FOUNDATIONS
ADVANCES

Volume 71 (2015)

Supporting information for article:

Image definition evaluation functions for X-ray crystallography: A new perspective on the phase problem

Hui Li, Meng He and Ze Zhang

Supporting Information

Image definition evaluation functions for X-ray crystallography: A new perspective on the phase problem

Hui LI*, Meng HE* and Ze ZHANG

Beijing University of Technology, Beijing 100124, People's Republic of China

CAS Key Laboratory of Nanosystem and Hierarchical Fabrication, National Center for Nanoscience and Technology, Beijing 100190, People's Republic of China

Zhejiang University, Hangzhou 310014, People's Republic of China

Correspondence e-mail: huilicn@yahoo.com, mhe@nanoctr.cn

In memoriam of Ms. Guang-Lian Tian

1. Parseval's theorem and corollaries for electron density images reconstructed by Fourier synthesis.

Parseval's theorem:

Suppose that *A(x)* and *B(x)* are two square integrable complex-valued functions of period 2π with Fourier series

$$A(x) = \sum_{n=-\infty}^{\infty} a_n \exp(inx)$$

and

$$B(x) = \sum_{n=-\infty}^{\infty} b_n \exp(inx)$$

respectively. Then

$$\sum_{n=-\infty}^{\infty} a_n \overline{b_n} = \frac{1}{2\pi}\int_{-\pi}^{\pi} A(x)\overline{B(x)}dx$$

where *i* is the imaginary unit and horizontal bars indicate complex conjugation.

The structure factor $F(H)$ is the Fourier transform of the electron density distribution $\rho(r)$,

$$F(H) = \sum_{r=0}^{n} \rho(r) \exp(-2\pi i H \cdot r)$$

then

$$\overline{F(H)} = \sum_{r=0}^{n} \overline{\rho(r)} \exp(2\pi i H \cdot r)$$

where $\overline{F(H)}$ and $\overline{\rho(r)}$ are complex conjugate functions of $F(H)$ and $\rho(r)$, respectively.

According to Parseval's theorem, we have

$$\frac{1}{2H} \int_{-H}^{H} F(H)\overline{F(H)} dH = \sum_{r} \rho(r)\overline{\rho(r)}$$

Because $\overline{\rho(r)}$ = $\rho(r)$ ($\rho(r)$ is a real function),

then

$$\int_{-H}^{H} \left|F(H)\right|^2 dH = 2H \times \sum_{r} \rho(r)^2$$

Since $F(H)$ is discrete, the above equation can be re-written as

$$\sum_{H} \left|F(H)\right|^2 = M \sum_{r} \rho(r)^2$$

where M is the number of structure factors.

For any given group of $|F(H)|^2$ (diffraction data), $\sum_{H} \left|F(H)\right|^2$ is a constant, then $\sum_{r} \rho(r)^2$ is also a constant.

Because

$$\sum \rho_{+}^2 + \sum \rho_{-}^2 = \sum_{r} \rho(r)^2$$

where ρ_+ and ρ_- are the non-negative and negative electron density, respectively, at the grids across the unit cell, then

$$\sum \rho_{+}^2 + \sum \rho_{-}^2 = \text{constant}$$

2. Iterative process of reconstructing the electron density image with maximum entropy method (MEM).

The entropy of the electron density distribution is defined as

$$H = -\sum\left(\frac{\rho_i}{Z}\right)\ln\left(\frac{\rho_i}{Z}\right)$$

where ρ_i is the number of electrons at the *i*th grid of the unit cell, Z is total number of electrons in the unit cell, $Z = \sum \rho_i$.

For any phase set which is assigned to the observed moduli, the consistence between the calculated and "observed" structure factors can be measured by the residual factor, which is defined as

$$\mathrm{R}_{\mathrm{mem}} = \frac{1}{M}\sum\frac{(F_{obs}^{H} - F_{cal}^{H})^2}{(\sigma F_{obs}^{H})^2}$$

where F_{obs}^{H} is the "observed" structure factor, F_{cal}^{H} is the calculated structure factor, σF_{obs}^{H} is the standard deviation of F_{obs}^{H}, and *M* is the number of structure factors. The "observed" structure factor is the combination of observed modulus and assigned phase. F_{cal}^{H} is generated by calculating the Fourier transform of the electron density distribution.

Build a function

$$G(\rho) = H + \lambda R_{mem} + Z$$

where λ is a disposable constant to be evaluated. The goal of iterative process of MEM is to maximize the entropy *H* subject to the conditions that R_{mem} is minimized and Z remains unchanged. This was achieved by setting

$$\frac{\partial G}{\partial \rho_i} = \frac{\partial H}{\partial \rho_i} + \lambda\frac{\partial R_{mem}}{\partial \rho_i} + \frac{\partial Z}{\partial \rho_i} = 0$$

Because

$$\frac{\partial R_{mem}}{\partial \rho_i} = -2\times\frac{1}{M}\sum\frac{(F_{obs}^{H} - F_{cal}^{H})}{(\sigma F_{obs}^{H})^2}\times\frac{\partial F_{cal}^{H}}{\partial \rho_i}$$

and

$$F_{cal}^{H} = \sum_{r=0}^{n}\rho(r)\exp(-2\pi i H\cdot r)$$

then

$$\frac{\partial R_{mem}}{\partial \rho_i} = -2\times\frac{1}{M}\sum\frac{(F_{obs}^{H} - F_{cal}^{H})}{(\sigma F_{obs}^{H})^2}\exp(-2\pi i H\cdot r)$$

Since

$$\frac{\partial H}{\partial \rho_i} = \frac{(-\ln(\rho_i)+1)}{Z}$$

and

$$\frac{\partial Z}{\partial \rho_i} = 1$$

then

$$\frac{\partial G}{\partial \rho_i} = -\frac{\ln(\rho_i)}{Z} + \frac{1}{Z} - \frac{2\lambda}{M} \times \sum \frac{(F_{obs}^H - F_{cal}^H)}{(\sigma F_{obs}^H)^2} \exp(-2\pi i H \cdot r) + 1 = 0$$

This leads to

$$\ln(\rho_i) = Z + 1 - \frac{2Z\lambda}{M} \times \sum \frac{(F_{obs}^H - F_{cal}^H)}{(\sigma F_{obs}^H)^2} \exp(-2\pi i H \cdot r)$$

and then

$$\rho_i = \exp\left[Z + 1 - \frac{2Z\lambda}{M} \times \sum \frac{(F_{obs}^H - F_{cal}^H)}{(\sigma F_{obs}^H)^2} \exp(-2\pi i H \cdot r) \right]$$

In the iterative process of MEM, the value of ρ_i in the $(n+1)$th iterative cycle, ρ_i^{n+1}, is derived from the electron density image of the nth cycle using the above equation, namely

$$\rho_i^{n+1} = \exp\left[Z + 1 - \frac{2Z\lambda}{M} \times \sum \frac{(F_{obs}^H - F_{cal,n}^H)}{(\sigma F_{obs}^H)^2} \exp(-2\pi i H \cdot r) \right]$$

where $F_{cal,n}^H$ is the calculated structure factor obtained by Fourier transforming the electron density image of the nth iterative cycle.

Before the iterative process of MEM, a positive initial value is assigned to ρ_i. In most cases, a uniform positive value is assigned to each grid as the initial value, which is known as the uniform model. As ρ_i is an exponential function, it is always positive in the iterative process.

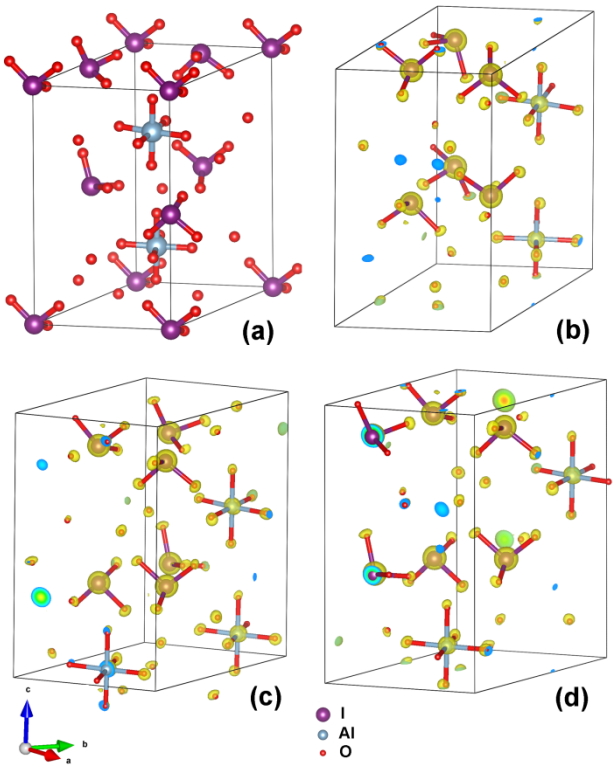

Figure S1. (a) The atomic structure model of $Al(IO_3)_3(H_2O)_8$ and the electron density images reconstructed with (b) the algorithm based on Tian1, (c) the error reduction and (d) charge flipping algorithms. The original atomic structure model was superimposed upon the electron density images to show the consistence between them. The atomic coordinates in the structure model are translational shifted and/or inverted to make a direct comparison with the electron density images in (b), (c) and (d).