Cation Binding by the Phenolate Group in Small Molecules and Proteins†

Pinak Chakrabarti*

Physical Chemistry Division, National Chemical Laboratory, Pune-411008, India

Barbara T. Hsu

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received August 4, 1993®

The geometry of interactions of metal ions with the phenolate group in proteins and small molecules has been examined using coordinates listed in structural databases. Cations are found to avoid the sp^2 lone pair directions of the ligand oxygen; in small-molecule structures most of the cations are clustered close to the aromatic plane in the region between the sp^2 and the C–O vectors, whereas in proteins the ions move out of the plane. Such a spatial distribution is different from that observed for hydrogen-bonded neighbors interacting with a neutral tyrosine residue. The environment of a tyrosinate anion is such that it has a cation on one side of its plane and a hydrogen-bond donor on the other. Metal binding can disturb the normal conformation of the tyrosine side chain.

Introduction

Compounds with phenolate ligand bound to metal ions have been studied for their catalytic properties.[1] The corresponding amino acid, tyrosine, is also found at the enzyme active site coordinated to cations; besides, phenolic derivatives serve as the substrate or the product of many metalloenzymes.[2] The tyrosinate anion may have a role in the light-driven proton translocation in bacteriorhodopsin[3] and the photosynthetic process in reaction centers.[4] A metal center can stabilize a nearby tyrosine radical[5] and can exhibit phenolate-to-cation charge transfer.[6] The stereochemistry of metal–tyrosine interaction is important in understanding the function of the metal ion in these biological processes.

In recent years antibodies capable of binding cations using histidine ligands have been produced.[7] As tyrosines are predominant among the antigen-contacting residues in antibodies,[8] such groups can be utilized for metal ligation, provided the various interactions that stabilize the tyrosinate anion can be incorporated into the structure. In this paper we analyze the geometry of interactions of the phenolate group with various cations and hydrogen-bond donors on the basis of the data bases of small molecules and protein crystal structures. This and similar studies involving other ligands[9-12] aim at establishing a set of stereochemical blueprints for metal binding in proteins. These geometric features can be used to develop biosensors for cations and to introduce metal-binding sites into peptides and proteins for structural stability,[13,14] spectroscopic labeling,[15] X-ray crystallography,[16] and regulation of catalytic function.[17]

Methodology

All metal–phenolate interactions were transformed into a common orientation given by the coordinate system as illustrated in Figure 1. The ligand oxygen atom is at the origin with the phenyl–oxygen bond direction representing the x axis. The xy plane is coincident with the aromatic ring, normal to which is the z axis. The position of the metal (M) ion with respect to the phenolate group is determined by three parameters: the M–O distance; the angle θ between the M–O direction and the z axis; the angle ϕ between the x axis and the projection of the O–M direction onto the xy plane. The phenolate group has mm symmetry, with one mirror being in the plane of the ring and another perpendicular to it and containing the phenolic bond. The cationic coordinates were reflected through these mirrors to bring them into the $(\pm x, +y, +z)$ sector (i.e., θ and ϕ in the range 0–90°). However, the symmetry is broken if one considers the C_α position of the tyrosine group. As a result, the disposition

† The paper is dedicated to Professor C. N. R. Rao on the occasion of his 60th birthday. Support for the work (NCL Communication No. 5408) was provided by a grant from the Department of Science and Technology (India).
® Abstract published in *Advance ACS Abstracts,* February 1, 1994.

(1) (a) Malhotra, K. C.; Martin, R. L. *J. Organomet. Chem.* **1982,** *239,* 159–187. (b) Kim, Y.; Osakada, K.; Takenaka, A.; Yamamoto, A. *J. Am. Chem. Soc.* **1990,** *112,* 1096–1104. (c) Chisholm, M. H. *Polyhedron* **1983,** *2,* 681–721.
(2) Hughes, M. N. In *Comprehensive Coordination Chemistry*; Wilkinson, G., Gillard, R. D., McCleverty, J. A., Eds.; Pergamon: Oxford, U.K., 1987; Vol. 6, pp 541–754.
(3) (a) Braiman, M. S.; Mogi, T.; Stern, L. J.; Hackett, N. R.; Chao, B. H.; Khorana, H. G.; Rothschild, K. J. *Proteins: Struct. Funct. Genet.* **1988,** *3,* 219–229. (b) Roepe, P.; Ahl, P. L.; Das Gupta, S. K.; Herzfeld, J.; Rothschild, K. J. *Biochemistry* **1987,** *26,* 6696–6707. (c) Dollinger, G.; Eisenstein, L.; Lin, S.-L.; Nakanishi, K.; Termini, J. *Biochemistry* **1986,** *25,* 6524–6533. (d) McDermott, A. E.; Thompson, L. K.; Winkel, C.; Farrar, M. R.; Pelletier, S.; Lugtenburg, J.; Herzfeld, J.; Griffin, R. G. *Biochemistry* **1991,** *30,* 8366–8371.
(4) de Groot, H. J. M.; Raap, J.; Winkel, C.; Hoff, A. J.; Lugtenburg, J. *Chem. Phys. Lett.* **1990,** *169,* 307–310.
(5) Prince, R. C. *Trends Biochem. Sci.* **1988,** *13,* 286–288.
(6) Que, L., Jr. *J. Chem. Educ.* **1985,** *62,* 938–943.
(7) (a) Roberts, V. A.; Iverson, B. L.; Iverson, S. A.; Benkovic, S. J.; Lerner, R. A.; Getzoff, E. D.; Tainer, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1990,** *87,* 6654–6658. (b) Iverson, B. L.; Iverson, S. A.; Roberts, V. A.; Getzoff, E. D.; Tainer, J. A.; Benkovic, S.J.; Lerner, R. A. *Science* **1990,** *249,* 659–662.
(8) (a) Padlan, E. A. *Proteins: Struct. Funct. Genet.* **1990,** *7,* 112–124. (b) Mian, I. S.; Bradwell, A. R.; Olson, A. J. *J. Mol. Biol.* **1991,** *217,* 133–151.
(9) Chakrabarti, P. *Biochemistry* **1989,** *28,* 6081–6085.
(10) Chakrabarti, P. *Biochemistry* **1990,** *29,* 651–658.
(11) (a) Chakrabarti, P. *Protein Eng.* **1990,** *4,* 49–56. (b) Chakrabarti, P.; Pal, S. *Chem. Phys. Lett.* **1993,** *201,* 24–26.
(12) Chakrabarti, P. *Protein Eng.* **1990,** *4,* 57–63.
(13) (a) Ghadiri, M. R.; Choi, C. *J. Am. Chem. Soc.* **1990,** *112,* 1630–1632. (b) Handel, T.; DeGrado, W. F. *J. Am. Chem. Soc.* **1990,** *112,* 6710–6711. (c) Ruan, F.; Chen, Y.; Hopkins, P. B. *J. Am. Chem. Soc.* **1990,** *112,* 9403–9404. (d) Krizek, B. A.; Amann, B. T.; Kilfoil, V. J.; Merkle, D. L.; Berg, J. M. *J. Am. Chem. Soc.* **1991,** *113,* 4518–4523.
(14) Pace, C. N. *Trends Biochem. Sci.* **1990,** *15,* 14–17.
(15) (a) Williams, R. J. P. *Chem. Br.* **1978,** *14,* 25–29. (b) Legg, J. I. *Coord. Chem. Rev.* **1978,** *25,* 103–132.
(16) (a) Hatfull, G. F.; Sanderson, M. R.; Freemont, P. S.; Raccuia, P. R.; Grindley, N. D. F.; Steitz, T. A. *J. Mol. Biol.* **1989,** *208,* 661–667. (b) Klein, C.; Vogel, W.; Bender, H.; Schulz, G. E. *Protein Eng.* **1990,** *4,* 65–67. (c) Dao-Pin, S.; Alber, T.; Bell, J. A.; Weaver, L. H.; Matthews, B. W. *Protein Eng.* **1987,** *1,* 115–123. (d) Tucker, A. D.; Baty, D.; Parker, M. W.; Pattus, F.; Lazdunski, C.; Tsernoglou, D. *Protein Eng.* **1989,** *2,* 399–405.
(17) Higaki, J. N.; Haymore, B. L.; Chen, S.; Fletterick, R. J.; Craik, C. S. *Biochemistry* **1990,** *29,* 8582–8586.

0020-1669/94/1333-1165$04.50/0

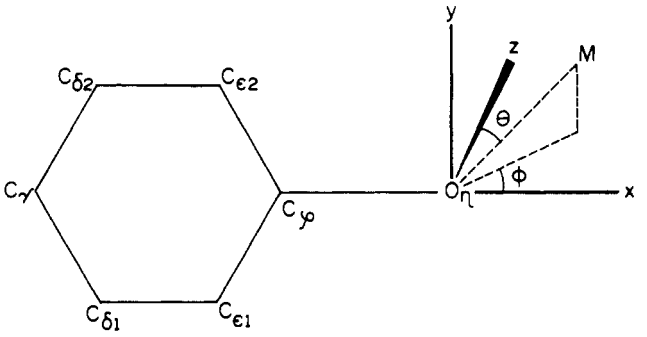

Figure 1. Molecular reference frame used to express metal (M)–phenolate interactions. The atom labels are those for a tyrosine residue, for which the substituent C_β at C_γ is planar to the aromatic ring, but the position of the main-chain atom, C_α is not fixed in space.

of the bound ion relative to the C_α location can be obtained by using the axial system depicted in Figure 1 or by reversing the directions of the y and z axes so as to always have the C_α atoms in the hemisphere with positive z values.

Structures containing the cation–phenolate moiety were retrieved from the Cambridge Structural Database (CSD)[18] for small molecules. This was achieved by first extracting the fragment shown in Figure 1 (where M is any metal ion) using the CSD connectivity search program CONNSER. Only the error-free entries with the estimated standard deviation of bonds less than 0.02 Å and no disorder were considered. In order to exclude the bias in the result emanating from the steric interaction with ortho substituents, all such structures were excluded. This also removed many ligands that bind cations by chelation, an interaction that imposes severe restriction on the geometry and might skew the result. Our survey did not bar bridging phenolates (i.e., a single ligand binding two or more cations), as such systems may have biochemical significance.[19] Atomic coordinates of the matched fragments were retrieved and calculations of geometric parameters performed with the program GSTAT89. All the entries in the resulting output were inspected individually to determine the metal ion involved and if the binding mode was bridging. A few cases of coordination by phenolic ethers were picked up. These and symmetrically redundant entries were eliminated. The final list contains 168 fragments (about 50% of which has bridging ligands) in 58 structures. The individual metal counts are as follows: lithium, 10; sodium, 5; potassium, 3; strontium, 9; aluminum, 4; tin, 4; titanium, 9; zirconium, 4; vanadium, 3; niobium, 11; tantalum, 6; molybdenum, 20; tungsten, 9; rhenium, 9; iron, 21; ruthenium, 8; cobalt, 1; rhodium, 4; iridium, 1; nickel, 1; palladium, 4; platinum, 2; copper, 6; zinc, 3; uranium, 11. All the phenolate fragments were superposed to derive the metal coordinates with reference to a common set of inertial axes. Details of input files to various programs, codes of the structures used in the analysis, and a list of the calculated geometric parameters and coordinates of the cation for each entry have been deposited as supplementary material.

Among the coordinates of the macromolecules, the refined coordinates for catalase[20] were taken from the Brookhaven Protein Data Bank (PDB).[21] The others were kindly provided by the respective authors. These are the coordinates for protocatechuate 3,4-dioxygenase (3,4-PCD),[22] lactoferrin,[23] and galactose oxidase.[24]

Results and Discussion

Geometry of Cation Coordination in Small Molecules. Consideration of the available three-dimensional space suggests that

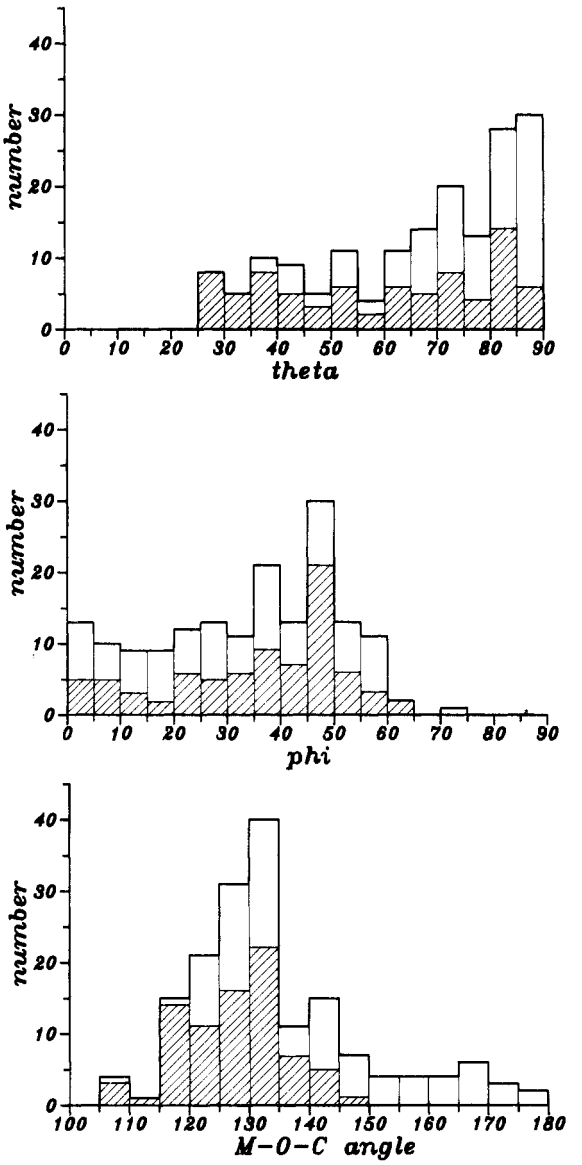

Figure 2. Histograms representing the distribution of various parameters in degrees: (a, top) θ, the out of plane deviation; (b, middle) ϕ, the deviation in the plane from the C–O direction; (c, bottom) M–O–C angle. The shaded area corresponds to bridging ligands.

there should be more cases with cations near the aromatic plane ($\theta = 90°$) than perpendicular to it ($\theta = 0°$),[25] the expected average angle for a random distribution being ~57°. This is approximately what is observed (Figure 2a), with a mean value 67(19)° for θ. However, the planar location has a chemical significance also. A concomitant value of $\phi = 60°$ indicates an interaction along the sp^2 lone pair on the phenolate oxygen. Although there exist a number of examples with ϕ close to 60° (Figure 2b), only a few are exactly along this direction, the distribution about which is very unsymmetrical (mean $\phi = 34(17)°$ and not many with $\phi > 60°$). Some of the cations, especially those belonging to transition metals from the 3d series, are believed to have a significant covalent character in their interaction with organic ligands. However, irrespective of the nature of the ion, the lone pair direction is generally avoided. Electrostatic (discussed in the next section) and steric interactions with the proximal ortho hydrogen may be responsible for this. A scatterplot of cations superimposed upon a reference phenoxide is shown in Figure 3.

The mean of the distribution of the M–O–C angles (Figure 2c) is 135(14)°. However, there are a few cases with a more

(18) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummulink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr.* **1979**, *B35*, 2331–2339.
(19) Solomon, E. I.; Penfield, K. W.; Wilcox, D. E. *Struct. Bonding (Berlin)* **1983**, *53*, 1–57.
(20) Fita, I.; Rossmann, M. G. *J. Mol. Biol.* **1985**, *185*, 21–37.
(21) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535–542.
(22) Ohlendorf, D. H.; Lipscomb, J. D.; Weber, P. C. *Nature* **1988**, *336*, 403–405.
(23) Anderson, B. F.; Baker, H. M.; Norris, G. E.; Rice, D. W.; Baker, E. N. *J. Mol. Biol.* **1989**, *209*, 711–734.
(24) Ito, N.; Phillips, S. E. V.; Stevens, C.; Ogel, Z. B., McPherson, M. J.; Keen, J. N.; Yadav, K. D. S.; Knowles, P. F. *Nature* **1991**, *350*, 87–90.
(25) Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc.* **1983**, *105*, 5761–5766.

Figure 3. Stereodiagram of superimposed metal–phenolate interactions in small-molecule structures.

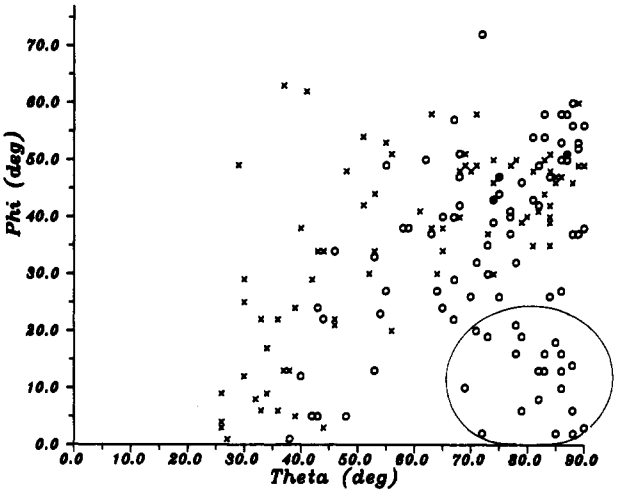

Figure 4. Plot of ϕ vs θ. Terminal and bridging phenolates are represented by circles and crosses, respectively. The points in the enclosure have close to a linear M–O–C arrangement.

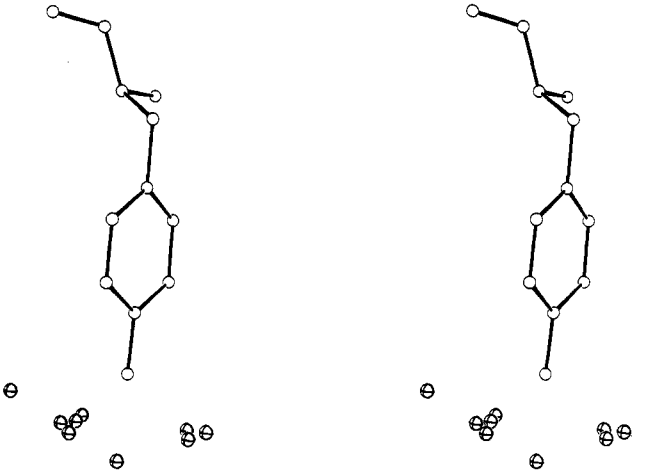

Figure 5. Scatterplot of metal-tyrosinate interactions in proteins. As discussed under Methodology, the orientation of the aromatic ring with C_α going down has been chosen; the main-chain atoms in the diagram are taken from catalase.

linear arrangement. It has been suggested that this geometry permits a greater degree of $d\pi$–$p\pi$ overlap, and a shorter M–O distance occurs due to the resultant multiple bonding.[1c,26,27] It seems in such cases the deviation from a linear M–O–C unit is not more than 35°, and ϕ and $(90° - \theta)$ values are within 25°. Barring these points in Figure 4, there is not much difference between the binding geometries of bridging and terminal phenoxide ligands (although when both types are present in the same molecule,[26,28] the former is reported to be at a longer distance from the cation than the latter). Roughly, as the metal ions move away from the aromatic plane, the angle ϕ also decreases in magnitude. The majority of the triply bridging ligands[29] have two cations bound at approximately $\theta = 70°$ and $\phi = \pm 50°$ (i.e., close to the two oxygen sp^2 lone pairs), and the third one is roughly associated with the angles 25 and 0°, respectively.

In nine cases, the phenoxide, in addition to metal coordination, also forms a very strong hydrogen bond with an uncoordinated phenol, alcohol, or water molecule.[1b,30] These exhibit coordination close to the sp^2 lone pair of the oxygen atom with average values of 125(4), 75(14), and 51(11)° for M–O–C, θ, and ϕ, respectively. However, when the phenol is also a ligand[26b,29d,30c] (usually at a distance longer than that to a phenoxide), the constraint of two simultaneous interactions causes a somewhat greater deviation from the lone-pair direction (the three angles based on five examples are 131(6), 62(15), and 38(13)° respectively).

Relative Position of the Cations in Proteins. The distribution of cations around the tyrosine group is shown pictorially in Figure 5 and expressed quantitatively in terms of the spherical polar coordinates in Table 1. All the bond lengths are fairly similar, except the one involving residue 495 in galactose oxidase. The M–O–C angle shows some variation, with an average of 131(16)°. In contrast to the case of small-molecule structures (with a large number of cations lying in, or near, the aromatic plane), there is a distinct preference for the metal ions to be away from the plane (average $\theta = 49(13)°$) as well as from the trigonal lone-pair direction at the oxygen atom (mean $\phi = 27(19)°$). This geometry also differs from the stereochemistry of hydrogen bonding involving the tyrosine side chain. Such interactions are clustered symmetrically about directions corresponding to sp^2

(26) (a) Lewis, L. N.; Garbauskas, M. F. *Inorg. Chem.* **1985**, *24*, 363–366. (b) Svetich, G. W.; Voge, A. A. *Acta Crystallogr.* **1972**, *B28*, 1760–1767.
(27) (a) Coffindaffer, T. W.; Rothwell, I. P.; Huffman, J. C. *Inorg. Chem.* **1983**, *22*, 2906–2910. (b) Handy, L. B.; Fair, C. K. *Inorg. Nucl. Chem. Lett.* **1975**, *11*, 496–500. (c) Lubben, T. V.; Wolczanski, P. T. *J. Am. Chem. Soc.* **1987**, *109*, 424–435.
(28) Coffindaffer, T. W.; Niccolai, G. P.; Powell, D.; Rothwell, I. P.; Huffman, J. C. *J. Am. Chem. Soc.* **1985**, *107*, 3572–3583.
(29) (a) Bhaduri, S.; Sapre, N.; Sharma, K.; Jones, P. G.; Carpenter, G. *J. Chem. Soc., Dalton Trans.* **1990**, 1305–1311. (b) Boersma, J.; Spek, A. L.; Noltes, J. G. *J. Organomet. Chem.* **1974**, *81*, 7–15. (c) Zozulin, A. J.; Moody, D. C.; Ryan, R. R. *Inorg. Chem.* **1982**, *21*, 3083–3086. (d) Drake, S. R.; Streib, W. E.; Chisholm, M. H.; Caulton, K. G. *Inorg. Chem.* **1990**, *29*, 2707–2708.
(30) (a) Braga, D.; Sabatino, P.; Di Bugno, C.; Leoni, P.; Pasquali, M. *J. Organomet. Chem.* **1987**, *334*, C46–C48. (b) Kegley, S. E.; Schaverien, C. J.; Freudenberger, J. H.; Bergman, R. G.; Nolan, S. P.; Hoff, C. D. *J. Am. Chem. Soc.* **1987**, *109*, 6563–6565. (c) Fraser, M. E.; Fortier, S.; Markiewicz, M. K.; Rodrigue, A.; Bovenkamp, J. W. *Can. J. Chem.* **1987**, *65*, 2558–2563. (d) Calderazzo, F.; Marchetti, F.; Dell'Amico, G.; Pelizzi, G.; Colligiani, A. *J. Chem. Soc., Dalton Trans.* **1980**, 1419–1424. (e) Darensbourg, D. J.; Sanchez, K. M.; Reibenspies, J. H.; Rheingold, A. L. *J. Am. Chem. Soc.* **1989**, *111*, 7094–7103.

Table 1. Geometric Parameters for Metal (M)–Tyrosine Interactions

protein	code[a]	metal[b]	residue no.	M–O (Å)	M–O–C (deg)	angle (deg) θ	angle (deg) ϕ	torsion (deg)[c] χ_1	torsion (deg)[c] χ_2	disposition[d]
catalase	7CAT	Fe	357	1.84	136.3	45.2	1.35	−78.6	67.8	+
protocatechuate 3,4-dioxygenase[e]	PCD	Fe	108	2.00	129.0	45.4	28.5	−79.6	87.7	−
			147	1.96	162.0	79.2	15.0	58.2	80.7	−
lactoferrin	LF	Fe(1)	92	2.19	118.6	41.1	40.2	58.7	97.8	−
			192	1.89	142.0	58.8	23.5	−54.3	110.2	+
		Fe(2)	435	1.85	126.6	46.7	37.3	40.7	106.1	−
			528	2.03	132.8	45.3	15.0	−71.4	134.0	+
galactose oxidase	GOX	Cu	272	1.93	130.7	43.0	16.6	−68.3	15.2	+
			495	2.69	104.7	38.5	65.6	−166.0	58.0	+

[a] Proteins have been identified by these codes in other tables. For catalase this corresponds to the PDB reference code. [b] Numbers in parentheses have been assigned to distinguish multiple metal sites within the same structure. All the cations are in the ferric or cupric oxidation state. [c] χ_1 = N–C_α–C_β–C_γ; χ_2 = C_α–C_β–C_γ–C_δ, where C_δ stands for either atom in Figure 1, depending on which gives a value of χ_2 in the range 0–180°. [d] The relative orientation of M and C_α with respect to the aromatic plane; + indicates that they are on the same side, and − specifies distal location. [e] The ligands belong to the β subunit.

hybridization at the phenolate oxygen.[31,32a] In general, protein ligands may have different geometrical requirements for binding metal ions and hydrogen-bond donors like water. These have two distinct regions of binding at the peptide carbonyl.[10] Even carboxylate[11a] and cysteine[9,32a] ligands show subtle differences in the distribution of bonded cations and protons. The binding patterns of these species are indistinguishable only in the case of histidine.[12]

The avoidance of the trigonal lone-pair direction by cations may be partly electrostatic. In both small-molecule and macromolecule structures, the spatial distribution of oxygen atoms in the aromatic ring environment demonstrates a significant tendency for the atoms to be located in the equatorial positions, where the electronegative oxygen atoms can favorably associate with ring hydrogen atoms carrying a partial positive charge.[33] By the same token, a cation bearing a positive charge, while approaching the phenolate oxygen along the sp^2 vector, will experience a considerable repulsion from the ring hydrogen atom at the proximal ortho position to push it closer to the C–O direction (i.e., $\phi < 60°$) or to force it out of the aromatic plane ($\theta < 90°$). A small magnitude of ϕ coupled with a value of θ close to 45° (in proteins) suggests a location between two lone-pair orbitals (one perpendicular to the aromatic ring and the other extending along the C–O bond) corresponding to sp hybridization at the oxygen atom—this may allow a lone-pair orbital to overlap with a metal d orbital of suitable symmetry.

An interesting feature of tyrosine ligation is its occurrence in pairs when binding centers other than the heme group; however, the two residues are widely separated along the sequence. Except in galactose oxidase, the two aromatic rings are very nearly perpendicular to each other (Table 2), but no clear pattern is discernible in their orientations.[34]

Hydrogen-Bond Interaction of the Tyrosinate Ligand. In addition to cation coordination, the tyrosinate anion is also stabilized by hydrogen bonding. The geometry of such an interaction, as given in Table 3, is such that the hydrogen-bond donor is not along the sp^2 lone pair (at $\theta = 90°$ and $\phi = 60°$), a direction that is preferred not only for a neutral tyrosine residue[31,32a] but for other oxygen ligands also.[32b] Steric reasons can account for this change in directionality. It is possible that the location of the cation (with other bound ligands) on one side of the ring imposes a spatial constraint on the possible approach of a hydrogen-bond donor along the trigonal orientation; such a group can only occupy a position on the other side of the aromatic ring (Figure 6). Indeed, if the cation is located in the (+x, +y, +z) octant (Figure 1), it is likely that the hydrogen-bonded group occupies the (+x, −y, −z) section (Table 3). In addition, the excess negative charge on the phenolate oxygen as compared to that on a phenolic oxygen may make the electrostatic interaction

Table 2. Relative Orientation of the Ligand Tyrosines within a Coordination Sphere

code	ligand pair	angle (deg) between planes	angle (deg) O–M–O	dist[a] (Å)
PCD	108, 147	85.2	95.7	5.46
LF	92, 192	93.5	109.1	6.27
	435, 528	88.9	102.7	6.28
GOX	272, 495	60.1	75.2	7.10

[a] Between the centroids of the two aromatic rings.

Table 3. Geometry of Hydrogen Bonding by Tyrosinates

code	ligand residue	hydrogen bonded to[a]	dist (Å)	angle (deg)[b] θ	angle (deg)[b] ϕ	disposition[c]
7CAT	357	NH2, R353	2.72	11.8	6.3	−
PCD	108	W	2.82	13.7	0.3	+
	147	W[d]	3.31	76.5	45.2	−
LF	92	W	2.50	55.2	47.9	−
	192	NH2, R210	3.11	33.7	63.1	−
	435	W	2.56	52.2	63.8	−
	528	W	2.86	42.9	67.7	−
GOX	272	O1, acetate	2.45	47.5	36.1	−
	495	W	3.28	36.7	60.7	+
		OH, Y405	2.87	82.5	71.3	−

[a] The name (PDB convention) of the atom donating a proton, followed by the one-letter amino acid code and the residue number, is given. W represents a water molecule. [b] The same convention of spherical polar angles, as used for defining the cation positions, has been utilized. [c] The relative orientation of the projections of the cation and the hydrogen-bond donor onto the xy plane with respect to the x axis (Figure 1). They are on the same side if the sign is + and opposite if −. An equivalent description can be provided in terms of the sign associated with the ϕ values of the cation and the hydrogen-bond partner bound to the tyrosinate oxygen: + and − indicate the same and the opposite signs, respectively, for the two ϕ values. [d] Could be an oxygen molecule.

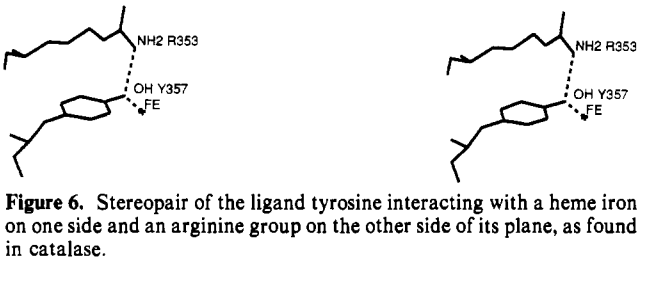

Figure 6. Stereopair of the ligand tyrosine interacting with a heme iron on one side and an arginine group on the other side of its plane, as found in catalase.

(31) Thanki, N.; Thornton, J. M.; Goodfellow, J. M. *J. Mol. Biol.* **1988**, *202*, 637–657.
(32) (a) Ippolito, J. A.; Alexander, R. S.; Christianson, D. W. *J. Mol. Biol.* **1990**, *215*, 457–471. (b) Murray-Rust, P.; Glusker, J. P. *J. Am. Chem. Soc.* **1984**, *106*, 1018–1025.
(33) (a) Thomas, K. A.; Smith, G. M.; Thomas, T. B.; Feldmann, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4843–4847. (b) Gould, R. O.; Gray, A. M.; Taylor, P.; Walkinshaw, M. D. *J. Am. Chem. Soc.* **1985**, *107*, 5921–5927.
(34) (a) Singh, J.; Thornton, J. M. *FEBS Lett.* **1985**, *191*, 1–6. (b) Burley, S. K.; Petsko, G. A. *Science* **1985**, *229*, 23–28.

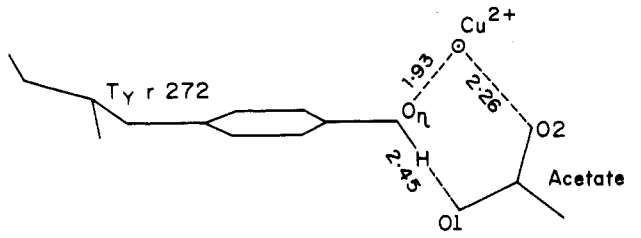

Figure 7. Illustration of the hydrogen-bonded cyclic structure involving the ligand tyrosine, acetate oxygens, and the cation in galactose oxidase. Relevant distances (Å) are indicated.

with a proton equally stable in orientations other than the lone-pair direction.

Galactose oxidase provides an interesting case, where there is a carboxylate oxygen within the hydrogen-bond distance of the phenolate oxygen of Tyr272 (Table 3). As the crystallization was set at pH 4.5,[24] it is likely that the phenolic hydroxyl is not deprotonated and the proton forms a hydrogen bond with an oxygen atom of the acetate anion, whose other oxygen coordinates to copper ion. This gives rise to a cyclic, planar arrangement (Figure 7), which is found in abundance in other systems involving carboxylate and hydroxyl ligands.[11] As discussed earlier, there are examples of phenolic ligands in small-molecule structures.

Secondary Structural Features. Consideration of the secondary structure in which a particular ligand is located shows that a metal-binding tyrosine can occur in a helix (catalase; residues 192 and 528 in lactoferrin) or β-sheet (residues 92 and 435 in lactoferrin; residue 147 in 3,4-PCD).[20,22-23] In fact, as same types of cations are bound, the propensity to be in a particular secondary structure should be similar for histidine and tyrosine, both of which use aromatic side chains for coordination. Nevertheless, the ligating atom in tyrosine is further from its main-chain atoms. As a result, a tyrosine ligand is under less severe steric constraint and indeed can bind a heme group, even when it is located in the middle of a helix 20 residues long in catalase, whereas ligand histidines occur at helix termini.[35]

Two ligands four residuals apart constitute an important cation-binding entity.[12] Even when a single histidine coordinates, its side chain at position i can form a hydrogen bond with the main-chain CO group at $i-4$, thereby stabilizing the helical structure.[35] Interestingly, catalase[20] also demonstrates that the metal coordination is facilitated by the interaction between two helical residues; in this structure there is a hydrogen bond between the tyrosine and an arginine four residues away (Figure 6). It appears that these two residues on the same side of the helix can be a particularly efficient motif for binding cations. Conversely, a polypeptide chain containing such a pair may be induced to adopt a helical conformation in the presence of a cation.

Side-Chain Conformation. The conformations of the ligand side chains as given by the two torsional angles are presented in Table 1. For tyrosine, χ_2 usually adopts a value very close to 90° in protein and peptide structures.[36] However, metal coordination perturbs the normal χ_2 distribution, the magnitude of the deviation being dependent on the relative positions of the cation and the C_α atom with respect to the aromatic plane. When they are on the same side, χ_2 can be 39(22)° (the average of five data points from Table 1) away from the expected value; however, when they are on opposite sides (four cases), the deviation is only 9(6)°. The variation in the torsion angle is a compromise to satisfy the stereoelectronic requirement of metal binding with respect to the filled orbitals at the tyrosinate oxygen. Results for cysteine and histidine ligands also suggest that the steric factors favoring some particular value of χ can be effectively overcome by metal coordination.[9,12] Incidentally, the greatest deviation of χ_2 is for residue 272 (galactose oxidase), which is involved in an unusual thioether bond with Cys228,[24] besides participating in a hydrogen-bonded unit (Figure 7).

Table 4. Basic Residues near Tyrosine Ligands in the Sequence[a]

protein	ligands	basic residues[b]
7CAT	357	R353
PCD	108, 147	R107, R109, H110, K111, R150
LF	92, 192	R89, H91
	435, 528	
GOX	272, 495	R270, R493, H496[c]

[a] Those within three residues of a ligand are enumerated (the only exception is 7CAT where R353 and Y357, although four residues apart, are close spatially because they are part of a helix. [b] One-letter amino acid code is used in front of the residue number. [c] Coordinates to the same metal center.

Preference for a Basic Amino Acid as a Neighboring Residue. Besides being within the hydrogen-bond distance of the ligand tyrosine in a few cases, one or more arginine or some other positively charged residues are close to the ligand in the primary structure (Table 4). It has been suggested that the function of a spatially close basic chain is to reduce the pK_a (10.0) of a tyrosine residue.[23,37] However, the presence of a positively charged residue in sequential proximity may also have a structural connotation. In the free state and at neutral pH, these ligand residues are not deprotonated. As such, tyrosines have a tendency to be found in protein interiors.[8a,38] The location near a charged residue with a long side chain may maintain the optimum solvent exposure of the uncharged ligand side chain in the folded state of the apo protein.

Summary and Conclusions

Earlier studies have shown that various ligands have distinct orientations for cation ligation.[9-12,39] Likewise, ions have non-random distribution around the phenolate oxygen. Steric and electrostatic elements make interaction along the sp^2 lone pair unfavorable; ions move away from such a direction toward the C–O vector (mostly in small-molecule structures) or away from the aromatic plane (proteins). Insofar as the sp^2 lone pair is averted, the stereochemistry of cation binding is different from the way a hydrogen-bond donor interacts with the tyrosine residue in proteins. The tyrosinate anion also makes hydrogen-bonded contact, not along the trigonal lone pair as shown by a neutral tyrosine, but on the side of the ring opposite to the metal ion. The geometrical pattern of binding shown by various ligand groups indicates the inadequacy of the isotropic force field parameters used to describe hydrogen bonding or metal coordination in proteins.[40] Any improvement along this direction should take into account not only the difference in the geometry of hydrogen bonding and metal ligation but also a possible change in the directionality of hydrogen-bonding interaction depending on the ionization state of the ligand. Moreover, metal coordination can alter the side-chain conformation of the tyrosine residue, as has also been found with cysteine and histidine ligands.

The role of arginine in binding anions is well recognized.[41] However, arginine and other positively charged residues may

(35) Chakrabarti, P. *Arch. Biochem. Biophys.* **1991**, *290*, 387–390.
(36) (a) Janin, J.; Wodak, S.; Levitt, M.; Maigret, B. *J. Mol. Biol.* **1978**, *125*, 357–386. (b) Cody, V. In *The Chemistry and Biochemistry of Amino Acids*; Barrett, G. C., Ed.; Chapman and Hall: London, 1985; pp 625–653. (c) Ponder, J. W.; Richards, F. M. *J. Mol. Biol.* **1987**, *193*, 775–791.
(37) Murthy, M. R. N.; Reid, T. J., III; Sicignano, A.; Tanaka, N.; Rossmann, M. G. *J. Mol. Biol.* **1981**, *152*, 465–499.
(38) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204*, 155–164.
(39) (a) Glusker, J. P. *Adv. Protein Chem.* **1991**, *42*, 1–76. (b) Einspahr, H.; Bugg, C. E. *Met. Ions Biol. Syst.* **1984**, *17*, 51–97. (c) Rosenfield, R. E., Jr.; Parthasarathy, R.; Dunitz, J. D. *J. Am. Chem. Soc.* **1977**, *99*, 4860–4862. (d) Chakrabarti, P.; Dunitz, J. D. *Helv. Chim. Acta* **1982**, *65*, 1482–1487.
(40) Vedani, A.; Huhta, D. W. *J. Am. Chem. Soc.* **1990**, *112*, 4759–4767.
(41) Riordan, J. F.; McElvany, K. D.; Borders, C. L., Jr. *Science* **1977**, *195*, 884–886.

also have an indirect role in maintaining the structural integrity of the metal-binding sites involving tyrosine ligands. Because of the hydrophilic nature of their long, flexible side chains, their sequential location near the ligands may keep the neutral ligands in the folded apoprotein at optimum distances from the protein surface for metal coordination.

Acknowledgment. We are grateful to Dr. D. C. Rees for encouragement and to Drs. E. N. Baker, N. Ito, and D. H. Ohlendorf for providing us with the atomic coordinates based on their work. We thank Mr. A. K. Gangopadhyay for assistance in the preparation of the manuscript.

Supplementary Material Available: Listings of inputs to programs used (CSD entries: acronyms and their journal references) and the GSTAT89 output (geometrical parameters and cation coordinates expressed in terms of a common axial system) (10 pages). Ordering information is given on any current masthead page.