# XAFS Debye-Waller factors for Zn metalloproteins

Nicholas Dimakis and Grant Bunker

*Illinois Institute of Technology, 3300 South Federal Street, Chicago, Illinois 60616–3793, USA*

An accurate and practical method for the calculation and use of thermal x-ray absorption fine structure (XAFS) Debye-Waller factors (DWFs) in active sites of metalloproteins is presented. These factors are calculated on model clusters within the local density functional approximation with nonlocal corrections. The DWFs are mapped out and parametrized as a function of the first shell distance and an angle (where applicable), for all significant single and multiple scattering paths, as well as the sample temperature. This approach is applied to the biologically essential but spectroscopically silent $Zn^{+2}$ active sites composed of histidines, cysteines, and carboxylate ligands in homogeneous and heterogeneous environments. Detailed analysis of the relative scattering paths for Zn metalloproteins using projected vibrational density of states further explain why these paths are not detectable by XAFS for first shell metal-ligand distances above a "cutoff" value.

## I. INTRODUCTION

### A. Background

X-ray absorption fine structure (XAFS) spectroscopy[1,2] is an important tool that can be used to obtain structural and electronic information of a sample in both crystalline and noncrystalline forms. In recent years substantial progress has been reported[3,4] in calculating the photoelectron single (SS) and multiple-scattering (MS) aspects of XAFS. Calculation of the essential vibrational properties in XAFS has generally lagged because of the difficulty of performing sufficiently accurate calculations of the forces between atoms. We have developed practical and generalizable models using the *ab initio* density functional theory (DFT)[5–7] to calculate accurate XAFS SS and MS parameters for active sites of metalloproteins. These models here are applied to $Zn^{+2}$ active sites. $Zn^{+2}$ possesses filled *d* orbital ($d^{10}$) and thus is spectroscopically silent to most spectroscopies such as ultraviolet[8]-visible[9] and electron paramagnetic resonance.[10] In the absence of a crystal structure, characterization of $Zn^{+2}$ sites is extremely difficult. However, since in XAFS there are no spectroscopically silent atoms and all sufficiently heavy atoms can be studied, XAFS has found frequent application in the characterization of biological $Zn^{+2}$ active sites.

The XAFS Debye-Waller factor (DWF) accounts for the structural and thermal disorder of a sample. In the simplest case (small or Gaussian disorder) it is expressed in the form $e^{-2k^2\sigma_j^2}$ where $\sigma_j^2$ is the mean square variance (MSV) of the *j*th photoelectron half-scattering path, and $\hbar k$ is the photoelectron momentum. Evidently its influence is most important at higher *k*. The DWF of a given path is a function of the normal mode phonon spectrum and depends on the vibrational eigenfrequencies and eigenvectors of the system under investigation.

The actual number of SS and MS paths depends on the local interatomic distribution in the sample. In cases where the absorbing atom is almost collinear with two scattering atoms, the MS scattering is strongly enhanced ("focusing effect") and will contribute substantially to the XAFS spectrum. This often occurs in active sites of metalloproteins,

causing the total number of scattering paths to be on the order of several hundred. It is well known that experimental XAFS data can only support a limited number of fitted parameters, i.e., $2\Delta k\Delta R/\pi + 2 \simeq 20{-}30$,[11] where $\Delta k \simeq 10{-}12$ Å$^{-1}$ and $\Delta R \simeq 3{-}4$ Å are the *k*- and *R*-space data ranges. For metal foils like Pt and Cu this number could slightly increase up to 32–34 parameters.[12] Additionally, in many cases where multiple scattering dominates these parameters are highly correlated with each other. Therefore these factors must be accurately known from an alternative source in order to quantitatively fit XAFS experimental data. Our work is an attempt to directly address this problem. Ideally the electronic structure calculation and the vibrational calculations would be carried out in the same programs, but until such codes are available the hybrid approach shown here demonstrates the potential viability of such an approach and provides usable results.

Metalloproteins appear in various natural states such as solutions and membranes, and also in crystalline forms. The term "active site" of a metalloprotein refers to an area in the vicinity of the metal atom of a metalloprotein that includes ligands from the amino-acid residues of the protein coordinated to the metal atom.[13,14] Various important protein functions such as oxygen transport, electron transfer, and oxidation depend on the structure of the active site. Binuclear or multinuclear metal sites and metal binding to other cofactors is also common, but are a topic of future studies and beyond the scope of this article. DWF calculations on the $Zn^{+2}$-tetraimidazole complex have been reported using both self-consistent (SCF) methods,[15,16] and force field models (FFM).[17–19] These calculations were an alternative to the single parameter Einstein or Debye models, which fail to describe either SS scattering and MS scattering paths in systems with heterogeneous bond strengths.[20]

We have found DWF calculations using DFT to be preferable over other SCF or FFM methods because of DFTs consistency and accuracy. However, DFT calculations on large active sites are not always practical or even possible. Poor SCF energy/geometry convergence, high CPU demand (execution times typically scale as $N^3$ where $N$ is number of basis functions used) and possible saddle points in the poten-
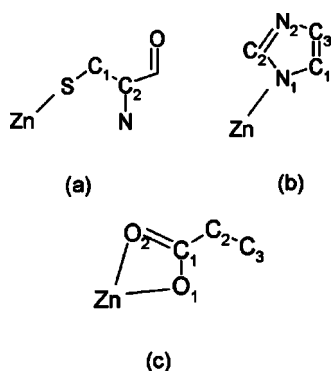
FIG. 1. Modeled structures used: (a) Zn-cysteine, (b) Zn-histidine, and (c) Zn-carboxylate model. An extra atom that is directly coordinated to the central metal atom and opposite to the ligand is not shown here.

tial energy surfaces for arbitrary geometries lead us to investigate alternative approaches for the DWF calculation.

A practical method that enables DWFs to be expressed as a function of first shell distance $R$ and temperature $T$ has been reported.[21] It is based on the observation that a reduced structure can be used to estimate the DWF parameters of a large active site (Ref. 15). This model described sites that consisted of Zn coordinated to histidine ligands, and provided accuracy sufficient for use with XAFS data analysis. The parametrizations permit use of these functions in a fitting procedure to analyze XAFS data, without the need for experimenters to perform any DFT calculations themselves. Since, apart from the temperature, the DWFs are expressed using only one additional parameter (the first shell distance), we refer to this as a one parameter model for calculation of the DWFs.

Zn amino-acid structures appear as $Zn^{+2}$, most often binding histidines (His), cysteines (Cys), and aspartic (Asp)/glutamic (Glu) acid ligands (Ref. 13). The latter two amino acids residues form mono- or bidentate conformations[22,23] with Zn binding to oxygen atom(s) of the carboxylate groups. With the exception to histidine ligand, in the other three amino-acids, the angle defined by the positions of the metal absorbing atom and two scatterers in the ligand may vary significantly due to steric hindrances, hydrogen bonding between residues, and other effects. In these cases the one parameter model will not be adequate to describe the DWFs of this particular model-ligand conformation. A two parameter model is needed to describe the dependence of the DWF on the first shell distance and an angle. Parametrization in this case is more complicated. In this paper we evaluate such a model for Zn-cysteine coordination [Fig. 1(a)] whereas Zn-carboxylates (Asp/Glu) [Fig. 1(c)] are represented by the same one parameter model accounting for symmetric bidentate conformations. The more complicated monodentate model will be presented elsewhere.

Another situation discussed here is when a metalloprotein has at least two different types of ligands (heterogeneous ligation). In this case it may not be possible to clearly resolve the first shell distance, thus the first shell DWFs might not be known from experimental XAFS spectra. The latter was previously used as an additional refinement parameter in the one

parameter Zn-histidine model (Ref. 21) to compensate for ring deformations due to different dynamics among the one ring and four ring case. An improved Zn-histidine model that predicts the first shell DWF at high accuracy is also presented here, and together with the Zn-cysteine and Zn-carboxylate models, they are applied to a series of DFT optimized structures, including active sites with homogeneous and heterogeneous ligation.

### B. Multiple scattering XAFS Debye-Waller factors

The mean square variance $\sigma_j^2$ of an arbitrary SS or MS half scattering path $j$ with respect to a central 0th absorbing atom at Gaussian disorder is given by (Ref. 20):

$$\sigma_j^2 = \frac{1}{2} \sum_n \left\{ \frac{1}{4} \sum_{\substack{k,l=0 \\ k \neq l}}^{j} \hat{R}_{kl} \cdot \left[ \frac{\epsilon_k(n)}{\sqrt{m_k}} - \frac{\epsilon_l(n)}{\sqrt{m_l}} \right] \right\}^2 \langle Q_n^2 \rangle, \quad (1)$$

where $\hat{R}_{kl} = \mathbf{R}_{kl}/|\mathbf{R}_{kl}|$ are the unit vectors for the equilibrium interatomic distance between atoms $k$ and $l$, $\epsilon_k(n)$ are the normal mode eigenvectors, $m_k$ is the mass of the $k$th atom in the cluster, and the mean square variance $\langle Q_n^2 \rangle$ in the $n$th normal mode is given by

$$\langle Q_n^2 \rangle = \frac{\hbar}{2\omega_n} \coth\left( \frac{\hbar\omega_n}{2K_B T} \right) \quad (2)$$

and $\omega_n$ is the corresponding phonon normal mode frequency for mode $n$. The first equation can be written as

$$\sigma_j^2 = \frac{1}{2\mu_j} \sum_n \rho_j(n) \langle Q_n^2 \rangle, \quad (3)$$

where $\rho_j(n)$ is the projected vibrational density of states (VDOS) (Ref. 17) that contributes to the $\sigma_j^2$, defined as

$$\rho_j(n) \equiv \mu_j \sum_m \left[ \frac{1}{4} \sum_{\substack{k,l=0 \\ k \neq l}}^{j} \hat{R}_{kl} \cdot \left[ \frac{\epsilon_k(m)}{\sqrt{m_k}} - \frac{\epsilon_l(m)}{\sqrt{m_l}} \right] \right]^2 \delta(\omega_m - \omega_n) \quad (4)$$

and $\mu_j$ is the reduced mass of the path used to unit normalize $\rho_j(n)$, i.e., $\Sigma_j \rho_j(n)^2 = 1$. The projected VDOS as defined by Eq. (4) can be interpreted as the probability that the displacement state defined by the vectors $\hat{R}_{kl}$ is in the vibrational mode $n$ and therefore it expresses the relative contribution of particular phonon mode to the $\sigma_j^2$. This is a temperature independent parameter and will be used to analyze the individual phonon mode contribution on the DWF of a particular scattering path.

## II. METHOD

The process starts by building a model metal-ligand structure as shown in Fig. 1, SCF geometrically optimizing it, and calculating the phonon normal mode vibrational spectrum. This was done on a series of Zn-ligand structures at various first shell distances and angles (where applicable) using DFT within the Unichem/DGAUSS 5.0[24] quantum chemistry

package, running on 4-processor Silicon Graphics Origin 200 server. The generalized gradient approximation (GGA) under gradient-corrected spin density functional of B88-PW91 was employed, which consists of the exchange B88[25–27] and correlation functional of Perdew, and Wang (PW91).[28] Use of the B88-PW91 functional and the replacement of the hydrogen atom opposite to the ligand and directly binding to the metal with the same atom as the first shell scatterer improved ligand geometry and accuracy of the calculated DWFs. The double zeta[29] basis set was used, which is specifically optimized for DFT.

Molecular modeled and large reference structures were optimized at high spin states that correspond to the ground state energy configuration, e.g., for Zn-histidine modeled-compounds of Fig. 1(b), and at the high spin configuration [spin multiplicity $(2S+1)=3$, where $S$ is the spin of the compound] the conformation energy was reported at $-56.019$ keV, lower than the corresponding energy value reported at $-56.017$ keV for the low-spin configuration ($2S+1=1$). DFT calculated phonon normal mode spectra, in the form of eigenfrequencies and eigenvectors, was input to a program written by the authors to calculate MSVs for all relevant SS and MS paths as shown by Eq. (1). For each metal-ligand model a primary geometrically optimized "equilibrium" configuration is obtained. In turn the first shell metal-ligand distance is altered by freezing the metal at the previously equilibrium position and translating the ligand along the metal-ligand axis, thus maintaining the internal geometry of the ligand. This new structure is geometrically optimized and with the exception of the hydrogen atoms and the atom that binds to the metal opposite to the ligand, all remaining atoms in the structure are fixed at these new positions. In order to ensure that modeled structures are at a stable configuration, i.e., all phonon normal mode frequencies are real, the variance of the distance/angle parameter is restricted. This ensures that clusters had converged to a conformational energy minimum rather than to a saddle point.

The $\sigma_j^2$ of an arbitrary scattering path can be expanded as a power series in terms of the distance $R$ between the metal and the ligands nearest atom, the angle $\Theta$ defined as the angle formed by the metal and two atoms in the ligand, and the temperature $T$. A general formula that expresses $\sigma^2(\Delta R, \Delta \Theta, T)$ was identified as

$$\sigma^2 = \sum_{ijk} D_{ijk} \Delta \Theta^j \Delta R^i T^k + C e^{\sum_{i'j'k'} D'_{i'j'k'} \Delta \Theta^{j'} \Delta R^{i'} T^{k'}}, \quad (5)$$

where $\Delta R = R_0 - R$, $\Delta \Theta = \Theta_0 - \Theta$, $R_0$, $\Theta_0$ are the equilibrium parameters, coefficients $D_{ijk}$ and $D'_{i'j'k'}$ are determined by fitting a single or two parameter expression with the DFT-calculated MSVs, and $C$ is 1 Å². The units of $D_{ijk}$ and $D'_{i'j'k'}$ coefficients are such that the product (1) $D_{ijk} \Delta \Theta^j \Delta R^i T^k$ is measured in Å², and (2) $D'_{i'j'k'} \Delta \Theta^{j'} \Delta R^{i'} T^{k'}$ is dimensionless. The first term of the series expansion in Eq. (5) (i.e., for $i,j,k=0$) denoted by $\sigma_0^2$ is MSV independent of $\Delta R$ and $\Delta \Theta$ parameters. For $\Delta R$ ranges used here, the exponential nature of the $\sigma_j^2$ only appears in some paths of the Zn-carboxylate model, stating that these paths will be undetectable by XAFS due to the rapid increase of the corresponding DWFs at first

shell distances greater than a path-dependent "cutoff" value. For example for the Zn-carboxylate model perturbing the first shell Zn-O$^{(1)}$ distance from 2.11 to 2.16 Å will substantially increase the $\sigma^2$ of the SS Zn-O$^{(2)}$ path by more than a factor of 10 (from 0.006 Å² to approximately 0.07 Å² at 300 K) which in turn will will decrease the XAFS $\chi(k)$ amplitude (at $k=10$ Å$^{-1}$) by

$$\frac{\chi^{R=2.16}(k)}{\chi^{R=2.11}(k)} \simeq \frac{e^{-k^2 \sigma_{R=2.16}^2}}{e^{-k^2 \sigma_{R=2.11}^2}} \simeq e^{-6.4} = 1.66 \times 10^{-3} \quad (6)$$

with respect to its value at $R=2.11$ Å. The minimum acceptable polynomial degree on the $\Delta \Theta$ parameter of the two parameter Zn-Cys model, is either one or two, depending on the appropriate scattering path. For the one parameter Zn-histidine and Zn-carboxylate models, the degree of $\Delta R$ depends on the choice of the ligand: two for Zn-histidines and carboxylates, and three for Zn-cysteines. Temperature $T$ is fitted using a third order polynomial and varies from 1 up to 400 K.

Coefficients for Zn-histidine, Zn-carboxylate models (Asp/Glu) and Zn-cysteine are presented in Tables I–III, respectively. The earlier equations can be used to either calculate MSV factors directly, or as input for program such as FEFFIT[30] as mathematical expressions that determine numerical path parameters.

An example is given for expressing $\sigma^2$ for the SS Zn-O$^{(1)}$ path of the Zn-carboxylate model at an arbitrary temperature $T \leq 80$ K and at the model equilibrium first shell distance $R_{\text{Zn-N}^{(1)}}^0$ (i.e., $\Delta R = 0$ Å). Using information from Table II $\sigma^2$ is written as

$$\sigma^2 = \left( 3.854 - \frac{2.31 \times 10^{-3}}{K} T \right.$$
$$+ \frac{6.61 \times 10^{-5}}{K^2} T^2 \frac{-6.91 \times 10^{-8}}{K^3} T^3$$
$$\left. + e^{-12.50 - \frac{6.67 \times 10^{-2}}{K} T + \frac{6.13 \times 10^{-4}}{K^2} T^2} \right) \times 10^{-3} \text{Å}^2 \quad (7)$$

where this expression is in the form of Eq. (5).

### III. RESULTS AND DISCUSSION

The reference structures that are used to test the previously described Zn models are large, DFT-optimized biomolecules. Their normal mode phonon spectrum approximates experimental infra-red or Raman spectra of corresponding real structures with high accuracy.[31] Using these structures as reference instead of experimental XAFS spectra has the advantage of performing a path-by-path MSV analysis rather than estimating the average MSV per shell. The $k$ range of the Fourier transforms of all XAFS $\chi(k)$ spectra reported later were taken from 2 to 18 Å$^{-1}$.

#### A. Zn-cysteines

##### 1. Zn-Cys model

The equilibrium parameter values $R_{\text{Zn-S}}^0$ and $\Theta_{\text{Zn-S-C}^{(1)}}^0$ of the Zn-Cys model were observed at 2.246 Å and 106.539°,

TABLE I. Calculated polynomial coefficients for SS and MS $\sigma^2$s for Zn-histidine group. The model equilibrium for the first shell distance Zn-N$^{(1)}$ is $R_0 = 1.96$ Å.

| Zn Path | $\Delta R$ | Coefficient ($\times 10^{-3}$) | | | $\Delta R^2$ | | | | $\sigma_0^2$ | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | $\times 10^{-3}$ | $\times 10^{-4}$ | $\times 10^{-7}$ | | $\times 10^{-2}$ | $\times 10^{-4}$ | $\times 10^{-6}$ | | $\times 10^{-3}$ | $\times 10^{-5}$ | $\times 10^{-8}$ |
| -N$^{(1)}$ | 5.860 | −3.43 | 2.29 | −2.75 | 2.918 | 5.32 | 4.96 | −0.700 | 2.552 | −1.98 | 2.91 | −2.88 |
| -N$^{(2)}$ | 3.968 | −1.25 | 2.07 | −2.61 | 4.668 | 5.55 | 4.43 | −0.616 | 3.126 | −2.65 | 4.57 | −4.72 |
| -C$^{(1)}$ | 3.664 | −1.40 | 1.73 | −2.14 | 5.521 | 6.64 | 5.53 | −0.743 | 3.956 | −2.60 | 7.62 | −8.38 |
| -C$^{(2)}$ | 3.522 | −2.54 | 1.64 | −2.02 | 5.073 | 6.06 | 4.93 | −0.688 | 4.092 | −2.70 | 7.48 | −8.14 |
| -C$^{(3)}$ | 4.569 | −2.93 | 2.43 | −3.06 | 5.137 | 5.16 | 5.33 | −0.741 | 3.368 | −2.95 | 4.92 | −5.06 |
| -N$^{(1)}$-N$^{(2)}$ | 4.159 | −2.30 | 2.13 | −2.66 | 4.131 | 5.46 | 4.15 | −0.581 | 3.000 | −2.56 | 3.89 | −3.88 |
| -N$^{(1)}$-N$^{(2)a}$ | 4.227 | −2.11 | 2.18 | −2.73 | 3.984 | 5.43 | 4.09 | −0.573 | 3.053 | −2.54 | 3.71 | −3.65 |
| -N$^{(1)}$-C$^{(3)}$ | 4.701 | −3.66 | 2.39 | −2.98 | 4.485 | 6.68 | 4.59 | −0.643 | 3.246 | −2.78 | 4.15 | −4.12 |
| -N$^{(1)}$-C$^{(3)a}$ | 4.777 | −2.50 | 2.42 | −3.02 | 4.475 | 5.70 | 4.61 | −0.646 | 3.309 | −2.76 | 3.99 | −3.91 |
| -N$^{(1)}$-C$^{(1)}$ | 4.474 | −2.43 | 2.04 | −2.49 | 4.681 | 5.07 | 4.95 | −0.690 | 3.192 | −2.28 | 4.26 | −4.41 |
| -N$^{(1)}$-C$^{(1)a}$ | 5.184 | −3.33 | 2.24 | −2.72 | 3.785 | 5.00 | 4.70 | −0.660 | 3.705 | −2.12 | 3.06 | −2.91 |
| -N$^{(1)}$-C$^{(2)}$ | 4.300 | −2.88 | 1.95 | −2.39 | 4.540 | 4.82 | 4.67 | −0.652 | 3.480 | −2.37 | 4.31 | −4.32 |
| -N$^{(1)}$-C$^{(2)a}$ | 5.029 | −3.26 | 2.21 | −2.69 | 3.757 | 4.93 | 4.44 | −0.625 | 3.985 | −2.18 | 3.13 | −2.79 |
| -C$^{(1)}$-C$^{(3)}$ | 4.215 | −3.11 | 2.09 | −2.60 | 4.500 | 6.93 | 4.69 | −0.659 | 3.691 | −2.80 | 5.89 | −6.23 |
| -C$^{(2)}$-N$^{(2)}$ | 3.821 | −3.62 | 1.91 | −2.38 | 4.272 | 6.22 | 4.45 | −0.626 | 3.431 | −2.75 | 5.67 | −6.03 |
| -C$^{(2)}$-N$^{(2)}$-N$^{(1)}$ | 4.045 | −4.08 | 2.03 | −2.53 | 3.526 | 6.34 | 4.06 | −0.577 | 3.110 | −2.60 | 4.59 | −4.77 |
| -C$^{(1)}$-C$^{(3)}$-N$^{(2)}$ | 4.120 | −3.85 | 2.09 | −2.60 | 3.270 | 6.43 | 3.89 | −0.547 | 2.877 | −2.41 | 3.84 | −3.92 |

[a]Triple scattering.

respectively, whereas $\Delta R$, $\Delta \Theta$ vary from 0 to 0.3 Å and −20 ° to +15 °, respectively, and counterclockwise is the positive direction for $\Theta_{\text{Zn-S-C}^{(1)}}$ angular rotation [Fig. 1(a)]. Stable Zn-Cys modeled configurations may still be obtained for larger $\Delta \Theta$ values with an increase of the corresponding Zn-S distance. In this case SS and MS paths beyond the first Zn-S shell will be undetectable by XAFS due to large DWFs and photoelectron scattering path lengths. Similarly, Zn-Cys modeled structures with $\Delta \Theta < -20°$ are unstable due to electrostatic repulsion between Zn and $C^{(1)}$ atoms. MSVs as a function of $\Delta R$ for the most important SS and MS paths of Zn-Cys model structures at 150 K are shown in Fig. 2, and can be further analyzed by observing their corresponding VDOS projected onto the central Zn atom. The VDOS that corresponds to the first shell Zn-S SS path can be seen in Fig. 3. At the equilibrium $R^0_{\text{Zn-S}}$ value three normal mode frequencies at approximately 146, 273, and 370 cm$^{-1}$ contribute the most to the corresponding MSV for this path. As $\Delta R$ is increased (a) all modes are downshifted (in cm$^{-1}$), and (b) the probability of the lowest mode is increased followed by a decrease on the probability of the highest mode. Since there an is inverse relation between $\sigma_j^2$ and normal mode frequencies [Eq. (2)], these statements confirm its increase proportional to $\Delta R$.

### 2. Comparison with reference structure

The two parameter Zn-Cys model is tested against the almost planar Zn-three cysteine structure (Fig. 4). The three sulfur atoms of the reference structure are positioned at 2.31, 2.33, and 2.36 Å away from the central Zn$^{+2}$ atom whereas the average Zn-S-C$^{(1)}$ angle was observed at 107.13° ±1.2°.

No imaginary frequencies appear in the phonon spectrum of the geometrically optimized reference structure. However, for structures of this size a real phonon spectrum does not always guarantee that the reference structure, at its optimal geometry configuration, has converged to a global energy minimum. It was observed that inclusion of the lowest detected normal mode frequency of 9.04 cm$^{-1}$ in the calculation of the $\sigma_j^2$ of SS Zn-C$^{(1)}$, Zn-C$^{(2)}$, and DS Zn-C$^{(1)}$-C$^{(2)}$ paths would have artificially increased their value due to the inverse relation between $\sigma_j^2$s and normal mode frequencies (Table IV). Other paths did not seem to depend on this particular frequency. Therefore for consistency reasons this frequency should not included in the DWF calculations.

Calculated MSVs using the Zn-Cys model and the DFT optimized reference structure for the most important scattering paths at three different temperatures and with no angular variation (since $\Delta \Theta$ is minimal in this case) are presented in Table IV. For this particular metal-ligand conformation the value of Zn-S-C$^{(1)}$ angle is difficult to be predicted accurately by experimental XAFS data, especially at high temperatures. Therefore the $\Delta \Theta$ parameter will be used to define an acceptable range for the corresponding high shells SS and MS paths. These MSVs values are inputted as parameters in the FEFF8$^{32}$ simulated $\chi(k)$ XAFS spectra of the reference structure and their Fourier transforms is shown in Fig. 5. The agreement between our model and the reference structure indicates that ligand coordination number although will affect the first shell Zn-S distance will not explicitly alter the DWFs themselves. The VDOS for the SS Zn-S path of the model at $R_{\text{Zn-S}}=2.33$ Å and DFT optimized reference structure is shown in Fig. 6. The three major peaks that contribute most to the DWF of this path appear at similar frequency

TABLE II. Calculated polynomial coefficients for SS and MS $\sigma^2$s for Zn-carboxylate group. The model equilibrium value for the first shell distance Zn-$O^{(1)}$ is $R_0=2.06$ Å. The first column stands for the power of $\Delta R$ as dictated by Eq. (5).

| Term | | | | | | | | Zn path | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | -$O^{(1)}$ | -$O^{(2)}$ | -$O^{(1)}$ | -$C^{(2)}$ | -$C^{(1)}$ -$C^{(2)}$ | TS$^a$ | -$O^{(1)}$ -$O^{(2)d}$ | -$C^{(1)}$ -$O^{(1)}$ | TS$^a$ | -$C^{(1)}$ -$O^{(2)}$ | TS$^a$ | TS$^e$ | TS$^f$ |
| ($\times 10^{-3}$) | | | | | | | | Coefficient | | | | | | |
| $\sigma_0^2$ | | 3.854 | 4.114 | 2.681 | 3.237 | 3.236 | 3.238 | 2.802 | 2.711 | 5.287 | 2.812 | 5.265 | 4.079 | 3.926 |
| | $\times 10^{-3}$ | −2.31 | −2.00 | −2.08 | −2.67 | −2.69 | −2.72 | −0.99 | −1.75 | −1.71 | −1.57 | −1.92 | −1.40 | −1.64 |
| | $\times 10^{-5}$ | 6.61 | 7.79 | 2.78 | 3.67 | 3.73 | 3.59 | 1.28 | 2.85 | 2.65 | 3.31 | 2.54 | 3.65 | 3.00 |
| | $\times 10^{-8}$ | −6.91 | −8.37 | −2.61 | −3.35 | −3.31 | −3.33 | −7.35 | −2.77 | −2.41 | −3.33 | −2.60 | −3.60 | −2.79 |
| $\Delta R$ | | 14.65 | 17.841 | 4.310 | 3.830 | 3.854 | 3.864 | 4.535 | 5.474 | 4.574 | 7.216 | 4.49 | 9.015 | 7.256 |
| | $\times 10^{-3}$ | 22.76 | −45.00 | −6.71 | −5.27 | −5.58 | −5.41 | −11.81 | 4.27 | −5.85 | 4.73 | −7.43 | 9.31 | −1.29 |
| | $\times 10^{-4}$ | 7.23 | 9.27 | 1.87 | 1.92 | 1.90 | 1.91 | 1.34 | 2.20 | 1.85 | 3.37 | 1.78 | 4.19 | 3.03 |
| | $\times 10^{-7}$ | −9.05 | −11.74 | −2.27 | −2.39 | −2.35 | −2.37 | −1.46 | −2.60 | −2.22 | −4.16 | −2.12 | −5.21 | −3.65 |
| $\Delta R^2$ | | 88.04 | 114.84 | 18.84 | 18.54 | 19.05 | 19.50 | 19.39 | 13.91 | 14.09 | 38.06 | 17.52 | 50.90 | 34.89 |
| | $\times 10^{-2}$ | 96.70 | 124.88 | 10.92 | 7.13 | 8.78 | 12.29 | −3.92 | 34.18 | 10.01 | 29.43 | 3.31 | 39.74 | 22.54 |
| | $\times 10^{-4}$ | 56.11 | 74.35 | 12.72 | 12.45 | 12.85 | 13.07 | 9.30 | 5.11 | 8.90 | 24.14 | 11.94 | 33.11 | 20.32 |
| | $\times 10^{-6}$ | −7.38 | −9.73 | −1.67 | −1.64 | −1.71 | −1.73 | −1.14 | −0.55 | −1.15 | −3.16 | −1.57 | −4.41 | −2.60 |
| $e^{\sigma_0^2}$ | ($T$ $\leqslant 80$ K)$^b$ | −12.50 | −12.41 | | | | | −13.71 | −13.09 | −13.62 | | | −13.27 | −13.51 |
| | $\times 10^{-2}$ | −6.67 | −6.25 | | | | | −7.09 | −1.08 | −6.62 | | | −6.65 | −7.02 |
| | $\times 10^{-4}$ | 6.13 | 5.97 | | | | | 6.20 | 8.93 | 6.23 | | | 6.30 | 6.23 |
| | ($T$ $\geqslant 80$ K) | −14.62 | −14.33 | | | | | 16.12 | −16.37 | −15.66 | | | 15.26 | −15.87 |
| | $\times 10^{-2}$ | 1.01 | 1.06 | | | | | 1.02 | 1.00 | 1.07 | | | 1.01 | 1.04 |
| | $\times 10^{-5}$ | −1.11 | −1.20 | | | | | −1.12 | −1.07 | −1.22 | | | −1.09 | −1.18 |
| $e^{\Delta R}$ | ($T$ $\leqslant 80$ K) | 75.76 | 76.15 | | | | | 70.75 | 47.11 | 74.62 | | | 74.65 | 72.64 |
| | | 1.00 | 1.09 | | | | | 1.03 | 1.38 | 1.02 | | | 1.02 | 1.03 |
| | $\times 10^{-3}$ | −7.06 | −9.03 | | | | | −7.40 | −9.91 | −7.63 | | | −7.74 | −7.44 |
| | ($T$ $\geqslant 80$ K) | 109.05 | 108.55 | | | | | 106.10 | 94.50 | 106.95 | | | 107.00 | 107.70 |

$^a$Triple scattering (TS) of the corresponding double scattering path at the previous column.
$^b$The exponential terms of the $\sigma^2$ are described differently for $T \leqslant 80$ K and $T \geqslant 80$ K.
$^c$Blanks denote zeros.
$^d$Additional terms $\Delta R^3$ apply: $54.62+(2.08/K)T-(5.44/K^2)\times 10^{-3}T^2-(7.64/K^3)\times 10^{-6}T^3$.
$^e$Zn-$C^{(1)}$-$O^{(1)}$-$O^{(2)}$ is a four-atom triple scattering path.
$^f$Zn—$C^{(1)}$—$O^{(2)}$—$O^{(1)}$ is a four-atom triple scattering path.

ranges as discussed previously. Peaks between modeled and reference structure differ by less than 15 cm$^{-1}$ which explains the agreement between the model and the reference structure for this path.

## B. Zn-histidines

The one parameter Zn-histidine model that predicts the first shell SS $\sigma_j^2$ at higher accuracy than the one reported in Ref. 21 is compared with previous XAFS (Ref. 19) and FFM (Ref. 18) references and with the Zn-tetraimidazole structure optimized under DFT B88-PW91 functional. The equilibrium parameter $R_{Zn-N^{(1)}}^0$ of the Zn-histidine model is 1.96 Å. In the Zn-histidine model, histidine amino acid is approximated by the imidazole ring [Fig. 1(b)]. In histidines addi-

tional atoms (a carbon atom, amine, and carboxylate groups) are bound to $C^{(3)}$ atom of the imidazole ring.

### 1. Zn-tetraimidazole structure under DFT B88-PW91 functional

DFT calculated optimal geometry and phonon normal mode spectrum of the Zn-tetraimidazole molecule under the local spin density approximation had been reported (Ref. 21). Since all metal-amino acid model and reference structures presented in this work have their structural and vibrational properties obtained under GGA (B88-PW91 functional), for consistency reasons we recalculated the Zn-tetraimidazole properties under the B88-PW91 functional.

The four nitrogen atoms of the geometrically optimized Zn-tetraimidazole molecule under B88-PW91 functional are symmetrically positioned at approximately 2.046 Å away

TABLE III. Calculated polynomial coefficients for SS and MS $\sigma^2$s for Zn-cysteine group. The model equilibrium values are $R_0 = 2.246$ Å and $\Theta_0 = 106.539°$. The first column stands for the power of $\Delta R$ and $\Delta\Theta$ as dictated by Eq. (5).

| Term | | -S | -C$^{(1)}$ | -C$^{(2)}$ | -S-C$^{(1)}$ | Zn Path -S-C$^{(2)}$ | -C$^{(1)}$-C$^{(2)}$ | -C$^{(1)}$—C$^{(2)\dagger}$ | -S—C$^{(1)}$-C$^{(2)}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Coefficient ($\times 10^{-3}$) | | | | |
| $\sigma_0^2$ | | 2.395 | 4.485 | 4.244 | 3.472 | 4.013 | 4.234 | 4.890 | 3.952 |
| | $\times 10^{-3}$ | −2.38 | 13.63 | 9.96 | 2.50 | 0.242 | 9.73 | 12.37 | 1.61 |
| | $\times 10^{-5}$ | 3.38 | 11.13 | 10.25 | 5.06 | 7.83 | 10.35 | 11.90 | 5.41 |
| | $\times 10^{-8}$ | −3.19 | −13.29 | −12.13 | −5.46 | −8.89 | −12.36 | −14.31 | −5.87 |
| $\Delta R$ | | 3.047 | 0.658 | 1.385 | 2.355 | 2.090 | 1.059 | 0.806 | 2.172 |
| | $\times 10^{-3}$ | 31.58 | −4.47 | −44.01 | 24.67 | −31.57 | −8.43 | 21.16 | 16.53 |
| | $\times 10^{-4}$ | 1.49 | 0.30 | 0.86 | 1.39 | 1.37 | 0.68 | 0.60 | 1.31 |
| | $\times 10^{-7}$ | −1.86 | −0.37 | −1.15 | −1.79 | −1.81 | −0.90 | −0.80 | −1.72 |
| $\Delta R^2$ | | 1.476 | −2.071 | −0.220 | −4.408 | −2.063 | 1.40 | −0.19 | −1.216 |
| | $\times 10^{-2}$ | −47.44 | −12.73 | 5.52 | −36.73 | −39.81 | −14.02 | −36.31 | −30.25 |
| | $\times 10^{-4}$ | −1.95 | −2.45 | −1.45 | −4.44 | −2.78 | | −0.82 | −2.63 |
| | $\times 10^{-6}$ | 0.33 | −0.35 | −0.23 | −0.61 | 0.41 | | −0.14 | −0.40 |
| $\Delta R^3$ | | 43.68 | 20.87 | 28.95 | 42.51 | 38.83 | 24.06 | 26.32 | 36.96 |
| | | 1.832 | 1.50 | 1.08 | 1.67 | 1.80 | 1.28 | 1.79 | 1.53 |
| | $\times 10^{-3}$ | 3.84 | 1.80 | 2.45 | 3.55 | 3.24 | 2.10 | 2.23 | 3.18 |
| | $\times 10^{-6}$ | −5.32 | −2.50 | −3.42 | −4.89 | −4.47 | −2.91 | −3.10 | −4.39 |
| $\Delta\Theta$ | $\times 10^{-3}$ | [a] | 5.566 | −6.59 | 14.32 | 5.10 | −4.267 | 13.12 | 1.726 |
| | $\times 10^{-5}$ | | 5.54 | 53.90 | 3.90 | −7.07 | −27.17 | −6.73 | 3.72 |
| | $\times 10^{-7}$ | | 2.22 | −7.20 | 5.65 | 3.52 | −3.07 | | 4.00 |
| | $\times 10^{-10}$ | | −2.85 | 9.66 | −6.92 | −4.60 | 3.89 | | −5.17 |
| $\Delta\Theta^2$ | $\times 10^{-4}$ | | | −2.02 | | 0.04 | −1.36 | | 0.1 |
| | $\times 10^{-5}$ | | | 1.20 | | 0.46 | 0.06 | | 1.21 |
| | $\times 10^{-8}$ | | | −2.58 | | | −0.95 | | 0.32 |
| | $\times 10^{-11}$ | | | 3.47 | | | −1.26 | | −0.46 |
| $\Delta R\Delta\Theta$ | $\times 10^{-2}$ | | 0.06 | 14.73 | 6.40 | 0.963 | 8.30 | −6.79 | 0.23 |
| | $\times 10^{-3}$ | | 0.15 | −11.29 | 0.76 | 3.538 | 6.78 | 4.31 | 1.66 |
| | $\times 10^{-5}$ | | | 1.46 | 0.45 | | 0.578 | 0.44 | |
| | $\times 10^{-8}$ | | | −1.97 | −0.59 | | −0.785 | −0.58 | |
| $\Delta R\Delta\Theta^2$ | $\times 10^{-3}$ | | | 6.547 | | −0.528 | 4.181 | | |
| | $\times 10^{-5}$ | | | −1.15 | | −0.40 | 13.45 | | −8.84 |
| | $\times 10^{-7}$ | | | 7.24 | | 0.67 | 2.74 | | |
| | $\times 10^{-10}$ | | | −9.69 | | −1.01 | −3.70 | | |
| $\Delta R^2\Delta\Theta$ | $\times 10^{-1}$ | | | −7.57 | −5.47 | −0.88 | −3.95 | −4.14 | 0.19 |
| | $\times 10^{-3}$ | | −19.21 | −79.79 | −9.85 | −26.77 | −53.20 | −37.70 | 20.00 |
| | $\times 10^{-5}$ | | | −8.35 | −3.75 | −0.84 | 2.94 | −2.61 | |
| | $\times 10^{-8}$ | | | 11.34 | 4.97 | 1.20 | 4.07 | 3.48 | |
| $\Delta R^2\Delta\Theta^2$ | $\times 10^{-2}$ | | | −3.80 | | −2.29 | | | |
| | $\times 10^{-5}$ | | | −53.68 | | −40.00 | −13.21 | | 12.78 |
| | $\times 10^{-6}$ | | | −4.57 | | −0.64 | −1.64 | | |
| | $\times 10^{-9}$ | | | 6.13 | | 1.00 | 2.25 | | |
| $\Delta R^3\Delta\Theta$ | | | 0.656 | 2.00 | 1.451 | 0.432 | 1.397 | 1.406 | 0.30 |
| | $\times 10^{-2}$ | | 7.97 | 19.63 | 5.02 | 8.87 | 14.88 | 11.48 | 7.83 |
| | $\times 10^{-5}$ | | 5.56 | 20.63 | 10.80 | 4.15 | 11.08 | 10.00 | 2.55 |
| | $\times 10^{-8}$ | | −7.55 | −28.09 | −14.58 | −5.89 | −15.37 | −13.55 | −3.56 |
| $\Delta R^3\Delta\Theta^2$ | $\times 10^{-2}$ | | | 7.77 | | 1.70 | 5.39 | | 1.12 |
| | $\times 10^{-3}$ | | | 2.53 | | 2.05 | 3.98 | | 1.76 |
| | $\times 10^{-6}$ | | | 9.32 | | 2.39 | 4.19 | | 1.15 |
| | $\times 10^{-8}$ | | | −1.25 | | −0.35 | −0.58 | | −0.16 |

[a] Blanks denote zeros.

from the central Zn atom. This is an overestimation by 0.057–0.046 Å of the corresponding $R_{\text{Zn-N}^{(1)}}$ first shell parameter of the Zn-tetraimidazole fluoroborate/perchlorate crystalline structure (Ref. 18) as observed experimentally.

Overestimation of the metal-imidazole distance is expected since our structure corresponds to 300 K temperature leading to an increased metal-imidazole distance when compared to same molecule under low temperature conditions.
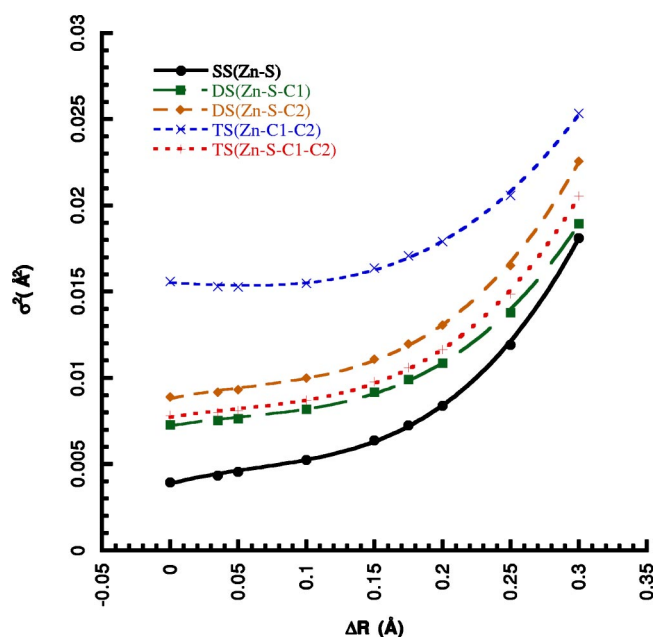
FIG. 2. (Color online) DFT-calculated $\sigma^2$ for various SS and MS paths of Zn-cysteine model at 300 K. Lines represent third order polynomial fits.

Two imaginary frequencies appear in the phonon normal mode spectrum: $i35.7$ and $i25.56$ cm$^{-1}$. The presence of the imaginary frequencies in the phonon normal mode spectrum does not change the $\sigma^2$s themselves: for arbitrary $j$ scattering path VDOS $\rho_j(n) \simeq 0$ where $n$ is any of the imaginary modes of the phonon spectrum [Eq. (3)]. However, because the DFT-geometry corresponds to a saddle point rather than an energy minima, an overestimation of the Zn-C$^{(1,2)}$ "bending" metal-imidazole mode is expected.
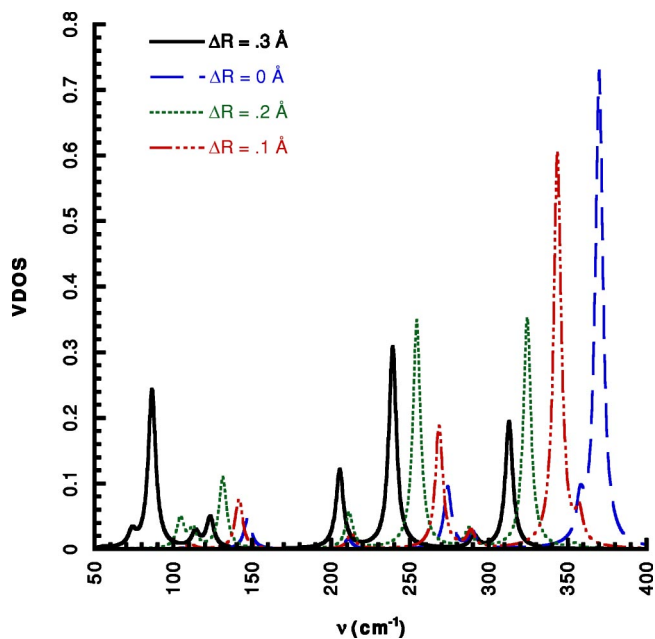


FIG. 3. (Color online) Projected VDOS for the SS Zn-S path at various perturbed first shell $\Delta R_{\text{Zn-S}}$ distances.
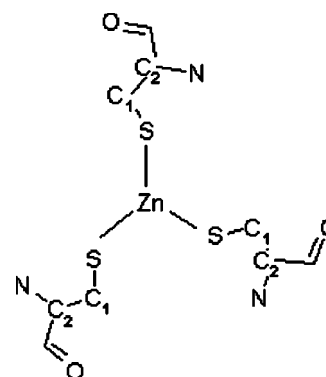


FIG. 4. Zn coordinated to three cysteines.

### 2. Comparing the Zn-imidazole model with references

The Zn-imidazole model predicts the first shell SS $\sigma_j^2$ at higher accuracy than the last reported model. At 20 K the $\sigma_j^2$ obtained using the Zn-imidazole model for the first shell SS Zn-N$^{(1)}$ is $2.65 \times 10^{-3}$ Å$^2$ whereas the corresponding values at the same temperature reported from FFM models and XAFS results using the tetrafluoroborate compound $[R_{\text{Zn-N}^{(1)}} = 2.00$ Å$]$ are $(2.61 \pm 0.01) \times 10^{-3}$ Å$^2$ and $(2.50 \pm 0.2) \times 10^{-3}$ Å$^2$, respectively (Refs. 18 and 19). This is also an improvement when compared to the value reported by the previous Zn-imidazole model ($2.77 \times 10^{-3}$ Å$^2$ in Ref. 21). At room temperature (300 K), and for the same path our improved model predicts $4.24 \times 10^{-3}$ Å$^2$ versus $4.26 \pm 0.03$

TABLE IV. Calculated $\sigma^2$ for the Zn-cysteine model using information from Table III vs reference for 20, 150, and 300 K for the most important scattering paths.

| | | | Reference | |
| --- | --- | --- | --- | --- |
| | *T* model | | Corrected | Raw |
| Path | K | | $\sigma^2 (\times 10^{-3}$ Å$^2)$ | |
| Zn-S | 20 | 2.643 | 2.673 ± 0.015 | 2.690 ± 0.01 |
| | 150 | 3.237 | 3.240 ± 0.04 | 3.329 ± 0.03 |
| | 300 | 4.925 | 4.915 ± 0.08 | 5.095 ± 0.07 |
| Zn-C$^{(1)}$ | 20 | 4.799 | 4.976 ± 0.02 | 5.356 ± 0.03 |
| | 150 | 8.433 | 9.431 ± 0.15 | 12.33 ± 1.3 |
| | 300 | 14.70 | 16.44 ± 0.3 | 22.23 ± 2.5 |
| Zn-C$^{(2)}$ | 20 | 4.497 | 4.482 ± 0.08 | 4.654 ± 0.1 |
| | 150 | 7.110 | 7.560 ± 0.25 | 8.922 ± 0.8 |
| | 300 | 12.12 | 13.092 ± 0.95 | 15.57 ± 1.2 |
| Zn-S-C$^{(1)}$ | 20 | 3.947 | 3.754 ± 0.02 | 3.840 ± 0.08 |
| | 150 | 5.163 | 5.376 ± 0.03 | 6.123 ± 0.05 |
| | 300 | 8.015 | 8.366 ± 0.08 | 9.860 ± 0.1 |
| Zn-S-C$^{(2)}$ | 20 | 4.252 | 4.101 ± 0.02 | 4.134 ± 0.02 |
| | 150 | 5.910 | 5.970 ± 0.1 | 6.204 ± 0.15 |
| | 300 | 9.228 | 9.602 ± 0.4 | 10.07 ± 0.4 |
| Zn-C$^{(1)}$-C$^{(2)}$ | 20 | 4.540 | 4.596 ± 0.04 | 4.876 ± 0.1 |
| | 150 | 7.594 | 8.203 ± 0.3 | 10.11 ± 1.5 |
| | 300 | 13.10 | 14.099 ± 0.5 | 17.912 ± 3.0 |
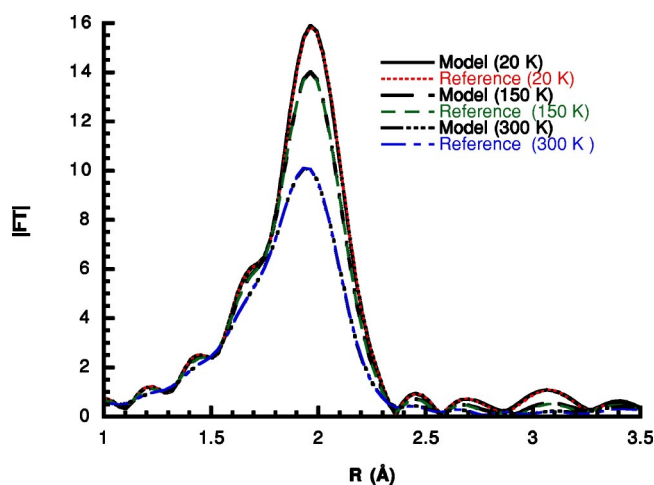
FIG. 5. (Color online) Zn absorption line for which the Fourier transform of the $\chi(k)$ XAFS spectra is calculated for the DFT-optimized Zn-three-cysteine structure at 20, 150, and 300 K.

$\times 10^{-3}$ Å$^2$ for the FFM model. The corresponding value reported by the previous model in Ref. 21 is $4.57 \times 10^{-3}$ Å$^2$. Similar improvement is obtained when either model is tested under the B88-PW91 optimized Zn-tetraimidazole structure $[R_{\text{Zn-N}^{(1)}} = 2.046$ Å$]$.

The accuracy of the model presented here for $\sigma_j^2$s is not limited to first shell SS but is also extended to higher shell SS and MS paths. For instance due to the presence of the focusing effect the double scattering path Zn-$N^{(1)}$-$N^{(2)}$ is of significant importance in the higher shell region of the Fourier transform. Its $\sigma_j^2$ at 300 K is predicted at $5.015 \times 10^{-3}$ Å$^2$ versus $4.93 \times 10^{-3}$ Å$^2$ for the FFM model (Refs. 18 and 19) again an improvement with the value of $5.31 \times 10^{-3}$ Å$^2$ reported by Ref. 21. This information together with $\sigma_j^2$ for other important paths is summarized in Table V. What is also of importance is similar to the Zn-cysteine model, the Zn coordination number does not affect the $\sigma_j^2$s.

One other issue that we discuss here is the imidazole ring deformation. Using the B88-PW91 Zn-tetraimidazole structure as reference the $N^{(1)}$-$N^{(2)}$ distance is shorter with respect to the same distance found in the Zn single histidine structure by 0.004 Å and the $N^{(1)}$-$C^{(3)}$ distance is also shorter by about 0.002 Å. This small ring deformation of the model structures mostly affects the DS/TS Zn-$N^{(1)}$-$C^{(3)}$ and DS Zn-$N^{(2)}$-$C^{(2)}$ paths whereas its effect on the other SS/MS paths is minimal. This systematic overestimation/underestimation of the corresponding MSV of the scattering path may be corrected by decreasing/increasing the first shell distance by an amount comparable to the one mentioned earlier. However, since the histidine amino acid had been approximated by the imidazole ring and the presence of additional atoms bound to $C^{(3)}$ atom of the histidine amino acid will cause additional deformation of the ring, no such correction of the MSVs was employed here.

### C. Zn-carboxylate

#### 1. Zn-carboxylate model

Carboxylate ligands appear in the form of aspartic and glutamic amino-acid structures (Fig. 7). A single Zn-
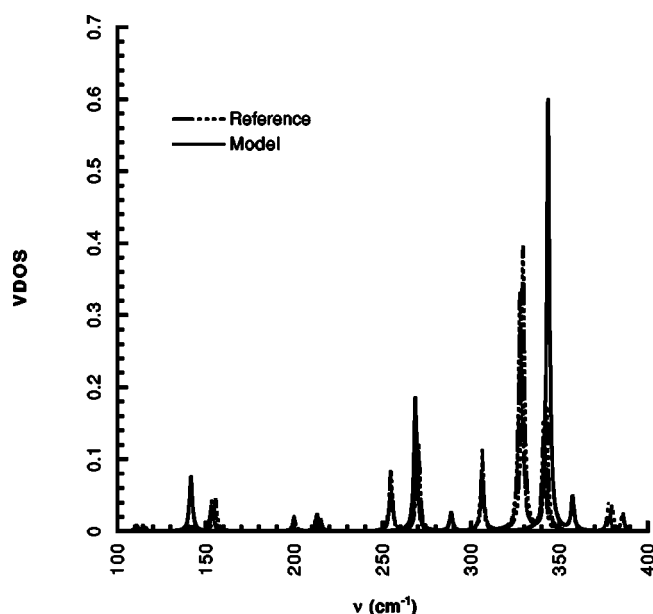


FIG. 6. Projected VDOS vs normal mode frequency for the SS Zn-S path using the planar three cysteine structure as a reference and the Zn-Cys model.

carboxylate model of a reduced structure is used for the DWF calculation of either carboxylate amino acid residue (Asp/Glu). This is justified since (a) no atoms more distant than the $C^{(2)}$ are XAFS detectable, and (b) Glu and Asp only differ by an extra carbon atom that is also undetectable by XAFS. A central metal atom is coordinated with carboxylate residues forming mono- or bi-dentate complexes. Due to the exponential dependence of most of the $\sigma_j^2$s in the Zn-carboxylate configuration on the first distance (Table II), the bidentate metal-oxygen conformation is considered separately from the monodentate case, and will be presented elsewhere. To preserve the symmetry, the Zn bi-dentate carboxylate model was translated along the Zn-$C^{(1)}$ axis. The values of the equilibrium Zn-oxygen interatomic distances were 2.06 Å Zn-$O^{(1)}$ and 2.08 Å Zn-$O^{(1)}$. Parameter $\Delta R$ for Zn
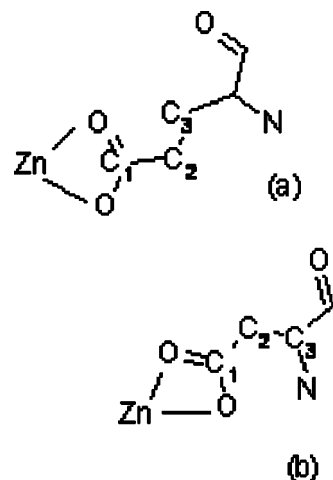


FIG. 7. Reference structures used (a) Zn-Glu and (b) Zn-Asp.

TABLE V. Calculated $\sigma^2$ for the Zn-imidazole model using information from Table I vs the last reported Zn-imidazole model (Ref. 16) and other references for the most important scattering paths.

| Path | $R_{Zn-N^{(1)}}$ Å | T K | Model | | Other reference |
|------|------|------|------|------|------|
| | | | New | Ref. 21 $\sigma^2(\times 10^{-3}$ Å$^2)$ | |
| ZnLIN[1] | 1.983 | 20 | 2.65 | 2.77 | 2.61±0.01 (Ref. 19) |
| | | | | | 2.046±0.2 (Ref. 18) |
| | 2.046 | | 3.06 | 3.21 | 2.79±0.01[a] |
| | 1.983 | 300 | 4.24 | 4.63 | 4.26±0.03 (Ref. 19) |
| | 2.046 | | 5.68 | 6.30 | 4.87±0.04 |
| Zn—C[1,2] | 1.983 | 20 | 4.07±0.06 | 4.25±0.06 | 4.12±0.11 (Ref. 19) |
| | 2.046 | | 4.35±0.05 | 4.56±0.04 | 4.30±0.07 |
| | 1.983 | 300 | 8.11±0.01 | 8.81±0.5 | 8.59±0.23 (Ref. 19) |
| | 2.046 | | 9.26±0.04 | 10.08±0.5 | 10.52±0.7 |
| Zn—N[2] | | 150 | 4.40 | 4.23 | 4.52±0.14 |
| Zn—C[3] | | | 4.77 | 5.33 | 4.62±0.1 |
| Zn—N[1]—N[2] | 1.983 | 20 | 3.12 | 3.18 | 2.83 (Ref. 19) |
| | 2.046 | | 3.43 | 3.53 | 3.23±0.03 |
| | 1.983 | 300 | 5.09 | 5.43 | 4.93 (Ref. 19) |
| | 2.046 | | 6.37 | 6.93 | 6.11±0.05 |
| Zn—N[1]—C[3] | 1.983 | 20 | 3.38 | 3.50 | 3.17±0.03 (Ref. 19) |
| | 2.046 | | 3.73 | 3.89 | 3.36±0.02 |
| | 1.983 | 300 | 5.45 | 6.11 | 5.56±0.1 (Ref. 19) |
| | 2.046 | | 6.88 | 7.80 | 6.13±0.08 |
| Zn—N[1]—C[1,2] | | 150 | 4.61±0.13 | 4.87±0.05 | 4.40±0.07 |
| Zn—C[1]—C[3] | | | 5.21 | 5.80 | 5.45±0.15 |
| Zn—N[2]—C[2] | | | 4.83 | 5.03 | 5.32±0.32 |

[a]When no reference is shown these factors correspond to the optimized B88-PW91 Zn-tetraimidazole structure.

-O[1] distance varied from −0.05 to 0.09 Å with any further $\Delta R$ variation leading to unstable energy conformations. Various $\sigma_j^2$ vs $\Delta R$ are plotted in Fig. 8.

In some cases, beyond a certain distance, the MSVs increase exponentially as a function of distance. In this case, the $\sigma_j^2$ for the most of SS and MS paths will increase rapidly even at small positive perturbations of the $\Delta R$ parameter. An increased DWF suppresses or totally eliminates the relative contribution of the corresponding scattering paths of the XAFS $\chi(k)$ amplitude. Since the DWF of the SS Zn-O scattering paths in Zn-carboxylate ligation increase exponentially, carboxylates binding to Zn will not be detected by XAFS at about 2.1–2.2 Å regardless of the sample temperature. This important exponential behavior can be described using the projected VDOS. For the SS Zn-O[2] projected VDOS is plotted for various values of $\Delta R$ (Fig. 9). Focusing our attention on the low frequency part of the normal mode spectrum, two modes at 138 and 303 cm$^{-1}$ contribute most to the corresponding $\sigma_j^2$ value of this path. Similarly to Zn-Cys case, as $\Delta R$ increases, the 138 cm$^{-1}$ mode is downshifted, reaching approximately 55 cm$^{-1}$ at $\Delta R = 0.0855$ Å with an increase of its probability amplitude. This also verifies the rapid increase of the DWF for Zn-carboxylate formations.
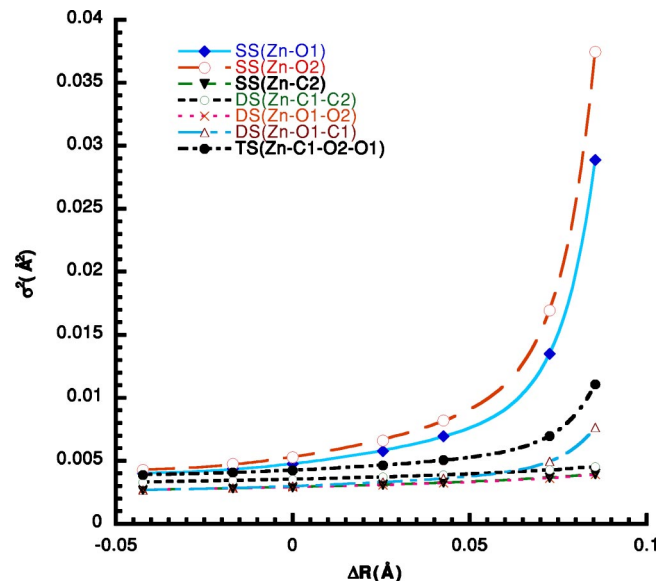


FIG. 8. (Color online) DFT-calculated $\sigma^2$ for SS and MS paths of Zn-carboxylate model at 150 K. Lines represent second order polynomial with an exponential term where applicable.

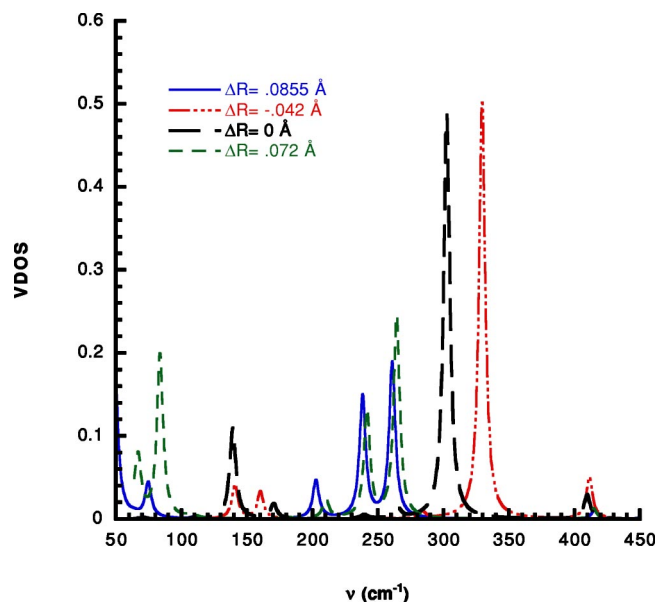FIG. 9. (Color online) Projected VDOS for the second shell SS Zn-$O^{(2)}$ path at various perturbed first shell $\Delta R_{\text{Zn-O}^{(1)}}$ distances.

### 2. Zn-Asp/Glu structures

The proposed model is tested against DFT optimized Zn-Glu and Zn-Asp structures at 150 K. For Zn-Asp and Zn-Glu Zn-O interatomic distances are observed as: 2.04, 2.07 Å for Zn-Asp and 2.03, 2.05 Å for Zn-Glu structures. Another oxygen atom that binds directly to Zn and opposite to the carboxylate ligand, located at 1.88–1.89 Å away from Zn for either configuration is also included in the reference and modeled structures. In metal-carboxylate structures under bidentate conformation, $O^{(1)}$ and $O^{(2)}$ are hardly resolved. Fitting the Zn-carboxylate model with Zn-Asp/Glu structures furnished an average distance of both oxygen shells. Therefore, Zn—O combined distance for either Asp/Glu case was found about $2.044 \pm 0.006$ Å. This distance was used for adjusting the $\Delta R$ model parameter. Calculated MSVs using the Zn-carboxylate mode versus corresponding parameters from Zn-Asp and Zn-Glu structures are summarized in Table VI. The Fourier transforms of the simulated $\chi(k)$ spectra of these reference structures with respect to the model is shown in
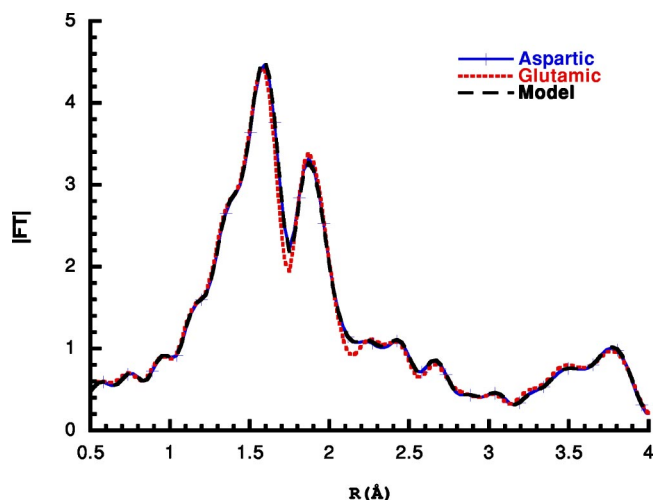


FIG. 10. (Color online) Zn absorption line for which the Fourier transform of the $\chi(k)$ XAFS spectra is calculated for DFT-optimized Zn-carboxylate (Zn-Asp, Zn-Glu) structures at 150 K. The first peak corresponds to a SS path with Zn and another oxygen bound at about 1.89 Å.

Fig. 10 and this verifies the validity of the model under homogeneous Asp or Glu ligation.

### D. Heterogeneous ligation

In most metalloprotein active sites, the metal binds to different types of amino acid residues. We refer to this as "heterogeneous ligation." Due to energy/geometry convergence difficulties the reference DFT optimized structures are of low ligand coordination number.

### 1. Zn-His/Cys structures

A hypothetical structure is built that consists of the Zn directly bound to a Cys and a His residue. For this structure the distance Zn-S was varied. This allows us to examine the chemical transferability of the histidine DWFs, since the histidine ligand is bound at the same distance in either Zn-His/Cys configuration. This distance Zn-S was locked at two different positions, 2.247 and 2.307 Å, respectively, whereas all other atoms were free to move during the geometry opti-

TABLE VI. Calculated $\sigma^2$ for the Zn-carboxylate model using information from Table II vs Zn-Asp and Zn-Glu DFT-optimized reference strcutures at 150 K for the most important scattering paths.

| Structure | -$O^{(1)}$ | -$O^{(2)}$ | -$C^{(1)}$ | -$C^{(2)}$ | -$C^{(1)}$-$C^{(2)}$ | TS[a] | -$O^{(1)}$-$O^{(2)}$ $\sigma^2(\times 10^{-3}$ Å$^2)$ | -$C^{(1)}$-$O^{(1)}$ | TS[a] | -$C^{(1)}$-$O^{(2)}$ | TS[a] | TS[b] | TS[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | 4.30 | 4.96 | 2.84 | 3.47 | 3.45 | 3.46 | 2.83 | 2.82 | 5.47 | 3.10 | 6.07 | 5.19 | 5.08 |
| Zn-Asp | 4.41 | 4.78 | 2.89 | 3.52 | 3.50 | 3.50 | 2.88 | 2.88 | 5.56 | 3.07 | 5.95 | 4.15 | 4.35 |
| Zn-Glu | 4.60 | 4.78 | 2.91 | 3.53 | 3.51 | 3.51 | 2.84 | 2.97 | 5.74 | 3.05 | 5.96 | 5.65 | 4.77 |

[a]TS of the corresponding double scattering path at the previous column.

[b]Zn—$C^{(1)}$—$O^{(1)}$—$O^{(2)}$ is four-atom triple scattering path.

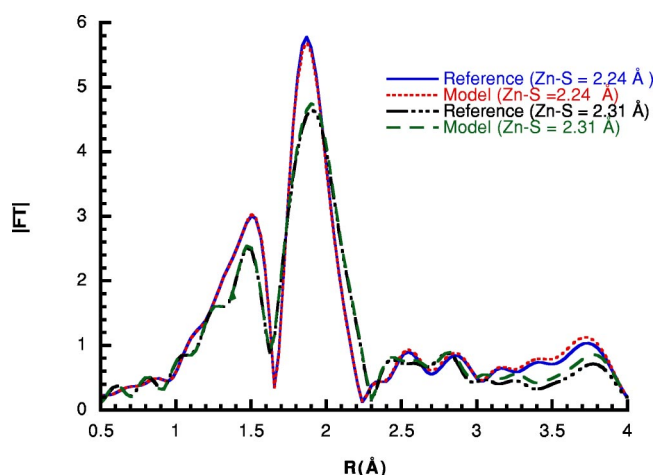[c]Zn—$C^{(1)}$—$O^{(2)}$—$O^{(1)}$ is a four-atom triple scattering path.

FIG. 11. (Color online) Zn absorption line for which the Fourier transform of the $\chi(k)$ XAFS spectra is calculated for DFT-optimized Zn-cysteine-histidine at 150 K for two distinct Zn—S distances.
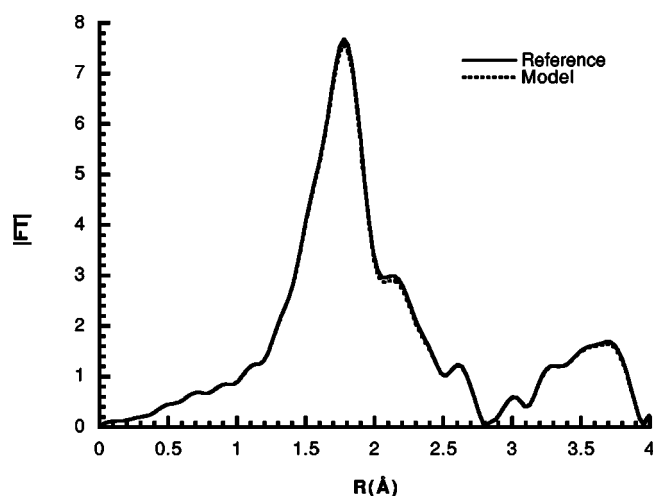


FIG. 12. Zn absorption line for which the Fourier transform of the $\chi(k)$ XAFS spectra is calculated for DFT-optimized Zn-Asp-Glu structure at 150 K.

mization procedure [$R_{\text{Zn-N}^{(1)}}=1.90$ Å for either Zn-His/Cys configuration].

Obtaining a nonimaginary normal mode frequency spectrum for this reference structure was not an easy task. For 2.247 Å Zn-S distance, one imaginary frequencies was observed at $i35.38$ cm$^{-1}$ whereas for 2.307 Å distance an imaginary frequency at $i37.44$ cm$^{-1}$ was also observed. In either configuration these small imaginary frequencies are due to geometric instabilities of the histidine tail, and they do not appear to alter the DWF of any of the SS and MS paths.

At 150 K MSVs obtained using the new Zn-imidazole model (Table I) versus the Zn-His/Cys structure are as follows [$R_{\text{Zn-N}^{(1)}}=1.90$ Å]: for the first shell SS Zn-N$^{(1)}$ $\sigma_j^2$ is $2.31\times10^{-3}$ Å$^2$ vs $2.25\times10^{-3}$ Å$^2$; for the second shell SS Zn-C$^{(1,2)}$ $\sigma_j^2$ is $(4.75\pm0.05)\times10^{-3}$ Å$^2$ vs $(4.86\pm0.29)\times10^{-3}$ Å$^2$; for the DS Zn-N$^{(1)}$-N$^{(2)}$ $\sigma_j^2$ is $2.84\times10^{-3}$ Å$^2$ vs $2.97\times10^{-3}$ Å$^2$; and for the DS Zn-N$^{(1)}$-C$^{(3)}$ $\sigma_j^2$ is $3.20\times10^{-3}$ Å$^2$ vs $2.99\times10^{-3}$ Å$^2$ for the reference Zn-His/Cys structures. The corresponding Fourier transforms using DWFs from the Zn-Cys and Zn-His models and directly from the Zn-His/Cys structure can be seen in Fig. 11. Excellent agreement is achieved for all shells.

### 2. Zn-Asp/Glu structures

The reference structure consists of an Asp and a Glu acid residue directly bound to Zn in a bidentate formation. At the optimal geometry no imaginary frequencies were observed in the normal mode spectrum. The Zn-O interatomic distances of the reference structure are distributed as follows: one oxygen atom of the Glu ligand appears at 2.067 Å, two oxygen atoms, one at Asp and one of Glut at 2.70 Å, and the last oxygen of the Asp ligand at 2.10 Å. Fitting this reference simulated $\chi(k)$ XAFS spectra with spectra of Zn-S complexes of known distances revealed that in the reference structure three oxygen atoms are located at $2.063\pm0.01$ Å and one at $2.13\pm0.04$ Å. These distances were used to adjust the $\Delta R$ parameter for the oxygen paths and the carboxylate positions.

When a metal binds to more than one carboxylate ligand, ligand-ligand scattering is observed. This effect was either minimal or not at all observed for structures with histidine and/or cysteine ligands. Intraligand MS paths that will contribute to the XAFS spectrum of Zn-carboxylate structures involve double/triple scattering paths of oxygen-oxygen, and oxygen-carbon atoms where the first/second atom belongs to the first/second ligand. Also the triple scattering paths defined as metal-scatterer-metal-scatterer-metal of either same or different ligands are also observed. In case of oxygen-oxygen intraligand scattering, the oxygen-metal-oxygen angle is about 130°. In such cases the MSV of these DS and TS paths will be nearly equal to the sum of the corresponding SS MSVs, i.e., $\sigma_{\text{O}^{(a)}\text{-Zn-O}^{(b)}}^2 \simeq \sigma_{\text{Zn-O}^{(a)}}^2 + \sigma_{\text{Zn-O}^{(b)}}^2$.

The Fourier transforms of the spectra of the reference Zn-Asp/Glu structure and also that using the modeled DWFs are shown in Fig. 12. Similarly to the previous paragraphs, spectra obtained using modeled DWFs highly agree with the corresponding DFT reference at all shells.

### IV. CONCLUSION

Various one- and two-parameter models have been used to estimate all relevant SS and MS XAFS DWFs for Zn active sites that contain cysteines, histidines, and carboxylate ligands, which are typical ligands found in Zn metalloproteins. All models were tested under model environments of varying coordination number and both homogeneous and heterogeneous ligation, which provided acceptable accuracy for XAFS data analysis. We also describe the behavior of metal-ligand binding as a function of shell interatomic distance variation and angle.

Our next step involves applying the two parameter model to carboxylates ligands such glutamic and aspartic acid residues in which substantial MS scattering is expected. This work strongly suggests that the vibrational DWF problem in XAFS analysis is soluble in practice, and it provides practical results that can be used by experimentalists in biological XAFS.

[1] E.A. Stern, in *X-Ray Absorption*, edited by D.C. Koningsberger and R. Prins (Wiley, New York, 1988), Chap. 1.

[2] P.A. Lee and J.B. Pendry, Phys. Rev. B **11**, 2795 (1975).

[3] J.J. Rehr, R.C. Albers, and S.I. Zabinsky, Phys. Rev. Lett. **69**, 3397 (1992).

[4] A. Filipponi and A. Di Cicco, TASK Q. **B4**, 575669 (2000).

[5] P. Hohenberg and W. Kohn, Phys. Rev. **136**, 864 (1964).

[6] W. Kohn and L.J. Sham, Phys. Rev. **A140**, 1133 (1965).

[7] R.G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).

[8] B.J. Clark, T. Frost, and M.A. Russell, *UV Spectroscopy* (Kluwer Academic, Dordrecht, 1993).

[9] M.J.K. Thomas and D.J. Ando, *Ultraviolet and Visible Spectroscopy: Analytical Chemistry by Open Learning* (Wiley, New York, 1997).

[10] R.G. Sajfutdinov, L.I. Larina, T.I. Vakulskaya, and M.G. Voronkov, *Electron Paramagnetic Resonance in Biochemistry and Medicine* (Kluwer Academic, Dordrecht, 2001).

[11] E.A. Stern, Phys. Rev. B **48**, 9825 (1993).

[12] S.I. Zabinsky, J.J. Rehr, A. Ankudinov, R.C. Albers, and M.J. Eller, Phys. Rev. B **52**, 2995 (1995).

[13] J.M. Castagnetto, S.W. Hennessy, V.A. Roberts, E.D. Getzoff, J.A. Tainer, and M.E. Pique, Nucleic Acids Res. **30**, 379 (2002).

[14] K.N. Degtyarenko, A.C.T. North, and J.B.C. Findlay, Protein Eng. **10**, 183 (1997).

[15] N. Dimakis and G. Bunker, J. Synchrotron Radiat. **8**, 297 (2001).

[16] N. Dimakis and G. Bunker, J. Synchrotron Radiat. **6**, 266 (1999).

[17] A.V. Poiarkova and J.J. Rehr, Phys. Rev. B **59**, 948 (1999).

[18] P.W. Loeffen and R.F. Pettifer, Phys. Rev. Lett. **76**, 636 (1996).

[19] P.W. Loeffen, P.F. Pettifer, and J. Tomkinson, Chem. Phys. **208**, 403 (1996).

[20] N. Dimakis and G. Bunker, Phys. Rev. B **58**, 2467 (1998).

[21] N. Dimakis and G. Bunker, Phys. Rev. B **65**, 201103 (2002).

[22] Mono/bidentate ligand forms one/two bonds with the central atom.

[23] M.M. Harding, Acta Crystallogr., Sect. D: Biol. Crystallogr. **57**, 401 (2001).

[24] Unichem/DGAUSS, Ver. 5.0, Oxford Mol. Group (2000).

[25] A.D. Becke, J. Chem. Phys. **84**, 4524 (1986).

[26] A.D. Becke, Phys. Rev. A **38**, 3098 (1988).

[27] A.D. Becke, J. Chem. Phys. **88**, 2547 (1988).

[28] J.P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[29] N. Godbout, D.R. Salahub, J. Andzelm, and E. Wimmer, Can. J. Chem. **70**, 560 (1992).

[30] E.A. Stern, M. Newville, B. Ravel, Y. Yakoby, and D. Haskel, Physica B **208&209**, 117 (1995).

[31] Depending on the functional and the basis set used, the DFT calculated normal mode frequencies are on the average 95% of the corresponding experimental value.

[32] A.L. Ankudinov, B. Ravel, J.J. Rehr, and S.D. Conradson, Phys. Rev. B **58**, 7565 (1998).