# We are IntechOpen,
# the world's leading publisher of
# Open Access books
# Built by scientists, for scientists

## 4,800
Open access books available

## 122,000
International  authors and editors

## 135M
Downloads

## 154
Countries delivered to

Our authors are among the

## TOP 1%
most cited scientists

## 12.2%
Contributors from top 500 universities

**WEB OF SCIENCE**™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

# Interested in publishing with us?
# Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

# Simulated Annealing of Two Electron Density Solution Systems

Mario de Oliveira Neto[2], Ronaldo Luiz Alonso[1], Fabio Lima Leite[2], Osvaldo N. Oliveira Jr[2], Igor Polikarpov[2] and Yvonne Primerano Mascarenhas[2]

[1] *University of São Paulo, Institute of Mathematical Sciences and Computing, São Paulo*
[2] *University of São Paulo, Institute of Physics of São Carlos, São Carlos, São Paulo*
[1,2] *Brazil*

## 1. Introduction

Resolving the structure of a macromolecule such as a protein or synthetic polymer is important to predict its properties, especially in the case of proteins whose structure determines their function in a living cell. With crystallography and nuclear magnetic resonance (NMR) techniques, the protein structure may be solved at a high atomic resolution. However, these high resolution methods apply only in rather specific conditions, for low molecular weight proteins in NMR and when a crystal may be formed in crystallography. In order to obtain structural information for systems that do not satisfy the requirements above, one has to resort to methods such as X-ray and neutron small angle scattering (SAS), in which macromolecules in solution can yield only low-resolution information (from 1–100 nm), but are applicable to a broader range of conditions and particle sizes (Feigin and Svergun (1987)).

In this chapter we shall concentrate on small angle X-ray scattering (SAXS), and elaborate on the theory and possible applications for proteins and conjugated polymers. In order to obtain low resolution models of such complex systems, several assumptions must be made. For instance, the system needs to be monodisperse, diluted and with particles possessing an electron density that is considerably different from the density of the medium. Of particular relevance will be the use of global optimization techniques, such as simulated annealing, for proteins. The aim is to find the structure of a monodisperse diluted system of protein solution from one-dimensional SAXS data (Glatter and Kratky (1982)). The Chapter is organized as follows: In Section 2 the main principles of X-ray scattering are introduced, including scattering of X-ray by free electrons and a pair of electrons, in addition to scattering of X-ray by atoms and a group of $n$ atoms. Section 3 deals with analysis of SAXS curves, whereas the method of simulated annealing is discussed in Section 4. In Section 5 the results of simulated annealing in two electron densities systems are discussed. Concluding remarks close the chapter in Section 6.

## 2. Principles of small angle X-ray scattering

### 2.1 Scattering of X-ray by free electrons

The use of X-rays in structural characterization of materials explores essentially their interaction with matter through the electrons of atoms and molecules. The electrons are

sensitive to the sinusoidal *electric and magnetic fields* of the impinging X-rays, which for a monochromatic electromagnetic wave may be represented by periodic functions of the type:

$$\vec{E} = \vec{E}_0 \cos 2\pi\omega t$$
$$\vec{H} = \vec{H}_0 \cos 2\pi\omega t \tag{1}$$

where $\vec{E}_0$ and $\vec{E}$ are the incident and scattered electric fields, respectively, $\vec{H}_0$ and $\vec{H}$ are the incident and scattered magnetic fields, respectively, and $\omega$ is the radiation frequency.

The direction of X-ray propagation can be given by the *Poynting vector*: $\vec{P} = (c/4\pi)\vec{E} \times \vec{H}$.

Let us consider an electron located at the origin of a system of coordinates xyz shown in figure 1, on which a parallel beam of polarized and monochromatic X-rays impinges. Under the electric field of the incident wave the electron suffers a force:

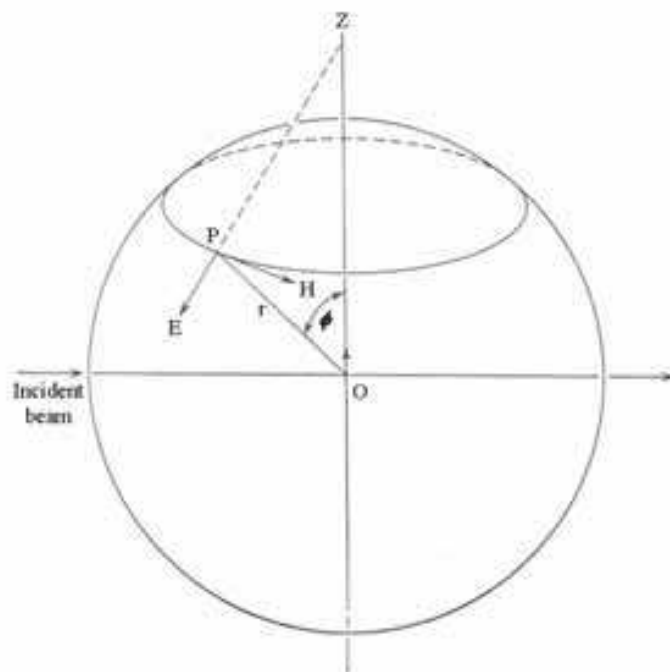$$\vec{F} = e\vec{E}_0 \cos 2\pi\omega t \tag{2}$$



Figure 1. Schematic view of the X-ray scattering by free electrons. An electron is located at the origin *O* on which a parallel beam of polarized, monochromatic X-rays is incident.

The electron is then accelerated according to:

$$\vec{a} = \frac{\vec{F}}{m} = \frac{e\vec{E}_0}{m} \cos 2\pi\omega t \tag{3}$$

The electron thus accelerated becomes a spherical source of electromagnetic waves of the same frequency of the incident radiation (elastic scattering). So, an electron will oscillate around its average position under this electric field. The electron scattering was analyzed by J. J. Thomson, who derived the field at *P* in figure 1:

$$E = \frac{ea\sin\phi}{4\pi R\varepsilon_0 c^2} \tag{4}$$

Substituting equation 3 in equation 4:

$$E = \frac{e^2 E_0}{m}\cos 2\pi wt \frac{\sin\phi}{4\pi R\varepsilon_0 c^2} \tag{5}$$

where $\phi$ is the angle between the scattered beam direction and the electric field of the incident wave, $\varepsilon_0$ is the vacuum permittivity, $R$ is the distance between the electron and the observation point, $c$ is the speed of light, $m$ is the electron mass and $e$ is the electron charge. The expression $\dfrac{e^2}{4\pi\varepsilon_0 c^2 m}$ has the dimension of length, being the scattering length in the classical electromagnetism, referred to as electron radius ($2.82.10^{-15}m$). To simplify the notation the amplitude of the scattered wave by a free electron will henceforth be designated as $A_e$.

$$A_e = \frac{e^2}{4\pi\varepsilon_0 c^2 m} \tag{6}$$

For a non-polarized incident beam, the electric field may be decomposed into two mutually perpendicular components. Figure 2 depicts a parallel, non-polarized beam propagating in the $Ox$ direction, impinging on an electron located at $O$. To determine the scattered electric field in the direction $OP$, the incident field can be decomposed in two components: a parallel one to the plane $OXP$, **E$_{//}$**, and a perpendicular component, **E$_\perp$**.
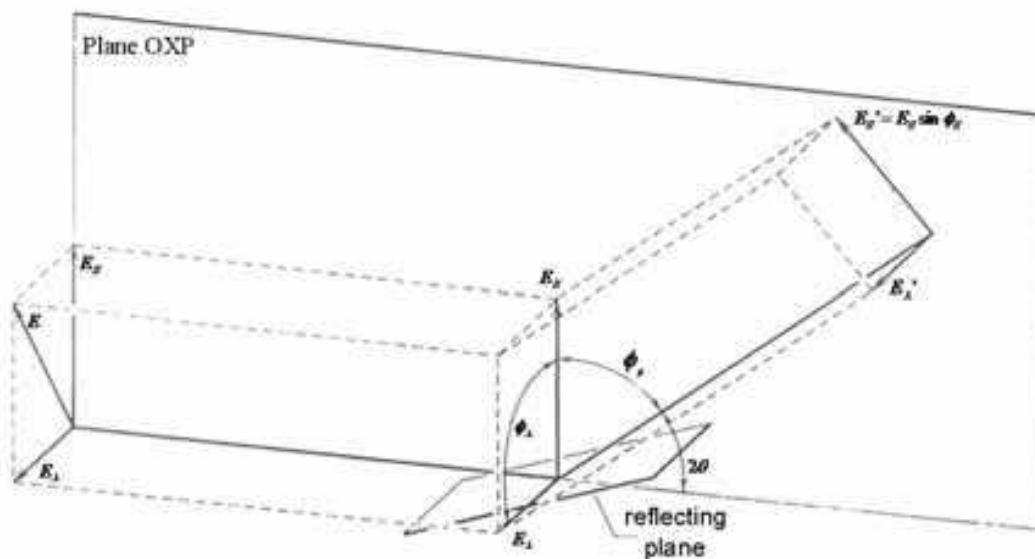


Figure 2. Schematic view of an electron located at the origin $O$ on which a parallel, monochromatic and non-polarized X-ray beam is incident in the $Ox$ direction. The incident field has a component parallel to the plane $OXP$, E$_{//}$, and another perpendicular to this plane, E$_\perp$.

Under the action of those fields the electron acquires acceleration with parallel and perpendicular components given by $a_\perp = eE_\perp / m$ and $a_{//} = eE_{//} / m$. Using these expressions in equation 5, one obtains the electric field of the scattered wave:

$$E_\perp ' = \frac{e^2 E_\perp}{4\pi\varepsilon_0 Rc^2 m} = \frac{A_e E_\perp}{R}$$

$$E_{//} ' = \frac{e^2 E_{//} \cos 2\theta}{4\pi\varepsilon_0 Rc^2 m} = \frac{A_e E_{//} \cos 2\theta}{R} \tag{7}$$

The intensity $I_o$ of the incident beam can also be decomposed into a perpendicular and a parallel component, proportional to $E_\perp^2$ and $E_{//}^2$, respectively:

$$E_{//}^2 = E_\perp^2 \propto \frac{I_0}{2} \tag{8}$$

The total intensity scattered at point $P$ is:

$$I(2\theta) = [E_{//}'^2 + E_\perp'^2]$$

$$I(2\theta) = \frac{I_0}{R^2}\left(\frac{e^2}{4\pi\varepsilon_0 c^2 m}\right)^2 \frac{(1+\cos^2 2\theta)}{2} \tag{9}$$

$$I(2\theta) = \left(\frac{I_0}{R^2}\right)(A_e)^2 \frac{(1+\cos^2 2\theta)}{2}$$

Since $I_0$ and $R$ can be taken as constant during the measurement, the scattered intensity of the electron is:

$$I_e(R, 2\theta) = A_e^2 \frac{(1+\cos^2 2\theta)}{2} \tag{10}$$

From equation 9 and 6 we infer that only the electrons contribute to the scattered intensity, which decreases with the square of the particle mass. Therefore, even though the nuclei also suffer the action of the impinging electric field, the effect is negligible because of the heavy mass of a nucleus (mass of the proton = 1837 times the mass of the electron). The term $(1 + \cos^2 2\theta)/2$ is the polarization factor, and indicates that the scattered wave is partially polarized even for a non-polarized incident beam. In SAXS experiments, the maximum scattering angle is usually around 5 degrees and $(1 + \cos^2 2\theta)/2$ is practically 1.

### 2.2 Scattering of X-ray by a pair of electrons, interference

To introduce the interference concept, we consider two electrons in a particle spaced by a distance $r$, immersed in a parallel monochromatic X-rays beam, as illustrated in figure 3a. The scattered waves will be coherent. Incoherent scattering may also occur but it can be neglected at small angles.
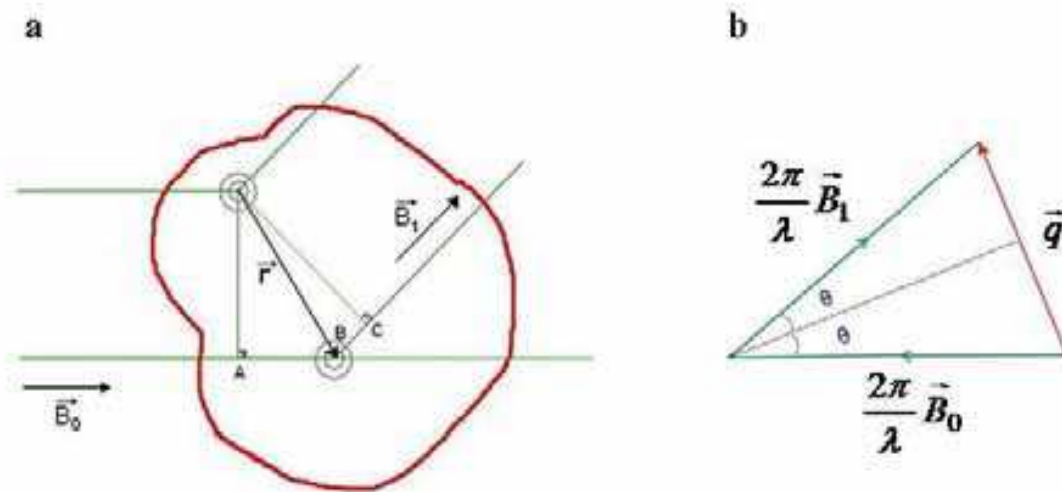
Figure 3. Schematic view for X-ray scattering by a pair of electrons. **a.** An electron is at the origin and the other is at a distance $\vec{r}$. There is a difference (AB + BC) in the optical paths between the two scattered beams, leading to a phase difference. **b.** Geometric relationships among the vectors $\vec{B}_1$, $\vec{B}_0$, and $\vec{q}$.

The phase difference due to the distinct optical paths for the electrons at the origin and at a distance $\vec{r}$ is

$$\Delta\varphi = \frac{2\pi\Delta l}{\lambda} \tag{11}$$

Since the difference in optical path $\Delta l = AB + BC$, the phase difference is

$$\Delta\varphi = \frac{2\pi(AB + BC)}{\lambda} \tag{12}$$

The directions of the incident and scattered beams are given by $\vec{B}_0$ and $\vec{B}_1$ unit vectors respectively.

$$AB = \vec{r} \cdot \vec{B}_0 \quad \text{and} \quad BC = -\vec{r} \cdot \vec{B}_1 \tag{13}$$

The phase difference is:

$$\Delta\varphi = \frac{-2\pi\vec{r}(\vec{B}_1 - \vec{B}_0)}{\lambda} \tag{14}$$

Where:

$$\vec{q} = \frac{2\pi(\vec{B}_1 - \vec{B}_0)}{\lambda} \tag{15}$$

Then:

$$\Delta\varphi = -\vec{r}\cdot\vec{q} = -rq\cos 2\theta \tag{16}$$

where $2\theta$ is the angle between $\vec{r}$ and $\vec{q}$.

Only the product of the components $\vec{r}\cdot\vec{q}$ is relevant to $\Delta\varphi$. It might now be possible to obtain the resulting amplitude by summing up all secondary waves, considering the scattering phase factor $e^{i\Delta\varphi}$ among them. Due to the enormous number of electrons in a macromolecule, it is convenient to introduce the electron density $\rho(\vec{r})$. The scattering amplitude will be then given by:

$$A(\vec{q}) = \int \rho(\vec{r})A_e e^{i\Delta\varphi} \tag{17}$$

Substituting Equation 16 into 17

$$A(\vec{q}) = \int A_e \rho(\vec{r})e^{-i\vec{r}\cdot\vec{q}}dV \tag{18}$$

Figure 3b displays the geometric relationships among the vectors $\dfrac{2\pi\vec{B_1}}{\lambda}$, $\dfrac{2\pi\vec{B_0}}{\lambda}$, and $\vec{q}$. It can be seen easily that $\vec{q}$ is perpendicular to the bisector of the angle between $\dfrac{2\pi\vec{B_1}}{\lambda}$ and $\dfrac{2\pi\vec{B_0}}{\lambda}$, with a module equal to:

$$q = \frac{4\pi\sin\theta}{\lambda} \tag{19}$$

## 2.3 Reciprocity law

Any scattering process is characterized by a reciprocity law that gives an inverse relationship between particle size and scattering angle. Let us consider in figure 4 the case of two spheres with constant electron density $\rho(\vec{r})$ and radius $R_1$ and $R_2$ under the same X-ray beam; $dI$ is the scattered intensity for an element of volume $dV$ for two different points from each sphere. There is a difference in optical path to the two spheres. If $2\theta$ is zero, so is $q$, and $\Delta\varphi$ for all the volume elements is null. The same path difference of one $\lambda$ occurs at a higher angle for the smaller sphere. The scattering curves for the two spheres clearly reveal a reciprocity relationship between the real and the scattering spaces, referred to as reciprocal space.

## 2.4 The phase problem

The scattering intensity $I(\vec{q})$ is the measurable quantity, proportional to the square of $A(q)$:

$$I(\vec{q}) \propto A(\vec{q})\cdot A(\vec{q})^* \tag{20}$$

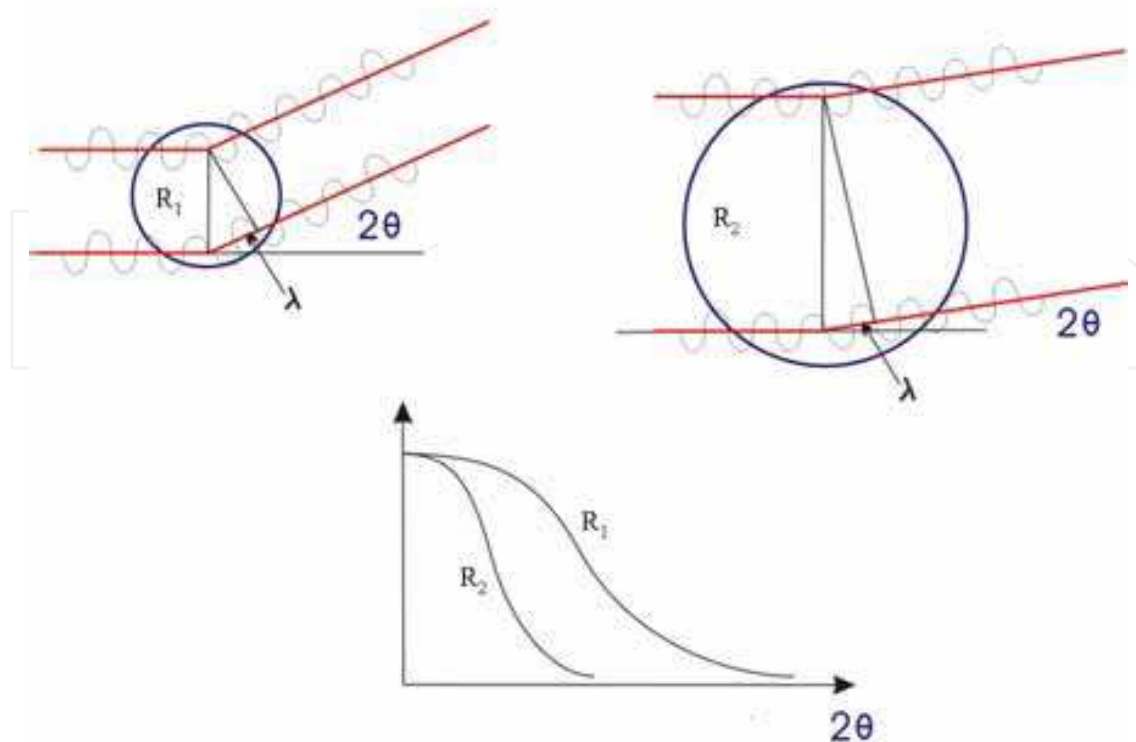where $A(\vec{q})^*$ it is the complex conjugate of $A(\vec{q})$.

Figure 4. **Reciprocity law.** Two spheres with constant electron density $\rho(\vec{r})$ and radius $R_1$ and $R_2$ under the same X-ray beam.

According to the properties of Fourier transforms, the inverse transform of $A(\vec{q})$, equation 18, allows for the calculation of $\rho(\vec{r})$, which is a real, positive quantity. However, to obtain $\rho(\vec{r})$ the phase and module of the wave vector $A(\vec{q})$ are required:

$$A(\vec{q}) = \left| A(\vec{q}) \right| e^{i\phi} \tag{21}$$

The intensity will be then:

$$I(\vec{q}) \propto \left( A(\vec{q}) \right)^2 = \left| A(\vec{q}) \right|^2 e^{i\phi} e^{-i\phi} = \left| A(\vec{q}) \right|^2 \tag{22}$$

Hence, as from the measurement only $I(\vec{q})$ is obtained the knowledge of the phase being lost.

### 2.5 Scattering of X-rays by atoms

Equation 18 can be calculated for any atom, provided that $\rho(\vec{r})$ is known, which is usually obtained with quantum methods assuming spherical symmetry. For $\vec{q} = 0$ the integral in equation 18 gives the total number of electrons of a neutral atom, *i.e.*, its atomic number Z. As $\vec{q}$ grows, the phase differences lead to a progressive decrease in $A(\vec{q})$. Curves of the type shown in figure 5 are obtained for carbon and oxygen.
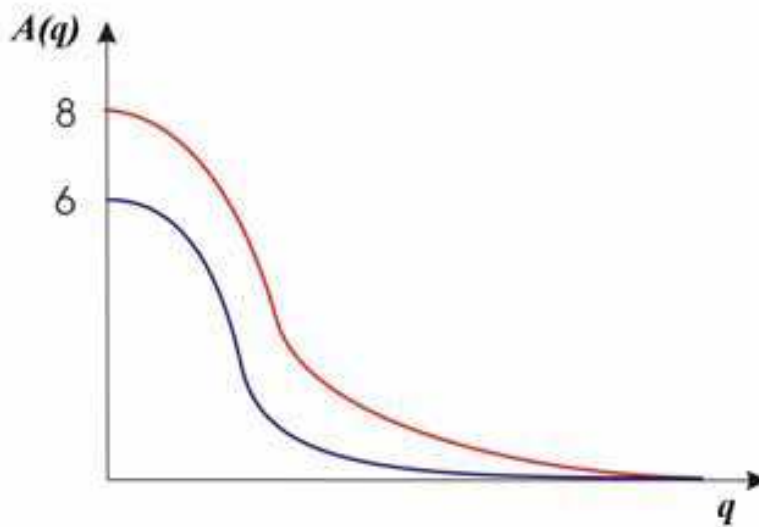
Figure 5. Schematic representation of the atomic scattering factor for carbon (blue line) and oxygen (red line).

The atomic scattering factor is defined as the ratio between the scattered amplitude for the atom and the scattered amplitude for one electron.

$$f(\vec{q}) = \frac{A(\vec{q})_{atom}}{A_e} \tag{23}$$

### 2.6 Scattering by a group of n atoms

With the atomic scattering factors, equation 18 can be rewritten considering the position vector $\vec{r}_j$ of each atom, and making a sum over all the n atoms, each one with a scattering factor $f_j$:

$$A(\vec{q}) = \sum_{j=1}^{n} f_j e^{-i\vec{q}\cdot\vec{r}_j} \tag{24}$$

The scattered intensity of the particle can be obtained if the atomic coordinates of the atoms are known. For n atoms with scattering factors $f_j(\vec{q}) = f_j$ and coordinates $\vec{r}_j$ for the center of each particle:

$$I(\vec{q}) = \sum_{j=1}^{n} \sum_{k=1}^{n} f_j f_k e^{-i\vec{q}\cdot\vec{r}_{jk}} \tag{25}$$

### 2.7 Inverse Fourier transform of the intensity

Because of the Phase Problem, it is not possible to invert $A(\vec{q})$ directly to obtain $\rho(\vec{r})$. Patterson (1935) proposed a method for calculating the Fourier transform taking as coefficients the intensities of scattered beams, obtained from the square of the amplitude in equation 18:

$$I(\vec{q}) = \int\int dV_1 dV_2 \rho(\vec{r_1})\rho(\vec{r_2})e^{-i\vec{q}(\vec{r_1}-\vec{r_2})} \tag{26}$$

Integration of equation (26) may be performed in two steps: first integrating over all the pairs of points of equal distance in the particle and then integrating on all relative distances in the particle. These steps are summarized as follows:

1st step: Mathematically it corresponds to calculating the auto-correlation function $\tilde{\rho}(\vec{r})$, making $\vec{r} = \vec{r_2} - \vec{r_1}$ constant.

$$\tilde{\rho}(\vec{r}) = \rho(\vec{r_1})\rho(\vec{r_2})dV \tag{27}$$

This function, known as Patterson function, is defined on a new space $C(r)$ (correlation space) in which each $\vec{r}$ corresponds to a distance $(\vec{r_2} - \vec{r_1})$ taken over the whole scattered object and whose value is the average of the products between electron densities in points $\vec{r_1}$ and $\vec{r_2}$.

2nd step: It consists in integrating the remaining function in the space $C(r)$ to obtain $I(\vec{q})$:

$$I(\vec{q}) = \int dV \tilde{\rho}^2(\vec{r})e^{-i\vec{q}\cdot\vec{r}} \tag{28}$$

This integral depends only on the relative distances between the elements in the volume and on the product of the corresponding electron densities.

For a concentrated system, the scattering intensity is $I_*(\vec{q}) = I(\vec{q}).S(\vec{q})$, where $I(\vec{q})$ is the form factor, and should be related to intra-particles distances, while $S(\vec{q})$ is the structure factor, associated with inter-particles correlations. For diluted systems, $S(\vec{q})$ can be taken as 1, and only the form factor contributes to the scattering curve (Craievich (2005)). In this chapter we will discuss only diluted systems.

According to equation 28, the distribution of the scattered intensities in reciprocal space is a function of electron density distribution in the scattering particle through its auto-correlation function. It can be obtained directly from the scattered intensities by calculating the inverse Fourier transform:

$$\tilde{\rho}^2(\vec{r}) = (\tfrac{1}{2\pi})^3 \int I(\vec{q})dV^* e^{-i\vec{q}\cdot\vec{r}} \tag{29}$$

where $dV^*$ is the volume element of the reciprocal space.

An important property of the auto-correlation function is the existence of an inversion center, because $\rho(\vec{r_1})\rho(\vec{r_2})$ is equal $\rho(\vec{r_2})\rho(\vec{r_1})$, resulting in the same value for $\vec{r}$ as for $-\vec{r}$, regardless of whether the scattering particle has an inversion center. Furthermore, there is a reciprocity relationship between equations 28 and 29, because their values just depend on the product $\vec{q}\cdot\vec{r}$. Therefore, an increase in $\vec{r}$ leads to a smaller $\vec{q}$.

### 2.8 Isotropic and diluted systems

A consequence of the reciprocity relationships between the direct and reciprocal spaces is that large particles will lead to smaller intervals for 2θ in which scattering is observed. For analyzing SAXS curves we need to adopt simplifying restrictions:

(a) The system is statistically isotropic either due to the particle scattering form, or to its spatial distribution and even random movement in the medium.

(b) There is no long range order among the particles, i.e., they should be sufficiently apart.

With restriction (a), $\tilde{\rho}^2(\vec{r})$ is centro-symmetric depending only on the module of $\vec{r}$. This may not be true in the real space, and we can substitute the phase factor $e^{-i\vec{q}\cdot\vec{r}}$ by its average value taken around $\vec{r}$, that is, applying Debye's formula:

$$\left\langle e^{-i\vec{q}\cdot\vec{r}} \right\rangle = \frac{\sin qr}{qr} \tag{30}$$

This allows us to rewrite equations 25 and 29:

$$I(q) = \sum_{i=1}^{n}\sum_{j=1}^{n} f_i f_j \frac{\sin qr_{ij}}{qr_{ij}} \tag{31}$$

$$I(q) = \int dV \, \tilde{\rho}^2(r) \frac{\sin qr}{qr} \tag{32}$$

Again, the expression of reciprocity between $r$ and $q$ is apparent in Debye's formula. If the product $qr$ is kept constant, an increase in $r$ causes a decrease in $q$.

Because $\sin qr$ is a periodic function, the denominator $qr$ is a dampening factor, generating smaller maxima. The first zero should happen for $qr = 2\pi$. We can thus verify, recalling that $q = (4\pi\sin\theta)/\lambda$, that $\theta \approx 2.5^\circ$ for r = 50 Å and λ = 1.54 Å. Therefore, the useful interval for measuring the scattering curve is approximately from 0 to 2.5º.

The consequence of the second restriction is that at large $r$ the electron densities become independent, and might be replaced by the average $\langle\rho\rangle$. The auto-correlation function must tend towards a constant value $V\langle\rho\rangle^2$, while at the origin $\tilde{\rho}^2(0)$ takes the value $V\overline{\rho^2}$ (the maximum, of course). So, in space $C(r)$ relevant structural information will be only in the area such that ρ is significantly different from its final constant value. For that reason the auto-correlation function is redefined to $\eta = \rho - \langle\rho\rangle$, expressing in fact only the fluctuations of electron density responsible for scattering:

$$\tilde{\eta}^2(r) = (\rho(r) - \langle\rho\rangle)^2 = V\gamma(r) \tag{33}$$

With a convenient change of notation, the correlation function $\gamma(r)$ is introduced, which is associated with the average of the density fluctuations for two electrons separated by $r$, where $r = |r_1 - r_2|$.

$$\gamma(r) = \left\langle \eta(r_1)\eta(r_2) \right\rangle \tag{34}$$

The fluctuations in electron density, relative to the medium that contains the scattering centers, can be negative or positive. For instance, pores in a material lead to a negative $\eta$

fluctuation. However $\gamma(r)$ will be positive as it is given (cf. eq. 33) by $\tilde{\eta}^2(r)$, with $\gamma(0) = \overline{\eta^2}$ and $\gamma \to 0$ for large $r$. Adding $\gamma(r)$ in eq. 32, with integration limits from 0 to $\infty$ and changing the integration parameter to $dr$, the intensity is:

$$I(q) = V \int_0^\infty 4\pi r^2 dr \gamma(r) \frac{\sin qr}{qr} \tag{35}$$

For $q = 0$

$$I(0) = V \int_0^\infty 4\pi r^2 dr \gamma(r) \tag{36}$$

$\gamma(r)$ is found from the inverse Fourier transform of eq. 35:

$$V\gamma(r) = \frac{1}{2\pi^2} \int_0^\infty q^2 dq I(q) \frac{\sin qr}{qr} \tag{37}$$

For $r = 0$

$$V\gamma(0) = \frac{1}{2\pi^2} \int_0^\infty q^2 dq I(q) = V\overline{\eta^2} \tag{38}$$

It is not possible to measure $I(0)$ in equation 36 because it coincides with the incident beam direction. However, it can be obtained by extrapolation of the curve I($q$), thus allowing at least an estimate of the number of electrons in the volume of the particle.

From equation 38 the integral of the intensity in the reciprocal space is constant. Even if a given particle has its form altered, but remaining as a whole intact, the integral is constant, equal to $V\overline{\eta^2}$, though the diffraction pattern or scattering may be changed. This constant is the so-called "invariant", given by:

$$Q = \int_0^\infty q^2 dq I(q) \tag{39}$$

The restricting condition of isotropic systems allows one to obtain averages that are treated as scalar quantities in the distribution of the autocorrelation function in the C($r$) space.

## 2.9 Redefinition of correlation function

Because the square electron density difference is always positive and constant, it is convenient to separate $\gamma(r)$, defined in equation 33, in the form

$$\gamma(r) = (\Delta\rho)^2 \gamma_0(r) \tag{40}$$

Where $\gamma_0(r)$ is a new correlation function, just for the geometry of the particle, with $\gamma_0(0) = 1$ and $\gamma_0(r \geq D_{max}) = 0$ (where $D_{max}$ is the maximum intra-particle distance). $\gamma_0(r)$ is called "characteristic function" and has a more intuitive meaning.

Thus, it is convenient to rewrite equations 35, 36, 37, 38 and 39:

$$I(q) = V\left(\Delta\rho\right)^2 \int\limits_0^D 4\pi r^2 dr \gamma_0(r) \frac{\sin qr}{qr} \tag{41}$$

$$I(0) = V\left(\Delta\rho\right)^2 \int\limits_0^D 4\pi r^2 dr \gamma_0(r) \approx \left(\Delta\rho\right)^2 V^2 \tag{42}$$

$$V\gamma_0(r) = \frac{1}{2\pi^2 \left(\Delta\rho\right)^2} \int\limits_0^\infty q^2 dq I(q) \frac{\sin qr}{qr} \tag{43}$$

$$V\gamma_0(0) = \frac{1}{2\pi^2 \left(\Delta\rho\right)^2} \int\limits_0^\infty q^2 dq I(q) \tag{44}$$

$$Q = \int\limits_0^\infty q^2 dq I(q) = V 2\pi^2 \left(\Delta\rho\right)^2 \tag{45}$$

### 2.10 Isotropic, diluted and monodisperse systems

Let us assume that the system of interest is a diluted solution of identical particles (monodisperse) with constant electron density $\rho$, embedded in a medium of constant $\rho_0$ *(solvent)*. Thus only $\Delta\rho = (\rho - p_0)$ is relevant for scattering. The condition of diluted system guarantees that each particle makes independent contributions to the scattering intensity, so that only one single particle needs to be considered.

A more complete data analysis for the determination of the particle's geometry can still be made through the calculation, starting from the experimental data, of the pair distribution function *p(r)*, where $p(r) = 4\pi r^2 \gamma_0(r)$. Thus, equation 41 can be rewritten as

$$I(q) = V\left(\Delta\rho\right)^2 \int\limits_0^D p(r) dr \frac{\sin qr}{qr} \tag{46}$$

Therefore, *p(r)* can be obtained from the inverse Fourier transform of *I(q)*. The *p(r)* function is zero for the maximum particle dimension $D_{max}$. It should be reminded that we assumed restrictive conditions such that "the solution needs to be monodisperse and sufficiently diluted to avoid inter-particle effects". It is also worth noting that a good contrast in electron density is needed between the solute and the solvent. The distance distribution function *p(r)* contains the same information as the scattering intensity *I(q)*, but the real space representation is more intuitive. Furthermore, information about particle shape can often be deduced by straightforward visual inspection of *p(r)*. Spherical particles have a Gaussian p(r) with maximum at $D_{max}/2$. Departures from a Gaussian curve are indicative of more anisotropic particles in solution.

## 3. Analysis of SAXS curves

To analyze the SAXS curve it is convenient to distinguish three regions related to different distances in real space. With this procedure, it is possible to calculate the radius of gyration, the relation *I(0)/Q* and the surface/volume ratio.

### 3.1 Small *q* in SAXS curves, determination of the radius of gyration

Assuming the ideal case of non-interacting, dilute spherical particles and isotropic solutions, A. Guinier showed for $q \to 0$ that the intensity curve can be described by an exponential function:

$$I(q) = I(0)e^{\frac{-q^2 R_g{}^2}{3}} \tag{47}$$

Where $R_g$ is the radius of gyration corresponding to the quadratic average distance from the electron to the center of gravity of the electron density, analogously to the radius of inertia in mechanics. Similar approximations, not shown here, can be considered for rod-like and flat particles. For ideal monodisperse systems, the Guinier plot $\ln I(q) \times q^2$ should be a straight line whose intercept gives *I (0)* and the slope yields the radius of gyration $R_g$. One should, however, always bear in mind that the Guinier approximation is valid for very small angles only, namely in the range $q < 1.3/ R_g$, and fitting a straight line beyond this range is unphysical. It is also possible to obtain $R_g$ as the normalized second moment of the pair distribution function *p(r)* of the particle (Svergun & Koch (2003)).

### 3.2 Central slope of the SAXS curve, determination of the volume

Dividing equation 42 by 45, the term corresponding to the absolute intensity is canceled out and:

$$\frac{I(0)}{Q} = \frac{V}{2\pi^2} \tag{48}$$

Because the data normally appear in arbitrary units, *I*(0) and Q are given in a relative scale. Hence, $I(0)/Q$ is also valid in an arbitrary scale. To obtain V the data are extrapolated to $I(0)$ using the Guinier plot.

### 3.3 Final slope of the SAXS curve, determination of the surface/volume ration

An analysis of the slope in the final region of the SAXS curve should contain information on finer aspects of the particle's structure, expressed by the behavior of $\gamma_0(r)$ at smaller $r_s$. Porod showed that a relationship exists between this part of the curve and the fourth power of *q* given by:

$$I(q) \to (\Delta\rho)^2 V \frac{8\pi}{q^4} = (\Delta\rho)^2 \frac{2\pi}{q^4} S \tag{49}$$

The asymptotic value for the curve *I(q) x q⁴* is expected to be proportional to the total surface of the particle:

$$\frac{S}{V} = \frac{\pi}{Q} \lim_{q\to\infty} q^4 I(q) \tag{50}$$

The data at large angles are assumed to follow a linear plot in $q^4 I(q)$ against $q^4$ coordinates. Nevertheless, if there are heterogeneities in electron densities in the scattering particles a relation of the form $q^4 I(q) \approx B q^4 I(q) + A$ may appear. By subtracting the constant *B (Porod constant)* from *I (q)*, the scattering corresponds to that of a homogeneous body.

In summary, three characteristic parameters of the particle are obtained, viz. radius of gyration, volume and surface, which can be used to design a possible low resolution model for the scattering particle.

## 4. Simulated annealing

The first paper suggesting the use of simulated annealing for minimization of a function with no obvious physical correspondence was the Kirkpatrick procedure for minimizing printed circuit board line crosses (Kirkpatrick *et al.*, (1983)). This paper constructs a non physical function based on crossing circuits and chooses a random line to uncross, as the equivalent of a physical atom. The situation in this case is quite different from the true simulated annealing, because the function used has no physical correspondence, but the ideas work in the same way. The simulated annealing algorithm in this case may be described by the following macro procedure (Svergun (1999)):

Start from a random configuration $X_0$ at a high temperature $T_0$. In this case $T_0$ may be a function of $X_0$.

Select an atom at random, randomly change its phase (configuration $X'$), and compute $\Delta = f(X') - f(X)$.

If $\Delta < 0$, move to $X'$

Else if $\exp(-\Delta/T) < random$

move to $X'$, else continue on $X$.

Hold T constant for 100 N reconfigurations or 10 N successful reconfigurations, whichever comes first, then cool the system ($T' = 0.9T$)

Continue cooling until no improvement in $f(X)$ is observed.

As the temperature decreases, these modifications become less random and sharper because the system is freezing. Note that only one dummy atom is changed per move so that only a single summand in equation must be updated to calculate the partial amplitudes. This summand is the most time consuming operation. It is exactly this acceleration that makes it possible to use simulated annealing, because it causes the evaluation of f(x) to be N times faster.

### 4.1 *Ab initio* reconstruction based on simulated annealing

The reconstruction of a three-dimensional model of an object from its one-dimensional scattering pattern is not easy. In addition, its uniqueness is not guaranteed, as different models may yield the same SAS curve with nearly the same accuracy (Vladimir *et al.*, (2003)). To simplify the description of the low-resolution models that can legitimately be obtained, data interpretation is often performed in terms of homogeneous bodies (the influence of internal inhomogeneities for single component particles can largely be eliminated by subtracting the Porod constant). In the past, shape modeling was done by trial-and-error using Debye's formula, computing scattering patterns from different shapes and comparing them with the experimental data. The models were either three-parameter geometrical bodies like prisms, triaxial ellipsoids, elliptical or hollow circular cylinders, etc, or shapes built from assemblies of regularly packed spheres (beads). The first *ab initio* shape determination method was proposed by Stuhrmann (1970). The particle shape was represented by an angular envelope function *r = F(ω)* describing the particle boundary in spherical coordinates *(r, ω)*. The use of the angular envelope function was, however, limited to relatively simple shapes (in particular, without holes inside the particle).

A more comprehensive description is achieved with the bead methods (Chacon *et al.,* (1998), Svergun (1999)). Such an approach uses the tremendous power of modern computers, and is based on the same idea used in the past for the trial and error with Debye's formula. Initially a spherical region with diameter $D_{max}$ is filled with N subunits (spheres) with hexagonal packing. Each of these subunits belongs either to the particle (index=1) or to the solvent (index=0). The geometric form composed by these subunits can be viewed as a vector with N components, and each of the components is either a zero or one. This model is known as the dummy atom model (DAM). The idea is to randomly modify this model by a Monte-Carlo procedure for obtaining a chain, i.e., a geometric configuration, for which the simulated scattering curve fits the experimental data. This approach was implemented on Dammin program, which works as follows: A model of a *K*-phase particle $K \geq 1$ is generated and its scattering is calculated. The next step is to define a spherical shape enclosing the particle. This corresponds to the step cited above, in which a sphere with diameter $D_{max}$ is defined and filled up with N dummy spheres ($N \approx (R/r_0)^3$, where $R$ is sphere radius and $r_0$ is the dummy atom radius.

Each dummy atom is assigned an index $X_j$ indicating the phase to which it belongs ($X_j$ ranges from 0 (solvent) to K). Given the fixed atomic positions, the shape and structure of the dummy atom model are completely described by a phase assignment vector *X* (configuration).

The dummy atoms of the k-th phase are assumed to have contrast $\Delta\rho_k$, and the scattering intensity from the Dummy Atom Model (DAM) is:

$$I(q) = \left\langle \left[ \sum_{k=1}^{K} \Delta\rho_k A_k(q) \right] \right\rangle_{\Omega} \tag{51}$$

where $A_k(q)$ is the scattering amplitude from the volume occupied by the $k$th phase. The scattering amplitude in the formula above can be given in terms of spherical harmonic functions $Y_{lm}(\Omega)$ as:

$$A_k(q) = \sum_{l=0}^{\infty} \sum_{m=-1}^{1} A_{lm}^{(k)}(q) Y_{lm}(\Omega) \tag{52}$$

The terms $A_{lm}^{(k)}(s)$ are obtained by mathematical manipulation, i.e., rearrangement of terms, given by:

$$A_{lm}^{(k)}(q) = i^l \sqrt{2/\pi} f(q) \sum_{j=1}^{N_k} j_1(qr_j) Y_{lm}^*(\omega_j) \tag{53}$$

The intensity is (Stuhrmann (1970b)):

$$I(q) = 2\pi^2 \sum_{l=0}^{\infty} \sum_{m=-1}^{1} \left[ \sum_{k=1}^{K} [\Delta\rho_k A_{lm}^{(k)}(q)]^2 + 2\sum_{n>k} \Delta\rho_k A_{lm}^{(k)}(q) \Delta\rho_n [A_{lm}^{(n)}(q)]^* \right] \tag{54}$$

where the sum runs over the dummy atoms of the $k$th phase, $r_j$, $v_j$ are their polar coordinates, $j_1(x)$ is the spherical Bessel function, and $f(q)$ is the scattering from a single atom (form factor).

The looseness criterion is applied to a set of $M \geq 1$ experimental curves $I_{\exp}^{(i)}(q), i = 1,...,M$, and the procedure in the Dammin program tries to minimize the discrepancy:

$$\chi^2 = \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{N(i)}[(I_{\exp}^{(i)}(q_j) - I^{(i)}(q_j))/\sigma(q_j)]^2 \tag{55}$$

In the formula above, $N(i)$ is the number of points of the i-th curve and $\sigma(s)$ denotes the experimental errors. For an adequate description of a structure the number of dummy atoms must, however, be as large as the number of true atoms ($N \approx 10^3$). On the other hand, if the resolution is low, the uniqueness of such a model cannot be meaningfully discussed.

The program assumes an hexagonal packing $N_c = 12$, except for the border atoms. The connectivity is defined by an exponential form $C(N_e) = 1 - P(N_e) = 1 - [e^{-0.5N_e} - e^{-0.5N_c}]$, where $C(12) = 1$ for ideal connectivity, and smaller values for $N_e < 12$ are assumed to emphasize loosely connected dummy atoms. The compactness of a given configuration $X$ can be computed as an average connectivity of all nonsolvent atoms $\langle C(N_e) \rangle$. Then, a configuration is characterized by the average looseness $P(X) = 1 - \langle C(N_e) \rangle$. The final step is to define a function to be minimized and to run the simulated annealing procedure. In this way, the Dammin program adds a penalty term *P(X)* and the function to be minimized becomes:

$$f(X) = \chi^2 + \alpha P(X) \tag{56}$$

where $\alpha > 0$, is the weight of the looseness penalty. The purpose of the penalty term is to guarantee the compactness of the resulting form.

## 5. Results of simulated annealing in two electron density solution systems

For the application of the simulated annealing the studied system must be monodisperse, diluted and with basically two electron densities. ($<\rho>$- $\rho_0$), were $<\rho>$ is the average electron density of particles in solution and $\rho_0$ is the solvent electron density. The models were generated by the simulated annealing procedure implemented with the Dammin program (Svergun (1999)). Other programs such as Gasbor (Svergun *et al.*, (2001)) may be used to evaluate models by simulated annealing. To exemplify the usefulness of the simulated annealing application, some *ab initio* three-dimensional models of proteins in solution generated from SAXS data were chosen, according to Figueira *et al.*, (2007). Determination of the molecular shapes and oligomeric forms of the thyroid hormone nuclear receptor (TR) by SAXS can be shown as an example. Thyroid hormone (TH) plays important roles in cell differentiation, growth, and metabolism and is a major regulator of mitochondrial activity. In its physiologically most relevant form of triiodothyronine (T3), TH exerts most of its effects by binding to thyroid hormone receptors (TRs), which are members of the nuclear receptor (NR) family of transcription factors. Crystallographic structures of separate DNA and ligand binding domains (DBD and LBD) of TR have yielded significant insights into TR action but up to now no crystallographic structures of the complete TR structure or even of the construct containing both DBD and LBD were resolved. Low-resolution X-ray structures of the isoform *β* of the TR DBD-LBD were reconstructed form SAXS data measured in solution. SAXS data (figure 6), the overall parameters (table) and simulated annealing modeling (figure 7) reveal significant changes in the oligomeric state of the receptor,

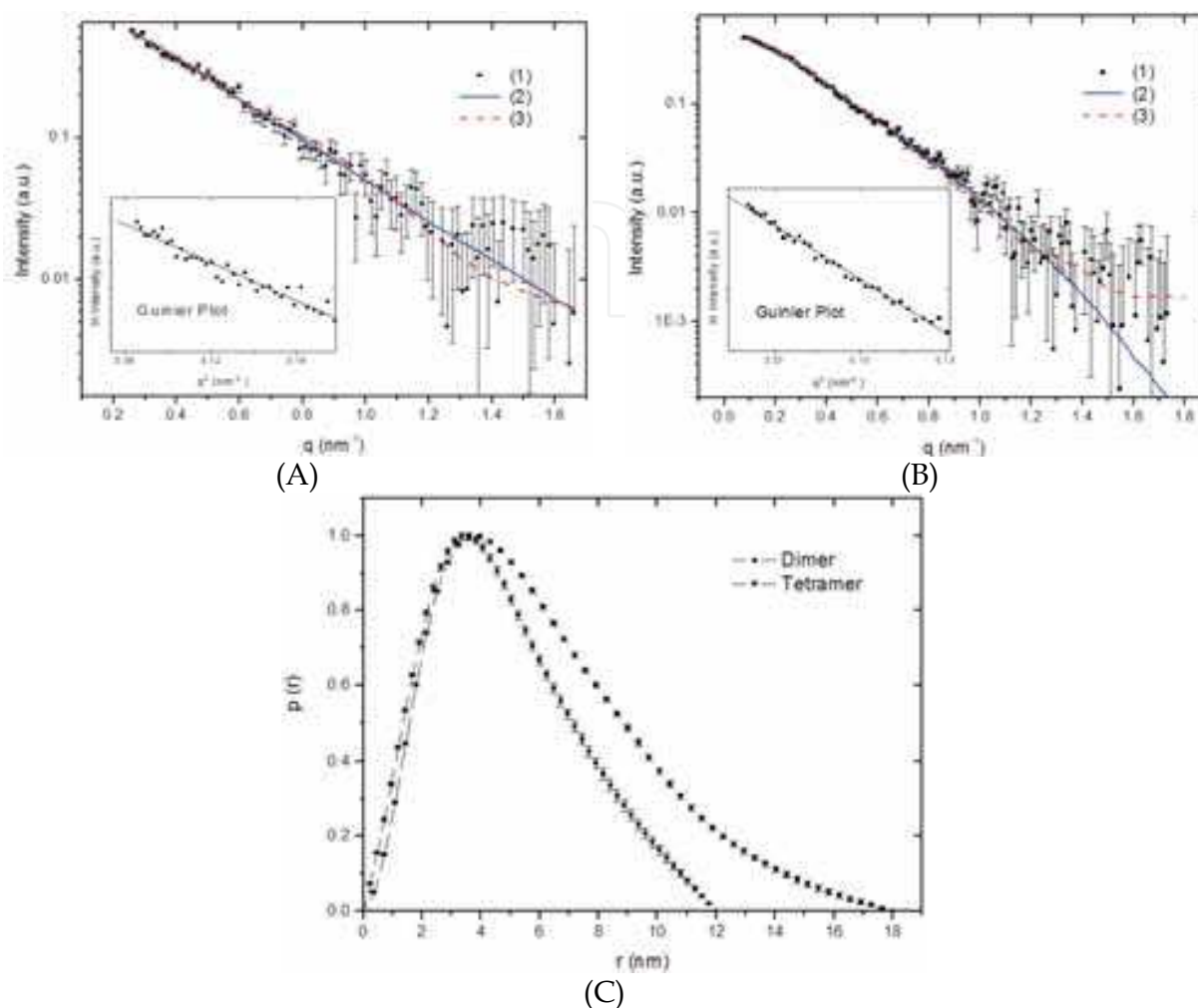suggesting that apo TRs form tetramers in solution which dissociate into dimers upon hormone binding.



(A)                                                    (B)



(C)

Figure 6. Experimental scattering curves for hTR$\beta$1 DBD-LBD construct in solution, which were fitted with the low and high-resolution models, and the distance distribution functions: (A) dimer and (B) tetramer. Log $I$ vs $q$ focusing on the fitting of the experimental curve at high $q$ values with an inset containing the corresponding Guinier plots (log $I$ vs $q^2$): (1) experimental curve, (2) scattering intensity from the DAMs [Dammim], and (3) scattering intensity from the high resolution models. (C) Distance distribution functions of hTR$\beta$1 DBD-LBD dimers (●) and tetramers (■) are given.

This methodology was also applied to other nuclear receptors and protein systems (Fischer *et al.*, (2003), Garcia *et al.* (2006), Grimm *et al.* (2006), Calgaro *et al.* (2007), Nascimento *et al.* (2007), Mario Oliveira Neto *et al.* (2008)), in an addition to polymers in solutions (Leite *et al.* (2007)). For the latter, the initial part of the curve, probably due to larger particles, had to be disregarded, thus considering the polymer system as approximately monodisperse. Figure 8 shows molecular models for the particle shape of poly(o-ethoxyaniline) at distinct pHs, obtained through an *ab initio* procedure based on simulated annealing using the dummy atom model (DAM). A less-packed, coiled structure is observed for pH 3, while at pH 10 blobs are formed.
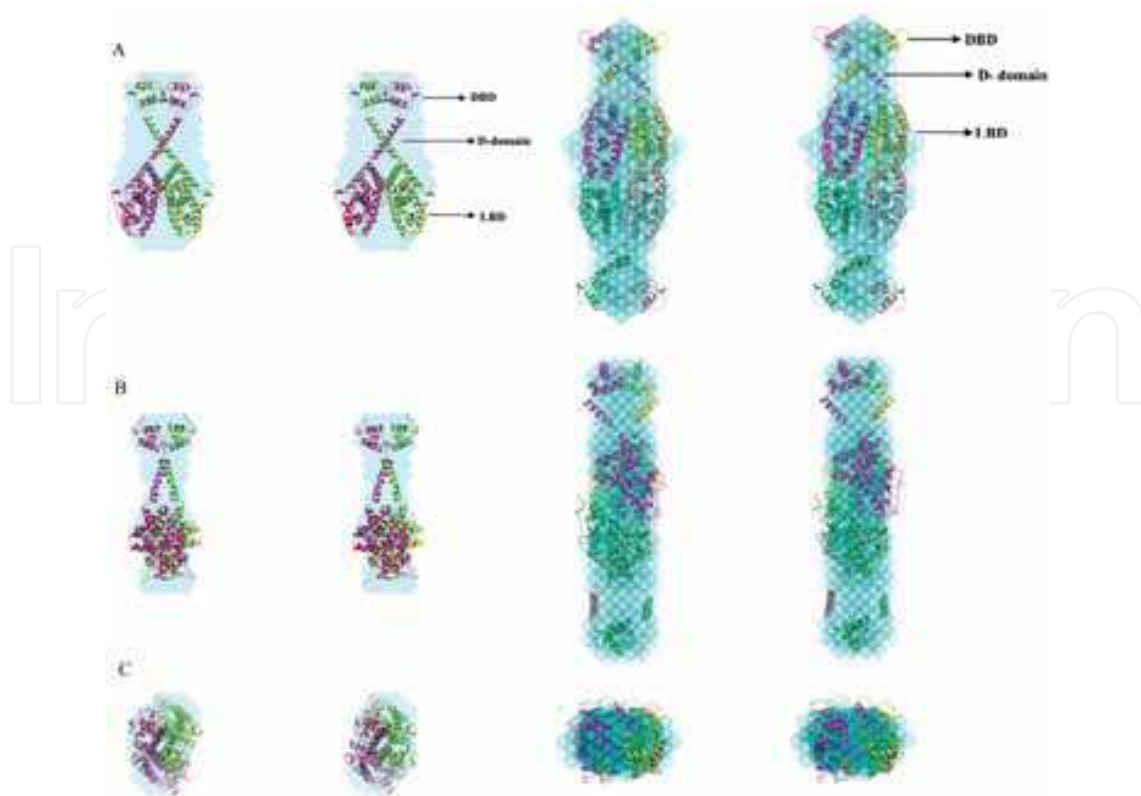
Figure 7. Steroviews showing the superposition of DAM and crystallographic structures for TR$\beta$ DBD-LBD dimer (left) and hTR$\beta$ DBD-LBD tetramer (right). Panels A-C display three orthogonal views. *Ab initio* simulations were performed using protocols implemented in Dammin.

| Parameters/ sample | hTR$\beta$1 DBD-LBD dimer | | | hTR$\beta$1 DBD-LBD tetramer | | |
|---|---|---|---|---|---|---|
| | exp[a] | mod[b] | DAM[c] | exp[a] | mod[b] | DAM[c] |
| D$_{max}$ (nm) | 12.00 ± 1.00 | 12.80 | 12.10 | 18.00 ± 1.00 | 18.61 | 17.84 |
| R$_g$ (nm) | 3.79 ± 0.50 | 3.82 | 3.66 | 4.97 ± 0.50 | 4.96 | 4.84 |
| Discrepancy $\chi$ | - | 0.9 | 0.9 | - | 1.2 | 1.1 |
| Resolution (nm) | 3.8 | - | 3.8 | 3.5 | - | 3.5 |

[a] Calculated from the experimental data.

[b] Parameters of the dimer and tetramer models.

[c] Parameters of the dummy atom models averaged over 20 models.

Table 1. structural parameters derived from SAXS data

## 6. Conclusion

Many structural studies have been performed with a combination of SAXS and simulated annealing to reconstruct three dimensional models. Simulated annealing is suitable for the study of monodisperse, diluted and two-electron densities systems. In this chapter we showed how the simulated annealing procedure can be used to minimize the discrepancy between two functions: the simulated intensity and the experimental one-dimensional SAXS curve. The goal was to find the most probable form for a protein molecule in a monodisperse dilute solution. In the past, this simulated intensity was obtained using

Debye's formula, in time-consuming trial-and-error procedures. Today, with the power of modern computers, it can be applied quite straightforwardly generating low resolution models of proteins in an efficient way. The main advantage of solution scattering is its ability to probe the structure of native particles in nearly physiological conditions and to analyze structural changes in response to variations in external parameters. Then, the oligomerization state of proteins and large conformational changes may be monitored. Moreover, the approach can be extended to conjugated polymers in solution, as described in this Chapter.



|        (a)        |        (b)        |        (c)        |
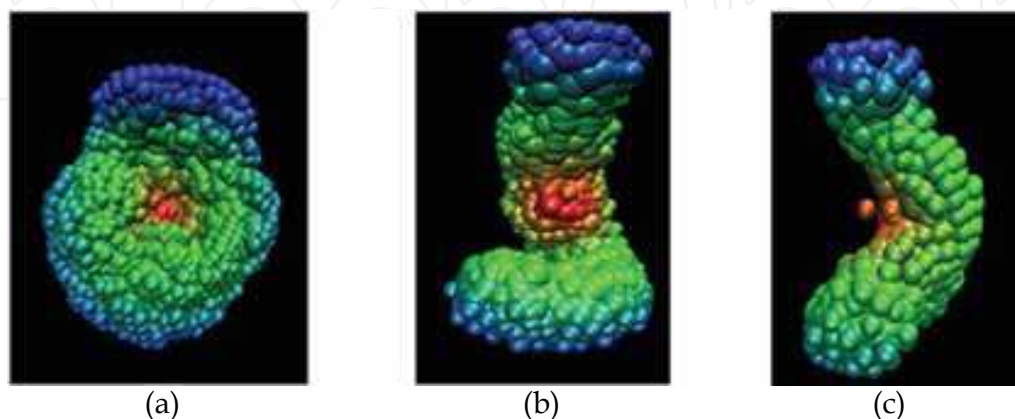
Figure 8. Average DAM for POEA in pH (a) 10.0, (b) 5.0 and (c) pH 3.0 (HCl) (adapted from Leite *et al.* (2007)).

## 7. Acknowledgments

## 8. References

Calgaro, M. R.; Oliveira Neto, M.; Figueira, A. C. M.; Santos, M. A. M.; Portugal, R. V.; Guzzi, C. A.; Saidemberg, D. M.; Bleicher, L.; Vernal, J.; Fernandez, P.; Terenzi, H.; Palma, M. S. & Polikarpov, I. (2007). Orphan nuclear receptor NGFI-B forms dimers with nonclassical interface. *Protein Sci.* Vol. 16, pp. 1762-1772.

Chacon, P.; Moran, F.; Diaz, J. F.; Pantos, E. & Andreu, J. M. (1998). Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. *Biophys J*, Vol. 74, pp. 2760-75.

Craievich, A. F. (2005). Small-Angle X-ray Scattering by Nanostructured Materials . In: *Handbook of Sol-Gel Science and Technology*. A. Sakka; R. Almeida. (Ed.), pp. 161-189, Kluwer Academic Publishers, ISBN 978-1402079696, Norwell.

Feigin, L. A. & Svergun, D. I. (1987). *Structure Analysis by Small-Angle X-Ray and Neutron Scattering,* Plenum Press, ISBN 978-0306426292, New York.

Figueira, A. C. M.; Oliveira Neto, M. O.; Bernardes, A.; Dias, S. M. G.; Craievich, A. F.; Baxter, J. D.; Webb, P. & Polikarpov, I. (2007). Low resolution structures of thyroid hormone receptor dimers and tetramers in solution. *Biochemistry*, Vol. 46, pp. 1273-1283.

Fischer, H.; Dias, S. M. G.; Santos, M.A.; Alves, A.C.; Zanchin, N; Craievich, A. F.; Apriletti, J. W.; Baxter, J. D.; Webb, P.; Neves, F.A.; Ribeiro, R.C. & Polikarpov, I. (2003). Low Resolution Structures of the Retinoid X Receptor DNA-binding and Ligand-binding Domains Revealed by Synchrotron X-ray Solution Scattering. *J. Bio. Chem.* Vol. 278, pp. 16030-16038.

Garcia, W.; de Araújo, A. P. U.; Oliveira Neto, M.; Ballestero, M. R. M.; Polikarpov, I.; Tanaka, M.; Tanaka, T. & Garratt, R. C. (2006). Dissection of a Human Septin: Definition and Characterization of Distinct Domains within Human SEPT4. *Biochemistry*; Vol. 45, pp. 13918-13931.

Glatter, O. & Kratky, O. (1982). *Small Angle X-ray Scattering,* Academic Press, ISBN 978-0122862809, London.

Grimm, E. D.; Portugal, R. V.; Oliveira Neto, M.; Martins, N. H.; Polikarpov, I.; Zaha, A. & Ferreira, H. B. (2006). Structural Analysis of an Echinococcus granulosus Actin-Fragmenting Protein by Small-Angle X-Ray Scattering Studies and Molecular Modeling. *Biophys. J.* Vol. 90, pp. 3216-3223.

Kirkpatrick, S.; Gelatt Jr., C. D. & Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science,* Vol. 220, pp. 671 – 680.

Leite, F. L.; Oliveira Neto, M.; Paterno, L. G.; Ballestero, M. R. M.; Polikarpov, I.; Mascarenhas, Y. P.; Herrmann, P. S. P.; Mattoso, L. H. C. & Oliveira Jr., O. N. (2007). Nanoscale conformational ordering in polyanilines investigated by SAXS and AFM. *J. Colloid Interface Sci.* Vol. 316, pp. 376–387.

Matozo, H. C.; Santos, M. A. M.; Oliveira Neto, M.; Bleicher, L.; Lima, L. M. T. R.; Iuliano, R.; Fusco, A. & Polikarpov, I. (2007). Low-resolution structure and fluorescence anisotropy analysis Of protein tyrosine phosphatase η catalytic. *Biophys. J.* Vol. 92, pp. 4424–4432.

Nascimento, A. S.; Dias, S. M. G.; Nunes, F. M.; Aparício, R.; Ambrosio, A. L. B.; Bleicher, L.; Figueira, A. C. M.; Santos, M. A. M.; Oliveira Neto, M.; Fischer, H.; Togashi, M.; Craievich, A. F., Garratt, R. C.; Baxter, J. D.; Webb, P. & Polikarpov, I. (2006). *Structural Rearrangements in the Thyroid Hormone Receptor Hinge Domain and Their Putative Role in the Receptor Function. J. Mol. Biol.* Vol. 360, pp. 586-98.

Nascimento, A. S.; Catalano-Dupuy, DL.; Bernardes, A.; Oliveira Neto, M.; Santos, M. A. M.; Ceccarelli, E. A. & Polikarpov, I. (2007). Crystal structures of Leptospira interrogans FAD-containing ferredoxin-NADP+ reductase and its complex with NADP+BMC. *Structural Biology*; Vol. 7, pp. 69-81.

Oliveira Neto, M.;Ferreira Júnior, J. R.; Fischer, H.; Nascimento, A. S.; Craievich, A. F.; Colau, D.; Dumoutier, L.; Renauld, J.-C. & Polikarpov, I. (2008). Interleukin-22 forms dimers which are recognized by two interleukin-22R1 receptor chains. *Biophys. J.,* Vol. 94, pp. 1754–1765.

Patterson, A. L. (1935). A direct method for the determination of the components of interatomic distances in crystals. *Z. Kristallogr. A* Vol. 90, pp. 517-542.

Stuhrmann, H. B. (1970). Ein neues Verfahren zur Bestimmung der Oberfla chenform und der inneren Struktur von gelosten globularen Proteinen. *Zeitschr. Physik. Chem. Neue Folge,* Vol. 72, pp. 177–198.

Stuhrmann, H. B. 1970b. Interpretation of small-angle scattering of dilute solutions and gases. A representation of the structures related to a one-particle scattering functions. *Acta Crystallogr.,* Vol. A26, pp. 297–306.

Svergun, D. I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J,* Vol. 76, pp. 2879-86.

Svergun, D.I.; Petoukhov, M.V. & Koch, M.H.J. (2001). Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.* Vol. 80, pp. 2946-2953.

Svergun, D.I. & Koch M.H.J. (2003). Small-angle scattering studies of biological macromolecules in solution. Rep. Prog. Phys. Vol. 66, pp. 1735–1782.

Volkov, V. V. & Svergun, D. I. (2003). Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Cryst.* Vol. 36, pp. 860-864.
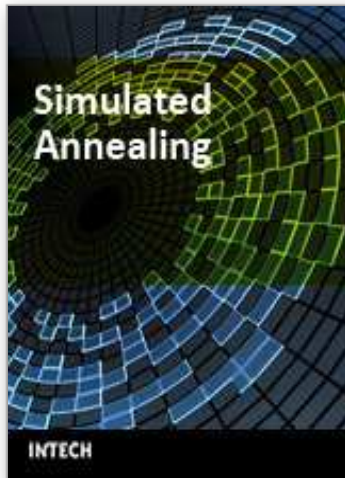
**Simulated Annealing**

Edited by Cher Ming Tan

ISBN 978-953-7619-07-7

Hard cover, 420 pages

**Publisher** InTech

**Published online** 01, September, 2008

**Published in print edition** September, 2008

This book provides the readers with the knowledge of Simulated Annealing and its vast applications in the various branches of engineering. We encourage readers to explore the application of Simulated Annealing in their work for the task of optimization.

**How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

# INTECH
open science | open minds