Solvent model in bio-molecular scattering calculation

A review

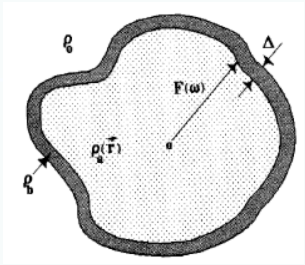
Background

- Scattering properties (scattering length, electron density) difference between solute, solvent, and solvent layer

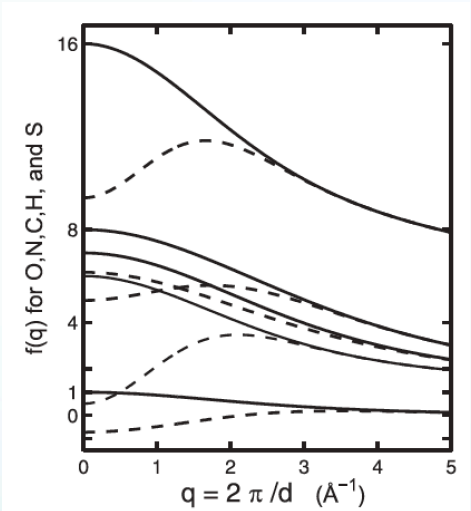
- $$I(q) = \langle |A_m(\mathbf{q}) - \rho_s A_s(\mathbf{q}) + \Delta\rho_b A_b(\mathbf{q})|^2 \rangle_\Omega$$

- Crysol $$I_{pred}(q) = \left\langle \left| \mathbf{F}_{mol} - \mathbf{F}_{disp} + \delta\rho \mathbf{F}_{surf} \right|^2 \right\rangle_\Omega$$

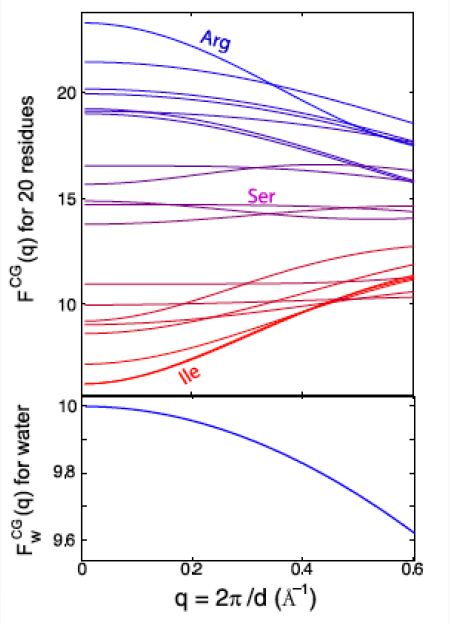
- Criticism of the implicit model



Structure factor correction

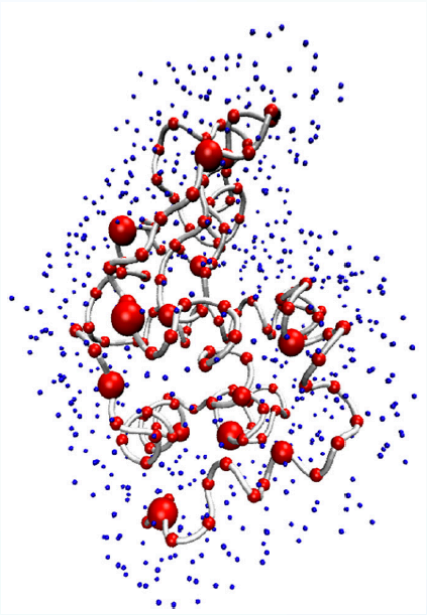


$$f_i'(q) = f_i(q) - v_i\rho_s \exp\left(-\pi v_i^{2/3} q^2\right)$$



$$F^{CG}(q) = \left\langle \sum_{i,i=1}^{n} f_i'(q) f_j'(q) \frac{\sin(qr_{ij})}{qr_{ij}} \right\rangle_{PDB}^{\frac{1}{2}}$$

$$F_w^{CG}(q) = \left[\sum_{i,j=1}^{3} f_i(q) f_j(q) \frac{\sin(qr_{ij})}{qr_{ij}} \right]^{\frac{1}{2}}$$



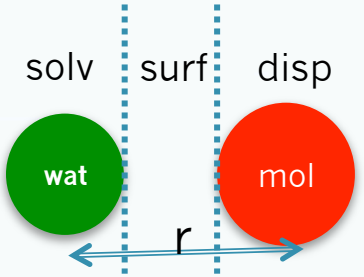
$$I^{CG}(q) = \sum_{i,j=1}^{N+M} F_i^{CG}(q) F_j^{CG}(q) \frac{\sin(qr_{ij})}{qr_{ij}}$$

A Rapid Coarse Residue-Based Computational Method for X-Ray Solution Scattering Characterization of Protein Folds and Multiple Conformational States of Large Protein Complexes. Sichun Yang et al, *Biophysical Journal*, 2009, **96**, 4449–4463

Solvent molecule assignment

- Solvent molecule assignment

  - $r < r_{vdw}$: displaced
  - $r_{vdw}+r_{wat} < r < r_{vdw}+r_{wat}+3$: surface
  - $r_{vdw} < r < r_{vdw}+r_{wat}$: $P_{surf} = \prod_j \left(\frac{r_j - r_{vdW}^j}{r_{wat}} \right)^\gamma$ *Avoid cavity*



Improved Fitting of Solution X-ray Scattering Data to Macromolecular Structures and Structural Ensembles by Explicit Water Modeling. Alexander Grishaev et. al., *J. AM. CHEM. SOC.* 2010, **132**, 15484–15486

Spherical method

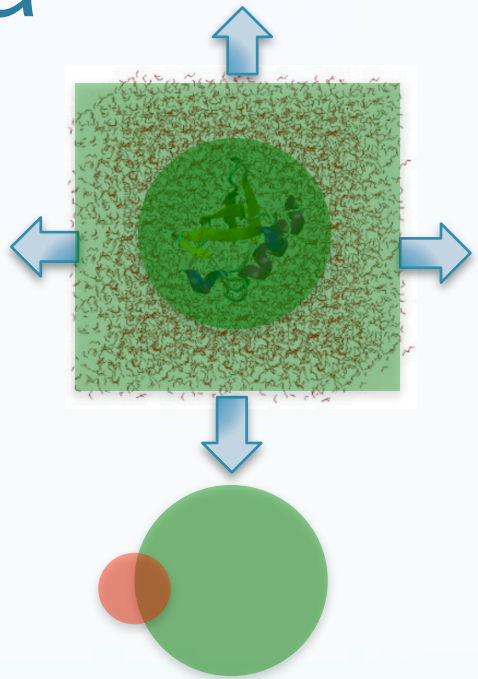
- Babinet's principle: expansion

$$f_{\text{outside}}(q) = -f_{\text{inside}}(q) + \delta(q)$$

- Boundary effects

$$F[f \otimes g] = F[f]F[g]$$

f/g: the electron density of the water/protein envelop sphere



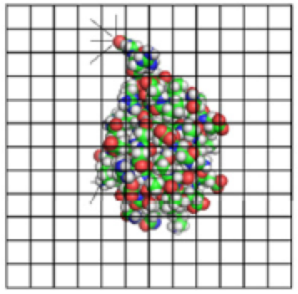
Modeling the Hydration Layer around Proteins: Applications to Small- and Wide-Angle X-Ray Scattering. Jouko Juhani Virtanen et al. *Biophysical Journal* 2011, **101** 2061–2069

Cube method

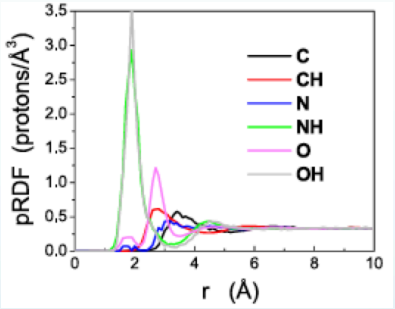
- Cube method

$$I(q) = \langle I(\mathbf{q}) \rangle = \left\langle \sum_{n=1}^{N} f_n(\mathbf{q}) e^{i\mathbf{q}\mathbf{r}_n} \sum_{m=1}^{N} f_m^*(\mathbf{q}) e^{-i\mathbf{q}\mathbf{r}_m} \right\rangle$$

$$f(\mathbf{q}) = 8\rho \frac{\sin\left(\frac{q_x a}{2}\right)\sin\left(\frac{q_y a}{2}\right)\sin\left(\frac{q_z a}{2}\right)}{q_x q_y q_z}$$



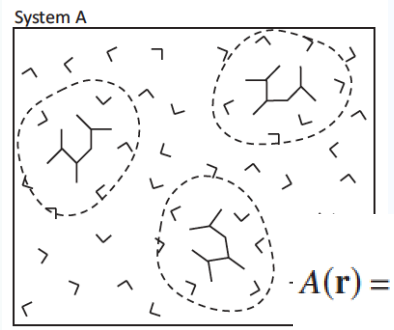
- HyPred



Modeling the Hydration Layer around Proteins: Applications to Small- and Wide-Angle X-Ray Scattering. Jouko Juhani Virtanen et al. *Biophysical Journal* 2011, **101** 2061–2069

Modeling the Hydration Layer around Proteins: HyPred. Jouko Juhani Virtanen et al. *Biophysical Journal* 2010, **99,** 1611–1619

Difference scattering



$$A(\mathbf{r}) = \sum_{n=1}^{N} A_n(\mathbf{r}) + A_0(\mathbf{r})$$

$$B(\mathbf{r}) = \sum_{n=1}^{N} B_n(\mathbf{r}) + B_0(\mathbf{r})$$

Assuming no electron density correlation between proteins

Assuming $\alpha_n(\mathbf{r},\mathbf{r}') = \beta_n(\mathbf{r},\mathbf{r}')$

$$\langle A_n(\mathbf{r})A_0(\mathbf{r}')\rangle = \langle A_n(\mathbf{r})\rangle\langle A_0(\mathbf{r}')\rangle + \alpha_n(\mathbf{r},\mathbf{r}')$$

$$D(\mathbf{q}) = \sum_{n=1}^{N} D_{nn}(\mathbf{q}).$$

$$D_{nk}(\mathbf{q}) := \langle \tilde{A}_n(\mathbf{q})\tilde{A}_k^*(\mathbf{q})\rangle - \langle \tilde{B}_n(\mathbf{q})\tilde{B}_k^*(\mathbf{q})\rangle - \langle \tilde{A}_n(\mathbf{q})\rangle\langle \tilde{B}_k^*(\mathbf{q})\rangle - \langle \tilde{B}_n(\mathbf{q})\rangle\langle \tilde{A}_k^*(\mathbf{q})\rangle + 2\langle \tilde{B}_n(\mathbf{q})\rangle\langle \tilde{B}_k^*(\mathbf{q})\rangle. \quad (22)$$

homogenous

$$\frac{\Delta I(q)}{N} = \frac{1}{4\pi}\int d\Omega_q D_{11}(\mathbf{q}),$$

$$D_{11}(\mathbf{q}) = |\langle \tilde{A}_1(\mathbf{q})\rangle - \langle \tilde{B}_1(\mathbf{q})\rangle|^2 + [\langle |\tilde{A}_1(\mathbf{q})|^2\rangle - |\langle \tilde{A}_1(\mathbf{q})\rangle|^2] - [\langle |\tilde{B}_1(\mathbf{q})|^2\rangle - |\langle \tilde{B}_1(\mathbf{q})\rangle|^2]. \quad (26)$$

$$\Delta I(q) = I_A(q) - I_B(q) \Rightarrow \Delta I(q) = \langle D(\mathbf{q})\rangle_p \qquad D(\mathbf{q}) := \langle |\tilde{A}(\mathbf{q})|^2\rangle_w - \langle |\tilde{B}(\mathbf{q})|^2\rangle_w$$

Simulated x-ray scattering of protein solutions using explicit-solvent models. Park *et al. J. Chem. Phys.* 2009, **130**, 134114

Difference scattering

$$\Delta I(q) = I_{\mathrm{A}}(q) - I_{\mathrm{B}}(q)$$

$$\Delta I_{\mathrm{scaled}}(q) = I_{\mathrm{A}}(q) - (1-f) I_{\mathrm{B}}(q)$$

f: volume fraction of protein in solution
(:=0 at low scattering angle)

$$I(q) = \sum_{i,j} f_i(q) f_j(q) [\Delta I_{ij}(q) + v I_{ij}(q)]$$

$$I_{ij}(q) = \delta_{ij}\rho_i + \rho_i \rho_j 4\pi \int_0^\infty r^2 [g_{ij}(r) - \gamma_{ij}] \frac{\sin(qr)}{qr} dr$$

$$\Delta I_{ij}(q) = \delta_{ij} \Delta N_i + \int_0^{2R} \Delta H_{ij}(r) \frac{\sin(qr)}{qr} dr$$

$$\Delta N_i = N_i - \rho_i V$$

$$\sim N^{0.33}$$

Atomic-resolution structural information from scattering experiments on macromolecules in solution. Jurgen Kofinger and Gerhard Hummer. PHYSICAL REVIEW E 87, 052712 (2013)

Summary

- Explicit vs Implicit
- Implicit solvent subtraction
- Experimental background subtraction
- Fully atomistic model
- SASSIE?
  - Sophisticated solvent layer construction
  - Scattering factor correction
  - Spherical or cube method?