Structural analysis by single crystal X-ray diffraction

El-Eulmi BENDEIF & Sébastien PILLET

Laboratoire de Cristallographie, Résonance Magnétique et Modélisations Université de Lorraine













N ~ 10^{18} atoms



Neumann principle : « the symmetry elements of any **physical property** of a crystal **should include all the symmetry elements of the symmetry point group of that crystal** »

Fundamental Concepts

Two pillars of crystallography

1- The symmetry of ideal crystalline structures: a better understanding of crystalline architectures and the consequences for their physical properties

2- The phenomenon of diffraction: the investigative tool of choice for fine, detailed structural analysis of atomic and molecular arrangements in ordered crystalline solids





Catalytic



P

Structure-properties relationship

Electrical







Reciprocal space



Symmetry Bond distances Bond angles disorder Thermal smearing

At thermodynamic equilibrium or out of equilibrium



Structural model









- Symmetry : determination of point group
- twins?
- Polymorphism?





- Crystal quality
- Microstructure
- Diffuse scattering (=disorder)
- twins?
- Symmetry (determination of certain translational symmetry elements, lattice mode)
- Chemical composition
- Symmetry (determination of space group)
- Structural organisation
- Studies as a function of T, P, ...



N ~ 10^{18} atoms



1. Asymmetric unit = motif





Asymmetric unit = motif Space group







Principle of single crystal x-ray diffraction



Thomson elastic scattering Radial process

$$\lambda [\text{Å}] = \frac{12.398}{\varepsilon [\text{keV}]}$$
 $k_i = k_d (\lambda_i = \lambda_d)$



Periodic system

Diffraction

(constructives interferences between scattered waves by all the electrons)



Unit cell structure factor



Scattering object: $\rho_{unit \, cell}(\vec{r})$



$$\rho_{unit \, cell}(\vec{r}) = \sum_{j}^{unit \, cell} \rho_{j}(\vec{r})$$
FT
$$F_{unit \, cell}(\vec{H}) = \sum_{j}^{unit \, cell} f_{j} e^{-2i\pi \vec{H} \cdot \vec{r}_{j}}$$

 f_j : atomic scattering factor

Unit cell structure factor



Argand diagram

Accounting for anomalous dispersion

Structure factor and interference function



Interference function



Finite size effects

Influence of finite size on the scattering process :

$$\rho_{crystal}(\vec{r}) = \rho_{crystal \infty}(\vec{r}) \cdot [boite(\vec{r})] = \left[\sum_{u,v,w=-\infty}^{+\infty} \delta(\vec{r} - \vec{r}_{u,v,w})\right] * \left[\sum_{unit \ cell} \rho_{atomic}(\vec{r})\right] \cdot [boite(\vec{r})]$$

$$FT \int F_{crystal}(\vec{H}) = \Gamma(\vec{H}) \cdot F_{unit \ cell}(\vec{H}) * B(\vec{H})$$

$$B(\vec{H}) = \frac{\sin(\pi L_a q_x)}{\pi L_a q_x} \frac{\sin(\pi L_b q_y)}{\pi L_b q_y} \frac{\sin(\pi L_c q_z)}{\pi L_b q_z} \frac{500}{400} \frac{100 \ mailles}{500 \ mailles}$$

$$G_{ij} = \frac{300}{-000} \frac{100 \ mailles}{-000} \frac{100 \ mailles}{-000}$$

Reciprocal lattice associated to a single crystal



Principle of a single crystal diffraction experiment



















Bragg condition : the path difference between two waves must be equal to n times the wavelength









Collecting diffraction frames





Collecting diffraction frames







Indexing and data integration



Indexing : determine orientation matrix and unit cell parameters a^* , b^* , c^* , α^* , β^* , γ^* so that the Miller indices (h k l) which define the position of each Bragg peak in reciprocal space are integers :

$$\vec{H} = h\vec{a} * + k\vec{b} * + l\vec{c} *$$



h	k	1	I(H)	σ(I(H))
0	0	1	5.10	0.2
0	0	1	6.30	0.4
0	0	2	47.40	2.3
-5	2	-9	159.39	7.8

 $\vec{H}, I(\vec{H}), \sigma(I)$

- □ Pixel-by-pixel summation over an integration volume
- □ Application of a 2D or 3D profile / profile fitting and learnt profile
- □ Background noise estimation

Collecting diffraction frames

(0kl) layer



Space group C2/c

Systematic absence : (Okl) with k=2n+1 (OOl) with l=2n+1



Collecting diffraction frames

(0kl) layer



Space group C2/c

Systematic absence : (Okl) with k=2n+1 : C lattice mode (OOl) with l=2n+1 : mirror c perp. To b















 $\square \qquad \left\{ \vec{H}, \ I(\vec{H}), \ \sigma(I(\vec{H})) \right\}$

Structure solution methods

Objective : determine the distribution of atomic electron density in the unit cell from diffraction data :

14

2

0 -2 -4 -6 -8

$$F_{unit cell}(\vec{H}) = TF(\rho(\vec{r}))$$

$$\rho(\vec{r}) = \frac{1}{V} \sum_{h,k,l=-\infty}^{+\infty} F(\vec{H}) \exp\left[-2i\pi \vec{H} \cdot \vec{r}\right]$$
with $F(\vec{H}) = \left|F(\vec{H})\right| \exp\left[i\varphi(\vec{H})\right]$

fondamental problem = phases are lost

$$I\left(\vec{H}\right) \propto F\left(\vec{H}\right) \bullet F^*\left(\vec{H}\right) \propto \left|F\left(\vec{H}\right)\right|^2$$



Structure solution methods



edited by Charles W. Carter, Jr., and Robert M. Sweet, pp. 110-128. New York: Academic Press.

R. Blessing, HWI, Buffalo

Structure solution methods

• Molecular replacement method:

- prior knowledge of molecular fragments

• Heavy atom method / Patterson : Dirdif, Shelx

- prior knowledge = the structure contains at least one heavy element. Use of contrast effect

• direct methods: Shelx, Sir

- prior knowledge = the crystal structure consists of discret atoms : accumulation of electron density in certain region of space (atomicity), chemical composition, symmetry

• Charge flipping : Superflip

- prior knowledge = the crystal structure consists of discret atoms : accumulation of electron density in certain region of space (atomicity)

Direct methods

Direct determination of the phases of structure factors from structure factor moduli using mathematical and probabilist relations that come from **some obvious features of the electron density**,

- **positivity** of the electron density $\rho(\mathbf{r}) > 0$ (Hauptman & Karle, 1953) : gives a restraint on the set of physically acceptable phases (Hauptman determinant ≥ 0)
- **atomicity** of the electron density (Sayre, 1951): the electron densities of different atoms do not overlap
- symmetry of the crystal structure (origine)
- chemical composition of the unit cell

Normalized structure factors :

$$E(\vec{H}) = \sum_{j=1}^{N_{at}} \frac{f_j}{\sqrt{\sum f_j^2}} exp[2i\pi\vec{H}.\vec{r}_j] = \frac{F(\vec{H})}{\sqrt{\langle F^2(\vec{H}) \rangle_{resolution}}}$$

Allows to get rid of thermal smearing, and consider point atoms

Direct methods : detection of inversion center

$$E(\vec{H}) = \sum_{j=1}^{N_{at}} \frac{f_j}{\sqrt{\sum f_j^2}} exp[2i\pi\vec{H}.\vec{r_j}] = \frac{F(\vec{H})}{\sqrt{\langle F^2(\vec{H}) \rangle_{resolution}}_{slices}}$$

P(E) allows to distinguish a centrosymmetric structure from a non-centrosymmetric one



We can also analyze the various moments of the distribution E : $|E|^2$, $|E|^3$, $|E|^4$, $|E|^5$, $|E|^6$

Direct methods : Sayre equation



For strong values of $|F(\vec{H})|$, $|F(\vec{H})|^2$ is also strong, real and positive. It is thus highly probable that the most important terms of the Fourier summation are also real and positive.

Thus if $F(\vec{K})$ and $F(\vec{H}-\vec{K})$ are also high,

 $\varphi\left(-\vec{H}\right)+\varphi\left(\vec{K}\right)+\varphi\left(\vec{H}-\vec{K}\right)\approx 0$

$$P\left(\varphi(-\vec{H})\middle|\varphi(\vec{K}),\varphi(\vec{H}-\vec{K})\right)$$

Direct methods : tangent formula

If we rewrite the Sayre equation

$$F\left(\vec{H}\right) = \frac{\gamma\left(\vec{H}\right)}{V} \sum_{\vec{K}} F\left(\vec{K}\right) F\left(\vec{H} - \vec{K}\right)$$

$$|F(\vec{H})|\exp[i\varphi(\vec{H})] = \frac{\gamma(\vec{H})}{V} \sum_{\vec{K}} |F(\vec{K})F(\vec{H}-\vec{K})|\exp[i(\varphi(\vec{K})+\varphi(\vec{H}-\vec{K}))]$$

 $|F(\vec{H})|\exp[i\varphi(\vec{H})] = A(\vec{H}) + iB(\vec{H})$

$$A(\vec{H}) = \frac{\gamma(\vec{H})}{V} \sum_{\vec{K}} \left| F(\vec{K}) F(\vec{H} - \vec{K}) \right| \cos\left(\varphi(\vec{K}) + \varphi(\vec{H} - \vec{K})\right)$$

$$B(\vec{H}) = \frac{\gamma(\vec{H})}{V} \sum_{\vec{K}} \left| F(\vec{K}) F(\vec{H} - \vec{K}) \right| \sin(\varphi(\vec{K}) + \varphi(\vec{H} - \vec{K}))$$

Tangent formula

$$\tan \varphi(\vec{H}) = \frac{\sum_{\vec{K}} \left| F(\vec{K}) F(\vec{H} - \vec{K}) \right| \sin(\varphi(\vec{K}) + \varphi(\vec{H} - \vec{K}))}{\sum_{\vec{K}} \left| F(\vec{K}) F(\vec{H} - \vec{K}) \right| \cos(\varphi(\vec{K}) + \varphi(\vec{H} - \vec{K}))}$$

Direct methods : tangent formula

Tangent formula

$$\tan \varphi \left(\vec{H} \right) = \frac{\sum_{\vec{K}} \left| F\left(\vec{K} \right) F\left(\vec{H} - \vec{K} \right) \sin \left(\varphi \left(\vec{K} \right) + \varphi \left(\vec{H} - \vec{K} \right) \right)}{\sum_{\vec{K}} \left| F\left(\vec{K} \right) F\left(\vec{H} - \vec{K} \right) \right| \cos \left(\varphi \left(\vec{K} \right) + \varphi \left(\vec{H} - \vec{K} \right) \right)}$$

Allows to explore the phase space by correlating phase largely distributed : iterative method of refinement and extension of phases



- Phases are refined in reciprocal space
- Direct space impose a strong atomicity constraint on phases

Direct methods : Advantages & Disadvantages

Advantages:

- □ This set of methods is **very efficient** and most of the small molecules crystal structures are determined with this approach.
- Quick in practice, direct methods give a high proportion of the atomic positions, facilitating the refinement step.
- □ Figure of merits are associated to the results, this allows the result to be appreciated with **hindsight**.
- □ Direct methods are a matter of equations, they are fully automated, thus **easy to use** especially because decades of experience have allowed to have defaults parameters well optimized.

Disadvantages:

- **Difficult** to understand due to the mathematical aspects behind. Very often used as black-boxes.
- □ The **space-group must be determined** prior to applying direct methods.
- The phase relationships become a minor feature as the number of atoms in the structure increases.
 Above 2000 atoms in the unit-cell, direct methods are inefficient. This method is not applicable for proteins and large macromolecules.

Initial references for direct methods : Karle, J. and Hauptman, H., *Acta Cryst.* 3 (1950), 181-187

Charge Flipping (CF) method (2004)

General concept : the idea is to calculate anyway the inverse Fourier Transform of the structure factors and to correct the **wrong phases** obtained by a succession of iterations based on the fact that an **electron density must be positive**

- ab-initio determination of crystal structure (periodic and aperiodic structures, single crystal and powder diffraction, X and N diffraction)



- $|\mathbf{F}|$: experimental amplitude of structure factors, without phase
- ${\bf F}:$ structure factors with phase
- $\boldsymbol{\rho}~:$ electron density with potential negative values
- ${\bf g}$: electron density with only positive values
- ${\bf G}:$ structure factors with phase

Sequence of the CF algorithm

Step 0 : random phases are attributed to the experimental data.

Step 1 : electron density in the unit-cell is calculated from F.

Step 2 : the electron densities that are negative are inverted to
positive. This is the "charge flipping" that names the method

Step 3 : new structure factors, G, with their phases are calculated.

Step 4 : the phases of G are kept while their amplitudes are replaced by the experimental ones, giving new F,

Charge Flipping (CF) method (2004)

Advantages:

- CF needs no preliminary information on the crystal, except the Unit-cell parameters. The space group is NOT required since all calculation are made in P1, the space group is deduced at the end of the process. No need of information on the chemical composition.
- □ CF is tolerant to imperfects experimental data, though high noise may be problematic.
- □ CF proved to **be very efficient** in the case of aperiodic structures. Superspace crystallographers use it.
- □ CF is also of help in powder diffraction since the space group is often difficult to obtain in powder and CF does not need the space group.
- **CF** is easy to understand

Disadvantages:

- □ Data with bad resolution **may not be pertinent** for CF.
- □ Not much parameters that the user can adjust to converge towards the result

Basic references for CF : The method was proposed in G. Oszlanyi and A. Suto, Acta Cryst A 60, 134, 2004 The implementation was developed in L. Palatinus, Acta Cryst B 69, 1, 2013

Improvement (refinement) of the structure



Structural analysis

- Bond distances
- Bond angles
- Torsion angles

- Thermal smearing ellipsoids
- Intermolecular contacts



Thermal smearing and Debye Waller factor

Born-Hoppenheimer approximation : the electron density follows the displacement of nucleus



Probability density function : $p(u) = \frac{1}{\sqrt{2\pi U}} \exp\left(\frac{-u^2}{2U}\right)$ in the spherical harmonic approximation

 $U = \langle u^2 \rangle$ Is the quadratic mean displacement of the atom from its equilibrium position

Thermal smearing and Debye Waller factor

 $\begin{pmatrix} U_{11} & U_{12} & U_{13} \\ & U_{22} & U_{23} \\ & & & U_{33} \end{pmatrix}$

$$\rho_{dynamic}(\vec{r}) = \int \rho_{static}(\vec{r} - \vec{u}) p(\vec{u}) d(\vec{u}) = \rho_{static}(\vec{u}) * p(\vec{u})$$

$$FT \qquad f(\vec{H}) = F_{static}(\vec{H}) \cdot T(\vec{H})$$

In the isotropic harmonic approximation :

$$T(\vec{H}) = \exp\left(-2\pi^2 U H^2\right) = \exp\left(-8\pi^2 U \frac{\sin^2 \theta}{\lambda^2}\right)$$
$$= \exp\left(-B \frac{\sin^2 \theta}{\lambda^2}\right) \qquad B \sim 3-10 \text{ Å}^2 (U \sim 0.02 - 0.12 \text{ Å}^2)$$

In the anisotropic harmonic approximation :

$$T\left(\vec{H}\right) = \exp\left(-2\pi^{2}\vec{H}^{T}\cdot\vec{U}\cdot\vec{H}\right) = \exp\left(-\sum_{i,j}2\pi^{2}h_{i}h_{j}a_{i}^{*}a_{j}^{*}\left(U_{ij}\right)\right)$$

 \widetilde{U} Is the atomic displacement tensor (symmetric)

Structural disorder



Absolute structures

Friedel law :
$$I(hkl) = I(-h - k - l)$$

In the absence of anomalous dispersion, the diffraction pattern is centrosymetric : Two structures containing two different enantiomers are not distinguishable

In presence of anomalous dispersion : Resonance phenomenon (EXAFS, XANES)









Absolute structures and optical activities

$$F^{2}(hkl, x) = (1-x)|F(hkl)|^{2} + x|F(-h-k-l)|^{2}$$

Flack parameter : x

requires a heavy atom and an appropriate choice of wavelength $\boldsymbol{\lambda}$



Non centrosymmetric space group: Flack parameter x P1 P1 0.012(1) 0.018(15)

Wang et al. Chem. Commun., (2009), 6940

Polymorphism



Moliner *et al.*, (2001), *Inorg Chem*, **40**, 3986. Sheu *et al.*, (2008), *Inorg Chem*, **47**, 10866.



Single crystal combined to powder XRD

Single crystal diffraction as a funciton of T



displex





10K < T < 70K

300K < T < 1000K





Symmetry breaking : development of a surstructure



Symmetry breaking : development of a surstructure



Symmetry breaking : development of a surstructure



Bréfuel et al. Angew. Chem. (2009), 48, 9304

Diffuse scattering and disorder



Photo-crystallography: principles



Photo-chromism

Diarylethen derivatives







Single crystal diffraction under pressure







Single crystal diffraction under pressure





P. Guionneau et al., Phys. Rev. (2005) B72, 214408

Isobaric expansion



Kusz et al. J Appl Cryst., 34, 229, (2001)



Average structure and local structure



References

• diffraction geometry, 4 cercles diffractometer, indexing

W. A. Paciorek, M. Meyer and G. Chapuis, (1999), Acta Cryst. A55, 543W. Kabsch, J. Appl. Cryst. (1988), 21, 67-71

• data integration

W. Kabsch, J. Appl. Cryst. (1988), 21, 916-924
Z. Otwinowski & W. Minor, (1997), Methods in Enzymology, Vol. 276, pp. 307-326
A. G. W. Leslie, Acta Cryst. (1999), D55, 1696-1702
R. Diamond, (1969), Acta Cryst., A25, 43-55

• data scaling

W. Hamilton, J.S. Rollet, and R.A. Sparks (1965). Acta Cryst. 18, 129-130

• TDS correction

M. J. Cooper, K. D. Rouse, Acta Cryst. (1968), A24, 405

• polarisation correction

L. V. Azaroff, Acta Cryst. (1955), 8, 701-704 R. Kahn et al. J. Appl. Cryst. (1982), 15, 330

• absorption correction

Semi-empirical

R.H. Blessing (1995). *Acta Cryst.* **A51**, 33-38 A. C. T. North, D. C. Phillips and F. S. Mathews (1968), Acta Cryst. A24, 351 N. Walker and D. Stuart (1983), Acta Cryst. A39, 158

Numerical and analytical

P. Coppens, L. Leiserowitz & D. Rabinovich, (1965), Acta Cryst., 1035-1038

References

• treatment of outliers

R.H. Blessing (1997). J. Appl. Cryst. 30, 421-426
R.H. Blessing and D.A. Langs (1987). J. Appl. Cryst. 20, 427-428

• variance analysis

R.H. Blessing (1987). *Crystallogr. Rev.* **1**, 3-58 R.H. Blessing (1989). *J. Appl. Cryst.* **22**, 396-397

• Normal probability plot

S. C. Abrahams & E. T. Keve (1971). Acta Cryst. A27, 157.

• bayesian analysis of weak reflections

S. French and K. Wilson (1979). Acta Cryst. A34, 517-525.

• direct methods

D. Harker & J. S. Kasper (1948), Acta Cryst. 1, 70
J. Karle & H. Hauptman (1950), Acta Cryst. 3, 181
H. Hauptman & J. Karle (1956), Acta Cryst. 9, 45
H. Hauptman & J. Karle (1959), Acta Cryst. 12, 93
D. Sayre (1952), Acta Cryst. 5, 60

• charge flipping methods

G. Oszlanyi & A. Süto, Acta Cryst (2008), A64, 123-134

G. Oszlanyi & A. Süto, Acta Cryst (2004), A60, 134-141

G. Oszlanyi & A. Süto, Acta Cryst (2005), A61, 147-152

G. Oszlanyi & A. Süto, Acta Cryst (2007), A63, 156-163