EXCURVE Tutorial

Martin C. Feiters

Dept. Synthetic Organic Chemistry, Institute for Molecules and Materials, Faculty of Science, Radboud University Nijmegen

Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

tel. +.31.24.3652016, Email m.feiters@science.ru.nl

Ref: Feiters & Meyer-Klaucke (2013) BioXAS

Chapter 6

Practical Approaches to Biological Inorganic Chemistry, 1st Edition. http://dx.doi.org/10.1016/B978-0-444-56351-4.00005-1.

X-ray Absorption Spectroscopy in Biology (BioXAS)

Practical Approaches to Biological Inorganic Chemistry (R. R. Crichton and R. Louro, Editors; 2013)

Martin C. Feiters [a] and Wolfram Meyer-Klaucke [b]

[a] *Department of Organic Chemistry, Institute for Molecules and Materials, Faculty of Science, Radboud University Nijmegen, Heyendaalseweg 135, Nijmegen, The Netherlands,* [b] *Faculty of Science, Department of Chemistry, University of Paderborn, Warburger Straße 100, Paderborn, Germany*

Chapter Outline

Why EXCURVE ?

> Pros:

- Well suited for users from biology (non-experts)
- Comes with ligand data-base (amino acids)
- Allows for automation of analysis / scripts (towards end of presentation)
- Small number of parameters to adjust prior start of refinement (defaults work well)
- k-space refinement (no filtering, considers the noise in your EXAFS)

> Cons:

- GUI not very helpful (reduces functionality, not covered today)
- To beginners the commands and options are sometimes not obvious
- In case defaults are not working well no instructive comment is given on how to proceed

My personal experience with EXCURVE

- Post-doc at the SRS Daresbury Laboratory (UK) in 1984-86 when the programme was being developed
- Adequate theory: 'curved wave of the electron' more accurate than 'plane wave approximation'

$$\chi(k) = {}_j\Sigma \underbrace{N_j . S_i(k) . F_j(k)}_{\text{amplitude}} . \underbrace{e^{-2k^2\sigma_j^2} . e^{-2r_j/\lambda}}_{\text{damping}} . \underbrace{\sin(2kr_j + \phi_j(k))/kr^2_j}_{\text{oscillatory}}$$

- Multiple scattering within units adequately treated
- Restrained refinement effectively reduces the number of parameters allowed to float

S. J. Gurman, N. Binsted, I. Ross, A rapid, exact curved-wave theory for EXAFS Calculations. J Phys C. Solid State Phys 1984; 17: 143 -151;
S. J. Gurman, N. Binsted, and I. Ross, A rapid, exact, curved-wave theory for EXAFS calculations: II. The multiple-scattering contributions. J Phys C. Solid State Phys 1986; 19: 1845-1861

Current status availability DL_EXCURVE

- On:http://www.cse.clrc.ac.uk/cmg/EXCURV/ the option:

Request a download is temporarily out of order, since the 'CCP3 project' currently has no funding

- Use **Contact us** (Email: ccp3@stfc.ac.uk) and ask Barry Searle: barry.searle@stfc.ac.uk
- Official site for DL-Excurv: http://www.cse.scitech.ac.uk/cmg/EXCURV/
- EXCURVE was developed for UNIX systems ('VAX')
- The graphics in the standard version Daresbury Laboratory Visualize package **DLV** can be made to work on Linux or cygwin
- Windows-compatible version provided under license by Norman Binsted up to 2011

Structure of EXCURVE

- Command lines
- Read experiments, parameters, phaseshifts
- Define atom types, calculate potentials, and then phaseshifts
- ‘Macro’ for set of commands
- Menus for theory and plot options
- Fourier filtering of shells
- Define models in shells, combine shells in units with polar coordinates
- Multiple scattering in units
- Constrained vs. Restrained refinement

To start the program and read in experiment files

ENTER COMMAND:

Read Exp (this reads in the experimental spectrum. The command **Read** on its own is used to input experiment, a parameter file and phaseshift files. **Read PAR** is used to input just parameters, and **Read PH** for phaseshifts only)

Filename for Experiment 1?

f01000.fin (this is the name of the experiment file of Zn K edge EXAFS of [Zn(imidazole)4] (ClO4)2 which has its background removed)

Experimental spectrum in f01000.fin

Point frequency [1] ?

1 (this reads in every point of the spectrum. A value of 2 would read in every 2nd point)

Column combination [12] ?

12 (this means that the program will read in energy (in eV) from column 1 of the background subtracted file, and Chi from column 2. Data taken/processed at other synchrotrons /other programmes may require a different column combination, such as 32 or 23)

Edge ? [CU K]

Zn K (The edge that is being analysed. Can be K, L3, L2, L1, e.g. Rb L3 or W L1)

Sequence number in polarization set [0]

0 (this is used when reading in a series of spectra of single crystals at different angles. The default 0 should be used for all other cases)

Number of clusters for this experiment [1]

1 (used when defining different structural clusters around the central atom for multiple scattering calculations. If you are using only single scattering, or there is only one type of cluster in the sample use the default of 1)

Plane wave/single scattering approximation of EXAFS

χ is a sum of j (resolved) shells of backscatters of type i

$$\chi(k) = {}_j\Sigma \underbrace{N_j . S_i(k) . F_j(k)}_{\text{amplitude}} . \underbrace{e^{-2k^2\sigma_j^2} . e^{-2r_j/\lambda}}_{\text{damping}} . \underbrace{\sin(2kr_j + \phi_j(k))}_{\text{oscillatory}} / kr^2_j$$

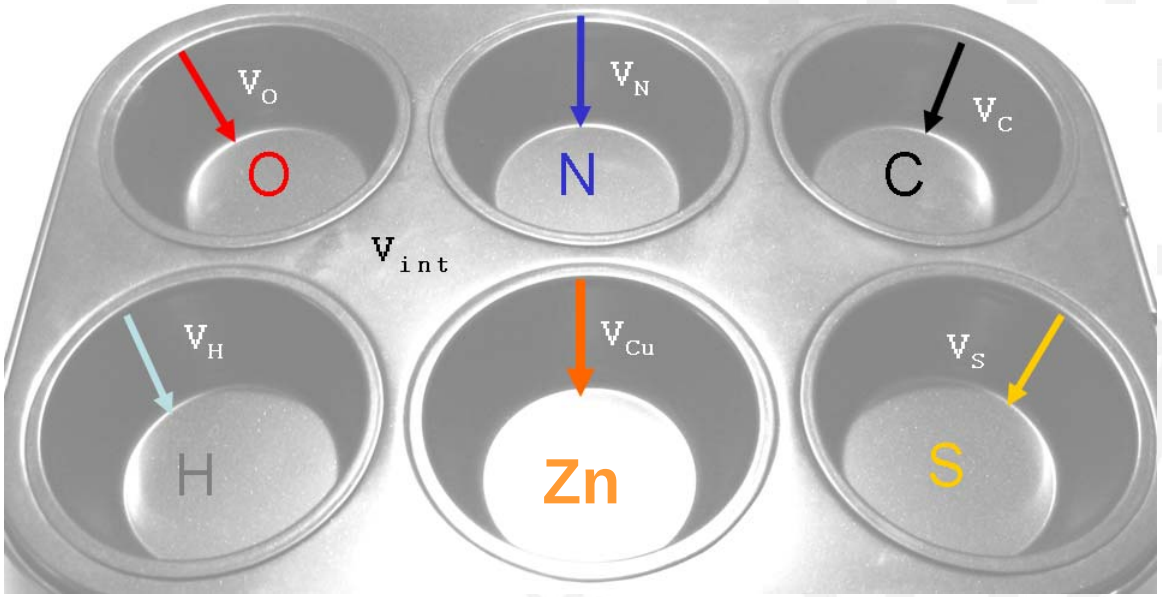
k, electron wave vector ($Å^{-1}$) defined as $k = \sqrt{\frac{2m_e(E-E_0)}{\hbar^2}}$

F_j backscattering amplitude of each of the
N_j backscattering atoms of type i with
σ_j Debye-Waller-type factor r_j distance
ϕ_j total phaseshift S_i amplitude reduction factor
λ electron mean free path $(\Delta)E_0$ threshold energy

The (bio)chemist is interested in the type of atom (choose F_j) and its number N_j and distance r_j (refine in simulation).
Unfortunately, refinement of (and correlation with !) the Debye-Waller-type factor and ΔE_0 cannot be avoided.

Calculation of Backscattering Amplitude and Phase Shift

- ‘Phase shift calculations’ in EXCURVE based on a Muffin-Tin Potential



- Backscattering amplitude F_i and phase shift ϕ_j can be accurately calculated for absorber-backscatterer
- In modern phase shift calculations, we do not adjust the *amplitude reduction factor* S_j nor the *electron mean free path* λ_j

Defining Atom Types

ENTER COMMAND:

Change ATOM1 Zn; Change ATOM2 N; C ATOM3 C; CATOM4 H; C ATOM5 O; C ATOM6 Cl; C ATOM7 Zn; C ATOM8 S (define an atom type for every element in the sample. You should include your central atoms twice as in the potential and scattering calculations an excited atom is not the same as the normal atom). Atom1 must be the central (absorber) atom. The order of the others does not matter, but if you save phaseshifts and parameter files you must read the phaseshift into Excurve in the same order.

ATOM8 changed from: (0) to S (16)

Calculating potentials

ENTER COMMAND:

CALculate POTentials

Calculating potentials for all atoms:

Enter: G (graphics), T(Terminal Output), M (Charge densities) or C (Continue)

C

Atom: 1 (Zn). Enter neighbouring atom [3 (C)]

(The general advice is to choose an atom type which most resembles the surrounding of the atom or the lightest neighbouring atom, in the case of Zn imidazole it is **N**. The program default is shown in square brackets. Press <**return**> to accept the default, or enter the required scatterer.)

Select Code For Excited Atom [1] :

No Correction (0)

1S Core Hole (K-edge) (relaxed approximation) (1)

1S Core Hole (K-edge) (Z+1 approximation) (-1)

-1 (Generally use Z+1 for 1st row transition metals and lighter and relaxed for heavier central atoms. This makes very little difference anyway, but try to be consistent.)

ZN (N) Rho0: 0.4497 Efermi: -4.087 V0: -18.228 Electrons: 30.113 (Z= 30)

Atom: 2 (N). Enter neighbouring atom [3 (C)]

(The program now asks for atoms surrounding each of the scatterers etc. In this Zn example, choose Zn as the neighbouring atom for N; choose N for C; choose C for H; choose Cl for O, and O for Cl; choose N for the 2nd Zn, and Zn for S)

S (ZN) Rho0: 0.4364 EFermi: -3.665 V0: -17.525 Electrons: 15.657 (Z= 16)

Calculating potentials and phaseshifts

(The interstitial potential V0 should be the same to within 2 eV for all the elements. If this is not the case a target V0 can be set:)

ENTER COMMAND:

Set COMmon V (This program takes an average of the potentials calculated, and sets this as a target potential – Set CONSTANT V in older versions of EXCURVE. Recalculate potentials by using the **CAL POT** routine again. The muffin tin radii of the atoms will be refined to give the target potential V0 for all the atoms.)

S (ZN) Rho0: 0.4568 EFermi: -3.662 V0: -17.950 Electrons: 15.642 (Z= 16)

Average values of V0,RHO0, FE0 : -17.960 0.5318 -2.147

(The calculated potential can now be used to calculate phaseshifts for the central atoms and all the scattering atoms.)

ENTER COMMAND:

CALculate PHaseshifts

Calculating phaseshift for all atoms.

Do you require the atomic absorption [no]?

(only required for XANES simulations – generally take the default of NO)

Enter core width (FWHM eV) [1.936]

(again, use default generally)

(Having calculated a set of phaseshifts, these can be saved to files which can be read into Excurve at a subsequent session. This saves having to recalculate every time and also ensures that the same set of phaseshifts is used in all your analyses.)

Printing phaseshifts (i)

Having calculated a set of phaseshifts, these can be saved to files which can be read into Excurve at a subsequent session. This saves having to recalculate every time and also ensures that the same set of phaseshifts is used in all your analyses.

ENTER COMMAND:

PRint PHaseshifts

Atom number 1

Filename: [exphsa1.zn] ?

Choose a suitable filename. Use the same filename with different suffixes for all the phaseshifts. (When you choose a filename for the first phaseshift this becomes the default for the rest.) E.g. enter the filename **znimsph.zn <return>**

Phase shifts printed in znimsph.zn

Atom number 2

Filename: [znimsph.n] ?

<return>

Phase shifts printed in znimsph.n

Atom number 3

Filename: [znimsph.c] ?

<return>

Phase shifts printed in nicnphs.c

Atom number 4

Filename: [znimsph.h] ?

<return>

Phase shifts printed in znimsph.h

Printing phaseshifts (ii)

Atom number 5
Filename: [znimsph.o] ?
<return>
Phase shifts printed in znimsph.o
Atom number 6
Filename: [znimsph.cl] ?
<return>
Phase shifts printed in znimsph.cl
Atom number 7
Filename: [znimsph.zn] ?
znimsph.zn2 <return> (do not overwrite the absorber atom phaseshift znimsph.zn with the backscatterer phaseshift, choose extention .zn2 to distinguish)
Phase shifts printed in znimsph.zn2
Atom number 8
Filename: [znimsph.s] ?
<return>
Phase shifts printed in znimsph.s

Phaseshifts that have been saved can be read into Excurve in a subsequent session using the command **Read PH**. (Be on guard with specific assignment problems when reading parameters from a Brookhaven Protein Data Bank (.pdb) file).

All commands given so far in macro: %macrozn.txt

```
s tit;Zn in HIV integrase;q
!Read the Zn K-edge data
r ex 1;f01000.fin;1;12;zn k;;;;
!Set up the atom types
c emin 0; c emax 3000; c emin1 0; c emax1 3000;
c atom1 zn*;c atom2 n;c atom3 c;c atom4 h;c atom5 o;c atom6 cl;c atom7 zn;c atom8 s;
!Calculate potentials
set common none;
ca pot;c;n;-1;zn;n;c;cl;o;n;zn;
!set potential to average of atom types involved
set common v;
!Calculate potentials
ca pot;c;n;-1;zn;n;c;cl;o;n;zn;
!Recalculate phaseshifts
ca ph;;;;;;;;;;;;;;;;;;;;
!Set some initial value of EF
c ef -5;
!Set up a feasible first shell
c ns 1; c a1 .001; c n1 4; c r1 2.0; c t1 2;
!Set up graphics settings like weighting and type of Fourier transform
set w 4; gs pw 4; ft t 2;
!calculate the simulation and plot it
cp;
```

Theory, Graphics and FT Parameters Sub Menus (i)

- (a) Settings menu: - The current (default) settings can be seen by entering **Set List**. Each item in the list has a Keyword and up to four Options. The current option is displayed in the final column, labeled Now. To return to the main Excurve command line type **Quit**.
- This menu controls the type of theory, ground state exchange and refinement weighting etc.
- The only setting which you would normally need to change from the default is the weighting. The default is to analyse data with k2 weighting (Option 3) which should be changed to k3 weighting (Option 4).
- To do this you can enter **Set Weight 4** or **Set Weight K3** from the Excurve command line, or enter **Set**, then **Weight 4** (or **Weight K3**), then enter **Quit** to return to the command line.
- Any other setting can be changed in a similar way, e.g. **Set THeory 1** changes from Small Atom Theory (default) to Curved Wave Theory.

Theory, Graphics and FT Parameters Sub Menus (ii)

- (b) Graphics menu: - The current (default) settings can be seen by entering **GS List** (GS, Graphics Settings). The menu works in the same way as the Settings menu and there are up to fice Options for each Keyword. To return to the main Excurve command line type **Quit**. The Spectrum default (Option 1) shows both experiment and theory in the EXAFS plots and Fourier transforms. It can be changed to show experiment only (Option 2), theory only (Option 3), or the difference (Option 4). This is also the menu where the weighting (k2 by default) needs to be adjusted to be consistent with that of the theory (k3): **GS PW 4**.
- (c) Fourier transform menu: - The current (default) settings can be seen by entering **FT List**. The menu works in the same way as the Settings and Graphics menu, and there are up to five Options for each Keyword. To return to the main Excurve command line type **Quit**. This menu controls FT weight, window type and transform type. The Transform Keyword can be altered if desired to display the Sine+Modulus in the Fourier transforms (**FT T 2**) which includes the imaginary part rather than just the modulus (default). Otherwise it is best to keep the defaults.

Fourier filtering (i)

As part of the learning process and in the initial phase of a simulation analysis it can be useful to isolate shells from the EXAFS spectrum by the subsequent steps of Fourier transformation, selection of a range of R in the Fourier transform, and back-transformation (Fourier filtering). (It is recommended however that final results are presented as much as possible on the basis of simulations of the unfiltered data based on molecular models rather than shell-by-shell models).

For the Zn example f01000.fin:

FF

A graphics window appears with a non-phaseshift corrected Fourier transform, in which windows for the back-transformation must be selected. This must be entered numerically in the Output window.

Enter RMIN

1.3

Enter RMAX

2.0

Enter MOVE, NEXT or CONTINUE

C

Enter shell radius or CONTINUE

Fourier filtering (ii)

Enter shell radius or CONTINUE

1.65 (an average value is entered)

Enter PRINT, REPLACE, NEXT, END

PRint

Filename: [exffta1.dat] ?

Enter a sensible name, **<return>** to accept the default.

Spectrum printed in exffta1.dat

Enter PRINT, REPLACE, NEXT, END

End

The Fourier filtered shell in the data set Exffta1.dat may be read in Excurve at any time using **READ EXP** with point frequency 1 and column combination 12.

In the same example, the second shell may also be isolated by Fourier filtering, selecting the range 2.4-3.0 with shell radius 2.7.

The selected shells or the raw data may now be simulated following the instructions in the next section.

Refine Ef, R1, A1, N1 for isolated shell

- It is convenient to show the effect of refinement of EF, R1, A1, and N1 on an isolated shell
- NOTE on Debye-Waller factor: Excurve uses a = 2 σ^2, Feff uses σ^2
- 'Absolute' Debye-Waller factor reflects thermal and/or static disorder, value should never be negative
- Can be kept within boundaries: c dwmin 0.002; c dwmax 0.030
- The Fourier-filtered shell may be compared to the unfiltered data
- The result of the simulation of the filtered shell may be compared to the unfiltered data
- The small shells can be isolated and analyzed analogously, but this is probably not useful since they are part of a imidazole unit

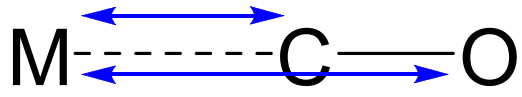
Introduce unit from amino acid data base

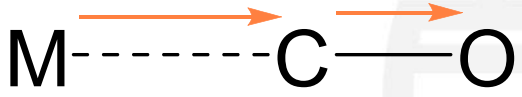
- A unit can be constructed by defining polar coordinates to a shell x and assigning it to unit y with the command c unx y.
- When data are read from a .pdb file the shells are also automatically assigned to a unit
- The imidazole unit can be introduced from the amino acid data base as ‘his’ (histidine)
- Plot the unit and inspect the polar coordinates
- Sort the shells, simulate the data, single scattering adequate ?
- No contribution above 4.5 Å, maintain only atoms of imidazole ring

Multiple Scattering

Important in the EXAFS of rigid ligand systems where the angle A-B-R-A approaches 180 º (> 140 º):

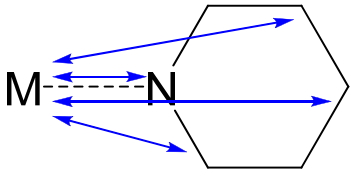
- Coordinating carbon monoxide

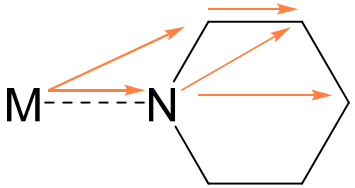




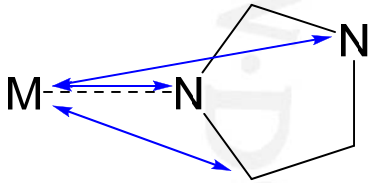
- Coordinating rigid heteroatomic ligand:

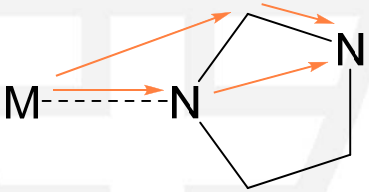
e. g. pyridine,



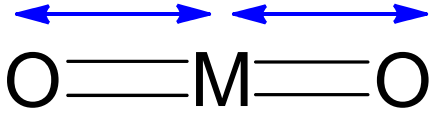


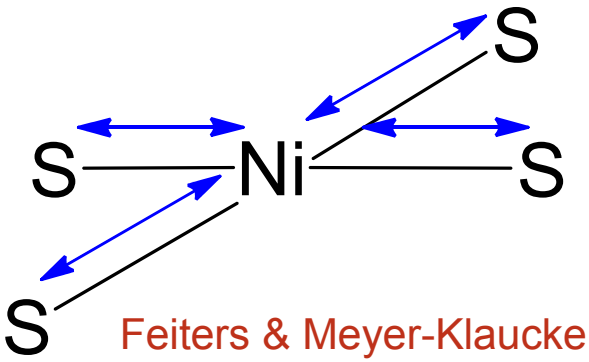
imidazole,

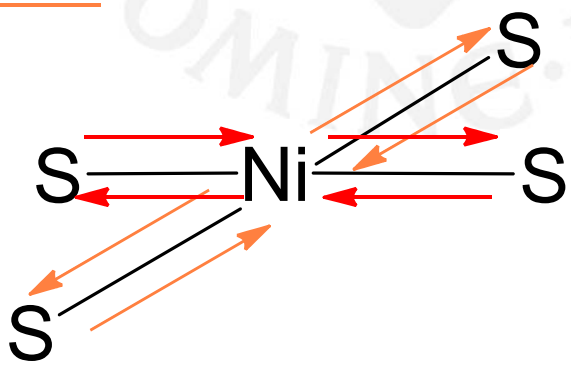




- Metal at center of unit: octahedral, square planar







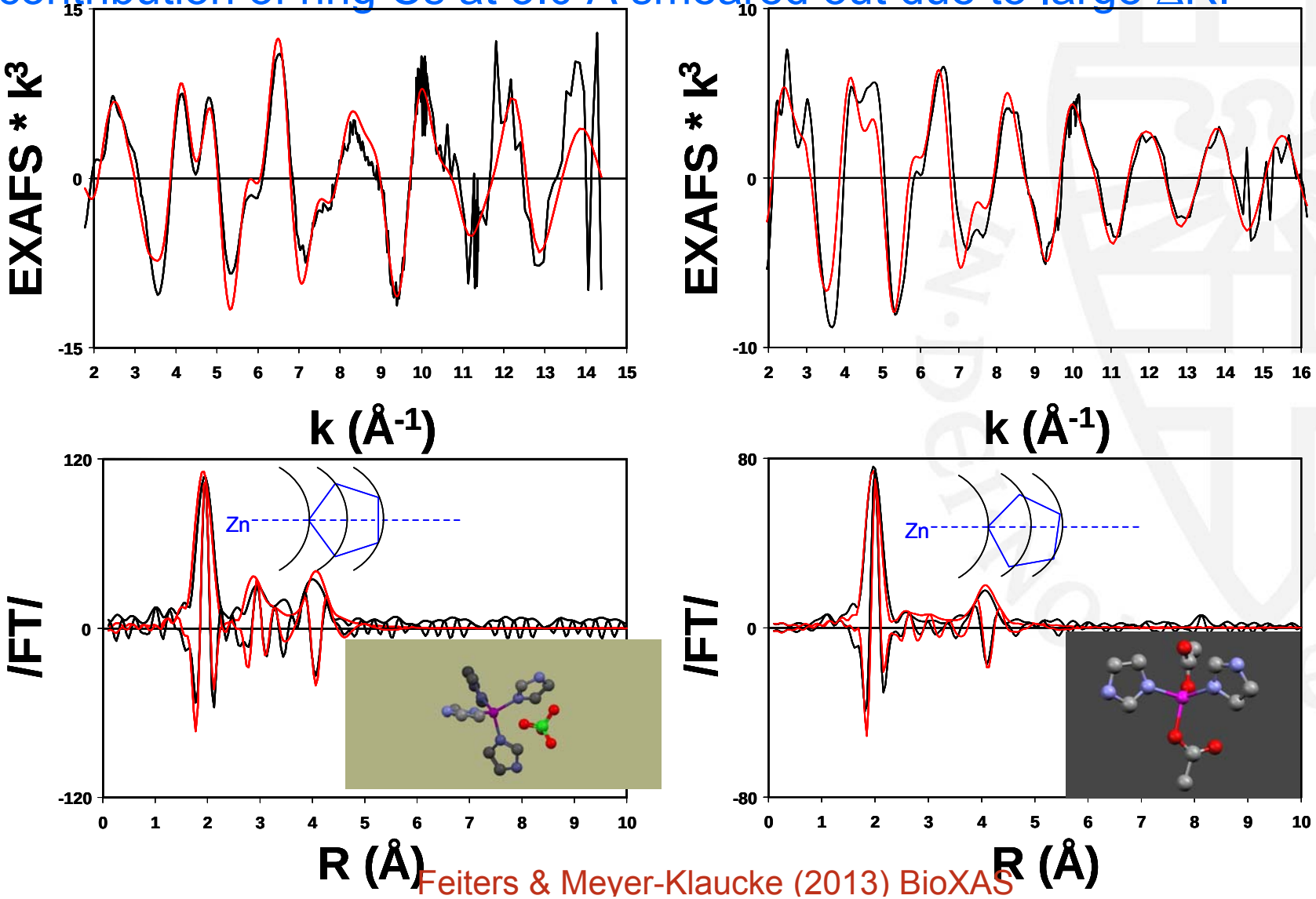
Multiple scattering refinements with restraints

- Setting the multiple scattering on units improves the simulation very much !
- Very nice agreement with the crystal structure imznpc.pdb, only by adjusting a and maybe EF
- For refinement one now attempts to reduce the number of parameters, r[1,2] and a[1,2] will refine symmetry-related parameters with the same value
- Constraints: distances within ring kept at a certain value, usually too rigid an approach
- Restraints: a penalty is added to the fit index for deviations from an idealized value (analogous to protein crystallography)
- l d; l w, fi

Binsted, N., Strange, R. W. & Hasnain, S. S. (1992). Constrained and Restrained Refinement in EXAFS Data Analysis with Curved Wave Theory. *Biochemistry* 31, 12117-12125.

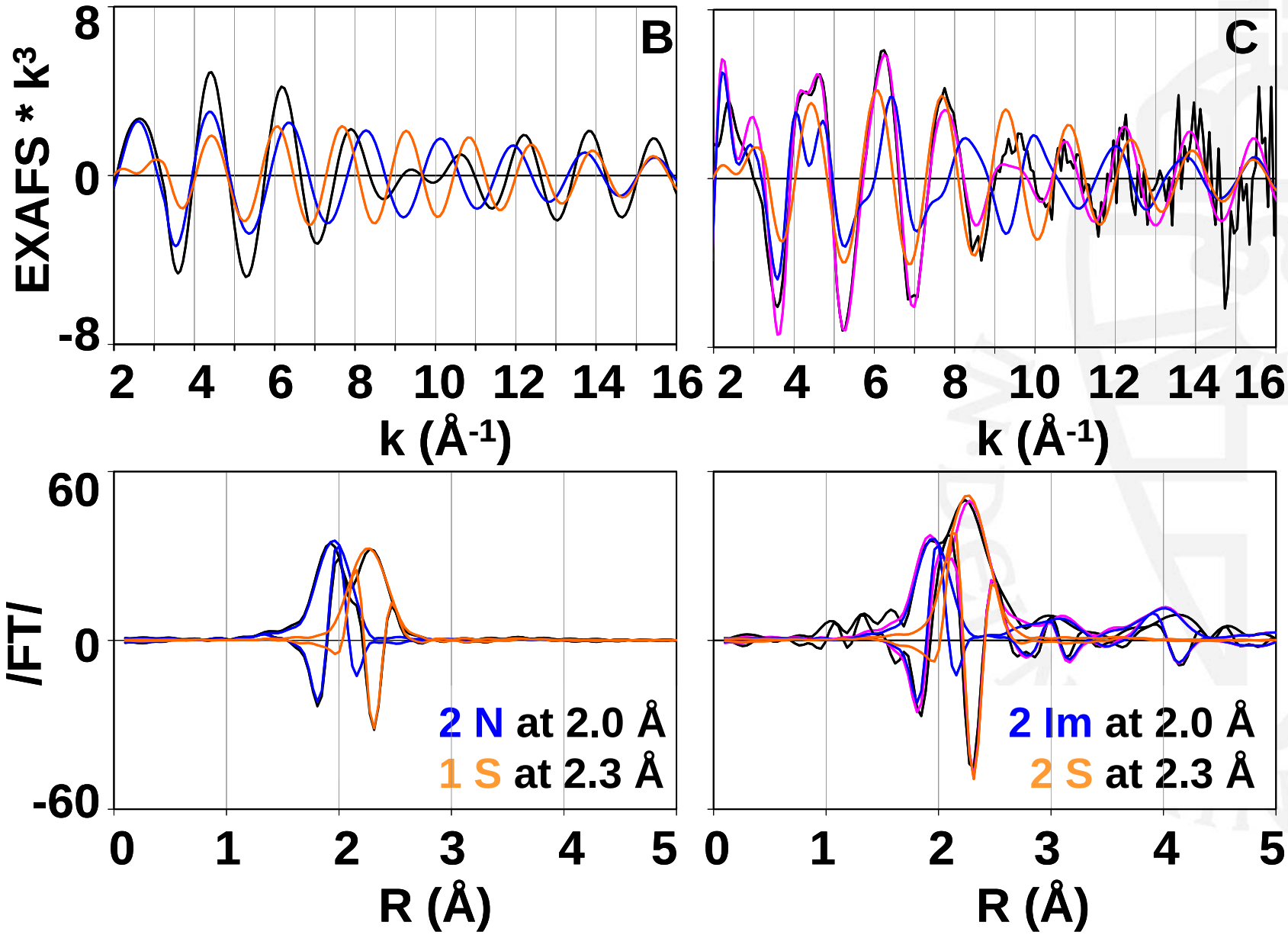
Simulations of Zn imidazole complexes

- left) Simulation of 'straight' $[Zn(im)_4](ClO_4)_2$
- right) 'tilted' imidazole in $Zn(im)_2(OAc)_2$: *camel back* less pronounced, contribution of ring Cs at 3.0 Å smeared out due to large ΔR.



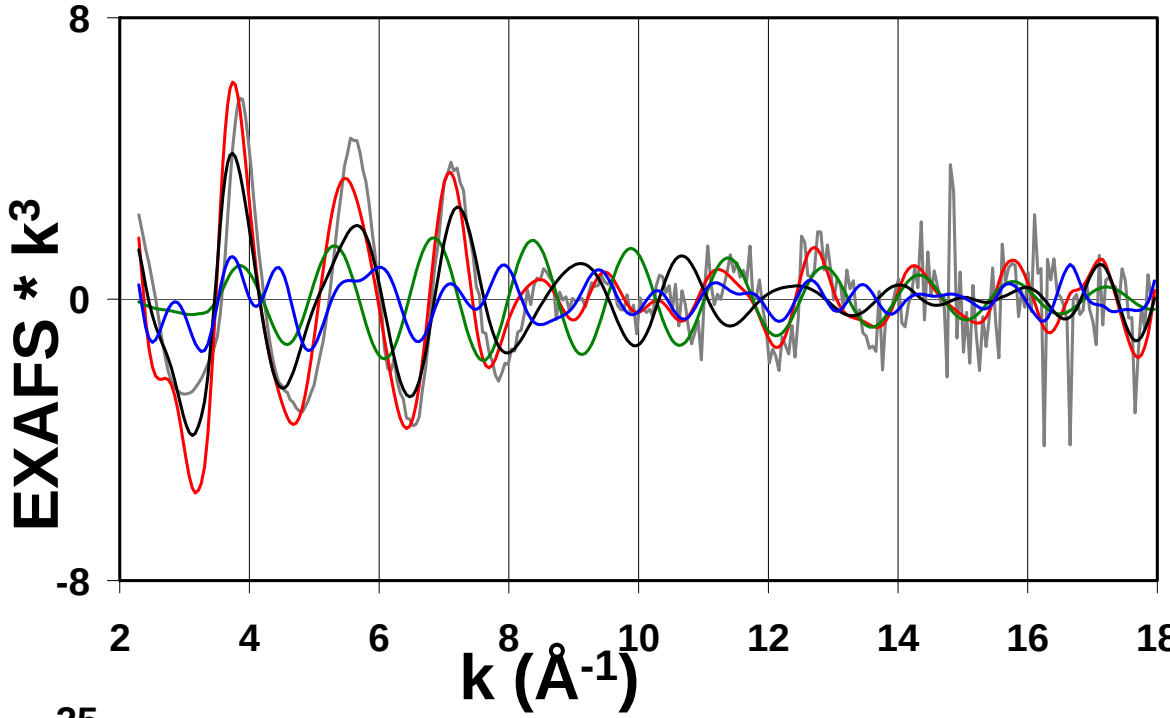
Feiters & Meyer-Klaucke (2013) BioXAS

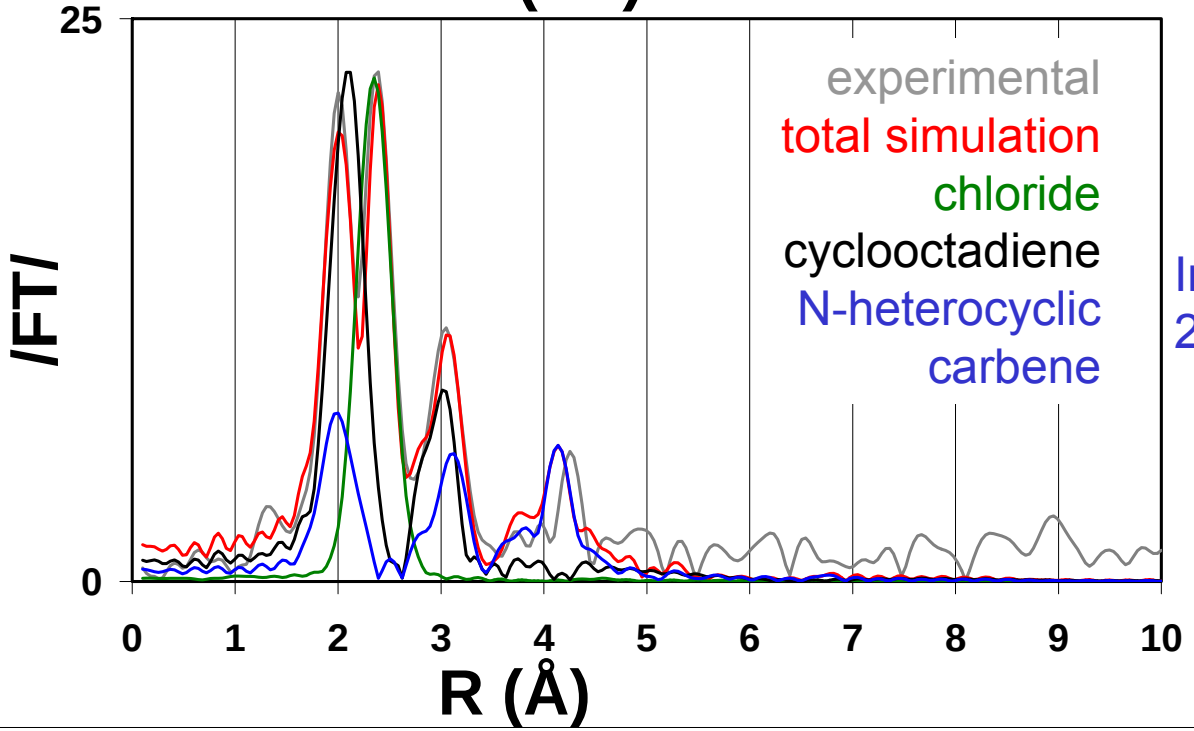
S and Im ligands to Zn resolved for HIV-2 integrase



M. C. Feiters et al. J. Synchr. Rad. 10 (2003) 86; Feiters & Meyer-Klaucke (2013) BioXAS

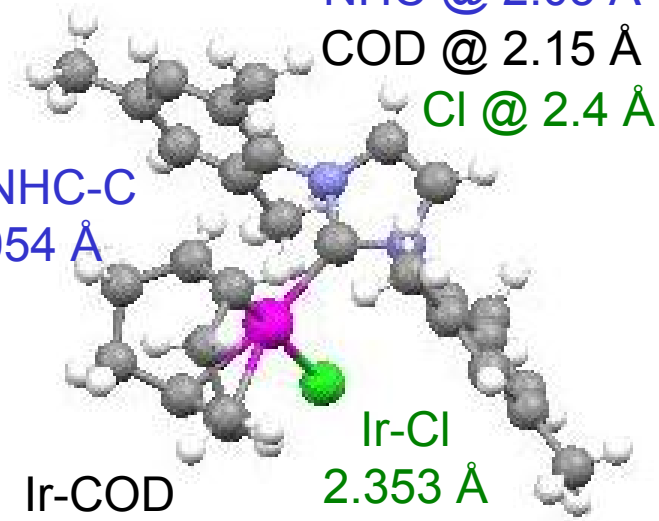
Simulation on basis of Xtal structure





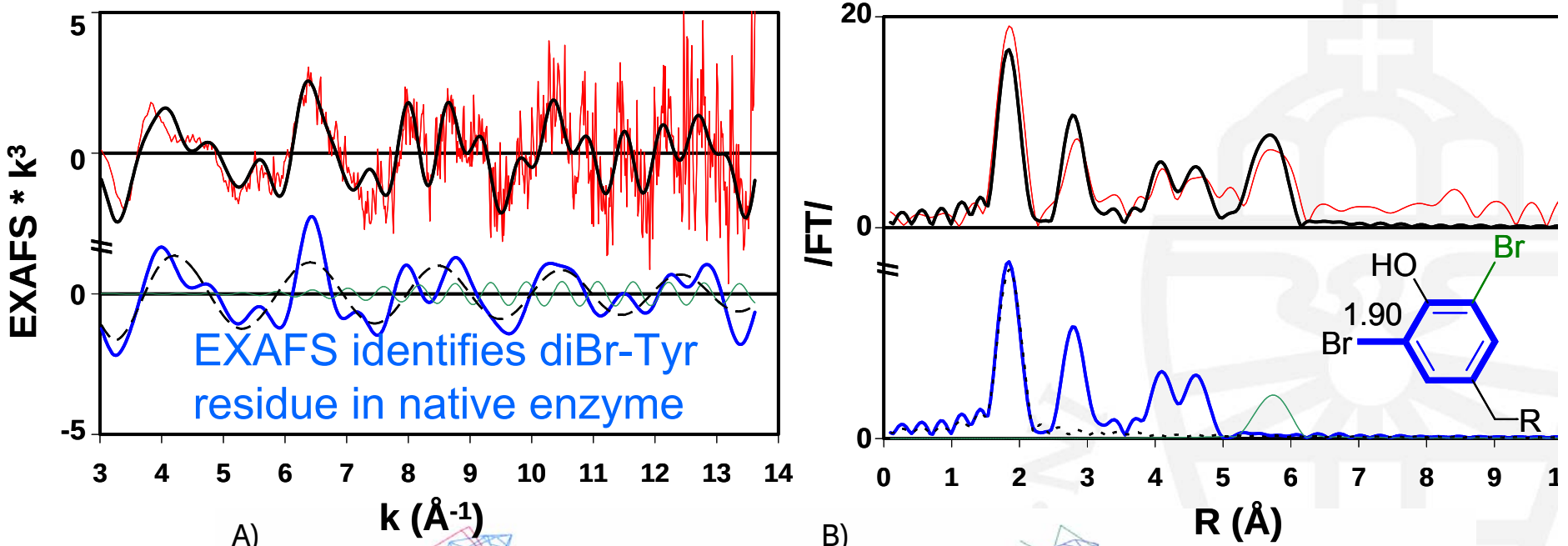
- Ir L3 EXAFS of NHC Ir(I) (cod) Cl
- Simulated with EXCURVE, pdb file, COD and NHC units created, multiple scattering
- Only slight adjustment of ΔEF, r, a needed

EXAFS
NHC @ 2.05 Å
COD @ 2.15 Å
Cl @ 2.4 Å

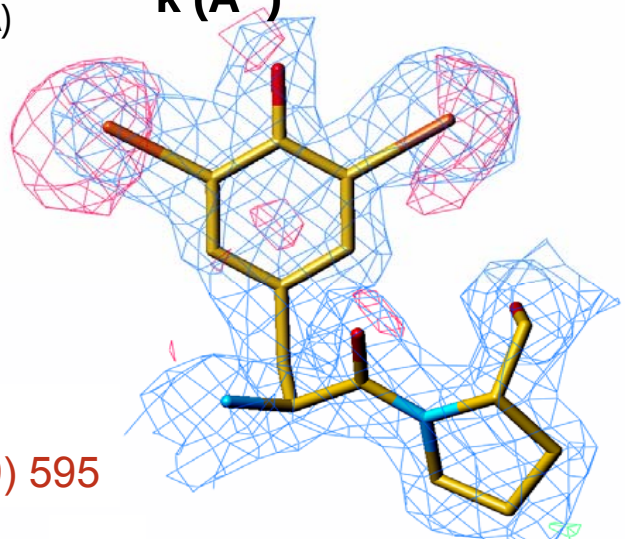


Ir-COD
2.107, 2.128, 2.176, 2.195 Å
3.003, 3.026, 3.062, 3.106 Å

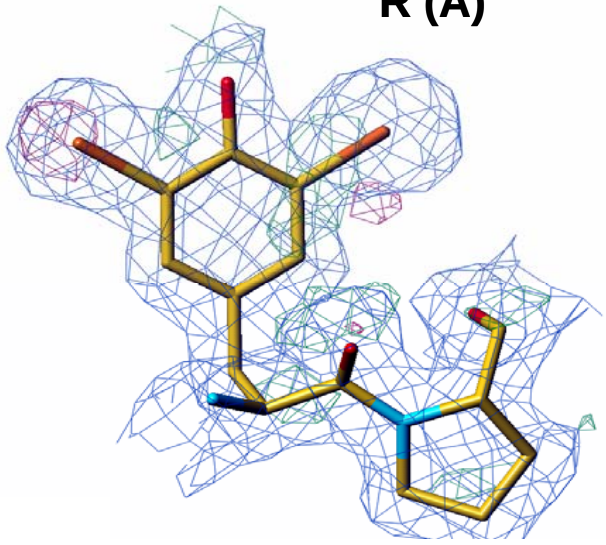
Br in native VBPO (*Ascophyllum nodosum*)



A)



B)



Weyand et al.
JMB 293 (1999) 595

Blue, electron density; red, negative Fourier-difference map contoured at a 3σ level.

Original coordinates:
Di-iodo with 0.5 occupancy

Feiters et al.
JACS 127 (2005) 15340

Revised coordinates:
Di-bromo

XAS Data Analysis Overview

- **Step 1 -Collect all available information**
- studies on other enzymes with similar functionality
  - biochemical studies on enzyme
  - **Step 2 –XANES analysis**
  - possible coordination/geometry of the Fe site
    - **» Step 3 –EXAFS analysis**
  - visual inspection of the spectrum and its FT
- test different structural (physically reasonable!) models
- validate the best refined model by checking its consistency
- with XANES and available knowledge

What happens if we use other structural models ?

>We could compare the fit index for each model to do a meta-analysis

>Strategy of *Automatic BioXASRefinement & Analysis* (ABRA)

Table 1

Comparison of published Zn-EXAFS data with *ABRA*'s results (1σ errors on the last digit extracted by meta-analysis are given in parentheses).

S, His and O are the zinc donor groups and their numbers are given for the different samples and structural models. The sum of His and O donor groups given by the meta-analysis is abbreviated as Low-*Z* ligands.

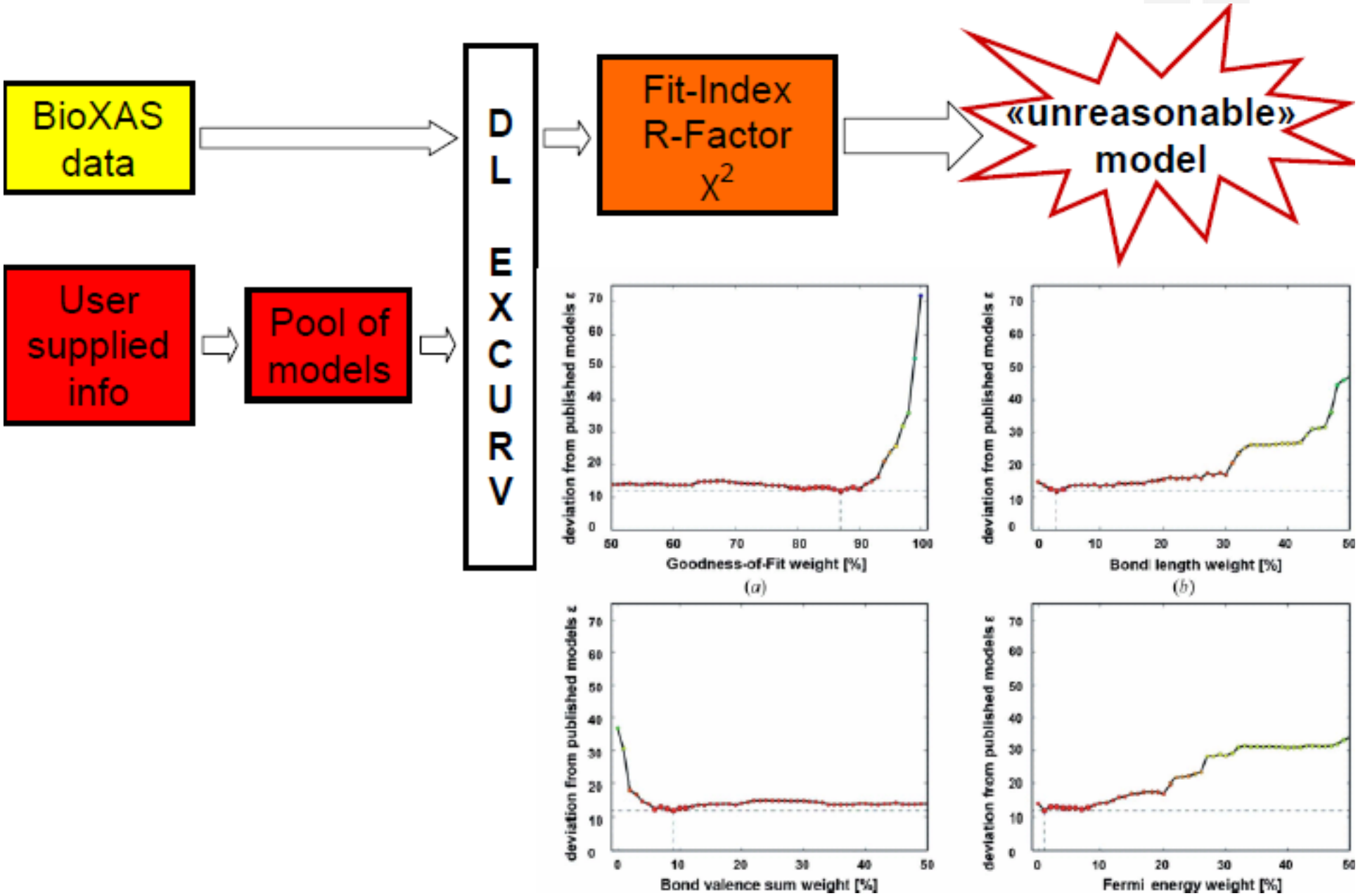
	Number of ligands							
	Published results			Best model by *ABRA*			Average of good models (meta-analysis)	
	S	His	O	S	His	O	S	Low-*Z*
Zinc finger proteins								
HPV E7	4.0			4.0			4.0 (4)	0.2 (2)
GCM	3.0	1.0		3.0	1.0		3.0 (0)	1.2 (2)
ZnF-UBP	2.67	1.33		2.67	1.00	0.33	2.8 (3)	1.2 (3)
Catalytic active sites								
Ec ZiPD		2.0	2.5		3.0	1.5	0.0 (0)	4.2 (2)
At Glx2-2		2.5	2.5		4.0	0.5	0.0 (0)	4.2 (2)
Bc bla	0.5	3.0	1.0		3.0	1.0	0.1 (2)	4.0 (4)
Model compounds								
Bis(acetato)bis(imidazol)zinc(II)		2.0	2.0		2.0	2.0	0.0 (0)	4.2 (2)
Tetrakis(imidazole)zinc(II)-perchlorate		4.0			3.5	0.5	0.0 (0)	4.3 (2)

What can we learn from the analysis of other structural models ?

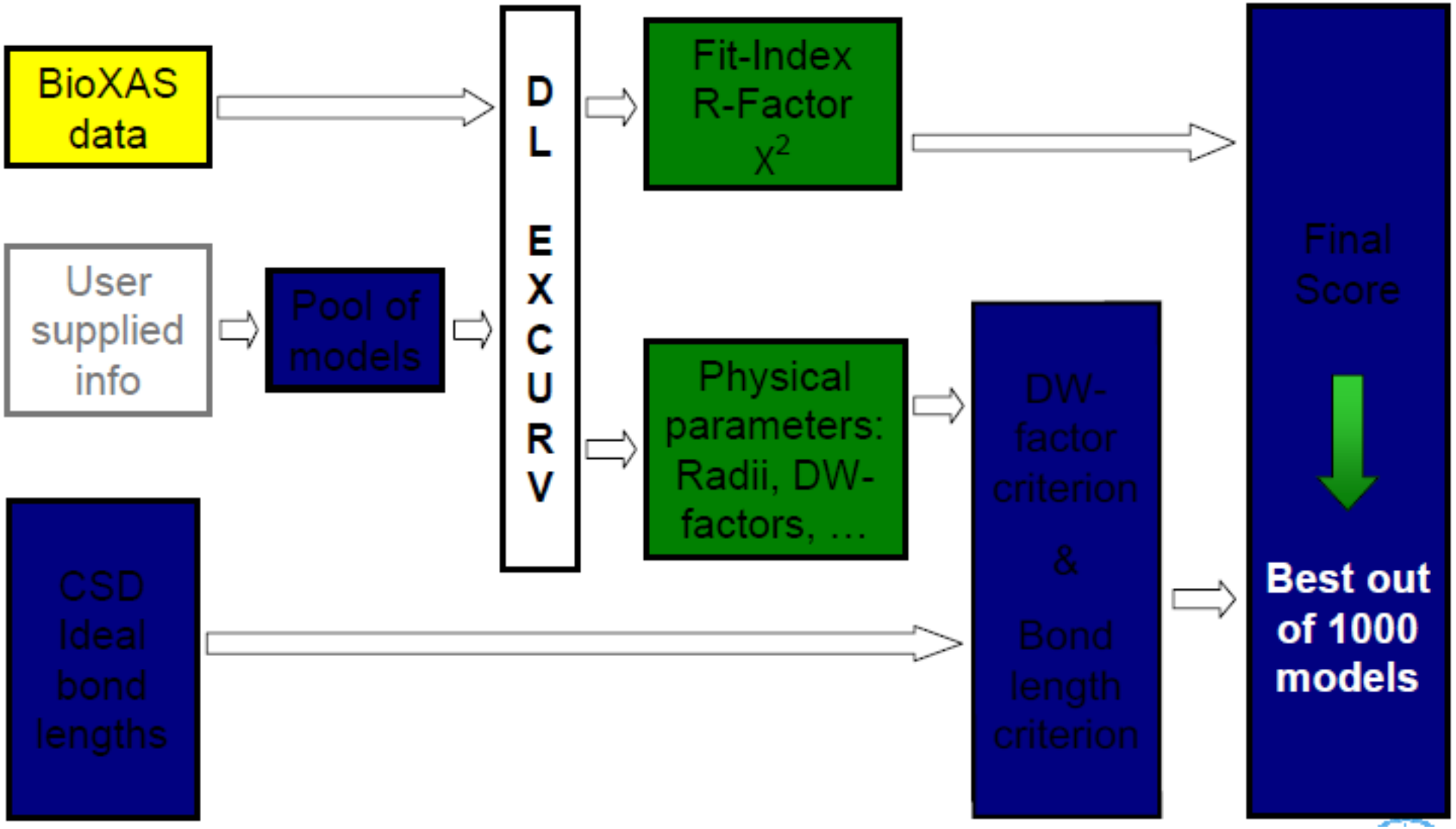
> Meta-analysis

	Average of good models (meta-analysis)	
	S	Low-Z
Zinc finger proteins		
HPV E7	4.0 (4)	0.2 (2)
GCM	3.0 (0)	1.2 (2)
ZnF-UBP	2.8 (3)	1.2 (3)
Catalytic active sites		
Ec ZiPD	0.0 (0)	4.2 (2)
At Glx2-2	0.0 (0)	4.2 (2)
Bc bla	0.1 (2)	4.0 (4)
Model compounds		
Bis(acetato)bis(imidazol)zinc(II)	0.0 (0)	4.2 (2)
Tetrakis(imidazole)zinc(II)-perchlorate	0.0 (0)	4.3 (2)

EXAFS data analysis: The standard way



ABRA: Automated biological EXAFS refinement and analysis



Validation of EXAFS results

- Criteria helping you to judge the quality of your refinement in the corresponding publications
- Most important:
- typical metal-ligand distances
  - Harding, M. M. (1999). ActaCryst. D55, 1432–1443.
  - Harding, M. M. (2000). ActaCryst. D56, 857–867.
  - Harding, M. M. (2001). ActaCryst. D57, 401–411.
  - Harding, M. M. (2002). ActaCryst. D58, 872–874.
  - Harding, M. M. (2004). ActaCryst. D60, 849–859.
  - Harding, M. M. (2006). ActaCryst. D62, 678–682.
- Bond valance sum analysis (keep in mind you obtain information on the oxidation state of the absorber from the edge position)
- Neither concept considers the spin-state at present. Take care!
- Energy shift (in Excurve called EF) should be similar for similar samples (If not you might have assumed the wrong neighbor atom (turn metals into gold))
- >ABRA

Wellenreuther, G., Parthasarathy, V., and Meyer-Klaucke, W., *Towards a Black-Box for Biological EXAFS Data Analysis –II. Automatic BioXASRefinement & Analysis (ABRA).*J Synchrotron Radiat, 2010. 17(1): p. 25-35.

Thank you for your attention

- Take home message:
- >Running Excurve is straight forward

- Errors in EXAFS
- >coordination number for first shell: ~ 20%

- >distances: ~ 0.02Åfor first shell

- >element type: about ±1

- Increase significance of results by including all available information

Graphics submenu

- Be careful when playing with the theory options as a non-expert.
- Keep in mind that any changes made in theory options have to be communicated in publications.
- **If not, you are cheating**.
- In case you think changes are required, you need to explain them.