TOWARDS IMPROVEMENTS IN SMALL-ANGLE X-RAY SCATTERING DATA MODELING OF MESOPOROUS SBA-15

Master project

Written by:

Xiangyin Tan

May 19, 2022

Supervised by:

Heloisa Nunes Bordallo

Pedro Leonidas Oseliero Filho

UNIVERSITY OF COPENHAGEN

UNIVERSITY OF
COPENHAGEN

FACULTY:	Science

Niels Bohr Institute

AUTHOR:	Xiangyin Tan	KU-ID: jcv814
TITLE:	Towards improvements in small-angle X-ray scattering data modeling of mesoporous SBA-15
SUPERVISOR:	Heloisa Nunes Bordallo	bordallo@nbi.ku.dk
CO-SUPERVISOR:	Pedro Leonidas Oseliero Filho	pedro.filho@nbi.ku.dk
HANDED IN:	19th of May 2022
DEFENDED:	17th of June 2022

NAME _____

SIGNATURE _____

DATE _____

Abstract

This project aims to use Monte Carlo (MC) method to improve the model proposed originally by Sundblom et al [1] for quantitative analysis of small-angle X-ray scattering (SAXS) data from the ordered mesoporous silica (OMS) material named SBA-15. The nanostructured mesoporous silica SBA-15, formed by around $20\mu m$ particles with hexagonal ordered mesopores with a diameter of 10 nm and macropores larger than 50 nm, has the ability to protect antigens from the harsh stomach environment, improve the immune response, and in the future allow for the use of needless vaccines. Depending on their size, the antigens can be located in the macropores (HBsAg, with a radius around 22 nm, used in the Hepatitis B vaccine) or in the mesopores (dANA, with a radius around 4 nm, used in Diphtheria vaccine). Thus, it is essential to know the size distribution of the mesopores in the SBA-15, and SAXS is a experimental technique that allows us answering this question. However, more advanced models to analyse the experimental data are needed, and MC is a good alternative.

Abbreviation List

dANA diphtheria anatoxin.

HBsAg Hepatitis B surface antigen.

MC Monte Carlo.

OMS Ordered Mesoporous Silica.

PEO-PPO-PEO poly (ethylene oxide) - poly (propylene oxide) - poly (ethylene oxide).

RF radio frequency.

SANS Small-Angle Neutron Scattering.

SAS Small-Angle Scattering.

SAXS Small-Angle X-Ray Scattering.

SBA Santa Barbara Amorphous.

TEOS tetraethoxilane.

TMOS tetramethoxysilane.

TPOS tetrapropoxysilane.

WAXS Wide-Angle X-Ray Scattering.

XRD X-Ray Diffraction.

Acknowledgement

Contents

Chapter 1

Introduction

Porous materials have raised a very large interest in the material science community since they appeared in 1990s [2]. This attention comes from their many advantages, such as large surface areas and pore volume, as well as narrow distribution of pore sizes. The International Union of Pure and Applied Chemistry (IUPAC), classifies the porous material into three different categories, according to their pore size; micropores being smaller than 20Å, macropores larger than 500Å, and the mesopores in between [3]. Accordingly, zeolites and microporous molecular sieves, which are widely applied as catalyst for oil refining and petrochemistry, are representatives of the microporous material family. However, due to their pore size, there is an intensive limitation in microporous material when the processes involving large molecules. Thus, researchers have focused on the enlargement into the mesopore range [2]. As a result, in 1998, Santa Barbara Amorphous (SBA)-15 materials are Ordered Mesoporous Silica (OMS), were firstly synthesized by Zhao et al [4]. by using amphiphilic triblock copolymers to direct the arrangement of polymerizing silica species (Fig.1.1). Since then, SBA-15 has attracted wide interest of scientists due to their highly ordered mesoporous structure, large pore diameter, inactive framework, thick pore wall and high thermal and hydrothermal stability, which means they can be applied widely in catalysis, absorption, immobilization, drug delivery and vaccine carrier [2, 5, 6]. More recently, it was
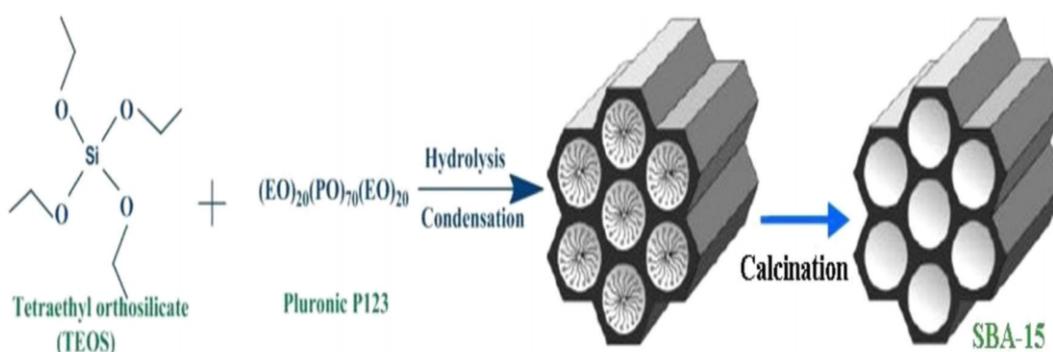
Figure 1.1: Sketch of the general procedures for the SBA-15 preparation [2]

shown that, due to its low toxicity, stable thermal and hydrothermal properties and hierarchical porous structure, SBA-15 can be used as nanocarriers for protecting the integrity of oral vaccine from the harsh gastrointestinal environment [5, 6]. The reason is that SBA-15 is composed by pores with hierarchical sizes, and proteins can be protected into these pores. For instance, diphtheria anatoxin (dANA) with around 4 nm diameter is small enough to enter the 10 nm mesopores, while Hepatitis B surface antigen (HBsAg), with a diameter of about 22 nm, is too large to be protected in the mesopores. However, the complex SBA-15 + HBsAg induces immune response to hepatitis, because HBsAg is encapsulated in the macropores of SBA-15 [6]. Moreover, it is now known that the morphology of SBA-15 could affect the efficiency of SBA-15 absorbing the antigens, and the level of immune response [6]. Based on this knowledge, it is vital to characterize the structure of SBA-15 sample used in oral vaccine delivery i.e. size, shape, size distribution and spatial ordering in advance. In this case, Small-Angle X-Ray Scattering (SAXS) is a powerful technique because it can provide the spatial information in the range of 10-1000 $\mathring{A}$ (1-100 nm) that fits well with the dimensions of the sample of interest. Moreover, it is a convenient technique, since the experiments are easy to perform, accurate, data collection is fast and, in general, it is non-destructive. However, as in most situations, the overall structural sizes in a SBA-15 sample can present fluctuation, and advanced analysis through data modeling is needed in order to retrieve "hidden" structural information contained in the experimental curves. Apart several structural parameters, size distributions between cylindrical pores must be taken into account. The models currently available to fit SAXS curves of SBA-15 assume a particle size distribution, such as Gaussian [7, 8] or Schulz-Zimm distribution [1] for the mesopores. Since this assumption affects the form factor describing their morphology, and consequently the scattering pattern, this may compromise the fitting quality.

In this project, we propose to apply the Monte Carlo (MC) method to improve Sundblom et al model [1], because compared with the other two models [7, 8], this one analyses the whole process of synthesising the SBA-15, and better fit the results. MC methods are thought to be the suitable methods in this work, because they are based on random sampling and the trial and error principle. In fact, the application of MC in polydisperse sphere particle system has been successfully done to describe a reasonable size distribution in the SAS pattern of boehmite (AlOOH) particles [9, 10].

To achieve the results presented in this thesis the consecutive steps were followed:

1. The first step was to write new program to simulate the SAXS curves of different shapes of mono and polydisperse particles: sphere, cylinder and core-shell sphere. To validate the new program,the results were compared with patterns generated by SasView, a widely used Small-Angle Scattering (SAS) data analysis software [11]. Results for sphere are shown in Appendix, together with all the related codes.

2. Then, I applied MC methods in a polydisperse sphere, in a long cylinder and in a long core-shell cylinder system and compared the retrieved distribution to the theoretical one. This is presented in Chapter 4.2, together with the analysis of a real data set.

3. The next step was to write a new Python code to build the mathematical model de-

scribed in Ref [1] for pure SBA-15. This result is shown in Chapter 4.1.

4. Finally, the MC method was applied to the SBA-15 model, and the radius of the mesopores distribution extracted. This first attempt is promising, but improvements are necessary as described in the Conclusions and Perspectives Chapter.

Chapter 2

Materials and Methods

2.1 Santa Barbara Amorphous: SBA-15

SBA-15 is one of the most promising members of the SBA family, due to its chemical and physical properties including low toxicity, thermal and hydrothermal stability and an hierarchical porous structure [5]. This nanostructured mesoporous silica is formed by 20 μm particles with hexagonal ordered mesopores with a radius of 50 $\mathring{A}$ and macropores larger than 500 $\mathring{A}$ as showed in Fig. 2.1.
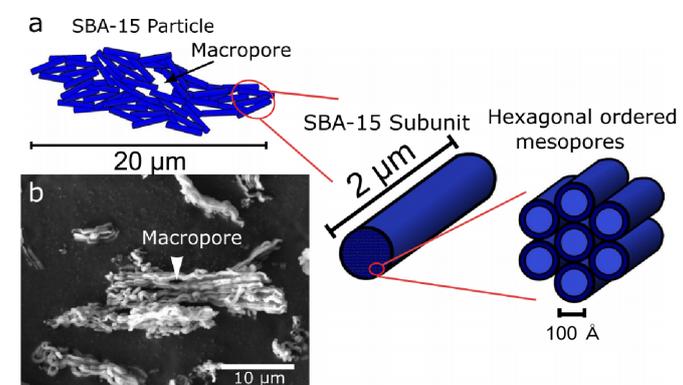
Figure 2.1: (a) Schematic of SBA-15 particle showing hexagonal ordered mesopores.(b) Scanning electron microscopy (SEM) image of the SBA-15 with some macropores [6].
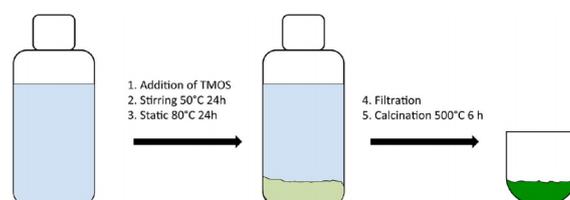
Figure 2.2: The standard procedure to produce SBA-15, and the morphology can be changed by varying the conditions. The acid solution including the structure director, poly (ethylene oxide) - poly (propylene oxide) - poly (ethylene oxide) (PEO-PPO-PEO), is heated to 50°C. The silica source, TEOS, is added, and the solution is stirred at this temperature for 24h. The reagent flask is then transfer to an oven, held at 80°C for another 24h. Then, the powder is filtered out from the solution and wash with water. Finally the material is calcined for 6h at 500°C [12].

SBA-15 is usually synthesised by PEO-PPO-PEO, also known as Pluronic P123, as a structure directing agent and appropriate silica sources, such as tetraethoxilane (TEOS), tetramethoxysilane (TMOS), and tetrapropoxysilane (TPOS) in a range of temperature from 35° to 140°, followed by high temperature calcination to degrade the copolymer-template, as shown in Fig.2.2 [2, 4, 12, 13].

The kinetic process of how the SBA-15 forms is shown in Fig. 2.3 [14]. The unique characteristic of PPO block in Pluronic P123, which is hydrophobic at temperatures above 288 K and is soluble in water at temperatures below 288 K, leads to the formation of micelles consisting of PEO-PPO-PEO triblock copolymers (Fig. 2.3 a). After addition of silica source, TEOS starts to hydrolyse, and then the shape of their micelles starts changing from spherical to cylinder-like (Fig.2.3 b-d). Afterwards, the cylinder-like micelles form a hexagonal phase due to gathering, and keep growing (Fig.2.3 e and f).
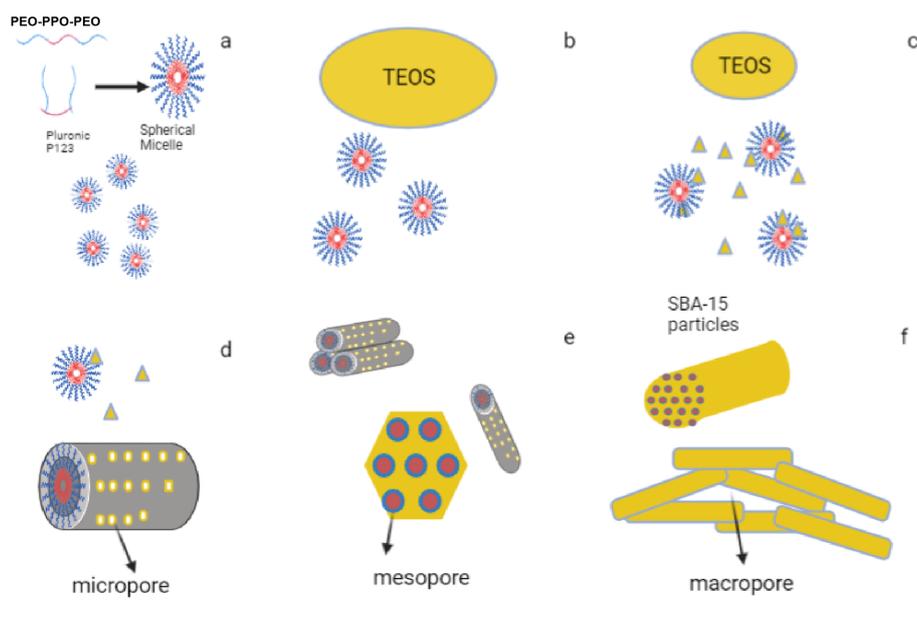
Figure 2.3: Schematic of how the structure of SBA-15 is formed. a: The formation spherical micelles; b: Addition of TEOS; c: Beginning of hydrolysis; d: Formation of cylinder shape; e: Nucleation of the 2D-hexagonal phase; f: Growth of 2D-hexagonal phase [14].

There are three SBA-15 samples studied in this work, and we name them SBA-15_a, _b and _c. All of them were synthesized at the Institute of Physics, University of São Paulo in Brazil. SBA-15_a and SBA-15_c were synthesized by Luis Cides and Pedro Leonidas Oseliero Filho. SBA-15_b was synthesized by Tereza Martin and it is a mix of multiple batches. The SAXS data of SBA-15_a was collected in Brazil on a laboratory instrument, and the SAXS of SBA-15_b and SBA-15_c were collected using CoSAXS [15, 16], at MAXIV in Lund. All three samples are pure calcined SBA-15, and their information are summarised in the table 2.1.

Table 2.1: The synthesized and measured information about three data sets

	a	b	c
Synthesis	synthesized by Luis and Pedro	mix batches synthesized by Tereza	synthesized by Luis and Pedro
Measurement	Xeuss 2.0 with $\lambda = 1.54\mathring{A}$	COSAXS,MAXIV, Lund, Sweden, with $\lambda = 1\mathring{A}$	COSAXS, MAXIV, Lund, Sweden, with $\lambda = 1\mathring{A}$

2.2 X-Rays
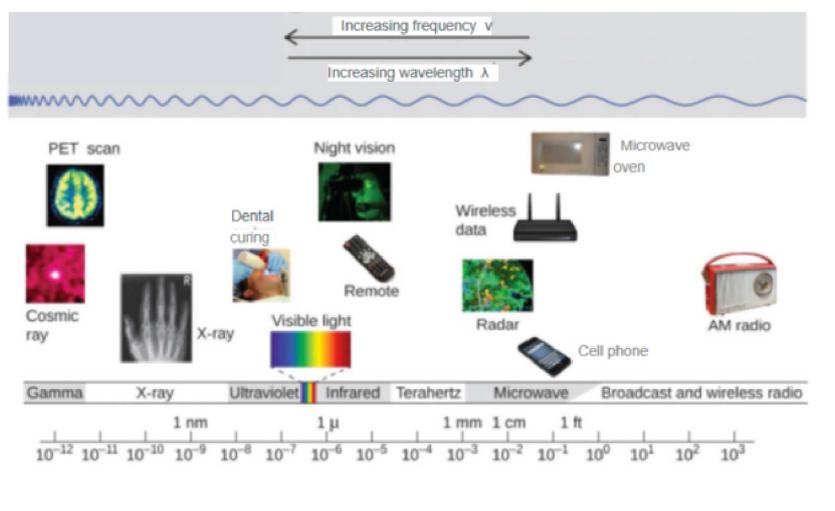
Figure 2.4: The wavelength of common radiations, showing that the X-rays wavelength enable the study of atomic systems [17].

X-ray is a high energy electromagnetic radiation discovered by the German physicist Wilhelm Conrad Röntgen in 1895. The standard X-ray tube showed in Fig. 2.5 consists mainly of three parts: filament, cathode and anode. By heating the filament and applying high voltage (around 30-60 kV) between the cathode and the anode, electrons will be accelerated and collide the metal target at a high speed. This process will create X-rays, which energies or wavelengths are defined by the anode material.
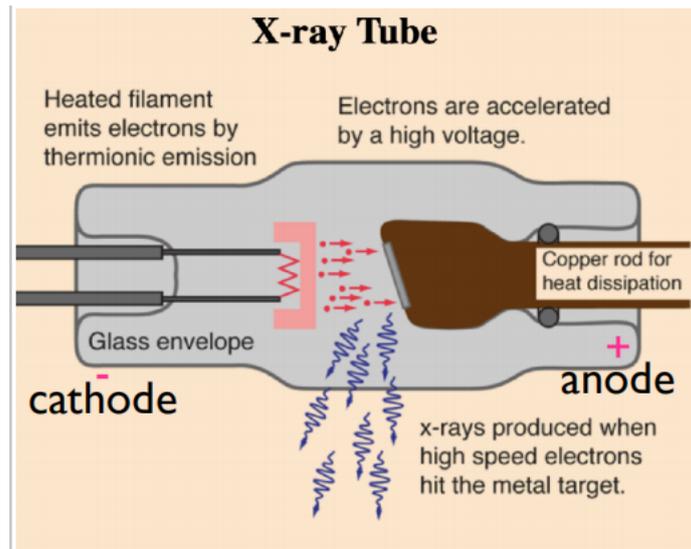
Figure 2.5: Schematic of X-Ray tube. The cathode and the anode are providing the electric field to accelerate the electrons emitted from the heated filament through thermionic emission, and when the electrons collide with the target metal, the X-rays are produced [17].

The raw X-ray spectrum is continuous because while the electron hits the anode, it is decelerating, and this radiation is called *Bremsstrahlung*, *bremsen* means brake in German. The highest energy of *Bremsstrahlung* depends on the voltage applied on the tube. The sharp peaks on the *Bremsstrahlung* spectrum in Fig.2.6 are called characteristic energy, which occurs when the incident electron hits the inner atomic electron, and removes it from the inner shell. The electron from the outer shell will fall down to the vacancy and this process will emit the fluorescent radiation K_α or K_β, which depends on the anode material. The energy of K_α is equal to the difference between L and K shell, while K_β is the difference between M and K shell.

 In a X-ray scattering experiment, we need to utilize monochromatic and highly intense radiation, and K_α is several order of magnitude more intense. Thus, we need to filter out K_β and only the K_α will be used in the scattering experiment. The filter is usually made of a metal having one proton less than the anode material. The condition of choosing the filter material is that the K excited energy of the filter should be in between the energy of anode's K_α and K_β. Thus, K_α line is left and the radiation with energy higher than K excited energy gets absorbed. Or we can use crystal monochromators to select the wanted wavelength of the X-rays. It's principle is to use Bragg's scattering as a "filter", and only the radiation suits the Bragg's law that can be reflected. This way could obtain a more narrow band of the wavelengths.
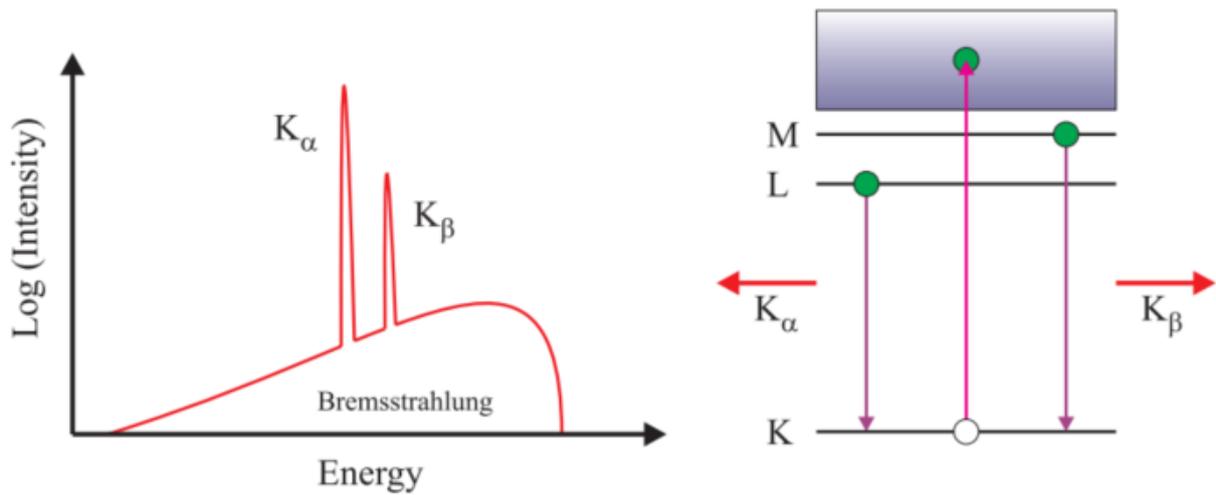
Figure 2.6: The spectrum from an X-ray tube has sharp fluorescent lines superimposed on the continuous bremsstrahlung radiation (left). Schematic atomic energy level diagram (right): the K_α line results from transitions between an L and K shell, whereas the K_β comes from an M to K transition [18].

2.3 Synchrotron Radiation

It is known that the electrons inside the synchrotron magnets emit electromagnetic radiation which is known as synchrotron radiation. This happens because the electrons are constantly accelerated, since they move in curved trajectories, even though their speed remains the same. The construction of a synchrotron facility is shown in Fig.2.7b.

(a) synchrotron generation. (b) planar view of synchrotron (c) storage ring
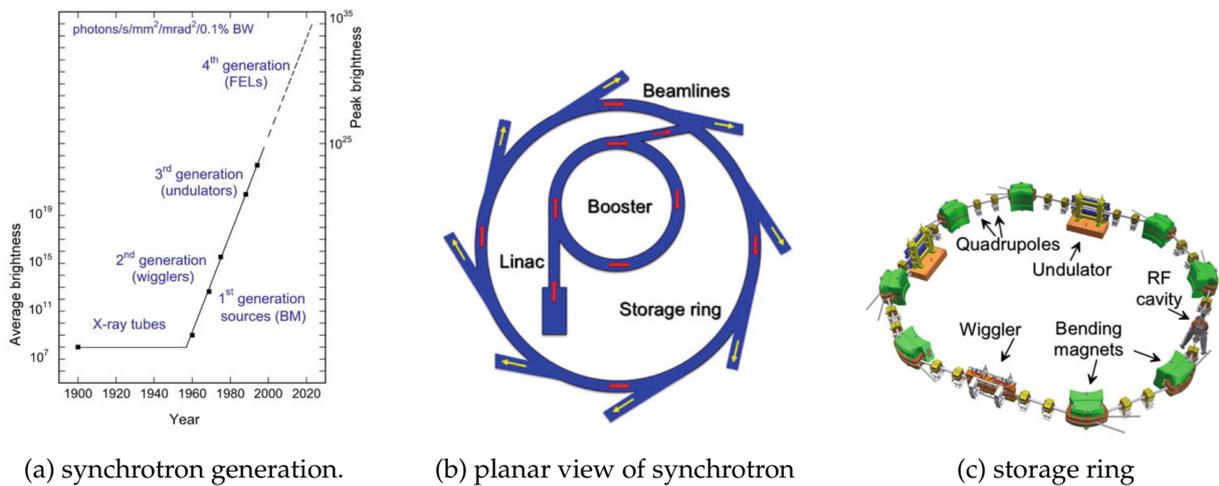
Figure 2.7: a is the comparison between the average brightness of storage rings for different synchrotron generations. b is the schematic planar view of a synchrotron radiation facility, where mainly is consisted of a linear accelerator (linac), booster, storage ring and beamlines. c is the schematic of a storage ring where some main elements like bending magnets, quadrupoles, wiggler, undulator and RF cavity are shown. [19]

At first, even though the electrons might have reached the expected energy, they lose some of it as electromagnetic radiation due to their movement. However, this energy is given back to them through radio frequency (RF) electric fields when the electrons pass

through RF cavities. Those cavities are installed inside linear sections of the storage rings. The RF cavities were the first way to keep the electron velocity under control. Afterwards, more techniques were developed like the wigglers and the undulators. Those techniques are what cause the synchrotron to have a continuous production of energy and at such high levels. As we can see from the following diagram the more advanced the technique used the greater the spectral brightness (Fig.2.7a). The definition of the brightness or brilliance is photons per second (s), per unit source size (mm^2) and angular divergences (mr^2) in a given bandwidth (BW), and its function [18] is:

$$Brightness = \frac{Photons}{s \cdot mr2 \cdot mm2 \cdot (0.1\%BW)} \tag{2.3.1}$$

where the angular divergences are set by the horizontal and vertical apertures (in milliradian), and the relative bandwidth of the monochromator crystal relative to 0.1%.

2.4   X-Rays Scattering Theory

When we use X-rays to irradiate the sample, X-rays interact with all the atoms in the irradiated volume. A fraction of X-rays get absorbed and part of X-rays get scattered. X-rays interact mostly with the electrons in the atoms, whereas neutrons interact with nuclei. In terms of how X-rays interact with electrons, it can be explained from two aspects, particles (photon) and wave. If we consider X-rays as photons, photons collide with electrons and their trajectory is changed by the electrons. If we consider X-rays as waves, electrons absorb the wave to a excited state, and another X-rays is emitted because the excited electrons decay to the stable state.

Under certain energy limits ($\approx 7 - 12KeV$, according to $E = \frac{hc}{\lambda}$, i.e. $\lambda \approx 1.03 - 1.77\text{Å}$), this scattering process can be thought as elastic scattering or Thomson scattering, where the energy and wavelength (λ) of the incident and scattered radiation are identical [20]. It is true that the interaction between electron and X-rays can be inelastic, called Compton scattering. In a elastic scattering experiment, scattering unit is an electron, which is thought to be structureless, and the simplest cases consists of two electrons with a displacement **r** between them. Determining the structure of this system is to determine the value of **r**. For elastic scattering $|\mathbf{k_0}| = |\mathbf{k_F}|$, the magnitude of scattering vector **q**(Fig.2.8) is equal to:

$$|\mathbf{q}| = 2k\sin\theta = \frac{4\pi}{\lambda}\sin\theta \tag{2.4.1}$$

The scattering amplitude is defined as the Fourier Transform (FT) of the distribution of the electrons within the irradiated system, and the scattering intensity is equal the square of the scattering amplitude.

The FT of the two electrons system (Fig.2.8) is written as

$$F(\mathbf{q}) = \int \rho(\mathbf{r})e^{i\mathbf{q}\cdot\mathbf{r}}dr \tag{2.4.2}$$
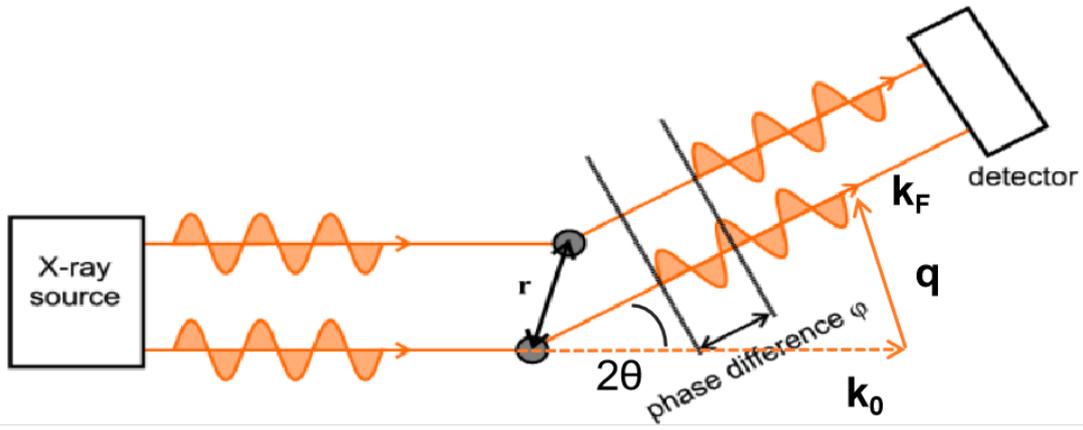
Figure 2.8: Representation of how two electrons get scattered. The lag phase or phase difference is ϕ, and the scattering angle is 2θ. $\mathbf{q}$ is the scattering vector.

where the $\rho(\mathbf{r})$ is the electron density function for X-rays or called scattering length density, which is in real space, and $\mathbf{q}$ is a coordinate in Fourier space, which is also called reciprocal space. Thus, the unit of q is $\mathring{A}^{-1}$, when r is in $\mathring{A}$. In this two electrons system, if we put the origin of coordinate at one of the electrons, the other is at a position $\mathbf{r}$ (Fig.2.8). With the FT, we can get the scattering amplitude

$$A(\mathbf{q}) = \int \rho(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} dr = \int (\delta(0) + \delta(\mathbf{r})) e^{iqr} dr = 1 + e^{i\mathbf{q}\cdot\mathbf{r}} \tag{2.4.3}$$

and the intensity becomes

$$I(q) = A(q)A(q)^* = 2 + e^{i\mathbf{q}\cdot\mathbf{r}} + e^{-i\mathbf{q}\cdot\mathbf{r}} \tag{2.4.4}$$

If $\mathbf{r}$ is randomly oriented, then the intensity is given by the orientational average of phase factor [18]:

$$< e^{i\mathbf{q}\cdot\mathbf{r}} >_\Omega = \frac{\int e^{iqr\cos\theta} \sin\theta d\theta d\phi}{\int \sin\theta d\theta d\phi} = \frac{\sin(qr)}{qr}. \tag{2.4.5}$$

These ideas can be extended to more than two electrons, and the amplitude of a system with any electrons is

$$A(\mathbf{q}) = \sum_i e^{i\mathbf{q}\cdot\mathbf{r_i}} \tag{2.4.6}$$

where $\mathbf{r_i}$ is the position of i'th electron.

The electrons bounded in atoms can not seen as a point like, and they are treated as a cloud surrounding the nucleus with a distribution $\rho(\mathbf{r})$. The integral of $\rho(\mathbf{r})$ is equal to the total number of electrons Z in the atom. The scattering amplitude of an atom is the atomic form factor and it is

$$F_a(\mathbf{q}) = \int \rho(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \tag{2.4.7}$$

For $\mathbf{q} \to 0$, the phase factor is equal one, so the integral is equal the number of electrons in the atom i.e. $F_a(\mathbf{q}) = Z$. When $\mathbf{q} \to \infty$, the wavelength of the radiation becomes small compared to the atom, the interference between the waves scattered from differ-

12

ent electrons in the atom is destructive, so $F_a(\mathbf{q}) = 0$.

2.5   Small-Angle X-ray Scattering (SAXS)
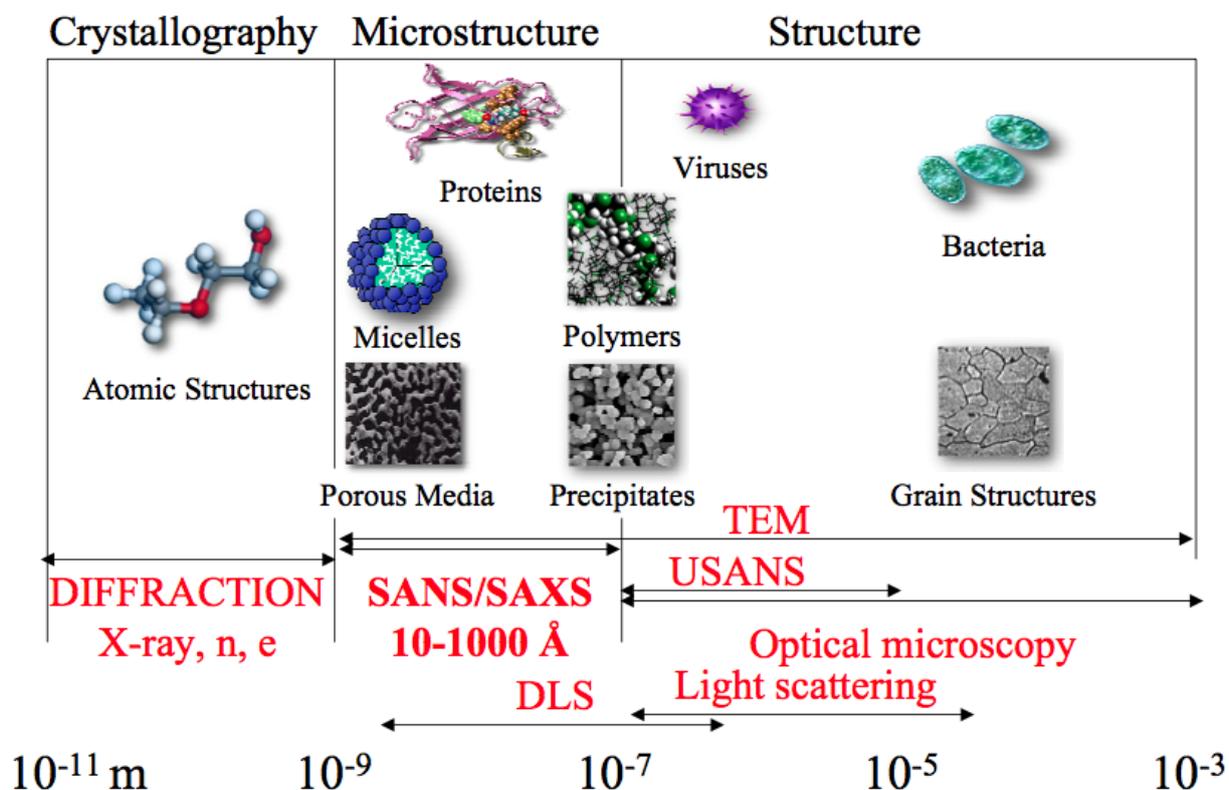
Figure 2.9: Different methods cover different size ranges for our study the suitability of SAXS is clearly shorwn. Here, n and e stands for neutrons and electrons scattering, and DLS for dynamic light scattering. USANS and TEM stand for Ultra Small-Angle Neutron Scattering, and Transmission Electron Microscopy, respectively [17].

Elastic X-ray scattering includes a number of material characterization techniques, for instance, X-Ray Diffraction (XRD), SAXS, and Wide-Angle X-Ray Scattering (WAXS) as shown in Fig.2.9. The term XRD traditionally applies to a well-ordered crystalline material and thus used for determination of crystal structures. This project deals with larger structures and therefore uses the small angle scattering technique, which allows us analysing size between $10 - 1000\mathring{A}$, and can be used to quantified nanoscale density differences in a sample.

The schematic of a SAS experiment is showed in Fig.2.10. Depending on the radiation used, the technique is called either Small-Angle X-ray Scattering (SAXS) or Small-Angle Neutron Scattering (SANS). During data collection a collimated beam is shoot to the sample in a straight trajectory, and after interacting with the sample, the beam will be scattered in a angle, 2θ. The scattering angle of SAS is in the range of $0.1\text{-}10°$, and the measured scattering curve, I(q), covers a momentum transfer q range from 0.007 - 0.7 $\mathring{A}^{-1}$ (in the most frequently used CuK_α line of $1.54\mathring{A}$). The resolved range can be extended on both sides by measuring smaller (Ultra Small-Angle Scattering, USAXS or USANS) or larger angles than the typical values, because distance between sample and

detector is easy to change [21].

In case of small angle scattering, the objects that we are concerned are very large com-
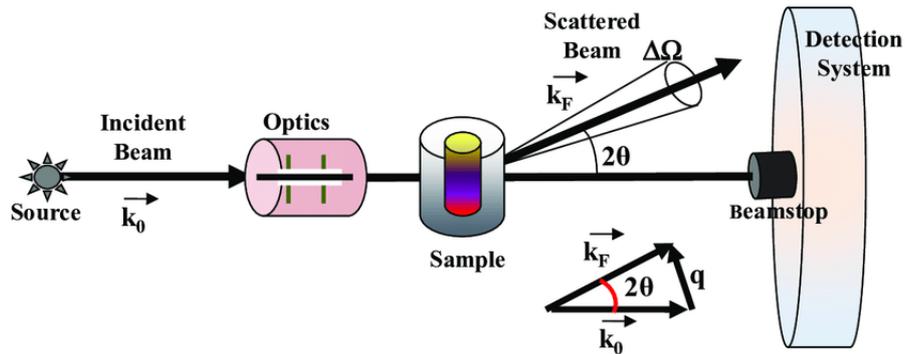
Figure 2.10: Schematic of a SAS setup showing the incident (k_0) and scattered (k_F) beam, as well as the source, optics, sample, beamstop and the detection system. 2θ is the scattered angle, q is the scattering vector, and the beamstop is used to protect the detection system from direct intensive beam [22].

pared with typical inter-atomic distances, and it will be elaborated later. The approaches to analyse the SAXS data will be described in next section.

In SAXS experiment, when the incident radiation suits Bragg's law:

$$\Delta l = n\lambda \tag{2.5.1}$$

$$n\lambda = 2d\sin\theta \tag{2.5.2}$$

there is a intensity increase in the scattering pattern. The second incident radiation (the lower arrows in Fig.2.11) will have half lag path $\frac{1}{2}\Delta l$ and the scattered radiations will have the same lag path. When the total lag path is an integer multiple of the wavelength, there is a constructive interference in these two scattered radiation.
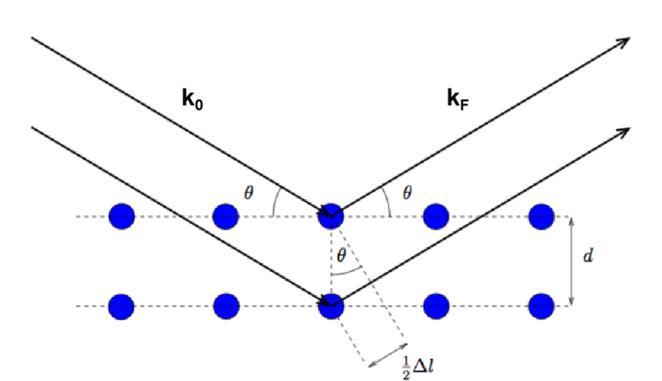
Figure 2.11: The schematic of Bragg's law. **k** and **k** are the incident vector and scattered vector, respectively. θ is the incident angle. d and $\frac{1}{2}\Delta l$ are lattice distance and half lag path.

2.6 Reducing and Modeling SAXS data

2.6.1 Data Reduction

There are basically four data reduction steps in SAXS experiment, corrections, scattering vector q calibration, intensity calibration and averaging. The corrected intensity for the radiation absorbed by the sample is

$$T = f_{sample}/f_0 \tag{2.6.1}$$

$$I_{corr} = \frac{I}{T} \tag{2.6.2}$$

, where T is the transmission of the sample, and f_{sample} and f_0 are the flux of radiation with and without the sample. In the experiment, it is necessary to measure two intensities, one is the system (sample holder + sample) and the other is sample. The corrected intensity of the sample is

$$I_{corr,sample} = I_{corr,total} - I_{corr,holder} \tag{2.6.3}$$

In our case, the sample holder is the empty quartz capillary tube. If the investigated system is liquid, the solvent should be included in the subtraction.

Silver behenate is often used in the scattering vector q calibration. The first order reflection of silver behenate is at $q = 0.1078 Å^{-1}$. It can be used to determine the beam center from the ring, and adjust q-scale so that the first peak is at $q = 0.1078 Å^{-1}$ (see Fig. 2.12a).

Integration is used to transform 2D scattering pattern into a 1D scattering curve. There are two ways to implement, radial and azimuthal integration. Azimuthal integration can always be used, either considering the full 2D data or only sectors, the former strategy being applied to analyse anisotropic patterns. For anisotropic patterns, radial integration can additionally be used in order to evaluate ordering features of the system. What azimuthal integration does is showed in Fig.2.12b. Performing a 2π integration to the intensity points that lies on a certain radius circle. As the radius increases, the value of q also increases, and then obtain the scattering curve.

2.6.2 SAXS Models

The scattering amplitude for a fixed particle is the FT of the particle electron density $\rho(\mathbf{r})$:

$$f(\mathbf{q}) = \frac{1}{4\pi} \int_V \rho(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \tag{2.6.4}$$

(a) q calibration

(b) Schematic of performing azimuthal integration.
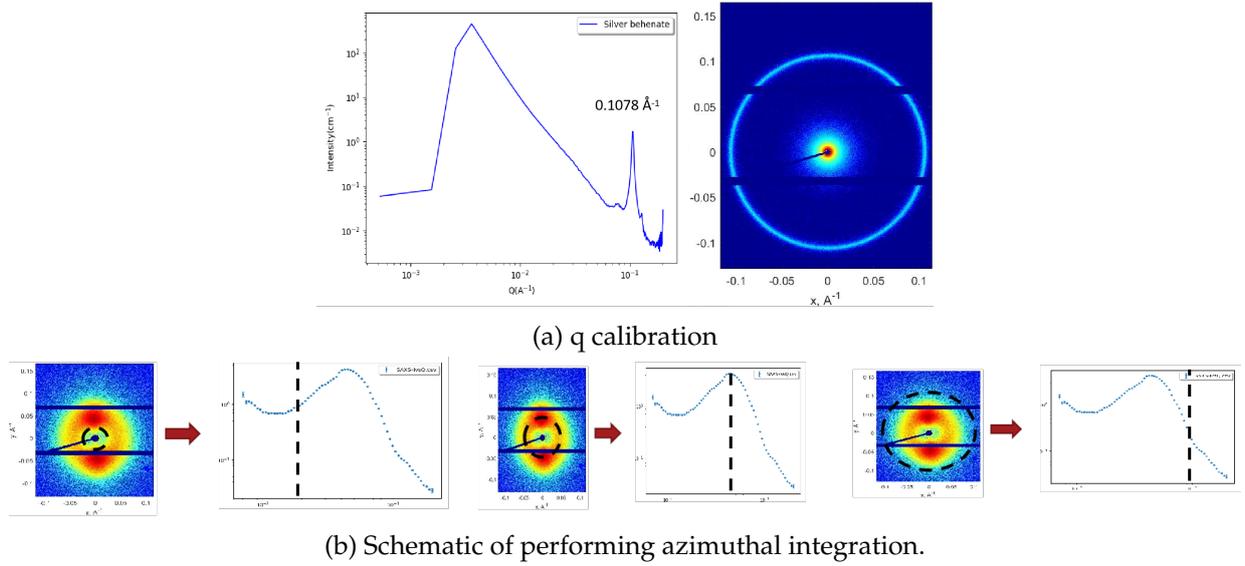
Figure 2.12: 2.12a is schematic of calibrating scattering vector q. The scattering curve of silver behenate is on the left, and the 2D scattering pattern is on the right. 2.12b is the schematic of performing azimuthal integration.

, and intensity $I_1(\mathbf{q})$:

$$I_1(\mathbf{q}) = \int_V \gamma(\mathbf{r})e^{-i\mathbf{q}\cdot\mathbf{r}}d\mathbf{r} \tag{2.6.5}$$

$$\gamma(\mathbf{r}) = \int_V \rho(\mathbf{r}^*)\rho(\mathbf{r}^* - \mathbf{r})d\mathbf{r}^* \tag{2.6.6}$$

Where $\gamma(\mathbf{r})$ is defined as self correlation function. If the particle is randomly oriented, then the orientational average of the intensity must be performed:

$$I_1(q) = < f(q)^2 >_\Omega = 4\pi \int_0^\infty p(r)\frac{\sin(qr)}{qr}dr \tag{2.6.7}$$

$$\gamma(r) = < \rho(\mathbf{r})\rho(\mathbf{r}^*) >_\Omega \tag{2.6.8}$$

$$p(r) = r^2\gamma(r) \tag{2.6.9}$$

where the p(r) is called pair distance distribution function [20]. Thus, the intensity of the a system with N particles is

$$I(q) = N < f(\mathbf{q})^2 >< S(\mathbf{q}) > \tag{2.6.10}$$

Where $< f(\mathbf{q})^2 >$ is the particle form factor, and $< S(\mathbf{q}) >$ is called the structure factor which considers possible interparticle interactions or system ordering effects.

The scattering length density of the particles and solvent are ρ_p and ρ_o, respectively. For particles distributed in the solvent, only the electrons density difference, $\rho_p - \rho_o$ i.e. contrast, is effective, and if $\rho_p = \rho_o$, the X-rays can not "see" the particles (see Fig.2.13) [21, 23]. In a dilute system, all the distances between particles are considered to be large enough or completely uncorrelated, and it means that inter-particle correlations can be neglected. if all the particles in the system are the same, this system we call it

16
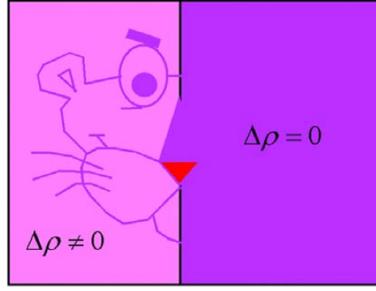
Figure 2.13: The contrast in SAXS is the difference of electron densities ($\Delta\rho = \rho_p - \rho_o$) between particle and environment (left). The right in the figure is invisable due to zero contrast [21].

monodisperse system. Scattering length is a term that describe the ability of an electron to scatter X-rays or other radiation. Because X-rays interact with the electrons in the particle, the scattered intensity function of a isolated particle [18] can be written in this way:

$$I_1^{SAXS}(q) = (\rho_p - \rho_o)^2 < \int_{V_p} e^{i\mathbf{q}\cdot\mathbf{r}}\, dV_p >^2 \qquad (2.6.11)$$

where the $\mathbf{q}$ and $\mathbf{r}$ are the scattering vector and displacement between electrons in the single particle. The integral in 2.6.11 is the sum of all the electrons in the particle and assuming the electrons are continuously distributed in the particle. The magnitude expression of the scattering vector is

$$|\mathbf{q}| = \frac{4\pi}{\lambda}\sin\theta \qquad (2.6.12)$$

where λ and θ are the wavelength of incident beam and the half of the scattering angle(Figure 2.1). Every displacement $\mathbf{r}$ is measured relative to the wavelength λ of the applied radiation, so in order to be independent from the λ, the scattering intensity is usually presented as function of $\mathbf{q}$ [21]. By introducing the normalised particle amplitude form factor,

$$F(q) = \frac{1}{V_p} < \int_{V_p} e^{i\mathbf{q}\cdot\mathbf{r}}\, dV_p >, \qquad (2.6.13)$$

the scattered intensity function of a fixed particle can be indicated

$$I_1^{SAXS}(q) = \Delta\rho^2 V_p^2 |F(q)|^2, \qquad (2.6.14)$$

and $\Delta\rho = (\rho_p - \rho_o)$ is the scattering contrast between particles and solvent.
Form factor is the scattering caused by the morphology of the particle, and it fulfils the condition: $F(q = 0) = 1$. Only few cases have the analytical mathematical expression in the integral in (2.3.1), and in most situation, the integrals only have numerical solution. When the particle is sphere with the radius R, probably the simplest case. Due to its symmetry and it is not necessary to consider orientational averaging. Therefore, the

form factor of sphere can be calculated as [18]:

$$F(q) = \frac{1}{V_p} \int_0^R \int_0^{2\pi} \int_0^{\pi} e^{iqr\cos\theta} r^2 \sin\theta \, d\theta \, d\phi \, dr = \frac{1}{V_p} \int_0^R 4\pi \frac{\sin(qr)}{qr} r^2 \, dr$$
$$= 3\left[\frac{\sin(qr) - qr\cos(qr)}{(qr)^3}\right]. \tag{2.6.15}$$

The form factor for a cylinder with radius R and length L has no analytical expression and it is numerical expression is [24]:

$$|F(q)|^2 = \int_0^{\pi/2} \left[\frac{2B_1(qR\sin\alpha)}{qR\sin\alpha} \frac{\sin(qL\cos\alpha)/2}{(qL\cos\alpha)/2}\right]^2 \sin\alpha \, d\alpha \tag{2.6.16}$$

, where $B_1(x)$ is the first order of the Bessel function,and the α is the angle angle between the axis of the cylinder and $\mathbf{q}$.

Some systems are intrinsically consisted of particles in different sizes, which are polydisperse system. We can use distribution functions to describe the variation of sizes. For particles with spherical shape, the number radii distribution D(R) and the scattering intensity is given by:

$$I_1^{SAXS}(q) = \Delta\rho^2 \int_0^{\infty} D(R) V_p(R)^2 |F(q,R)|^2 \, dR \tag{2.6.17}$$

The number particle radii distribution is normalised such that $\int_0^{\infty} D(R) \, dR = 1$. For cases that the interparticle interactions can not be ignored, the structure factor $S(q)$ need to be introduced. Using a local monodisperse approximation for including the effects of polydispersity [25], the intensity function changes to be

$$I_1^{SAXS}(q) = \Delta\rho^2 \int_0^{\infty} D(R) V_p(R)^2 |F(q,R)|^2 S(q,R) \, dR. \tag{2.6.18}$$

Depending on the system to be investigated, it is necessary to choose different distributions. Actually, according to data analysis experience, there are some suggested application. If the polydispersion is to describe particle sizes, people often apply the Lognormal or Schulz distributions. If it is to describe interfacial thicknesses, the Gaussian or Boltzmann distributions are suggested [26]. In a SAXS experiment, the total intensity of the illuminated sample is:

$$I^{SAXS}(q) = \sum_{n=1}^{N} I_n^{SAXS}(q) = \eta\Delta\rho^2 V_p(R)^2 |F(q,R)|^2 S(q,R) \tag{2.6.19}$$

$$= \phi\Delta\rho^2 V_p(R) |F(q,R)|^2 S(q,R) \tag{2.6.20}$$

where N is the total number of particles in the sample, η ($\eta = \frac{N}{V_{sample}}$) is the number density of the particles in the sample (i.e. concentration if the sample is solution), and ϕ is the volume fraction, ($\phi = \frac{V_{total\,particles}}{V_{sample}} = \frac{NV_{particle}}{V_{sample}} = nV_{particle}$).
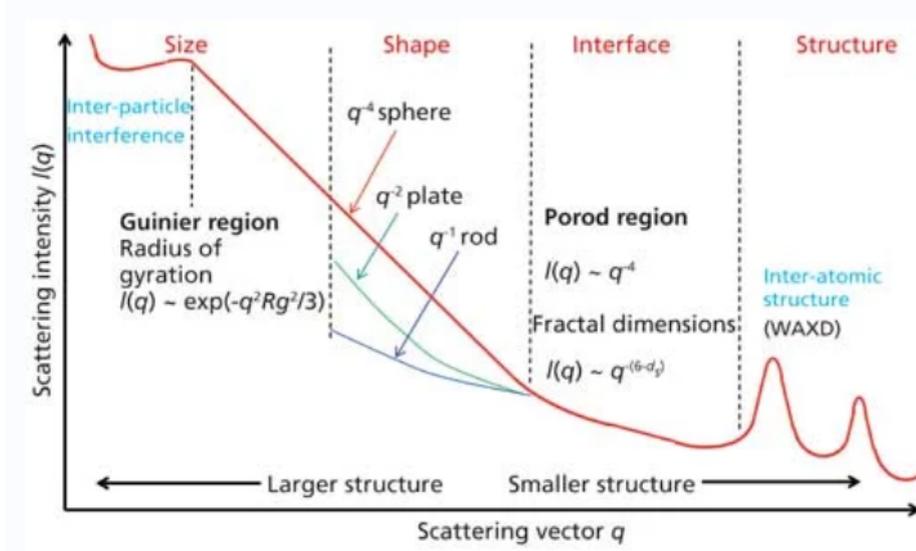
Figure 2.14: The schematic of different q corresponding to different size structure. Guinier and Porod region are the long wavelength and short wavelength limit, respectively. [27]

Form factor, P(q), in the long and short wavelength limit has different approximation as shown in Fig.2.14.

Long Wavelength Limit (i.e. $qr \rightarrow 0$)

In this limit, the form factor, P(q), can be approximated by [18]:

$$P(q) \approx a_0 \cdot e^{\frac{-R_g^2 q^2}{3}} \tag{2.6.21}$$

where R_g is the radius of the gyration. This equation was first derived by Guinier in 1939, so this region also called Guinier region. The radius of gyration R_g of a partilce is the root-mean-squared distance from the particle's center of gravity. If the scattering length density is homogeneous, spherically symmetric and the radius of gyration squared is

$$R_g^2 = \frac{\int_{V_p} r^2 dV_p}{V_p} \tag{2.6.22}$$

If the scattering length density is spatially distributed, and the radius of gyration is given by

$$R_g^2 = \frac{\int_{V_p} \rho_p(\mathbf{r}) r^2 dV_p}{\int_{V_p} \rho_p(\mathbf{r}) dV_p} \tag{2.6.23}$$

Short Wavelength Limit (i.e. $qr \rightarrow \infty$)

In the short wavelength limit, in which the wavelength is small compared with the particle size, but still large compared with the inter-atomic spacing, the SAXS intensity

for a spherical particle becomes [14]

$$I_1^{SAXS}(q) = |F(q)|^2 = \frac{2\pi\Delta\rho^2}{q^4}S_p \qquad (2.6.24)$$

which is proportional to the surface area of the particle, and inversely proportional to the fourth power of q. According to the shape of particle, q has a different Porod exponent n. The Form factor depend sensitively on particle dimension. We can have a roughly understanding of this fact though the element of integration in E.q.2.6.13. When the particle is sphere shape, the element of integration is $dV_p = 4\pi r^2 dr$, so the integrand varies as r^2. The intensity is the square of the form factor and therefore it varies as r^4. When the particle is a two-dimensional object, an infinitesimally thin plate, with radius R, and the element of integration is an area $dA_p = 2\pi r dr$. When considering one-dimensional object, an infinitesimally thin rod of length L, the element of integration is a constant independent of r. For the rough interface, the Porod exponent would be in the format of $-(6-d)$, with the help of fractal mathematics. [18] All these expressions are provided in table 2.2.

Table 2.2: Variation with dimension d of SAXS single particle form factor square P(q), radius of gyration, R_g, and Porod exponent. Results are given for a sphere of radius r, and infinitely thin plate of radius r, infinitely thin rod of length L, and the particle volume V_p. Here Si(x) is sine integral $\int_0^x \sin t / t dt$, $B_1(x) = (\sin x) - x\cos x)/x^2$ is a Bessel function of the first kind, and $<>_\Omega$ means performing a orientational average [18].

	P(q)	Radius of gyration R_g	Porod exponent n
Sphere (d = 3)	$(\frac{3B_1(qr)}{qr})^2$	$\sqrt{\frac{3}{5}}r$	-4
Plate (d=2)	$\frac{2}{q^2r^2}(1 - \frac{B_1(2qr)}{qr})$	$\sqrt{\frac{1}{2}}r$	-2
Rod (d =1)	$\frac{2Si(qL)}{qL} - \frac{4\sin^2(\frac{qL}{2})}{(qL)^2}$	$\sqrt{\frac{1}{12}}L$	-1
Generalized Form	$< \frac{1}{V_p}\int_{V_p} e^{i\mathbf{q}\cdot\mathbf{r}}\,dV_p >_\Omega^2$	$\sqrt{\frac{\int_{V_p} r^2 dV_p}{V_p}}$	$-(6-d)$

2.7   Modeling Method

There are mainly two approaches to analyse small-angle scattering data, model independent and model dependent. Both approaches need to apply least-squares methods [20, 24]. In the model-independent approach, we basically perform a Fourier Transformation (FT) of the experimental scattering curve I(q), and the obtained curve is the pair distance distribution function, p(r). FT is usually done by the Indirect Fourier Transformation (IFT) that was initially proposed by Glatter [23]. Because this project is related to the improvement of a model dependent method, the focus will be direct to it from now on.

Model dependent methods use mathematical models to describe the experimental
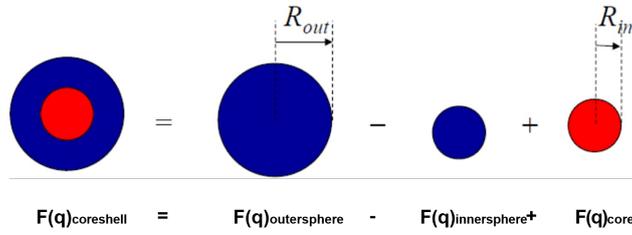
Figure 2.15: The schematic of constructing the amplitude form factor of core shell sphere, and it is equal the amplitude form factor of a sphere with larger radius (R_{out}) minus the amplitude form factor of inner sphere and plus core sphere with different electron density ρ_{core}.

SAXS data. We can have analytical or semi-analytical expressions of the form factor of simple shape particle to calculate the scattering intensity with E.q. 2.6.19. The main advantage of this is that there are only some parameters to adjust to fit the experimental data, allowing the determination of structural information with reasonable reliability. If the model does not fit the data correctly, this means that the real shape of particle is different from the assuming shape [20]. For the form factor of a complicated shape, it is possible to use known subunits to construct the particle shape. For example, form factor of core shell sphere can be built up in the method shown in Fig.2.15.

Least square method is employed to find the best fit for the SAXS experimental data, and chi-squared (χ^2) function or chi-squared test is used as a measure for the deviation between the experimental data and the model [20, 24]. Given a set of experimental intensity points $I_{exp}(q_i)$, i = 1,...,N, and q_i is the modulus of the scattering vector corresponding to the measured intensity. The counting statistics will bring the statistical uncertainties σ_i on the intensity point $I_{exp}(q_i)$. The χ^2 is given as:

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{I_{exp}(q_i) - I_{mod}(q_i)}{\sigma_i} \right)^2 \tag{2.7.1}$$

where $I_{mod}(q_i)$ is the model intensities which depends on the parameters in the model. It is common to use reduced χ^2, which is defined as:

$$Prob = \frac{\chi^2}{N - M} \tag{2.7.2}$$

where N is the number of data points, M is the number of independent fitting parameters in the model, and $N - M$ is called number of degrees of freedom (ndf). If a good fit is achieved, Prob should be closed to 1, which in the ideal case, Prob is equal to 1, which means the deviations $|I_{exp}(q_i) - I_{mod}(q_i)|$ are on average equal to the statistical uncertainties [24]. If Prob is greatly larger than 1, it might indicate important differences between the model and experimental data. It also indicates that the uncertainties are underestimated. On the other hand, if Prob is greatly lower than 1, the uncertainties may be overestimated [20]. The χ^2 function (i.e. 2.7.1) is minimized during the fitting, and this can be done through many different methods. For instance, making a reasonably initial guess on the values of the parameters, and then simply varying the values one by

one in a meaningful bound, so that consequently lower values of χ^2 are obtained. This simple grid search method could work but time consuming. In this project, a function curve_fit in scipy package [28] is applied to perform the least-squares to fit the model function.

Chapter 3

SBA-15 Model and MC Methods in Modeling SAXS data

3.1 The Analytical Model of SBA-15

Now we come back to SBA-15, which exhibits a hierarchical structure with micro, meso and macropores. This implies that to analyse the measured data, the SAXS model must consider these three different size pores contributing to the scattering intensity. Therefore, the scattered intensity per unit volume for the mesostructure in SBA-15 can be written in this:

$$I_{meso}(q) = (\Delta\rho)^2 n_d < F(q)^2 > S(q); \tag{3.1.1}$$

where $\Delta\rho$ is the scattering contrast between surrounding solvent and the cylinder material, n_d is the number density of the particles in the sample, F(q) is the Form factor of the particle, and S(q) is the structure factor.

$$S(q) = 1 + \beta(q)(< Z(q) > -1)G(q) \tag{3.1.2}$$

Z(q) is the lattice factor describing the spatial distribution of the particles, and the angular brackets$< ... >$means an average operation in terms of the particle size and spatial distribution of particles, which are assumed to be independent. These two equations imply a decoupling approach that allows one to factorize the scattering intensity into the contributions from the form factor and the structure factor.

The square of form factor $P(q) =< F(q) >^2$ for long cylinder can be approximately factorized into longitudinal and cross-section contribution of the particle [24, 29]. Thus $P(q)$ [1] becomes

$$P(q) = P_{rod}(q)P_{CS}(q) \tag{3.1.3}$$

Then, the longitudinal factor $P_{rod}(q)$ is taken as the form factor of an infinitely long thin rod:

$$P_{rod} = \frac{2Si(qL)}{qL} - \frac{4\sin^2\frac{qL}{2}}{(qL)^2} \tag{3.1.4}$$

where

$$Si(x) = \int_0^x t^{-1} \sin t \, dt \tag{3.1.5}$$

and L and R are the length and radius of the cylinder [24]. The cross-section contribution $P_{CS}(q)$ is

$$P_{CS}(q) = (\frac{2B_1(qR)}{qR})^2 \tag{3.1.6}$$

where $B_1(x)$ is the first order of the Bessel function. Using the Miller indices (h k l) for a given crystal lattice plane, the lattice factor for an ideal undistorted lattice, $Z(q)$ has the expression [29]:

$$Z(q) = \frac{(2\pi)^{d-1}c}{nv_d\Omega_d q^{d-1}} \sum_{hkl} m_{hkl} f_{hkl}^2 L_{hkl}(q) \tag{3.1.7}$$

where n is the number of particles per unit cell, v_d is the volume (d=3), surface (d=2), or long-period (d=1) of the d-dimensional unit cell, Ω_d is the d-dimensional solid angle, f_{hkl} is the symmetry factor that takes into account symmetry-related extinction rules, $L_{hkl(q)}$ is a normalized peak shape function, and m_{hkl} is the reflection multiplicity. The mesopores in the SBA-15 are in a p6mm space group, 2-D hexagonal lattice structure, so Z(q) can be written in this format:

$$Z_0(q) = c\frac{2}{\sqrt{3}a^2 q} \sum_{hk} m_{hk} L_{hk}(q) \tag{3.1.8}$$

Here, c is a constant of order unity, which ensures that the product of form factor and structure factor fulfills the equation for Porod invariant Q, a is the lattice parameter, and m_{hk} is the multiplicity factor of the hexagonal lattice with values of 12 ($for\, h \neq k \neq 0$) and 6. Porod invariant Q is the integral over all **q** of the intensity, and it is given by

$$Q = \int_0^\infty q^2 I(q) dq \tag{3.1.9}$$

In the q range of interest, there are five obvious peaks in the SAXS data of the SBA-15 corresponding Miller indies are (100), (110), (200), (210), and (300). L_{hk} is the peak shape functions, and they have the expression:

$$L_{hk}(q) = \frac{2}{\pi\delta} |\frac{\Gamma[\frac{v}{2} + \frac{i\gamma_v 2(q-q_{hk})}{\pi\delta}]}{\Gamma[\frac{v}{2}]}|^2, \int_{-\infty}^{+\infty} L_{hk}(q)\, dq = 1 \tag{3.1.10}$$

where $\Gamma[z]$ is a complex gamma function, q_{hk} determines the peak position, and v controls the shape of the peaks. In the limit $v \to 0$, the peak function L_{hk} is a Lorentzian form, and $v \to \infty$, it is a Gaussian form. δ is the width of the peak, which is related to the domain size D of the ordered domain through the Dybye-Scherrer function:

$$D = \frac{2\pi}{\delta} \tag{3.1.11}$$

G(q) is similar to the Debye-Waller factor, which describes the thermal disorder, and given by

$$G(q) = e^{-(\sigma_a aq)^2} \tag{3.1.12}$$

$\beta(q)$ includes the effect of the polydispersity of the cylinders in their radii to the interference term, and its expression is

$$\beta(q) = \frac{< F_{CS}(q) >^2}{< F_{CS}(q)^2 >} \tag{3.1.13}$$

$$< F_{CS}(q) > = \frac{\int_0^\infty D(r) F_{CS}(q,r)\, dr}{\int_0^\infty D(r)\, dr} \tag{3.1.14}$$

$$< F_{CS}(q)^2 > = \frac{\int_0^\infty D(r) F_{CS}(q,r)^2\, dr}{\int_0^\infty D(r)\, dr} \tag{3.1.15}$$

In the equation, the brackets $< ... >$ means the average of all possible cylinder sizes, weighted by a distribution D(r), which in our model is Schulz-Zimm distribution, which expression is given by

$$D(r) = (\frac{z+1}{\bar{r}})^{z+1} \frac{r^z}{\Gamma(z+1)} exp(-(z+1)\frac{r}{\bar{r}}) \tag{3.1.16}$$

Here, $\Gamma(x)$ is the real gamma function, $\bar{r}$ is the mean radius of the cylinder, and z is a measure of the spread in the radii. Here we consider that the relative polydispersity $pd = \sigma_r/r$, and $z = \frac{1}{pd^2} - 1$. In the implementation of this model, we only add the polydispersity into the inner radii, because the intensity pattern are very similar even if the polydispersity into the thickness t of the cylinder shell is added. Considering the smeared outer surface of the core-shell cylinder, the cross-section form factor $F_{CS}(q)$ is described by [24]

$$F_{CS}(q) = r_{out}^2 \frac{2J_1(qr_{out})}{qr_{out}} exp[-\frac{1}{2}q^2\sigma_{int}] - (1 - \frac{\Delta\rho_{in}}{\Delta\rho_{out}})r_{in}^2 \frac{2J_1(qr_{in})}{qr_{in}} \tag{3.1.17}$$

$$r_{out} = r_{in} + t \tag{3.1.18}$$

where $\Delta\rho_{in}/\Delta\rho_{out}$ is the ratio between the inner and outer electron density contrast of the cylinder, r_{out}, r_{in}, and t are the outer, inner radii, and thickness of the cylinder shell, respectively, and σ_{int} is the width of the smeared Gaussian interface.

In the high q range, the scattering intensity pattern is dominated by the microstructure, and it can be described by the Gaussian chain form factor: [1]

$$I_{micro}(q) = \frac{2[exp(-q^2R_g^2) - 1 + (qR_g)^2]}{(qR_g)^4} \tag{3.1.19}$$

Table 3.1: Summarizing of all parameters contained in the SBA-15 model.

Sc_1	scale for the mesoporous structure contribution
Sc_2	scale for the Guassian chain contribution
Sc_3	scale for the macrostructure contribution
c	correction for Porod invariant
a	lattice parameter
D	domain size
σ_a	disorder parameter
r_{in}	inner radius of the cylinder
t	thickness of the cylinder
$\Delta\rho_{in}/\Delta\rho_{out}$	ratio between inner and outer contrast of the cylinders
σ_{int}	smearing width of the outer shell
pd	relative polydispersity of radii
R_g	radius of gyration of the Gaussian chain
bkg	the backgroud intensity
L	length of the cylinder
ν	peak shape

where R_g is the radius of the gyration of the chain, and it is a fitting parameter in the model.

According to Porod's law, $I_{Porod}(q)$ is only contributing in the low q-range ($q < 0.01 \mathring{A}^{-1}$) and comes from interface between the grains [14]:

$$I_{Porod}(q) \simeq \frac{2\pi A}{q^4} \tag{3.1.20}$$

where A is the outer surface area.

Thus, the final expression for the SBA-15 model is

$$I(q) = Sc_1 P_{rod}(q) < F_{CS}(q)^2 > (1 + \beta(q)[< Z(q) > -1]G(q))$$
$$+ Sc_2 I_{micro}(q) + Sc_3 I_{Porod}(q) + bkg \tag{3.1.21}$$

where Sc_1, Sc_2, and Sc_3 are the scale factors for the ordered hexagonal mesopores, the microstructure and the macrostructure contribution, respectively. The bkg term accounts for the background intensity. The fitting parameters are summarised in the table 3.1 above.

3.2 MC in SAXS Model

To test the MC fit, we need to simulate the SAXS data set with known distribution, and try to recover it. Due to symmetric property of Gaussian distribution, we use it in the implementation, and we try to apply Monte Carlo (MC) method to extract the distribution. N_s random radii from Gaussian distribution with the mean radius $\bar{r}$ and standard deviation $\sigma = pd\ \bar{r}$, will be used to simulate the scattering intensity with the

function 2.6.19.

The MC method used here is essentially identical to that used by Pauw et al [9]. Before implementing the MC method in the sphere system, a reasonable radius range has to be estimated, and it needs the help of scattering curve. Scattering vector $q = \frac{4\pi \sin\theta}{\lambda}$, λ the wavelength of the radiation and 2θ the scattering angle, all can get from the SAXS experiment. A reasonable estimated maximum and minimum radii are $R_{max} \simeq \pi/q_{min}$ and $R_{min} \simeq \pi/q_{max}$ respectively, so in the MC fitting, the radius range should be set to $R_{min} < R \leq R_{max}$ or $0 < R \leq R_{max}$ [9].

The initial guess of the total scattering intensity I_{MC} is calculated through n_s sphere whose radii R are randomly sampled from a bound uniform distribution, and the range of the bound is in between R_{min} and R_{max} (i.e. $R \in [R_{min}, R_{max}]$). The initial guessed intensity is calculated using E.q. 2.6.19.

$$I_{MC}(q) = b + A \sum_{k=1}^{n_s} |F_{sph,k}(qR_k)|^2 (\frac{4}{3}\pi)^2 R_k^{(6-p_c)}, \tag{3.2.1}$$

$$A = \frac{\phi \Delta\rho^2}{\sum_{k=1}^{n_s} \frac{4\pi}{3} R_k^{(3-p_c)}}. \tag{3.2.2}$$

$F_{sph,k}(qR_k)$ is the sphere form factor for sphere k, which has analytical expression 2.6.15. R_k is the radius for sphere k. A is the scaling factor and b is a constant background term. p_c is a computational aid in the range $0 \leq p_c \leq 6$, biasing the volume weighting of the contributions, and the choose of p_c would not affect the result. It is because p_c is compensated in A and p_c only affect the computational speed. Because the scattering intensity is proportional to the volume square, it means the scattering intensity of a sphere with radius 10 Å is 10^6 stronger than the one with 1 Å.

$$\phi_i = \frac{A}{\Delta\rho^2} \frac{4\pi}{3} R_i^{(3-p_c)} \tag{3.2.3}$$

$$\phi = \sum_{k=1}^{n_s} \phi_i \tag{3.2.4}$$

The recommended value for p_c is 3, because this value is one of the fastest values [9]. In the function 3.2.1, A and b are adjusted to make the I_{MC} fits the measured scattering intensity through applying the least-square method [24]. There is package in SciPy, *scipy.optimize.curve_fit*, using non-linear least squares to fit a function [28], that is more convenient to use this package to fit E.q.3.2.1, and thus will be used in this thesis.

The iteration cycle runs as the following.

1. Randomly selecting the ns sphere radii from a uniform distribution in a reasonable bound that mention before.

2. The initial fitting intensity I, background b and the scale factor A can be calculated through *scipy.optimize.curve_fit* package.

3. The χ^2 between fitted and measured can be calculated. If the χ^2 is larger than the set criteria, then the MC iteration cycle start. If not, the fitting result will be exported.

4. A random sphere radius r is chosen from n_s spheres sample generated in the first step, and randomly select a new radius r_{new} for this sphere from the uniform distribution in the same bound. The new intensity after this change is calculated by subtracting the contribution of the previous sphere and adding the contribution of the new sphere radius, and a new scaling factor A and background term b are obtained by least-square fit.

5. Only when this radius change can decrease the χ^2 between measured and MC intensity, the change can be accepted. Otherwise, the choosing radius step will keep repeating. After enough iteration times, the χ^2 decreases to smaller than the set criteria, and the iteration will stop. Or when the iteration times, $i \leq 10^6$, the optimization would stop. Finally, the ns sphere radii data set, A, b, and the fitting curve will be exported. The MC optimization process when no clues about the distribution shows in the flowing chart 3.1.
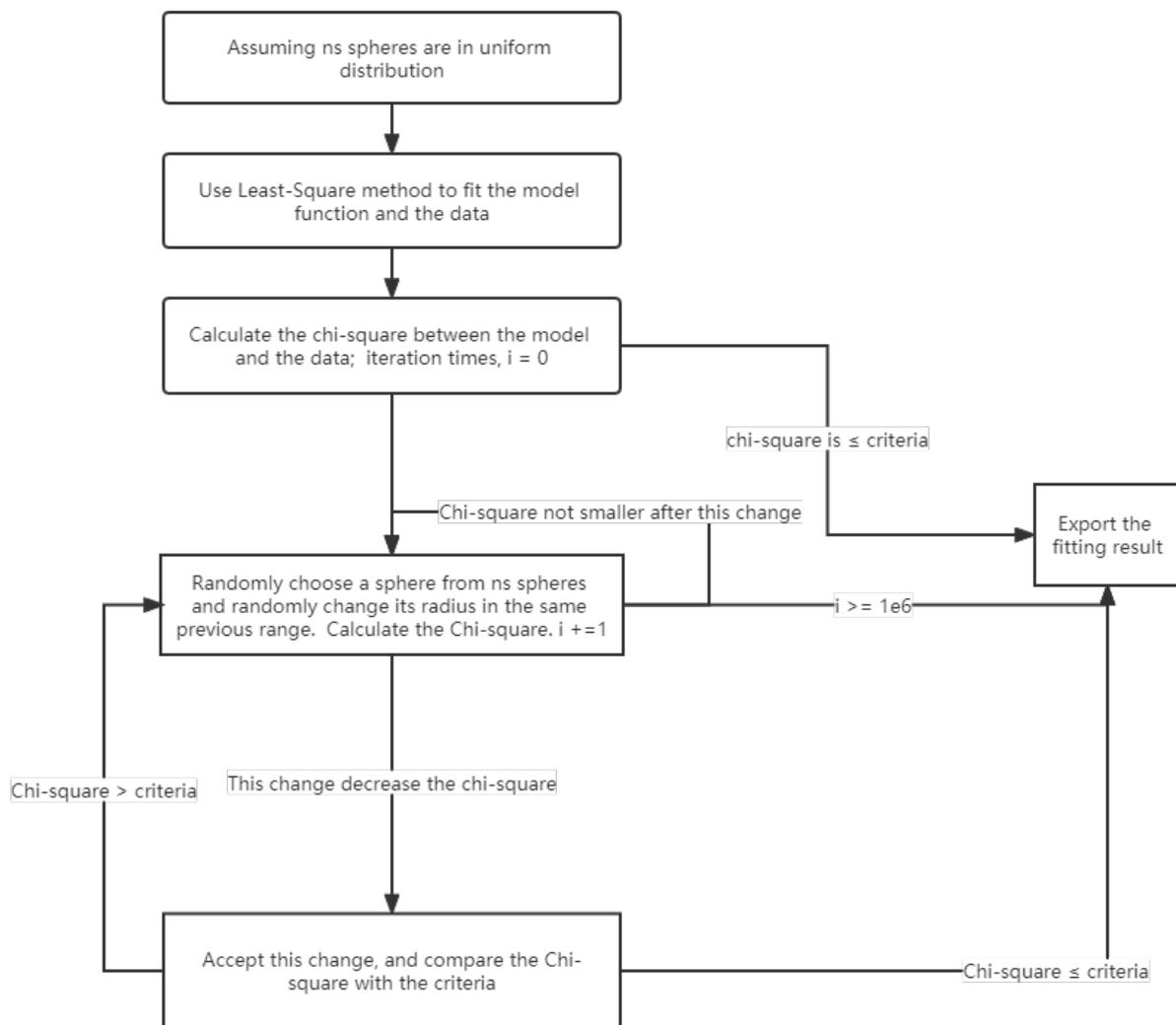
Figure 3.1: The flowing chart of the Monte Carlo optimization in fitting the sphere shape SAXS data

28

The MC functions for a long cylinder are

$$I_{MC}(q) = b + A \sum_{k=1}^{n_s} P_{cross-section}(q, R_k) P_{rod}(q, L) \pi^2 R_k^{(4-p_c)} L^2, \qquad (3.2.5)$$

$$A = \frac{\phi \Delta \rho^2}{\sum_{k=1}^{n_s} \pi R_k^{(2-p_c)} L}. \qquad (3.2.6)$$

where $P_{cross-section}(q, R)$ (E.q. 3.1.6) and $P_{rod}(q, L)$ (E.q. 3.1.4) are the cross-section and infinity thin rod form factor respectively. The MC functions of core-shell long cylinder are different from the long cylinder, and they are

$$I_{MC}(q) = b + A \sum_{k=1}^{n_s} P_{core-shell}(q, R_k) P_{rod}(q, L), \qquad (3.2.7)$$

$$A = \sum_{k=1}^{n_s} \phi_k \Delta \rho_{out}^2 \pi \qquad (3.2.8)$$

$$\phi = \sum_{k=1}^{n_s} \phi_k = \frac{A}{\pi \Delta \rho_{out}^2} \qquad (3.2.9)$$

where $P_{core-shell}(q, R_k)$ are the square of core shell form factor (E.q. 3.1.17). For MC fit to the SBA-15 the function becomes

$$I_{MC} = sc_1 \sum_{k=1}^{N_s} P(q, R_k, L) S(q, R_k) + sc_2 I_{micro} + sc_3 I_{porod} + bkg \qquad (3.2.10)$$

$$P(q, R_k, L) = P_{rod}(q, L) < F_{CS}(q, R_k)^2 > \qquad (3.2.11)$$

$$S(q, R_k) = 1 + \beta(q, R)[< Z(q) > -1]G(q) \qquad (3.2.12)$$

The loop to run the MC functions are the same as the sphere (i.e. Fig. 3.1).

Chapter 4

Results and discussions

4.1   Implementing the SBA-15 model in Python

In these work three SBA-15 SAXS data sets, a, b and c, which precedence and measurement information were given in the table 2.1, were analysed. All these SBA-15 samples have basically the same morphology.

From the CoSAXS experiment, we collected patterns in 20 different points along the quartz cylinder in b and c, and all scattering curves are shown in the appendix, and Fig. 4.1b, and Fig. 4.1c represents the data we discuss now. The Bragg reflections in SBA-15_c scattering curve, from left to right, represent the Miller index of (10), (11), (20), (21), and (30), respectively. SBA-15_b is highly inhomogeneous, therefore we observe double reflections (Fig. 4.1b). The "twin reflections" are the combination of at least two lattices with similar but not the same lattice parameter a.

From Fig. 4.1c, we can observe that the scattering curve is proportional to the q^{-1} for $q < 0.006 Å^{-1}$, and proportional to the q^{-4} for $q > 0.006 Å^{-1}$. The intensity curve reflects the particle size and surface area between $q < 0.006 Å^{-1}$ and $0.006 Å^{-1} < q < 0.01 Å^{-1}$, respectively. As q increases, there is a transition from macroscale to mesoscale. The implemented model in this project mostly takes into account the mesoscale features, i.e. the mesopores, and it does not mainly consider macroscale features, except for the domain size, D, which affects the peak width only. Thus, the q range from 0.01 to 0.35 $Å^{-1}$ is a trustful range for this model.

Implementing Sundblom's model [1] in a reasonable q range, to analyse SBA-15_a and SBA-15_c leads to reduced χ^2 equal to 25.9 and 30676, respectively. For the sample SBA-15_c, CoSAXS data, the reduced χ^2 between model the experimental is considerably larger than forSBA-15_a, and the reason may be the uncertainty for SBA-15_c is underestimated. From the fit of SBA-15_c (Fig. 4.1c), the model will start to break with the q raising. It is most likely because even the micropores are included in the model through the Gaussian chain form factor, but this is too simplistic compared to the way that we consider the mesopores. Also, we are assuming a continuum electron contrast in the model, which is not true for the reason that discrete the molecular feature.
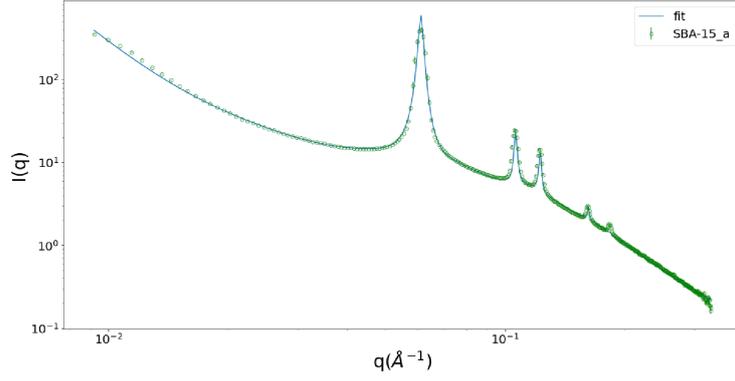
(a) The scattering curve of SBA-15_a and it's fit.
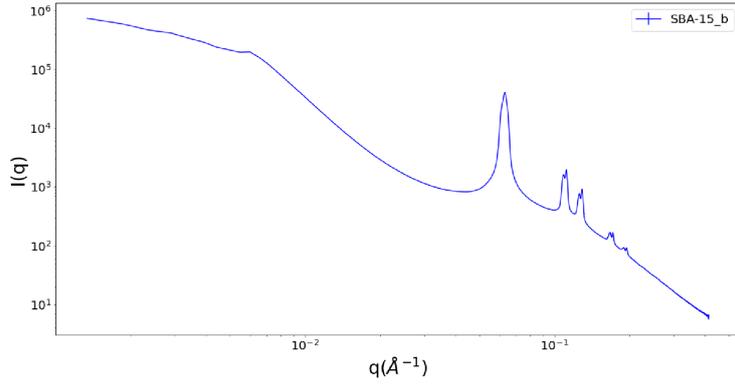
(b) The scattering curve SBA-15_b.
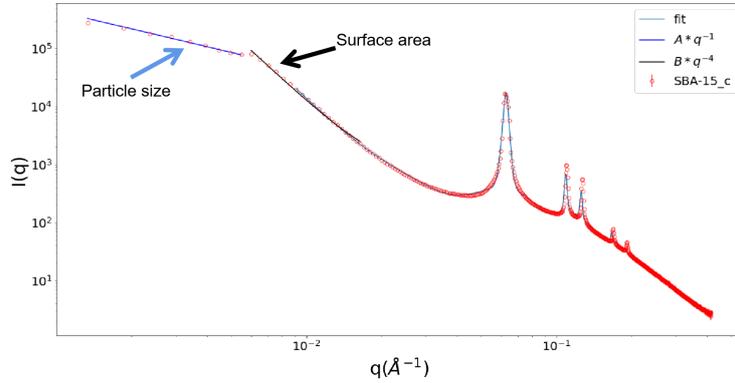
(c) The scattering curve SBA-15_c and it's fit.

Figure 4.1: a is the experimental scattering curves of SBA-15_a (green circles) and it's fit (blue line) for $0.01 < q < 0.33 Å^{-1}$. b is the experimental scattering curves of SBA-15_b. c is the experimental scattering curve (red circles) and it's fit (light blue line) with the Sundblom's model [1] for $0.01 < q < 0.33 Å^{-1}$. For $0.001 < q < 0.006 Å^{-1}$ and $0.006 < q < 0.01 Å^{-1}$ are the fit with the function $A * q^{-1} + bkg$ (dark blue line) and $B * q^{-4} + bkg$ (black line), respectively.

From the fit SBA-15_a, we can obtain the mean radii, r_{in} and the thickness, t, of the mesopores are $52.7 \pm 0.5 Å$, and $13 \pm 1.9 Å$. The polydispersity, pd, of the inner radius of the mesopores is 0.145 ± 0.002. The values of fitting parameters of SBA-15_a are concluded in the table 4.1. The value of domain size, D, in the fitting result is not the real one, because D is mainly affected by the peak width. The peak width is also easily affected by the instrumental smearing, and in our implementation, we did not include the smear-

31
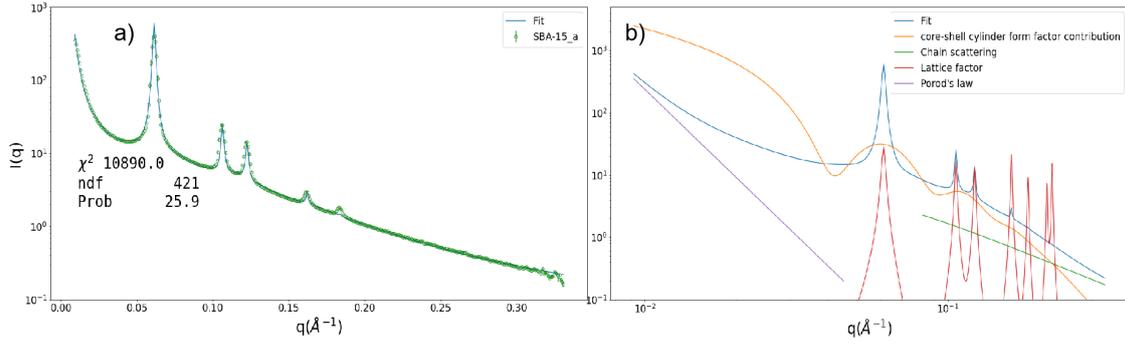
Figure 4.2: The green open circles are the experimental SBA-15_a SAXS data with the X-ray wavelength $\lambda = 1.54\text{Å}$, and the blue line is the fit (Fig.a). The χ^2, number of degree of freedom (ndf), and reduce χ^2 (Prob), are given in the figure. Different contributions to fit (blue line) are showed on the Fig.b. Porod's law (purple line) mostly contributes in the low q range, and the core shell cylinder form factor (orange line), and Gaussian chain form factor (green line) dominate in the middle q range, and high q range, respectively. The lattice factor are shown in red line.

ing effect using the resolution function in the model. The ratio $\Delta\rho_{in}/\Delta\rho_{out}$ is closed to 0, because these materials are empty SBA-15, where there is nothing but the air in the mesopores.

Table 4.1: Summarizing all parameters fitting values contained in the SBA-15_a model

parameters	values	$(\pm)$uncertainty
Sc_1	0.13	0.01
Sc_2	7.8	0.7
Sc_3	2.01×10^{-7}	0.0
c	9.5	0.9
a	118.21	0.08
D	3900	224
σ_a	0.081	0.002
r_{in}	52.7	0.5
t	13	1.9
$\Delta\rho_{in}/\Delta\rho_{out}$	0.10	0.01
σ_{int}	40	13
pd	0.145	0.002
R_g	28	1.8
bkg	0.00	0.03
L	5000	4000
ν	0.15	0.02

Fig. 4.2 a is the fit of the SBA-15_a data and the Fig. 4.2 b shows the contributions of different terms in the fit. From Fig. 4.2 b, we can observe in the low q range, the Porod's law (purple line) dominates the fit (blue line), and it rapidly decreases with the increasing q. When comes to the middle q range, the form factor of the long core shell cylinder (orange line), $P_{core-shell}(q)$, is the main contribution of the fit. As q raises, the scattering

of the micropores (green line) takes over the $P_{core-shell}(q)$, and finally dominates in the high q range. The scattering curve is the product of the structure (or lattice) factor S(q) with form factor P(q) of the particles, when the lattice reflection and the form factor minima collide, then a so-called systematic extinction follows and the peaks cannot be observed. This is the reason why the last two peaks with the miller indices (22), and (31) do not appear in the fit (Fig. 4.2 b). In this model, it only fits the first four peaks of the scattering curve well, and the fifth peak in the fit is weaker than the experimental (Fig. 4.2 a). The distribution of the radius of the mesopores would affect the form factor contribution, and assuming distribution (Schulz-Zimm distribution) may affect the form factor thereby decreasing the fitting quality.

Together these factors indicates that the complexity of this analytical model lead us towards the first attempt for developing a Monte Carlo (MC) approach to analyse SAXS data. The obtained results are discussed in the next chapter of the thesis. However, before achieving this goal, we needed to apply and validate our own MC models for simpler systems firstly. This is discussed in the next section.

4.2 Implementing MC models to retrieve distribution in Python

4.2.1 Retrieve Spheres' Distribution

Gaussian
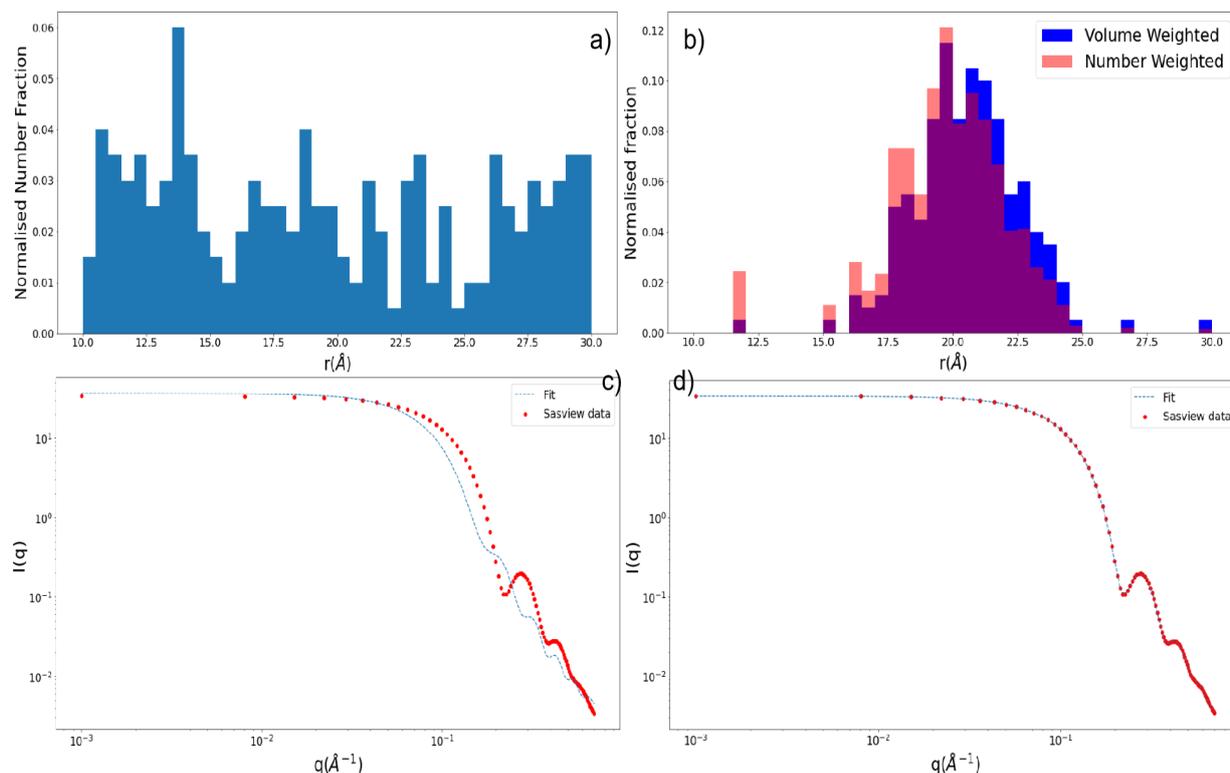
Figure 4.3: Comparison between before and after MC fit. a) is the initial guess distribution of the n_s radius, and it is a uniform distribution; b) is the distribution after MC fit showed in normalised number (pink) and volume (blue) weighted; c) is the fit under uniform distribution, i.e. before performing MC fit (blue dash line), and the red dots are the SasView intensity for $0.001 \leq q \leq 0.7 \mathring{A}$; d) is the MC fit (blue dash line) the SasView intensity (red dots).

Scattering curve includes the information about particle size distribution, and MC fit can help us to retrieve it. Before implementing the MC fit to the experimental scattering curve, we need to test it theoretically. More information about generating the theoretical SAXS intensity curve (red dots in Fig. 4.3 c) can be found in Appendix. In the theoretical test, assuming there is 1% uncertainty relative to every intensity value. As mentioned in Chapter 3.2, $n_s = 200$ sphere radii are randomly chosen from a uniform distribution that in a reasonable bound ($r \in [10, 30]$), and the normalised number weighted size distribution of the radius is the Fig. 4.3a. We use least-square method to fit the E.q.3.2.1, and then we can get the scattering curve (blue dash line in Fig. c). The fitting is not well, because the distribution in SasView intensity curve is not the uniform distribution. Now, we apply the MC fit to this intensity curve, and we set the criteria is 1, which means that for $\chi^2 < 1$, the MC fit stops, and the scaling factor $A = 1.07 \times 10^{-6}$, the

background term $b = 2.20 \times 10^{-5}$ values are exported. The distribution of the n_s radius can represent the distribution of the irradiated sample (Fig. 4.3 b). It can be seen that around 300 times successful changes in the n_s radii, and its distribution turns into the Gaussian distribution. Every radius R_k corresponds to a volume fraction ϕ_k, and every bin value is equal to the sum of volume fraction ϕ_k in the corresponding radius range (the blue bins in Fig.b). Thus, this particle size distribution is volume weighted. In terms of radius range, except for the last bin is closed on the left and closed on the right (i.e. $R_i \in (R_{i-1}, R_{i+1}]$), and the others are closed on the left and open on the right (i.e. $R_i \in [R_{i-1}, R_{max}]$). Before we calculate the mean radius, we need to transform into number weighted distribution (the pink bin in Fig.b) through dividing by the sphere volume $\frac{4\pi r^3}{3}$ in every bin, and we observed that the extracted mean radius $r = 19.87\text{Å}$ is quite closed to set mean radius 20Å. In SasView, the particle size distribution is implemented as number-weighted [26], but the way we calculating the scattering intensity in MC method is in volume-weighted. Thus, when comparing the extracted distribution with original one, we need to do a transformation.

In order to better compare the difference between extracted and origin distribution, we
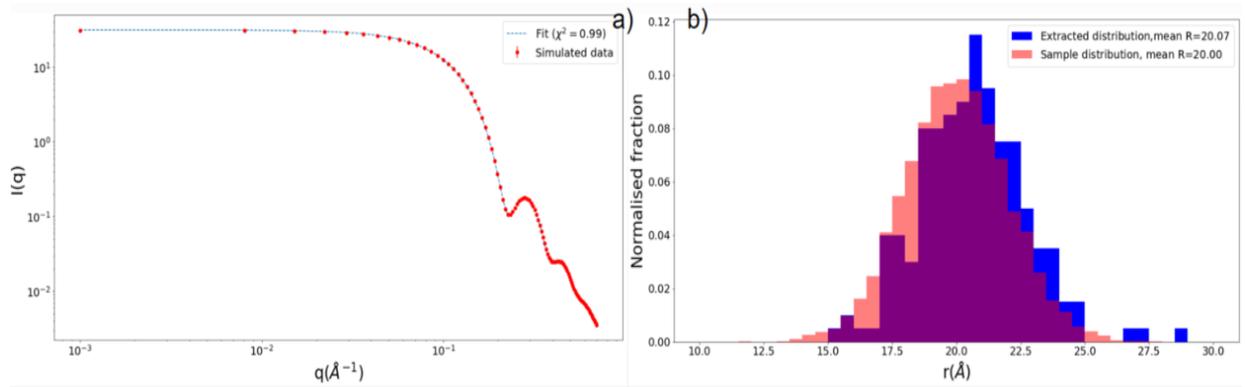
Figure 4.4: The MC fit of the simulated data, and the blue dash line is the MC fit and red dots are the simulated data(a).The blue histogram is the extracted distribution in one MC loop, and the pink one is sample distribution(b).

simulate SAXS data curve with a sample containing 10^4 sphere particles, which radii are in Gaussian distribution with the mean radius $R_{mean} = 20\text{Å}$ and $\sigma = 2\text{Å}$. The $pd = \frac{\sigma}{R_{mean}} = 0.1$. The volume fraction ϕ, scattering contrast $\Delta\rho$ of this system are 1 and $3 \times 10^{-6}/\text{Å}^{-2}$. The comparison between simulated and SasView intensity is in the Appendix.5 and it is to be sure that obtain the correct simulated intensity. We implement MC fit to the simulated scattering curve, and the extracted distribution (blue bins), and the fit to the intensity curve are presented in Fig. 4.4. The final scale factor A and background are 1.70×10^{-6} and 4.14×10^{-5} respectively.
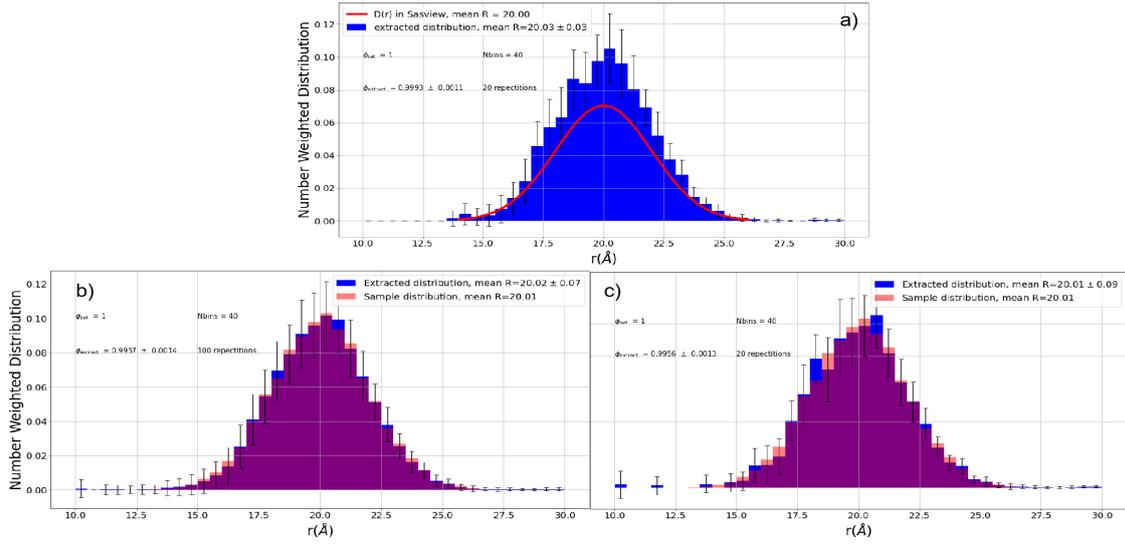
Figure 4.5: Normalised number weighted distribution are extracted by the MC fit the intensity curves. Fig.a shows the extracted (the blue histograms) and original (red line) distribution for 40 bins. Fig.b and c show the extracted (the blue histograms) and the sample (the pink histograms) distribution for 40 bins. Every bin value and error bar is the mean and standard deviation over 20 (a and c) or 100 (b) repetitions

In order to avoid arriving at the radius that is valid but unrelated during the MC process, we implemented 20 independent MC fits to the simulated and Sasview intensity, and obtained the Fig.4.5 a and c, respectively. Every bin and error bar value in Fig.4.5 are the mean and standard deviation over 20 or 100 repetitions. The extracted mean radius, and volume fraction are $R = 20.03 \pm 0.03 \mathring{A}$, and $\phi = 0.9993 \pm 0.0011$ closed to the set R = 20 and $\phi = 1$ in Sasview. The calculation of all 20 repetition takes around 6 minutes for Intel i7-8550U CPU. If we run the MC loop 100 times, we could obtain Fig.4.5 b. The retrieval volume fraction $\phi = 0.9957 \pm 0.0014$ is closed to the simulated $\phi = 1$. The standard deviation or error bar of the radius R only decreases very little compared with run the MC loop 20 times. However, the consuming time increase greatly, and it takes around 30 minutes to run 100 times.

Log-Normal

We also tested other distributions. Such as, the log-normal distribution D(r) in Sasview given by [11],

$$D(r) = \frac{1}{Norm} \frac{1}{r\sigma} exp(-\frac{1}{2}(\frac{ln(r) - \mu}{\sigma})^2) \tag{4.2.1}$$

where Norm is a normalisation factor which will be determined during the numerical calculation, $\mu = ln(x_{med})$ and x_{med} is the median value of the log-normal distribution, and $\sigma = pd$ is a parameter describing the width of the underlying normal distribution. We fit the intensity curve when pd = 0.1 and 0.2, assuming the uncertainty is $0.01 * I$, then we can get the Fig. 4.6. In Fig. 4.6 b and e, First chi2 are defined as the χ^2 value

before accepting the new radius, and the Second chi2 are define as the χ^2 value after accepting the new radius.
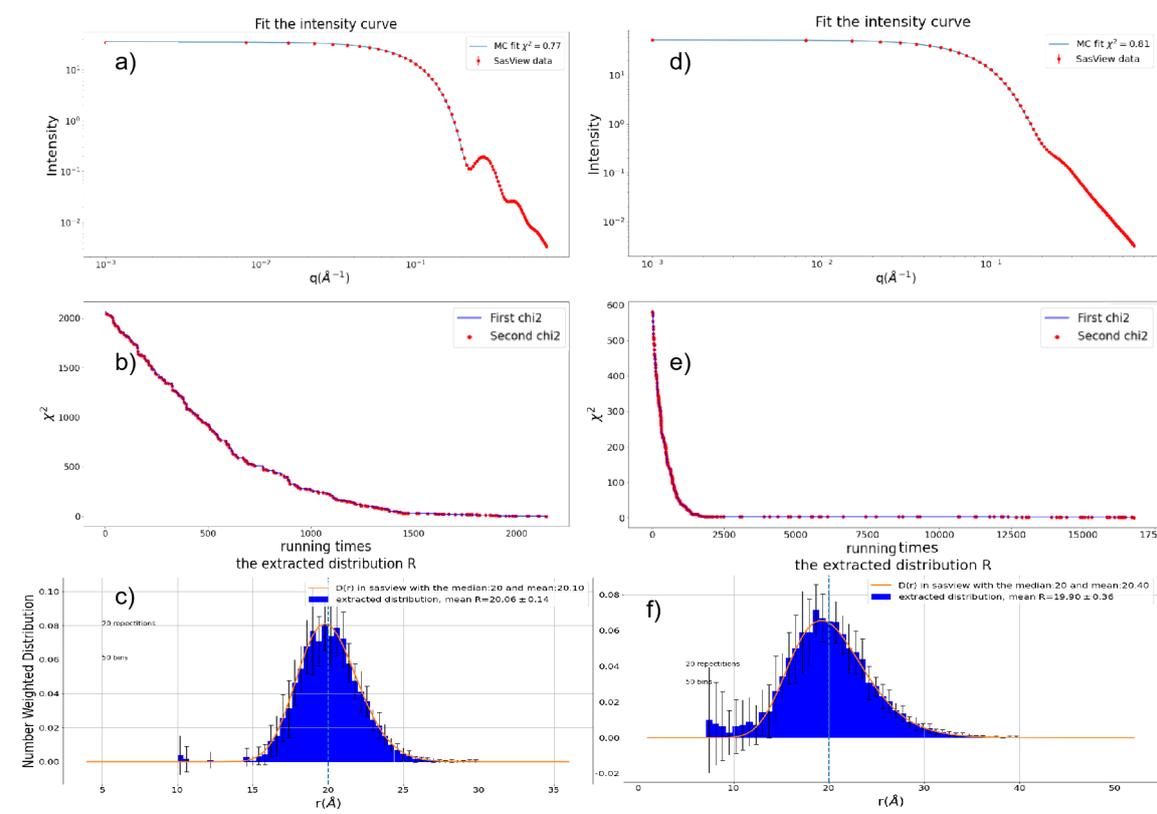
Figure 4.6: Fitting SAXS curve for the $0.001 \leq q \leq 0.7\AA^{-1}$ when D(r) is log-normal distribution. In Fig.a, the red circles are the sphere scattering intensity simulated with SasView in the log-normal distribution that pd = 0.1, and the blue line is the MC fit with the $\chi^2 = 0.77$. Fig.b shows how the χ^2 decreases during the MC fitting SasView data (pd = 0.1), the red dots are the second chi2, and the blue line is the First chi2. Fig.c shows the extracted distribution (blue histograms) from scattering intensity (red circles in Fig.a) and the distribution D(r) in SasView (orange line). Figs.d-f are similar but pd = 0.2.

Afterwards, we implemented 20 independent MC fit, and then the mean retrieved distributions of pd = 0.1 and 0.2 were obtained (Fig.4.6 c and f). For higher pd values the longer running time are needed (Fig.4.6 b and e). The higher pd gives rise to higher standard deviation. When applying the MC fit for the pd = 0.2, the assuming radius range should be more constrained. The radius range in Fig4.6 a is around between 5 and 40 $\AA$, and the assuming radius range set in MC can not be too wide especially for the minimum value. For example, if the set R range in MC is $R \in [0, 40]$, the MC still can fit the intensity curve, but we can not obtain a reasonable extracted distribution. The minimum radius that can be "seen" in this q range is $R_{min} \approx \frac{\pi}{q_{max}} = 4.5\AA$, so some small radii are impossible to be seen but they may be accepted during MC process.

Bimodal
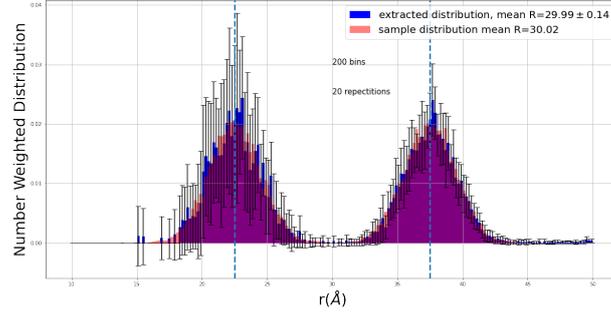
(a) CD = $15\mathring{A}$, $R_{ex} = 29.99 \pm 0.14\mathring{A}$, and $R_{sam} = 30.02\mathring{A}$
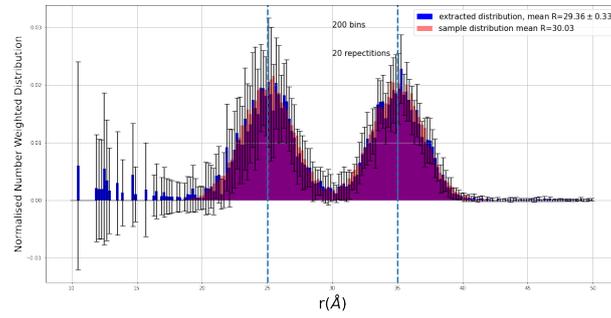
(b) CD = $10\mathring{A}$, $R_{ex} = 29.36 \pm 0.33\mathring{A}$, and $R_{sam} = 30.03\mathring{A}$
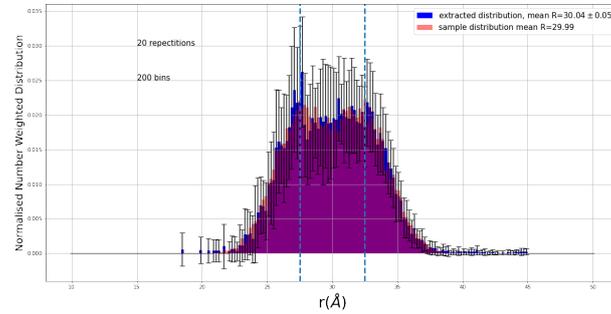
(c) CD = $5\mathring{A}$, $R_{ex} = 30.04 \pm 0.05\mathring{A}$, and $R_{sam} = 29.99\mathring{A}$

Figure 4.7: The recovered distribution (normalised number weighted distribution) over 20 independent MC fit, and the number of bins is 200. The blue bins and the error bars are the mean the standard deviation over 20 times, respectively. The pink bins are the sample distribution, and the blue dash lines show the center of the Gaussian distribution. The R_{ex} and R_{sam} are the extracted mean and the mean of the sample, respectively.

We also tested the bimodal distribution, which was built by summing Gaussian distributions with different means and same standard deviation ($\sigma = 2\mathring{A}$). For example, $r_1 = 22.5\mathring{A}$ and $r_2 = 37.5\mathring{A}$ and the center difference (CD) between these two distribution is $15\mathring{A}$ (i.e. CD = $15\mathring{A}$). The scattering contrasts is the same as previous and the sample particles are equal 2×10^4. We apply the MC fit the scattering intensity when CD = 15, 10, and 5 $\mathring{A}$, and the average recover distribution over 20 times are shown in the Fig.4.7 above. The R_{ex} is smaller than the R_{sam} and is outside more than one σ

(Fig.4.7b). The reason is that too many small but unwanted radii are accepted in the MC fit, and we can constrain the guessed radii range to get a better mean. The guessed radii range in Fig4.7b is $R \in [10, 50]$, and in Fig4.7c is $R \in [15, 45]$.

4.2.2 Test MC Fit With Real Data

We tested the MC method in real experimental data that was provided by Erik S. Brok in the students' Experimental X-rays course (UCPH, 2021). This is a SAXS curve of a protein solution. Through the fitting in SasView with Gaussian distribution function, the protein can be well described by sphere shape. Applying the MC fit to this protein solution assuming that the distribution is uniform, and the radius is in the range from 0 to 35 $\mathring{A}$. We obtained the results shown in the Fig.4.8. In the MC fit, we set the criteria is 1.2 considering the consuming time. Because the reduced χ_2 converges very slowly after 1.2. The contrast $\Delta \rho = 9.4 \times 10^{-6}/\mathring{A}^{-2}$ that we used is fitted value in SasView, and the volume fraction fitted in SasView is $\phi_{Sas} = 0.205\%$. Then, we use this $\Delta \rho$ to calculate the extracted volume fraction ϕ_{ext} of the protein in the solution, the mean $\phi_{ext} = 0.297 \pm 0.002\%$ through 20 independent MC fit. It takes around 40 minutes to complete the 20 repetitions. The scale in the red line is to normalised the D(r), and r^3 is to transform the number weighted distribution into volume weighted.
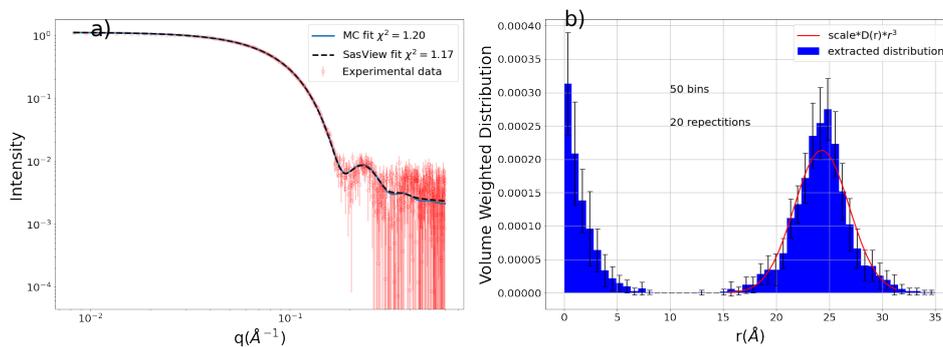
Figure 4.8: The figures are fit of a protein solution scattering intensity curve for $0.008 \leq q \leq 0.6\mathring{A}$(a) and the extracted volume fraction (b). In Fig.a, the blue solid line is the MC fit with the $\chi^2 = 1.20$, the black dash line is the fit in Sasview with the $\chi^2 = 1.17$, and the open red circles are experimental scattering intensity. Fig.b is the volume weighted size distribution for 50 bins. The blue bins are the average of 20 independent MC fit, and the error bars on the bin are standard deviation over 20 repetitions. D(r) is the Gaussian distribution with the mean 23.4 $\mathring{A}$ and $\sigma = 2.6\mathring{A}$ in SasView describing the polydispersity.

SasView still can fit the scattering curve well even it does not include the small particles (i.e. the particles are in the bins for r in the range of [0,10] in Fig4.8b.). The reason is that the scattering intensity is proportional to V_p^6, and the measured scattering intensity is mainly scattered by the "big" particles. This may be also one of the reasons that ϕ_{ext} is larger than ϕ_{Sas}. However, more information about the protein solution is needed to confirm the volume fraction. We found if the initial guessed distribution set to be Gaussian distribution and the selected radii distribution is uniform distribution, the

MC fit spends less time to converge down to the criteria.

4.2.3   Retrieve Long Cylinders' Distribution

Gaussian

We also applied MC in the long cylinder shape particles to retrieve the radius distribution of the cylinder. The form factor of the long cylinder(i.e. $L > 10R$) can be approximately well into two parts, cross-sectional and longitudinal contribution, and mesopores in SBA-15 are considered into a long core-shell cylinder shape[29]. In SasView, the scattering intensity of cylinder shape particle is calculated through E.q.2.6.16, and our simulated data calculate in a approximately (E.q.3.1.3). We use 10^4 radii to generate the simulated data, the radius R, length L, scattering contrast $\Delta\rho$, polydispersity pd in inner radius R and volume fraction ϕ are equal $30\mathring{A}$, $3000\mathring{A}$, $4 \times 10^{-6}\mathring{A}^{-2}$, 0.133, and 1 respectively. The intensity curve of SasView is used to validate the simulated scattering curve (see Fig.5.4a). We apply the MC fit in the simulated scattering curve. After applying the MC fit for 20 times, average distribution is showed in the Fig.4.9a.
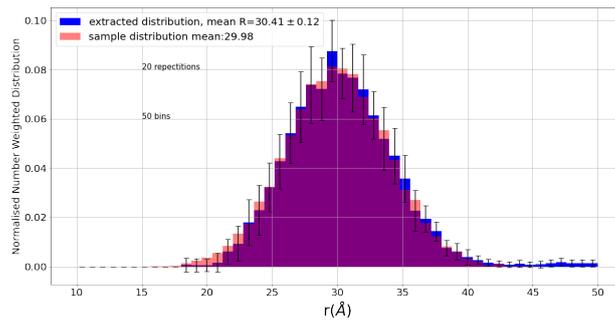
(a) Long cylinder.
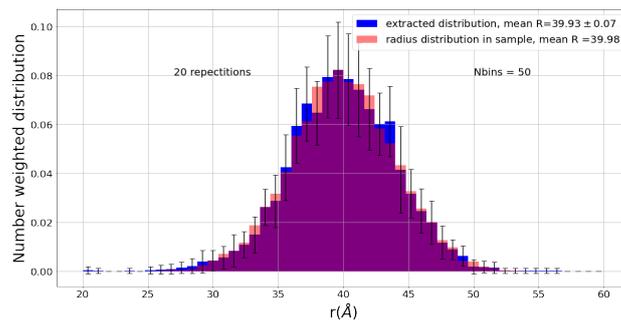
(b) Long core shell cylinder

Figure 4.9: 4.9a is the average of 20 times extracted distribution. The blue bins are the average of the extracted distribution, and the pink ones are the distribution in the simulated data. The mean of the extracted and the sample are $R = 30.41 \pm 0.12\mathring{A}$ and $29.98\mathring{A}$ respectively. 4.9b is the average of 20 times MC fit of the core-shell cylinder. The mean radius of extracted distribution is $39.93 \pm 0.07\mathring{A}$, and it in the sample is $39.98\mathring{A}$.

Considering the mesopores in SBA-15 is long core shell cylinder shape, we implement

the MC fit to the long core-shell cylinder, the mean inner radius R, thickness of the shell t, pd, length of the cylinder L, outer electron contrast $\Delta\rho_{out}$, and contrast ratio $\frac{\Delta\rho_{in}}{\Delta\rho_{out}}$ are 40Å, 20Å, 0.1, 4000Å, $3 \times 10^{-6}\text{Å}^{-2}$, and $1/3$ respectively. $\Delta\rho_{out}$ is the electron density difference between shell and the solvent, and $\Delta\rho_{in}$ is the electron density difference between core and the solvent. We run the MC fit for 20 times, and then we can get Fig. 4.9b.

Log-Normal

We also applied MC in the long cylinder with log-normal distribution (same median radius 30Å but different pd), and the process is similar to applying MC in sphere with log-normal distribution. The results are shown Fig. 4.10.
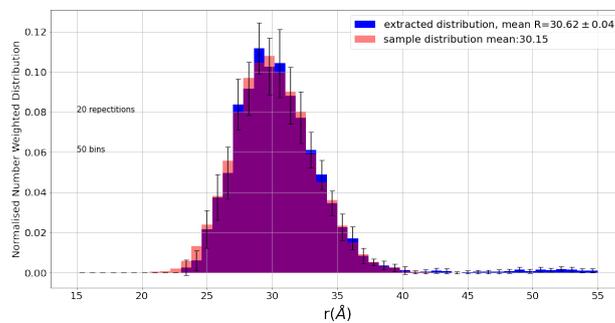
(a) pd = 0.1
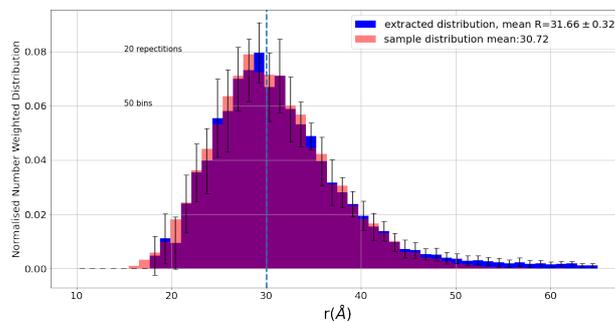
(b) pd = 0.2

Figure 4.10: The figures show the average over 20 times to recover the log-normal distribution. The blue bins are the average of the extracted distribution, and the pink ones are the distribution in the simulated scattering intensity curve. The mean of the extracted and the sample are $R = 30.62 \pm 0.04\text{Å}$ and 30.15Å respectively in Fig4.10a. The mean radius of extracted distribution is $30.66 \pm 0.32\text{Å}$, and it in the sample is 30.72Å in Fig. 4.10b. The dash line shows the median radii 30Å of sample.

4.2.4 First Attempt to SBA-15

Finally, We attempted to apply the MC method into SBA-15_a. Considering many parameters in the model, we only varied the Sc_1, Sc_2, Sc_3, and bkg parameters. All the

other parameters are fixed and their values are shown in table 4.1. We also assume n_s radii distribution change to be Gaussian distribution but in the selecting radii distribution keeps in the uniform distribution. It would take less time in the MC fit if the assuming distribution is similar to real distribution. The assuming radii are in range of 10 to 80 $\mathring{A}$, the criteria is set to be 5. However, we found the value of χ^2 fluctuate considerably, and did not converge (see Fig.4.11) during the MC fit.
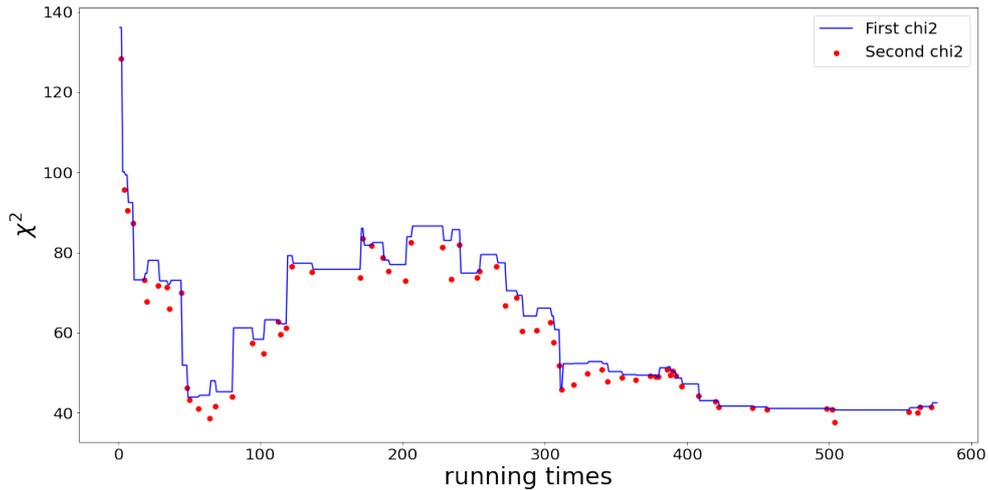
Figure 4.11: The χ^2 behaviour of the first attempt to use MC to SBA-15_a.

The first chi2 does not follow the second chi2 as expected during the MC fit, and the reason may be when including the structure factor S(q) in the MC fit, the previous way to calculate the intensity is not suitable. It is possible to combine the full SBA-15 model with the MC fit. We conclude that the MC optimization strategies need to be improved to speed up the process and to have the χ^2 under control.

Chapter 5

Conclusions and Perspectives

The goal of this thesis was the first attempt to incorporate the MC method in the Sundblom's model [1] in order to retrieve the SBA-15 mesopores size distribution from the corresponding SAXS data in a free modeling approach, i.e., without any assumption on the profile of such distribution. To achieve this goal, we wrote the codes of simple shape particle's SAXS models (i.e. sphere, long cylinder, and core shell long cylinder) in Python. After that, we validated these models against SasView, which is a widely used SAS data analysis software. Furthermore, we tested the MC method described on previous studies from the literature and used this approach to fit theoretical scattering curves of simple shape particles. With this method, We successfully recovered the distributions behind the theoretical scattering curves that were simulated in the previous step. Then, we started to implementing Sundblom's model [1] in Python and successfully applying it to fit SBA-15 experimental data. All the related Python scripts are in the link that shown in appendix. It required from the author of this thesis a deep understanding of the SAXS theoretical foundations of the models and the development of several computational skills.

Having hands on those, and considering that the scattering curves that we tested are too idealistic, so a "real case" was considered in which the Python code and algorithm were applied to satisfactorily fit a SAXS curve from a small globular protein in solution, yielding to the particle size distribution which agrees with the one obtained through the analysis with SasView software. Additionally, limitations of MC and validation of the results were extensively tested. MC was also adapted to deal with core-shell long cylinders, structures present in the SBA-15 model. Finally, in a first attempt, MC was incorporated to the SBA-15 model and used to fit the SBA-15 scattering curve. After several tests, a large χ^2 fluctuation was observed, preventing convergence of the method. Likely, the combination of the form factor calculation (where MC is) with the lattice structure factor is the source of such instability. Additionally, a much longer computing time was needed to process the fits compared to the previous tests using spheres and cylinders. This means that the complexity of the problem is greatly increased when dealing with SBA-15 model.

In conclusion, the study presented in this thesis gathered important experimental, theoretical and computational aspects of the considered problem and allowed to narrow it, providing basis for its continuation in the future.

In terms of perspectives, a deeper investigation on the source of the MC method instability is needed. For example, it would be interesting to test the MC against experimental scattering curve of long core shell cylinders. One way to speed up the data fitting is to combine the MC optimization into the least-square method. In this project the *Scipy* package used in the script may bring longer running time. Strategies to solve it along with the processing time speed up should be developed by experts in this field.

Bibliography

[1] Andreas Sundblom, Cristiano LP Oliveira, Anders EC Palmqvist, and Jan Skov Pedersen. Modeling in situ small-angle x-ray scattering measurements following the formation of mesostructured silica. *The Journal of Physical Chemistry C*, 113(18): 7706–7713, 2009.

[2] Vasu Chaudhary and Sweta Sharma. An overview of ordered mesoporous material sba-15: synthesis, functionalization and application in oxidation reactions. *Journal of Porous Materials*, 24(3):741–749, 2017.

[3] J Rouquerol, D Avnir, CW Fairbridge, DH Everett, JM Haynes, N Pernicone, JDF Ramsay, KSW Sing, and KK Unger. Recommendations for the characterization of porous solids (technical report). *Pure and applied chemistry*, 66(8):1739–1758, 1994.

[4] Dongyuan Zhao, Jianglin Feng, Qisheng Huo, Nicholas Melosh, Glenn H Fredrickson, Bradley F Chmelka, and Galen D Stucky. Triblock copolymer syntheses of mesoporous silica with periodic 50 to 300 angstrom pores. *science*, 279(5350):548–552, 1998.

[5] Martin Kjærulf Rasmussen, Heloisa N Bordallo, Marcela Aparecida Bordenalli, Milena Apetito Akamatsu, Aryene Góes Trezena, Milene Tino-De-Franco, Osvaldo A Sant'Anna, Tereza da Silva Martins, José Luiz de Souza Lopes, Márcia Carvalho de Abreu Fantini, et al. Assessing the efficiency of sba-15 as a nanocarrier for diphtheria anatoxin. *Microporous and Mesoporous Materials*, 312:110763, 2021.

[6] Martin K Rasmussen, Nikolay Kardjilov, Cristiano LP Oliveira, Benjamin Watts, Julie Villanova, Viviane Fongaro Botosso, Osvaldo A Sant'Anna, Marcia CA Fantini, and Heloisa N Bordallo. 3d visualisation of hepatitis b vaccine in the oral delivery vehicle sba-15. *Scientific reports*, 9(1):1–8, 2019.

[7] Cedric J Gommes, Gonzalo Prieto, and Petra E de Jongh. Small-angle scattering analysis of empty or loaded hierarchical porous materials. *The Journal of Physical Chemistry C*, 120(3):1488–1506, 2016.

[8] Adriana Zeleňáková, Pavol Hrubovčák, Ondrej Kapusta, Norbert Kučerka, Aleksander Kuklin, Oleksandr Ivankov, and Vladimír Zeleňák. Size and distribution of

the iron oxide nanoparticles in sba-15 nanoporous silica via sans study. *Scientific reports*, 9(1):1–9, 2019.

[9] Brian R Pauw, Jan Skov Pedersen, Samuel Tardif, Masaki Takata, and Bo Brummersted Iversen. Improvements and considerations for size distribution retrieval from small-angle scattering data by monte carlo methods. *Journal of applied crystallography*, 46(2):365–371, 2013.

[10] Ingo Breßler, Brian Richard Pauw, and Andreas F Thünemann. Mcsas: software for the retrieval of model parameter distributions from scattering patterns. *Journal of applied crystallography*, 48(3):962–969, 2015.

[11] SasView. about sasview, 2012. URL `http://www.sasview.org/`.

[12] Viveka Alfredsson and Hakan Wennerstrom. The dynamic association processes leading from a silica precursor to a mesoporous sba-15 material. *Accounts of Chemical Research*, 48(7):1891–1900, 2015.

[13] Marcello Crucianelli, Bruno Mattia Bizzarri, and Raffaele Saladino. Sba-15 anchored metal containing catalysts in the oxidative desulfurization process. *Catalysts*, 9(12):984, 2019.

[14] Sabine Manet, Julien Schmitt, Marianne Impéror-Clerc, Vladimir Zholobenko, Dominique Durand, Cristiano LP Oliveira, Jan Skov Pedersen, Christel Gervais, Niki Baccile, Florence Babonneau, et al. Kinetics of the formation of 2d-hexagonal silica nanostructured materials by nonionic block copolymer templating in solution. *The journal of physical chemistry B*, 115(39):11330–11344, 2011.

[15] Tomás S Plivelic, Ann E Terry, Roberto Appio, Keld Theodor, and Konstantin Klementiev. X-ray tracing, design and construction of an optimized optics scheme for cosaxs, the small angle x-ray scattering beamline at max iv laboratory. In *AIP Conference Proceedings*, volume 2054, page 030013. AIP Publishing LLC, 2019.

[16] Maik Kahnt, Konstantin Klementiev, Vahid Haghighat, Clemens Weninger, Tomás S Plivelic, Ann E Terry, and Alexander Björling. Measurement of the coherent beam properties at the cosaxs beamline. *Journal of Synchrotron Radiation*, 28 (6), 2021.

[17] Notes from experimental x-rays course ucph 2021.

[18] Jens Als-Nielsen and Des McMorrow. *Elements of Modern X-ray Physics*, chapter 4, 5. John Wiley Sons, West Sussex, 2 edition, 2011.

[19] Antonella Balerna and Settimio Mobilio. Introduction to synchrotron radiation. In *Synchrotron radiation*, pages 3–28. Springer, 2015.

[20] Cristiano Luis Pinto Oliveira et al. Investigating macromolecular complexes in solution by small angle x-ray scattering. *Current trends in X-ray Crystallography*, pages 367–392, 2011.

[21] Heimo Schnablegger and Yashveer Singh. *The SAXS Guide*, chapter 2,3. Anton Paar GmbH, Austria, 4 edition, 2017.

[22] Domenico Lombardo, Pietro Calandra, and Mikhail A Kiselev. Structural characterization of biomaterials by means of small angle x-rays and neutron scattering (saxs and sans), and light scattering experiments. *Molecules*, 25(23):5624, 2020.

[23] Ingrid Pilz, Otto Glatter, and Otto Kratky. [11] small-angle x-ray scattering. In *Methods in enzymology*, volume 61, pages 148–249. Elsevier, 1979.

[24] Jan Skov Pedersen. Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting. *Advances in colloid and interface science*, 70:171–210, 1997.

[25] Jan Skov Pedersen. Determination of size distribution from small-angle scattering data for systems with effective hard-sphere interactions. *Journal of applied crystallography*, 27(4):595–608, 1994.

[26] Sasview user documentation, 2015. URL `https://www.sasview.org/docs/user/qtgui/Perspectives/Fitting/pd/polydispersity.html?highlight=polydispersity`.

[27] N Sanjeeva Murthy. Recent developments in small-angle x-ray scattering. *Recent Developments in Small-Angle X-Ray Scattering*, 2017.

[28] scipy.optimize.curve_fit. URL `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve_fit.html`.

[29] S Förster, A Timmann, M Konrad, C Schellbach, A Meyer, SS Funari, P Mulvaney, and R Knott. Scattering curves of ordered mesoscopic materials. *The Journal of Physical Chemistry B*, 109(4):1347–1360, 2005.

[30] Particle size distribution. URL `https://wiki.anton-paar.com/cn-cn/particle-size-distribution/`.

[31] Bruce B. Weiner. What is particle size distribution weighting: How to get fooled about what was measured and what it means?, 2011. URL `https://www.brookhaveninstruments.com/wp-content/uploads/2020/02/what-is-particle-size-distribution-weighting-brookhaven-instruments.pdf`.

Appendix

Python script

```
https://drive.google.com/drive/folders/1BNmdTBijQqYdLHk3rGlIR_FUHc48sOQN?usp=
sharing
```

Number and Volume Weighted Distribution
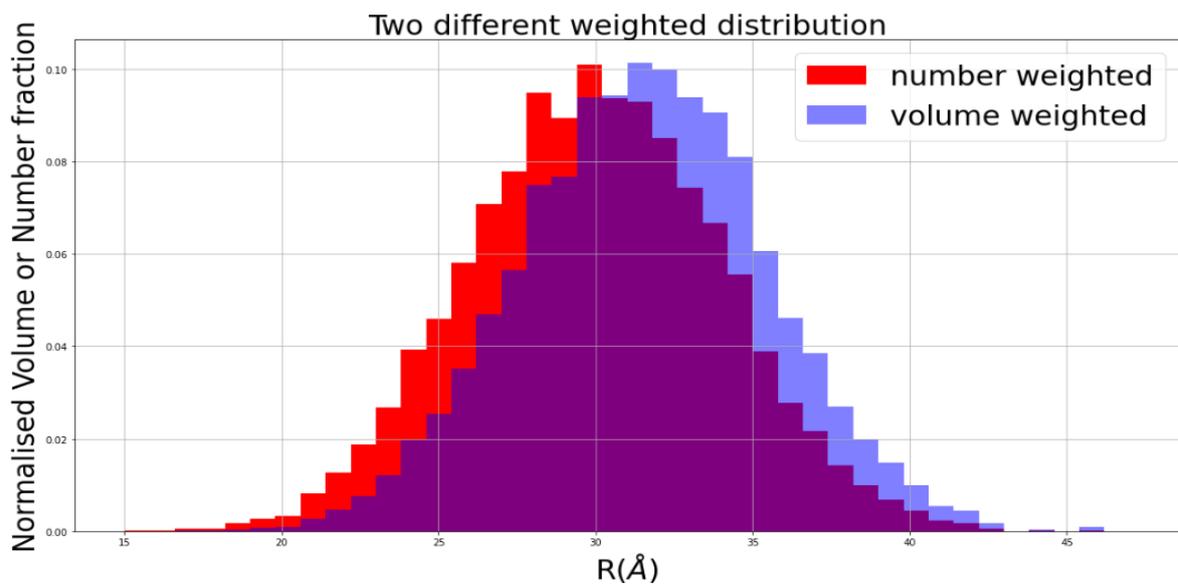
Figure 5.1: Two different weighted distribution but they are the same radii sample from a same normal distribution with the mean and σ are 30 and 4 Årespectively.
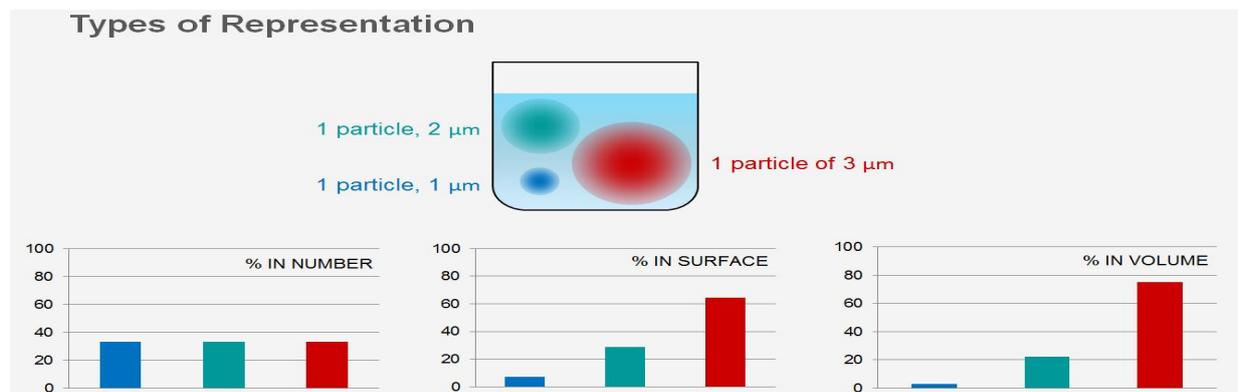
Figure 5.2: Theoretical particle size distribution of a simple mixture in different weightings.[30]

Particle size distributions can look very different even they are same data set, if they are weighted in different ways. Fig.5.1 shows the distribution of same radii data set but two different weighted. In number weighted the median value of the distribution is equal to the mean, but in volume weighted the median value would larger than the mean[31]. It reflects on the distribution graph is the volume weighted would skew to the large values. Different measurement techniques "see" the particles in a different way, which translates to a different weighting as the result. As an example, as a microscope sees the diameter of each particle, using this technique will deliver a number-weighted result. In contrast, the diffraction of light is proportional to the volume of the particle, therefore techniques such as laser diffraction or X-ray diffraction deliver volume-weighted results. The difference between these three weighted shows in Fig.5.2. It is easier to calculate the mean using the number weighted distribution, and the way transform volume weighted to number weighted is dividing by the particle volume.

Validating the SAXS Model

Before applying MC method, we need to be sure that we can simulate the correct data set. Here what we are doing is comparing our sphere and long cylinder shape particle SAXS model with the one in SasView [11], which is software to analyse SAXS data. The goal of this part is to be sure that the model is built correctly.

Sphere SAXS Model

The radius R of the sphere is 20 $\mathring{A}$, and the scattering length density of the sphere and the solvent are $4 \times 10^{-6}/\mathring{A}^{-2}$ and $1 \times 10^{-6}/\mathring{A}^{-2}$, respectively. Therefore, the scattering contrast $\Delta\rho$ is $3 \times 10^{-6}/\mathring{A}^{-2}$. Using E.q.2.6.15 under the monodisperse condition, the range of the scattering vector q divides evenly from 0.001 to 0.7 $\mathring{A}$ into 100 points, and we can plot the figure a).
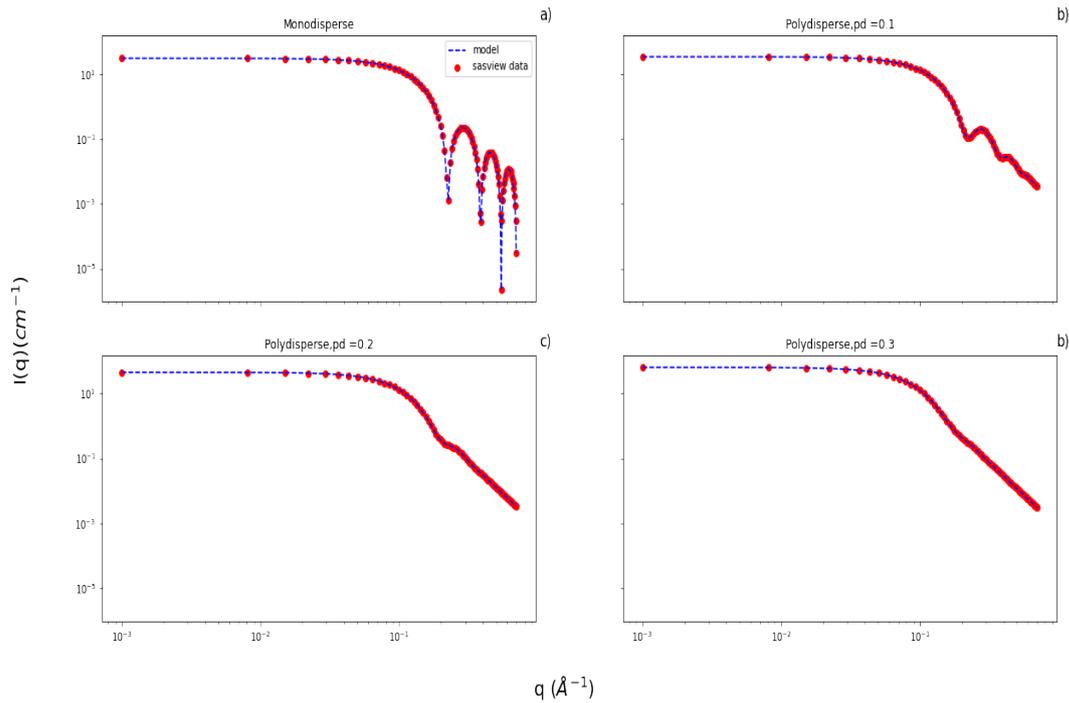
Figure 5.3: The SAXS curve of sphere particle. The red dots are the data exported from SasView, and the blue dash line is plotted in the same condition as set in the SasView. a): all the sphere are identical, or monodisperse; b),c),d) are in the polydisperse system, and the polydispersity is different, 0.1, 0.2, and 0.3 respectively.

pd is a parameter to describe how polydisperse the system is, and the larger pd the more polydisperse. Figure b),c),d) are in the polydisperse system, applying the E.q.(2.3.7) can plot the scattering curve. Here the particle size distribution D(R) is the Gaussian distribution, with the mean radius R_{mean} is 20 Å, and the $\sigma = pd \cdot R_{mean}$, so the σ for b),c),d) in Figure 3.1 are 2nm, 4nm, and 6nm, respectively. The radius R can get a infinite larger or negative value from a complete Gaussian distribution, although its possibility is very small. In order to avoid these extreme values, in actual modeling, a incomplete Gaussian will be applied. In this case, we only use the range of $[R_{mean} - 3\sigma, R_{mean} + 3\sigma]$ in the Gaussian distribution. Qualitatively, polydispersity plays a role to smear out the sharps feature in the SAXS curve. In Fig.3.1, we can see that introducing a small spread in the sphere radius of pd = 0.1 (Fig.a.) leads to rapid damping of the intensity oscillations. Keep increasing the pd, the smearing effect become more obvious, and the intensity oscillation almost disappears when $pd = 0.3$.

Now we are going to use SasView to simulate the scattering data and try to apply the MC method to extract the "hidden" distribution. The settings of the simulated data are mean radius $R_{mean} = 20\text{Å}$, pd = 0.1, and the distribution function D(R) is Gaussian distribution in $+-3\sigma(\sigma = pd\,R_{mean})$.

Long Cylinder SAXS Model
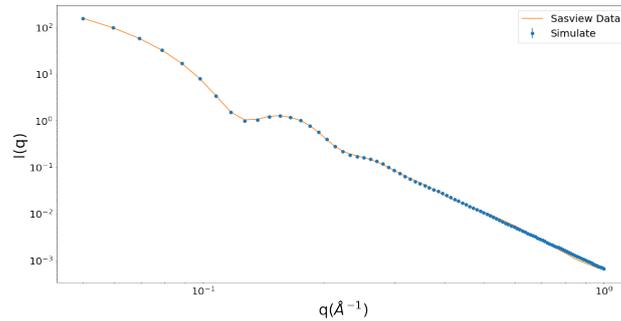
(a) Gaussian distribution, pd = 0.133
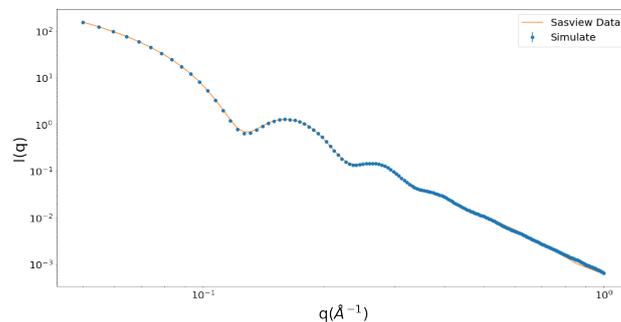
(b) Log-normal distribution, pd = 0.1

Figure 5.4: The comparison of the simulated and SasView data of long cylinder with Gaussian ($mean = 30\AA$, and $\sigma = pd \cdot mean$) and Log-normal ($\mu = ln(30)$ and $\sigma = pd$) distribution for q in the range of $0.05 \leq q \leq 1\AA^{-1}$. The blue dots are the simulated data, and the orange line is the SasView data.

The form factor F(q) of the cylinder in SasView is E.q.2.6.16, and the simulated intensities are calculated approximately with E.q.3.1.3. The length of the cylinder L, and the contrast $\Delta\rho$ are $3000\AA$ and $4 \times 10^{-6}/\AA^{-2}$. We can see the model is matched the SasView data very well, which suits our expectation, because the data exported from SasView is the theoretical data. If there is nothing wrong in our model functions, and settings in the model are the same, the scattering patterns should be the same.

Vertical scanning of the sample during the CoSAXS experiment

We perform the SAXS experiment of SBA-15_b and SBA-15_c with 20 different vertical points in CoSAXS, and the energy of the radiation is 12.4 keV (i.e. $\lambda = 1\AA$), the sample-detector distance is 3032 mm, and the exposure time is 1 second. Five representative scattering curves are shown in the figure below.
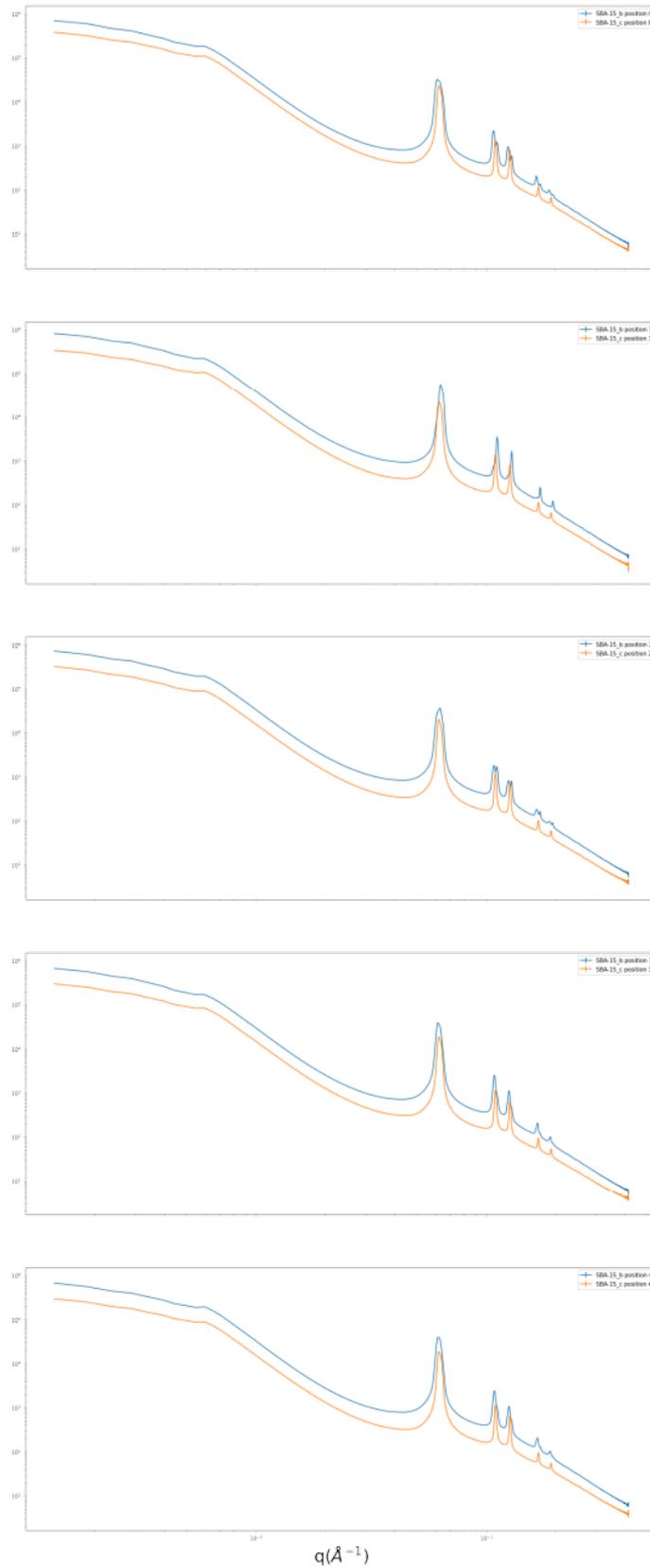
Figure 5.5: The scattering curves of different vertical points from position 0 to 4 for SBA-15_b (blue line) and SBA-15_c (orange line) for q in the range of $0.001 \leq q \leq 0.4 Å^{-1}$