Scattering from Biological Macromolecules

DALAL_GA Reconstruction of Solution Structures from X-ray Scattering using a Genetic Algorithm

EMBO Practical Course on Solution
**Reciprocal space**

\[ I(S) = 2 \int_0^\infty p(r) \frac{\sin(2\pi Sr)}{Sr} \, dr \]

**Real space**

\[ p(r) = \frac{1}{\pi} \int_0^\infty I(S) \, Sr \, \sin(2\pi Sr) \, dS \]
Maximal distance ($D_{\text{max}}$)
\[ p(r) = 0 \quad \text{para} \quad r \geq D_{\text{max}} \]

Radius of Giration ($R_g$)
\[ R_g^2 = \frac{\int_{0}^{\infty} p(r)r^2 \, dr}{2\int_{0}^{\infty} p(r) \, dr} \]
\[ I(S) = (\Delta n) \frac{2}{(4\pi S R_g)^2} e^{-\frac{(4\pi S R_g)^2}{2}} \]

Volumen ($V$)
\[ V = \frac{1}{4\pi} \frac{I(0)}{Q} \]
\[ Q = \int_{0}^{S_{\text{exp}}} I(S)S^2 \, ds + \frac{K_1}{S_{\text{exp}}} \]
Maximal distance of the particle

\[ D_{\text{max}} \leq \frac{1}{2S_{\text{min}}} \]

Resolution

\[ D_{\text{min}} \geq \frac{1}{2S_{\text{max}}} \]

Amount of information

\[ n_{\text{max}} \leq S_{\text{max}} \cdot 2 \cdot D_{\text{max}} \leq \frac{S_{\text{max}}}{S_{\text{min}}} \]
Debye equation

\[ I(S) = I_\circ(S) \left[ N + 2 \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\sin(2\pi S r_{ij})}{2\pi S r_{ij}} \right] \]

\[ I_\circ(S) = V^2 \Delta \rho^2 \left( \frac{3 \sin 2\pi S R - 2\pi S R \cos 2\pi S R}{(2\pi S R)^3} \right)^2 \]

Rayleigh formula
DALAI_GA1

Uses to generate uncontinuous models

DALAI_GA2

1) Adequate the $S$ range of the profile to the size of the model spheres $S_{\text{max}} = 1/4R$

2) Automatization of the mask strategy.

3) Introduce the spatial relationship in the genetic operators.
3D recombination restricted to model

Genetic Operators
3D Mutation restricted to model

Exchange or transposition

Genetic Operators
Synthetic profiles

\[ 2bb2 \quad \beta b2 - crystallin \]
\[ g2c \quad \delta - crystallin \]
\[ 1bnh \quad \text{Ribonuclease inh.} \]
\[ lyc \quad \text{lysozyme} \]
\[ 1rcm \quad 2xlysozyme \]

\[ \rightarrow \text{DALAI\_GA2} \]
Fitting of $\beta$-crystallin profile at 6, 5, 4, 3 and 2 Å
Evolution of $\rho_2$-crystallline fitting
Synthetic Profiles

Synthetic profiles + noise
Problems of the method

DEGENERATION
Problems of the method

ESTIMATION OF ERROR IN REAL SPACE

Implies superposition of structures at different resolution

SITUS ➔ http://www.scripps.edu/mb7wriggers/situs
[Wriggers et al., 1998; Wriggers and Schulten, 1999].

SUPERCOMB ➔ http://www.embl-hamburg.de/ExternalInfo/Research/Sax/supcomb.html
Experimental profiles

Scattering profiles of known structures measured by ourselves (Daresbury (SRS)) : myoglobin, psp-I/psp-II, Chymotripsinogen, ovoalbumin, catalase and tubulin

Scattering profiles of known structures kindly provided by other groups: superoxid dismutase, nitrite reductase (Dr. G. Grossmann, Daresbury Lab, UK) y troponin C (Prof. Fujisawa Spring 8, Japón).

Scattering profiles of unknown structures measured by ourselves : C-lyta, RepA- 133, y β4 fragment of α4β4 integrin
Common features of the obtained models

4 Their scattering profiles are coincident with the experimental ones (rms ~ noise). Their Rg values are within the experimental error with the crystallographic ones.

4 The number of spheres of the models can be correlated with the molecular mass of the crystallographic structures.

4 They converge to a given shape, which is compatible with the crystallographic structure, the degree of similarity holds on the resolution employed.

4 The volumes of the models are always larger than the Van der Waals volumes of the corresponding crystallographic structures.
Nitrite Reductase

Catalase
Unknown structures

RepA-Δ133

C-lyta
β₄ fragment of α₂β₄ integrin
de Pereda et al. 1999 EMBO, 18, 4087-95 PDB 1qg3
Future development

Implementation of automatic methods for evaluation of the models in real space. (Situs algorithm)

Improvement of search efficiency

Aplication

Aplication of the method to crystallographic data.
Aplication of the method to ordered polymers.
Servidor Web
http://akilonia.cib.csic.es/DALAI_GA2

CIB Madrid
- J.M Andreu
- P. Chacón
- J.M. De Pereda

DL, Warrington
- J. Bordas
- G. Diakun
- E. Pantos
- L. Towns-Andrews

UCM, Madrid
- F. Morán