Pure & Appl. Chem., Vol. 69, No. 10, pp. 2205–2212, 1997.
Printed in Great Britain.

X-ray absorption spectroscopic studies of metal centres in biology and developments of relevant synthetic analogues

C. David Garner, Pippa Baugh, David Collison, E. Stephen Davies, Andrew Dinsmore, John A. Joule, Elna Pidcock and Clare Wilson.

The University of Manchester, Department of Chemistry, Manchester, M13 9PL, U.K.

Abstract: X-ray absorption spectroscopy (XAS) has been used extensively to probe the local environment of transition metal centres in biological systems. The synergy between such XAS studies and co-ordination chemistry is well developed. Examples where XAS studies of metal centres in biological systems have stimulated developments in co-ordination chemistry include Amavadin and representatives of the oxymolybdoenzymes. Bond Valence Sum Analysis, using structural information available for co-ordination complexes, is a useful tool to inspect the structure obtained from XAS and/or protein crystallography for a d-transition metal site in a biological system.

INTRODUCTION

"Inorganic" elements are essential for the normal development and health of all living systems. Although the majority of these elements are usually only present in trace amounts, this does not diminish their significance. A considerable number of these 'trace elements' have been shown to play key rôles in biological processes, including the d-transition metals:- vanadium, manganese, iron, cobalt, nickel, copper, zinc, molybdenum and tungsten (ref. 1).

Crystallography is a most powerful tool capable of revealing the overall molecular architecture of proteins, together with the location of an individual d-transition metal atom (or atoms) or a cluster of such atoms. Developments in the techniques of protein crystallisation, the provision of new and more powerful X-ray sources (including the availability of synchrotron radiation as an intense and tunable source), improved detectors, and tremendous gains in computational power for data collection and analysis, have significantly advanced the applicability and precision of protein crystallography. Thus, we are moving to a situation where for a metalloprotein we expect not only a full structural characterisation, but also that this will be achieved for several derivatives - *e.g.* oxidised, reduced, inhibitor- or substrate analogue-bound, and site- or loop-directed mutants. Furthermore, NMR structure determination of protein structures in solution is advancing rapidly, and for metalloproteins the presence of one (or more) paramagnetic centres is advantageous for the definition of the local environment of the metal centre(s) (ref. 2).

Despite the beauty and significance of the information provided by protein crystallography and NMR spectroscopy, the structural data provided should be taken to represent the *beginning* of a systematic investigation of the rôle of the d-transition metal (or metals) in the biological function of the enzyme, not the *end* of such a study. An important aspect of the systematic investigation is to define the co-ordination of the metal atom(s) and how

this varies in different states of a redox active protein and/or metalloenzyme. Thus, *the chemistry of transition metals is the chemistry of their complexes* and we need to know the immediate and overall environment of metals in a protein to understand the properties of these centres. In respect of the structure of the metal site, it is important to note that, at present, protein crystallography will not define metal-ligand distances to a precision of better than ±0.1Å and often the resolution achieved is significantly inferior to this. Uncertainties of this magnitude are chemically significant. Therefore, there is a clear case for using metal-edge X-ray Absorption Spectroscopy (XAS) to probe metal centres in proteins (ref. 3), especially as the associated Extended X-ray Absorption Fine Structure (EXAFS) can provide metal-ligand bond lengths to a precision of ±0.03Å. The combination of protein crystallography and XAS is particularly powerful, especially when the same state of the system is studied by each. XAS is a versatile technique with no requirement for crystallinity and, thus, can be used independent of protein crystallography.

X-RAY ABSORPTION SPECTROSCOPY

Introduction

Since the first experiment on a synchrotron radiation source (ref. 4), X-ray absorption spectroscopy (XAS) has become established as a technique for investigating the local environment of metal atoms in proteins. XAS has many advantages as a probe of metal centres in biological systems and numerous studies have been reported in the literature (ref. 3). The X-ray absorption spectrum of Cu_5Zn_3-Metallothionein (Fig. 1) shows the element specificity of the technique and regions within the profile of an absorption edge.

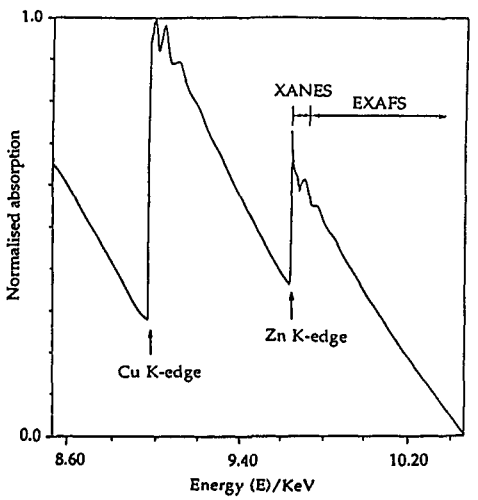


Fig. 1 X-ray absorption spectrum of Cu_5Zn_3-Mellatothionein
(XANES = X-ray Absorption Near Edge Structure;
EXAFS = Extended X-ray Absorption Fine Structure).

Beyond the absence of a requirement for crystalline material, major attractions encouraging the application of XAS include the specificity and sensitivity of the technique, so that information can be obtained for a particular metal centre presented at (say) 1 atom in 10^6 in a large and complicated protein. Also, XAS allows several different states of a system to be probed with relative ease, thereby providing a profile of how a metal centre responds, *e.g.*, to the oxidation level of an enzyme, the presence of substrates, or substrate analogues of inhibitors.

Information Content of XAS

The actual energy of a particular X-ray absorption edge of an element depends upon the oxidation state and the nature of the immediate chemical environment of that element. Typically, one unit increase of oxidation state increases the 3d or the 4d element edge

position by 1-3 eV, *e.g.* the shift in the position of the iron *K*-edge and its associated structure of *ca.* 1.5 eV to higher energy upon oxidation of cytochrome *c* (ref. 5). The sense and magnitude of this shift are consistent with a redox process which is concentrated at the iron centre. Excitation of a core electron into the continuum may be convoluted with transitions to the valence levels. These promotions give rise to pre-edge and edge features which can provide information concerning the chemical nature and electronic structure of the primary absorber (Fig. 2).

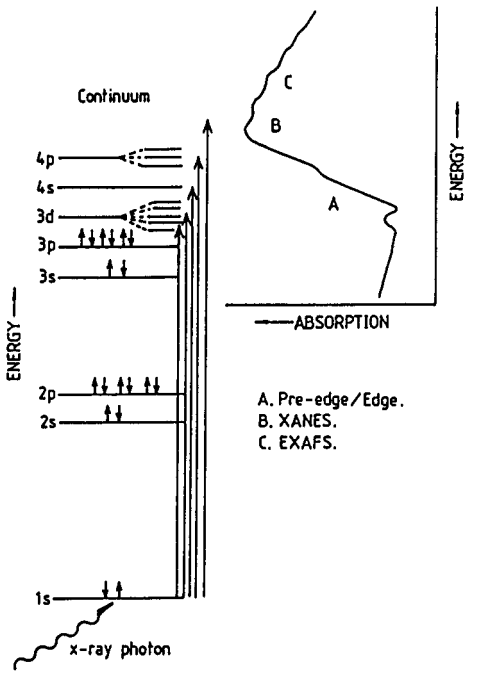


Fig. 2 Schematic representation of the origin of the X-ray absorption edge for promotion of a K-shell electron of a 3d element (*e.g.* V(V),d°).

Beyond the X-ray absorption edge, oscillations in the amplitude of the absorption can extend for up to 1,000 eV. Historically, and more recently because of the different theoretical treatment necessary to interpret the data, it has been customary to classify the oscillations within *ca.* 50 eV of the edge as the X-ray Absorption Near Edge Structure (XANES) and those which extend beyond this region as the Extended X-ray Absorption Fine Structure (EXAFS). The theoretical basis of the latter is considered mature and interpretations of EXAFS data have been reported confidently for *ca.* 20 years. In contrast, progress in the understanding and, therefore, the application of XANES has been relatively slow and data in this spectral region are generally not interpreted but used qualitatively to "fingerprint" a metal site. EXAFS provides structural information concerning the distance, nature, number and coherence of the shells of atoms around the primary absorber. However, confidence in the interpretation of the EXAFS of a metalloprotein centre is increased significantly by the successful completion of the corresponding exercise for a related and structurally characterised chemical analogue.

The primary sources of uncertainty in the structural parameters determined from EXAFS analyses arise from the correlation between the co-ordination number (N) and the Debye-Waller factor (σ^2) of individual shells of backscatterers. The usual accuracy which can be expected for the primary co-ordination shell of a metal centre in a protein is: $R \pm 0.03$Å, $N \pm 25\%$ and $Z \pm 1$. The present limitations in defining N and Z are especially frustrating and emphasise the need to integrate the information available from other spectroscopic and structural techniques into EXAFS analyses. Amino acid sequence data and/or site-directed mutagenesis studies can provide clear indications of the nature of the groups binding metals in metalloproteins. Furthermore, as illustrated by studies on rubredoxin (ref. 6) and Cu,Zn-superoxide dismustase (ref. 7), XAS and protein crystallography are especially complementary techniques for characterising metal centres in proteins.

When contemplating the use of XAS in general, and EXAFS in particular, to probe a metal centre within a protein or when considering the results of such studies, the following should be noted:

(i) angular information is not obtained unless oriented samples are investigated;

(ii) rarely does the structural information extend beyond 4Å from the primary absorber (the metal atom);

(iii) the spectrum sums data for all atoms of a particular element and, if the element of interest is present in more than one chemical form, an *average* environment is obtained from the analysis of the data;

(iv) the possibility of radiation damage must be anticipated and the integrity of samples should be monitored after, and if possible during, measurement.

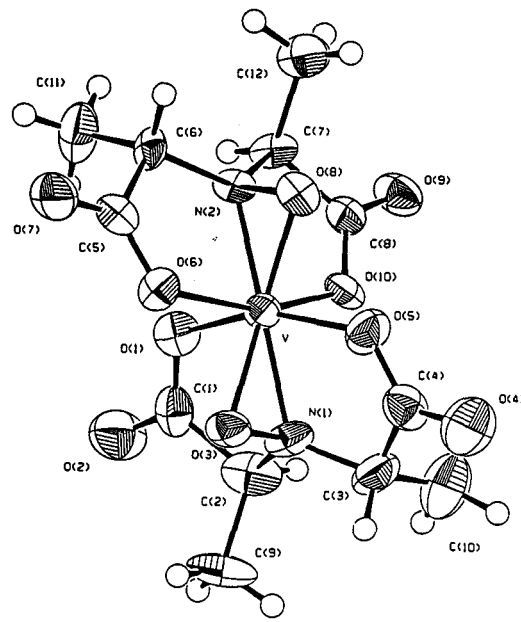


Fig. 3 The structure of $[\Delta\text{-V}((S,S\text{-hidpa})_2]^-$.

AMAVADIN

Vanadium is concentrated by fungi of the *Amanita* genus to remarkably high levels, with the highest concentrations (up to 400 ppm dry weight) being generally found in *Amanita muscaria* (ref. 8). The low molecular weight, blue V(IV) Amavadin is readily isolated from *Amanita muscaria* (ref. 9) and is constituted as a 1:2 complex of the metal with the proligand *(S,S)-N*-hydroxy-2,2′-iminodipropionic acid (H_3hidpa). Amavadin is reversibly oxidized to the V(V) level (ref. 10).

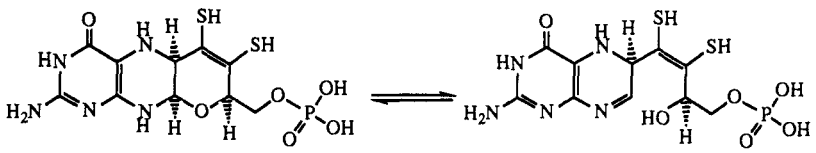
We have accomplished (ref. 11) an extensive series of spectroscopic studies for Amavadin in its V(IV) and V(V) oxidation states and a series of close chemical relatives. These studies have shown that Amavadin, as isolated, consists of an approximately equal mixture of $[\Delta\text{-V}\{(S,S)\text{-hidpa}\}_2]^{2-}$ and $[\Lambda\text{-V}\{(S,S)\text{-hidpa}\}_2]^{2-}$, each of these anions involving eight-co-ordinated V(IV) but *no* V=O group (see Fig. 3). The metal-ligand donor atom distances of Amavadin, obtained from an analysis of the vanadium K-edge EXAFS recorded for *Amanita muscaria*, are in good agreement with the dimensions obtained from the X-ray crystallographic determinations of the structures of the chemical relatives.

OXOMOLYBDOENZYMES AND THEIR TUNGSTEN COUNTERPARTS

This is an exciting time in the development of our understanding of the catalytic centres of the oxomolybdoenzymes and their close relatives, the tungsten enzymes *(tunzymes)*. Crystal structures of representatives of Mo and W families of enzymes have recently become available (refs. 12-14) and this new information, linked to previous spectroscopic studies and analogue chemistry, provides an excellent platform for further investigations.

Considerable EPR (ref. 15) and XAS (ref. 16) evidence has demonstrated that, in the oxmolybdoenzymes, molybdenum is co-ordinated to 1-2 oxoligands and 3-4 S-donor ligands, and *is* the catalytic centre. These enzymes contain cofactors (Moco (ref. 17) and Tuco), each of which involves either one or two molecules of a special pterin (molybdopterin (MPT)) bound to the metal. All eukaryotic molybdoenzymes contain MPT unmodified, prokaryotic forms usually contain a nucleotide appended to the MPT *via* a pyrophosphate linkage. Tunzymes contain MPT, either alone or modified. Based on the X-ray crystallographic determinations MPT can be represented as (1).

(1)

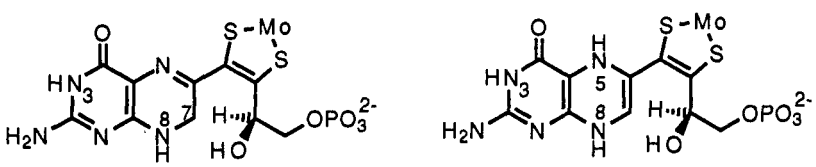


MPT ring-opened tautomer of MPT

The nature of MPT is remarkably close to the proposals of Rajagopalan *et al.* (ref. 18). Protein crystallography shows that the pterin rings of MPTs are not planar and, therefore, the structure corresponds to a reduced pterin. However, protein crystallography cannot accurately delineate the level of saturation/unsaturation at the pyrazine ring or the side-chain, sulfur-bearing, carbon atoms. In Moco and Tuco, both the pterin and the metal centre are "potentially" redox active. The redox changes at the metal centre can be linked to the oxidation levels of the pyrazine ring and the pterin may be considered as an electron transfer route to the metal (see ref. 13). Also, the different redox states of the pterin may stabilise different oxidation levels of the metal centre.

For redox co-operativity between the molybdenum centre and the pterin moiety in the operation of the oxomolybdoenzymes, there must be some mechanism for electronic communication between the metal centre and the pterin and/or the nitrogen(s) of the pyrazine ring. Accordingly, we favour (ref. 19) postulate (2) for active Moco in which C-6 is sp^2 hybridised.

(2)

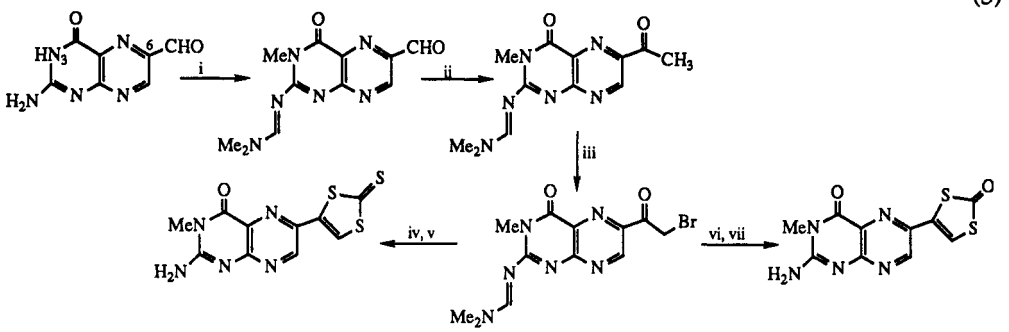


7,8-dihydropterin, or 5,8-dihydropterin

Towards the Synthesis of Molybdopterin

Molybdopterin as a free molecule is unlikely to be isolated - the ene-1,2-dithiol unit is a particularly sensitive moiety . In the actual cofactors, the two sulfur atoms are linked to a metal centre; in any synthetic plan this labile unit must be carried through in some protected form. Much of our work has utilised 6-formylpterin, available *via* a one-pot degradation of folic acid to produce unambiguously 6-substituted-pterins. A key to increasing the solubility of pterins in organic solvents is the removal of hydrogen-bonding possibilities, especially the hydrogen at N-3 (3).

(3)



Reagents: i, DMFDA, DMF, 110°C (61%); ii, CH_2N_2, CH_2Cl_2, rt (55%); iii, 2-pyrrolidone, HBr, Br_2, 80°C, AcOH (43%); iv, NaS_2CNMe_2, MeOH, CH_2Cl_2 (64%); v, c. H_2SO_4, 40°C, then NaSH, aq. AcOH, rt (10%); vi, KSC(S)O-i-Pr, i-PrOH, CH_2Cl_2, rt (98%); vii, c. H_2SO_4, rt (65%).

We have achieved the synthesis of $[MoO\{S_2C_2H(pterin)\}_2]^{2-}$ as a minimal model for Moco (Fig. 4).

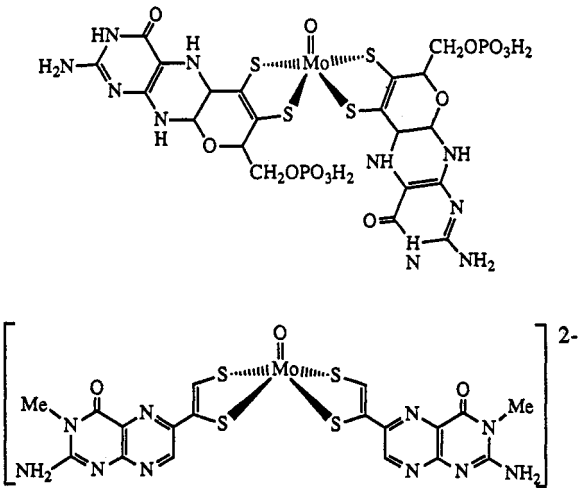


Fig. 4 Comparison of the postulated minimal model of the molybdenum centre in some oxymolybdoenzymes with the proposed structure of $[MoO\{S_2C_2H(pterin)\}_2]^{2-}$.

BOND VALENCE SUM ANALYSIS (BVSA)

Introduction

The Bond Valence Sum evolved from the efforts to formalise the "properties" of the chemical bond. Pauling initiated the process when he introduced the idea of the bond strength of a bond in one of his "principles determining the structure of complex ionic crystals" (ref. 20). He reasoned that the electric charge of each cation must balance the strength of the electrostatic valence bonds reaching it from the surrounding anions. Thus, for each cation, equation (4) applies where V_i is the valence (the formal oxidation state) of the cation, N_i is the co-ordination number of the cation, and s_{ij} is often referred to as the "Pauling bond strength".

$$s_{ij} = V_i / N_i \qquad (4)$$

Brown and Altermatt (ref. 21) proposed the use of equation (5) based upon Pauling's electrostatic valence principle, to describe the relationship between bond valence and bond length; s_{ij} is the bond valence of the bond r_{ij} between cation i and anion j, r_o is a constant calculated for a particular cation-anion pair and is defined as the length of a bond of unit valence. The bond valence sum parameters r_o and B were refined (using bond length data contained in the Inorganic Crystal Structure Database (ICSD) and the parameters were calculated by fitting equation (6), relating the oxidation state of the cation calculated from the sum of the contributions from each bond between the cation and the ligands, to the actual oxidation state of the metal centre.

$$s_{ij} = \exp((r_o - r_{ij})/B) \quad (5)$$

$$V_i = \Sigma_j s_{ij} \quad (6)$$

We have investigated an alternative application of the correlations provided by BVSA (ref. 22). If reasonably reliable values of r_o can be obtained from structural data available for co-ordination complexes contained within the Cambridge Structural Database (CSD), equations (5) and (6) can be used to obtain **co-ordination numbers** of metals in metalloenzymes in cases where the oxidation state and metal-ligand distances are known. This approach is especially complementary to XAS which provides information concerning the oxidation state of a metal atom (edge position) and the dimensions of the inner co-ordination sphere (from EXAFS).

Distortion Theorem

BVSA can be used to comment on the metal-ligand distances obtained from protein crystallographic studies of metalloproteins, since these are rarely defined with a precision better than ± 0.1Å. The distortion theorem states that the displacement of a metal atom from the centre of a polyhedron will result in an increase in the calculated bond valence sum (ref. 23). Therefore, the *minimum* value of the bond valence sum is calculated for the metal atom at the centre of a polyhedron. The distortion theorem has been applied, for example, to assess possible positions for metal sites in proteins (ref. 24).

As a refinement of this procedure, we have proposed (ref. 22) an approach which permits the position of a metal atom to be determined upon the basis of bond valence sum calculations, combined with a "check" on the bond lengths generated by a consideration of the position of the metal atom. If the calculated bond length lies within ±2σ of the mean bond length determined from the relevant cation-anion pair data set obtained from the CSD, a score of 1 is give to the bond length. If the bond lies to within ±(2-3)σ of the mean, the bond is given a score of 10, and if it is greater or less than the mean of the bond length ±3σ, the bond is accorded a score of 20. ***Thus, for a proposed position of a metal centre, ideally, the total bond length score should be no greater than the co-ordination number of the metal centre.*** This approach has been illustrated by a consideration of the position of a zinc atom in an environment of 4 cysteinyl sulfur atoms in rat liver metallothionein II (ref. 22).

CONCLUSIONS

The study of the co-ordination chemistry of d-transition metal atoms in proteins is clearly advancing rapidly. The synergy which arises by combining the results of crystallographic and spectroscopic investigations, together with information gained from studies of relevant co-ordination compounds, is increasing our understanding of the manner in which chemical behaviour is controlled at transition metal centres in proteins to produce particular biological functions. We advocate the use of Bond Valence Sum Analysis,

either simply or *via* a Distortion Theorem approach, to inspect the structure obtained from EXAFS analysis and/or protein crystallography for a d-transition metal site in a biological system; ***BVSA should be used as an aid to good judgement but not as a substitute for this.***

ACKNOWLEDGEMENTS

We thank The Royal Society for the award of a Fellowship (DC), the University of Manchester for the award of a Scholarship (EP) and the EPSRC for studentships (PB, CW) and the BBSRC for a PDRA award (AD).

REFERENCES

1. J.J.R., Fraústo da Silva and R.J.P. Williams. *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life.* Clarendon Press, Oxford (1991).
2. L. Banci, I. Bertini, K.L. Bren, M.A. Cremoni, H.B. Gray, C. Luchinat and P. Turano. *JBIC.* **1**, 117 (1996).
3. C.D. Garner. *Adv. Inorg. Chem.* **36**, 303 (1991) and references therein.
4. B.M. Kincaid and P. Eisenberger. *Phys Rev. Lett.* **34**, 1361 (1975).
5. A. Labhart and C. Yuen. *Nature (London),* **277**, 150 (1979).
6. R.G. Shulman, P. Eisenberger, B.H. Teo, B.M. Kincaid and G.S. Brown. *J. Mol. Biol.* **124**, 305 (1978).
7. N.J. Blackburn, S.S. Hasnain, N. Binstead, G.P. Diakun, C.D. Garner, and P.F. Knowles. *Biochem. J.* **219**, 985 (1984).
8. H-U. Meisch, W. Reinle and J.A. Schmitt. *Naturwissenschaften.* **66**, 620 (1979).
9. E. Bayer and H. Kneifel, Z. *Naturforsch. Teil B,* **27**, 207 (1972).
10. M.A. Nawi and J.L. Riechel. *Inorg. Chim. Acta* . **136**, 33 (1987).
11. E.M. Armstrong, R.L. Beddoes, L.J. Calviou, J.M. Charnock, D. Collison, N. Ertock, J.H. Naismith and C.D. Garner. *J. Am. Chem. Soc.* **115**, 807 (1993).
12. M.K. Khan, S. Mukund, A. Kletzin, M.W.W. Adams and D.C. Rees. *Science* . **267**, 1463 (1995).
13. M.J. Romao, M. Archer, I. Moura, J.J.G. Moura, J. LeGall, R. Engh, M. Schneider, P. Hof and R. Huber. *Science* . **270**, 1170 (1995).
14. H. Schindelin, C. Kisker, J. Hilton, K.V. Rajagopalan, and D.C. Rees. *Science* . **272**, 1615 (1996).
15. R.C. Bray, *Quart. Rev. Biophysics* . **21**, 299 (1988).
16. S.P. Cramer, *Advances in Inorganic and Bioinorganic Mechanisms* , Ed. A.G. Sykes, Academic Press, London, 259, (1983). B. Hedman, K.O. Hodgson and C.D. Garner. *'Synchrotron Radiation and Biophysics'* . Ed. S.S. Hasnain, Ellis Horwood, Chichester, 43 (1990).
17. IUPAC Glossary, *JBIC* . **1**, G20 (1996).
18. S. Gardlik and K.V. Rajagopalan, *J. Biol. Chem.* **265**, 13047 (1990).
19. D. Collison, C.D. Garner and J.A. Joule, *Chem. Soc. Rev.* **25**, 25 (1996).
20. L. Pauling. *J. Am. Chem. Soc.* **51**, 1010 (1929).
21. I.D. Brown and D. Altermatt. *Acta. Cryst.* **B41**, 244 (1985).
22. C.D. Garner, D. Collison and E. Pidcock. *Phil. Trans. R. Soc. Lond. A.* **354**, 325 (1996).
23. I.D. Brown and R.D. Shannon. *Acta. Cryst.* **A29**, 266 (1973).
24. D.S. Gregory, A.C.R. Martin, J.C. Cheetam, J.C. and A.R. Rees. *Protein Engineering.* **6**, 29 (1993).