

CRYSTALFF – from crystallography to EXAFS multiple-scattering calculations with FEFF

Karine Provost,^{*} Frédéric Champloy and Alain Michalowicz

Groupe de Physique des Milieux Denses, Université Paris XII – Val de Marne, 61 av. du Général de Gaulle, 94010 Creteil Cedex, France. E-mail: provost@univ-paris12.fr

CRYSTALFF is an alternative utility to ATOMS containing most of the features of this standard program for converting crystallographic data to FEFF input. In addition, it offers an interface with molecular modelling programs via the PDB format and new coordination sphere analysis options.

Keywords: EXAFS; FEFF input; crystalline structures; molecular models.

1. Introduction

For coordination and organometallic chemistry XAS studies, it is often necessary to combine crystal structures, molecular modelling and *ab initio* XAS calculations. Indeed, a complete XAS study of an inorganic compound family requires multiple-scattering calculations on many structures, only part of them being solved by crystallography. Among the three standard multiple-scattering EXAFS codes, EXCURVE (Binsted *et al.*, 1991; Gurman *et al.*, 1986), FEFF (Rehr *et al.*, 1992) and GNXAS (Filipponi & DiCicco, 1995), only the first provides a molecular modelling interface. Although Ravel's ATOMS code (Ravel, 1999) provides a very convenient tool for the study of mineral compounds, it is less adapted to molecular structures. Using this program, we came across two types of difficulties: (i) interfacing molecular modelling outputs with FEFF inputs; (ii) superimposing different structures in order to compare their features of paths contributions. In order to overcome these difficulties, we wrote some new software, CRYSTALFF, which includes most of the crystallographic ATOMS features and new functions for an easier molecular modelling–FEFF interface. Other XAFS programs referenced in the IXS catalogue (http://ixs.itt.edu/catalog/XAFS_Programs) such as AUTOFIT, LASE and WINXAS provide an FEFF interface. However, their crystallographic and molecular modelling facilities can be completed by the use of CRYSTALFF.

2. General program organisation

The program is organised into six modules: crystal or molecular models input or generation; molecular models output; coordination sphere calculations; coordination sphere input or output; FEFF input files generation; bonds analysis. These modules are accessible interactively by hierarchical menus. In CRYSTALFF, we distinguish the model, containing the whole structure expressed in atomic cartesian coordinates, and the coordination sphere containing only the atoms inside a sphere defined by its central atom and radius.

2.1 Crystal or molecular models input or generation

Models can either be read in the Brookhaven Protein Data Bank (PDB) format (Bernstein *et al.*, 1977) available as input–output of

most molecular modelling programs, in CRYSTALFF format (see §2.2) or in free format, or calculated from crystallographic data (cell parameters, space group and the content of the asymmetric unit expressed in fractional coordinates). The latter option provides identical results to those of ATOMS. The number of generated cells in each direction is user-defined. The number of input and generated atoms is only limited by the free disk space since all information is kept in scratch files. Each atom is then characterised by its index in the asymmetric unit, the applied symmetry operation index and the three applied translations. The program works for any set of symmetries including the standard 230 space groups. The symmetry operations are entered as text files containing the equivalent positions coordinates as defined in *International Tables for Crystallography* (1996). The advantage of using external text files is that they allow a better control on the origin choice and axis definition for non-conventional space groups, as often encountered in old inorganic crystallographic data. For PDB inputs, it is also possible to apply symmetry operations defined by the user.

2.2 Model outputs

Models read or generated can be exported in two formats: CRYSTALFF format (text) contains all information concerning the model generation (cell parameters, space group and, for each atom, its index in the asymmetric unit and the symmetry operation used); the PDB format does not contain crystallographic information; it is suitable as an interface with molecular modelling programs.

2.3 Coordination sphere calculations

This is the central part of the program containing three main functions: coordination sphere generation, reorientation and analysis.

(i) *Coordination sphere generation.* The simplest way to generate a coordination sphere is to define its central atom and its radius. In the case of crystal structures generation, the program checks whether the sphere radius is not too large according to the model size: the distances between the central atom and all the faces of the simulated crystal are calculated; when atoms of the coordination sphere may be missing (not included in the crystal model), a warning message is displayed.

(ii) *Coordination sphere reorientation.* The comparison of two coordination spheres calculated from crystallographic data is frequently difficult because the molecules have different orientations in the crystal cell. The orientation function of CRYSTALFF allows one to modify the sphere orientation as required, either manually, by defining two new axes on selected atoms, or by fitting the molecular orientation to a reference structure. In the fitting option, the program proceeds in two steps: a rough estimation of the orientation matrix, and a refinement of this matrix. The orientation is defined by the Euler angles (ψ , θ , ϕ). These angles can be calculated on the basis of the central atom and two neighbours in each coordination sphere. Two sets of N atoms are selected by the user, in the coordination sphere and in the reference structure, respectively. The program checks the consistency of the selected atoms indexes in both selections. A rough orientation is obtained by averaging Euler angles calculated on all atom pairs. The Euler angles are then refined by a standard minimization of the quantity:

$$\chi^2 = \sum_{\text{selected atoms}} \left[(x - x_{\text{ref}})^2 + (y - y_{\text{ref}})^2 + (z - z_{\text{ref}})^2 \right] \quad (1)$$

where (x , y , z) (functions of the Euler angles) and (x_{ref} , y_{ref} , z_{ref}) are the Cartesian coordinates of the selected atoms. The CRYSTALFF

code is not machine specific, and thus does not include any graphic interface. Nevertheless, there is a quantitative control of the fit *via* the distances between the atoms of the refined sphere and the reference, the final χ^2 and its derivatives. For a graphical control, we recommend the use of the PDB output and external programs.

(iii) *Coordination sphere analysis*. *CRYSTALFF* is not only designed for creating *FEFF* input files, but also offers tools for structural analysis of the coordination sphere in relation with XAFS. The decomposition of an EXAFS signal is based on the notion of shells. This is the reason why we provide a shell analysis tool. Two atoms of the same type are considered in the same shell if their distances to the centre differ from less than a 'thickness' parameter. *CRYSTALFF* provides statistics for each shell: number of atoms and minimum, maximum and average distance. If the number of shells is less than eight, the shell index can be used as potential type in *FEFF*. The alignment analysis provides complementary information for multiple-scattering *FEFF* analysis. It is well known that alignment of atoms enhances considerably the multiple-scattering amplitudes. *CRYSTALFF* calculates all the scattering angles of the coordination sphere. The scattering paths corresponding to angles equal to 0° or 180° within a tolerance are listed.

The program being fully interactive, it is possible to perform successive coordination sphere generations and analysis from the same model. This option is particularly useful when there are many different types of scatterers in the structure.

2.4 Coordination sphere input–output

Coordination spheres can be input or output in PDB or *CRYSTALFF* formats.

2.5 FEFF input files generation

FEFF input files are saved interactively under a user defined name, in the *FEFF6–7* format, with the atomic potential types defined from the coordination shells calculations. The only *FEFF* cards generated by *CRYSTALFF* are TITLE, POTENTIALS, CONTROL and ATOMS.

2.6 Bonds analysis

The interface between *CRYSTALFF* and molecular modelling programs is greatly helped by the possibility to draw the atom connections. In order to improve this interactivity between molecular modelling programs and *CRYSTALFF*, the program uses a bond-types library, based on the distances described in *International Tables of Crystallography* (1996). Each bond type is defined by the atomic numbers of the two atoms and a distance range. At any time, the user can modify interactively this library. Connections calculations can be performed either on the model or on the coordination sphere and can be saved either when creating a complete PDB output or in a separate file. The use of the connections tables is planned in further versions.

3. Example: study of cobalt(III) complexes related to nitrile hydratase metallic site

We chose an example in which the interface between crystallographic data, molecular modelling and *FEFF* is essential. This example deals with cobalt containing nitrile hydratases, enzymes which catalyse the hydration of nitriles to amides. It is possible to model the active site of these enzymes by a square planar

cobalt(III) complex with an N_2S_2 di-*N*-carboxamidodithiolato-tetradentate ligand (Heinrich *et al.*, 2001).

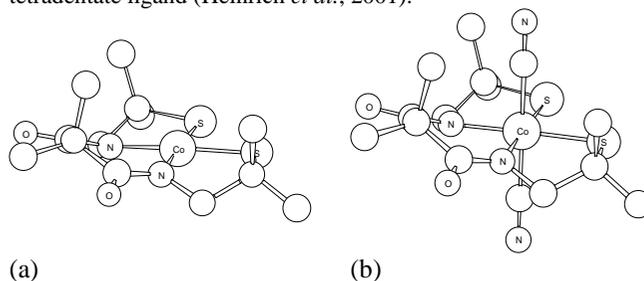


Figure 1

Crystallographic structure of the square planar complex (a) and *Chem3D* final model of the dicyano hexacoordinate complex (b).

The properties of the complex are dramatically modified by the addition of axial ligands. Hexacoordinated species correspond to the active form of the enzyme. We studied the structural modifications induced by the addition of two axial cyanide ligands (Fig. 1). Such modifications in the tetradentate ligand could not be modelled without using a molecular modelling software and thus the interface between crystallographic data, molecular modelling and *FEFF*.

The crystal structure could be solved only in the case of the square planar complex. In a first step, the *FEFF* model of the square planar complex was compared with the experimental EXAFS curve (Fig. 2), using the crystallographic coordinates converted to an *FEFF* input file by *CRYSTALFF*.

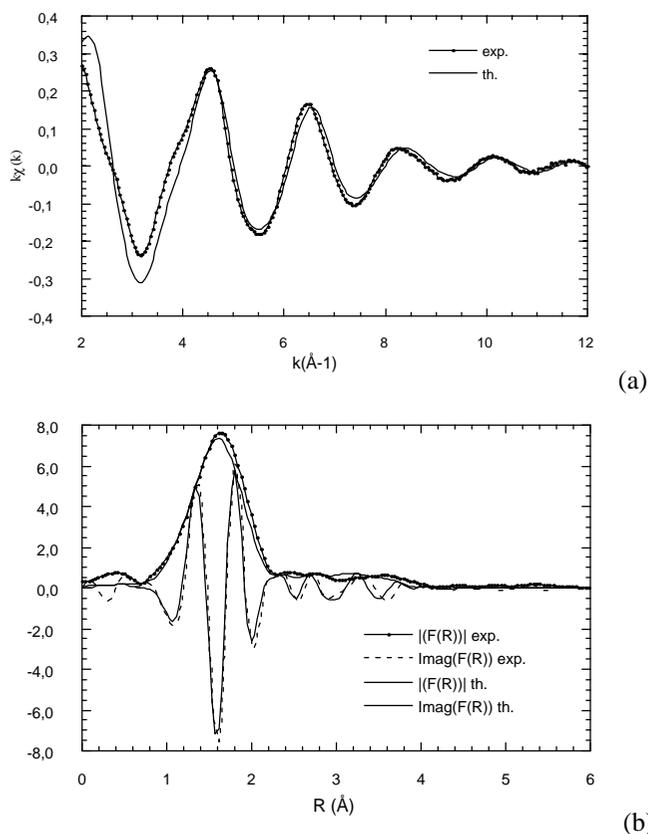


Figure 2

Comparison between experimental and *FEFF* simulated spectra of the square planar complex based on its structure: a) EXAFS spectra, b) Fourier transform modules and imaginary parts.

In order to build the first hexacoordinated model, we converted the crystal structure into a PDB file. Two cyanide ligands were added using *Chem3D*, at conventional distances and orientation, without any modification of the equatorial ligand geometry. Since the resulting *FEFF* calculation could not fit the experimental data (Fig. 3), we fitted the first coordination shell together with the multiple-scattering signals of the cyanide second neighbours. We used the fitting program included in 'Exafs pour le Mac' (Michalowicz, 1991, 1997), but any other EXAFS fitting program could be used (IXS). Fitting ligand distances can considerably modify the multiple-scattering signals. It was thus necessary to validate the fit by a final modelling of the whole structure. This implied new modifications in *Chem3D*, and *FEFF* input file generation via the PDB format of *CRYSTALFF*. The EXAFS fitting conditions are detailed elsewhere (Heinrich *et al.*, 2001). As shown in Figs. 3 and 4, the resulting *FEFF* model is considerably improved.

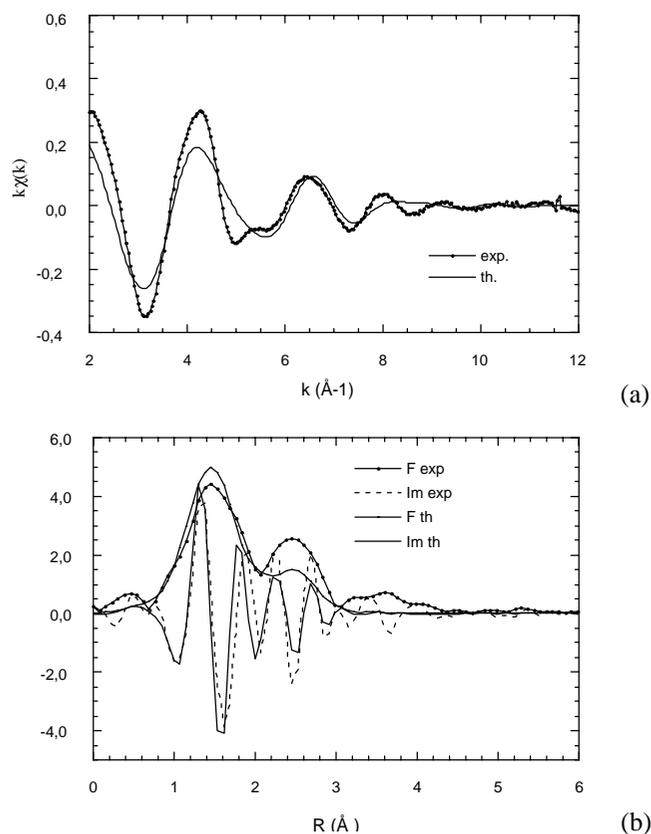


Figure 3 Comparison between experimental and *FEFF* simulated spectra of the dicyano complex based on *Chem3D* model before refinement: a) EXAFS spectra, b) Fourier transform modules and imaginary parts. The Co–N and Co–S distances are kept to 1.88 Å and 2.13 Å, as in the square planar complex.

Such modelization implies an iterative procedure using many times molecular modelling and *FEFF* calculations alternatively. The interface provided by *CRYSTALFF* significantly simplified this process. Other examples can be found in the literature (Curis *et al.*, 2001; Van Koningsbruggen *et al.*, 2000).

4. Program availability and future development

The program was written in *FORTRAN77* with the *FORTRAN90* wide-format extension on an iMac Apple computer. The code is standard and can be exported on any computer with a few modifications specified in the manual. On a Macintosh, the program runs with 400 ko RAM, but requires disk space for temporary files depending on the structure size. The pack contains the source code, a compiled version for PPC Macintosh, libraries (some of the 230 space groups and the bond library) and some examples. A pack for MSWindow-driven computers is also available.

The package is free of charge and can be obtained as an e-mail attached document under request to the authors. The last version should soon be available on the computer programs web page of LURE, the French synchrotron radiation facility (http://www.lure.u-psud.fr/log_scient.htm).

Further versions will include tools for *FEFF* results analysis involving the use of the structure, and input of a model in CIF format. Any other modification (new input/output format, new options) suggested by users are welcome.

The authors would like to thank Dr J. Moscovici and Dr E. Fonda for intensive critical use of the program leading to very useful suggestions. We also thank Dr Heinrich, Dr Li and Professor Chottard who allowed us to use the cited example.

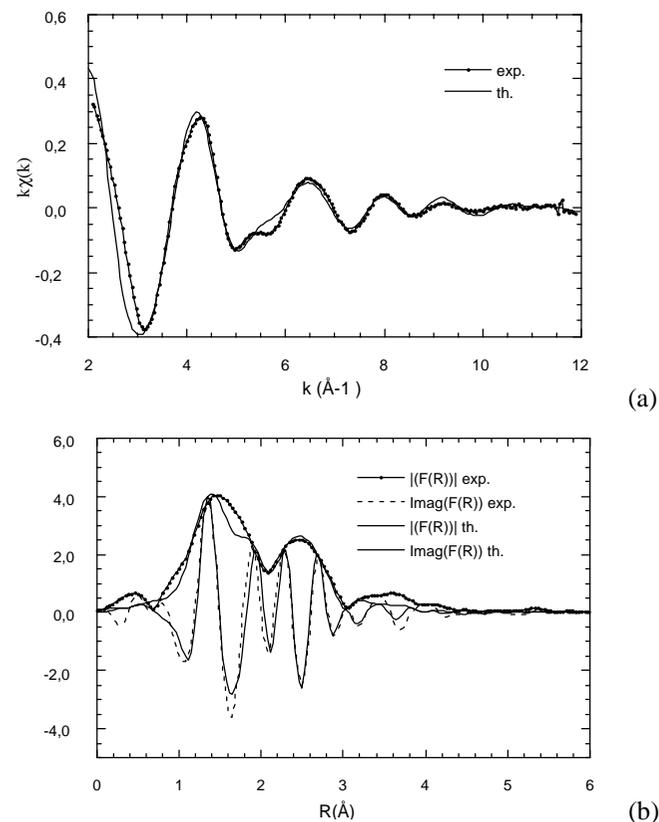


Figure 4 Comparison between experimental and *FEFF* simulated spectra of the dicyano complex based on the *Chem3D* model after refinement: a) EXAFS spectra, b) Fourier transform modules and imaginary parts. The Co–N and Co–S are fitted to 1.95 Å and 2.22 Å, leading to considerable structure changes in the equatorial ligand.

References

- Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
- Binsted, N., Campbell, J. W., Gurman, S. J. & Stephenson, P. C. (1991). *EXCRV92*. EPSRC Daresbury Laboratory, Warrington, UK.
- Curis, E., Provost, K., Bouvet, D., Nicolis, I., Crauste-Manciet, S., Brossard, D. & Bénazeth, S. (2001). *J. Synchrotron Rad.* **8**, 716–718.
- Filliponi, A. & Di Cicco, A. (1995). *Phys. Rev. B*, **52**, 15135–15149.
- Gurman, S. J., Binsted, N. & Ross, I. (1986). *J. Phys. C*, **19**, 1845–1861.
- Heinrich, L., Li, Y., Vaissermann, J., Provost, K., Michalowicz, A. & Chottard, J. C. (2001). *Inorg. Chem. Acta*. In the press.
- International Tables for Crystallography* (1996). Vol. A, 4th rev. ed., edited by Th. Hahn. Dordrecht: Kluwer.
- Michalowicz, A. (1991). *Logiciels pour la Chimie*, p. 102. Paris: Société Française de Chimie.
- Michalowicz, A. (1997). *J. Phys. IV*, **7**, 235–236 (http://ixs.itt.edu/catalog/XAFS_Programs).
- Ravel, B. (1999). <http://leonardo.phys.washington.edu/~ravel/atoms/>.
- Rehr, J. J., Zabinsky, S. I. & Albers, R. C. (1992). *Phys. Rev. Lett.* **69**, 3397–3400 (<http://leonardo.phys.washington.edu/feff/>).
- Van Koningsbruggen, P. J., Garcia, Y., Kahn, O., Fournès, L., Kooijman, H., Spek, A. L., Haasnoot, J. G., Moscovici, J., Provost, K., Michalowicz, A., Renz, F. & Gülich, P. (2000). *Inorg. Chem.* **39**, 1891–1900.