J. Synchrotron Rad. (1999). **6**, 266–267

Rapid single- and multiple-scattering EXAFS Debye–Waller factor calculations on active sites of metalloproteins

Nicholas Dimakis,[a]* M-Ali Al-Akhras[a] and Grant Bunker[b]

[a]Illinois Institute of Technology, Chicago, IL 60616 USA, and [b]Jordan University of Science and Technology, Irbid 22110 Jordan. E-mail: dimakis@biosgo2.bio.aps.anl.gov

This paper describes recent results using our approach to calculating self-consistently single (SS) and multiple-scattering (MS) Debye-Waller factors (DWF) on active sites of metalloproteins. The calculation of MS DWF, together with the Feff7 program allows us to simulate *ab-initio* EXAFS spectra for a given temperature systems with no adjustable parameters. In our latest report (Dimakis N., and Bunker G., 1998) we calculate, using density functional and semi-empirical approaches, the SS and MS DWF for small molecules and compared them to Raman, infrared and EXAFS spectra. In this report calculation of DWFs is done for tetrahedral Zn imidazole, a complex containing thirty two atoms that is similar in certain respects to active sites of many metalloproteins. *Ab-initio* calculation, although it is a more accurate and reliable scheme, it is not at present practical on desktop computers; computation times are weeks. Therefore as an alternative we have tried the semiempirical MNDO Hamiltonian, which is at least three orders of magnitude faster than *ab-initio*, and can be expected to be of reasonable accuracy because it is parameterized for organic compounds. Our approaches take advantage of commercially available molecular orbital programs. We have written additional programs which, using normal mode calculations, calculate the MS paths, and transparently interface with Feff7 to produce the EXAFS spectra. Results are in very good agreement with experimental data tested.

Keywords: EXAFS; Data Analysis; Debye-Waller factor.

1. Introduction

The XAFS DWF is an essential term appearing in the harmonic approximation as $e^{-2k^2\sigma^2}$ in the XAFS $\chi(k)$ equation, accounting for the structural and thermal disorder of a given sample (Stern, E.A., Koningsberger, D.C. & Prins, R. 1988, Lee, P.A. & Pendry, J.B. 1975). The parameter σ^2 is the Mean Square Variation (MSV) of a given scattering path. The DWF is a k^2-dependent term; its importance is enhanced as k is increased. Therefore for $k < 3$-4Å^{-1} the effect of this factor in the $\chi(k)$ is minimal and often can be ignored.

In order to perform a quantitative agreement with EXAFS spectra for the purpose of fitting the SS and in most cases MS DWFs must be known (Rehr, J.J., Albers, R.C. & Zabinsky, S.I. 1992, Rehr, J.J., De Leon, J., Zabinsky, S.I. & Albers, R.C. 1991). Importance of the relevant MS paths and implications of the estimation of these factors using either Einstein and/or correlated Debye-model is discussed elsewere (Dimakis N. & Bunker G., 1998).

In our last report (Dimakis N. & Bunker G., 1998) we presented the observations that *ab-initio* and in some cases semi-empirical MNDO (Dewar, M.J.S. & Thiel W., 1977) can be used to calculate self-consistently the normal mode eigenfrequencies and eigenvectors for a given sample. Although the *ab-initio* Density-Functional Theory (DFT) method (Parr, R.G. & Yang W. 1989) is far more accurate than any semi-empirical Hamiltonian, its required CPU time scales as N^3 where N is the number of atoms in the molecule, rendering such calculations impractically slow (on present desktop machines) for large biomolecules like active sites of metalloproteins. Since the semi-empirical MNDO is parameterized for organic compounds, including those containing rings it tends to provide accurate normal mode frequencies.

The Zn tetrahedral imidazole, shown in Fig.1, is an analogue of an active site of a metalloprotein that, due to the fact that the angle between the central absorber the first and last shell atoms is approximately 10°, forward MS scattering dominates in the second and third shells (Bunker G., Stern, E.A., Blankenship, R.E. & Parson, W.W. 1982). SS and MS DWFs for the "surviving" paths are calculated using a program written by the authors. The calculation is consistent with the data within the experimental errors.

Figure 1
Zn Tetrahedral imidazole.

2. Methods

The Zn tetrahedral imidazole molecule was built using both the CAChe Molecular *Editor* Ver. 4.0 and also Unichem Molecular Editor Ver. 4.1, both from Oxford Molecular Group. Both of these editors have built-in fragment libraries so a user can drag a ring instead of building everything from scratch. The structure was self-consistently optimized by minimizing the conformational energy using MNDO. A dielectric sphere surrounding the complex mimics the liquid (imidazole) phase of the sample. We found the dielectric constant can be between 20-30 with insignificant change in the σ^2s themselves. The radius of the sphere should be just enough to cover the whole molecule. Normal mode analysis was performed using this semiempirical MNDO method. It should be mentioned that although MNDO provide a good guess for normal mode frequencies, radial distances and angles may not be quite satisfactory. This is because semiempirical Hamiltonians are parameterized using infrared and Raman frequencies than interatomic distances for organic compounds. Therefore atomic coordinates in *feff.inp* file were taken from a DFT, double zeta basis set.

A program written by the authors reads normal mode eigenfrequencies and eigenvectors and calculates single scattering, and up to eight-legged path multiple scattering Debye-Waller factors. Four atom cluster calculations are also included.

3. Results and Discussion

The Zn tetrahedral imidazole complex, at 120°K, was used for reasons discussed before, as a test analog of a Zn metalloprotein active site. Since experimental data available were fluorescence data, a fill-gas correction was performed. Two experimental scans have been used here to provide a rough measure of the noise. The relative DWF between these two scans for the first shell is about

$7 \times 10^{-4} \text{Å}^2$. Due to monochromator calibration a shift in energy of -8.3 eV in theoretical $\chi(k)$ was performed.

The background was removed using cubic spline method. The experimental $\chi(k)$ was also McMaster normalized [1]. This correction is necessary when theoretical data are compared with experimental ones.

Fourier Transforms of the $\chi(k)$ data were performed to filter the high frequency noise contribution in our EXAFS spectrum. The relevant noise contribution in the pass-band is estimated by the scan to scan fluctuation. Three shells can be seen here. Double and triple scattering will be present in the last two shells whereas the first shell is dominated by SS. By applying the Ratio Method for the first shell we were able to determine the ${S_0}^2$ and the radial distance r_{Zn-N}. This was found to be approximately 2Å. DWFs are averaged over three of the four rings. The fourth ring is not included in the calculation of the σ^2s since its radial distance is much larger than this of the rest rings. The relative error on the calculated DWFs varies from $4 \times 10^{-4} \text{Å}^2$ to $3 \times 10^{-5} \text{Å}^2$. The next step was to input the DWF calculated by our program back to *list.dat* file and re-execute feff7. First shell SS DWF was found to be $2.23 \times 10^{-3} \text{Å}^2$ and the experimental one approximately $2.7 \times 10^{-3} \text{Å}^2 (\pm 7 \times 10^{-4} \text{Å}^2)$. Since MS scattering dominated in the higher shells, traditional fitting routines will fail here. In order to determine the other radial distances and angles we used Unichem Molecular builder. Finally the Fourier transform, real and imaginary part, of mean experimental, mean plus standard deviation and MNDO is given at Fig.2(a) and (b).

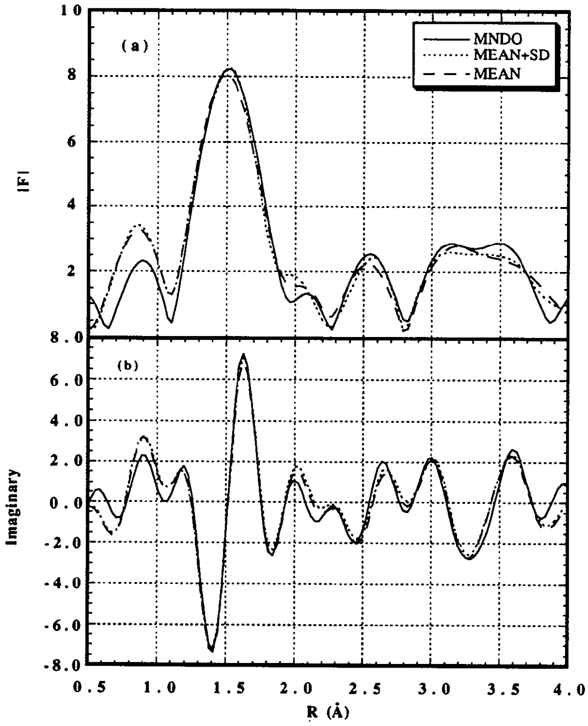

Figure 2

Fourier Transform of Mean Experimental *vs.* Mean plus Standard Deviation, vs MNDO for Zn Tetrahedral imidazole. *(a)* Magnitude and *(b)* Imaginary Part.

Results using MNDO method lie within the experimental error bars.

4. Conclusion

The semiempirical MNDO Hamiltonian provides a good and fast way to calculate single and multiple-scattering EXAFS Debye-Waller factors for active sites of metalloproteins. More accurate and reliable analysis can be obtained by means of Density Functional methods under the double zeta basis sets, but at a cost of at least three orders of magnitude of CPU time. Given the pace of advancements in computer technology it is clear that even full scale DFT calculations on XAFS-scale clusters will be practical on desktop computers in coming years. This will help to solve the parameter to data ratio problem in multiple scattering fitting of the XAFS of biomolecules.

This project is supported by NIH-RR08630. We are grateful to Dr. Ke Zhang for Zn imidazole data.

References

Bunker, G., Stern, E.A., Blankenship, R.E., and Parson, W.W. (1982). *Biophys. J.* **37**, 539-551.

Dewar, M.J.S. & Thiel W. (1977). *J. Amer. Chem. Soc.* **99**, 4899–4906.

Dimakis N. & Bunker G. (1998). *Phys. Rev.* **B58**, 2467-2474.

Lee, P.A. & Pendry, J.B. (1975). *Phys. Rev.* **B11**, 2795-2811.

Parr, R.G. & Yang W. (1989). *Density-Functional Theory of Atoms and Molecules*, ed., Oxford Univ., New York.

Rehr, J.J., Albers, R.C. & Zabinsky, S.I. (1992). *Phys. Rev. Lett.* **69**, 3397–3400.

Rehr, J.J., De Leon, J., Zabinsky, S.I. & Albers, R.C. (1991). *J. Amer. Chem. Soc.* **113**, 5135–5140.

Stern, E.A., Koningsberger, D.C. & Prins, R. (1988). eds., Wiley & Sons, New York, ch. 1.

(*Received 10 August 1998; accepted 11 January 1999*)

---

[1] Experimental data were divided by $\Delta\mu^{McMaster}(k)/\Delta\mu^{McMaster}(k=0)$