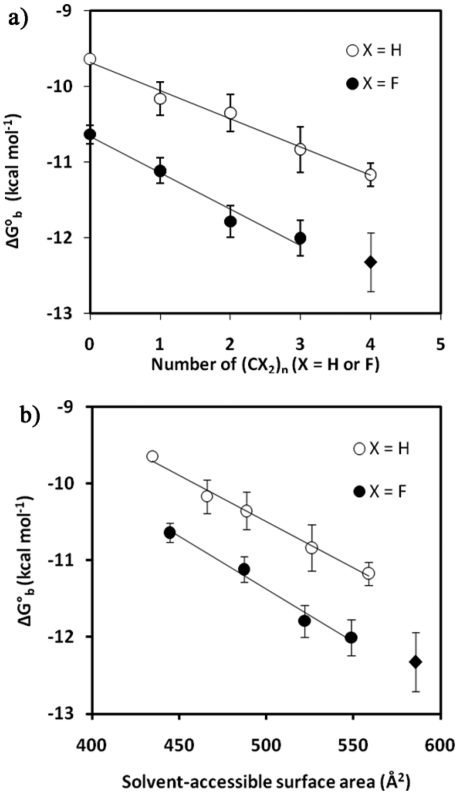


Figure 2. (a) Dependence of $\Delta G°_b$ for binding of benzenesulfonamide ligands ($H_2NSO_2C_6H_4-CONHCH_2(CX_2)_nCX_3$, X = H, F) with alkyl (○) and fluoroalkyl (●) tails on their chain length. The slope of the regression line through the alkyl data is $\Delta\Delta G°_b = -366 \pm 30$ cal mol^{-1} $-CH_2$-$^{-1}$ ($R^2 = 0.99$), and that for fluoroalkyl data is $\Delta\Delta G°_b = -479 \pm 35$ cal mol^{-1} -CF_2-$^{-1}$ ($R^2 = 0.97$) ($N = 7$). (b) Dependence of $\Delta G°_b$ for binding of benzenesulfonamide ligands with alkyl (○) and fluoroalkyl (●) tails on their solvent-accessible surface area in the fully extended conformation. The slope of the regression line for the alkyl data is $\Delta\Delta G°_b = -12 \pm 1$ cal mol^{-1} Å^{-2} ($R^2 = 0.99$), and that for the fluoroalkyl data is $\Delta\Delta G°_b = -14 \pm 1$ cal mol^{-1} Å^{-2} ($R^2 = 0.98$) ($N = 7$). The longest fluoroalkyl ligand (♦) is not included in the linear regression, because, in contrast to other fluoroalkyl ligands, it causes a flip in the orientation of Gln136 of HCA II (see the text). The analysis of data with all fluoroalkyl ligands ($n = 0-4$) gives $\Delta\Delta G°_b = -418$ cal mol^{-1} -CF_2-$^{-1}$ ($R^2 = 0.96$) and $\Delta\Delta G°_b = -12$ cal mol^{-1} Å^{-2} ($R^2 = 0.98$).

with those reported by Gao et al.[11] $\Delta\Delta G°_b$ values (defined as $\Delta G°_b/\Delta n$) are -366 ± 30 cal mol^{-1} per CH_2 group and -479 ± 35 cal mol^{-1} per CF_2 group (Gao et al. used fluorescence-based assay in their studies of binding of greasy tail ligands to BCA II and reported values of −560 cal mol^{-1} per CH_2 and −860 cal mol^{-1} per CF_2).

On the basis of our X-ray crystallographic analysis of HCA II−ligand complexes (see later in this paper), we excluded the data for the longest fluoroalkyl ligand (X = F, n = 4) from the analysis of thermodynamics of interactions for R_F (for free energy, enthalpy and entropy of binding; see below). The crystal structure of HCA II in the complex with the fluoroalkyl ligand that contains the longest side chain (n = 4), in contrast to structures with all other fluoroalkyl and alkyl ligands, shows that Gln136 flips its orientation (the gauche conformation in the case of X = F, n = 4, and the anti conformation in all other cases).

Both Series of Ligands Display a Favorable Incremental Entropy of Binding. Figure 3 shows that the free energies of

Table 2. Thermodynamic Parameters ΔG°_b, ΔH°_b, $-T\Delta S^\circ_b$, and ΔC_p for Binding of Ligands $(H_2NSO_2C_6H_4CONHCH_2(CX_2)_nCX_3$, $n = 0-4$, X = H, F) to HCA II[a]

n	X	SASA (Å^2)	K_d (nM)	ΔG°_b (kcal mol^{-1})	ΔH°_b (kcal mol^{-1})	$-T\Delta S^\circ_b$ (kcal mol^{-1})	ΔC_p (cal mol^{-1} K^{-1})	PDB ID
0	H	435	86	−9.6	−9.2	−0.4	−70 ± 8	3RYV
1	H	466	38	−10.2	−9.4	−0.8	−82 ± 4	3RYY
2	H	489	28	−10.4	−9.4	−1.0	−88 ± 7	3RZ0
3	H	526	13	−10.8	−9.6	−1.2	n.d.[b]	3RZ5
4	H	559	6.7	−11.2	−9.8	−1.3	n.d.[b]	3RZ8
0	F	445	16	−10.6	−9.7	−0.9	−72 ± 5	3RYJ
1	F	487	7.3	−11.1	−10.0	−1.1	−80 ± 6	3RYX
2	F	522	2.4	−11.8	−10.3	−1.5	−91 ± 5	3RYZ
3	F	549	1.8	−12.0	−10.5	−1.6	n.d.[b]	3RZ1
4	F	586	1.1	−12.3	−10.4	−1.9	n.d.[b]	3RZ7

[a] The solvent-accessible surface area (SASA) of ligands was calculated using the Molecular Operating Environment (MOE) suite. [b] n.d. = not determined.

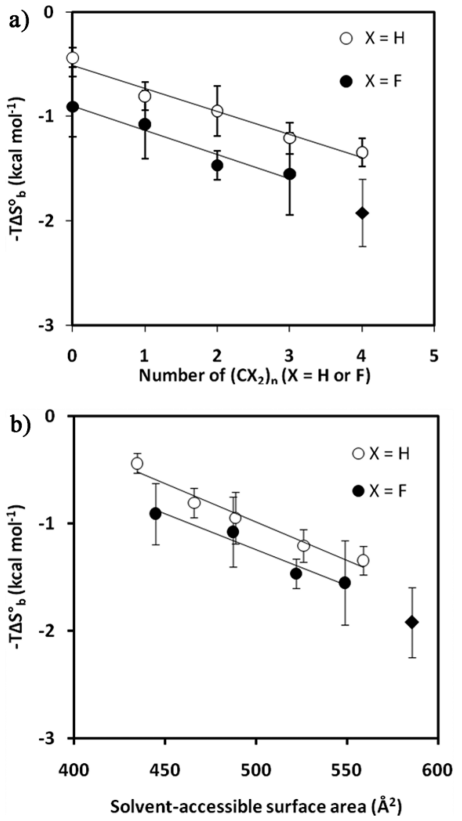


Figure 3. (a) Plots for $-T\Delta S^\circ_b$ versus chain length for alkyl (○) and fluoroalkyl (●) tails of benzensulfonamide ligands ($H_2NSO_2C_6H_4$-$CONHCH_2(CX_2)_nCX_3$, X = H, F). The slope of the regression line through the alkyl data is $-T\Delta\Delta S^\circ_b = -217 \pm 23$ cal mol^{-1} -CH_2-$^{-1}$ ($R^2 = 0.97$), and that for fluoroalkyl data is $-T\Delta\Delta S^\circ_b = -232 \pm 48$ cal mol^{-1} -CF_2-$^{-1}$ ($R^2 = 0.94$) ($N = 7$). (b) Plots for $-T\Delta S^\circ_b$ versus solvent-accessible surface area for alkyl (○) and fluoroalkyl (●) tails. The slope of the regression line for the alkyl data is $-T\Delta\Delta S^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2} ($R^2 = 0.98$), and that for the fluoroalkyl data is $-T\Delta\Delta S^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2} ($R^2 = 0.95$) ($N = 7$). The longest fluoroalkyl ligand (♦) is not included in the trendline, because it causes the conformational change of Gln136 of HCA II (see the text). The analysis of data with all fluoroalkyl ligands ($n = 0-4$) gives $-T\Delta\Delta S^\circ_b = -234$ cal mol^{-1} -CF_2-$^{-1}$ ($R^2 = 0.88$), $-T\Delta\Delta S^\circ_b = -7$ cal mol^{-1} A^{-2} ($R^2 = 0.90$).

binding for both series of ligands (R_H and R_F) have a favorable entropic contribution that increases with the length of the hydrophobic tail (for alkyls $-T\Delta\Delta S^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2}, for fluoroalkyls $-T\Delta\Delta S^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2}). This contribution could be (i) related to changes in the structure of the network of waters that hydrate the ligand and/or the active site of the protein, (ii) the result of a change in the conformational degrees of freedom of the tail on binding, or (in principle) (iii) the result of changes in the conformational degrees of freedom of amino acid side chains on binding of ligand.

Both Series of Ligands Also Display a Favorable Incremental Enthalpy of Binding. Extending the chain length of alkyl and fluoroalkyl tails also contributed to the free energy of binding through a favorable enthalpic term (Figure 4). The slope of enthalpy as a function of chain length (or ligand solvent-accessible surface area) is indistinguishable for alkyl and fluoroalkyl tails (from the alkyl data, $\Delta\Delta H^\circ_b = -5 \pm 1$ cal mol^{-1} Å^{-2}, and from the fluoroalkyl data, $\Delta\Delta H^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2}; the uncertainties in these values overlap).

Both Series Have Indistinguishable Heat Capacities. We measured heat capacities (ΔC_p) for binding of the first three ($n = 0-2$) R_H and R_F ligands to HCA II in the temperature range between 283 and 303 K. Heat capacities were obtained from the slope of the enthalpy of binding as a function of temperature.

ΔC_p values for binding of alkyl and fluoroalkyl ligands are indistinguishable (for alkyls -9 ± 1 cal mol^{-1} K^{-1} per CH_2, for fluoroalkyls -10 ± 1 cal mol^{-1} K^{-1} per CF_2).

Crystallization of Protein−Ligand Complexes. We grew crystals of HCA II in conditions reported by McKenna and coworkers (1.15 M sodium citrate, 100 mM Tris hydrochloride, pH = 7.8) because under these conditions, crystals of the native protein diffract X-rays to ~1.0 Å resolution.[24] We performed soaking experiments using the ligands with short ($n = 0$ or 1) tails by transferring crystals from their mother liquor to a fresh drop that contained sodium citrate (1.32 M), Tris (100 mM), and ligand (20−450 μM). The ligands with $n \geq 2$ were insoluble in sodium citrate, which prohibited its use as the medium for soaking experiments.

We expected that the solubility of the ligands would be higher in solutions containing polyethylene glycol (PEG, 30−35%) than in sodium citrate, but were unable to grow crystals of HCA II in solutions of PEG. We chose, thus, a solution condition (PEG 1500, 20%; HEPES 100 mM) that was slightly higher in concentration of PEG than conditions reported previously to crystallize HCA II in the same polymorph as our crystals, and we transferred crystals of native HCA II, grown in sodium citrate, into drops

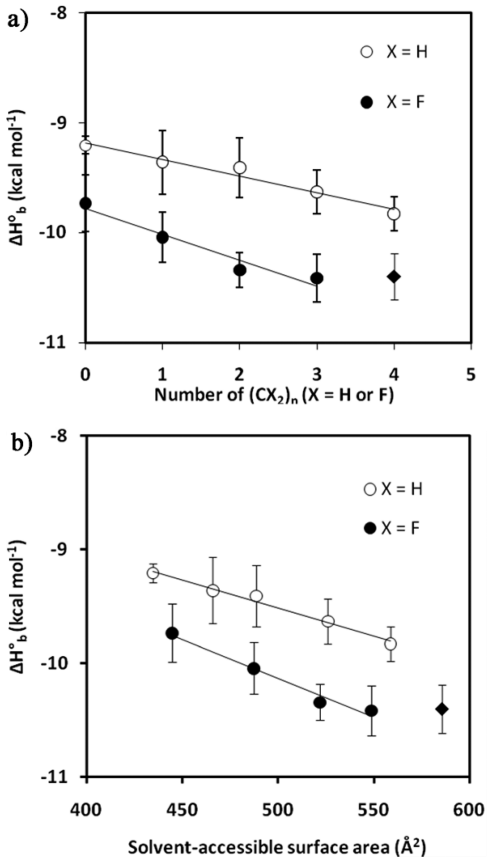


Figure 4. (a) Dependence of ΔH°_b for benzenesulfonamide ligands ($H_2NSO_2C_6H_4CONHCH_2(CX_2)_nCX_3$, X = H, F) containing alkyl (○) and fluoroalkyl (●) tails on the length of the chain. The slope of the regression line through the alkyl data is $\Delta\Delta H^\circ_b = -150 \pm 30$ cal mol^{-1} -CH_2-$^{-1}$ ($R^2 = 0.96$), and that for fluoroalkyl data is $\Delta\Delta H^\circ_b = -247 \pm 37$ cal mol^{-1} -CF_2-$^{-1}$ ($R^2 = 0.95$) ($N = 7$). (b) Dependence of ΔH°_b for ligands containing alkyl (○) and fluoroalkyl (●) tails on the solvent-accessible surface area. The slope of the regression line for the alkyl data is $\Delta\Delta H^\circ_b = -5 \pm 1$ cal mol^{-1} Å^{-2} ($R^2 = 0.99$), and that for the fluoroalkyl data is $\Delta\Delta H^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2} ($R^2 = 0.98$) ($N = 7$). The longest fluoroalkyl ligand (♦) is not included in the trendline, because it causes the conformational change of Gln136 of HCA II (see the text). The analysis of data with all fluoroalkyl ligands ($n = 0-4$) gives $\Delta\Delta H^\circ_b = -184$ cal mol^{-1} -CF_2-$^{-1}$ ($R^2 = 0.80$), $\Delta\Delta H^\circ_b = -5$ cal mol^{-1} A^{-2} ($R^2 = 0.81$).

containing PEG and saturated with ligands with longer ($n \geq 2$) tails. The strategy was successful in that the resulting crystals diffracted X-rays to 1.5−1.8 Å resolution, and the maps of electron density derived from molecular replacement indicated the presence of ligand (Figure S4 in Supporting Information). We refined the crystal structures of each of the ten HCA II−ligand complexes at high resolution (1.83−1.05 Å) data (Table 2, Table S1 in Supporting Information).

Crystallography Indicates That the Structure of the Protein Is Invariant in Most of the Protein−Ligand Complexes. To determine whether the thermodynamic trends in binding were the result of structural changes to the protein, we aligned the 10 structures and calculated the root-mean-squared deviation (rmsd) for all atoms of the proteins. The average value for rmsd for these structures was 0.091 Å, a result that indicated that the conformation of the protein did not depend on the identity of ligand bound in the active site.

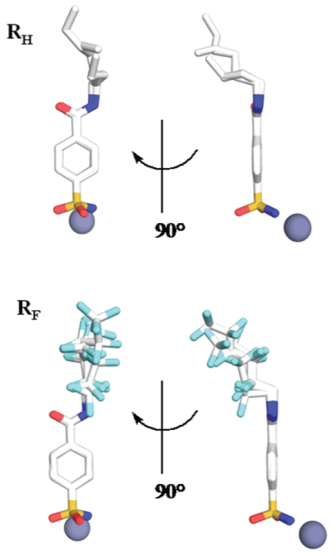


Figure 5. Alignment of the atoms of the ligands. Aligned structures for 10 ligands determined by X-ray crystallography appear as ball and stick representations. The Zn^{2+} cofactor appears as a silver sphere. Individual images of each ligand appear in Figure S6 in Supporting Information.

To verify our assumption that the geometry of the ligands in the active site of HCA II was conserved for each complex, we aligned the atoms of the 10 HCA II−ligand complexes and calculated the RMSDs for the atoms of the ligands and the Zn^{2+} ion (Figure 5). The 10 ligands had the same geometries of binding: the average value of rmsd for the alignment of the heavy atoms of the ligand, the Zn^{2+} ion, and the heavy atoms of residues His94, His96, His119, Phe131, Thr200 (chosen arbitrarily to allow the three-dimensional alignment) was 0.064 Å, the data that justified our assumption that the carboxybenzenesulfonamide group, the Zn−N bond, and the interaction between the carboxamide group and the protein-bound water at Thr200 were indistinguishable for the 10 complexes.

The structure of HCA II is invariant in 9 of 10 crystal structures of ligand complexes that we solved, the exception being the structure of HCA II in complex with the longest fluoroalkyl tail (X = F, $n = 4$, Figure 6). In this case, the side chain of Gln136, which is at the edge of the hydrophobic shelf, flips to contact the terminal CF_3 group of the tail. This conformational flip induces a gauche conformation of the Gln136 side chain (Figure 6B). The number of crystallographically defined water molecules within 4 Å of the $CONH_2$ group of Gln136, however, did not change: each of the two possible conformations of Gln136 showed four contacts with ordered molecules of water in the crystal structures (Figure S7 in Supporting Information). Crystallography of the protein−ligand complexes verified that the ligands bind to the active site in the same geometry, and that extension of the tails increased the putative surface of contact between the ligand and the protein.

For the R_H series, all of the contacts between the tail and the hydrophobic wall occurred between methylene and methyl groups of the tails and methyl groups of Val135 and Leu204, methylene groups of Pro202, and methine groups of Phe131. This observation validated our first-order analysis of hydrophobic effects in the linear trends in ΔH°_b and $-T\Delta S^\circ_b$ across the series. For the R_F series, similarly, crystallography validated our first-order analysis of linear trends in ΔH°_b and $-T\Delta S^\circ_b$ for the

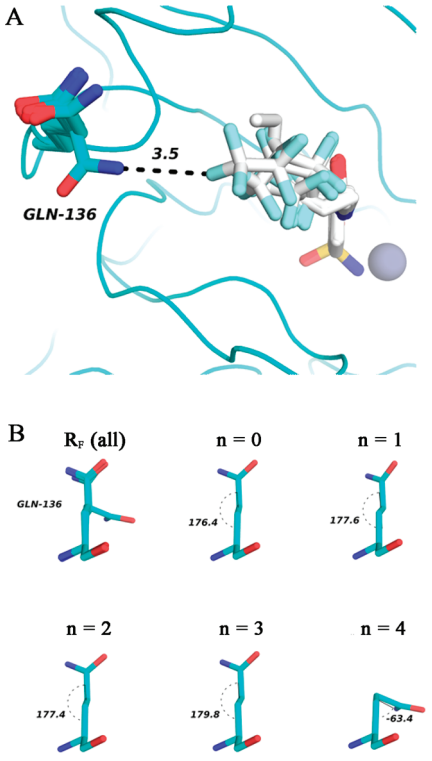


Figure 6. Crystal structures of fluorinated "greasy tails" complexed with HCA II. (A) Superimposition of all ligands and the Gln136 side chain. The distance between Gln136 and one fluorine atom of the ligand (R_F, $n = 4$) appears as a dashed line with its length labeled in angstroms. (B) Conformational analysis for Gln136. Superimposition of ligands (top left) and individual ligands. Gln136 possesses the anti conformation in cases of $n = 0-3$, while the gauche conformation is observed in the case of $n = 4$. Values of dihedral angles are labeled.

ligands $n = 0-3$ and indicated potential causes (i.e., polar contacts, steric interactions, changes in the hydration of protein, and change in the conformation of Gln136) for deviations from those trends for the ligand $n = 4$.

Crystallography provides no direct evidence concerning the hydration of the hydrophobic wall or the structure of the network of hydrogen bonds among molecules of water at the surface. It is interesting, however, to analyze both the regions of the active site in which crystallographic waters appear and those in which they do not. We analyzed three recent structures of HCA II that were solved by high resolution (~1.0 Å) X-ray diffraction and by neutron diffraction.[25] These structures show that more than 90% of the observable (crystallographic) waters are in indistinguishable positions. Interestingly, however, no crystallographically identifiable molecules of water appear within 3 Å of the hydrophobic wall. This observation provides no positive support for the idea of structured water near a hydrophobic surface in HCA II (although it also does not demonstrate the *absence* of such structure).

DISCUSSION

Negative values of $\Delta\Delta G^\circ_b$ ($\Delta G^\circ_{b,n+1} - \Delta G^\circ_{b,n} = \Delta G^\circ_{CX2,protein} - \Delta G^\circ_{CX2,solv}$; $\Delta G^\circ_{b,n+1}$ and $\Delta G^\circ_{b,n}$ are free energies of binding for greasy tails with the chain length of $n + 1$ and n, respectively. $\Delta G^\circ_{CX2,protein}$ and $\Delta G^\circ_{CX2,solv}$ are the free energy of binding of the CX_2 group to the protein and the free energy of hydration of the CX_2 group in the unbound state, respectively) could be the result of (i) favorable desolvation of the protein and/or the ligand (i.e., $-\Delta\Delta G^\circ_{CX2,solv} < 0$), (ii) favorable interactions (e.g., dispersion interactions) between the alkyl and fluoroalkyl tails of the ligand and the hydrophobic wall (i.e., $\Delta\Delta G^\circ_{CX2,protein} < 0$), (iii) an increase in conformational degrees of freedom of the ligand or protein on binding (i.e., $\Delta\Delta G^\circ_{CX2,protein} < 0$), or (iv) favorable solvation of the protein−ligand complex (i.e., $\Delta\Delta G^\circ_{CX2,protein} < 0$). The analysis we describe below indicates that desolvation of the greasy tails and of the hydrophobic wall determine the thermodynamics of binding in this system.

Dehydration of the Ligand Accounts for the Favorable Entropy of Binding. Our ITC data show that the entropy of binding becomes more favorable with larger alkyl and fluoroalkyl tails (for alkyls $-T\Delta\Delta S^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2}, for fluoroalkyls $-T\Delta\Delta S^\circ_b = -7 \pm 1$ cal mol^{-1} Å^{-2}, Table 1). There are at least four ways to explain the favorable incremental entropy of binding for R_H and R_F:

First, as Kauzmann and Tanford would have predicted, the desolvation of R_H and R_F could be entropically favorable (i.e., $T\Delta\Delta S^\circ_{CX2,solv} > 0$). We assume, based on the large number of calorimetric data for the transfer from octanol into aqueous phase, and for transfer from aqueous phase to vacuum of homologous alkyl-alcohols, and alkyl-amides[26,27], that the incremental entropy of desolvation of the R_H tails is favorable (−0.9 kcal mol^{-1} CH_2^{-1}).[28] This value, which is ~0.6 kcal mol^{-1} more favorable than $-T\Delta\Delta S^\circ_b$ for R_H and HCA II, indicates that dehydration of the tail makes the dominant *favorable* contribution to $-T\Delta\Delta S^\circ_b$. Our crystallographic data indicate that the CH_2 group of the ligand is only partially dehydrated upon binding to HCA II. Partial dehydration of the CH_2 group would give a value of the entropy of dehydration that is less favorable than the entropy of dehydration of the whole CH_2 group; thus, dehydration of the alkyl tail may account for the observed values of $-T\Delta\Delta S^\circ_b$. It is also plausible that the conformational flexibility of the tail in the unbound state is greater than that of the tail in the bound state: loss in conformational flexibility of the tail on binding to HCA would rationalize this difference between $-T\Delta\Delta S^\circ_{CH2,solv}$ and $-T\Delta\Delta S^\circ_b$. We are not aware of a calorimetric study of the solvation of aliphatic fluorocarbon compounds, presumably because their low solubility in water makes their solvation inherently difficult to study by calorimetry.

Second, Homans et al. showed in studies with major urinary protein (MUP) that increasing the number of methylene groups in ligands made unfavorable contributions to $-T\Delta\Delta S^\circ_b$:[29] for the series of *n*-alkyl alcohols, they report a value for $-T\Delta\Delta S^\circ_b$ = +412 cal mol^{-1} CH_2^{-1}. Values of $-T\Delta S^\circ_b$, however, include the difference in conformational mobility of the ligand between the unbound and bound states. Although the conformational mobility of alkyl chains on *n*-alkyl alcohols in the unbound state is likely to be similar to that of our alkyl tails in the *unbound* state, the conformational mobility of these groups in the *bound* state is determined by the structure of the binding site of the protein. The structures of the binding sites of HCA II and MUP are very different: the binding site of MUP is a narrow groove lined with Leu, Tyr, and Phe residues that completely surround its ligands, while that of HCA II is an open, conical cleft in which bound ligands retain 30−50% of their solvent-accessible surface area. It is difficult to estimate the conformational mobility of ligands in the bound state. It is plausible, however, that alkyl tails in the active site of HCA II retain more conformational mobility than

do alkyl alcohols bound in the active site of MUP.[30] Such a difference in conformational mobility would rationalize, at least in part, the difference in values of $-T\Delta\Delta S^{\circ}_{b}$ between alkyl tails binding to HCA II and to MUP.

The third contribution to the overall $-T\Delta\Delta S^{\circ}_{b}$ could derive from changes from conformational degrees of freedom for amino acids in the active site upon binding. Our crystallographic data indicate that this contribution is unlikely to be important, because all ligands except the longest fluorinated one (X = F, $n = 4$) do not change the conformation of side chains of amino acids in the proximity of ligands (the conformation of all amino acids of HCA II remain the same in complexes with R_H and R_F ligands, the only exception being Gln136 in the case of R_F, $n = 4$).

Fourth, there is no indication from high-resolution X-ray and neutron diffraction studies of native HCA II that molecules of water are localized near the hydrophobic wall, although configurational restriction of molecules of water that are not observable by crystallography could, at least in principle, make favorable contributions $T\Delta\Delta S^{\circ}_{CX2,protein}$.

Our results, an increasing favorable entropy with increasing area of alkyl and fluoroalkyl tails, are compatible with Kauzmann−Tanford's hypothesis for the origin of the hydrophobic effect: the burial of hydrophobic surface area, in both cases, is entropically favorable. The entropic contribution to $\Delta\Delta G^{\circ}_{b}$ appears to be dominated by the entropically favorable dehydration of the greasy tail. Although it is plausible that association with HCA II restricts the conformational flexibility of the alkyl tails, this restriction could be less unfavorable for the association of greasy tails with HCA II than it is for that of *n*-alkyl alcohols with MUP because of the differences in the structures of the active sites of the two proteins. The favorable contribution of dehydration of the hydrophobic surface of ligands to the overall free energy of binding is also illustrated by negative values of ΔC_p. The heat capacity of desolvation of a CH_2 group is around -13 cal mol^{-1} K^{-1}.[28,31] This value is slightly larger than what we have observed for binding of R_H ligands to HCA II (-9 cal mol^{-1} K^{-1} per CH_2). The observed difference presumably results from a partial dehydration of R_H ligands in the bound state; this proposal is well supported by X-ray crystallographic data for HCA II−R_H complexes.

Is the Favorable Enthalpy of Binding a Result of Dispersion Interactions or Dehydration of the Hydrophobic Wall of HCA II? Our ITC results show that the enthalpy of binding to HCA II becomes more favorable with larger alkyl and fluoroalkyl tails (for alkyls $\Delta\Delta H^{\circ}_{b} = -5 \pm 1$ cal mol^{-1} Å^{-2}, for fluoroalkyls $\Delta\Delta H^{\circ}_{b} = -7 \pm 1$ cal mol^{-1} Å^{-2}, Table 1). The values we measure for $\Delta\Delta H^{\circ}_{b}$ for the alkyl tails are quite different from those reported by Homans for the aliphatic alcohols binding to MUP ($\Delta\Delta H^{\circ}_{b} = -150 \pm 30$ cal mol^{-1} for HCA II with alkyl tails; $\Delta\Delta H^{\circ}_{b} = \sim -1350$ cal mol^{-1} for MUP with aliphatic alcohols). Homans et al., however, did not compare thermodynamics of binding for aliphatic alcohols and fluorinated aliphatic alcohols, and they did not correlate the thermodynamic parameters of binding with solvent-accessible surface area of ligands.

Binding of primary alcohols ($CH_3(CH_2)_nOH$, $n = 4-9$) to MUP appears to conflict with Kauzmann−Tanford's view on the hydrophobic effect: Homans et al. found that enthalpy of binding becomes more favorable and the entropy of binding becomes less favorable with the increasing chain length of the ligands. They rationalized that trend by invoking dispersion interactions, which contribute favorably to $\Delta\Delta H^{\circ}_{b}$, between the alkyl groups of the alcohols and the active site of the MUP.[29]

Previous studies of the dehydration of aliphatic compounds (and of other model compounds) suggest that the dehydration of aliphatic surface area is enthalpically unfavorable ($-\Delta\Delta H^{\circ}_{CH2,solv} = \sim 0.7$ kcal mol^{-1}).[28,29] An unfavorable value for partial dehydration of CH_2 group $-\Delta\Delta H^{\circ}_{CH2,solv}$ paired with an overall favorable value for $\Delta\Delta H^{\circ}_{b}$ requires the sum of the remaining contribution to be favorable (for R_H $\Delta\Delta H^{\circ}_{CH2,protein} = \sim -0.6$ kcal mol^{-1}). This requirement, in turn, suggests at least three possibilities to obtain the overall favorable value of $\Delta\Delta H^{\circ}_{b}$ that we observe experimentally: (i) noncovalent interactions between the tail moieties and the hydrophobic wall make $\Delta\Delta H^{\circ}_{CX2,protein} < 0$, (ii) dehydration of the hydrophobic wall of HCA II is enthalpically favorable ($\Delta\Delta H^{\circ}_{CX2,protein} < 0$), or (iii) hydration of protein−ligand complex is enthalpically favorable ($\Delta\Delta H^{\circ}_{CX2,protein} < 0$).

We propose, from the comparison of thermodynamics of binding for R_H and R_F, and from crystallography, that partial dehydration of the hydrophobic wall of HCA II upon ligand binding releases loosely bound molecules of water that are involved in hydrogen bonding interactions that are weaker than those in bulk water. This release results in the formation of stronger hydrogen bonds among waters in the bulk, an enthalpically favorable process.

Dispersion Interactions Presumably Do Not Account for the Enthalpic Contributions to Binding for Alkyl and Fluoroalkyl Tails. Dispersion interactions are considerably weaker in fluorocarbon liquids than they are in hydrocarbon liquids; these interactions are determined by a low polarizability of fluorocarbons inferred from correlating the index of refraction with electronic polarizability. The difference in polarizabilities would predict that the dispersion interactions between the alkyl tails (R_H) and the hydrophobic wall would be more enthalpically favorable than the analogous interactions of fluoroalkyl tails (R_F), a difference that is *not* observed in our measurements (for alkyls $\Delta\Delta H^{\circ}_{b} = -5 \pm 1$ cal mol^{-1} Å^{-2}, for fluoroalkyls $\Delta\Delta H^{\circ}_{b} = -7 \pm 1$ cal mol^{-1} Å^{-2}). We conclude, therefore, that the difference in dispersion interactions between fluorocarbons and hydrocarbons does *not* account for the value of $\Delta\Delta H^{\circ}_{b}$ for the two series of ligands.

Favorable Enthalpic Contribution Results from Nonoptimal Hydration of the Hydrophobic Wall of HCA II. It is plausible that the hydrophobic wall of HCA II is solvated with molecules of water that, because of the structure of the active site, are connected by weaker hydrogen bonds than are water molecules of the bulk. The displacement of such near-surface molecules of water (or, alternatively, of molecules of water disordered in the conical active site) by the ligand would be enthalpically favorable regardless of the structure and chemical composition of the ligand that is doing the displacement. In addition, larger ligands should displace more of these water molecules from the hydrophobic wall, and the favorable enthalpic contribution to binding would be proportional to the surface (and also volume) of the ligand.

The absence of observable, crystallographically bound water molecules could indicate, but does not prove, that hydration of the hydrophobic wall does not occur with the ordering of water molecules that Kauzmann and Tanford suggested as being the origin of an entropy-dominated hydrophobic effect.

Crystal structures support thermodynamic data in suggesting that the molecular driving forces for binding of R_H and R_F to HCA II are indistinguishable. The change in conformation for Gln136 on binding correlates with a deviation in the trends of enthalpy and entropy of binding for the ligand with the longest

fluoroalkyl tail compared to the other members of the series. This result suggests that the binding of alkyl and fluoroalkyl tails (except for X = F, n = 4) to HCA II is determined by similar molecular interactions (on an area- and volume-corrected basis) at the hydrophobic wall of HCA II.

The differences in polarizability of R_F and R_H could plausibly make the enthalpy of binding of R_H slightly more favorable than R_F. In contrast to other alkyls (n = 0−4) and fluoroalkyls (n = 0−3), crystallographic analysis of the longest fluorinated sulfonamide (n = 4) shows a major structural difference compared to other tails. Gln136 possesses the gauche conformation in the case of this fluoroalkyl (n = 4), while in the presence of all other ligands, the anti conformation of Gln136 is observed (Figure 6). Overall, we believe that this difference contributes to the lower value of $\Delta H°_b$ for the longest fluorinated ligand (n = 4) and explains the inconsistency (~0.4 kcal mol^{-1} less than what would be predicted using the least-squares linear regression fit obtained from n = 0−3) in its value of $\Delta H°_b$ with those of the other fluoroalkyl ligands (Figure 4).

CONCLUSIONS

R_F and R_H Have Indistinguishable Hydrophobicities. ITC demonstrates that the increasingly favorable binding of hydrophobic tails (R_H and R_F) to HCA II with increasing length of R_H or R_F chain results from favorable contributions from *both* enthalpy and entropy. These thermodynamics also show that alkyl and fluoroalkyl tails have indistinguishable thermodynamic signatures after correction for the differences in solvent-accessible surface area. This similarity indicates that the molecular basis for increasing affinity with increasing surface area of the tail group is similar for both. Our data are consistent with the hypothesis that the hydrophobic effect, in this case, results from the exclusion of water molecules from the contact region between the hydrophobic surface of the ligand, the hydrophobic wall of HCA II, and from the active site cavity, and *not* from different physical properties of R_H and R_F. On the basis of the thermodynamics of binding, we conclude that hydrocarbons and fluorocarbons are virtually indistinguishable when interacting with the hydrophobic surface of HCA II. Apparent differences between fluorocarbons and hydrocarbons in this study result primarily from differences in their hydrophobic surface area and not from differences in dispersion interactions.

The Origin of the Hydrophobic Effect Is Release of Water Molecules from the Protein Binding Pocket and from the Surface of the Ligand. Most of the favorable free energy of binding is gained from interactions of water with nonpolar surfaces. In this particular case, dehydration of the ligand (which presents a convex hydrophobic surface area) results in a favorable change in entropy of binding, and dehydration of the hydrophobic wall of HCA II (which is a concave hydrophobic surface area) results in a favorable change in enthalpy of binding. Theoretical studies by Rossky, Berne, Abel, Friesner, and others have, time and again, suggested that the free energy of water molecules that solvate hydrophobic surfaces depend on the shape of the surface.[24,26,27,32−36] Our experimental observations are compatible with this view and indicate that favorable contributions to $\Delta G°_b$ can arise simultaneously from the entropy of dehydration of the greasy tail of the ligand and from the enthalpy of dehydration of the hydrophobic wall of the protein. Whether or not there are general relationships between the curvature of hydrophobic surfaces and the thermodynamic parameters for the dehydration of them remains to be determined, although understanding these relationships may be centrally important for designing ligands that bind tightly to proteins.

Rational Ligand Design May Require the Explicit Consideration of Water. A better understanding of the thermodynamics of water interacting with nonpolar surfaces in various biological systems is required for generating predictive algorithms in rational ligand design. Knowledge of the hydration of the active sites of medicinally relevant proteins might be useful in designing high affinity ligands. Releasing water molecules of an active site that is partially (nonoptimally) hydrated by a ligand would provide an enthalpically favorable component to the free energy of binding. In this respect, ligands with larger solvent-accessible surface area (and also larger volume) would release more water molecules upon binding than ligands with smaller SASA (or smaller volume). Thus, designing new inhibitors would involve an approach where ligands with larger SASA (and perhaps also volume) would be better targets than those with smaller SASA and volume. Incorporation of fluorine instead of hydrogen is one way to achieve larger SASA, but there are other functional groups (e.g., CH_3 instead of H) that could provide a similar effect.

Our results demonstrate that water must be considered when designing ligands to bind tightly to proteins. Structural characterization of proteins by crystallography describes only part of the influence of water on molecular recognition. Nuclear magnetic resonance may fill in some of the details,[37,38] but its application to proteins much larger than about 25 kDa has not yet provided detailed information about locations of water molecules. Rational ligand designers need theoretical approaches that predict accurately the structure of water in and around the active sites of proteins. Toward that aim, we believe it is important to provide the theoretical community with the integrated structural and thermodynamic characterization of well-defined model systems of ligands and proteins against which to validate theoretical models. Carbonic anhydrase and arylsulfonamides are particularly well suited for this purpose.

ASSOCIATED CONTENT

Supporting Information. Background about the hydrophobic effect, additional thermodynamic analysis, and crystallographic data. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
gwhitesides@gmwgroup.harvard.edu

ACKNOWLEDGMENT

This work was supported by the National Institutes of Health (GM051559, GM030367) and a predoctoral fellowship from Eli Lilly (K.A.M.). Crystallography data for this study were measured at beamlines X25 and X29 of the National Synchrotron Light Source. Financial support comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the US Department of Energy, and from the National Center for Research Resources of the National Institutes of Health grant number P41RR012408.

REFERENCES

(1) Kauzmann, W. *Adv. Protein Chem.* **1959**, *14*, 1–63.
(2) Tanford, C. *Science* **1978**, *200*, 1012–1018.
(3) Tanford, C. *J. Mol. Biol.* **1972**, *67*, 59–74.
(4) Ball, P. *Chem. Rev.* **2008**, *108*, 74–108.
(5) Southall, N. T.; Dill, K. A.; Haymet, A. D. J. *J. Phys. Chem. B* **2002**, *106*, 521–533.
(6) Blokzijl, W.; Engberts, J. B. F. N. *Angew. Chem., Int. Ed.* **1993**, *32*, 1545–1579.
(7) Dalvi, V. H.; Rossky, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 13603–13607.
(8) Muller, K.; Faeh, C.; Diederich, F. *Science* **2007**, *317*, 1881–1886.
(9) Montclare, J. K.; Son, S.; Clark, G. A.; Kumar, K.; Tirrell, D. A. *ChemBioChem* **2009**, *10*, 84–86.
(10) Yoder, N. C.; Kumar, K. *Chem. Soc. Rev.* **2002**, *31*, 335–341.
(11) Gao, J.; Qiao, S.; Whitesides, G. M. *J. Med. Chem.* **1995**, *38*, 2292–2301.
(12) Lee, A.; Mirica, K. A.; Whitesides, G. M. *J. Phys. Chem. B* **2011**, *115*, 1199–1210.
(13) Dunitz, J. D. *ChemBioChem* **2004**, *5*, 614–621.
(14) Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. *Chem. Rev.* **2008**, *108*, 946–1051.
(15) Krishnamurthy, V.; Bohall, B.; Kim, C. Y.; Moustakas, D.; Christianson, D.; Whitesides, G. *Chem. Asian J.* **2007**, *2*, 94–105.
(16) Böhm, H.-J.; Banner, D.; Bendels, S.; Kansy, M.; Kuhn, B.; Müller, K.; Obst-Sander, U.; Stahl, M. *ChemBioChem* **2004**, *5*, 637–643.
(17) Kwon, O.-H.; Yoo, T. H.; Othon, C. M.; Van Deventer, J. A.; Tirrell, D. A.; Zewail, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 17101–17106.
(18) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L.-N. *Anal. Biochem.* **1989**, *179*, 131–137.
(19) Boriack-Sjodin, P. A.; Zeitlin, S.; Christianson, D. W.; Chen, H.-H.; Crenshaw, L.; Gross, S.; Dantanarayana, A.; Delgado, P.; May, J. A.; Dean, T. *Protein Sci.* **1998**, *7*, 2483–2489.
(20) Nair, S. K.; Calderone, T. L.; Christianson, D. W.; Fierke, C. A. *J. Biol. Chem.* **1991**, *266*, 17320–17325.
(21) Burton, R. E.; Oas, T. G.; Fierke, C. A.; Hunt, J. A. *Protein Sci.* **2000**, *9*, 776–785.
(22) Khalifah, R. G.; Strader, D. J.; Bryant, S. H.; Gibson, S. M. *Biochemistry* **1977**, *16*, 2241–2247.
(23) Turnbull, W. B.; Daranas, A. H. *J. Am. Chem. Soc.* **2003**, *125*, 14859–14866.
(24) Berne, B. J.; Weeks, J. D.; Zhou, R. *Annu. Rev. Phys. Chem.* **2009**, *60*, 85–103.
(25) Pratt, L. R.; Pohorille, A. *Chem. Rev.* **2002**, *102*, 2671–2692.
(26) Wang, L.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 1326–1330.
(27) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 808–813.
(28) Plyasunov, A. V.; Shock, E. L. *Geochim. Cosmochim. Acta* **2000**, *64*, 439–468.
(29) Malham, R.; Johnstone, S.; Bingham, R. J.; Barratt, E.; Phillips, S. E. V.; Laughton, C. A.; Homans, S. W. *J. Am. Chem. Soc.* **2005**, *127*, 17061–17067.
(30) Krishnamurthy, V. M.; Bohall, B. R.; Semetey, V.; Whitesides, G. M. *J. Am. Chem. Soc.* **2006**, *128*, 5802–5812.
(31) Syme, N. R.; Dennis, C.; Phillips, S. E. V.; Homans, S. W. *ChemBioChem* **2007**, *8*, 1509–1511.
(32) Cheng, Y.-K.; Rossky, P. J. *Nature* **1998**, *392*, 696–699.
(33) Giovambattista, N.; Lopez, C. F.; Rossky, P. J.; Debenedetti, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2274–2279.
(34) Liu, P.; Huang, X.; Zhou, R.; Berne, B. J. *Nature* **2005**, *437*, 159–162.
(35) Zhou, R.; Huang, X.; Margulis, C. J.; Berne, B. J. *Science* **2004**, *305*, 1605–1609.
(36) Abel, R.; Salam, N. K.; Shelley, J.; Farid, R.; Friesner, R. A.; Sherman, W. *ChemMedChem* **2011**, *6*, 1049–1066.
(37) Wider, G.; Wuthrich, K. *Curr. Opin. Struct. Biol.* **1999**, *9*, 594–601.
(38) Wuthrich, K.; Otting, G.; Liepinsh, E. *Faraday Discuss.* **1992**, *93*, 35–45.