Protein Engineering vol 4 no.1 pp 57–63, 1990

Geometry of interaction of metal ions with histidine residues in protein structures

P.Chakrabarti

Department of Chemistry and Biochemisry, University of California, Los Angeles, CA 90024, USA

Present address: Division of Physical Chemistry, National Chemical Laboratory, Pune 41108, India

An analysis of the geometry and the orientation of metal ions bound to histidine residues in proteins is presented. Cations are found to lie in the imidazole plane along the lone pair on the nitrogen atom. Out of the two tautomeric forms of the imidazole ring, the NE2-protonated form is normally preferred. However, when bound to a metal ion the ND1-protonated form is predominant and NE2 is the ligand atom. When the metal coordination is through ND1, steric interactions shift the side chain torsional angle, χ_2 from its preferred value of 90 or 270°. The orientation of histidine residues is usually stabilized through hydrogen bonding; ND1-protonated form of a helical residue can form a hydrogen bond with the carbonyl oxygen atom in the preceding turn of the helix. A considerable number of ligands are found in helices and β-sheets. A helical residue bound to a heme group is usually found near the C-terminus of the helix. Two ligand groups four residues apart in a helix, or two residues apart in a β-strand are used in many proteins to bind metal ions.

Key words: binding geometry/histidine ligand/metals/protein secondary structure/side chain conformation

Introduction

In many metalloproteins histidine (His) residues act as ligands to metal centers, which play important roles in catalysis, oxidation–reduction and transport processes (Sundberg and Martin, 1974). Covalent attachment of ruthenium ions to the His side chain has been used to introduce a second redox center at the surface of proteins and then to study electron transfer reactions (Mayo *et al.*, 1986; Osvath *et al.*, 1988). The geometry and the orientation of metal ions with respect to the imidazole ring of histidyl residues should be useful in understanding many of the properties associated with metal centers and are analyzed here. The geometry of such interactions can also be useful as a model of how a His residue acting as a nucleophile interacts with an electrophile at the active center of different enzymes (Zvelebil and Sternberg, 1988). The side chain conformation and the secondary structural features of ligand residues are also analyzed.

Materials and methods

Atomic coordinates from the Brookhaven Protein Data Bank (PDB) (Bernstein *et al.*, 1977) were used in this analysis; those for glutamine synthetase were kindly provided by the authors (Yamashita *et al.*, 1989). The structure of cytochrome *b*562 (Lederer *et al.*, 1981) has not yet been refined and was not considered. Also excluded from the analysis were some recently determined structures like alkaline phosphatase (Sowadski *et al.*, 1985), haemocyanin (Volbeda and Hol, 1989), phospholipase C (Hough *et al.*, 1989), protocatechuate 3,4-dioxygenase (Ohlendorf *et al.*, 1988) and a blue copper protein (Guss *et al.*, 1988). Only one protein was used to represent a family of homologous molecules.

The imidazole side chain of histidine in its neutral form has been known to exist in the two tautomeric forms (Reynolds *et al.*, 1973), depicted in Figure 1. The ND1-protonated-form (tautomer I) can bind metal ions through NE2 (the distal N) and the NE2-protonated form (tautomer II) can coordinate through ND1 (the proximal N).

The relative position of the metal ion was expressed in the coordinate system depicted in Figure 2. The ligand N atom is at the origin. When the coordination is through NE2, the y axis is given by the vector from CD2 to CE1 (Figure 1) and the x axis by the vector from the middle of these points to NE2; the z axis is along the normal to the imidazole plane. Similarly, when the coordination is through the other nitrogen, ND1, CE1 and CG atoms are used to define x and y axes. The geometry of metal–ligand bonding is given by the following spherical polar coordinates: the M–N distance; the acute angle θ between the M–N direction and the z axis; the angle ϕ between the x axis and the projection of the N–M direction on the xy plane. The secondary structural features of all ligand residues were defined according to the method of Kabsch and Sander (1983).

Results and discussion

Tautomeric ratio

Of all the ligand residues considered here, 43 coordinates through NE2 (tautomer I in Figure 1) and 15 through ND1 (tautomer II). Spectroscopic studies have shown that for histidine in basic aqueous solution, tautomer II is energetically more stable than tautomer I and the population ratio of tautomer II to I is ~8:2 (Reynolds *et al.*, 1973; Wasylishen and Tomlinson, 1977; Ashikawa and Itoh, 1978; Tanokura, 1983). Steric factors may be responsible for reversing the preference of the two tautomers when bound to metal ions (see later). All iron atoms have so far been found to be coordinated through NE2.

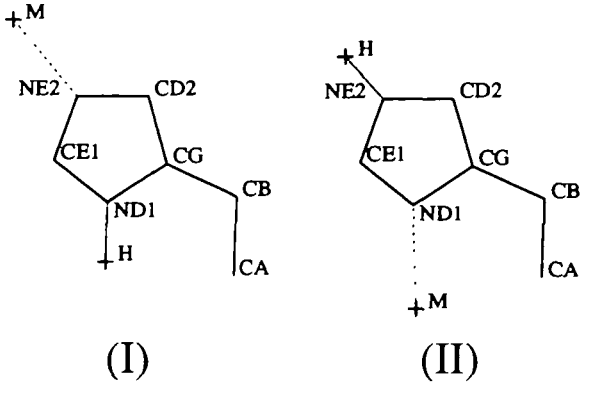

Fig. 1. Two tautomeric forms of the neutral imidazole side chain of histidine. The nomenclature for the atoms is that used in PDB.

Geometry and orientation

All the relevant parameters are given in Table I. Average values for various bond distances are Fe–N 2.1(1), Cu–N 2.09(7) and Zn–N 2.06(8) Å. The deviation of metal ions from the imidazole plane is given by the angle θ; a value of 90° indicates a position on the plane. The angle ϕ indicates the orientation of the metal ions in the imidazole plane. The distribution shown in Figure 2 suggests a clear preference for metal ions to lie close to the C1–N–C2 bisector, the direction of the lone-pair on the N atom. Protons show a similar lone-pair directionality when approaching the nitrogen atom in a five-or six-membered aromatic ring (Vedani and Dunitz, 1985). Metals interacting with carboxylic or peptide carbonyl groups can show deviation from the lone-pair direction (Chakrabarti, 1990a,b). However, in the case of histidine a significant deviation will bring the metal ion closer to a ring carbon atom or to aromatic π electrons. As a result, the orientation of the metal ion with respect to the imidazole ring does not change significantly with the oxidation state or the coordination number of the cation.

Carboxylate anions are known to bridge two different metal ions (Chakrabarti, 1990a). However, only in 2SOD (proteins are referred to by their PDB codes as given in Table I), a His residue as an imidazolate anion has been found to bridge a Cu and a Zn ion.

Location in secondary structure

From an analysis of metal–carboxylate interactions it has been suggested that when a metal ion binds four or more protein residues, it is not likely for such a residue to be located in the middle of a long helix (Chakrabarti, 1990a). However, for a metal like Zn with a low coordination number, the binding site can have a residue from anywhere in the helix. As shown in Table I, a ligand His can occur in the middle of a helix (in 1HMQ and 3TLN) or a β-sheet (3BCL, 2CAB and 2SOD). However, a helix residue binding a heme iron is usually from near the C-terminus

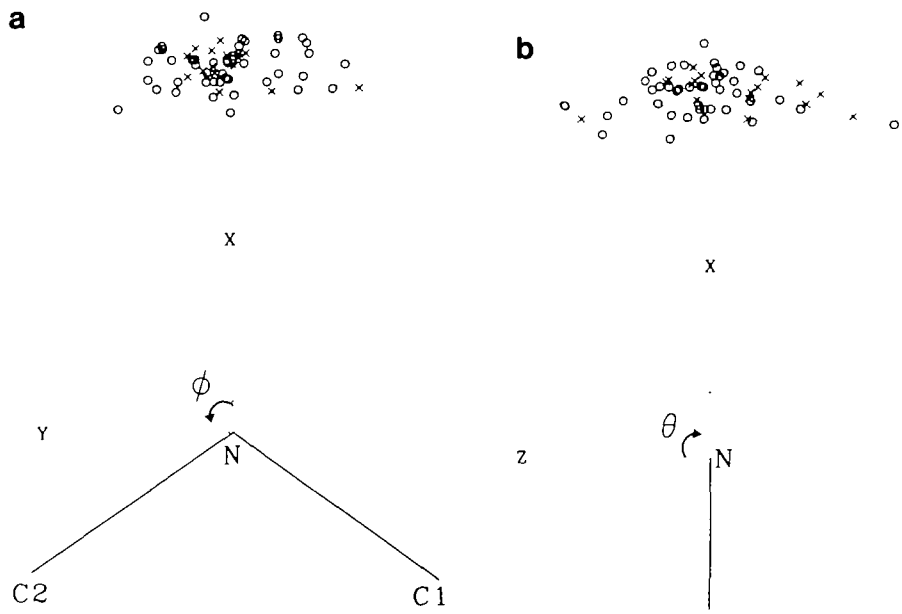

Fig. 2. Distribution of metal ions as viewed from the top (**a**) and the side (**b**) of the imidazole ring. Circles and crosses represent metal ions bound to NE2 and ND1 atoms respectively. The diagrammatic representation of the spherical polar co-ordinates used to express metal–histidine interactions is also shown. When N = NE2 (structure I in Figure 1), C1 is CD2 and C2 is CE1; when N = ND1 (structure II in Figure 1), C1 is CE1, C2 is CG. When θ is >90°, it is made 180° – θ.

Table I. Geometric parameters for metal(M)–histidine interactions

PDB code[a]	Protein name	Metal[b]	Residue[b]	Structure[c]	M–N distance (Å)	Angle (°)[d]		Hydrogen bonded to[e]
						ϕ	θ	
4ADH	alcohol dehydrogenase	Zn(1)	67,NE2	C	2.19	9.5	81 3	OD2,D49
2AZA	azurin	Cu(A)	46(A),ND1	B	2.08	5.0	81 9	O,N10
			117(A),ND1	C	2 01	3 9	86.8	W
		Cu(B)	46(B),ND1		2.09	−1.5	85.6	
			117(B),ND1		1 99	2.1	89.4	
3BCL	bacteriochlorophyll-A protein	Mg(1)	105,NE2	E(4,3)	2.15	12.6	67.6	O,A106
		Mg(3)	290,NE2	H(1)	2.13	−1.5	87.5	O,C345
		Mg(4)	282,NE2	C	2.11	0 0	67.8	O,P283
		Mg(6)	140,NE2	E(3,5)	2.16	−11.7	79.3	–
		Mg(7)	289,ND1	H(2)	2 11	−2.5	82.7	O,A186
2CAB	carbonic anhydrase B	Zn	94,NE2	E(4,4)	1 93	−11.2	86.6	OE1,Q92
			96,NE2	E(4,2)	1 94	−6 1	85 2	O,N244
			119,ND1	E(5,4)	1 90	−7 0	83 5	OE1,E117
5CPA	carboxypeptidase A	Zn	69,ND1	S	2 13	2 4	87 2	OD1,D142
			196,ND1	E(4,2)	2 07	−1.4	85.4	W

Table I. *Continued*

PDB code[a]	Protein name	Metal[b]	Residue[b]	Structure[c]	M–N distance (Å)	Angle (°)[d]		Hydrogen bonded to[e]
						ϕ	θ	
3CNA	concanavalin	Mn	24,NE2	E(2,1)	2 07	18.5	65.7	O,Y22
1ECD	erythrocruorin	Fe	87,NE2	H(1)	2.18	−1.9	84.8	O,F83
3CYT	cytochrome *c*	Fe(A)	18(A),NE2[f]	C	1.96	0 5	89.3	O,P30
		Fe(B)	18(B),NE2		2.04	0.1	88.9	
2B5C	cytochrome *b*5	Fe	39,NE2	C	1.90	8.6	87 5	O,G42
			63,NE2	C	1.78	−0.2	81.7	O,F58
2CCY	cytochrome *c'*	Fe(A)	122(A),NE2[f]	H(4)	2 02	−11.5	80.8	–
2CDV	cytochrome *c*3	Fe(A)	70,NE2	H(1)	1.93	1.3	88.5	O,Y66
			106,NE2[f]	C	1.94	3.6	89.1	W
		Fe(B)	35,NE2	C	2.06	−2.3	85.4	O,P36
			52,NE2[f]	S	2.05	0.2	88.0	O,A62
		Fe(C)	22,NE2	S	2.02	3.4	85.4	W
			34,NE2[f]	S	1.88	−0.9	87.3	W
		Fe(D)	25,NE2	G	2.12	−17.2	82.4	O,N21
			83,NE2[f]	H(5)	1.95	11.5	87.4	O,L97
2CYP	cytochrome *c* peroxidase	Fe	175,NE2	G	1.95	8.0	86.8	OD1,D235
2C2C	cytochrome *c2*	Fe	18,NE2[f]	C	1.99	2.4	79.9	O,P30
351C	cytochrome *c*551	Fe	16,NE2[f]	B	1.99	0.9	84.2	O,P25
1CC5	cytochrome *c*5	Fe	23,NE2[f]	H(1)	2.08	5.0	82.7	O,P31
GS	glutamine synthetase	Mn(2)	269,ND1	E	2.15	−20.3	68.4	–
2HHB	hemoglobin (deoxy)	Fe(A)	87(A),NE2	H(3)	2.15	9.5	89.6	O,L83
		Fe(B)	92(B),NE2	H(3)	2.15	10.1	88.9	O,L88
1HMQ[g]	hemerythrin (Met)	Fe(1)	73,NE2	H(4)	2.30	3.2	88 4	–
			77,NE2	H(8)	2.16	−2.3	89.6	NE2,Q59
			101,NE2	H(4)	2.20	−7.0	86.2	W
		Fe(2)	25,NE2	H(7)	2.20	−11.1	83.7	W
			54,NE2	H(11)	2 19	−7.1	83.2	OE2,E24
1INS	insulin	Zn(1)	10(B),NE2	H(2)	2.06	−0.7	88.4	–
			10(D),NE2		2.05	2.3	85.0	
1MBO	myoglobin (oxy)	Fe	93,NE2	H(3)	2.07	5.4	89.6	O,L89
1PCY	plastocyanin	Cu	37,ND1	B	2.04	−0.9	83.9	O,A33
			87,ND1	G	2.10	2.4	75.8	W
1PAZ	pseudoazurin	Cu	40,ND1	C	2.16	1.1	87.3	OD1,N9
			81,ND1	T	2.13	4.8	88.9	W
1PPT	pancreatic polypeptide	Zn	34,NE2[h]	C	2.08	−0.7	89.5	–
2SOD[g]	superoxide dismutase	Cu	44,ND1	E(4,3)	2.01	1.5	69.7	OD2,D122
			46,NE2	E(4,1)	2.11	11.9	85.5	O,G59
			61,NE2	B	2.21	−10.4	86.3	
			118,NE2	E(3,1)	2.10	2.6	61 1	O,G139
		Zn	61,ND1		2.09	0.2	82.8	
			69,ND1	C	2.14	6.1	77.2	OD1,D122
			78,ND1	C	2.04	6.4	73.1	ND2,N63
3TLN	thermolysin	Zn	142,NE2	H(6)	2.11	−6.9	84.0	OD2,D170
			146,NE2	H(6)	2.08	8 5	87.0	OD1,N165
1TON	tonin	Zn	57,NE2	G	2.04	5.0	84 7	OD2,D102
			97,NE2	C	2.05	−7.8	76.3	O,P95
			99,NE2	S	2.05	−16.4	72.0	–

[a]4ADH, Colonna-Cesari *et al.* (1986), 2AZA, Baker (1988); 3BCL, Tronrud *et al.* (1986), 2CAB, Kannan and Ramanadham (1981); 5CPA, Rees *et al.* (1983), 3CNA, Hardman *et al.* (1982); 1ECD, Steigemann and Weber (1979); 3CYT, Takano and Dickerson (1980); 2B5C, Mathews *et al.* (1972); 2CCY, Finzel *et al.* (1985); 2CDV, Higuchi *et al.* (1984); 2CYP, Finzel *et al.* (1984); 2C2C, Salemme *et al.* (1973); 351C, Matsuura *et al.* (1982); 1CC5, Carter *et al* (1985); GS, Yamashita *et al.* (1989); 2HHB, Fermi *et al.* (1984); 1HMQ, Stenkamp *et al.* (1983); 1INS, Bordas *et al.* (1983); 1MBO, Phillips (1980); 1PCY, Guss and Freeman (1983); 1PAZ, Petratos *et al.* (1988); 1PPT, Blundell *et al.* (1981); 2SOD, Tainer *et al.* (1982); 3TLN, Holmes and Matthews (1982); 1TON, Fujinaga and James (1987).

[b]1, 2 or A, B, etc. are used to distinguish different metal atoms or subunits in the same protein molecule.

[c]As defined by Kabsch and Sander (1983): B, residue in isolated β-bridge; C, non-regular structure; E, extended strand; G, 3_{10}-helix; H, α-helix; S, bend; T, H-bonded turn. For a residue in a helix, its position from the nearest end of the helix is given in parenthesis. Similarly, a residue in a β-sheet is given by E(m,n), where m denotes the strand to which it belongs, counting from the nearest peripheral strand and n is the position of the residue from the nearest end of the strand. When the same residue appears in the table more than once, the secondary structural feature is assigned to one entry only.

[d]As defined in the text.

[e]When the coordination is through NE2, ND1-H can form a bond with a hydrogen bond acceptor, which is given. Similarly for ND1-coordination, the hydrogen-bonded neighbor of NE2 is tabulated W represents a water molecule.

[f]Two Cys residues in the peptide sequence -Cys-X-Y-Cys- preceding the ligand His are involved in covalent bonding to the heme group.

[g]Values for the first subunit are given only.

[h]Crystallographic-symmetry-related position.

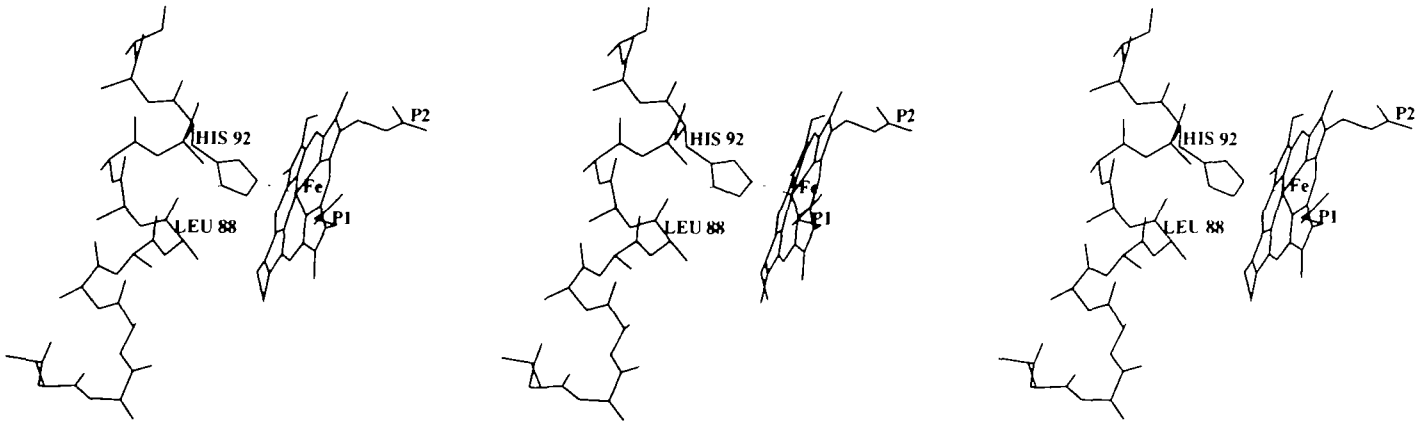

Fig. 3. Stereoview showing the binding of a heme group near the C-terminus of a helix (except for the ligand group, side chain atoms beyond C_β are not shown) in hemoglobin (chain B). NE2 of His92 coordinates (shown by dashed line) to the heme iron. The hydrogen bond involving ND1 of the ligand residue and the O of Leu88 is shown by thin dashed line. Two heme propionates are indicated by P1 and P2.

of the helix, as illustrated by 2HHB (Figure 3). Heme groups in 1ECD, 2CDV [the heme containing Fe(A)] and 1MBO are similarly oriented with respect to a helix, with the heme plane being parallel to the helix axis. The disposition of the heme plane with respect to the helix is given by the side chain torsional angles of the ligand residue and these angles are very similar in all structures (see later). His residues have been found in general to concentrate at the last turn of a helix (Richardson and Richardson, 1988). This has been explained by the electrostatic interaction involving the helix dipole (Ptitsyn, 1969; Shoemaker *et al.*, 1987). However, steric factors may be responsible for the occurrence of heme ligands at the carbonyl ends of helices. When placed near the middle of a helix, heme propionates (Figure 3) may not be properly exposed to the solvent or hydrogen bonded to other protein residues.

In many cytochrome structures the heme is covalently attached to the protein through two cysteine groups in the peptide sequence -Cys-X-Y-Cys-His-, where His is the ligand residue (Mathews, 1985). In 2CCY and 2CDV (residue 83) such a ligand is from a helix. The presence of two covalent bonds makes the angle between the helix axis and the heme plane widen from being close to zero in Figure 3. In all structures where a helical His binds a cation, the coordination is through NE2; coordination through ND1 will put the metal much closer to the helix to be able to bind to other ligand residues; residue 289 in 3BCL is an exception—in this structure both 289 and 290 act as ligands to two different bacteriochlorophyll *a* molecules.

It is interesting to note that in the photosynthetic reaction center (Deisenhofer *et al.*, 1985; Allen *et al.*, 1987), ligands for four bacteriochlorophyll *b* molecules are provided by two transmembrane and two periplasmic helices. In all the cases the ligand His is within three residues from the N-terminal end of the helix. However, none of the histidine ligands for the nonheme iron are from the first or the last turn of a helix. This again suggests that a heme or a chlorophyll molecule is likely to bind to a helical histidine only at the first or the last turn of the helix.

As in the case of other ligands (Chakrabarti, 1989, 1990a,b), many His ligands are from turns and regions with no regular secondary structure. However, a considerable number of ligands are also provided by helices and β-sheets. In 1HMQ and 3TLN, two His ligands from the consecutive turns of a helix bind to the same metal ion. Similarly, in hemocyanin two helical His residues at i and $i + 4$ positions bind to a Cu ion (Volbeda and Hol, 1989) and a similar arrangement has been proposed for the zinc-binding domain from transcription factor IIIA and related proteins (Berg, 1988). In phosphilipase C (Hough *et al.*, 1989), a His at position i and a Glu or an Asp at $i + 4$ bind to the same Zn center; H54 and E58 in 1HMQ are involved in a similar binding. Also His groups two residues apart along a strand can bind to the same cation. This is observed in 2CAB, 2SOD and protocatechuate 3,4-dioxygenase (Ohlendorf *et al.*, 1988). The position of the metal ion with respect to the β-sheet in 2CAB is shown in Figure 4. It appears that two ligand groups (one or both being His) four residues apart in a helix or two residues apart in a β-strand can constitute a part of a metal-binding center.

Table II. Position of a proline within two residues of the ligand histidine

PDB code	His residue	Pro residue[a]
2AZA	117	−2
3BCL	282	+1
3CNA	24	−1
1ECD	87	+2
2B5C	39	+1
2CDV	35	+1
1PCY	37	−1
	87	−1
1PAZ	81	−1
2SOD	61	−1
1TON	97	−2

[a]− and + signs indicate positions preceding and following the ligand residue.

In many structures, a Pro is found within two residues of the ligand group (Table II). The role of Pro appears to be to stabilize the turn conformation (Wilmot and Thornton, 1988) near the metal center, as shown in Figure 5 for 1PCY.

Hydrogen bonding

It has been suggested that hydrogen bonding can stabilize the orientation of a ligand residue (Argos *et al.*, 1978), modulate the oxidation state of the metal center (Valentine *et al.*, 1979) and affect the strength of the metal−ligand bond (Smulevich *et al.*, 1988). Christianson and Alexander (1989) have analyzed the hydrogen bonding between the carboxylate group and ligand His bound to a zinc ion. However, when all cations in known protein structures are considered it is found that a peptide carbonyl is

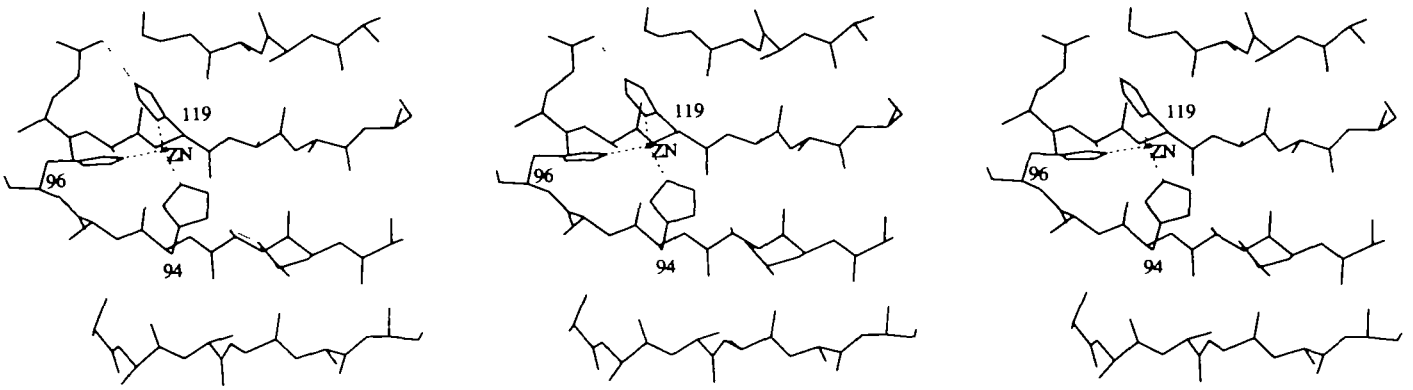

Fig. 4. Stereoview showing the coordination sphere of Zn in carbonic anhydrase B (except for the ligand groups and their hydrogen-bonded neighbors, side chain atoms beyond C_β are not shown). Ligand residues are indicated by their sequence numbers. Metal–ligand bonds are represented by dashed lines. Hydrogen bonds within the β-sheet involving two ligand residues are shown by thin dashed lines.

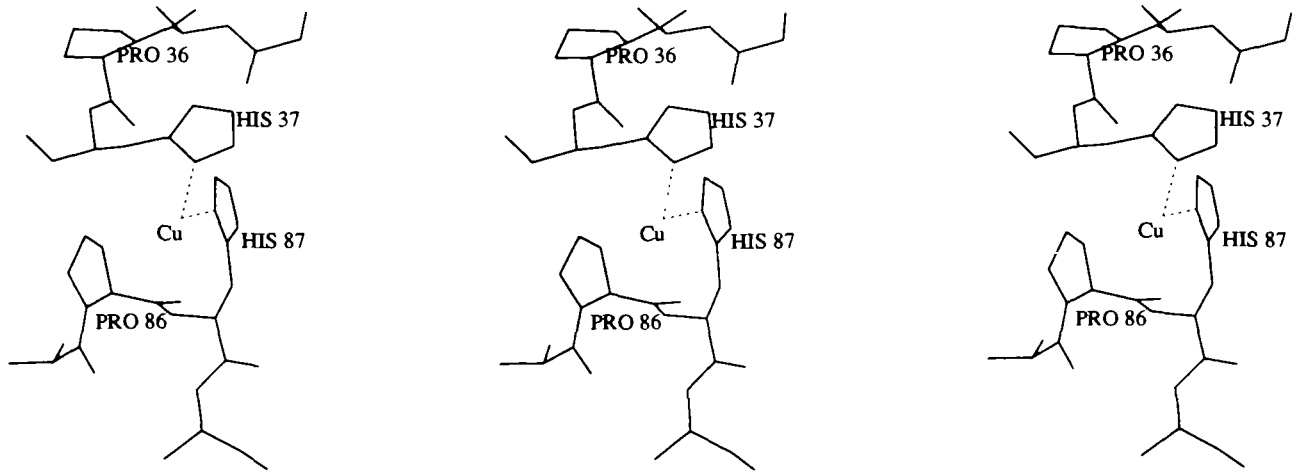

Fig. 5. Stereoview showing ligand histidines and nearby prolines in plastocyanin (other side chain atoms beyond C_β are not shown) Cu–ND1 bonds are represented by dashed lines. Pro86 and His87 are two central residues in a type I β-turn.

also likely to be a hydrogen-bonded neighbor to a ligand His. As can be seen from Table I, when the coordination is through NE2, a peptide carbonyl group usually forms a hydrogen bond to ND1 (22 cases). A few ND1 atoms are hydrogen bonded to side chain atoms (9) or to water molecules (5) and in six cases there are no hydrogen-bonded neighbors. When the coordination is through ND1, the hydrogen bond is through NE2 and since it is farthest from the main chain it always finds a hydrogen-bonded neighbor, usually a polar side chain atom. A group hydrogen bonded to the ligand residue can be labeled as an 'orienter'. In cytochrome *c*-type structures the orienter is a proline carbonyl group. This residue does not belong to any regular secondary structure. A Pro at this position reduces the conformational freedom of the main chain and thereby properly orients the carbonyl group toward the ligand His residue. This Pro is highly conserved in all cytochrome *c* structures (Mathews, 1985).

Some of the ligand histidines located in strands are hydrogen bonded to the side chain of the neighboring residue pointing in the same direction (Figure 4) or even to some nearby peptide carbonyl (residue 105 in 3BCL, 24 in 3CNA).

Side chain conformation

The conformation of the side chain of a His residue is specified by two torsional angles χ_1 (involving N-CA-CB-CG) and χ_2 (CA-CB-CG-ND1). For amino acids in general, the χ_1 distribution is trimodal, the preferred conformation being around χ_1 = 60° (g^-), 180° (t) and 300° (g^+); g^- conformation is the least favorable. For His, the χ_1 distribution is similar to those of other aromatic residues (Trp, Tyr, Phe) and show a preference for the t conformation when the residue belongs to a helix (McGregor *et al.*, 1987). χ_2 shows a preference for values near 90 or 270° (Figure 6); at these angles the imidazole ring is perpendicular to the CA-CB-CG plane and the steric interactions between the δ-atoms on the one hand, and the α-carbon atom and the hydrogen atoms attached to the β-carbon atom on the other, is a minimum (Bhat *et al.*, 1979; Janin *et al.*, 1978). The χ_1/χ_2 map for ligand His residues is shown in Figure 7.

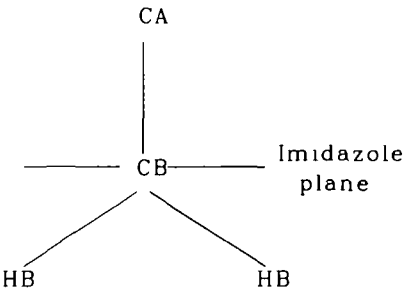

Fig. 6. The favorable conformation of His side chain with the imidazole ring being perpendicular to the plane defined by atoms CA, CB and CG. For this orientation, χ_2 is 90 or 270°

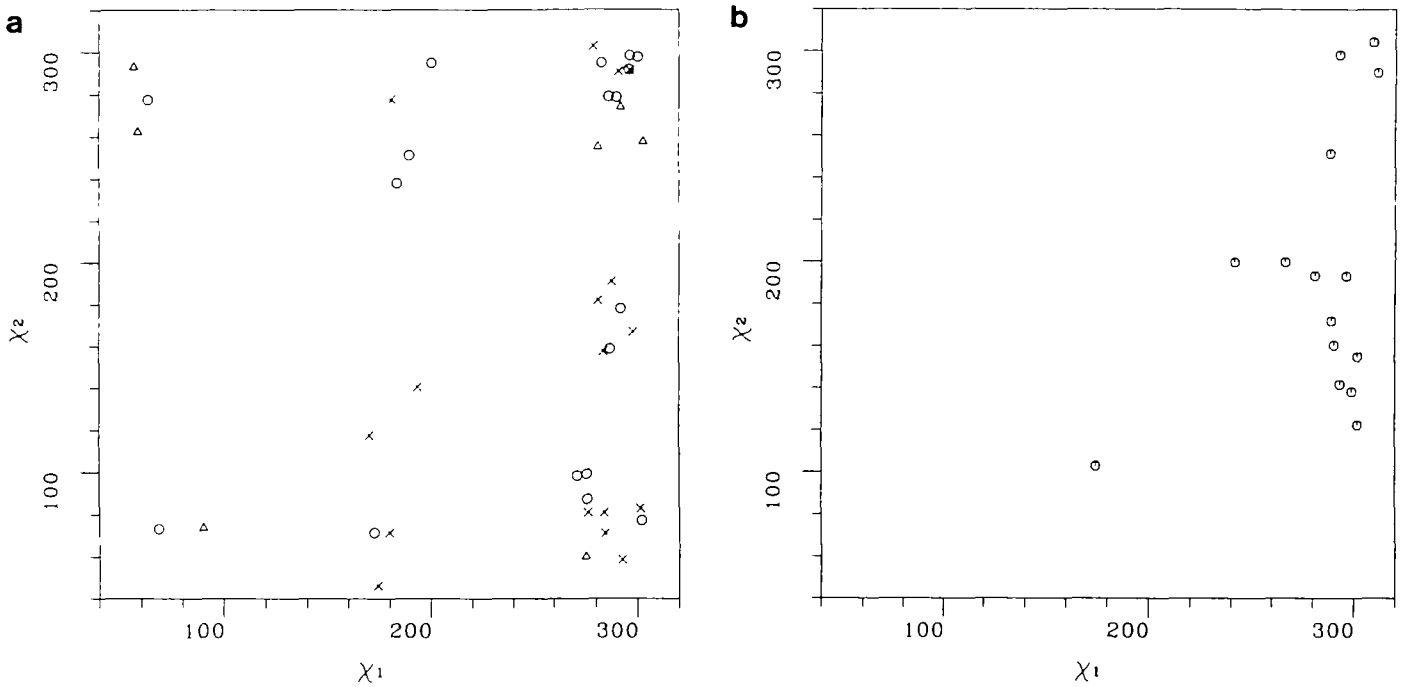

Fig. 7. Joint distribution of χ_1 and χ_2 in ligand histidines Values for residues coordinating through NE2 and ND1 are plotted in (**a**) and (**b**) respectively Residues from α-helix, β-sheet and non-α/non-β structures are represented by crosses, triangles and circles respectively in (a), no such distinction is made in (b).

Helical ligands coordinating through NE2, like those in 1ECD, 2CDV (residue 70), 2HHB and 1MBO have χ_1 in g^+ conformation and χ_2 around 90°. In this conformation a His residue at position i in the tautomeric form (I) (Figure 1) can form a hydrogen bond involving its ND1-H atom and the carbonyl group at position $i - 4$ in a helix. Such a structure, shown in Figure 3, is similar to those formed by Ser and Thr by making hydrogen bonds to carbonyl oxygen atoms in the preceding turn of the helix (Gray and Matthews, 1984). However, not all helical residues are involved in this type of hydrogen bonding. In fact two ligands (2CCY and 2CDV, residue 83) to heme iron, where the heme group is covalently attached to the same helix, have χ_2 angles around the unfavorable value of 180°. Three residues from β-sheet hydrogen bonded to adjacent groups have χ_1 in the less favorable g^- conformation.

For all the cases involving ND1-coordination, χ_1 conformation is predominantly g^+ (Figure 7B). However, the stable χ_2 values of 90 and 270° (Figure 6) are rarely observed. At these angles the ND1 atom is pointing in toward the main chain. A metal bound at this position will be close to main chain atoms and other ligands may not approach it in proper orientation. As a result, χ_2 takes up less stable conformations. This may be a factor why so few ND1-coordinated metals are found in nature.

This and other similar studies (Chakrabarti, 1989, 1990a,b) have shown that there are distinct orientations in which ligand groups bind metal ions in protein structures and these are independent of the type of metal ions. The geometry observed in such interactions can be used as constraints during the refinement (Hendrickson and Konnert, 1980; Brünger *et al.*, 1987) of structures with poor resolution of data. Secondary structural features of ligands should be useful in introducing metal-binding sites in proteins and also to predict such sites in unknown structures, as has been done for the zinc finger (Berg, 1988). We are analyzing now how anions are bound in proteins. Understanding the nature of ion-binding sites in protein structures will be an important step towards understanding molecular recognition (Rebek, 1987).

Acknowledgements

I am grateful to Prof. D Rees for discussion and to Prof D Eisenberg and Dr M Yamashita for providing coordinates of glutamine synthetase.

References

Allen,J.P., Feher,G., Yeates,T.O., Komiya,H. and Rees,D.C. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6162–6166.
Argos,P., Garavito,R.M., Eventoff,W., Rossmann,M G. and Brandén,C.I. (1978) *J Mol. Biol.*, **126**, 141–158.
Ashikawa,I. and Itoh,K. (1978) *Chem. Lett.*, 681–684.
Baker,E.N (1988) *J. Mol. Biol.*, **203**, 1071–1095.
Berg,J M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 99–102
Bernstein,F.C., Koetzle,T.F., Williams,G.J B., Meyer,E F ,Jr, Brice,M.D , Rodgers,J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) *J. Mol. Biol.*, **112**, 535–542
Bhat,T.N., Sasisekharan,V. and Vijayan,M. (1979) *Int. J Peptide Protein Res.*, **13**, 170–184
Blundell,T L., Pitts,J E , Tickle,I.J , Wood,S.P and Wu,C.-W (1981) *Proc. Natl. Acad Sci. USA*, **78**, 4175–4179.
Brünger,A.T., Kuriyan,J. and Karplus,M. (1987) *Science*, **235**, 458–460
Bordas,J., Dodson,G.G., Grewe,H., Koch,M.H.J., Krebs,B. and Randall,J. (1983) *Proc. R. Soc. Lond., Ser B*, **219**, 21–39
Carter,D.C., Melis,K.A., O'Donnell,S.E., Burgess,B.K., Furey,W.F ,Jr. Wang,B. and Stout,C.D. (1985) *J. Mol. Biol.*, **184**, 279–295.
Chakrabarti,P. (1989) *Biochemistry*, **28**, 6081–6085
Chakrabarti,P. (1990a) *Protein Engng*, **4**, 49–56.
Chakrabarti,P. (1990b) *Biochemistry*, **29**, 651–658
Christianson,D W. and Alexander,R.S (1989) *J. Am. Chem. Soc.*, **111**, 6412–6419
Colonna-Cesari,F., Perahia,D., Karplus.M., Eklund,H , Brändén,C I and Tapia,O. (1986) *J. Biol. Chem.*, **261**, 15273–15280.
Deisenhofer,J., Epp,O., Miki,K., Huber,R. and Michel,H. (1985) *Nature*, **318**, 618–624
Fermi,G , Perutz,M F and Shaanan,B. (1984) *J. Mol Biol.*, **175**, 159–174.
Finzel,B.C , Poulos,T.L. and Kraut,J. (1984) *J. Biol. Chem.*, **259**, 13027–13036
Finzel,B.C , Weber P C , Hardman,K D and Salemme,F.R. (1985) *J Mol Biol*, **186**, 627–643.

Fujinaga,M and James,M.N.G. (1987) *J. Mol Biol.*, **195**, 373–396
Gray,T.M. and Matthews,B.W. (1984) *J. Mol. Biol.*, **175**, 75–81
Guss,J.M. and Freeman,H.C. (1983) *J. Mol. Biol.*, **169**, 521–563.
Guss,J.M., Merritt,E.A., Phizackerley,R.P., Hedman,B., Murata,M., Hodgson,K.O. and Freeman,H.C. (1988) *Science*, **241**, 806–811.
Hardman,K D., Agarwal,R.C. and Freiser,M J (1982) *J Mol Biol*, **157**, 69–86.
Hendrickson,W A. and Konnert,J.H. (1980) In Diamond,R, Ramaseshan,S. and Venkatesan,K (eds), *Computing in Crystallography*. Indian Academy of Sciences, Bangalore, pp. 13.01–13.23.
Higuchi,Y., Kusunoki,M., Matsuura,Y, Yasuoka,N. and Kakudo,M (1984) *J. Mol Biol.*, **172**, 109–139.
Holmes,M.A. and Matthews,B.W. (1982) *J. Mol. Biol.*, **160**, 623–639.
Hough,E., Hansen,L K., Birknes,B., Jynge,K., Hansen,S., Hordvik,A, Little,C., Dodson,E. and Derewenda,Z. (1989) *Nature*, **338**, 357–360
Janin,J., Wodak,S., Levitt,M. and Maigret,B. (1978) *J. Mol. Biol.*, **125**, 357–386.
Kabsch,W. and Sander,C (1983) *Biopolymers*, **22**, 2577–2637
Kannan,K.K. and Ramanadham,M. (1981) *Int. J. Quantum Chem*, **20**, 199–209
Lederer,F., Glatigny,A., Bethge,P H., Bellamy,H D. and Mathews,F S. (1981) *J. Mol Biol.*, **148**, 427–448.
Mathews,F.S. (1985) *Prog. Biophys. Mol. Biol.*, **45**, 1–56.
Mathews,F S., Levin,M. and Argos,P. (1972) *J. Mol. Biol.*, **64**, 449–464.
Matsuura,Y., Takano,T. and Dickerson,R.E. (1982) *J. Mol. Biol*, **156**, 389–409
Mayo,S.L., Ellis,W.R.,Jr, Crutchley,R.J. and Gray,H.B. (1986) *Science*, **233**, 948–952
McGregor,M J., Islam,S A and Sternberg,M.J.E. (1987) *J. Mol. Biol.*, **198**, 295–310.
Ohlendorf,D.H., Lipscomb,J.D. and Weber,P.C. (1988) *Nature*, **336**, 403–405.
Osvath,P., Salmon,G.A. and Sykes,A.G. (1988) *J. Am. Chem. Soc.*, **110**, 7114–7118.
Petratos,K., Dauter,Z. and Wilson,K S. (1988) *Acta Crystallogr, B*, **44**, 628–636.
Phillips,S E.V. (1980) *J. Mol. Biol.*, **142**, 531–554.
Ptitsyn,O.B. (1969) *J. Mol. Biol.*, **42**, 501–510.
Rebek,J.,Jr (1987) *Science*, **235**, 1478–1484
Rees,D.C., Lewis,M. and Lipscomb,W.N. (1983) *J. Mol. Biol*, **168**, 367–387.
Reynolds,W.F., Peat,I.R., Freedman,M.H. and Lyerla,J.R.,Jr (1973) *J Am. Chem. Soc.*, **95**, 328–331
Richardson,J.S. and Richardson,D C. (1988) *Science*, **240**, 1648–1652
Salemme,F R., Freer,S.T., Xuong,N.H., Alden,R.A. and Kraut,J. (1973) *J Biol. Chem.*, **248**, 3910–3921
Shoemaker,K.R., Kim,P S., York,E.J., Stewart,J.M. and Baldwin,R L (1987) *Nature*, **326**, 563–567.
Smulevich,G., Mauro,J.M, Fishel,L.A, English,A.M, Kraut,J and Spiro,T.G. (1988) *Biochemistry*, **27**, 5477–5485
Sowadski,J.M., Handschumacher,M.D, Murthy,H.M.K., Foster,B.A and Wyckoff,H.W. (1985) *J. Mol Biol*, **186**, 417–433.
Steigemann,W. and Weber,E. (1979) *J. Mol. Biol.*, **127**, 309–338.
Stenkamp,R.E., Sieker,L C and Jensen,L.H (1983) *Acta Crystallogr B*, **39**, 697–703.
Sundberg,R J and Martin,R.B. (1974) *Chem. Rev.*, **74**, 471–517.
Tainer,J.A., Getzoff,E.D., Beem,K.M, Richardson,J S and Richardson,D.C. (1982) *J. Mol. Biol.*, **160**, 181–217.
Takano,T. and Dickerson,R.E. (1980) *Proc. Natl. Acad Sci. USA*, **77**, 6371–6375
Tanokura,M. (1983) *Biochim. Biophys. Acta*, **742**, 576–585
Tronrud,D.E., Schmid,M.F. and Matthews,B W (1986) *J. Mol. Biol*, **188**, 443–454.
Valentine,J.S., Sheridan,R.P, Allen,L.C and Kahn,P C. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1009–1013.
Vedani,A. and Dunitz,J.D (1985) *J Am. Chem Soc.*, **107**, 7653–7658.
Volbeda,A and Hol,W.G.J. (1989) *J. Mol. Biol.*, **206**, 531–546.
Wasylishen,R E. and Tomlinson,G. (1977) *Can. J Biochem.*, **55**, 579–582
Wilmot,C.M. and Thornton,J M. (1988) *J. Mol. Biol.*, **203**, 221–232.
Yamashita,M.M., Almassy,R.J., Janson,C.A., Cascio,D. and Eisenberg,D (1989) *J. Biol. Chem.*, **264**, 17681–17690.
Zvelebil,M.J.J.M. and Sternberg,M.J.E (1988) *Protein Engng*, **2**, 127–138.

Received on March 1, 1990, accepted on June 15, 1990