



Published in final edited form as:

J Am Chem Soc. 2007 February 7; 129(5): 1312–1320. doi:10.1021/ja066780e.

Dependence of Effective Molarity on Linker Length for an Intramolecular Protein-Ligand System

Vijay M. Krishnamurthy, Vincent Semetey, Paul J. Bracher, Nan Shen, and George M. Whitesides*

Department of Chemistry and Chemical Biology, Harvard University 12 Oxford Street, Cambridge, MA 02138

Abstract

This paper reports dissociation constants and “effective molarities” (M_{eff}) for the intramolecular binding of a ligand covalently attached to the surface of a protein by oligo(ethylene glycol) (EG_n) linkers of different lengths ($n = 0, 2, 5, 10,$ and 20), and compares these experimental values with theoretical estimates from polymer theory. As expected, the value of M_{eff} is lowest when the linker is too short ($n = 0$) to allow the ligand to bind noncovalently at the active site of the protein without strain, is highest when the linker ($n = 2$) is the optimal length to allow such binding to occur, and decreases monotonically as the length increases past this optimal value (but, only by a factor of approximately eight from $n = 2$ to $n = 20$). These experimental results are not compatible with a model in which the single bonds of the linker are *completely* restricted when the ligand has bound non-covalently to the active site of the protein, but are quantitatively compatible with a model that treats the linker as a random-coil polymer. Calorimetry revealed that enthalpic interactions between the linker and the protein are not important in determining the thermodynamics of the system. Taken together, these results suggest that the manifestation of the linker in the thermodynamics of binding is exclusively entropic. The values of M_{eff} are, theoretically, intrinsic properties of the EG_n linkers, and can be used to predict the avidities of multivalent ligands with these linkers for multivalent proteins. The weak dependence of M_{eff} on linker length suggests that multivalent ligands containing flexible linkers that are longer than the spacing between the binding sites of a multivalent protein will be effective in binding, and that the use of flexible linkers with length somewhat greater than the optimal distance between binding sites is a justifiable strategy for use in the design of multivalent ligands.

Introduction

The primary motivation for this paper was to determine the influence of the “linker”—the structural element that connects the binding moieties in a multivalent ligand—on the binding of a multivalent ligand to a multivalent protein (Figure 1E). As our model system, we adopted a ligand covalently tethered to the surface of a protein by oligo(ethylene glycol) linkers (EG_n) (Figure 1D), because this design allows us to interrogate the system without complications from separate binding events of the two ends of the ligand; complications that

* Author to whom correspondence should be addressed. Telephone: (617) 495–9430 Fax: (617) 495–9857 E-mail: gwhitesides@gmwhgroup.harvard.edu.

Supporting Information Available

Synthetic procedures and characterization, assay protocols, SDS-PAGE and assessment of activity of HCA**, mass spectra and SE-HPLC chromatograms of modified HCA** proteins, fluorescence titration curves for the binding of DNSA to control HCA** proteins, fluorescence calibration plot for determining purity of HCA**-SSEG_nSA proteins, fluorescence titration curves for determination of K_d of SA-OMe and Ethox to HCA**-SCH₂CO₂⁻, and ITC thermograms for the binding of Ethox and SA-OMe to HCA**-SCH₂CO₂⁻ and to HCA**-SSEG₁₀SA. This material is available free of charge via the Internet at <http://pubs.acs.org>.

do arise when both ends are free. We report dissociation constants (and effective molarities, see below) for the intramolecular binding of the tethered ligand to the active site of the protein as a function of the linker length (n), and discuss the data in the context of four possible models (Figure 2). We find that only one of these models (Figure 2B) is consistent with the following results: the linker plays an exclusively entropic role in the thermodynamics of the system, and it has significant conformational mobility even after the ligand is bound at the active site of the protein. The results are quantitatively well-explained by a model that describes the linker as a random-coil polymer (Figure 3); this model is one originally proposed by Lees and coworkers for a related problem, and estimates the probability of intramolecular binding as the probability that the ends of the linker are separated by a distance (d) equal to the separation of the sites of covalent attachment and non-covalent binding of the ligand to the protein (Figure 1D).¹

Effective molarity (M_{eff}) is an empirical term that is commonly used to relate the kinetics and equilibria of intramolecular and intermolecular reactions (both covalent and non-covalent) in chemistry (eq 1).^{2, 3} $K_{\text{d}}^{\text{inter}}$, which has units of concentration (e.g., molarity), is the dissociation constant for an intermolecular reaction (Figure 1A) and $K_{\text{d}}^{\text{intra}}$, which is dimensionless, is the dissociation constant for an analogous intramolecular reaction (Figure 1B). M_{eff} depends on the length and flexibility of the “linker” between the two reactive groups.^{2, 3}

$$M_{\text{eff}} = K_{\text{d}}^{\text{inter}} / K_{\text{d}}^{\text{intra}} \quad (1)$$

Effective concentration (C_{eff} ; units of concentration) is a theoretical parameter that allows the estimate of the rate and/or equilibrium for ring closure for intramolecular reactions (Figure 1B) by assuming that the linker between the two reactive groups behaves as a random-coil polymer (Figure 3A).³⁻⁷ It is essentially the probability that the two reactive groups (the two ends of the polymer) will be within an infinitesimal distance of one another (eq 2); N_A is Avogadro's number and $\langle r^2 \rangle^{1/2}$ is the root-mean-squared distance (in dm) between the ends of the polymer.

$$C_{\text{eff}} = \frac{1}{N_A (\langle r^2 \rangle^{1/2})^3} \left(\frac{3}{2\pi} \right)^{\frac{3}{2}} \quad (2)$$

The parameter $\langle r^2 \rangle^{1/2}$ can be estimated from polymer chemistry assuming a three-dimensional random flight (eq 3) where n is the number of segments in the chain and C is proportional to the distance (l) between two segments of the chain (this value is equal to the bond length for a freely-jointed polymer) and is characteristic of a given chain structure.^{5, 89} The proportionality constant C/l gives a measure of the stiffness of the chain and takes into account effects such as bond angles and rotational barriers; it typically varies from 1.5 for a very flexible chain to 5.5 for a very stiff one.^{5, 6}

$$\langle r^2 \rangle^{1/2} = (C/l) l \sqrt{n} = C \sqrt{n} \quad (3)$$

Winnik and Mandolini have reviewed the literature for ring closing (macrocyclization) reactions in small molecule systems, and concluded that C_{eff} and M_{eff} are closely correlated for many of these reactions.^{3, 10}

In many situations (and, in particular, in protein-ligand binding) we are not concerned with the probability that the two ends of the linker are an infinitesimal distance apart. Rather, we require the probability that the two ends are a distance (d) apart, where d is set by the system (Figures 1D,E and 3B). Lees and coworkers derived eq 4 for this case, where $C_{\text{eff}}(0)$ is C_{eff} defined in eq 2.¹ The parameter p arises because of the presence of the protein: the ligand cannot occupy the same space as the protein and because it is excluded from this volume, its C_{eff} increases (a so-called “excluded volume” effect). Lees and coworkers proposed a value of 2 for this term;

this value assumes that the ligand has only a hemisphere of free access when constrained by the protein (Figure 1E), as compared to a sphere when it is free in solution.¹

$$C_{\text{eff}}(d) = pC_{\text{eff}}(0) \exp\left(-\frac{3}{2(\langle r^2 \rangle^{1/2})^2} d^2\right) \quad (4)$$

We have defined multivalency as multiple interactions (often of the same kind) between two different species.^{4, 11} Theoretical approaches to multivalency have assumed a step-wise pathway for dissociation, in which the first step is intramolecular (Figure 1E). The value of the dissociation constant ($2K_d^{\text{intra*}}$) for this step has been challenging to estimate theoretically.⁴ Without an estimate for this parameter, we cannot *predict* multivalent avidities from component monovalent affinities (although the work by Lees, Reinhoudt, and others represents a significant advance toward this capability)^{1, 6, 7}. The study presented here clarifies and simplifies this problem because we obtain empirical estimates for the dissociation constants for intramolecular protein-ligand binding that are applicable to the thermodynamics of the intramolecular step in multivalent binding (Figure 1D,E).

We have previously argued that flexible oligomers should not function effectively as linkers in multivalent ligands because of the severe loss in conformational entropy of the linker ($T\Delta S^\circ \sim RT \ln 3 \sim 0.7 \text{ kcal mol}^{-1}$ per freely rotating single bond of the linker) when it is bound at both ends.^{4, 11, 12} Flexible linkers (e.g., oligo(ethylene glycol)) have, however, been used with success in multivalent ligands.^{4, 7, 11, 13} Determining how flexible linkers could work in multivalent ligands, when this simple theoretical model argues that they should not, was a key motivation for this paper.

Experimental Design

We selected the combination of human carbonic anhydrase II (HCA, EC 4.2.1.1) and *p*-substituted benzenesulfonamides as our model system, because it is the simplest one that we know for studying protein-ligand interactions: the conserved mode of binding of sulfonamides to HCA has been well-established by X-ray crystallography, there are a number of well-characterized assays to use in following binding, and there are a number of commercially available, high-affinity arylsulfonamides to use for competition experiments.¹⁴⁻¹⁷ Further, HCA is easy to over-express in, and purify from, *E. coli* culture in high yield; this ease allows the generation of HCA mutants with a chemical “handles” (i.e., reactive sites) to which to couple the ligands.¹⁸⁻²⁰

An examination of crystal structures of HCA complexed with *p*-substituted benzenesulfonamides containing oligo(ethylene glycol) linkers (PDB: 1CNY, 1CNX, 1CNW)²¹ established that Lys-133 was spatially close to the terminus of the ligand but outside of the conical cleft of the enzyme (Figure 4). Mutating this residue to Cys would generate a chemical handle for thiol-selective coupling.^{18, 20} HCA has an endogenous Cys at position 206. In order to preclude side reaction at this site, we mutated it to Ser to generate a double mutant (Cys-206→Ser, Lys-133→Cys), which we refer to as HCA** in the remainder of this paper. Krebs and Fierke demonstrated that the C206S mutant of HCA is active,²² and Mårtensson et al. determined that this mutant is as stable as wild-type HCA.²⁰

We decided to use *p*-substituted benzenesulfonamides with a conserved alkyl chain (*p*-H₂NSO₂C₆H₄CH₂NHCO(CH₂)₄CO-) as our model ligands because benzenesulfonamides bind CA in an invariant orientation, and the alkyl chain should saturate the hydrophobic surface of the conical cleft of CA and orient the linker into solution.^{14, 21, 23-25} We selected oligo(ethylene glycol) (EG_{*n*}) as the linker to tether the ligand to HCA** because it is one of the most common linkers in multivalent ligands.^{4, 11, 13} We wanted to understand how these

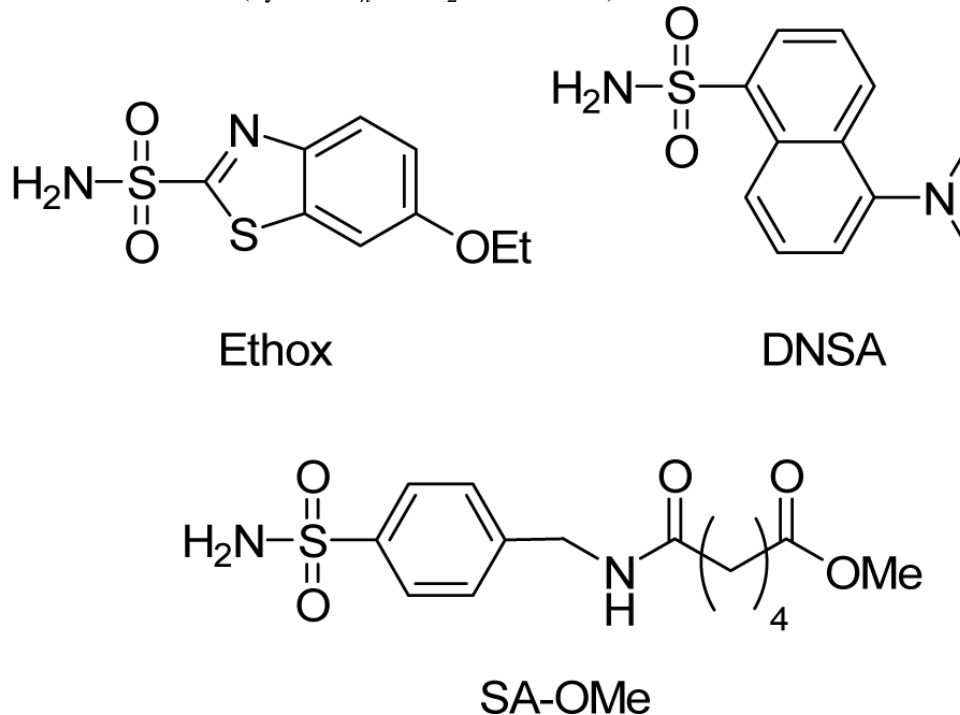
flexible oligomers could be effective as linkers in multivalent ligands. Oligo(ethylene glycol) has been shown to be resistant to the non-specific adsorption of proteins when displayed on surfaces;²⁶ this property should minimize complications from non-specific adsorption of the linker to the surface of the protein (Figure 2C,D).

We chose thiolate-disulfide interchange as a means of coupling the ligands to HCA** because this reaction is completely selective for thiols (vs. the various other chemical functionalities in proteins).²⁷

Results and Discussion

Synthesis of Benzenesulfonamides Containing Activated Disulfides for Covalent Tethering to HCA**

We synthesized disulfide-activated *p*-substituted benzenesulfonamides with oligo(ethylene glycol) linkers of different lengths (Pyr-SSEG_{*n*}SA; *n* = 0, 2, 5, 10, and 20; Table 1) using conventional amide-bond coupling reactions (see Supporting Experimental Procedures for details). We used ligands containing disulfides activated with 2-pyridylthiol because the aromatic thiol is an excellent leaving group; thus, incubating the activated disulfide with HCA** should generate only the desired products (HCA**-SSEG_{*n*}SA).²⁸ We also synthesized two control molecules (Pyr-SSEG_{*n*}CONH₂ and SA-OMe) for our studies.



Design and Purification of HCA II with a Surface-Accessible Cysteine Residue

We generated a vector encoding HCA** by conventional site-directed mutagenesis of the plasmid encoding the C206S mutant of HCA (pC206S), and inserted this mutated vector into the backbone of the plasmid encoding wild-type HCA (pACA; both pC206S and pACA were kind gifts of Professor Carol Fierke, University of Michigan).^{22, 29} We over-expressed HCA** in *E. coli* (BL21(DE3)) as described by Fierke and co-workers¹⁸ and purified it as described by Khalifah et al.¹⁹ We confirmed the purity of the enzyme by SDS-PAGE and its

activity (~100%) by measuring the amount of protein that could bind ethoxzolamide (Ethox), a specific inhibitor of CA (Figure S.1).¹⁷

Synthesis of Modified HCA** Proteins and Characterization of Conjugates

We treated HCA** with ~5 equiv of one of five disulfide-activated ligands (Pyr-SSEG_nSA, *n* = 0, 2, 5, 10, and 20) in Tris-sulfate buffer pH 8.0, and purified the conjugated proteins (HCA**-SSEG_nSA, *n* = 0, 2, 5, 10, and 20) by exhaustive dialysis. We also synthesized two control proteins (HCA**-SSEG₂CONH₂ and HCA**-SCH₂CO₂⁻) by allowing HCA** to react with ~10 equiv of Pyr-SSEG₂CONH₂ or iodoacetate (see the Experimental Section for details). The use of HCA**-SSEG₂CONH₂ established the influence of the linker minus the benzenesulfonamide moiety on binding. We used HCA**-SCH₂CO₂⁻ for control studies (see below).

We characterized all of the modified proteins by ESI-MS and, in all cases, observed peaks corresponding to the combined masses of the protein and the reacted ligand (Table S.1). We examined the HCA**-SSEG_nSA proteins by size-exclusion high-performance liquid chromatography (SE-HPLC) in order to determine whether a non-covalent dimer was present in any of the samples (Figure S.2). For HCA**-SSEG_nSA with *n* = 2, 5, 10, and 20, we saw no evidence for a dimer at a concentration of protein of 20 μM. For HCA**-SSEG₀SA (i.e., *n* = 0), we observed a peak (~14% of total protein) corresponding in size to a dimer; this peak quantitatively dissociated in the presence of a large excess of Ethox (a high-affinity arylsulfonamide; see next section), and thus represents a non-covalent dimer that is stable on the time-scale of the analysis (~25 min). As a stable, non-covalent dimer represents a minor contaminant for HCA**-SSEG₀SA, and is not present for any of the other HCA**-SSEG_nSA proteins, it should not complicate our analysis of dissociation constants (see next section).

We measured the affinity of a well-characterized fluorescent arylsulfonamide—dansylamide (DNSA)—for HCA**, HCA**-SCH₂CO₂⁻, and HCA**-SSEG₂CONH₂ (Figure S.3).¹⁶ The values of *K*_d for the three proteins were the same (Table 2); this observation suggests that the EG_n linker itself has no effect on the binding of arylsulfonamides to the modified HCA** proteins, and that we can use HCA**-SCH₂CO₂⁻ as a surrogate for HCA** for control studies because it is easier to handle (e.g., there is no need for added reductant and no possibility of dimer formation without the reductant).

In order to determine the purity of the sulfonamide-conjugate proteins (HCA**-SSEG_nSA), we examined the fluorescence of the protein (excitation wavelength = 290 nm, emission wavelength = 460 nm) when treated with 5 μM of DNSA; this concentration is ~16-fold greater than the *K*_d of DNSA or HCA** (Table 2). The only fluorescent species under these conditions is the HCA**-DNSA complex. We assumed that any fluorescence observed was due to residual, unmodified HCA**, because all of the modified proteins have dissociation constants for DNSA too high to bind it at the concentration of DNSA used in the experiment (see below). We constructed a standard curve for the fluorescence of 5 μM of DNSA as a function of the concentration of HCA** and used it to interpolate the concentration of unmodified HCA** remaining in the coupling reactions (Figure S.4). Using these values and the concentrations of total HCA** protein from UV spectroscopy,³¹ we obtained purities of 90–95% for HCA**-SSEG_nSA where *n* = 0, 2, 10, and 20 and ~80% for *n* = 5 (see Supporting Experimental Procedures for details). This amount of contamination will not affect values of dissociation constants for these proteins (see next section).

For calorimetric studies where completely homogeneous protein was required, we purified the sulfonamide-conjugated protein (HCA**-SSEG₁₀SA) over a column of benzenesulfonamide-conjugated agarose (see Supporting Experimental Procedures for details). Purities were ≥99% (assessed using the standard curve) after this procedure.

Measurement of Dissociation Constants of Ethoxzolamide (Ethox)

Ethox is a high-affinity ligand for CA that quenches ~70% of the fluorescence of the Trp residues of the enzyme (so-called “intrinsic” protein fluorescence), and thus offers a spectrophotometric method of following binding and determining values of K_d (Figure S.1B).¹⁷ We treated HCA^{**}-SSEG_{*n*}SA ($n = 0, 2, 5, 10, \text{ and } 20$) and HCA^{**} (incubated with 2 equiv of SA-OMe) with different concentrations of Ethox and followed the fluorescence of the free, unbound protein (excitation wavelength = 290 nm, emission wavelength = 340 nm) to determine dissociation constants (Figure 5). Figure 6 shows the normalized data and fits to the data assuming a single-site binding model (see Experimental Section). As linker length (n) increased, $K_d^{\text{comp.Ethox}}$ (Figure 5) for HCA^{**}-SSEG_{*n*}SA reached a peak when $n = 2$, and then decreased. The magnitude of the decrease in $K_d^{\text{comp.Ethox}}$ was only weakly dependent on n (only a ~8-fold decrease from $n = 2$ to $n = 20$; Table 3).

To account for the presence of unmodified HCA^{**} in the HCA^{**}-SSEG₅SA sample (see previous section), we normalized the fluorescence data (Figure 6) to the intensity observed when [Ethox] = 40 nM (this concentration is sufficient to saturate the binding sites of all of the residual unmodified HCA^{**}). Given the high value of $K_d^{\text{comp.Ethox}}$ (5.6 μM) observed for this protein, <1% of the HCA^{**}-SSEG₅SA in the sample will bind Ethox at this concentration.

The fluorescence of HCA^{**} that had been incubated with ~2 equiv of SA-OMe (the ligand that was covalently tethered to HCA^{**} in the modified proteins HCA^{**}-SSEG_{*n*}SA) decreased linearly with increasing concentrations of Ethox until a stoichiometric amount of Ethox was present (Figure 6). This result demonstrates that the presence of free, unbound SA-OMe could not explain the observation that the affinity of Ethox is much lower for HCA^{**}-SSEG_{*n*}SA than for HCA^{**}, and that the value of K_d^{Ethox} (Figure 5) for HCA^{**} is too low to be determined *directly* by this method (i.e., the value of K_d^{Ethox} for HCA^{**} is much lower than the concentration of protein, ~200 nM, required for a detectable signal from intrinsic fluorescence).

To determine the value of K_d^{Ethox} , we allowed Ethox and DNSA (5 μM) to compete for the active site of HCA^{**}-SCH₂CO₂⁻ and measured the decrease in fluorescence of the HCA^{**}-SCH₂CO₂⁻-DNSA complex (excitation wavelength = 290 nm, emission wavelength = 460 nm) with increasing concentration of Ethox (Figure S.5; Table 2).^{16, 24, 25} The value we obtained ($K_d^{\text{Ethox}} = 0.2$ nM) was in good agreement with one from the literature (Table 2).¹⁷ We used the same competition experiment with DNSA to determine the value of the dissociation constant of SA-OMe ($K_d^{\text{inter}} = 0.51$ μM ; Figure 1C and eq 1) for HCA^{**}-SCH₂CO₂⁻ (Figure S.5; Table 2).

Measurement of K_d^{intra} and Calculation of M_{eff}

Eq 5 gives the dissociation constant (K_d^{intra}) for the intramolecular binding of the tethered ligand to the active site of HCA for HCA^{**}-SSEG_{*n*}SA from a thermodynamic cycle for the binding of Ethox to HCA^{**}-SSEG_{*n*}SA (Figure 5). Values of $K_d^{\text{comp.Ethox}}$ (for HCA^{**}-SSEG_{*n*}SA) and K_d^{Ethox} (for HCA^{**}-SCH₂CO₂⁻) were measured as described in the previous section. Table 3 lists the calculated values of K_d^{intra} .

$$K_d^{\text{intra}} = K_d^{\text{Ethox}} / K_d^{\text{comp.Ethox}} \quad (5)$$

Table 3 also lists values of M_{eff} calculated using eq 1, and values of K_d^{intra} and K_d^{inter} (the dissociation constant for the binding of SA-OMe to HCA^{**}-SCH₂CO₂⁻; see previous section and Figure 1C,D). M_{eff} was only weakly dependent on linker length (n) when $n \geq 2$: a decrease of approximately a factor of 8 between $n = 2$ and $n = 20$. The model shown in Figure 2A anticipates a much greater dependence of M_{eff} on n than that observed: the simplest model, which assumes that each single bond in an EG unit of the linker has three possible conformations before binding and only one after binding, anticipates a >10-fold decrease in

M_{eff} per EG unit (see Introduction).^{4, 11, 12} This model is inconsistent with the weak dependence of M_{eff} on n observed experimentally, and can be eliminated.

Theoretical Values of C_{eff} for Intramolecular Binding of Sulfonamide-Conjugated HCA** (HCA** $-SSEG_nSA$)

As mentioned in the section on Experimental Design, the “ligand” in HCA** $-SSEG_nSA$ consists of the benzenesulfonamide moiety and the alkyl chain that is directly attached to it ($p\text{-H}_2\text{NSO}_2\text{C}_6\text{H}_4\text{CH}_2\text{NHCO}(\text{CH}_2)_4\text{CO}-$); thus, the linker consists of the EG_n chain and the remainder of the alkyl chain ($-\text{NH}(\text{CH}_2\text{CH}_2\text{O})_n\text{CH}_2\text{CH}_2\text{NHCOCH}_2\text{CH}_2\text{S}-$). Treating the linker as a copolymer allowed us to estimate a value for $\langle r^2 \rangle^{1/2}$ from the sum of the contributions from the EG_n and the non- EG_n portions of the linker (eq 6).⁸ For the EG_n portion of the linker, we used a literature value of 0.58 nm for the parameter C in eq 3.³² We assumed that the non- EG_n portion of the linker was “alkyl-like” and contributed a constant value of 0.40 nm to the linker length.³³

$$\langle r^2 \rangle^{1/2} = \sqrt{0.58^2 n + 0.40^2} \quad (6)$$

In order to estimate C_{eff} (eq 4), we need a value for the distance (d) between the sites of covalent attachment of the ligand to the protein and of the ligand bound at the active site (Figure 1D). There are no reported X-ray crystal structures of complexes of HCA with p -substituted benzenesulfonamides with alkyl tails, so we examined the crystal structure of a complex of HCA with a p -substituted benzenesulfonamide having a tri(ethylene glycol) tail ($\text{ArEG}_3\text{NH}_3^+$; Figure 4): we assumed that the alkyl chain of the ligand in HCA** $-SSEG_nSA$ interacts with the same hydrophobic patch of CA, and has the same orientation in the active site as the EG_3 tail in $\text{ArEG}_3\text{NH}_3^+$. From the crystal structure, we estimated a distance of ~ 0.9 nm between the first methylene group of the third glycol unit of $\text{ArEG}_3\text{NH}_3^+$ —a position that should correspond to the $-\text{NH}-$ group of the linker in HCA** $-SSEG_nSA$ —and the γ -carbon of Lys-133 of HCA—a position that should correspond to the thiol of HCA** (Figure 4).

Figure 7 shows our experimental values of M_{eff} (see previous section) and a fit to the data calculated using eq 4 for C_{eff} and eq 6 for $\langle r^2 \rangle^{1/2}$. We allowed the values of d and p to vary in order to optimize the fit. The value of d (0.82 ± 0.03 nm) that gave the best least-squares fit to the data is very close to the value (0.9 nm) that we determined from analyzing the X-ray crystal structure of the HCA- $\text{ArEG}_3\text{NH}_3^+$ complex (Figure 4). This observation, together with the excellent fit to the experimental data, suggests that the theoretical model based on polymer theory successfully (if unexpectedly) rationalizes the dependence of the intramolecular binding of the tethered ligand to the active site of the protein, as a function of the length of the variable part of the linker. Further, because the theory from which C_{eff} is derived explicitly ignores enthalpic contributions to ring closure, these observations suggest that entropy is dominant in intramolecular binding (see the next section for further discussion).

The value of p (0.12 ± 0.009 ; eq 4) that gave the best fit to the data is significantly lower than the value of 2 proposed by Lees and coworkers for the bivalent protein-ligand complex (Figure 1E).¹ Because the point of attachment of the benzenesulfonamide ligands to HCA for HCA** $-SSEG_nSA$ is at the periphery of the conical cleft (Figure 1D and 4), the ligand will encounter less steric repulsion with the protein (when it is not bound at the active site of CA) than the unbound ligand in the intermediate state in a bivalent protein-ligand complex (Figure 1E). This effect can decrease the value of p to no less than unity—the unbound ligand has a full sphere of volume to explore, and is in a similar environment to a polymer that is free in solution. At present, we do not have a clear understanding of what effect could contribute the additional 8-fold decrease in p . We speculate that the geometry of the catalytic cleft of CA (~ 1.5 nm deep and conical; Figure 4) could result in the polymer being excluded from the cleft because of steric interaction; this excluded volume effect would lower p (and C_{eff}).

Calorimetric Investigation to Determine Enthalpy and Entropy of Binding for K_d^{intra}

Our previous work suggested that EG_n tails interacted with a hydrophobic patch in the conical cleft of CA, albeit with no effect on the observed dissociation constants.^{21, 24} We have recently shown that increasing the length of these tails makes the enthalpy of association of these ligands with CA more favorable, and the entropy less favorable, in a way that perfectly balances to give no change in values of K_d .³⁰ As mentioned in the section on Experimental Design, we designed $\text{HCA}^{**}\text{-SSEG}_n\text{SA}$ to include a conserved alkyl chain to saturate this hydrophobic surface in the conical cleft of the enzyme. Thus, we would not anticipate that contacts between the EG_n linker and the surface of HCA could occur (Figure 2C,D).

To test this idea, we used isothermal titration calorimetry (ITC) to dissect the intramolecular dissociation constant (K_d^{intra} ; Figures 1D, 2, and 5) for $\text{HCA}^{**}\text{-SSEG}_{10}\text{SA}$ into its enthalpic ($\Delta H^{\circ}_{\text{intra}}$) and entropic ($\Delta S^{\circ}_{\text{intra}}$) components. We examined the binding of Ethox to $\text{HCA}^{**}\text{-SSEG}_{10}\text{SA}$ (Figure S.6) and to the control protein $\text{HCA}^{**}\text{-SCH}_2\text{CO}_2^-$ (Figure S.7A). We calculated $\Delta H^{\circ}_{\text{intra}}$ and $\Delta S^{\circ}_{\text{intra}}$ for $\text{HCA}^{**}\text{-SSEG}_{10}\text{SA}$ by making use of the thermodynamic cycle shown in Figure 5 (Table 2). The value of $K_d^{\text{comp, Ethox}}$ that we measured by ITC for the binding of Ethox to $\text{HCA}^{**}\text{-SSEG}_{10}\text{SA}$ was in good agreement (a difference of a factor of 2–3) to the one that we measured by fluorescence quenching (Tables 2 and 3).

The value of $\Delta H^{\circ}_{\text{intra}}$ for $\text{HCA}^{**}\text{-SSEG}_{10}\text{SA}$ ($-6.7 \pm 0.7 \text{ kcal mol}^{-1}$; Table 2) was the same, within error, as $\Delta H^{\circ}_{\text{obs}}$ for the binding of SA-OMe (a ligand containing the benzenesulfonamide moiety and alkyl chain of $\text{HCA}^{**}\text{-SSEG}_{10}\text{SA}$) to $\text{HCA}^{**}\text{-SCH}_2\text{CO}_2^-$ ($-6.1 \pm 0.12 \text{ kcal mol}^{-1}$; Figure S.7B and Table 2). This result suggests that there is little, if any, interaction of the EG_n linker with the protein: our previous results suggested that any interaction of the EG_n chain with CA should occur with a measurable change in enthalpy.³⁰ Thus, this result serves to rule out the model shown in Figure 2C.

There remains the possibility that the linker makes stabilizing contacts with the protein whether the ligand is bound or not bound at the active site (Figure 2D), and thus there would be no change in enthalpy for these contacts between the two states. We cannot rigorously rule out this possibility. The fact that a theory from polymer chemistry that is based entirely on entropic terms fits the experimental data very well (Figure 7), however, persuades us to prefer the simpler model shown in Figure 2B.

Conclusions

This paper reports values of effective molarity (M_{eff}) as a function of the length (n) of linkers of oligo(ethylene glycol) for a ligand covalently tethered to a protein (Figure 2). Values of M_{eff} reach a maximum when the linker is the optimal length to allow the ligand to bind to the active site of the protein, and then decrease weakly with increasing linker length (decrease in M_{eff} by only a factor of ~ 8 over the range $n = 2$ to $n = 20$) beyond this value (Figure 7). The experimental data are well-explained by a theoretical, entropic model from polymer chemistry based on the probability that the ends of the linker are separated by the same distance as the sites of covalent attachment and of noncovalent binding of the tethered ligand. Calorimetry revealed that enthalpic contacts of the linker with the protein are not important in this system, and that entropy is dominant (Table 2; Figure 2C,D).

Our results suggest that, at least in the case of oligo(ethylene glycol) linkers, contacts of the linker with the surface of the protein (outside of the active site) can be safely ignored. These results are consistent with previous studies that demonstrated that surfaces displaying oligo(ethylene glycol) groups did not adsorb proteins non-specifically.²⁶ These results can be reconciled with our previous findings that suggested an interaction of oligo(ethylene glycol) tails with CA: in these previous studies the tails interacted with a surface of the protein inside

the conical cleft of CA.^{21, 30} This possibility is precluded for the ligands studied here (HCA**_n-SSEG_nSA) because they have alkyl chains of sufficient length to saturate this hydrophobic surface.

Most importantly, our results are not consistent with a model in which the linker is *completely* restricted upon association of the ligand with the active site of the protein (when the linker is bound at both ends; Figure 2A). This model requires that M_{eff} decrease sharply with the length of the linker (>10-fold decrease with each ethylene glycol unit); an expectation that is not met by the experimental results. Our results support a model in which the linker has significant conformational mobility when bound at both ends (Figure 2B).

The values of M_{eff} reported here should be intrinsic properties of the oligo(ethylene glycol) linkers themselves. These values can, thus, be used to *predict* the avidities of multivalent ligands containing oligo(ethylene glycol) linkers for a multivalent protein, if the monovalent affinity is known (Figure 1E). Our results suggest, however, that the quantitative accuracy of these predictions is limited by difficulties in estimating the influence of such effects as the “excluded volume” term and the geometric and steric demands of the active site of the protein target on ligand binding (both of which appear in the parameter p in eq 4).

We did not anticipate the weak dependence of M_{eff} on the length of the linker (when it is longer than the optimal length). This *weak* dependence suggests that the most effective strategy for the design of multivalent ligands will connect the ligand moieties by a flexible linker that is significantly longer than the spacing between the multivalent sites of the target protein (Figure 1E). Our results suggest that the penalty in conformational entropy for such a linker in the association of these multivalent ligands with their multivalent proteins will be low (and much lower than we previously hypothesized^{4, 11, 12}). Indeed, the results presented here using oligo(ethylene glycol) linkers, which do not interact with the surfaces of proteins, can be seen as a worst-case scenario. Other types of linkers (e.g., alkyl chains) are likely to interact with the surfaces of proteins, and these interactions should effectively “shorten” the linker and make M_{eff} even less sensitive to the length of the linker than in the results presented here (Figure 2C,D).

Finally, ligands with flexible linkers can adopt a number of conformations without steric strain (unlike ligands with rigid linkers); this flexibility should allow the multivalent ligand to sample conformational space to optimize its binding to the multiple binding sites. This sampling of conformational space reduces the possibility of a sterically obstructed fit, a circumstance that can occur readily with rigid linkers that are not perfectly designed.⁴

Experimental Section

General Methods

All chemicals were obtained from Sigma (St. Louis, MO), Aldrich Milwaukee, WI, or Alfa Aesar (Ward Hill, MA) and used without further purification, unless otherwise noted. DNSA was recrystallized from ethanol prior to use. DNA sequencing was performed by the DNA sequencing facility at the Harvard University Department of Molecular and Cellular Biology. Fluorescence measurements were performed on a Molecular Devices SpectraMax M5 for detection at 340 nm (intrinsic protein fluorescence for Ethox binding assays), and on a Molecular Devices SpectraMax Gemini XS for detection at 460 nm (for DNSA binding assays). Isothermal titration calorimetry was performed using a VP-ITC microcalorimeter from MicroCal (Northampton, MA). Varian Inova spectrometers operating at 400 MHz and 500 MHz (¹H) were used for NMR experiments. UV-Vis spectroscopy was conducted on a Hewlett Packard 8453 spectrophotometer (Palo Alto, CA). UV spectroscopy was used to quantify

HCA** proteins ($\epsilon_{280} = 54,000 \text{ M}^{-1} \text{ cm}^{-1}$)³¹ and DNSA ($\epsilon_{326} = 4640 \text{ M}^{-1} \text{ cm}^{-1}$)²⁵. Ethox and SA-OMe were quantified by ¹H NMR spectroscopy.^{24, 25, 30, 34}

Preparation of Plasmid Encoding HCA**

The plasmids encoding wild-type HCA II (pACA) and a Cys-206→Ser mutant of HCA II (pC206S) were gifts from Professor Carol Fierke^{22, 29}. The Lys-133→Cys mutation was introduced into the pC206S plasmid using oligonucleotide site-directed mutagenesis with a Quickchange Mutagenesis Kit (Stratagene) and following the manufacturer's directions. This plasmid was digested with the restriction enzymes BamHI and KasI (New England BioLabs), purified by agarose gel electrophoresis using a Quick Gel extraction kit (Qiagen), and ligated into the pACA backbone, which had been digested with BamHI and KasI, treated with CIP (calf intestinal alkaline phosphatase from New England BioLabs; to prevent re-circularization of the vector without the insert), and purified by gel electrophoresis. The presence of the desired mutation was verified by sequencing the entire HCA II gene using the method of Sanger et al.³⁵

Expression and Purification of HCA**

BL21(DE3) cells were transformed with the HCA** plasmid and grown as described by Fierke and coworkers (see, for example, ref¹⁸). Cell lysates were prepared as described by Krebs and Fierke,²² and the HCA** protein was purified using sulfonamide-conjugated agarose (Sigma-Aldrich) as described by Khalifah et al.¹⁹ The purified HCA** protein was stored at -80°C in 50 mM Tris-sulfate buffer pH 8.0 with 0.2 mM ZnSO₄ and 1 mM DTT (to prevent oxidation of the thiol).

Preparation of Modified HCA** Proteins

HCA** was desalted over a NAP-10 column (Amersham) into 50 mM Tris-sulfate buffer pH 8.0 from which oxygen had been removed by bubbling Ar through it for >1 h. HCA** (50–100 μM) was treated with 5–10 equiv of the appropriate modifying agent (Pyr-SSEG_nSA, Pyr-SSEG₂CONH₂, or iodoacetate) to generate the desired modified HCA** (HCA**-SSEG_nSA, HCA**-SSEG₂CONH₂, or HCA**-SCH₂CO₂⁻, respectively). The reaction was allowed to proceed for 1 h at 25 °C and then for ~18 h at 4 °C. The reaction products were purified by exhaustive dialysis, and then characterized by ESI-MS (electrospray ionization mass spectrometry) to verify the appearance of a peak corresponding to the combined masses of HCA** and the chemical reactant (see Table S.1). For the HCA**-SSEG_nSA proteins, contamination by unmodified HCA** was assessed by examining the binding of DNSA (see Supporting Experimental Procedures and Figure S.4).

Determination of Dissociation Constants ($K_d^{\text{comp, Ethox}}$) for the Binding of Ethox to HCA**-SSEG_nSA

To the wells of a black microwell plate were added dilutions of Ethox and then the appropriate HCA** protein (200 nM) in a final volume of 200 μL of 20 mM sodium phosphate pH 7.5. The plate was allowed to incubate at 25 °C for 2 h, and then its fluorescence was measured with an excitation wavelength of 290 nm and an emission wavelength of 340 nm (with a 325 nm cut-off filter). Wells were read ~50 times. Fluorescence intensities were normalized to the fluorescence intensity of the protein with no added Ethox, and then corrected for the inner filter effect³⁶ by using a control experiment where soybean trypsin inhibitor (a protein with no affinity for sulfonamides)³⁷ was treated with Ethox. The corrected data were fit to eq 7 where $[\text{CA}]_{\text{total}}$ is the total concentration of modified HCA** in the assay (200 nM) and $[\text{Ethox}]_{\text{total}}$ is the total concentration of Ethox (the independent variable). Only the parameter $K_d^{\text{comp, Ethox}}$ was allowed to vary to optimize the fit by non-linear least squares fitting (Origin), with the other parameters held constant at their known values.

$$F=0.3+0.7 \left(1 - \frac{K_d^{\text{comp,Ethox}} + [\text{CA}]_{\text{total}} + [\text{Ethox}]_{\text{total}} - \sqrt{(K_d^{\text{comp,Ethox}} + [\text{CA}]_{\text{total}} + [\text{Ethox}]_{\text{total}})^2 - 4[\text{CA}]_{\text{total}}[\text{Ethox}]_{\text{total}}}}{2[\text{CA}]_{\text{total}}} \right) \quad (7)$$

Eq 7 assumes that Ethox quenches 70% of the fluorescence of HCA**⁺; this value was determined by examining the fluorescence of HCA**⁺ in the presence of Ethox (Figure S.1B).

The data were also fit to eq 8, which assumes that the concentration of Ethox not bound by CA is equal to the total concentration of Ethox ($[\text{Ethox}]_{\text{total}}$) and gives a sigmoid fit to the data when the concentration of Ethox is plotted on a logarithmic scale. This equality is not true for HCA**SSEG₀SA and HCA**SSEG₂₀SA, because in those cases $K_d^{\text{comp,Ethox}} \sim [\text{CA}]_{\text{total}}$ (200 nM).

$$F=0.3+0.7 \frac{K_d^{\text{comp,Ethox}}}{K_d^{\text{comp,Ethox}} + [\text{Ethox}]_{\text{total}}} \quad (8)$$

Isothermal Titration Calorimetry

In order to determine values of ΔH° and K_d , $\sim 6 \mu\text{M}$ HCA**⁺-SCH₂CO₂⁻ or $\sim 60 \mu\text{M}$ HCA**⁺-SSEG₁₀CA, which had been purified over sulfonamide-conjugated agarose (see Supporting Experimental Procedures), in 20 mM sodium phosphate buffer pH 7.5 (with 0.6% DMSO-*d*₆) was titrated with 120 or 440 μM , respectively, of Ethox in the same buffer at 25 °C. HCA**⁺-SCH₂CO₂⁻ ($\sim 6 \mu\text{M}$) was also titrated with 120 μM of SA-OMe. See Figures S.6 and S.7 for detailed information. After subtraction of background heats, the data were analyzed by a single-site binding model using the Origin software (provided by Microcal) and allowing the values of binding stoichiometry, ΔH° , and K_d to vary to optimize the fit.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

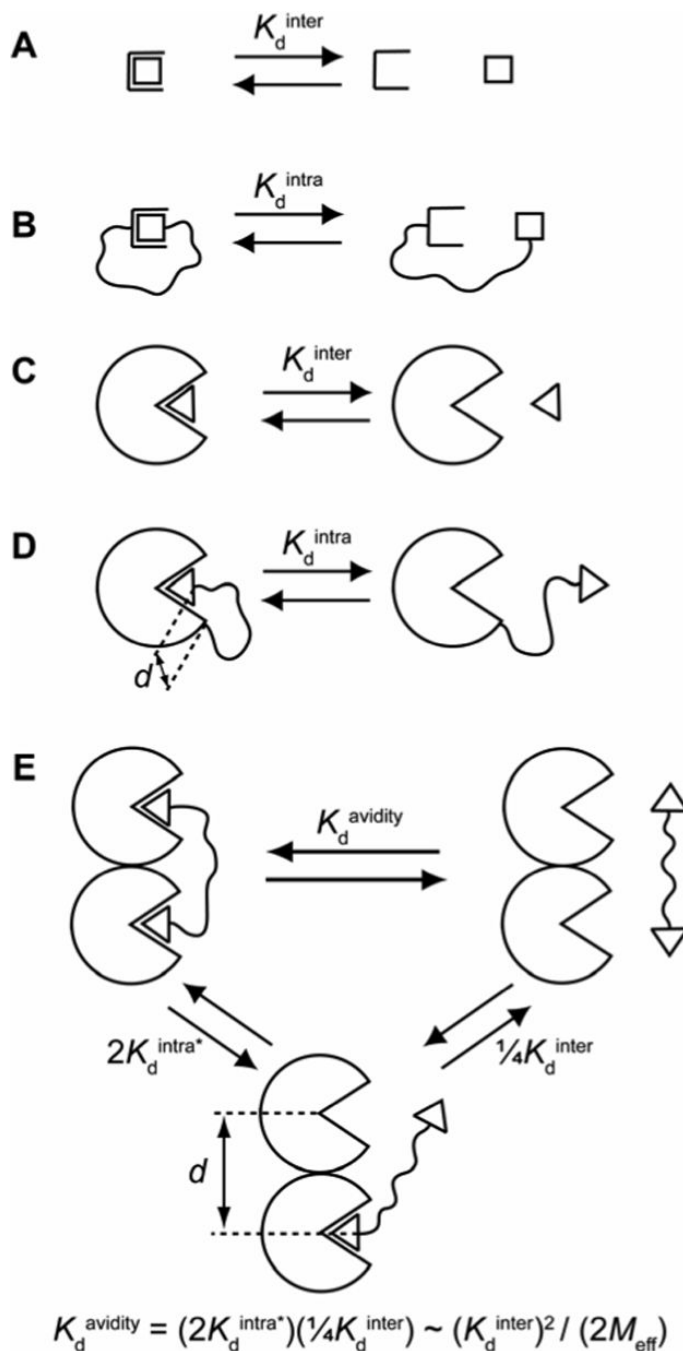
This work was supported by the National Institutes of Health (GM30367). V.M.K. and P.J.B. acknowledge support from NDSEG and NSF pre-doctoral fellowships, respectively. V.S. acknowledges support from La Ligue Contre Le Cancer (France). We thank Professor Carol Fierke (University of Michigan) for the kind gift of the HCA-encoding plasmids, and Andrea Stoddard (Fierke group, Michigan) for helpful discussions about protein purification. We thank Dr. Julian P. Whitelegge (The Pasarow Mass Spectrometry Laboratory, University of California, Los Angeles) for ESI-MS analysis. We thank Professor Greg Verdine (Harvard University) for use of the Spectramax M5, and Marie Spong (Verdine group, Harvard) for helpful discussions about molecular biology. We acknowledge the Bauer Center for Genomics Research (Harvard University) for the use of the Spectramax Gemini XS and Perspective Biosystems Voyager-DE PRO for MALDI.

References

- Gargano JM, Ngo T, Kim JY, Acheson DWK, Lees WJ. *J. Am. Chem. Soc* 2001;123:12909–12910. [PubMed: 11749553]
- a Page MI, Jencks WP. *Proc. Natl. Acad. Sci. U. S. A* 1971;68:1678–1683. [PubMed: 5288752]Page, MI. *Chem. Soc. Rev.* 2. 1973. p. 295-323.Kirby, AJ. *Adv. Phys. Org. Chem.* Gold, V.; Bethell, D., editors. 17. Academic Press; London: 1980. p. 183-278.
- Mandolini, L. *Adv. Phys. Org. Chem.* Gold, V.; Bethell, D., editors. 22. Academic Press; London: 1986. p. 1-111.
- Krishnamurthy, VM.; Estroff, LA.; Whitesides, GM. *Fragment-based Approaches in Drug Discovery.* Jahnke, W.; Erlanson, DA., editors. 34. Wiley-VCH; Weinheim: 2006. p. 11-53.
- Jacobson H, Stockmayer WH. *J. Chem. Phys* 1950;18:1600–1606.

6. Mulder A, Auletta T, Sartori A, Del Ciotto S, Casnati A, Ungaro R, Huskens J, Reinhoudt DN. *J. Am. Chem. Soc* 2004;126:6627–6636. [PubMed: 15161290]
7. Kramer RH, Karpen JW. *Nature* 1998;395:710–713. [PubMed: 9790193]
8. Flory, PJ. *Principles of Polymer Chemistry*. Cornell University Press; Ithaca, NY: 1953.
9. In order for the theory to apply, the polymer must be long enough to eliminate the possibility of non-ideal effects (e.g., the non-Gaussian nature of short chains, correlations between bond angles of different subunits, transannular steric effects; see refs⁵ and¹⁰). Jacobson and Stockmayer have argued that this minimum length is on the order of 15 atoms in the chain, but the exact length depends on the flexibility and structure of the polymer (see ref¹⁰).
10. Winnik MA. *Chem. Rev* 1981;81:491–524.
11. Mammen M, Choi S-K, Whitesides GM. *Angew. Chem. Int. Ed* 1998;37:2755–2794.
12. Mammen M, Shakhnovich EI, Whitesides GM. *J. Org. Chem* 1998;63:3168–3175.
13. Choi, S-K. *Synthetic Multivalent Molecules: Concepts and Biomedical Applications*. John Wiley & Sons, Inc.; Hoboken, NY: 2004.
14. Urbach, AR.; Whitesides, GM. in preparation b Christianson DW, Fierke CA. *Acc. Chem. Res* 1996;29:331–339.
15. Supuran, CT.; Scozzafava, A.; Casini, A. *Med. Res. Rev.* 23. 2003. p. 146-189. Supuran, CT.; Scozzafava, A.; *Carbonic Anhydrase: Its Inhibitors and Activators*. Supuran, CT.; Scozzafava, A.; Conway, J., editors. 1. CRC Press; Boca Raton, FL: 2004. p. 67-147.
16. Chen RF, Kernohan JC. *J. Biol. Chem* 1967;242:5813–5823. [PubMed: 4990698]
17. Kernohan JC. *Biochem. J* 1970;120:26P.
18. Burton RE, Hunt JA, Fierke CA, Oas TG. *Protein Sci* 2000;9:776–785. [PubMed: 10794421]
19. Khalifah RG, Strader DJ, Bryant SH, Gibson SM. *Biochemistry* 1977;16:2241–2247. [PubMed: 16641]
20. Mårtensson L-G, Jonsson B-H, Freskgård P-O, Kihlgren A, Svensson M, Carlsson U. *Biochemistry* 1993;32:224–231. [PubMed: 8418842]
21. Boriack PA, Christianson DW, Kingery-Wood J, Whitesides GM. *J. Med. Chem* 1995;38:2286–2291. [PubMed: 7608893]
22. Krebs JF, Fierke CA. *J. Biol. Chem* 1993;269:948–954. [PubMed: 8419374]
23. a Gao J, Qiao S, Whitesides GM. *J. Med. Chem* 1995;38:2292–2301. [PubMed: 7608894] b King RW, Burgen ASV. *Proc. R. Soc. London, B* 1976;193:107–125. [PubMed: 5728] c Cappalonga Bunn AM, Alexander RS, Christianson DW. *J. Am. Chem. Soc* 1994;116:5063–5068.
24. Jain A, Huang SG, Whitesides GM. *J. Am. Chem. Soc* 1994;116:5057–5062.
25. Jain A, Whitesides GM, Alexander RS, Christianson DW. *J. Med. Chem* 1994;37:2100–2105. [PubMed: 8027991]
26. a Prime KL, Whitesides GM. *Science* 1991;252:1164–1167. b Prime KL, Whitesides GM. *J. Am. Chem. Soc* 1993;115:10714–10721. Mrksich, M.; Whitesides, GM. *Poly(ethylene Glycol): Chemistry and Biological Applications*. 1st ed.. Harris, JM.; Zalipsky, S., editors. 680. American Chemical Society; Washington, D. C.: 1997. p. 361-373.
27. Gilbert, HF. *Adv. Enzymol. Relat. Areas Mol. Biol.* Meister, A., editor. 63. Wiley; New York: 1990. p. 69-172. b Lees WJ, Whitesides GM. *J. Org. Chem* 1993;58:642–647. c Erlanson DA, Wells JA, Braisted AC. *Annu. Rev. Biophys. Biomol. Struct* 2004;33:199–223. [PubMed: 15139811] d Erlanson DA, Braisted AC, Raphael DR, Randal M, Stroud RM, Gordon EM, Wells JA. *Proc. Natl. Acad. Sci. U. S. A* 2000;97:9367–9372. [PubMed: 10944209] e Erlanson DA, Hansen SK. *Curr. Opin. Chem. Biol* 2004;8:399–406. [PubMed: 15288250]
28. Carlsson J, Drevin H, Axén R. *Biochem. J* 1978;173:723–737. [PubMed: 708370]
29. Nair SK, Calderone TL, Christianson DW, Fierke CA. *J. Biol. Chem* 1991;266:17320–17325. [PubMed: 1910042]
30. Krishnamurthy VM, Bohall BR, Semetey V, Whitesides GM. *J. Am. Chem. Soc* 2006;128:5802–5812. [PubMed: 16637649]
31. Nyman PO, Lindskog S. *Biochim. Biophys. Acta* 1964;85:141–151. [PubMed: 14159292]
32. Knoll D, Hermans J. *J. Biol. Chem* 1983;258:5710–5715. [PubMed: 6853541]

33. There are a seven single bonds in the non-EGn portion of the linker (between the thiol of the linker and the –NH– of the carboxamide closest to the alkyl chain of the ligand). Using eq 3 with an estimate for l of 0.153 nm (the C-C bond length) and assuming an ideal, random-flight polymer ($C/l = 1$), gives an estimate of ~ 0.4 nm for the value of the non-EGn portion of the linker.
34. Krishnamurthy, VM.; Bohall, BR.; Kim, C-Y.; Moustakas, DT.; Christianson, DW.; Whitesides, GM. submitted
35. Sanger F, Nicklen S, Coulson AR. Proc. Natl. Acad. Sci. U. S. A 1977;74:5463–5467. [PubMed: 271968]
36. Lakowicz, JR. Principles of Fluorescence Spectroscopy. 2nd ed.. Kluwer Academic/Plenum; New York: 1999.
37. Colton IJ, Carbeck JD, Rao J, Whitesides GM. Electrophoresis 1998;19:367–382. [PubMed: 9551788]

**Figure 1.**

(A) A monovalent binding event between a receptor and ligand is characterized by a dissociation constant (K_d^{inter}) with units of concentration (M). (B) When the ligand is tethered to the receptor, the dissociation constant (K_d^{intra}) is now dimensionless and is related to K_d^{inter} by the effective molarity, M_{eff} (eq 1). (C) Analogous to (A) for a binding event between a protein and ligand. (D) Analogous to (B) for an intramolecular protein-ligand binding event. In this case, however, the distance (d) between the two ends of the linker in the bound state is non-zero (unlike the case in (B)). (E) Dissociation of a bivalent ligand from a bivalent receptor. This process can be conceptualized as occurring in two steps: the first step (with dissociation constant $2K_d^{\text{intra}^*}$) is intramolecular, and the second step (with dissociation constant

$\frac{1}{4}K_d^{\text{inter}}$ is intermolecular. The equation given provides a means of estimating M_{eff} from the observed dissociation constant (K_d^{avidity}) and the intermolecular dissociation constant (K_d^{inter}) by assuming that the dissociation constant for the first step is equivalent to the product of a statistical factor of 2 and K_d^{intra} (defined in (D)); i.e., this procedure assumes that $K_d^{\text{intra}^*} = K_d^{\text{intra}}$. This assumption is often quite poor because of complications arising from the influence of binding at one site on binding at the other site (e.g., cooperativity between binding sites, enthalpy/entropy compensation).⁴

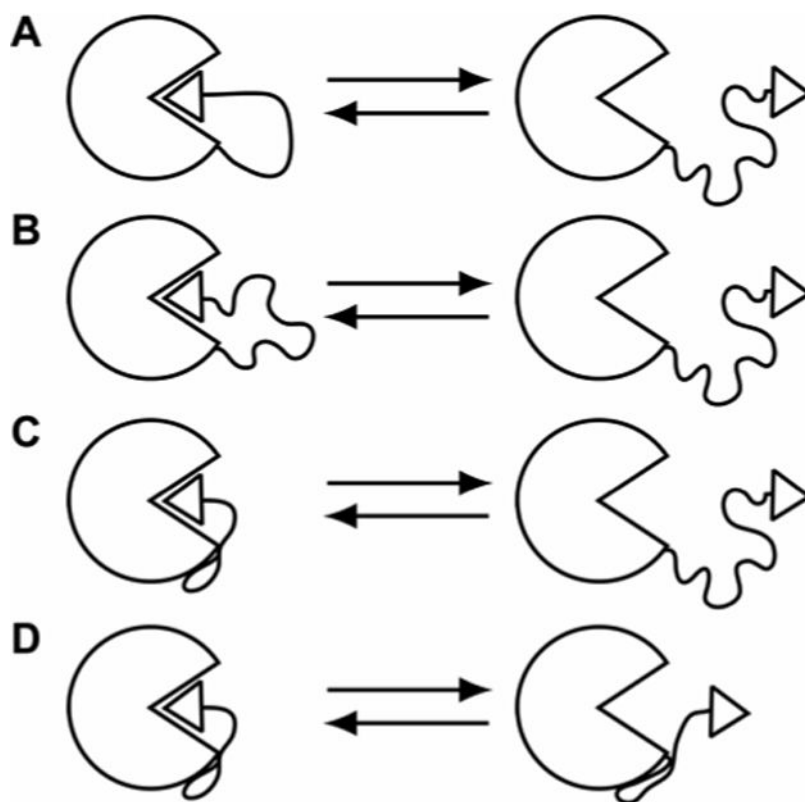


Figure 2. Possible models to explain the binding of a ligand, covalently attached to the surface of a protein by the “linker”, to the active site of that protein. When the ligand is bound at the active site, the linker: **(A)** has low conformational mobility (i.e., free rotations of the bonds of the linker are significantly restricted), **(B)** has significant conformational mobility, or **(C)** and **(D)** makes stabilizing contacts with the surface of the protein. In **(D)**, the linker makes contacts with the surface of the protein even when the ligand is not bound at the active site.

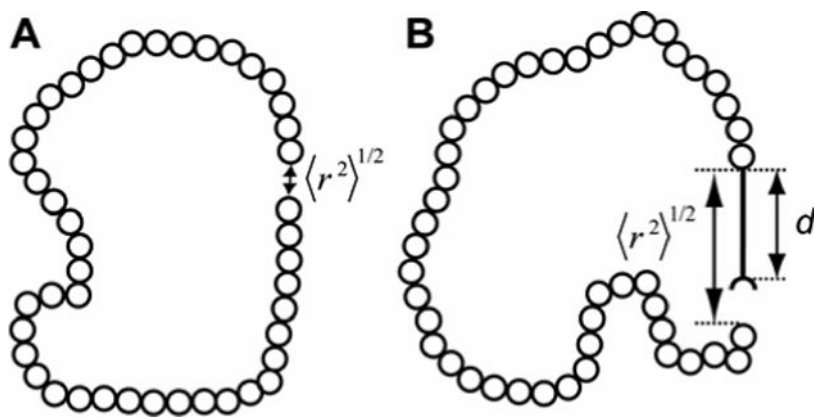


Figure 3. Random-coil polymer with the number (n) of repeat units (represented as open circles) equal to 45. (A) The polymer forms a closed loop: the distance $\langle r^2 \rangle^{1/2}$ between the ends of the polymer is near zero. (B) The polymer bears a physical pole of length (d) attached at one end. The distance $\langle r^2 \rangle^{1/2}$ between the ends of the polymer is close to this defined distance (d).

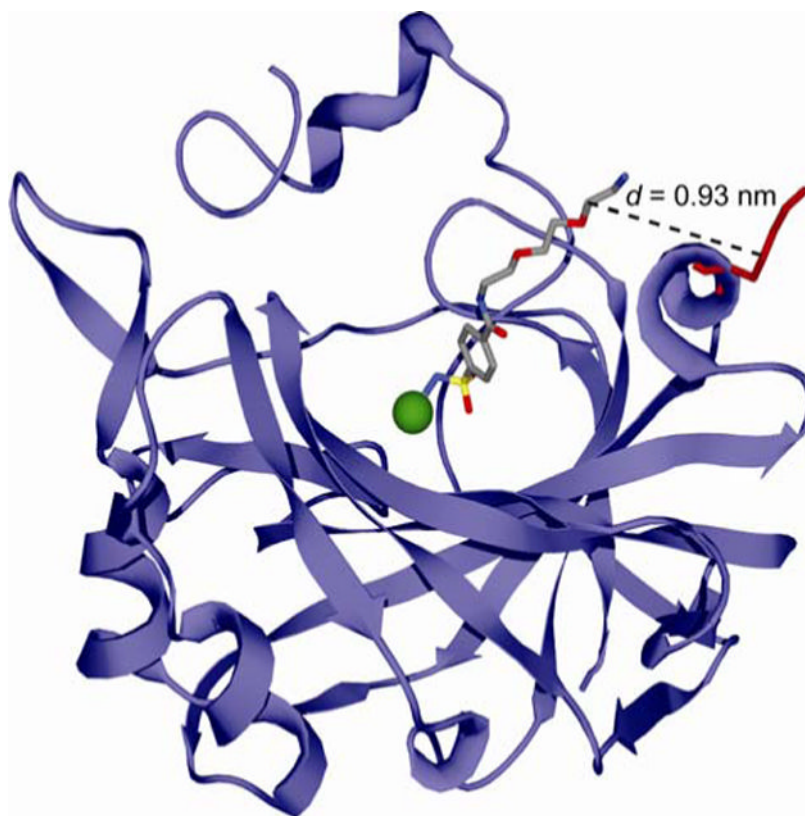


Figure 4. Model for the interaction of *p*-H₂NSO₂C₆H₄CONH(CH₂CH₂O)₂CH₂CH₂NH₃⁺ (ArEG₃NH₃⁺) with HCA based on the deposited X-ray crystallographic coordinates (PDB: 1CNX).²¹ The arylsulfonamide ligand is rendered as a ball-and-stick model in CPK color scheme. HCA is depicted as a light blue ribbon diagram with the catalytically essential Zn²⁺ cofactor shown as a green sphere. Lys-133 of HCA II is represented as a red ball-and-stick model. The distance (*d*) between the last glycol unit of the ligand and the γ-CH₂ group of Lys-133 (corresponding to the thiol in the Lys→Cys HCA** mutant) is indicated by the dashed line.

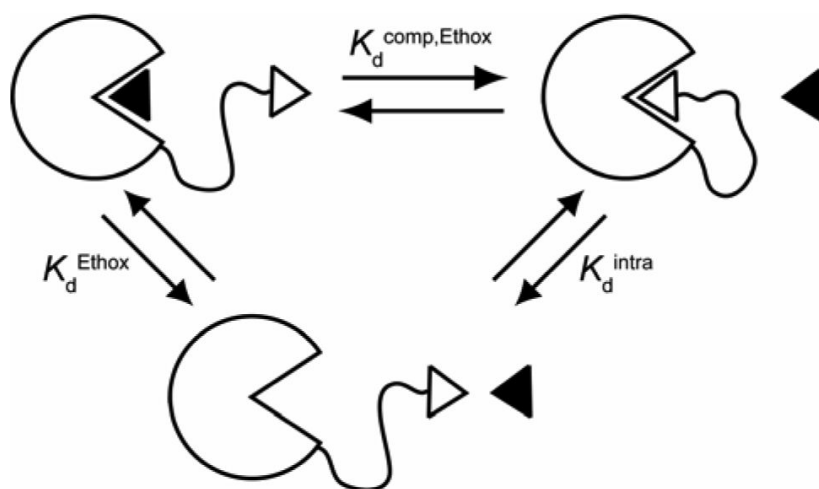


Figure 5. Schematic diagram depicting the strategy for the measurement of K_d^{intra} for HCA** - SSEG_nSA using the observed dissociation constant ($K_d^{\text{comp,Ethox}}$) for a competing ligand, Ethox (shown as black triangle). For this thermodynamic cycle, $K_d^{\text{intra}} = K_d^{\text{Ethox}} / K_d^{\text{comp,Ethox}}$ (eq 5).

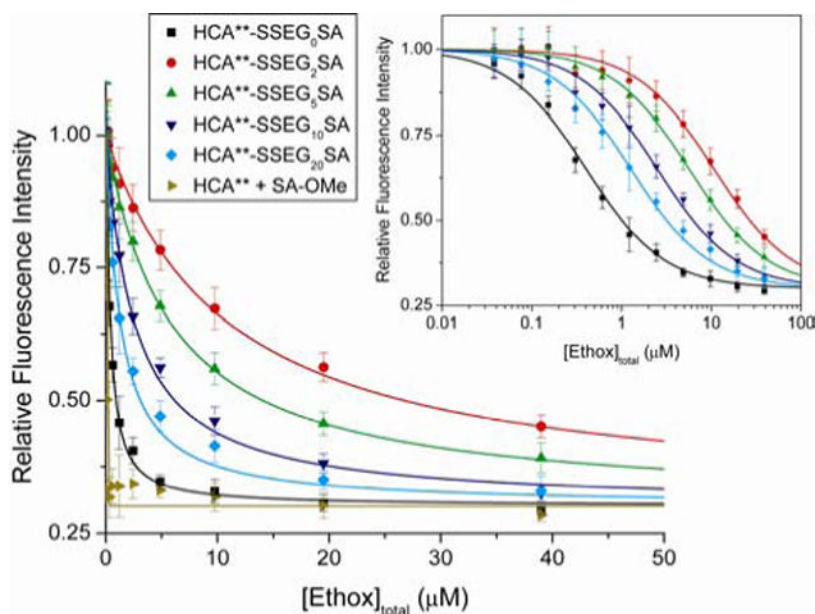


Figure 6.

Decrease in fluorescence of modified HCA** proteins ($\text{HCA}^{**}\text{-SSEG}_n\text{SA}$, $n = 0, 2, 5, 10,$ or 20) and HCA** (incubated with two equiv of SA-OMe) with increasing concentration of ethoxzolamide (Ethox). The intrinsic fluorescence of 200 nM of protein in 20 mM sodium phosphate pH 7.5 at 298 K with different concentrations of the fluorescence quencher Ethox was measured (excitation wavelength = 290 nm, emission wavelength = 340 nm). The data are shown after background subtraction, correction for the inner filter effect, and normalization to a maximum signal of unity (see Experimental Section). The solid curves are fits to the data using the full quadratic equation for binding (eq 7; this equation does not make the assumption that $[\text{Ethox}]_{\text{free}} \approx [\text{Ethox}]_{\text{total}}$), and using a value for the minimum in fluorescence intensity of all samples of 0.3 (Figure S.1B). Error bars represent the maximum variation of an independent measurement from the mean of four experiments (duplicates from two independent experiments). Inset: The data on a logarithmic scale for the x -axis. The solid curves are fits to the data using a single-site Langmuir binding model (eq 8), which makes the assumption that $[\text{Ethox}]_{\text{free}} \approx [\text{Ethox}]_{\text{total}}$ — a poor assumption when $n = 0$ or 20 (see Experimental Section).

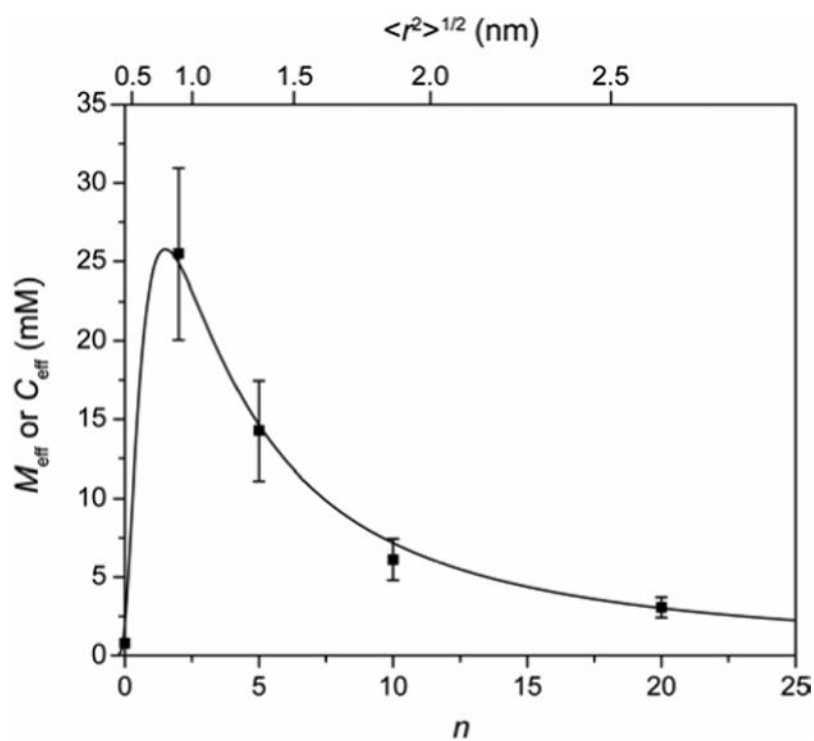
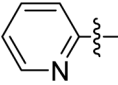
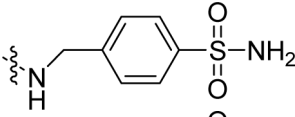
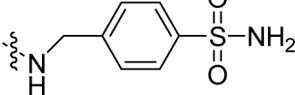
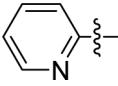


Figure 7. Variation of effective concentration (C_{eff}) and effective molarity (M_{eff}) with linker length (n), related to the root-mean-squared distance between the ends of the linker $\langle r^2 \rangle^{1/2}$, see eq 6) for HCA** -SSEG_n SA. The points are empirical values of M_{eff} (eq 1) from this study (Table 3). The solid curve is a fit to the data using the definition of C_{eff} (eq 4) with estimates of $\langle r^2 \rangle^{1/2}$ from eq 6. The best fit to the data was obtained with a value for d (the distance between the site of covalent attachment of the ligand to the protein and the active site of the protein; Figure 4) of 0.82 ± 0.03 nm, and for p of 0.12 ± 0.009 .

Table 1

Chemical labels and modified HCA** proteins used in this study.

Name	R	R'
Pyr-SSEG _n SA		
HCA**-SSEG _n SA	HCA**	
Pyr-SSEG _n CONH ₂		NH ₂
HCA**-SSEG _n CONH ₂	HCA**	NH ₂

Measured and calculated enthalpy and entropy values for intermolecular and intramolecular binding of sulfonamides to modified and unmodified HCA** proteins.

Table 2

Protein	Ligand	$K_d \times 10^6$	ΔG^0 kcal mol ⁻¹	ΔH^0 kcal mol ⁻¹	$-T\Delta S^0$ kcal mol ⁻¹
HCA**	DNSA	$0.30 \pm 0.014,^{abc}$	-8.90 ± 0.15	—	—
HCA**--SSEG ₇ CONH ₂	DNSA	$0.26 \pm 0.008,^{abc}$	-8.98 ± 0.10	—	—
HCA**--SCH ₂ CO ₂ ⁻	DNSA	$0.23 \pm 0.007,^{abc}$	-9.05 ± 0.10	—	—
	Ethox	$0.0002 \pm 0.00004,^{abcd}$	-13.2 ± 0.6	$-17.7 \pm 0.4,^{ef}$	$+4.5 \pm 0.7$
	SA-OMe	$0.51 \pm 0.018,^{ab}$	-8.58 ± 0.11	—	—
	SA-OMe	$0.55 \pm 0.03,^{ae}$	-8.54 ± 0.16	$-6.10 \pm 0.12,^{ef}$	-2.4 ± 0.2
HCA**--SSEG ₁₀ SA	Ethox	$6.1 \pm 0.01,^{aeg}$	-7.12 ± 0.08	$-11.0 \pm 0.6,^{ef}$	$+3.9 \pm 0.6$
	Intra ^h	33 ± 7^i	-6.1 ± 0.6	$-6.7 \pm 0.7,^{ef}$	$+0.6 \pm 0.9$

^aUnits of M.

^bMeasured by fluorescence either directly or through competition with DNSA.^{16, 17, 24, 25} Uncertainties are 95% confidence intervals from curve-fitting.

^cCompare to literature value of 0.3 μM for the binding of DNSA to a Cys-206→Ser mutant of HCA.²²

^dCompare to literature value of 0.2 nM for the binding of Ethox to bovine carbonic anhydrase II.¹⁷

^eMeasured by calorimetry.

^fUncertainties were assumed to be due primarily to errors in the quantitation of titrant.³⁰ An uncertainty of 2% of ΔH^0 was taken for the binding to HCA**--SCH₂CO₂⁻ and 5% of ΔH^0 for the binding to HCA**--SSEG₁₀SA.

^gCompare to value from fluorescence of 2.4 μM (Table 2).

^hThermodynamic parameters are those for the intramolecular equilibrium shown in Figure 1D.

ⁱDimensionless.

Table 3

Experimental and theoretical values for the intermolecular and intramolecular binding of sulfonamides to sulfonamide-conjugated HCA** proteins (HCA**-SSEG_nSA).

<i>n</i>	$K_d^{\text{comp.Ethox}} (\mu\text{M})^a$	$K_d^{\text{intra}} \times 10^{5b}$	$M_{\text{eff}}(\text{mM})^c$
0	0.31 ± 0.015	66 ± 14	0.8 ± 0.16
2	10 ± 0.7	2.0 ± 0.4	26 ± 5
5	5.6 ± 0.5	3.6 ± 0.8	14 ± 3
10	2.4 ± 0.17	8.3 ± 1.8	6.1 ± 1.3
20	1.2 ± 0.08	17 ± 4	3.1 ± 0.7

^a Observed dissociation constants for the binding of Ethox to HCA**-SSEG_nSA (Figure 5). Uncertainties are 95% confidence intervals from curve-fitting.

^b Calculated dissociation constants for the binding of the tethered ligand to the active site of HCA (Figure 5). Uncertainties were propagated from uncertainties in $K_d^{\text{comp.Ethox}}$ and K_d^{Ethox} (see Table 2).

^c Calculated using eq 1. Uncertainties were propagated from uncertainties in K_d^{intra} and K_d^{inter} (see Table 2).