X-ray Absorption Spectroscopy of Iron-Tyrosinate Proteins

A. L. Roe, D. J. Schneider, R. J. Mayer,‡ J. W. Pyrz,‡ J. Widom, and L. Que, Jr.*

Contribution from the Department of Chemistry, Baker Laboratory, Cornell University, Ithaca, New York 14853. *Received June 7, 1983*

Abstract: The iron K-edge absorption features of a series of 28 synthetic iron compounds and a total of 11 different complexes of ovotransferrin, catechol 1,2-dioxygenase, and protocatechuate 3,4-dioxygenase have been compared to determine the coordination number of the protein complexes. The intensity of the 1s to 3d preedge transition measured for the synthetic compounds varies inversely with coordination number. The normalized preedge peak areas averaged 8 units for six-coordinate compounds, 16 units for five-coordinate compounds, and 24 units for four-coordinate compounds. An extended-Hückel molecular orbital calculation showed a good correlation of the preedge peak areas with the total amount of iron 4p atomic orbitals mixed into the predominantly iron 3d molecular orbitals. This correlation is expected for a dipolar transition. The ovotransferrin complexes exhibited preedge peak areas consistent with the coordination numbers 6 and 7, while the dioxygenase complexes exhibited values that range from 8 to 13.4 units, indicating an active site that can be five- or six-coordinate.

The iron-tyrosinate proteins[1,2] are a diverse group consisting of the transferrins,[3-6] the catechol dioxygenases,[2,7-12] and the purple acid phosphatases.[13-15] These proteins are characterized by a distinctive visible absorption arising from tyrosinate-to-iron(III) charge transfer as demonstrated by resonance Raman spectroscopy. The presence of resonance-enhanced Raman vibrations characteristic of coordinated phenolate constitutes a criterion for classification into this subclass of ferric proteins. Although phenolate coordination is clearly demonstrated in these proteins, the overall environment of the iron remains to be elucidated.

X-ray absorption spectroscopy, in particular the extended X-ray absorption fine structure, has been very useful in the determination of the coordination chemistry of metal centers in a wide variety of materials.[16] In general, analysis of the EXAFS data provide the identity of the ligating atoms as well as the iron-ligand bond distances. An estimate of the coordination number can also be obtained. The uncertainty associated with this number is usually 20%, somewhat limiting its applicability.

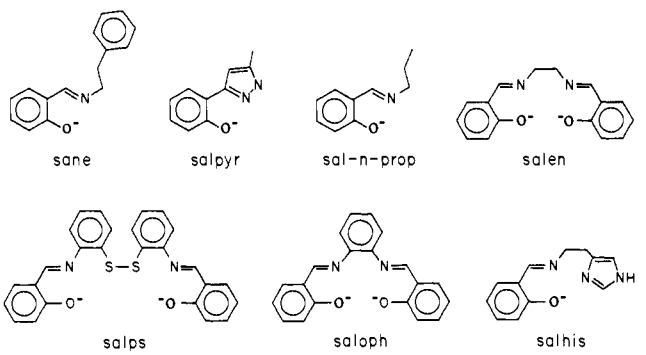
The X-ray absorption preedge and near-edge features are also informative, though there are fewer examples of their utility in biological systems.[17-21] A small preedge peak in the spectra of transition-metal complexes with incompletely filled d shells has been assigned to the 1s-3d transition.[17] This transition has different intensities for octahedral and tetrahedral complexes and has been used to infer site symmetry in metal complexes.[22]

In this paper we report the preedge features of an extensive list of Fe(III) complexes and apply this information to the interpretation of the X-ray absorption spectra of ovotransferrin, protocatechuate 3,4-dioxygenase (PCD), and catechol 1,2-dioxygenase (CTD) complexes. The synthetic complexes have been chosen to represent amino acid ligation and contain mainly oxygen and nitrogen ligands. These complexes represent a broad range of geometries whose exact nature has been determined by X-ray crystallography. Extended-Hückel calculations, using the atomic coordinates from the structure determinations, show a correlation between the observed preedge peak intensities and the amount of iron 4p character in the final states of the excited electron. We demonstrate that the preedge freatures provide insight into the coordination chemistry of the ferric centers in the iron-tyrosinate proteins.

Experimental Section

The model complexes were made following procedures in the references listed in Table I. Several of the multidentate ligands and their abbreviations are shown in Chart I. Abbreviations used for the other ligands are as follows: EDTA, ethylenediaminetetraacetate; dtc, dithiocarbamate; acac, acetylacetonate; $catH_2$, 1,2-dihydroxybenzene; H_4EHPG, *N,N'*-ethylenediaminebis(2-hydroxyphenylglycine); TPP, *meso*-tetraphenylporphyrin; HEDTA, (hydroxyethyl)ethylenediamine-triacetate. H_4EHPG was recrystallized from 95% ethanol, catechol was sublimed prior to use, and all other starting materials were reagent grade

Chart I



* To whom correspondence should be addressed at the Department of Chemistry, University of Minnesota, Minneapolis, MN 55455.

‡ NIH Predoctoral Trainees.

(1) Que, L., Jr. *Coord. Chem. Rev.* **1983**, *50*, 73-108.
(2) Keyes, W. E.; Loehr, T. M.; Taylor, M. L. *Biochem. Biophys. Res. Commun.* **1978**, *83*, 941-945.
(3) Carey, P. R.; Young, N. M. *Can. J. Biochem.* **1974**, *52*, 273-280.
(4) Gaber, B. P.; Miskowski, V.; Spiro, T. G. *J. Am. Chem. Soc.* **1974**, *96*, 6868-6873.
(5) Tomimatsu, Y.; Kint, S.; Scherer, J. R. *Biochemistry* **1976**, *15*, 4918-4924.
(6) Ainscough, E. W.; Brodie, A. M.; Plowman, J. E.; Bloor, S. J.; Loehr, J. S.; Loehr, T. M. *Biochemistry* **1980**, *19*, 4072-4079.
(7) Tatsuno, Y.; Saeki, Y.; Iwaki, M.; Yagi, T.; Nozaki, M.; Kitagawa, T.; Otsuka, S. *J. Am. Chem. Soc.* **1978**, *100*, 4614-4615.
(8) Que, L., Jr.; Heistand, R. H., II. *J. Am. Chem. Soc.* **1979**, *101*, 2219-2221.
(9) Que, L., Jr.; Heistand, R. H., II; Mayer, R.; Roe, A. L. *Biochemistry* **1980**, *19*, 2588-2593.
(10) Que, L., Jr.; Epstein, R. M. *Biochemistry* **1981**, *20*, 2545-2549.
(11) Felton, R. H.; Cheung, L. D.; Phillips, R. S.; May, S. W. *Biochem. Biophys. Res. Commun.* **1978**, *85*, 844-850.
(12) Bull, C.; Ballou, D. P.; Salmeen, I. *Biochem. Biophys. Res. Commun.* **1979**, *87*, 836-841.
(13) Gaber, B. P.; Sheridan, J. P.; Bazer, F. W.; Roberts, R. M. *J. Biol. Chem.* **1979**, *254*, 8340-8342.
(14) Antanaitis, B. C.; Strekas, T.; Aisen, P. *J. Biol. Chem.* **1982**, *257*, 3766-3770.
(15) Davis, J. C.; Averill, B. A. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4623-4627.
(16) Cramer, S. P.; Hodgson, K. O. *Prog. Inorg. Chem.* **1979**, *25*, 1-39.
(17) Shulman, R. G.; Yafet, Y.; Eisenberger, P.; Blumberg, W. E. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 1384-1388.
(18) Bunker, B.; Stern, E. A. *Biophys. J.* **1977**, *19*, 253-264.
(19) Hendrickson, W. A.; Co, M. S.; Smith, J. L.; Hodgson, K. O.; Klippenstein, G. L. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 6255-6259.
(20) Alagna, L.; Bianconi, A.; Desidari, A.; Tomlinson, A. A. G. *Ital. J. Biochem.* **1980**, *29*, 73-76.
(21) Elam, W. T.; Stern, E. A.; McCallum, J. D.; Sanders-Loehr, J. *J. Am. Chem. Soc.* **1982**, *104*, 6369-6373.
(22) Srivastava, U. C.; Nigam, H. L. *Coord. Chem. Rev.* **1972**, *9*, 275-310.
(23) Novozhilova, N. V.; Polynova, T. N.; Porai-Koshits, M. A. *J. Struct. Chem.* **1975**, *16*, 796-804.

0002-7863/84/1506-1676$01.50/0

Table I. X-ray Absorption Preedge Data for Synthetic Complexes

	first coord shell	1s-3d peak area [a] (normalized)	% Fe 4p (Hückel)	structure ref
7-coordinate				
$Na[Fe(EDTA)(H_2O)]$	2 N, 5 O	8.1	1.5	23
6-coordinate				
$Fe(diethyl\text{-}dtc)_3$	6 S	5.8	3.4	24
$Fe(pyrrolidine\text{-}dtc)_3$	6 S	6.8	1.6	25
$Fe(acac)_3$	6 O	6.9	0.2	26
$[Fe(salhis)_2]ClO_4$	4 N, 2 O	7.3	0.6	27
$K_3[Fe(cat)_3]$	6 O	7.7	0.4	28
$[Fe(salen)Cl]_2$	2 N, 3 O, Cl	7.9	3.5	29
$[Fe(salpyr)_2(CH_3OH)_2]NO_3$	2 N, 4 O	8.1	0.4	30
Fe(salen)acac	2 N, 4 O	8.4	2.2	31
Na[Fe(EHPG)]	2 N, 4 O	8.7	3.1	32
Fe(salps)Cl	2 N, 2 O, Cl, S	9.3	3.0	33
K[Fe(salen)cat]	2 N, 4 O	9.3	1.9	31
5-coordinate				
$Fe(pyrrolidine\text{-}dtc)_2Cl$	Cl, 4 S	12.4	13.2	34
Fe(saloph)SPh	2 N, 2 O, S	12.7	13.1	35
$Fe(sane)_2Cl$	2 N, 2 O, Cl	14.5	11.4	36
Fe(salen)Cl	2 N, 2 O, Cl	15.9	11.5	37
Fe(TPP)catH	4 N, O	16.4	12.9	38
$Fe(sal\text{-}n\text{-}prop)_2Cl$	2 N, 2 O, Cl	17.2	11.2	39
Fe(saloph)catH	2 N, 3 O	18.8	13.0	40
4-coordinte				
$(Et_4N)[Fe(SC_{10}H_{13})_4]$	4 S	23.1	21.6	41
$(Et_4N)FeBr_4$	4 Br	23.4	20.0	42
$(Et_4N)[Fe(OC_{10}H_{13})_4]$	4 O	24.1	21.8	43
$(Et_4N)FeCl_4$	4 Cl	25.0	20.6	44
oxo-bridged dimers				
$(C_2H_{10}N_2)[Fe(HEDTA)]_2O$	N, 5 O	13.2		45
$[Fe(TPP)]_2O$	4 N, O	23.9		46
$[Fe(salen)]_2O$	2 N, 3 O	24.1		47
$[Fe(saloph)]_2O$	2 N, 3 O	24.8		48
$(C_5H_6N)_2[FeCl_3]_2O$	O, 3 Cl	33.7		49

[a] Normalized area expressed in 10^{-2} eV units.

and were used without further purification. Poly(ethylene glycol) (MW 6000) was purchased from Sigma Chemical Co. Both tetraethylammonium tetrakis(2,3,5,6-tetramethylphenolato)ferrate(III) and tetraethylammonium tetrakis(2,3,5,6-tetramethylthiophenolato)ferrate(III) were generously provided by Drs. S. Koch and M. Millar.

CTD was obtained from *Pseudomonas arvilla* C-1 (ATCC 23974) by a modification of the published procedure.[50] After treatment with streptomycin sulfate, a poly(ethylene glycol) fractionation from 7.5% (w/v) to 20% was used in place of the ammonium sulfate fractionation. The pellet was redissolved in pH 8.5, 50 mM Tris-acetate buffer and applied to a 4.4 cm × 25 cm DEAE-Sephacel column. The column was washed with 1 L of 1% (w/v) ammonium sulfate, and the enzyme was eluted with 2 L of a linear gradient from 1% to 5% ammonium sulfate. Alternatively, the column can be washed with 0.1 M NaCl followed by a gradient from 0.1 to 0.3 M NaCl. Fractions with a specific activity greater than 15 were concentrated with an Amicon Diaflo apparatus and then applied to a 4.4 cm × 90 cm Sephacryl S-200 column.[51] Typical yields were 300 mg of CTD with a specific activity of 37 from 120 g of wet cells.

PCD from *Pseudomonas aeruginosa* (ATCC 23975) was prepared according to published procedures and had a specific activity of 78.[52,53] Ovotransferrin was purchased from Sigma Chemical Co. as type 1 conalbumin. A spectrophotometric assay for residual iron in the apoprotein indicated that less than 5% of the stoichiometric amount of iron was present and it was used without further purification.

The ovotransferrin complexes were made by using modifications of published procedures.[54] Carbonate buffer was used to make the carbonate complex. Carbon dioxide free buffer was made by boiling acidic buffer, degassing on a vacuum line, and adding solid Trizma-base under

(24) Leipoldt, J. G.; Coppens, P. *Inorg. Chem.* **1973**, *12*, 2269–2274.
(25) Sinn, E. *Inorg. Chem.* **1976**, *15*, 369–375.
(26) Roof, R. B., Jr. *Acta Crystallogr.* **1956**, *9*, 781–786.
(27) Averill, B. A., unpublished data.
(28) Raymond, K. N.; Isied, S. S.; Brown, L. D.; Fronczek, F. R.; Nibert, J. H. *J. Am. Chem. Soc.* **1976**, *98*, 1767–1774.
(29) Gerloch, M.; Mabbs, F. E. *J. Chem. Soc. A* **1967**, 1900–1908.
(30) Ainscough, E. W.; Brodie, A. M.; Plowman, J. E.; Brown, K. L.; Addison, A. W.; Gainsford, A. R. *Inorg. Chem.* **1980**, *19*, 3655–3663.
(31) Lauffer, R. B.; Heistand, R. H.; II; Que, L., Jr. *Inorg. Chem.* **1983**, *22*, 50–55.
(32) Bailey, N. A.; Cummins, D.; McKenzie, E. D.; Worthington, J. M. *Inorg. Chim. Acta* **1981**, *50*, 111–120.
(33) Bertrand, J. A.; Breece, J. L. *Inorg. Chim. Acta* **1974**, *8*, 267–272.
(34) Hoskins, B. F.; White, A. H. *J. Chem. Soc. A* **1970**, 1668–1674.
(35) Mayer, R. J.; Que, L., Jr., unpublished data.
(36) Magurany, C. J.; Strouse, C. E. *Inorg. Chem.* **1982**, *21*, 2348–2350.
(37) Gerloch, M.; Mabbs, F. E. *J. Chem. Soc. A* **1967**, 1598–1608.
(38) Rogers, M.; Que, L., Jr., unpublished data.
(39) Davies, J. E.; Gatehouse, B. M. *Acta Crystallogr., Sect. B* **1972**, *B28*, 3641–3645.
(40) Heistand, R. H., II; Roe, A. L.; Que, L., Jr. *Inorg. Chem.* **1982**, *21*, 676–681.
(41) Millar, M.; Lee, J. F.; Koch, S. A.; Fikar, R. *Inorg. Chem.* **1982**, *21*, 4105–4106.
(42) Sproul, G. D.; Stucky, G. D. *Inorg. Chem.* **1972**, *11*, 1647–1650.
(43) Koch, S. A.; Millar, M. *J. Am. Chem. Soc.* **1982**, *104*, 5255–5257.
(44) Cotton, F. A.; Murillo, C. A. *Inorg. Chem.* **1975**, *14*, 2467–2469.
(45) Lippard, S. J.; Schugar, H.; Walling, C. *Inorg. Chem.* **1967**, *6*, 1825–1831.
(46) Hoffman, A. B.; Collins, D. M.; Day, V. W.; Fleischer, E. B.; Srivastava, T. S.; Hoard, J. L. *J. Am. Chem. Soc.* **1972**, *94*, 3620–3626.
(47) Coggon, P.; McPhail, A. T.; Gross, P. M.; Mabbs, F. E.; McLachlan, V. N. *J. Chem. Soc. A* **1971**, 1014–1019.
(48) Lewis, J.; Mabbs, F. E.; Richards, A. *J. Chem. Soc. A* **1967**, 1014–1018.
(49) Drew, M. G. B.; McKee, V.; Nelson, S. M. *J. Chem. Soc., Dalton, Trans.* **1978**, 80–84.
(50) Fujiwara, M.; Golovleva, L. A.; Saeki, Y.; Nozaki, M.; Hayaishi, O. *J. Biol. Chem.* **1975**, *250*, 4848–4855.
(51) Sephacryl S-200 was modified by Pharmacia in 1979. Material obtained after that apparently removes iron from the protein due to the addition of thiol groups to the support material. Consequently, the specific activity of the protein is greatly decreased by use of the improved Sephacryl. Only the unmodified Sephacryl is effective and can be obtained upon request.
(52) Fujisawa, H.; Uyeda, M.; Kojima, Y.; Nozaki, M.; Hayaishi, O. *J. Biol. Chem.* **1972**, *247*, 4414–4421.
(53) May, S. W.; Phillips, R. S.; Oldham, C. D. *Biochemistry* **1978**, *17*, 1853–1860.
(54) Schlabach, M. R.; Bates, G. W. *J. Biol. Chem.* **1975**, *250*, 2182–2189.

Table II. X-ray Absorption Preedge Data for Protein Complexes

sample	1s-3d peak area (normalized)	activity retained, %
CTD Complexes		
phenol	10.3	>95
thiophenol	13.4	90[a]
4-methylcatechol	11.6	>95
benzoate	11.4	75
PCD Complexes		
native	9.0	70
p-hydroxybenzoate	11.0	95
terephthalate	10.4	>95
succinate	9.8	95
4-nitrocatechol	8.0	
transferrin[b]		
thioglycolate	4.9	
carbonate	7.7	

[a] Percent of original absorbance at 430 nm. [b] No change in iron coordination indicated by a comparison of initial and final scans on a particular sample.

carbon dioxide free nitrogen. All manipulations of the thioglycolate complex were done under carbon dioxide free nitrogen. The thioglycolate sample was frozen in liquid nitrogen during its transfer from a glovebag to the X-ray fluorescence detector.

The PCD and CTD inhibitor complexes were made by adding an aliquot of a concentrated inhibitor solution to the stock enzyme solution. The enzyme solutions contained 50 mM Hepes buffer at pH 8.0 and pH 7.5 for PCD and CTD, respectively. The pH of each concentrated inhibitor solution was adjusted to match the enzyme pH by the addition of sodium hydroxide. The final concentration of iron in all of the PCD inhibitor complexes was 3 mM and the concentrations of the inhibitors in the individual samples were 10 mM for 4-nitrocatechol and 50 mM for succinate, terephthalate, and *p*-hydroxybenzoate. For CTD samples the iron concentration was 4 mM and the inhibitor concentrations were 20 mM for phenol, 10 mM for thiophenol, and 100 mM for benzoate. The iron content of PCD and CTD was determined by spectrophotometric analysis with 2,4,6-tripyridyl-*s*-triazine (TPTZ). A protein sample was denatured in 1 N HCl and precipitated with 10% TCA. An excess of TPTZ was added to the supernatant, which was buffered with NH_4-OAc, and the iron concentration determined by the absorbance at 593 nm (extinction coefficient 22 600 M^{-1} cm^{-1}). The iron concentration was found to be 15 ± 5% in excess of protein concentration as determined by the absorbance at 280 nm.

The X-ray absorption data were obtained over several periods on the C-line stations at CHESS. All data were taken in the parasitic mode with the electron storage ring operating at an average current of 10 mA and at energies of 4.7, 5.0, and 5.4 GeV. All spectra were taken with a beam width of 1.5 cm and a resolution of 1–2 eV. A control experiment using beam heights ranging from 0.3 to 3.0 mm showed that the preedge peak area varied less than 10%. The energy range of the spectra was from ~50 eV below the edge to 375 eV above the edge, with a 0.5-eV step size near the edge and a 5-eV step size above the edge.

Transmission and fluorescence modes were used to determine the edge structure of the models and proteins, respectively. The absorption coefficients for the model complexes were determined by using gas ionization chambers as detectors and finely ground samples with an average absorption coefficient change of 0.8 at the iron K edge. The protein edge data were taken with a MnO_2 filter, Soller slits, and a large solid-angle collection ionization chamber filled with argon, generously provided by Dr. J. Wong.[55] The proteins were held in a slot cut in a copper plate with 0.25-mil of polypropylene for the front window and 2-mil Kapton (3M Co.) for the back window. The copper provided cooling to −10 °C by thermal contact to a Dewar containing a dry ice–acetone slush. The area around the sample was purged with helium to reduce atmospheric absorption of the X-rays and to prevent CO_2 diffusion into the ovotransferrin–thioglycolate sample.

The data represent both single scans and sums of scans. The model complexes produce data with good signal-to-noise ratios; consequently, the average counting time per data point is ~3 s. In contrast, the average counting times for the proteins range from 30 s to 2 min. The total exposure times were 2 h for the native PCD and CTD–benzoate complex, 4 h for all other dioxygenase complexes, and 8 h for the more stable ovotransferrin complexes. The proteins were checked for radiation damage by assaying for enzymatic activity, or, in the cases of CTD–thiophenol and the ovotransferrin complexes, by comparing the UV-vis spectra before and after irradiation. The results are listed in Table II. The EPR spectrum of the ovotransferrin–thioglycolate sample after irradiation was identical with the one taken before irradiation.

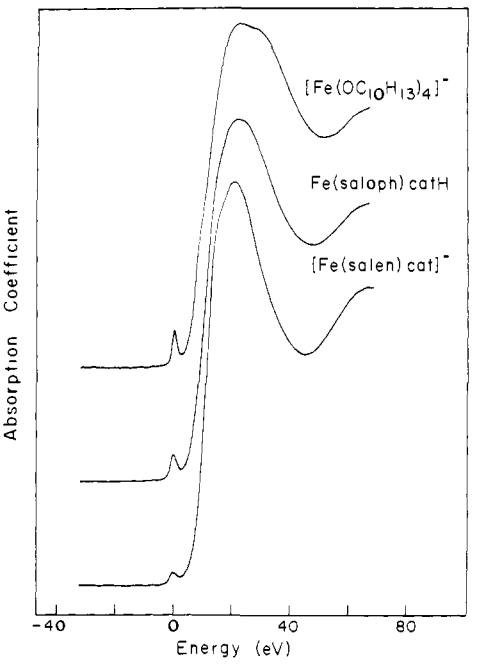


Figure 1. Iron K-edge absorption spectra of model complexes: top curve, four-coordinate [Fe(2,3,5,6-tetramethylphenolate)$_4$]$^-$; middle curve, five-coordinate Fe(saloph)catH; bottom curve, six-coordinate [Fe(salen)cat]$^-$.

The preedge peak areas were calculated by subtracting an arctangent function from the data and normalizing with respect to the edge jump height. The background function was determined by a least-squares fit of an arctangent and a first-order polynomial to the data below the inflection point of the edge: the fitting regions were from −30 to −5 eV and from +5 eV to the inflection point; 0 eV was defined as the preedge peak position. The area of the preedge peak after the background subtraction was obtained by integrating over a range of 10 eV. This range centered on the peak and any residual background function was interpolated over that range. The edge jump was determined by fitting first-order polynomials to the data from −30 to −5 eV and from +150 to +325 eV. The difference between these two lines at the inflection point of the edge was used as the normalization factor for the preedge peak area. For example, Na[Fe(EDTA)(H_2O)] had a normalized preedge peak area of 8.1 × 10^{-2} eV, which is abbreviated as 8.1 units.

The extended-Hückel calculations were done with the standard parameters for H, C, N, O, and S.[56-58] The parameters for chlorine and iron as the singly charged cation were taken from published calculations.[59-61]

Results and Discussion

Model Complexes. The preedge features in the K-edge spectra of a series of high-spin ferric complexes have been studied as models for non-heme iron proteins. These compounds all exhibit a sharp absorption band about 10 eV below the midpoint of the absorption edge, which is assigned to the 1s–3d transition.[17] Three representative curves in Figure 1 show that the amplitude of this band is dependent on coordination number. The lowest curve has the smallest amplitude and results from a six-coordinate complex. The middle curve has a larger amplitude and represents a five-coordinate complex. The largest amplitude is observed for a

(55) Stern, E. A.; Heald, S. M. *Rev. Sci. Instrum.* **1979**, *50*, 1579–1582.

(56) Hoffmann, R. *J. Chem. Phys.* **1963**, *39*, 1397–1412.
(57) Hoffmann, R.; Lipscomb, W. N. *J. Chem. Phys.* **1962**, *36*, 2179–2195.
(58) Hoffmann, R.; Lipscomb, W. N. *J. Chem. Phys.* **1962**, *36*, 3489–3493.
(59) Summerville, R. H.; Hoffmann, R. *J. Am. Chem. Soc.* **1976**, *98*, 7240–7254.
(60) Dedieu, A.; Albright, T. A.; Hoffmann, R. *J. Am. Chem. Soc.* **1979**, *101*, 3141–3151.
(61) Goddard, R. J.; Hoffmann, R.; Jemmis, E. D. *J. Am. Chem. Soc.* **1980**, *102*, 7667–7676.

four-coordinate complex. High-resolution spectra of some complexes show asymmetry in the 1s–3d peak. For Fe(salen)Cl, a five-coordinate complex, the peak is asymmetric with a shoulder on the high-energy side, which results from transitions to 3d orbitals of slightly different energies. In general, the 1s–3d peak line width decreases with a reduction of the coordination number. The decrease in line width in going from six-coordination to four-coordination is consistent with the 1s–3d assignment, since the line width should reflect the ligand field splitting of the 3d orbitals in the various geometries.

The amplitudes for the preedge peaks of synthetic ferric complexes having different coordination numbers are given in Table I. These amplitudes are expressed in terms of peak areas that have been normalized by dividing by the edge jump height. The largest single group of complexes are six-coordinate, with areas ranging from 5.8 to 9.3 units. The tris(chelate) complexes, which have higher symmetry than the other complexes studied, exhibit peak areas at the lower end of this range. As the complexes become more distorted, the peak areas tend to increase. For example, Fe(salps)Cl is at first glance an unexceptional six-coordinate complex.[33] Upon closer scrutiny, we note that, due to the weak Fe–S bond (2.54 Å), the complex is distorted toward a square pyramid. Consequently, the peak area for Fe(salps)Cl falls on the high end of the six-coordinate range.

The remaining models represent the other common coordination geometries for iron. The sole seven-coordinate complex has a peak area within the range of the six-coordinate complexes, indistinguishable from those of six-coordinate complexes. In contrast, the five-coordinate complexes are quite distinct from the six-coordinate complexes. Their range of peak areas, 12.4–18.8, is clearly separate from all other ranges. The four-coordinate complexes have the smallest range of peak areas, 23.1–25.0, reflecting the fact that all of the complexes are very close to tetrahedral. For the complexes studied, there are three separate groups of peak areas: the six- and seven-coordinate complexes exhibit areas from 5 to 10, the five-coordinate complexes from 12 to 18, and the four-coordinate complexes from 23 to 25.

In addition to the above complexes, oxo-bridged dimers have been investigated and found to have 1s–3d peak areas that are larger than those of their corresponding monomers. The six-coordinate dimer has a peak area that falls in the range of the five-coordinate monomers, and the five-coordinate dimers have peak areas in the range of four-coordinate monomers. There is a roughly constant increase in peak area (~7 units) in going from monomer to dimer for constant coordination number. This increase in intensity is very similar to the behavior of d–d transition intensities in oxo-bridged dimers and is ascribed to the presence of substantial ($-J > 80$ cm^{-1}) antiferromagnetic coupling.[62,63] Dimers with weaker coupling exhibit peak areas consistent with the monomers of the same coordination number. For example, $[Fe(salen)Cl]_2$ exhibits a $-J$ value of 7.5 cm^{-1} [64] and a preedge peak area consistent with six-coordination.

The complexes containing sulfur all have the lowest peak intensities of their particular geometries. This may be a result of poor background fitting for these complexes caused by other features in the absorption edge. For all of the complexes there is a small, unresolvable band centered around 5 eV above the 1s–3d band. When this band is large it can give rise to an edge with two inflection points, as in the case of $[Fe(SC_{10}H_{13})_4]^-$. In the sulfur-containing complexes this band increases in intensity, and the fitting routine places the background higher than the actual K-absorption edge because the base line is fitted to the +5-eV peak in addition to the underlying edge. This problem is not serious for the four-coordinate complexes since the peak height is much greater than the line width. The error introduced in peak areas is more significant for five- and six-coordinate complexes, because of the smaller peak heights.

Scheme I

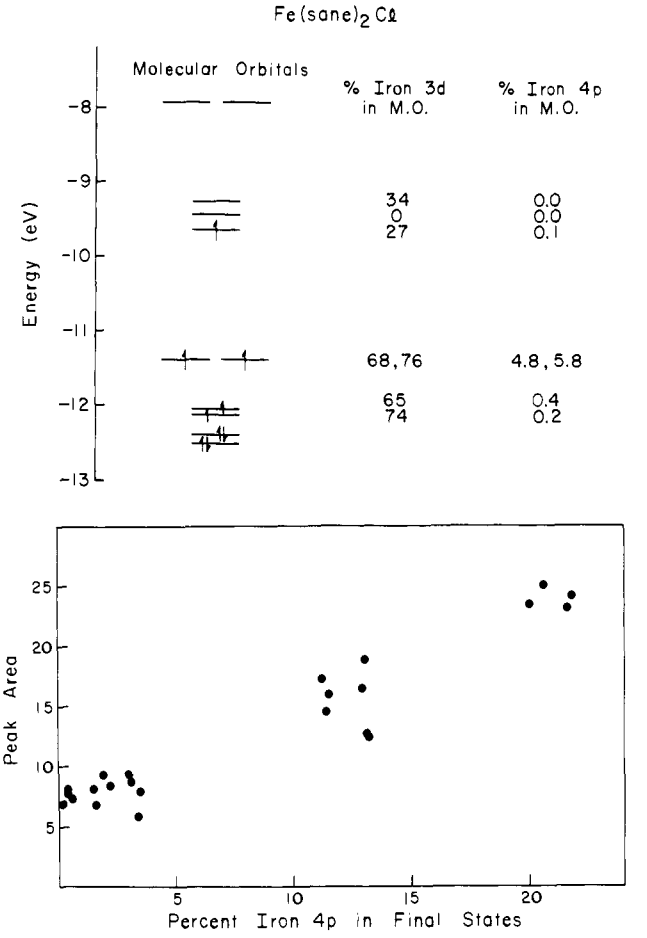


Figure 2. Correlation of 1s–3d preedge peak area with the total percentage of iron 4p atomic orbitals in the molecular orbitals that are final states for the excited electron.

Molecular Orbital Calculations. Extended-Hückel molecular orbital calculations were done to gain a qualitative understanding of the factors affecting the 1s–3d peak area. The electron occupation of the iron 3d molecular orbitals was constrained to be consistent with the high-spin ferric nature of the complexes. In the case of Fe(pyrrolidine-dtc)$_2$Cl the electron occupation was made consistent with three unpaired electrons since the magnetic moment of this complex shows it to be in an intermediate spin state.[65] The observed intensities for the 1s–3d peaks were assumed to arise principally from dipolar transitions from the iron 1s orbital to the iron 4p components of the molecular orbitals which have appropriate energies to be final states of the excited electron. The square of the transition moment between the iron 1s and the final state was assumed to be proportional to the iron 4p character in the final state. This depends on the transition moments between the iron 1s orbital and all ligand orbitals being considerably smaller than those for the iron 1s and 4p orbitals, because of the exponential decay of the atomic orbitals on the two separated atoms.

Scheme I shows the results of an extended-Hückel calculation on a representative model complex, Fe(sane)$_2$Cl. As expected, the highest occupied molecular orbitals consist mainly of iron 3d atomic orbitals and contain varying amounts of iron 4p atomic orbitals. To relate the iron 4p character in each molecular orbital to the area of the 1s–3d peak, it is necessary to sum the contributions from all the final states of the excited electron. For Fe(sane)$_2$Cl all available molecular orbitals between −9 and −12.2 eV were assumed to be potential final states, since monochromator resolution and lifetime broadening effects make the individual components unresolvable. The sum is expressed as the total of

(62) Gray, H. B. *Adv. Chem. Ser.* **1971**, *No. 100*, 365–389.
(63) Murray, K. S. *Coord. Chem. Rev.* **1974**, *12*, 1–35.
(64) Gerloch, M.; Lewis, J.; Mabbs, F. E.; Richards, A. *J. Chem. Soc. A* **1968**, 112–116.
(65) Martin, R. L.; White, A. H. *Inorg. Chem.* **1967**, *6*, 712–717.

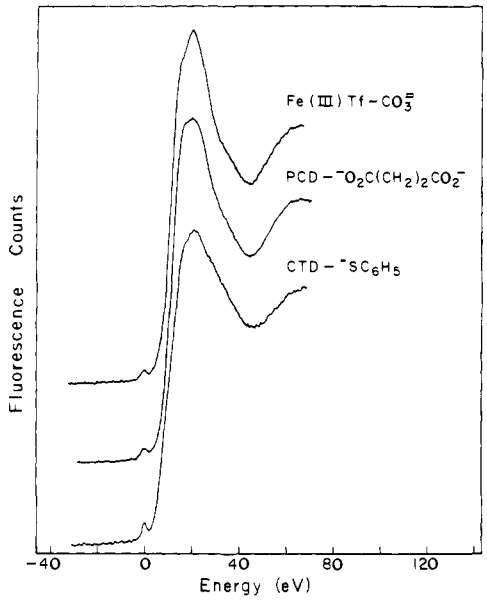


Figure 3. Iron K-edge absorption spectra of three iron proteins: top curve, iron–ovotransferrin–carbonate complex; middle curve, protocatechuate 3,4-dioxygenase–succinate complex; bottom curve, catechol 1,2-dioxygenase–thiophenolate complex.

all the individual percentages of iron 4p atomic orbitals in each molecular orbital; in this case, the total is 11.4 units. It is important to note that only a few of the molecular orbitals have substantial iron 4p character, and all of the these are close in energy.

The results of the calculations for all of the synthetic compounds are shown in Table I. Figure 2 shows that there is a good correlation between the total percentage of iron 4p and the observed peak areas and indicates that the peak intensity can be predicted rather accurately given a particular coordination geometry. On the basis of the assumption stated earlier, the percentage of iron 4p atomic orbitals in the final states should be directly proportional to the peak area. However, the best line fit does not pass through zero area at zero 4p content. There are three possible reasons for this behavior: vibronically allowed dipolar transitions, quadrupolar transitions, and dipolar transitions to the ligand atomic orbital components of the predominantly iron 3d molecular orbitals. A study of a centrosymmetric complex, $(C_4H_8N_3O)_2[CuCl_4]$, has shown that both vibronically allowed transitions and quadrupolar transitions can affect the intensity of the preedge peak.[66] Finally, the Hückel calculation cannot include antiferromagnetic interactions, and the correlation between 4p character and coordination number breaks down for oxo-bridged dimers.

Protein Complexes. Several complexes of each of the non-heme iron proteins were studied and three representative spectra are shown in Figure 3. All of the spectra have the same general shape, with a large peak around +20 eV and a relatively small 1s–3d band. The 1s–3d band areas for all of the complexes are given in Table II. The small 1s–3d areas observed for the protein complexes are distinctly different from the large peak areas observed for the four-coordinate models as well as the four- and five-coordinate oxo-bridged dimers. Previous studies have ruled out interacting iron atoms in these proteins,[67,68] and the magnitude of the 1s–3d band is in agreement with this. Therefore only five-, six-, or seven-coordinate isolated sites need be considered.

The 1st–3d peak areas for the ovotransferrin complexes of carbonate and thioglycolate are both very small. Ovotransferrin has been shown to have two slightly different iron binding sites,[68] and the peak areas show that neither of the sites has a coordination number lower than 6. This finding agrees with the work on human serotransferrin, which was also found to be six- or seven-coordinate.[20] Furthermore, the low value for the thioglycolate complex is consistent with sulfur binding to iron, since all of the sulfur-containing models exhibit the lowest 1s-3d areas of their particular geometries.

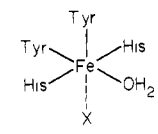
The interpretation of the dioxygenase data is not as straightforward. The preedge features of native PCD and its 4-nitrocatechol complex indicate a six-coordinate environment, the native enzyme having a somewhat more distorted site. The CTD–thiophenol complex, on the other hand, exhibits a peak area consistent with a five-coordinate environment; indeed, it matches well the features observed for Fe(saloph)SPh. The remaining dioxygenase complexes, however, exhibit peak areas higher than those for six-coordinate complexes but lower than those for five-coordinate ones.[69]

These intermediate areas can be interpreted as arising from a mixture of five- and six-coordinate sites or from a six-coordinate site distorted toward five-coordinate. A mixture can only result from an equilibrium between different coordination numbers at a single protein site since Mössbauer spectra of several dioxygenase complexes show a single iron environment at 4 K.[70] The second interpretation suggests a six-coordinate active site more distorted than any of the six-coordinate synthetic complexes.

One possible distortion was examined by using the correlation of peak areas and Hückel calculations. Starting with a hypothetical octahedral complex, $[Fe(OH)_6]^{3-}$, one of the Fe–OH bonds was lengthened and the remaining ligands were kept at the original bond length. The equatorial oxygens were rotated to follow the departing oxygen in order to minimize the point charge repulsion in the complex; this eventually produces a square pyramid. This calculation showed that the mixing of iron 4p orbital character into the d_{z^2} molecular orbital increased fairly linearly to 80% of its maximum value over a 0.7-Å increase in the unique bond length. The mixing of iron 4p orbital character into the E set of 3d orbitals was not significant until 0.5 Å of bond lengthening had occurred. It can be estimated that an increase of 0.2 Å in one of the iron–ligand bonds in an octahedron is the minimum required to produce a recognizable increase in the 1s–3d peak area by this mechanism.

Both interpretations of the X-ray absorption preedge features of the catechol dioxygenase complexes imply an active site with a weakly coordinated, or easily dissociable ligand. The active site would then have the flexibility to be five- or six-coordinate, depending on the circumstances. This would be consistent with its catalytic function. An easily dissociable ligand would provide a readily available site during the catalytic cycle to bind intermediate species. In the proposed mechanism of Que et al., such a site could serve to coordinate the postulated superoxide or peroxide formed during the reaction.[71]

Our present model for the dioxygenase active site is depicted below. Water has been demonstrated as a ligand with EPR



experiments using $H_2{}^{17}O$.[72] Ligation of histidine is suggested

(66) Hahn, J. E.; Scott, R. A.; Hodgson, K. O.; Doniach, S.; Desjardin, S. R.; Solomon, E. I. *Chem. Phys. Lett.* **1982**, *88*, 595–598.
(67) Que, L., Jr.; Lipscomb, J. D.; Zimmermann, R.; Munck, E.; Orme-Johnson, N. R.; Orme-Johnson, W. H. *Biochim. Biophys. Acta* **1976**, *452*, 320–334.
(68) Aisen, P.; Listowsky, I. *Annu. Rev. Biochem.* **1980**, *49*, 357–393.
(69) Uncertainties to the interpretation introduced by the presence of adventitious iron are minimal, assuming that this iron is in a six-coordinate environment. Taking the 15% adventitious iron into account would correct the observed protein values by 3–5%. If the advantitious iron were in a tetrahedral environment, a possibility we consider highly unlikely, the active-site contribution to the preedge peak area would be significantly smaller than the observed values.
(70) Munck, E.; Kent, T.; Que, L., Jr.; Widom, J., unpublished results.
(71) Que, L., Jr.; Lipscomb, J. D.; Munck, E.; Wood, J. M. *Biochim. Biophys. Acta* **1977**, *485*, 60–74.
(72) Lipscomb, J. D.; Whitakker, J. W.; Arciero, D. M. In "Oxygenases and Oxygen Metabolism"; Nozaki, M., Yamamoto, S., Ishimura, Y., Coon, M. J., Ernster, L., Estabrook, R. W., Eds.; Academic Press: New York, 1982; pp 27–38.

from EXAFS data,[73] and that of tyrosine from resonance Raman spectra.[2,7-12] The absence of thiolate ligation is deduced from Mössbauer, Raman, and EXAFS data,[67,73] while the presence of X, identity not presently known, is suggested by our X-ray absorption preedge studies.

The ligand arrangement proposed is clearly one of several permutations possible. X is the weakly coordinated or easily dissociable ligand postulated earlier to account for the intermediate preedge values observed for many of the dioxygenase complexes. The proposed cis arrangement of tyrosines with one trans to X allows the complex to distort toward a square pyramid, resulting in a structure with an apical and a basal tyrosine. The two tyrosines would thus exhibit charge-transfer bands of different energy, in agreement with the Raman excitation profiles of several dioxygenase complexes.[9,10,74] In this model, the CTD–thiophenol complex would be five-coordinate, if two ligands (presumably X and H_2O) were to dissociate when thiophenol binds. Similarly, the PCD–4-nitrocatechol complex would be six-coordinate, if two ligands were displaced by the chelating catechol.[75]

In conclusion, the X-ray absorption preedge studies we have reported, involving a broad sampling of coordination environments, clearly delineate their utility and some of their limitations in structural investigations of iron complexes. These studies are easily done along with EXAFS experiments and can provide valuable information on coordination number, independent of and complementary to that from EXAFS data. Indeed, the preedge information has aided in the interpretation of the EXAFS for iron–transferrin–carbonate and thioglycolate complexes.[76]

Acknowledgment. This work was supported by the National Institutes of Health (Grant GM-25422) and the National Science Foundation (Grant PCM-8213865). Synchrotron radiation time was provided by the Cornell High Energy Synchrotron Source, supported by the National Science Foundation (Grant DMR-8112822). We thank Drs. S. Koch and M. Millar for the generous gift of the four-coordinate iron phenolate and thiophenolate complexes and Dr. J. Wong for the generous loan of the EXAFS detector. We appreciate the invaluable assistance of the CHESS staff and Dr. J. Wong during the data collection. L.Q. is an Alfred P. Sloan Foundation fellow (1982–1984) and the recipient of an NIH Research Career Development Award (1982–1987).

Registry No. Na[Fe(EDTA)(H_2O)], 21626-24-4; Fe(diethyl-dtc)$_3$, 13963-59-2; Fe(pyrrolidine-dtc)$_3$, 21288-86-8; Fe(acac)$_3$, 14024-18-1; [Fe(salhis)$_2$]ClO_4, 88357-96-4; K_3[Fe(cat)$_3$], 59388-67-9; [Fe(salen)Cl]$_2$, 15308-73-3; [Fe(salpyr)$_2$(CH_3OH)$_2$]NO_3, 74877-75-1; Fe(salen)acac, 88357-97-5; Na[Fe(EHPG)], 16455-61-1; Fe(salps)Cl, 53317-28-5; K[Fe(salen)cat], 81293-72-3; Fe(pyrrolidine-dtc)$_2$Cl, 28010-69-7; Fe(saloph)SPh, 88357-98-6; Fe(sane)$_2$Cl, 42294-92-8; Fe(salen)Cl, 38586-93-5; Fe(TPP)catH, 88357-99-7; Fe(sal-n-prop)$_2$Cl, 40214-84-4; Fe(saloph)catH, 80041-63-0; (Et_4N)[Fe($SC_{10}H_{13}$)$_4$], 82741-95-5; (Et_4N)$FeBr_4$, 21279-19-6; (Et_4N)[Fe($OC_{10}H_{13}$)$_4$], 82742-05-0; (Et_4N)-$FeCl_4$, 14240-75-6; ($C_2H_{10}N_2$)[Fe(HEDTA)]$_2$O, 26998-37-8; [Fe(TPP)]$_2$O, 12582-61-5; [Fe(salen)]$_2$O, 18601-34-8; [Fe(saloph)]$_2$O, 18601-73-5; (C_5H_6N)$_2$[$FeCl_3$]$_2$O, 66512-55-8; PCD, 9029-47-4; Fe, 7439-89-6.

(73) Felton, R. H.; Barrow, W. L.; May, S. W.; Sowell, A. L.; Goel, S.; Bunker, G.; Stern, E. A. *J. Am. Chem. Soc.* **1982**, *104*, 6132–6134.
(74) Pyrz, J. W.; Roe, A. L.; Que, L., Jr., manuscript in preparation.
(75) Lauffer, R. B.; Que, L., Jr. *J. Am. Chem. Soc.* **1982**, *104*, 7324–7325.
(76) Schneider, D. J.; Roe, A. L.; Mayer, R. J.; Que, L., Jr. *J. Biol. Chem.*, in press.