Variation in Strength of an Unconventional C−H to O Hydrogen Bond in an Engineered Protein Cavity

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · SEPTEMBER 1997

Impact Factor: 12.11 · DOI: 10.1021/ja9716766

5 AUTHORS, INCLUDING:

Rabi Ann Musah
Albany State University
42 PUBLICATIONS **1,170** CITATIONS
SEE PROFILE

Robin J Rosenfeld
The Scripps Research Institute
21 PUBLICATIONS **861** CITATIONS
SEE PROFILE

Duncan E Mcree
The Scripps Research Institute
110 PUBLICATIONS **9,102** CITATIONS
SEE PROFILE

Variation in Strength of an Unconventional C−H to O Hydrogen Bond in an Engineered Protein Cavity

Rabi A. Musah, Gerard M. Jensen, Robin J. Rosenfeld, Duncan E. McRee, and David B. Goodin*

Department of Molecular Biology, MB8, The Scripps Research Institute, 10550 North Torrey Pines Road La Jolla, California 92037

Steven W. Bunte

U.S. Army Research Laboratory Aberdeen Proving Ground, Maryland 21005-5066

Received May 22, 1997

While hydrogen bonds of the type XH···Y (X, Y = N and/or O) are found in essentially all macromolecular structures, the role of CH···O hydrogen bonds in such systems is uncertain. Recent reports[1−5] have suggested that a significant number of the CH···O contacts observed in proteins, RNA, and carbohydrates represent cohesive interactions, which may, in sufficient numbers, be significant to structure, stability, and function. The existence of short intermolecular CH···O interactions is well established in many small molecule crystals.[6−9] In cases where the C−H−O angle is approximately linear and the C···O bond distance is less than the combined van der Waals radii (3.3 Å), these interactions have been labeled as hydrogen bonding. Theoretical calculations have estimated the CH···O bonding interactions to be worth 1−2 kcal/mol.[2,10,11] However, many computational approaches utilizing force field parameterizations do not explicitly consider such interactions as attractive, and CH···O hydrogen bonds are not generally considered in the analysis of structures. The strength of such interactions should vary considerably depending on the C−H bond polarity and may thus be different for the C_αH···O interactions observed in some protein β-sheets than those for others involving the more acidic C_ϵH protons of histidine.[1] While there is a correlation of the C···O distance with the C−H bond acidity for a number of organic compounds,[6,12] nothing is known experimentally about the strength of these interactions and how this strength varies with the polarity of the proton donor.

We have utilized the ligand binding properties of a buried cavity created in the interior of a protein to obtain direct information about the variation in the strength of CH···O interactions between ligand and protein. The cavity created by the W191G mutation of cytochrome *c* peroxidase (CCP) has been shown to bind a number of cationic heterocyclic compounds.[13−15] Two such compounds, 2,3,4-trimethylthiazole (234TMT) and 3,4,5-trimethylthiazole (345TMT) bind to the W191G cavity in essentially isosteric conformations (Figure 1).[16] Crystal structures of the protein soaked in solutions containing these compounds show clear evidence for binding in the omit electron density maps. No significant differences were observed in the structure of the protein between the ligand-free and either of the ligand bound states. Placement of a ligand-model within

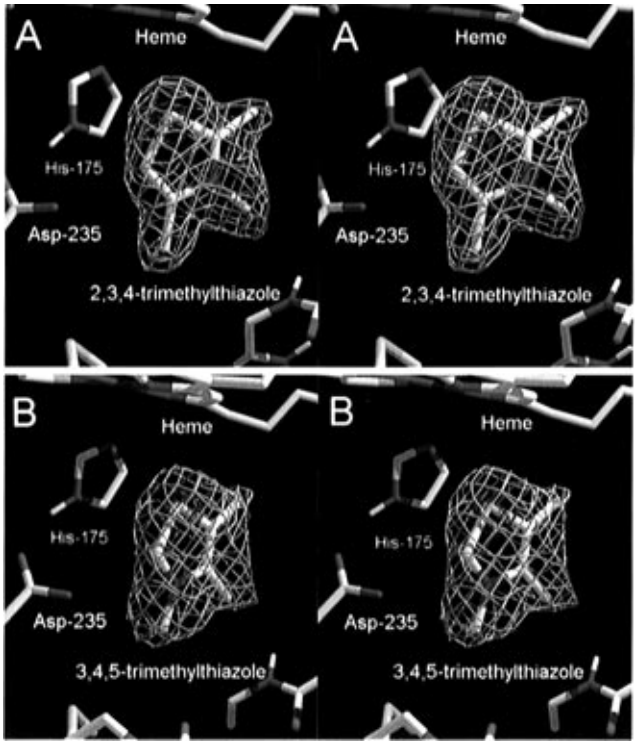

Figure 1. Electron density for 234TMT (A) and 345TMT (B) binding to the W191G cavity. Shown are stereoviews of $F_o - F_c$ Fourier omit maps contoured at +4 σ (white) and +9σ (red) superimposed on a model of the ligand which was placed into the omit density.[16]

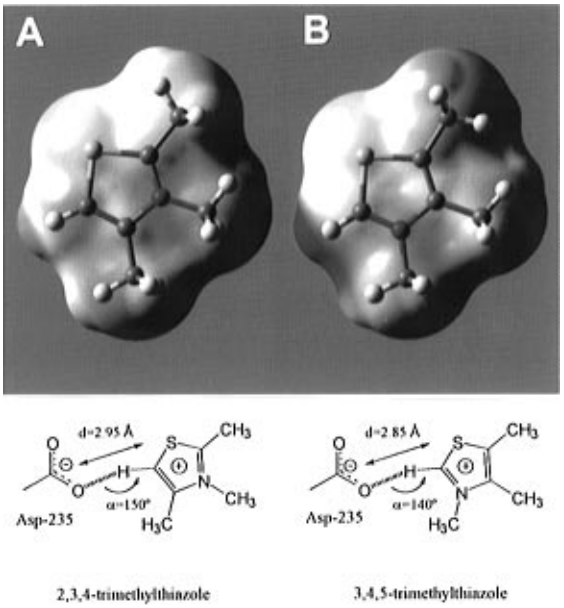

Figure 2. Electrostatic potential surfaces for 234TMT (A) and 345TMT (B). The color map illustrates the localized differences resulting from the increased polarity of the C2-H proton for 345TMT relative to the C5−H bond of 234TMT.[18] At the bottom are the geometries of the interactions with Asp-235 that are inferred from the placement of the geometry-optimized ligand into the omit electron density.

(1) Wahl, M. C.; Sundaralingam, M. *Trends Biochem. Sci.* **1997**, *22*, 97−102.
(2) Ornstein, R. L.; Zheng, Y.-J. *J. Biomol. Struct. Dyn.* **1997**, *14*.
(3) Bella, J.; Berman, H. M. *J. Mol. Biol.* **1996**, *264*, 734−742.
(4) Derewenda, Z. S.; Derewenda, U.; Kobos, P. M. *J. Mol. Biol.* **1994**, *241*, 83−93.
(5) Derewenda, Z. S.; Lee, L.; Derewenda, U. *J. Mol. Biol.* **1995**, *252*, 248−262.
(6) Sutor, D. J. *Nature* **1962**, *193*, 68−69.
(7) Hamilton, W. C.; Ibers, J. A. *Hydrogen Bonding in Solids*; Benjamin: New York, 1968.
(8) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 5063−5070.
(9) Steiner, T.; Saenger, W. *J. Am. Chem. Soc.* **1992**, *114*, 10146−10154.
(10) Kollman, P.; McKelvey, J.; Johansson, A.; Rothenberg, S. *J. Am. Chem. Soc.* **1975**, *97*, 955−965.
(11) Seiler, P.; Weisman, G. R.; Glendening, E. D.; Weinhold, T.; Johnson, V. B.; Dunitz, J. D. *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 1175−1177.
(12) (a) Desiraju, G. R. *J. Chem. Soc., Chem. Commun.* **1989**, *3*, 179−180. (b) Desiraju, G. R. *Acc. Chem. Res.* **1996**, *29*, 441−449. (c) Desiraju, G. R. *J. Chem. Soc., Chem. Commun.* **1990**, *6*, 454.
(13) Fitzgerald, M. M.; Trester, M. L.; Jensen, G. M.; McRee, D. E.; Goodin, D. B. *Protein Sci.* **1995**, *4*, 1844−1850.
(14) Fitzgerald, M. M.; Musah, R. A.; McRee, D. E.; Goodin, D. B. *Nat. Struct. Biol.* **1996**, *3*, 626−31.
(15) Fitzgerald, M. M.; Churchill, M. J.; McRee, D. E.; Goodin, D. B. *Biochemistry* **1994**, *33*, 3807−3818.

S0002-7863(97)01676-4 CCC: $14.00

Table 1. Thermodynamic Parameters for Trimethylthiazole Binding to the Buried Cavity of W191G[22]

	K_d (mM)	ΔH (kcal mol^{-1})	ΔS (cal mol^{-1} K^{-1})	ΔG (kcal mol^{-1})
234TMT	1.50 (0.16)	−14 (0.6)	−33 (2)	−3.9 (0.1)
345TMT	0.20 (0.03)	−15 (0.7)	−34 (3)	−5.1 (0.2)

the omit density implies a close contact between the C5 ring proton of 234TMT and one carboxylate oxygen of Asp-235. An analogous interaction involving the C2 ring proton is observed for 345TMT. Asp-235 is observed to hydrogen bond to other cavity binding ligands and helps determine the cation binding specificity of this cavity.[13,15,17] Thus, the absence of standard hydrogen-bond donors in 234TMT and 345TMT evidently results in the substitution of weaker alternative interactions that fulfill a similar role.

The geometries of the interactions implied by the crystal structures are consistent with a CH···O hydrogen bond.[8] The C···O distances between the ligands and the Asp-235 carboxylate oxygen were 2.95 and 2.85 Å for 234TMT and 345TMT, respectively, shorter than the combined van der Waals radii of C and O (3.3 Å). The C−H−O angles estimated from placing the geometrically optimized ligands[18] into the electron density were approximately 150° and 140° for 234TMT and 345TMT, respectively (Figure 2). The deviations of these angles from an ideal linear CH···O hydrogen bond are not outside the range (130−170°) of those observed in organic crystals.[8] Although the two compounds differ structurally only by the interchange of the N3 and C4 atoms in the thiazole ring, this difference serves to increase the acidity of the C2 proton in 345TMT (pK_a = 17−19),[19] relative to the C5 proton in 234TMT (p$K_a \approx$ 25−30). This difference in acidities is manifested in the electrostatic potential calculated for the two compounds and mapped onto the solvent accessible surface (Figure 2). Since the CH···O hydrogen bond is primarily electrostatic in nature, the calculated potentials indicate that 345TMT should form a significantly stronger interaction with Asp-235.

Thermodynamic parameters of 234TMT and 345TMT binding to W191G were determined by isothermal titration calorimetry (Table 1). The results show a 5-fold decrease in K_d for 345TMT relative to 234TMT, corresponding to 1.2 kcal/mol of additional binding energy for 345TMT. Due to the absence of structural changes in the protein, the similar contacts made with the protein, and the fact that both compounds desolvate the cavity to the same degree, the observed difference in the free energy of binding for the two compounds should result from differences in the electrostatic interactions between ligand and protein and from differences in their desolvation energies. However, estimates of ligand desolvation energies were very similar (Table 2), indicating that the observed difference in binding free energy for the two compounds can be attributed to protein−ligand interactions.

Table 2. Calculated Electrostatic Contribution to the Relative Binding Energy of 234TMT and 345TMT to the W191G Cavity[23] [a]

	$-\Delta V_{Q\mu}$	$-\Delta V_{Q\alpha}$	$-\Delta V_{QL}$	$-\Delta V_B$	$-\Delta V_{tot}$	$\Delta G_{solv(water)}$
234TMT	−56.88	3.82	−6.44	−7.26	−66.76	−26.112
345TMT	−57.83	2.98	−5.26	−7.13	−67.24	−26.096
$\Delta\Delta G_{345TMT-234TMT}$	−0.95	−0.84	1.18	0.13	−0.48	−0.016

[a] Values are given in kcal/mol.

The electrostatic contribution to the ligand−protein interaction was estimated for the two ligands bound to the protein cavity. Partial charges were calculated for each ligand (Figure 2) to account for the different C−H bond polarities, and the protein dipoles/Langevin dipoles (PDLD) method[20−23] was used to account for effects of microscopic atomic polarizabilities of atoms surrounding the ligands. The more favorable electrostatic interaction of 345TMT with the protein relative to that of 234TMT (Table 2) arises from both dipole ($\Delta V_{Q\mu}$) and induced dipole ($\Delta V_{Q\alpha}$) terms. These interactions are partially canceled by the solvent screening effects represented in the Langevin grid (ΔV_{QL}) and Born (ΔV_B), terms. The net electrostatic contribution to the binding free energy (ΔV_{tot}) is approximately 0.5 kcal/mol of additional stabilization of 345TMT relative to 234TMT. This value is in good agreement with recent *ab initio* quantum mechanical analysis of CH···O bond energies.[2] Thus, electrostatics alone form a significant contribution to the observed difference in the binding free energy of 1.2 kcal/mol.

In conclusion, the observation of untethered ligand binding to this engineered cavity shows that the CH···O interaction is made by choice and is thus a stabilizing interaction. While weak, summation of several such interactions may play a significant role in the stability of macromolecular structures. The strength of this interaction can be modulated by over 1 kcal/mol by changes in the C−H bond polarity. Thus, selective inclusion of polar CH···O interactions in refinement algorithms for NMR and crystal structures may be justified.

Acknowledgment. We acknowledge C. L. Brooks III, C. D. Stout, M. R. Ghadiri, J. Rebek Jr., D. A. Case, Y. Cao, S. K. Wilcox, P. Beroza, and B. N. Dominy for helpful discussions and critical reading of the manuscript. Support for this work was provided by grants GM-41049 to D.B.G., GM-17844 to R.A.M., and GM-48495 to D.E.M. from the NIH. S.W.B. acknowledges financial support from the ARL Director Research Initiative Program. Support was provided by a grant from the DoD HPC Center, U.S. Army Research Laboratory, for time on the SGI Power Challenge Array.

JA9716766

(16) Crystal structures of W191G soaked in 50 mM 234TMT or 345TMT were determined as previously described.[15] For each ligand, the orientation was unambiguous due to density observed for the methyl substituents and for the sulfur atom at the higher contour level. The maps were constructed by merging F_{obsd} for the soaked crystal with F_{calcd} from models of the W191G empty cavity. The structures of 234TMT and 345TMT bound to W191G have been submitted to the Brookhaven database (entries 1ac4 and 1ac8, respectively). Diffraction statistics for 234TMT and 345TMT, respectively, included unit cell parameters (103.76, 73.31, 44.64 Å and 106.03, 75.89, 51.13 Å), resolution (2.0 and 2.1Å), $I/\sigma_{I(av)}$ (18.2 and 12.4), $I/\sigma_{I(last\ shell)}$ (2.00 and 1.76), number of reflections (18 219 and 19 056), percent completeness (80 and 87%), and R_{sym} (0.048 and 0.062).

(17) Miller, M. A.; Han, G. W.; Kraut, J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11118−111122.

(18) Ligand partial charges were calculated as Kollman ESP charges from geometry-optimized structures using density functional theory. The calculations were performed *in vacuo* as previously described[21] using the program Gaussian 94, Becke3LYP functional, and 6-31G* basis set. ESP partial charges were used to generate an electrostatic potential grid with the program Delphi (MSI) using a uniform dielectric constant of 1, ionic strength of 0.0, and full Coulombic boundary conditions. This potential grid was mapped onto the solvent accessible surface of the molecule calculated with a 1.4 Å probe radius as a color spectrum from +135 kT/e (red) to +200 kT/e (blue).

(19) (a) Washabaugh, M. W.; Jencks, W. P. *Biochemistry* **1988**, *27*, 5044−5053. (b) Washabaugh, M. W.; Jencks, W. P. *J. Am. Chem. Soc.* **1989**, *111*, 674−683.

(20) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley-Interscience: New York, 1991.

(21) Jensen, G. M.; Goodin, D. B.; Bunte, S. W. *J. Phys. Chem.* **1996**, *100*, 954−959.

(22) Dissociation constants and enthalpies were measured at 25 °C in 100 mM Bis-Tris propane pH 4.5 by isothermal titration calorimetry (Microcal MC2 ITC Calorimeter). W191G (0.2−0.4 mM) was titrated with 5 μL injections of ligand (2−8 mM) equilibrated in the same buffer. Error estimates for K_d and ΔH are given as the standard deviation of multiple determinations, and those for ΔS and ΔG were obtained by propagation.

(23) Electrostatic calculations were performed using the protein dipoles/Langevin dipoles (PDLD) method with the program POLARIS.[20] ΔV_{tot} is the sum of four terms, $\Delta V_{Q\mu}$, $\Delta V_{Q\alpha}$, ΔV_{QL}, and ΔV_B. $\Delta V_{Q\mu}$ is the classical electrostatic interaction of the ligand charges with all of the charges on all protein atoms. $\Delta V_{Q\alpha}$ is the interaction energy from the dipole moments induced by the electric field as a result of atomic polarizabilities. ΔV_{QL} is the energy of interaction of the ligand with the field defined by the solvent dipoles. ΔV_B is the Born energy of interaction with bulk solvent beyond the radius used to define the Langevin dipole grid. Values of ΔV have the sign of electrostatic potential, and thus a more positive value for ΔV corresponds to a more negative ΔG. POLARIS calculations were performed with Asp235 and Arg48 charged, and all other amino acid residues neutral. An estimate of the electrostatic interaction of the ligand charges with the solvent reaction field ($\Delta G_{solv(water)}$) for the two compounds was obtained with a continuum dielectric model (Delphi, MSI) using $\epsilon_{water} = 80$ and $\epsilon_{internal} = 2$, with a grid spacing of 0.25 Å.