Mapping Protein–Peptide Affinity: Binding of Peptidylsulfonamide Inhibitors to Human Carbonic Anhydrase II

Ann M. Cappalonga Bunn, Richard S. Alexander, and David W. Christianson*

Contribution from the Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323

Received November 24, 1992. Revised Manuscript Received December 20, 1993®

Abstract: The complexes of human carbonic anhydrase II (CA; EC 4.2.1.1) with the competitive inhibitors *N*-(4-sulfamylbenzoyl)glycine methyl ester (SBG), *N*-(4-sulfamylbenzoyl)benzylamine (SBB), and *N*-[*N*-[*N*-(4-sulfamylbenzoyl)glycyl]glycyl]glycine benzyl ester (SBG_3) have been studied by X-ray crystallographic methods at limiting resolutions of 2.0, 2.3, and 2.4 Å, respectively. Although the sulfonamide–zinc binding modes are similar in all three complexes, the hydrophobic tail of SBB and the oligoglycine tails of SBG and SBG_3 wrap across a hydrophobic patch in the enzyme active site with varying degrees of van der Waals contact surface area, and differences in enzyme–inhibitor contact are correlated with differences in affinity (Jain, A.; Huang, S. G.; Whitesides, G. M. *J. Am. Chem. Soc.*, preceding paper in this issue). Importantly, this work illuminates some of the structural determinants of affinity in the CA active site. Moreover, this work demonstrates that it is possible to design an oligopeptide which binds to a protein surface not predesigned to do so.

Human carbonic anhydrase II (CA; EC 4.2.1.1) is a zinc metalloenzyme containing one metal ion bound to a single polypeptide chain of 260 amino acids.[1] The conical active site of the enzyme is about 15-Å deep, and it is roughly divisible into a hydrophobic wall and a hydrophilic cleft. The catalytically required zinc ion resides at the bottom of this cleft, and it is here where the hydration of carbon dioxide into products bicarbonate ion and a proton is catalyzed at a rate near that of diffusion control. Although this ubiquitous enzyme and its isozymes are found in a variety of tissues, CA plays an important role in the eye, since its ocular activity is linked to elevated intraocular pressure, a symptom of glaucoma.[2] Importantly, it has been demonstrated that the inhibition of CA by sulfonamides such as Diamox (acetazolamide; Figure 1) is effective in glaucoma therapy, and this has highlighted the pharmaceutical importance of CA as a target for rational drug design.[3]

Given the substantial pharmaceutical interest in CA, it is desirable to explore new inhibitor designs which may ultimately lead to differing specificities with regard to carbonic anhydrase isozymes and tissue localizations. To date, drug design efforts have taken advantage of the three-dimensional structures of the native enzyme refined at 2.0-Å resolution[4,5] and its complexes with Diamox[4,6] and 3-(acetoxymercurio)-4-aminobenzenesulfonamide.[4,7] More recently, the three-dimensional structures of two thienothiopyran-2-sulfonamides and a thienothiophene-2-sulfonamide have been reported in their complexes with CA.[8]

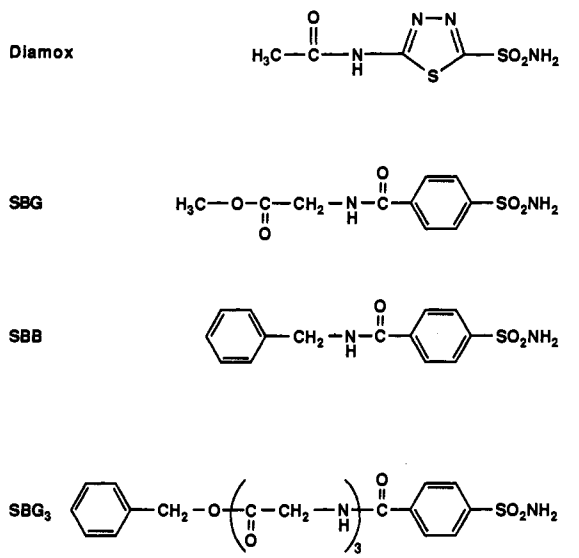

Figure 1. Molecular structures of Diamox (acetazolamide, K_i = 7 nM[3f]), *N*-(4-sulfamylbenzoyl)glycine methyl ester (SBG, K_d = 63 nM[9]), *N*-(4-sulfamylbenzoyl)benzylamine (SBB, K_d = 1.1 nM[9]), and *N*-[*N*-[*N*-(4-sulfamylbenzoyl)glycyl]glycyl]glycine benzyl ester (SBG_3, K_d = 75 nM[9]).

Here, we approach the structure-assisted design of CA inhibitors which are characterized by oligopeptidyl "tails" tethered to an arenesulfonamide group.[9] Three inhibitors are studied in their complexes with CA by X-ray crystallographic methods:

* Abstract published in *Advance ACS Abstracts*, April 1, 1994.

(1) (a) Coleman, J. E. *Annu. Rev. Pharmacol.* **1975**, *15*, 221–242. (b) Lindskog, S. In *Zinc Enzymes*; Bertini, I., Luchinat, C., Maret, W., Zeppezauer, M., Eds.; Birkhauser: Boston, 1986; pp 307–316. (c) Silverman, D. N.; Lindskog, S. *Acc. Chem. Res.* **1988**, *21*, 30–36. (d) Christianson, D. W. *Adv. Protein* Chem. **1991**, *42*, 281–355.

(2) (a) Kinsey, V. E. *Arch. Ophthalmol.* **1953**, *50*, 401–417. (b) Friedenwald, J. S. *Am. J. Ophthalmol.* **1949**, *32*, 9–27.

(3) (a) Mann, T.; Keilin, D. *Nature* **1940**, *146*, 164–165. (b) Becker, B. *Am. J. Ophthalmol.* **1954**, *37*, 13–15. (c) Grant, W. M.; Trotter, R. R. *Arch. Ophthalmol.* **1954**, *51*, 735–739. (d) Breinin, G. M.; Gortz, H. *Arch. Ophthalmol.* **1954**, *52*, 333–348. (e) Maren, T. H. *Drug. Dev. Res.* **1987**, *10*, 255–276. (f) Maren, T. H. In *Carbonic Anhydrase: From Biochemistry and Genetics to Physiology and Clinic Medicine*; Botre, F., Gros, G., Storey, B. T., Eds.; VCH Publishers: New York, **1991**; pp 186–207. (g) Maren, T. H. *Mol. Pharm.* **1992**, *41*, 419–426.

(4) Eriksson, A. E.; Jones, T. A.; Liljas, A. In *Zinc Enzymes*; Bertini, I., Luchinat, C., Maret, W., Zeppezauer, M., Eds.; Birkhauser: Boston, 1986; pp 317–328.

(5) Eriksson, A. E.; Jones, T. A.; Liljas, A. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 274–282.

(6) Eriksson, A. E.; Kylsten, P. M.; Jones, T. A.; Liljas, A. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 283–293.

(7) Vidgren, J.; Liljas, A.; Walker, N. P. C. *Int. J. Biol. Macromol.* **1990**, *12*, 342–344.

(8) (a) Baldwin, J. J.; Ponticello, G. S.; Anderson, P. S.; Christy, M. E.; Murcko, M. A.; Randall, W. C.; Schwam, H.; Sugrue, M. F.; Springer, J. P.; Gautheron, P.; Grove, J.; Mallorga, P.; Viader, M.-P.; McKeever, B. M.; Navia, M. A. *J. Med. Chem.* **1989**, *32*, 2510–2513. (b) Prugh, J. D.; Hartman, G. D.; Mallorga, P. J.; McKeever, B. M.; Michelson, S. R.; Murcko, M. A.; Schwam, H.; Smith, R. L.; Sondey, J. M.; Springer, J. P.; Sugrue, M. F. *J. Med. Chem.* **1991**, *34*, 1805–1818.

(9) (a) Jain, A.; Huang, S. G.; Whitesides, G. M. *J. Am. Chem. Soc.*, preceding paper in this issue. (b) Jain, A.; Whitesides, G. M.; Alexander, R. S.; Christianson, D. W. *J. Med. Chem.*, submitted.

0002-7863/94/1516-5063$04.50/0

Table 1. Data Collection and Refinement Statistics for Carbonic Anhydrase II–Inhibitor Complexes

	SBG	SBB	SBG_3
no. of crystals	4	2	1
no. of measured reflections	30 296	24 862	18 829
no. of unique reflections	12 877	7861	7967
max resolution (Å)	2.0	2.3	2.4
R_m[a]	0.085	0.082	0.080
no. of water molecules in final cycle of refinement	102	92	97
no. of reflections used in refinement (6.5 – max resolution, Å)	11 452	7233	7203
R factor[b]	0.173	0.158	0.200
RMS deviation from ideal bond lengths (Å)	0.020	0.024	0.011
RMS deviation from ideal bond angles (deg)	1.5	1.4	2.3
RMS deviation from ideal planarity (Å)	0.011	0.009	0.001
RMS deviation from ideal chirality (Å^3)	0.138	0.163	0.059

[a] R_{merge} for replicate reflections, $R = \sum|I_{hi} - \langle I_h\rangle|/\sum\langle I_h\rangle$; I_{hi} = intensity measured for reflection h in data set i. $\langle I_h\rangle$ = average intensity for reflection h from replicate data. [b] Crystallographic R factor, $R = \sum||F_o| - |F_c||/\sum|F_o|$; $|F_o|$ and $|F_c|$ are the observed and calculated structure factors, respectively.

N-(4-sulfamylbenzoyl)glycine methyl ester (SBG, K_d = 63 nM), *N*-(4-sulfamylbenzoyl)benzylamine (SBB, K_d = 1.1 nM; the benzylamine group is an analogue of glycine methyl ester), and *N*-[*N*-[*N*-(4-sulfamylbenzoyl)-glycyl]glycyl]glycine benzyl ester (SBG_3, K_d = 75 nM). The molecular structures of these compounds, along with the structure of Diamox (K_i = 7 nM[3f]) for reference, are illustrated in Figure 1. With the arenesulfonamide group of each inhibitor anchored to the active-site zinc ion, successively longer oligopeptidyl tails increase the potential contact surface area of the protein–peptide interface. The goal of this work is to design oligopeptides which bind to a protein surface not predesigned by Nature to do so and to illuminate the structural features that govern affinity by X-ray crystallographic analysis of protein–peptide complexes. As such, this work represents the first step of a structure-assisted approach toward the design of *de novo* protein–peptide complexes.

Experimental Section

The CA inhibitors SBG, SBB, and SBG_3 were generously provided by Professor George M. Whitesides, Harvard University, and native blood CA was purchased from Sigma and used without further purification. Crystals were grown by the sitting drop method, and enzyme–inhibitor cocrystallizations required the addition of a 5–10-μL drop containing the enzyme–inhibitor complex [0.3 mM CA, 50 mM Tris-HCl (pH 8.0 at room temperature), 150 mM NaCl, 3 mM NaN_3, and inhibitor (4.0 mM SBG or 4.3 mM SBB or ca. 1 mM SBG_3 dissolved in DMSO such that the final concentration of DMSO in the crystallization well was 5% (vol/vol))] to a 5–10-μL drop of precipitant buffer [50 mM Tris-HCl (pH 8.0 at room temperature), 150 mM NaCl, 3 mM NaN_3 with 1.75–2.5 M ammonium sulfate]; additionally, the surrounding crystallization well was charged with 1 mL of precipitant buffer. We note that the crystallization of the CA–SBG_3 complex required an adjustment of the sodium chloride concentration to 10 mM and the addition of 1 mM β-octyl glucoside[10] to the crystallization buffer.

Both enzyme and precipitant drops were saturated with methyl mercurioacetate in order to promote the growth of diffraction-quality parallelepipedons.[11] Crystals of typical dimensions 0.2 mm × 0.2 mm × 0.8 mm appeared within 2 weeks at 4 °C. Enzyme–inhibitor complexes crystallized in space group $P2_1$ and exhibited typical unit cell parameters of a = 42.7 Å, b = 41.7 Å, c = 73.0 Å, and β = 104.6°. Importantly, crystals of each enzyme–inhibitor complex were isomorphous with those obtained from the native enzyme.

Crystals of CA–inhibitor complexes were mounted and sealed in 0.5-mm glass capillaries with a small portion of mother liquor. A Siemens X-100A multiwire area detector, mounted on a three-axis camera and equipped with Charles Supper double X-ray focusing mirrors, was used for X-ray data acquisition. A Rigaku RU-200 rotating anode X-ray generator operating at 45 kV/65 mA supplied Cu Kα radiation. All data were collected at room temperature by the oscillation method; the crystal-to-detector distance was set at 14 cm, and the detector swing angle was fixed at 20° in at least four runs per experiment. Data frames of 0.0833° oscillation about ω were collected, with exposure times of 60 s/frame, for total angular rotation ranges about ω of at least 70° per run. Multiple data sets for each enzyme–inhibitor complex were collected, and diffraction intensities were measured to limiting resolutions of 2.0–2.4 Å. Raw data frames were analyzed using the BUDDHA package,[12] reflections with $I < 2\sigma$ were discarded, and replicate and symmetry-related structure factors were merged using PROTEIN;[13] final merging R factors for each enzyme–inhibitor complex, as well as other relevant data reduction statistics, are recorded in Table 1.

In the analysis of CA–inhibitor crystal structures, structure factors obtained from the corrected intensity data were used to generate difference electron density maps using Fourier coefficients $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ with phase calculated from the structure of refined human CA[4,5] (atomic coordinates were obtained from the Brookhaven Protein Data Bank[14]). Fast Fourier transform routines were employed in all electron density map and structure factor calculations.[15,16] Inspection of the electron density maps revealed that only minor adjustments to the protein model were required, and model building was performed with the graphics software FRODO[17] installed on an Evans and Sutherland PS390 interfaced with a VAXstation 3500. Atomic coordinates were then refined against the observed data by the reciprocal space least-squares method using the stereochemically restrained least-squares algorithm of Hendrickson and Konnert.[18,19] Neither the inhibitor atoms nor active-site water molecules were included in the initial stages of refinement for each enzyme–inhibitor structure. Residue conformations throughout the protein were examined during the course of refinement by using maps calculated with Fourier coefficients outlined above and phases derived from the in-progress atomic model. Only minimal adjustments of atomic positions were necessary, and the inhibitor, as well as active-site water molecules, was added when the crystallographic R factor dropped below 0.200 in the refinement of each enzyme–inhibitor complex. In the case of SBG_3, no density was observed for the C-terminal benzyl ester, so its occupancy was set to zero. Due to the noisy quality of electron density maps of the CA–SBG_3 complex, stereochemical restraints had to be additionally weighted during refinement in order to maintain good molecular geometry. For each enzyme–inhibitor complex, refinement converged smoothly to final crystallographic R factors of 0.15–0.20; each final model had excellent stereochemistry with root mean square deviations from ideal bond lengths and angles of 0.011–0.024 Å and 1.4–2.3°, respectively. Pertinent refinement statistics are recorded in Table 1.

For each refined CA–inhibitor structure, a difference electron density map calculated with Fourier coefficients $|F_o| - |F_c|$ and phases derived from the coordinates of the final model revealed that the highest peaks in the vicinity of the active site were under 3.5σ. Additionally, the root mean square error in atomic positions for each structure was estimated

(10) McPherson, A.; Koszelak, S.; Axelrod, H.; Day, J.; Williams, R.; Robinson, L.; McGrath, M.; Cascio, D. *J. Biol. Chem.* **1986**, *261*, 1969–1975.
(11) Tilander, B.; Strandberg, B.; Fridborg, K. *J. Mol. Biol.* **1965**, *12*, 740–760.
(12) Durbin, R. M.; Burns, R.; Moulai, J.; Metcalf, P.; Freymann, D.; Blum, M.; Anderson, J. E.; Harrison, S. C.; Wiley, D. C. *Science* **1986**, *232*, 1127–1132.
(13) Steigemann, W. Ph.D. Thesis, Max Plank Institut fur Biochemie, 1974.
(14) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535–542.
(15) Ten Eyck, L. F. *Acta Crystallogr., Sect. A* **1973**, *29*, 183–191.
(16) Ten Eyck, L. F. *Acta Crystallogr., Sect. A* **1977**, *33*, 486–492.
(17) Jones, T. A. *Methods Enzymol.* **1985**, *115*, 157–171.
(18) Hendrickson, W. A.; Konnert, J. H. In *Biomolecular Structure, Conformation, Function and Evolution*; Srinivasan, R., Ed.; Pergamon: Oxford, U.K., 1981; Vol. 1, pp 43–47.
(19) Hendrickson, W. A. *Methods Enzymol.* **1985**, *115*, 252–270.

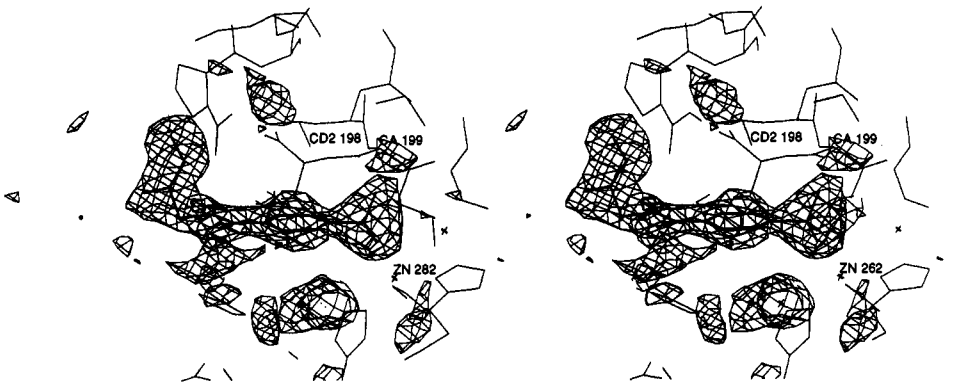

Figure 2. Difference electron density map of the CA–SBG complex, calculated with Fourier coefficients $|F_o| - |F_c|$ and phases derived from the final model less the inhibitor and active-site solvent molecules. The map is contoured at 2.5σ, and refined atomic coordinates are superimposed; Leu-198, Thr-199, and zinc are indicated.

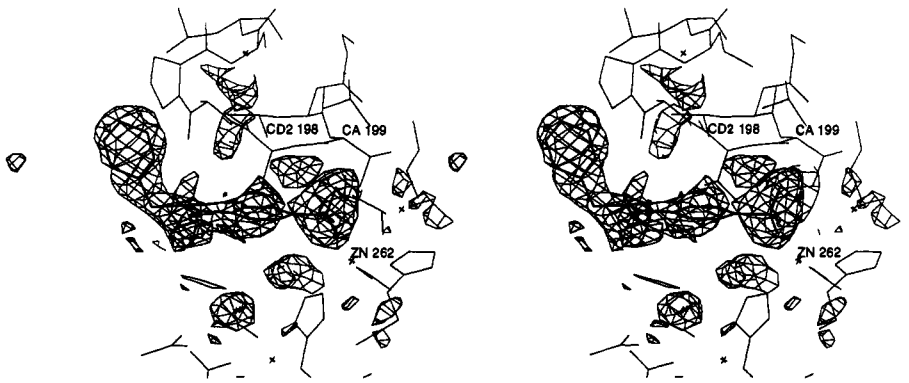

Figure 3. Difference electron density map of the CA–SBB complex, calculated with Fourier coefficients $|F_o| - |F_c|$ and phases derived from the final model less the inhibitor and active-site solvent molecules. Refined atomic coordinates are superimposed on the map (contoured at 2.1σ); Leu-198, Thr-199, and zinc are indicated.

to be ca. 0.25 Å on the basis of relationships derived by Luzzati.[20] The coordinates of each enzyme–inhibitor complex have been deposited in the Brookhaven Protein Data Bank.[14]

Results

The structure of CA in its complexes with SBG, SBB, or SBG_3 exhibits only minor differences when compared with uncomplexed CA, and these differences are largely confined to the side chains of occasional surface residues. Difference electron density maps of the CA–SBG and CA–SBB complexes are quite clear and reveal well-defined density for all inhibitors atoms (Figures 2 and 3, respectively). A difference electron density map of the CA–SBG_3 complex is less clear, but the course of the triglycyl tail of the inhibitor is nonetheless visible (Figure 4). No electron density is observed for the C-terminal benzyl ester of SBG_3, nor is density observed for the carbonyl oxygen of inhibitor Gly-1 and the backbone NH group of inhibitor Gly-3. However, the conformation of inhibitor Gly-1 in SBG_3 was satisfactorily modeled by using the single, well-characterized glycine residue observed in the CA–SBG complex as a guide.

Selected enzyme contacts with the three inhibitors are recorded in Tables 2–4. In each enzyme–inhibitor complex, an ionized inhibitor sulfonamide nitrogen displaces the zinc-bound hydroxide ion of the native enzyme such that tetrahedral metal coordination geometry is maintained. Each zinc-bound sulfonamide nitrogen in turn donates a hydrogen bond to the γ-hydroxyl group of Thr-199, which in turn donates a hydrogen bond to the ionized carboxylate of Gly-106. An additional solvent molecule is displaced by a sulfonamide oxygen atom in each enzyme–inhibitor complex. This water molecule is the so-called "deep" water molecule which resides at the mouth of the hydrophobic pocket in the native enzyme.[4,5] These features are in accord with the binding modes observed in other CA–sulfonamide complexes.[4,6-8]

Although the arenesulfonamide groups in the CA complexes with SBG, SBB, and SBG_3 are essentially superimposable, the aryl rings of each inhibitor are not coplanar with the thiadiazole ring of the prototypical CA inhibitor Diamox, as it is observed to bind in the enzyme active site; instead, the aryl rings are each tilted about an axis defined by the para substitutents by ca. 15°. To illustrate, the atomic coordinates of enzyme-bound Diamox are superimposed on the CA–SBG complex in Figure 5. The para substituted tails of the three inhibitors pack against the hydrophobic wall in the active-site cavity in a region largely defined by the side chains of Leu-198, Pro-201, and Pro-202. As seen in Figure 5, the peptidyl tail of SBG (and also SBB and SBG_3) makes different contacts with the enzyme site as compared with the acetamido tail of Diamox. The acetamido carbonyl group of Diamox accepts a hydrogen bond from the side chain nitrogen of Gln-92, and the acetamido methyl group is directed out toward the hydrophilic cleft. Considering the flexibility and the hydrophobic nature of the inhibitors SBG, SBB, and SBG_3, it is not too surprising that they most optimally associate with a hydrophobic patch, instead of binding across the hydrophilic cleft, in order to minimize their solvent-accessible surface area. Even where Nature provides a predesigned peptide binding cleft, such as in the zinc proteases carboxypeptidase A and thermolysin, it is found that inhibitors bearing N-terminal Cbz–Gly groups prefer to associate with hydrophobic clefts in the S_1 subsites rather than the well-defined peptide binding grooves of the protease active sites.[21,22]

(20) Luzzati, P. V. *Acta Crystallogr.* **1952**, *5*, 802–810.

(21) Holden, H. M.; Tronrud, D. E.; Monzingo, A. F.; Weaver, L. H.; Matthews, B. W. *Biochemistry* **1987**, *26*, 8542–8553.

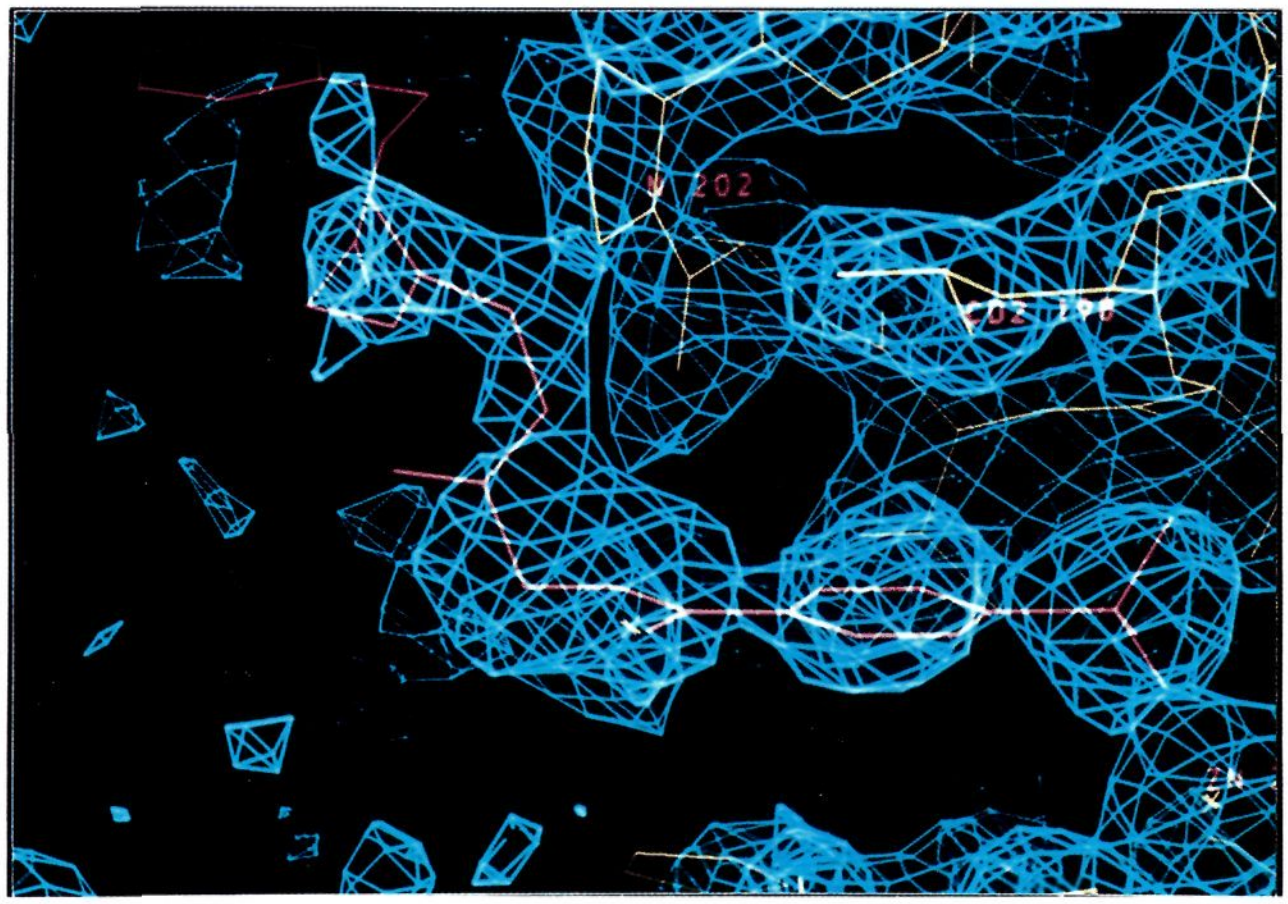

Figure 4. Difference electron density map of CA–SBG_3, calculated with Fourier coefficients $2|F_o| - |F_c|$ and phases calculated from the final model less the active-site solvent molecules and the inhibitor. Refined atomic coordinates are superimposed on the map, and Leu-198, Pro-202, and zinc are indicated. The inhibitor is represented by a red stick-figure model, and enzyme atoms are represented by a yellow stick-figure model.

Table 2. Selected Carbonic Anhydrase II–SBG Interactions[a]

inhibitor atom	enzyme residue	separation (Å)
sulfonamide O1	Thr-199 NH	3.2*
sulfonamide O2	Zn^{2+}	3.2
sulfonamide NH	Zn^{2+}	2.2
	Thr-199 O_γ	3.1*
ester O	Phe-131 C_ζ	3.6

[a] An asterisk denotes a possible hydrogen bond as judged from distance and stereochemical criteria.

Table 3. Selected Carbonic Anhydrase II–SBB Interactions[a]

inhibitor atom	enzyme residue	separation (Å)
sulfonamide O1	Thr-199 NH	3.6*
sulfonamide O2	Zn^{2+}	3.2
sulfonamide NH	Zn^{2+}	2.0
	Thr-199 O_γ	3.2*
ester O	Phe-131 $C_{\epsilon 1}$	3.6
	Phe-131 C_ζ	3.1

[a] An asterisk denotes a possible hydrogen bond as judged from distance and stereochemical criteria.

The inhibitor SBB binds more tightly than SBG to CA.[9] However, the glycine methyl ester of SBG occupies the same hydrophobic patch in the enzyme active site as the benzylamine group of SBB (Figure 6). Given that a methyl-protected glycine is somewhat hydrophobic in nature, it is reasonable to consider that a benzylamine group is a less flexible and more hydrophobic analogue of glycine methyl ester, if steric constraints permit. Correspondingly, the K_d values measured for SBG and SBB indicate that a benzylamine group is the preferred occupant of a hydrophobic cleft or patch. This analogy may be exploited in biomolecular design experiments with other peptide mimics, particularly where ligand–receptor affinity is substantially modulated by the hydrophobic effect: substitution of a benzylamine group for glycine may yield a binding enhancement of up to ca. 60-fold.

The triglycyl tail of SBG_3 wraps even further around the hydrophobic wall of the active-site cavity as compared with the case of its shorter analogue SBG. This result is not surprising

(22) Christianson, D. W.; Lipscomb, W. N. *J. Am. Chem. Soc.* **1988**, *110*, 5560–5565.

Table 4. Selected Carbonic Anhydrase II–SBG_3 Interactions[a]

inhibitor atom	enzyme residue	separation (Å)
sulfonamide O2	Zn^{2+}	2.9
sulfonamide O1	Thr-199 NH	3.2*
sulfonamide N1	Zn^{2+}	2.5
	Thr-199 O_γ	2.6*
	Thr-199 NH	3.6
ester O	Phe-131 C_ζ	3.3
	Phe-131 $C_{\epsilon 1}$	3.6
peptide Gly-1 NH	OHH	3.5
peptide Gly-3 NH	Pro-202 C_δ	3.3
	Pro-202 C_γ	3.5
peptide Gly-3 O	Pro-202 C_δ	3.5

[a] An asterisk denotes a possible hydrogen bond as judged from distance and stereochemical criteria.

considering the somewhat hydrophobic nature of the amino acid glycine; the surprising result is that the tripeptidyl inhibitor binds with good affinity yet denies itself of any possible intermolecular hydrogen bond interactions in the hydrophilic cleft. Here, too, enzyme–inhibitor association appears to be driven by the hydrophobic effect where the inhibitor binds so as to minimize its solvent-accessible surface area. The NH group of inhibitor Gly-1 makes a 3.5 Å contact with a solvent molecule, but this interaction exhibits poor hydrogen bond stereochemistry. Additionally, a close intrainhibitor contact is observed in enzyme-bound SBG_3, but this interaction is similarly characterized by poor hydrogen bond stereochemistry as judged by distance and stereochemical criteria.[23] Moreover, this interaction involves an atom which is not clearly defined in electron density maps of the CA–SBG_3 complex, the carbonyl oxygen of inhibitor Gly-1; therefore, a conservative interpretation is warranted. The backbone torsion angles of inhibitor glycine residues are as follows: $\phi_1 = 96°$, $\psi_1 = -19°$; $\phi_2 = -34°$, $\psi_2 = -15°$; and $\phi_3 = 97°$, $\psi_3 = -73°$. A search of main chain torsion angles characterizing protein structures contained in the Brookhaven Protein Data Bank[14] reveals no identical di- or tripeptide conformations which correspond to the inhibitor conformation. We note that the longest glycine polypeptide contained in reported protein structure

(23) (a) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97–179. (b) Ippolito, J. A.; Alexander, R. S.; Christianson, D. W. *J. Mol. Biol.* **1990**, *215*, 457–471.

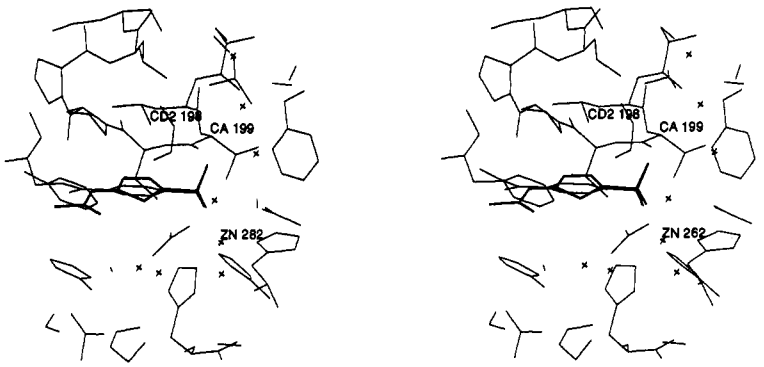

Figure 5. Superposition of the CA–Diamox and CA–SBG complexes (dark bonds and light bonds, respectively; the coordinates of the CA–Diamox complex were generously provided by Drs. Brian McKeever and James Springer, Merck, Sharp, and Dohme Research Laboratories, Rahway, NJ). Enzyme residues Leu-198 and Thr-199 and the active-site zinc ion are indicated. Note that the phenyl ring of SBG is almost coplanar (within ca. 15°) with the thiadiazole ring of Diamox as it is bound in the active site.

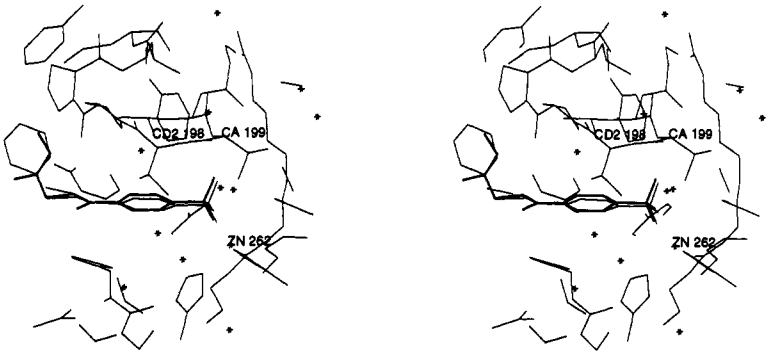

Figure 6. Superposition of the bound conformation of SBG (dark bonds) on the CA–SBB complex (light bonds). Enzyme residues Leu-198 and Thr-199 and the active-site zinc ion are indicated. Note that the benzylamine group of SBB binds in roughly the same manner as the methyl-protected glycine group of SBG.

coordinates is found in rizopuspepsin, where a glycine tetrapeptide is part of a β-sheet–β-turn motif.

The interpretation of peptidylsulfonamide dissociation constants in light of the structural data acquired from X-ray crystallographic studies indicates that, in a series of oligoglycine CA inhibitors synthesized by Jain and colleagues,[9] enzyme–inhibitor affinity is governed by the hydrophobic effect and attenuated by entropic effects. To illustrate, the triglycyl inhibitor SBG_3 binds to CA with affinity comparable to that of the monoglycyl inhibitor SBG.[9] The van der Waals contact surface areas for the peptidyl tails of SBG and SBG_3 in their complexes with CA are 137 Å^2 and 235 Å^2, respectively. Since the disordered benzyl group of SBG_3 does not contribute to the enzyme–inhibitor interface, we estimate that each additional glycine residue contributes an additional 49 Å^2 to the contact surface area of the CA–SBG_3 complex relative to the CA–SBG complex. Since no hydrogen-bond interactions occur between the enzyme and the peptidyl tails of each inhibitor, hydrophobic van der Waals contacts must govern enzyme–inhibitor association. Since glycine is somewhat hydrophobic in nature, and since the free energy of transfer of a hydrophobic group out of aqueous solution is reasonably estimated to be at least –0.02 kcal mol^{-1} Å^{-2} (some calculations suggest that this value may approach –0.047 kcal mol^{-1} Å^{-2}),[24] the additional contact surface area of SBG_3 relative to SBG should contribute ca. –2 kcal mol^{-1} to enzyme–inhibitor affinity.

However, the entropic cost of freezing the conformations of two additional glycyl groups of SBG_3 offsets the favorable free energy of increased contact surface area relative to the case of SBG. The cost of freezing a single glycyl group in a particular conformation is about +1 kcal mol^{-1};[25] it follows that the entropic cost of freezing two glycyl groups is about +2 kcal mol^{-1}. Hence, in the CA–SBG_3 complex the favorable free energy of at least –2 kcal mol^{-1} resulting from increased enzyme–inhibitor contact surface area is nearly perfectly attenuated by the entropic cost of +2 kcal mol^{-1} required to maintain that contact surface area.

The crystallographic disorder observed for the C-terminal benzyl ester of SBG_3 is consistent with the results of ^{1}H NMR spin echo spectroscopic studies of oligoglycine inhibitors in their complexes with CA.[9] Dipole–dipole interactions result in spin relaxation for Cα methylene protons of inhibitor residues Gly-1 through Gly-3 of a hexaglycyl sulfonamide bound in the CA active site, such that $T_2 < 20$ ms. However, inhibitor residues Gly-4 through Gly-6 do not experience relaxation ($T_2 \geq 175$ ms), thereby implying that these residues are disordered relative to the first three residues of the inhibitor. Presumably, these residues extend out of the active-site cleft into solution and cannot achieve the minimum contact surface area with the protein surface required to fix their conformation. Given that the structures of the CA–SBG and CA–SBB complexes reveal that the benzylamine (and, by analogy, the benzyloxy) group is an analogue of glycine, and since the C-terminal benzyloxy protecting group of SBG_3 is disordered, it follows that disorder is likewise expected for Gly-4 through Gly-6 of the longer oligoglycine inhibitor.

Three viable routes remain toward improving enzyme–inhibitor affinity for this series of peptidyl sulfonamide inibitors: (1) engineer stronger, more polar interactions, such as hydrogen bonds, between the enzyme and inhibitor; (2) minimize the entropic cost of maintaining an inhibitor conformation by incorporating an amino acid (or an amino acid analogue such as the benzylamine group) less flexible than glycine in the oligopeptidyl tail of the inhibitor; and (3) maximize the amount of

(24) (a) Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151–176. (b) Sharp, K. A.; Nicholls, A.; Friedman, R.; Honig, B. *Biochemistry* **1991**, *30*, 9686–9697.

(25) Nemethy, G.; Leach, S. J.; Scheraga, H. A. *J. Phys. Chem.* **1966**, *70*, 998–1004.

enzyme–inhibitor contact surface area attainable per conformational degree of freedom. For instance, inhibitors SBG and SBB have the same conformational degrees of freedom which must be frozen out for enzyme complexation. However, the benzylamine tail of SBB makes 151 Å^2, whereas the glycine methyl ester tail of SBG makes only 137 Å^2, of contact surface area with CA. Thus, SBB offers more contact surface area per conformational degree of freedom, and enhanced contact surface area presumably contributes to enhanced enzyme–inhibitor affinity.

Parenthetically, we note that oligoglycine sequences may play important structural and functional roles in other systems, and the structure of the CA–SBG_3 complex provides the first model of the intermolecular interactions of oligoglycine moieties. Quasi-repetitive oligoglycine "loops" have been identified by sequence analysis in filament proteins such as keratins and nuclear lamins, loricrins, and single-stranded RNA-binding proteins.[26] Generally, these loops are comprised of as many as 35 glycine residues (occasionally interspersed by serine, asparagine, arginine, or cysteine residues) and are flanked by aromatic or long-chain aliphatic residues such as valine or leucine. Although the molecular function of these loops remains unknown, it is hypothesized that they mediate intermolecular interactions by the hydrophobic effect.[26] Additionally, a recent report describes the construction of bis-sialosides which bind to the influenza virus.[27] The separation of the two sialic acid groups in the best inhibitor of hemagglutination, a bivalent sialoside, is 57 Å. Presumably, the oligoglycine portions of the inhibitor associate with hydrophobic patches on the surfaces of two trimers in a proposed $trimer_2$–inhibitor complex. The protein–oligoglycine binding mode observed in the CA–SBG_3 complex illustrates structural aspects that are applicable to these systems.

Summary and Conclusions

It is possible to design specific peptides which bind to protein surfaces not predesigned or preevolved to do so. The arenesulfonamide groups of the enzyme inhibitors SBG, SBB, and SBG_3 bind to the active site of CA in a similar manner as determined by X-ray crystallographic methods. However, the peptide (or peptide analogue) tails of these inhibitors associate with a hydrophobic patch in the enzyme active site and do not engage in any intermolecular hydrogen bonds with the enzyme. Second-generation peptidyl sulfonamides have been designed following the structural inspection of these enzyme–inhibitor complexes, and successful leads optimize hydrophobic interactions and actually steer the oligopeptide across different regions of the active site; this work will be reported separately.[9b]

In principle, the structural analysis of a protein–oligoglycine complex should lead to amino acid substitutions in the oligopeptide that will enhance protein–peptide affinity. This structure-assisted rationale comprises a general approach toward the design of an oligopeptide which spans two preselected points on a protein surface: (1) anchor an oligoglycine peptide between the two points on the protein surface and determine the structure of the protein–oligoglycine complex; (2) make rational substitutions in the amino acid sequence of the oligopeptide which optimize its contact with the protein surface; and (3) remove the anchoring functionalities from the optimized oligopeptide (if desired). We acknowledge that it is possible that amino acid substitutions made on the first-generation oligoglycine peptide may alter the preferred binding mode of the peptide. However, this illuminates an alternative design rationale: (1) anchor a single amino acid to a preselected point on a protein surface; and (2) construct an oligopeptide such that the interactions of its amino acid side chains with the protein steer the main chain of the oligomer across the protein surface to a second preselected point. Further biophysical and X-ray crystallographic experiments are necessary to evaluate these approaches.

In summary, this structure-assisted approach toward the design of peptide-based CA inhibitors may be applied toward the design of *de novo* protein–peptide complexes in other systems. We note that this rational approach is complementary to random approaches exploited in the optimization of antibody–peptide complexes,[28] and future work will probe the degree to which these complementary appraches can be combined.

Acknowledgment. We thank Profs. C. A. Fierke, V. J. Hruby, and T. H. Maren for helpful comments, and we thank Drs. Brian McKeever and James Springer for the atomic coordinates of the CA–Diamox complex. Additionally, we thank the National Institutes of Health for Grant GM45614 in support of this work. D.W.C. is an Alfred P. Sloan Research Fellow and a Camille and Henry Dreyfus Teacher–Scholar.

(26) Steinert, P. M.; Mack, J. W.; Korge, B. P.; Gan, S.-Q.; Haynes, S. R.; Steven, A. C. *Int. J. Biol. Macromol.* **1991**, *13*, 130–139.

(27) Glick, G. D.; Knowles, J. R. *J. Am. Chem. Soc.* **1991**, *113*, 4701–4703.

(28) (a) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82–84. (b) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84–86.