nitrogen atoms N(9) and N(12) lie within experimental error in the plane of their three substituents.

3,7-Bis(*p*-methoxyphenyl)-1,5-dithia-2,4,6,8-tetrazocine (5). Crystal data: a = 6.3967 (4) Å, b = 34.486 (5) Å, c = 7.3826 (4) Å, orthorhombic, space group *Pbca*, Z = 4, molecular symmetry C_i. Data were collected on a Syntex $P2_1$ diffractometer using graphite-monochromated Cu Kα radiation (λ 1.541 78 Å) in the θ-2θ mode in the range $3° \leq 2\theta \leq 135°$ at scan speeds of 2.93-29.30°/min, depending on the intensity of the reflection. Lorentz-polarization and empirical absorption corrections (μ = 29.71 cm^{-1} for Cu Kα) were applied. After the data reduction 951 unique reflections ($I \geq 3.0\sigma(I)$) were retained for the refinement of the structure.

The position of the sulfur atom was found by interpretation of a Patterson map; the residual atoms could be located in difference maps. The compound displays crystallographic C_i symmetry. The hydrogen atoms which were also located in difference maps were included in the refinement together with isotropic temperature factors. For all other atoms anisotropic temperature factors were introduced. After several cycles the refinement converged to a final value of R = 0.041.

The atomic numbering scheme is the same as that of Figure 1. Bond lengths and bond angles are listed in Tables III and IV. There are no significant differences in the geometry between the molecules **3** and **5**; the bonding parameters in the *p*-methoxyphenyl groups are well in the expected range. The phenyl rings are distorted by 9.1° out of the plane of the eight-membered ring; the carbon atoms of the methoxy substituents are only 0.05 Å off the phenyl plane.[14]

Supplementary Material Available: Final atomic coordinates of **3-5** (3 pages). Ordering information is given on any current masthead page.

The Active Site Electrostatic Potential of Human Carbonic Anhydrase

Robert P. Sheridan and Leland C. Allen*

Contribution from the Department of Chemistry, Princeton University, Princeton, New Jersey 08544. *Received November 16, 1979*

Abstract: The electrostatic potential fields in the active sites of the B and C isozymes of carbonic anhydrase are examined in vacuo. Both isozymes catalyze the reversible reaction $CO_2 \rightleftharpoons HCO_3^- + H^+$, but they have different activities and sensitivities to inhibitors. The proteins are modeled by point charges derived from ab initio molecular orbital calculations on amino acid residues. Hydration of ionized residues is simulated by the use of limiting-case models, and a discussion of solvent effects is included. The active site potentials of the two isozymes are considerably different and may be used to rationalize the functional differences.

Introduction

Molecular electrostatic potential is one method to correlate structure and activity of biomolecules.[1] The potential at a point in space near a molecule is the work needed to bring a unit positive charge from infinity to that point, assuming that the unit charge has no perturbing effect on the molecule. Maps of the electrostatic potential have recently been employed to determine sites of protonation and H-bond formation in small molecules.[2] Similar maps for enzymes can be useful in predicting qualitative features of enzyme-substrate and enzyme-inhibitor interactions. In this paper we construct and analyze the electrostatic potential of two isozymes of carbonic anhydrase in vacuo and take a first step in considering solvent effects.

Very few electrostatic potential maps of proteins exist in the literature. They are carboxypeptidase A,[3,4] lysozyme,[5,6] and cytochrome *c*.[7] (In the case of lysozyme, both the electrostatic potential and the energy of interaction with a single water molecule was studied.) In each case the protein has been modeled as a collection of point charges. For the first two proteins, one point charge was assigned per protein atom with the point charges derived from Mulliken population analysis of molecular orbital calculations on isolated residues. We[8] and others[1,2,9,10] have demonstrated that this model is able to mimic the more exact electrostatic potential calculated directly from molecular wave functions. It was also shown[8,10] that electrostatic potentials derived from wave functions are relatively insensitive to the presence of normal and strong hydrogen bonds or to most perturbing effects of nearby molecules. In addition, the qualitative features of the maps do not appear strongly dependent on the molecular orbital basis set.[1-3] Thus point charge libraries derived for isolated residues should be applicable to the study of polypeptides and proteins in which a large number of residues are in spatial proximity. Examples and extended discussion of electrostatic potential representation by point charges, its sensitivity to the presence of hydrogen bonds and to the choice of atomic orbital basis, and construction of the point charge library are given in the supplementary material.

The influence of solvent, substrate, or counterions has not been explicitly included in this investigation. We have, however, simulated a particularly important solvent effect, the screening of ionized residues near the protein surface, by use of limiting-case models. We also give a detailed discussion of how microscopic solvent effects may be taken into account and how they might influence the electrostatic potential. Furthermore it should be noted that the in vacuo treatment itself is useful for several purposes: (i) to obtain a first approximation of those active site regions where substrates and water molecules are most likely to

(1) Petrongolo, C. *Gazz. Chim. Ital.* **1978**, *108*, 445-478.
(2) Scrocco, E.; Tomasi, J. *Adv. Quantum. Chem.* **1978**, *11*, 116-193.
(3) Hayes, D. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1976**, *98*, 3335-3345.
(4) Hayes, D. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1976**, *98*, 7811-7816.
(5) Clementi, E.; Ranghino, G.; Scordamaglia, R. *Chem. Phys. Lett.* **1979**, *49*, 218-224.
(6) Ranghino, G.; Clementi, E. *Gazz. Chim. Ital.* **1978**, *108*, 157-170.
(7) Koppenol, W. H.; Vroonland, C. A. J.; Braams, R. *Biochim. Biphys. Acta* **1979**, *503*, 499-508.
(8) Sheridan, R. P. Ph.D. Thesis, Department of Biochemical Sciences, Princeton University, 1979.
(9) Bonaccorsi, R.; Petrongolo, C.; Scrocco, E.; Tomasi, J. *Theor. Chim. Acta* **1971**, *20*, 331-342.
(10) Scrocco, E.; Tomasi, J. *Top. Curr. Chem.* **1973**, *42*, 95-170. Pullman, A.; Berthod, H. *Theor. Chim. Acta* **1978**, *48*, 269-277.

0002-7863/81/1503-1544$01.25/0

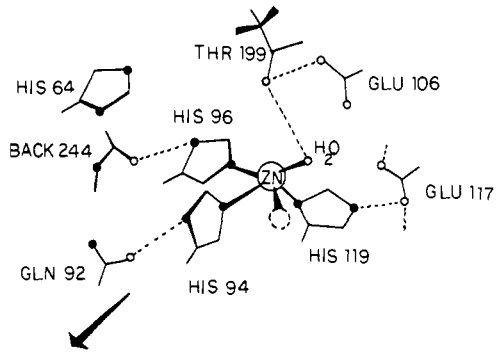


Figure 1. Invariant residues near Zn in HCAB and HCAC. A probable fifth coordination site is shown. The arrow points toward the opening of the active site cavity.

be attracted, (ii) to ascertain the relative electrostatic contributions of different residues, (iii) to compare isozymes, and (iv) to test simplified computational procedures. Because of these reasons and because our calculations do not explicitly include solvent molecules, discussion of solvation is put at the end of this paper.

Carbonic anhydrase contains an essential Zn^{2+} ion. It catalyzes the reversible hydration of CO_2 to $HCO_3^- + H^+$, the hydrolysis of esters, the hydrolysis of phosphates, and the hydration of aldehydes and ketones.[11] The hydration of CO_2 has the highest turnover of any known enzymatic reaction. Unlike those for carboxypeptidase or lysozyme, substrates for carbonic anhydrase tend to be small molecules. The active site, instead of being a cleft or groove as in proteases, is a conical depression about 12 Å deep. The Zn ion is located near the bottom of this cavity and is coordinated by three histidine ligands. Two of them, His-94 and His-96 (HCAB residue numbering), are bound by the Nε(NE2) and the Nδ(ND1). A fourth ligand appears to be a water or hydroxide molecule. There is a probable fifth Zn coordination site where the electronegative atoms of substrates or inhibitors can bind.[12,13] A number of invariant nonligand side chains are near Zn, among them are Thr-199, Glu-106, His-64, Gln-92, and Glu-117. The triad, Zn-bound H_2O, Thr-199, and Glu-106, has been postulated to be a catalytic charge relay.[12] See Figure 1.

Two isozymes have been studied by high-resolution X-ray crystallography: human erythrocyte carbonic anhydrase B (HCAB) and C (HCAC).[14] In both, the active site cavity can be almost perfectly divided into nonpolar and polar halves[14] on the basis of the nature of the side chains lining the cavity (see Table I). The boundary between the halves passes almost through Zn. From inhibitor-binding studies[12,13] and from structural analogy with other Zn enzymes,[15] it appears that the nonpolar portion of larger substrates will bind in the nonpolar half while the reactive groups coordinate to Zn and/or hydrogen bond to residues in the polar half. One of the purposes of the present work has been to determine the electrostatic consequence to this division of the active site.

The shape of the cavity and the nature of its side-chain lining show slight but significant differences in detail between the isozymes.[15] See Table I. The activity and inhibitor-binding properties also differ.[13,14,16,17] These differences give rise to the possibility of correlating the electrostatic potential with functional properties.

Table I. Residues with Side Chains Lining Active Site[a]

nonpolar			polar			
res no.	HCAB	HCAC	res no.	HCAB	HCAC	
91	Phe	Ile	7	Tyr	Tyr	
121	Ala	Val	61	Asn	Asn	
131	Leu	Phe	64	His	His	
141	Leu	Leu	65	Ser	Ala	b
143	Leu	Leu	67	His	Asn	
145	Gly	Gly	69	Asn	Glu	
201	Pro	Pro	92	Gln	Gln	
202	Pro	Pro	94	His	His	c
204	Tyr	Leu	96	His	His	c
206	Ser	Cys	106	Glu	Glu	d
207	Val	Val	117	Glu	Glu	d
211	Ile	Val	119	His	His	c
			199	Thr	Thr	
			200	His	Thr	

[a] Table modified from ref 14. [b] Originally listed as a nonpolar residue, residue 65 is well within the borders of the polar half. [c] Zinc ligand. [d] Added to the original list.

Pocker and Sarkanen[11] and Lindskog et al.[16] have comprehensively reviewed the enzymology of carbonic anhydrase. Notstrand et al.[14] compare the structure of HCAB and HCAC. Maren[18,19] has given an excellent review of the physiological roles of this enzyme.

Models and Methods

Ionization State of Residues. A number of reasonable assumptions might be made about the ionization state of exposed acidic and basic residues. Two extreme models have been taken. Hayes and Kollman in their study of carobxypeptidase A[3,4] assumed that all the residues were in their predominant state of ionization in aqueous solution at pH 7. In contrast, Clementi and co-workers,[5,6] in their studies on lysozyme, assumed that all residues were neutral. They reasoned that ionized residues would tend to bind counterions in solution, thus leading to effective neutrality for both ionized and unionized residues. An assumption of the first type gives electrostatic potential maps that are sensitive to the number and kind of charged residues, with the permanent dipoles of the protein making only a secondary contribution. The omission of hydration and counterions clearly exaggerates the importance of charged residues. On the other hand, the second type of assumption has counterions bound so tightly that no charge will be felt from an ionized residue. This also is too extreme.

In this study we treat three limiting-case models in order to bridge the two extreme assumptions illustrated by the Hayes and Kollman and Clementi treatments discussed above. We start by taking a pH slightly above neutrality (the enzyme is fully active at this pH) and adopting the ionization states used by Hayes and Kollman. We then consider limiting cases where various sets of the charged residues are included or omitted. Aspartate and glutamate side chains are thus anionic and arginine and lysine side chains, cationic. Histidine is neutral. Glu-106 and Glu-117 are partly buried and therefore require special consideration. Since Glu-106 is postulated to be part of a catalytic charge relay,[12] it must be anionic in the native active enzyme. Glu-117 is at the bottom of the active site 5 Å from Zn, and each of its carboxyl oxygens has at least one hydrogen bond from proton-donating groups. One of the proton donors is His-119, a Zn ligand which can be expected to have a partial positive charge. It seems reasonable, therefore, that Glu-117 is anionic in the native enzyme. The range of possible hydration and counterion effects is then represented by three limiting-case models.

(a) The first model includes all atoms of the protein plus Zn and Zn-bound H_2O. Provision is made for charge transfer onto the Zn ligands. All ionized residues are represented. This model

(11) Pocker, Y.; Sarkanen, S. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1978**, *47*, 149–274.
(12) Kannan, K. K.; Petef, M.; Fridborg, K.; Cid-Dresdner, H.; Lövgren, S. *FEBS Lett.* **1977**, *73*, 115–119.
(13) Kannan, K. K.; Vaara, I.; Notstrand, B.; Lövgren, S.; Borrell, A.; Fridborg, K.; Petef, M. In "Proceedings on Drug Action at the Molecular Level"; G. C. K. Roberts, Ed.; University Park Press: Baltimore, 1977, pp 73–91.
(14) Notstrand, B.; Vaara, I.; Kannan, K. K. In "Isozymes I"; Markert, C. L., Ed.; Academic Press: New York; 1975, pp 575–599.
(15) Argos, P.; Garavito, R. M.; Eventoff, W.; Rossmann, M. G. *J. Mol. Biol.* **1978**, *126*, 141–158.
(16) Lindskog, S.; Henderson, L. E.; Kannan, K. K.; Liljas, A.; Nyman, P. O.; Strandberg, B. *Enzymes, 3rd. Ed.* **1977**, *5*, 587–665.
(17) Maren, T. H.; Rayburn, C. S.; Liddel, N. E. *Science (Washington, D.C.)* **1976**, *191*, 469–472.
(18) Maren, T. H. *Physiol. Rev.* **1967**, *47*, 595–781.
(19) Maren, T. H. *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **1967**, *26*, 1097–1103.

represents the limit of no hydration, no counterions, nor solvent dielectric. (The maps in this case will be sensitive to the number and kind of charged residues but not very sensitive to the assumed polarity of hydrogen bond networks.)

(b) In the second model all atoms of all charged residues *except* those lining the active site are left out. This represents neutralization of residues on the outside of the enzyme, but full retention of effective charge of the residues not in contact with bulk solvent. There is, of course, a certain limited arbitrariness in distinguishing which charged residues are separated from solvent and which are not. The charged residues retained are listed in Table I.

(c) All atoms of all charged residues are left out, including Zn and Zn-bound H_2O (the ligand histidines revert to neutral). This case will bring out electrostatic features of the active site produced by permanent dipoles in the protein which would otherwise be masked by the presence of charged residues. Contributions from the backbone of the missing residues will also be left out. Thus the electrostatic field in regions between the carbonyl C and amino N of residues adjacent to a missing residue will obviously be artifactual. The problem which this might cause, however, is bypassed because the contour levels in the maps are such that the regions of artifact are well within the van der Waals envelope of the missing residues. This model is sensitive to the assumptions about the polarity of hydrogen bond networks but not to the charged residues.

Sequence and Coordinate Modification. All residue numbering is that from the HCAB sequence. Coordinates of nonhydrogen atoms are taken from the Brookhaven Protein Data Bank,[20] and we have employed their atom nomenclature throughout. It should be noted that the X-ray structures of HCAB and HCAC are at a somewhat preliminary stage of refinement, there being a number of close contacts and "holes" in the protein as it presently stands. However, further refinement is not expected to change the qualitative features of the electrostatic potential maps we show here. The sequence of HCAB and HCAC used for the calculations are listed in the supplementary material (Table II). The first four residues of HCAB and the first two and last residue of HCAC are not resolvable in the crystal structures, and coordinates are not provided; these residues are left out of the calculations. Since Asp-4 is left out of HCAB, the total charge on the molecule will appear too positive by one unit charge. For HCAC Lys-261 is left out. Since the positive charge of this side chain cancels the charge of the carboxy terminus, the total charge for HCAC as calculated will be correct. The sequence of HCAC was revised by Henderson et al.[21] after the crystal structure was solved. There are five Asp → Asn and Glu → Gln changes and one Gln → Glu change. Some of the discrepancies from earlier sequence studies are due to the fact that three Asn and Gln residues spontaneously deaminate in vitro. See supplementary material (Table II). For Asp → Asn and Glu → Gln the ambiguity as to which position in the side chain is N and which is O was resolved on the basis of which position would be more stereochemically reasonable as judged from the stereo pictures of the environment around the residues in the AMSOM atlas.[22]

The apoprotein coordinates have been modelbuilt with standard bond lengths and angles, but the Zn positions listed by the data bank were measured directly from the wire models, and this introduces unrealistic Zn–ligand distances. These were corrected, and a more nearly tetrahedrally coordinated Zn was used. Zn-bound H_2O positions were also assigned a more nearly tetrahedral arrangement around Zn. For a comparison of HCAB and HCAC, the HCAB coordinates for nonhydrogen atoms were rotated and translated into the HCAC coordinate frame by using the Rao and Rossman routine[23] (as modified by Dr. C. S. Wright). The minimized root-mean-square displacement of α-carbons was found to be 1.5 Å. The transformation for rotation and translation is given in the supplementary material (Table III).

Hydrogen Positions. Hydrogens have been attached to the other atoms with the subroutine HYDRO in the energy refinement package of Levitt.[24] HYDRO places hydrogens 1.08 Å from carbons, 1.03 Å from nitrogens, and 0.97 Å from oxygens in the best tetrahedral, trigonal, or bent configuration. Where rotation around single bonds is allowed, HYDRO chooses the least locally hindered conformation (i.e., usually staggered).

The hydrogens in Zn-bound H_2O, serine, threonine, tyrosine, cysteine, and histidine are capable of forming hydrogen bonds. An attempt was made to place the hydrogens such that the best hydrogen bonds were formed. The AMSOM atlas[22] provides a list of probable hydrogen bonds based on stereochemical criteria. Those hydrogen bonds listed for HCAB and for equivalent or different groups in HCAC are tabulated in the supplementary material (Table IV). For serine, threonine, tyrosine, and cysteine, the hydrogen on each proton donor was positioned at the standard C–O–H or C–S–H bond angle with a dihedral angle around the C–O or C–S bond such that the best hydrogen bond was made with the most likely electron donor. Which nitrogen of neutral histidine has the hydrogen was also decided from the most likely pattern of hydrogen bonding. Where a serine, threonine, tyrosine, or cysteine residue was not involved in an obvious hydrogen bond, the hydrogen was left in the sterically favorable position determined by HYDRO. Histidines not involved in hydrogen bonds were assumed protonated at ND1.

Point Charge Library. A Mulliken point charge library slightly modified from that of Hayes and Kollman[3,4] was used for the calculation of the electrostatic potential maps. One point charge is assigned to every atom in the protein. The library is in the supplementary material (Table V). We also note that (i) the charge on the nitrogen of the first visible residue (Trp-5 in HCAB and His-4 in HCAC) is taken as the sum of N and H charges (this obviates the problem of not knowing where the hydrogen is pointing), (ii) the atomic charges for the backbone of C-terminal Phe-260 in HCAB are taken from the charges in C-terminal Ala, the side chain charges from the standard Phe residue (the charges of CA and HA were adjusted to produce a total charge of −1.0), and (iii) the last residue carbonyl in HCAC was capped with an in-plane neutral $-NH_2$.

Provision was also made for the charge transfer onto the Zn ligands. We assume the fourth ligand is water (rather than hydroxide) consistent with NMR evidence.[11] It is not anticipated that our conclusions will be sensitive to the exact nature of the various assumptions described here.

Planes of Cross Section through the Active Site. The maps shown in this paper are in two planes of cross section. One plane, V, is taken through the axis of the conical active site and through Zn (average Zn position from HCAB and HCAC). The other plane, H, is perpendicular to the first. The axis is the line through the bisecting point between CA (C_α) 64 and CA 131 perpendicular to the plane containing CA 64, CA 69, and CA 131. (The CA positions are the average of the corresponding CA positions in HCAB and HCAC.) See Figure 2.

Calculation of the Maps. The cross section of the plane through the van der Waals envelope was generated by assuming all atoms have an average radius of 1.4 Å. (This is too large for hydrogen, but will not introduce a significant error, given the scale of the maps.)

The electrostatic contours were generated by using a modified version of the program PTCHG developed in our laboratory by A. L. Fisher.[25] The potential contour levels chosen (absolute value ≥5 kcal) are those which will be insensitive to interresidue interaction, so that the point charge library derived from isolated residues should be adequate for this study.

(20) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rogers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535–542.

(21) Henderson, L. E.; Henriksson, D.; Nyman, P. O. *J. Biol. Chem.* **1976**, *251*, 5457–5463.

(22) Feldman, R. A. "Atlas of Macromolecular Structure on Microfiche"; Tracor Jitco, Inc. 1976.

(23) Rao, S. T.; Rossmann, M. G. *J. Mol. Biol.* **1973**, *76*, 241–256.

(24) Levitt, M. *J. Mol. Biol.* **1974**, *82*, 393–420.

(25) Fisher, A. L. Senior Thesis, Department of Chemistry, Princeton University, 1978.

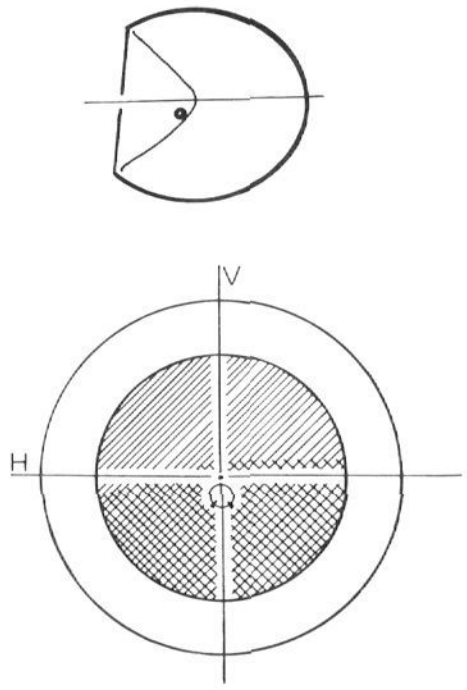


Figure 2. Top: schematic cross section of carbonic anhydrase showing the conical active site cavity. The dark circle represents the position of Zn. The horizontal line is an approximate axis of the active site. Bottom: schematic view looking down the axis into the active site. Zn ligands are below Zn. Two cross-section planes through the axis are shown: V (through Zn) and H (perpendicular to V). Cross-hatched area represents the nonpolar half of the active site and double-cross-hatched area, the polar half.

Results

The electrostatic potential maps for HCAB and HCAC are shown in Figures 3 and 4, respectively. We will confine our discussion to overall qualitative features of the maps.

Steric Restriction in the Active Site. The most striking feature of the active sites of both isozymes is that they are very narrow, contrary to the impression one usually gets from ORTEP drawings. Comparison of the site width with the size of a water molecule shows that there can be only a few water molecules in the site. As noted in the literature[13] the bottom of the active site, where substrates and inhibitors bind, is much more restricted in HCAC than in HCAB.

The Limiting-Case Models. (a) When all charged residues are present, the potential of the active site of HCAB (overall molecular charge 0.0) is strongly positive. That of HCAC (overall charge 0.0) is somewhat negative except in the immediate vicinity of Zn. In neither HCAB nor HCAC is there a region of negative potential outside the van der Waals envelope of the protein that can be attributed to Glu-106 or Glu-117. However, they reduce the positive potential in the vicinity of Zn.

(b) When all charged residues but those in the active site are removed, the only charge difference between HCAB (net charge 0.0) and HCAC (net charge −1.0) is due to Glu-69 in HCAC which maintains a region of negative potential near the mouth of the active site. The corresponding residue in HCAB is Asn. Thus, in two limiting cases HCAB appears to have a more positive potential at the active site than HCAC. We suspect that this is one feature which distinguishes the isozymes. Potential studies of the enzyme molecule as a whole shows that the placement of charged residues on the enzyme surfaces is responsible for the difference. The residues surrounding the active site of HCAB tend to be cationic and those around HCAC, anionic. See Figures 5 and 6.[26] The deamination of specific Asn's or Gln's in HCAC[21] would not change this conclusion. Neither would the presence of the missing Asp-4 in HCAB.

(c) When all charge-bearing groups are removed, only the contributions of permanent dipoles remain. In this limiting case the two isozymes are remarkably similar. Instead of a complex pattern of local positive and negative potential, one sees a separation into two major regions. Most of the active site has a negative potential. This is due to the lone pairs on the ligand histidines. There is, however, a wide crescent of positive potential extending from the outer lip of the active site cavity to about two-thirds of the way to the bottom. The crescent does not follow the division of nonpolar and polar halves. About half of it extends into the polar region; the remainder in the nonpolar region. See Figure 7. The placement is roughly the same for both isozymes (this can be seen properly only in three-dimensional models). The nonpolar portions of substrates and inhibitors bind to that part of the active site where the crescent appears. It is not clear at this time what groups produce the positive potential, although the groups are common to both isozymes. It is also not clear whether the positive potential region has any functional significance since the native active enzyme contains charge-bearing groups whose potential could swamp out this feature. However, it is clear that there is no distinction between the isozymes at the active site when charged groups are omitted. It should be noted that in none of the limiting case models is there any direct correspondence between electrostatic features of the active site with its division into polar and nonpolar halves.

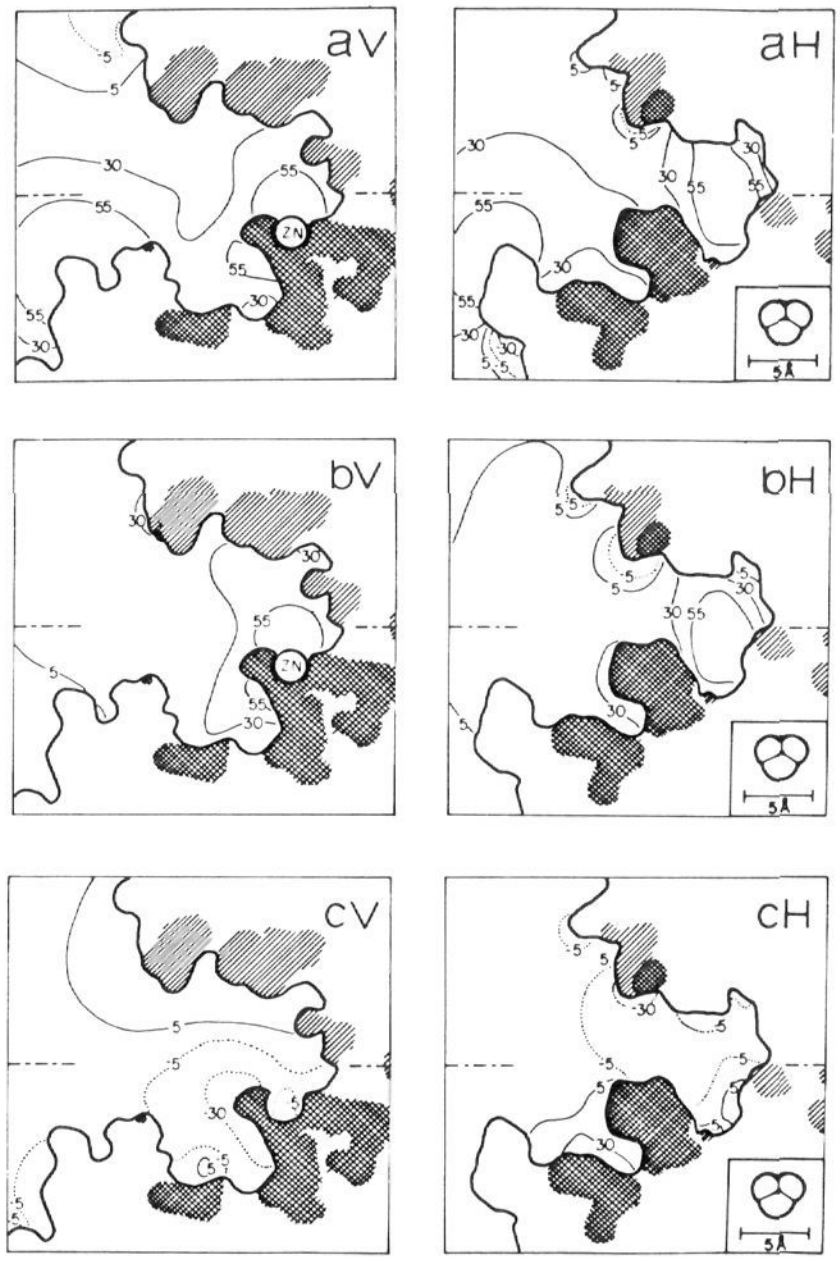


Figure 3. Electrostatic potential maps for the two cross-section planes (V and H) shown in Figure 2 for three limiting cases in HCAB: (a) all atoms of protein (plus Zn-bound H_2O) with residues in their ordinary states of ionization (this case is the limit of no hydration, no counterions, and no solvent dielectric); (b) all atoms of charged residues omitted except those lining the active site (this is a model which represents neutralization of charged residues on the outer surface of protein but retention of residue charges not in contact with the bulk solvent (i.e. those inside the conical active site); (c) all atoms of charged residues (including Zn and Zn-bound H_2O) omitted (this is the limiting case where the electrostatic potential is produced solely by permanent dipoles and higher multipole moments). Contour levels in kcal/mol. The cross-hatched areas are cross sections through the residues in the nonpolar half of the active site and double-cross-hatched areas, cross sections through the residues in the polar half. See Table I. Inset in H cross sections shows a water molecule at the same scale.

(26) The maps in Figure 5 were produced by placing a unit point charge (±1.0) at the atom nearest the centroid in each charged residue: at CG of aspartate, CD of glutamate, NZ of lysine, CH of arginine, and C of C-terminus. A charge of +2.0 was used for Zn.

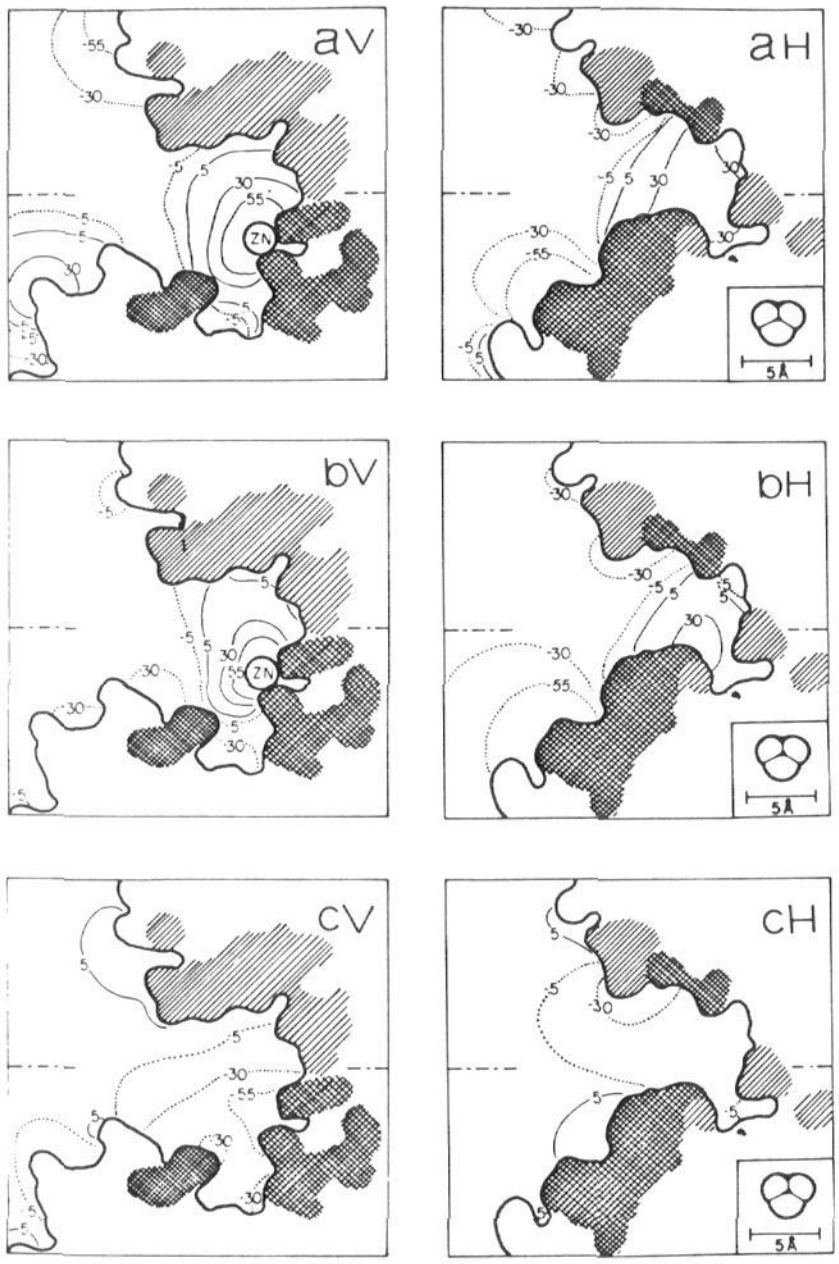


Figure 4. Electrostatic potential maps for HCAC. Same limiting cases and conventions as Figure 3.

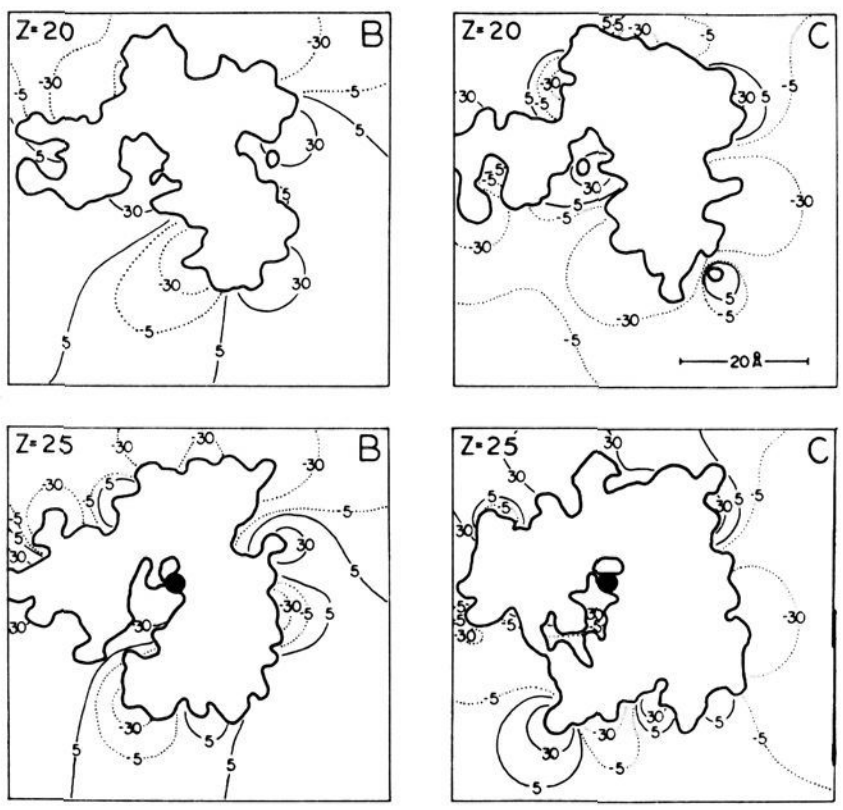


Figure 5. Electrostatic potential maps of HCAB (left) and HCAC (right) for two sections parallel to the *XY* plane in the coordinate system specified in the Brookhaven Protein Data Bank for HCAC (sections are in the planes $Z = 20$ and 25 Å). Contours in kcal/mol. The active site opens toward bottom left and out of page. Filled circle is Zn position. Potential was generated by using a unit point charge (±1.0) per ionized *residue* and Zn equal 2+.

Discussion

The electrostatic potential is a consequence of protein structure and as such may be used to address questions of function. In the following discussion we make some preliminary examination of the differences between isozymes where a correlation is evident. Some of the most noted function differences are the following:

(1) The turnover number for the hydration of CO_2 or dehydration of HCO_3^- is about 20 times higher for HCAC than HCAB.[17] HCAB is more active than HCAC in hydrolizing some esters.[17]

(2) The titration of kinetic and a variety of spectroscopic parameters is more complex for HCAB than HCAC.

(3) HCAC is more sensitive to inhibition by acetazolamide and other substituted sulfonamide inhibitors.

(4) HCAB is much more strongly inhibited (noncompetitively or mixed) for CO_2 hydration and (competitively) for HCO_3^- dehydration by anions, notably Cl^- and I^-.[17]

(5) Only HCAB binds imidazole, the only known competitive inhibitor of CO_2 hydration.[12]

Binding of Ions. A higher positive charge at the active site of HCAB may form the basis of the differential sensitivity of HCAB to inhibition by anions. HCAB is especially sensitive to HCO_3^- as a noncompetitive inhibitor for CO_2 hydration.[27]

One might also expect the K_M for HCO_3^- dehydration to be lower for HCAB than HCAC if the electrostatic attraction of the active site for substrate anion were the controlling parameter on K_M. As measured, the K_M for HCAB is indeed lower than for HCAC by a factor of 2 around neutral pH.[16]

Catalytic Steps. The hydration reaction requires the removal of a proton from the active site between the release of the HCO_3^- product and the binding of the next substrate CO_2. Kinetic isotope investigations for both the uncatalyzed[28] and catalyzed[29] hydration of CO_2 show that a proton transfer is always the rate-limiting step. NMR relaxation studies of Co^{2+} substituted bovine carbonic anhydrase and HCAB[27] indicate that the pK of the proton transfer

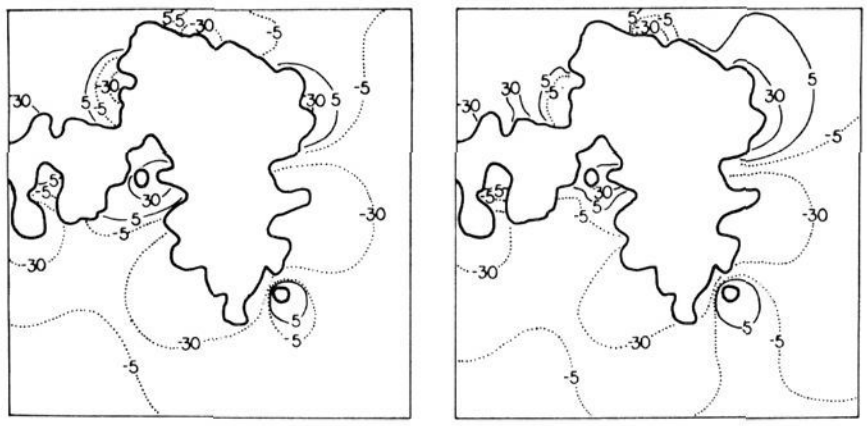


Figure 6. Comparison of electrostatic potential generated by single point charge per charged residue (as shown in Figure 5) with that obtained from Mulliken net charges *on each atom*. $Z = 20$ Å section. Left side: single point charge per residue model. Right side: Mulliken charges. The close similarity of the two maps is evidence that the monopole charges dominate the protein potential.

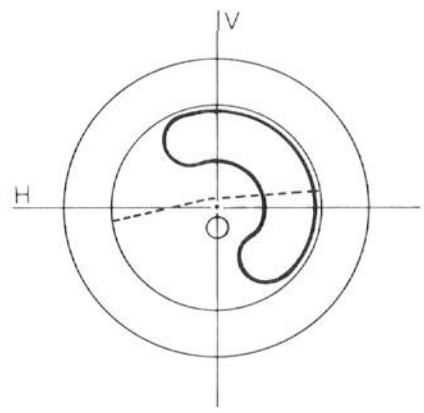
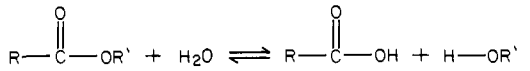


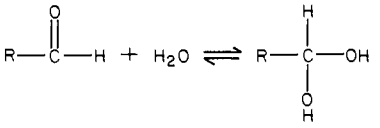
Figure 7. Schematic view looking into active site showing crescent of positive potential obtained when all charged residues, including Zn, are omitted (heavy outline). The rest of the active site has a negative potential. Dashed line is boundary between polar and nonpolar halves.

(27) Fabry, M. E. *J. Biol. Chem.* **1978**, *253*, 3568–3574.
(28) Pocker, Y.; Bjorkquist, D. W. *J. Am. Chem. Soc.* **1976**, *99*, 6537–6543.
(29) Jönsson, B. H.; Steiner, H.; Lindskog, S. *FEBS Lett.* **1976**, *64*, 310–314.

in the absence of buffer is always different from the pK of the optical shift with pH. However when the buffer is added, the pK of the transfer for bovine carbonic anhydrase approaches that of the optical pK. The pK of the transfer for HCAB is not dependent on buffers. The optical pK is assumed to be due to the ionization of a Zn ligand, presumably Co-bound H_2O. The activity pK is very close to the optical pK. These results imply two relatively slow proton transfers, one from the metal-bound water to a proton-accepting group on the enzyme and one from the group to bulk solution. The second step is aided by buffers and is rate limiting in their absence.[29] The proton accepting group has been identified with the hydrogen bond network of the "charge relay" Thr-199···Glu-106[12] or with the network His-107···Glu-117.[27] This group may as easily be far from the metal. Proton transfer obviously involves charge rearrangement in the active site. The shape of the electrostatic potential in the site will affect the energy profile of the transfer. At this time, since we do not know which residues are involved, we cannot say how the rate might be affected by the potential in the two isozymes. However we have at least one hint that electrostatics are probably important in distinguishing the rates of HCAB and HCAC. HCAB is slower than HCAC in hydration and dehydration where proton transfer between turnover is important. In contrast, in the esterase reaction



where there is no rate-limiting step involving net removal of a proton from the substrates or product to bulk solvent, HCAB and HCAC have comparable rates for many esters.[16] This should also prove true for the aldehyde hydration reaction[11]



which to our knowledge has not been studied for the human isozymes.

Binding of Substrate. We did not address ourselves to the problem of the interaction of substate with the isozymes. Hayes and Kollman[4] studied the interaction of carboxypeptidase with a peptide inhibitor. Unfortunately, no molecule that can be considered a model substrate has been crystallized with carbonic anhydrase. An important recent study on carbonic anhydrase by Pullman and Demoulin[30] has treated the interaction between a model active site (identical for both isozymes) and the CO_2 substrate. They have determined the substrate orientation corresponding to the minimum electrostatic energy between a tetraliganded Zn and CO_2 which are excluded from an approach closer than the sum of the van der Waals envelopes surrounding them. Our concern here has been the electrostatic potential of whole proteins, and we are currently investigating to what extent this potential may modify the CO_2 orientation and its path of approach.

Steric Restrictions in the Active Site. In the course of our study it became obvious that besides the electrostatic potential, the geometry of the active site itself would have consequences for catalysis and inhibition. First, the very narrowness of the active site leads one to expect that the solvent in the site would be rather ordered because of hydrogen bonding to the protein and the restriction in translation and rotation. Solvent molecules in the site of HCAC were reported in the original X-ray structure, but whether the electron densities interpreted as solvent are real or the result of phasing errors is not clear.[11,14]

Both isozymes of carbonic anhydrase are inhibited by most monoanions since they bind to Zn and K_I's for strongly inhibiting ions follow the order of anion binding to simple Zn complexes. Weakly inhibiting anions follow the Hofmeister series, with smaller anions (e.g., F^-) binding the least tightly. It has been suggested

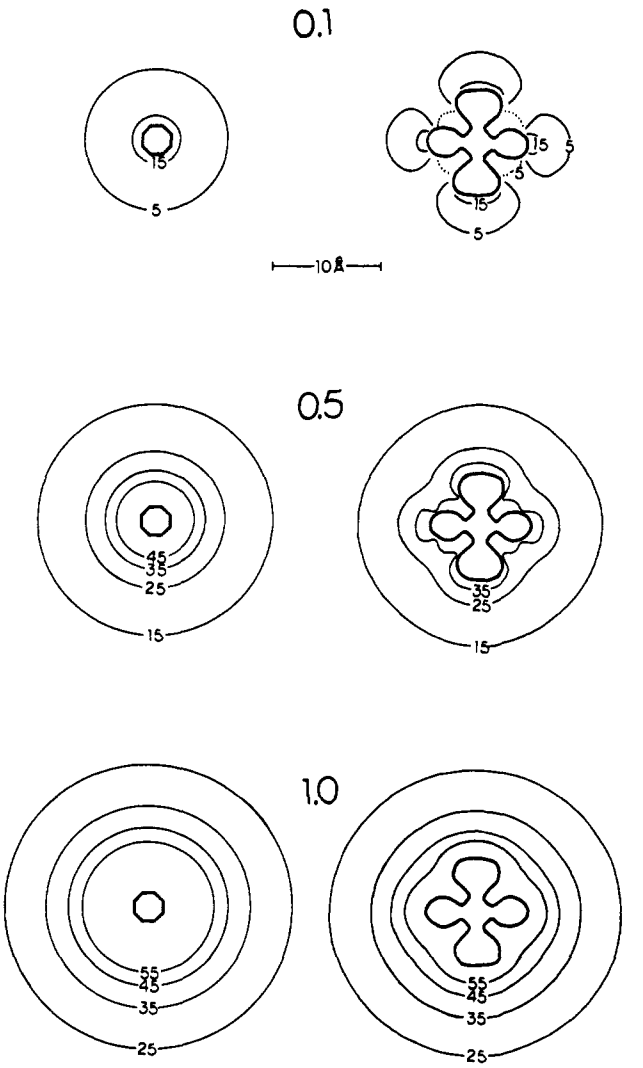


Figure 8. Electrostatic potential of isolated point charges of magnitude 0.1, 0.5, and 1.0 (left side) compared with that of same charges surrounded by cage of six water molecules (right side). Each water molecule is represented by three point charges.

that the specificity of the weak monoanion inhibition is based on the change of water structure in the active site accompanying the formation of the enzyme–anion complex[11] because the enthalpy increase for breaking the water shell around small ions is greater than that for larger ions. We can add an additional geometrical reason for the ordering. Anions with rigidly held hydration shells simply would not be able to fit into the narrow active site.

The fact that HCAB binds imidazole while HCAC does not may be explained by the narrowness of the cavity in HCAC relative to HCAB near the fifth coordination site where HCAB binds the inhibitor. A cyclic compound such as imidazole may simply be unable to approach within coordination distance.

Solvent Effects

The presence of solvent (i.e., water) can be expected to affect the electrostatic potential of a protein. The dipoles of oriented water molecules may, for instance, help "screen" charges. Some important work on protein–solvent interactions exists in the literature. This research has employed one of two methods. The first, used extensively by Clementi and co-workers,[5,6,31] finds the energy of interaction of a single water molecule at various points around a macromolecule. Solvent–solvent interactions and entropic effects are omitted, therefore exaggerating somewhat the degree of solvent orientation around the protein.

(30) Pullman, A.; Demoulin, D. *Int. J. Quantum. Chem.* **1979**, *16*, 641–653.

(31) Clementi, E.; Cavallone, F.; Scordamaglia, R. *J. Am. Chem. Soc.* **1977**, *99*, 5531–5545.

A more complete picture is given by Monte Carlo methods. The largest such calculation to data has been the study of solvated lysozyme done by Hagler and Moult.[32] In their treatment, a large number of water molecules surrounding the entire protein were included, a set of configurations was randomly generated, and their energy was calculated by using empirical force functions. The contribution of each configuration to the total energy was weighted by the Boltzmann factor. Hagler and Moult were able to show that this procedure correctly predicts the crystallographic *R* factors for a small peptide in an aqueous crystal. For lysozyme it appears that one or two layers of water molecules are strongly localized at the protein surface, especially near side chains. Water molecules further away from the surface than about 4 Å are not localized but are very much like "bulk" water. Clementi et al., in a Monte Carlo study relevant to the present work, calculated the solvent structure around a Zn^{2+} ion[33] and also around Zn plus 27 nearby residues for the active site of HCAB.[34] Their work indicates one or two strongly localized layers of waters near Zn both for the isolated ion and the active site.

The next logical step needed to follow a treatment such as that of Clementi et al. is calculation of the electrostatic potential in the enzyme active site using solvent coordinates obtained from the Monte Carlo calculations. Computer programs for such electrostatic calculations are not yet available for carrying this out, but we can make a preliminary estimate on the maximum amount by which the dipoles of sovlent molecules can modify the in vacuo active site potential assuming that only about one laye of water is localized. The rotational and translational average of the dipole of a molecule of bulk water is obviously zero. Most of the correction, therefore, will come from the single layer of water molecules at the protein surface and then only if the molecules are rotationally oriented as well as translationally fixed. The maximum possible effect will occur when each molecule dipole is at right angles to the isopotential contour lines. Our model consists of a central point charge surrounded by a cage of six water molecules. Three point charges are used to represent the atoms of each water molecule. The "oxygens" are placed 3 Å from the central point charge. The symmetry axis of the water passes through the charge; i.e., the dipoles of the waters are fully aligned. Polarization and charge transfer are neglected. Figure 8 compares the potential field emanating from central charges of magnitude 0.1, 0.5, and 1.0 unit surrounded by six waters. As can be seen, there is little change in the contours for central charges much greater than 0.1. In the case of protein potentials this would lead us to expect that the potential produced by a charged residue would not be affected very much by one shell of strongly oriented water. Only in regions of low-field strength would the contribution from oriented water dipoles be significant, and in that case the dipoles would not be strongly oriented. Thus, in general, we do not expect solvent alone to make gross changes in those in vacuo maps for which charged residues are included. Counterions, on the other hand, might make a strong difference.

Acknowledgment. The authors wish to thank Eric Henry and Dr. C. S. Wright for patient instruction in the use of HYDRO and the Rao and Rossmann routine. We are also grateful to Dr. D. M. Hayes, for providing his Mulliken point charge library. This work was supported by the National Institutes of Health (Grant GM 26462).

Supplementary Material Available: The primary sequence for HCAB and HCAC including the recent revisions of Henderson et al. (Table II); the matrix transform to rotate and translate coordinates from the HCAB frame to the HCAC frame (Table III); assumed hydrogen bond networks involving serine, threonine, tyrosine, and neutral histidine (Table IV); description and list (Table V) of the point charge library for amino acids (22 pages). Ordering information is given on any current masthead page.

(32) Hagler, A. T.; Moult, J. *Nature (London)* **1978**, *272*, 222-226.
(33) Clementi, E.; Corongiu, G.; Jönsson, B.; Romano, S. *J. Chem. Phys.* **1980**, *72*, 260-263.
(34) Clementi, E.; Corongiu, G.; Jönsson, B.; Romano, S. *FEBS Lett.* **1979**, *100*, 313-317.

The Reactions of Semi-Met Forms of Hemerythrin

P. C. Harrington and R. G. Wilkins*

Contribution from the Department of Chemistry, New Mexico State University, Las Cruces, New Mexico 88003. *Received August 1, 1980*

Abstract: The properties of $(semi\text{-}met)_O$, produced by one-electron oxidation of deoxyhemerythrin, and $(semi\text{-}met)_R$, produced by one-electron reduction of methemerythrin, are described. Both forms react with F^-, Br^-, SCN^-, CN^-, and N_3^- to give, with each anion, a semi-met adduct. Full kinetic data are given for the formation of the azide adduct from met, $(semi\text{-}met)_O$ and $(semi\text{-}met)_R$ forms of hemerythrin at pH 6.3 and 8.2 and 25 °C. The stability of the met-azide form is much higher than that of the semi-met form, mainly arising from a much smaller dissociation rate constant. The rate parameters for the disproportionation of the two semi-met forms at pH 8.2 and 25-35 °C were determined, and the role of disproportionation in reactions of the semi-met forms with O_2, $Fe(CN)_6^{3-}$, and $S_2O_4^{2-}$ is delineated. The processes are discussed in terms of the octomeric structure of the protein, and it is concluded that disproportionation results from an intramolecular electron transfer involving a rate constant of 2.7×10^{-3} s^{-1} over distances of 28-30 Å, between each (binuclear) iron unit. Data are for protein from *Themiste zostericola* and (limited) from *Phascolopsis gouldii* and *Themiste dyscritum*.

Hemerythrin occurs in the erythrocytes of certain marine worms in a polymeric, usually octameric, form. Each subunit, mol wt ~13 500, contains two linked nonheme irons, but there is still a question whether amino acids or oxy bridging is involved.[1-4] The form with both irons in the oxidation state +2 (deoxy) interacts rapidly and reversibly with oxygen.[5] It is easily oxidized to the met form containing irons only in the +3 oxidation state. This is no longer O_2 sensitive but does react with a number of anions

(1) D. M. Kurtz, Jr., D. F. Shriver, and I. M. Klotz, *Coord. Chem. Rev.*, **24**, 145 (1978).
(2) W. A. Hendrickson, *Naval Res. Rev.*, **31**, 1 (1978).
(3) R. E. Stenkamp and L. H. Jensen, *Adv. Inorg. Biochem.*, **1**, 219 (1979).
(4) J. S. Loehr and T. M. Loehr, *Adv. Inorg. Biochem.*, **1**, 235 (1979).
(5) D. J. A. deWaal and R. G. Wilkins, *J. Biol. Chem.*, **251**, 2339 (1976).

0002-7863/81/1503-1550$01.25/0