Readout system for delay line detectors with a time stamp TDC and a small angle scattering study of intermolecular interactions in protein solutions

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von
Marc Niebuhr
aus Celle

Hamburg
2004

Gutachter der Dissertation:	Prof. Dr. R.L. Johnson
	Prof. Dr. M.H.J. Koch
Gutachter der Disputation:	Prof Dr. R.L. Johnson
	Prof. Dr. E. Mandelkow
Datum der Disputation:	23.09.2004
Vorsitzender des Prüfungsausschusses:	Dr. H.D. Rüter
Vorsitzender des Promotionsausschusses:	Prof. Dr. R. Wiesendanger
Dekan des Fachbereichs Physik:	Prof. Dr. G. Huber

Wer gar zu viel bedenkt, wird wenig leisten.

Friedrich von Schiller

Abstract

The present work consists of two parts: The first part is a contribution to instrumentation and more particularly the development of a readout system for gas detectors used in small angle X-ray scattering (SAXS) and the second part covers the results of an investigation of interactions between proteins in solution.

One aim of the instrumentation part was to investigate the role of fluorescence in gas proportional detector with delay line readout, a phenomenon that leads to events, unrelated to the actual scattering pattern. For this purpose a time stamp (TS-) TDC readout was developed that is able to detect simultaneous events in the detector by using all timing information in the anode and cathode signals (check-sum).
It was found that the fluorescence of the gas clearly limits the spatial and time resolution in gas detection. The present observations confirm the influence of argon fluorescence on the point-spread function (PSF) and demonstrate that a rejection mechanism that analyses single events only cannot entirely eliminate these effects. The signal-to-noise ratio can, however, be improved by taking the drift time of the electrons and the check-sum of the signals into account. Drift time measurements are made possible by the unique ability of the TS-TDC to correlate the observation of scattered photons with individual radiation bunches in the storage ring. This feature could also facilitate time-resolved measurements with nanosecond-resolution.

In the second part of this thesis gas detectors were used in a contribution to an important topic in biophysics: the study of intermolecular interactions in protein solutions. This was done by measuring the structure factors of protein solutions.
According to the DLVO theory the main interactions between spherical particles are the hard-sphere interactions, a short range attraction, due to surface-surface forces, and a long range repulsion caused by the fact that the particles are charged.
A computer program was written, which calculates the structure factors from the pair potentials and was shown to reproduce some results in the literature as well as all the main observations in our experiments.
Several measurements were made on solutions of the proteins glucose oxidase and lysozyme under various conditions in order to investigate the interactions of these proteins in the presence of co-solutes such as salts, urea, TMAO, glycine or at different protein concentrations. The influence of variables like temperature and pH on the attractive and repulsive interactions and in particular the pair potentials was also studied.
The main features of the structure factors observed in the protein concentration series are correctly predicted with the program. The most interesting finding was that salt concentration series of structure factors and the change of the structure factor upon addition of urea and TMAO could be better described if the strength of the attractive potential decreases with increasing salt concentration. Previous work in the literature had relied on a constant attractive potential within a given series of measurements.

Kurzfassung

Die vorliegende Arbeit besteht aus zwei Teilen: Der erste ist ein Beitrag zur Instrumentierung und insbesondere zur Entwicklung eines Auslesesystems für Gasdetektoren, die in Kleinwinkelstreuexperimenten verwendet werden. Der zweite Teil beschäftigt sich mit der Untersuchung von Proteinwechselwirkungen in Lösung.

Ein Ziel des Instrumentierungsteils war es, die Auswirkungen von Fluoreszenz in Gasdetektoren mit Verzögerungsleitungsauslese zu untersuchen. Ein Phänomen, das zu detektierten Ereignissen führt, die nicht Teil des eigentlichen Streumusters sind. Zu diesem Zweck wurde ein Zeitmarken-TDC (TS-TDC) entwickelt, welcher in der Lage ist gleichzeitig stattfindende Ereignisse nachzuweisen, indem er von sämtlichen Zeitinformationen der Anoden- und Kathodensignale Gebrauch macht (Checksumme).
Es wurde festgestellt, dass die Fluoreszenz des Gases das räumliche und zeitliche Auflösungsvermögen des Detektors beschränkt. Die gemachten Beobachtungen bestätigen den vermuteten Einfluss der Argon-Fluoreszenz auf die Point-spread-function (PSF) und legen dar, dass selbst ein Selektionsmechanismus, der nur Einzelereignisse aufnimmt, diesen Effekt nicht völlig aufheben kann. Das Signal-Rausch-Verhältnis kann jedoch verbessert werden, indem die Driftzeit und die Checksumme der Signale gemessen und zum Filtern verwendet werden. Die Messung der Driftzeit wird durch die Fähigkeit des TS-TDCs einzelne gestreute Photonen ihrem Ursprungs-Photonenbunch des Speicherrings zuzuordnen ermöglicht. Diese Eigenschaft kann auch für zeitaufgelöste Messungen im Nanosekunden-Bereich ausgenutzt werden.

Im zweiten Teil dieser Arbeit wurden Gasdetektoren dazu verwendet einen Beitrag zu einem wichtigen Bereich der Biophysik zu leisten: die Untersuchung von intermolekularen Wechselwirkungen in Proteinlösungen. Dieses wurde durch Messungen des Strukturfaktors der Proteinlösungen erreicht.
Nach der DLVO-Theorie sind das Harte-Kugel-Potential, eine kurzreichweitige Anziehung, die durch Kräfte zwischen den Oberflächen erzeugt wird, und eine langreichweitige Abstoßung zwischen den geladenen Molekülen die wichtigsten Wechselwirkungen zwischen kugelförmigen Teilchen.
Ein Computerprogramm, das die Strukturfaktoren aus den Paarpotentialen ausrechnen kann, wurde geschrieben und war in der Lage sowohl die bekannten Ergebnisse aus der Literatur als auch die wichtigsten Beobachtungen unsere Experimente zu reproduzieren.
Zahlreiche Messungen von Lösungen der Proteine Glucose-Oxidase und Lysozym wurden unter verschiedenen Bedingungen durchgeführt, um die Wechselwirkungen der Proteine unter dem Einfluss von Substanzen wie Salzen, TMAO, Glycin oder bei unterschiedlichen Proteinkonzentrationen zu untersuchen. Der Einfluss von Veränderlichen wie Temperatur und pH-Wert auf anziehende und abstoßende Wechselwirkungen und insbesondere die Paarpotentiale wurde untersucht. Die wichtigsten Merkmale der Strukturfaktoren, die in den bei verschiedenen Proteinkonzentrationen gemessen wurden, konnten von dem Programm korrekt reproduziert werden. Das interessanteste Ergebnis war, dass die Messungen mit unterschiedlichen Salzkonzentrationen und die Messungen unter dem Einfluss von Harnstoff und TMAO mit einem attraktiven Potential, das mit steigender Salzkonzentration abnahm, am besten erklärt werden konnte. Bei ähnlichen Messungen, die in der Literatur dokumentiert sind, wurde immer mit einem konstanten anziehendem Potential gearbeitet.

Content

1 Introduction

X-rays are an important tool to study systems on the molecular and atomic level. Structural biology, with techniques such as protein crystallography, has contributed a precise description of the components of many biological systems. Synchrotron radiation plays a role in the investigation of biological systems because the high brilliance of the sources improves the conditions, the accuracy and the speed of crystallographic measurements. The brilliant sources opend up new possibilities for experiments, such as those on muscle contraction, which provided the initial drive for the use of X-ray synchrotron radiation [Holmes, 1998]. High-brilliance tunable radiation facilitated the study of biological macromolecules in solution (SAXS) and the spectroscopic properties (EXAFS) of the many metalloproteins. This is another example for the fact that all progress in science is connected to the progress in instrumentation or as the Spanish physician and Nobel prize laureate Santiago Ramón y Cajal wrote [Ramón y Cajal, 1999]: "Mastery of technique is so important that without fear of contradiction it may be stated that great discoveries are in the hands of the finest and most knowledgeable experts on one or more of the analytical methods." As a modest start in this direction the first part of this scientific work, which was done at the X33 beamline of the EMBL in HASYLAB, is a contribution to instrumentation and more particularly the development of a readout system for gas detectors and the second part covers the results of an investigation of interactions between proteins in solution.

Many X-ray diffraction or scattering problems require the quantitative evaluation of a linear or 2D-pattern. There are several types of position sensitive detectors (PSD) and for many applications gas filled counters are a good choice. For example in solution scattering experiments it is important to measure intensities very accurately. In addition the detector must be able to make short time frames and to handle count rates of up to several thousand counts per square millimetre. Spatial resolution and energy discrimination are of less importance here.
Gas proportional counters can count single photons. They have a large dynamic range (4-5 orders of magnitude: 0 – several ten thousand cps/mm^2) and low noise. By connecting a position sensitive single photon counting device to a histogramming memory one can make a system which is capable of reading time-resolved X-ray scattering measurements with sub-millisecond resolution.
A common way to read out position sensitive proportional counters is a delay line electrode, where each event creates two signals travelling in opposite directions, connected to a time-to-digital converter (TDC) [Gabriel, 1977]. The signal at one end starts a time measurement, the signal at the other end stops it. This start-stop method has some limitations at high count rates. When the pulses of more than one event travel at the same time on the delay line they may lead to false counts in the histogram. To circumvent this problem one can establish a rejection mechanism with the disadvantage that some counts will be lost.
One aim of this work was to investigate the role of fluorescence in gas proportional detectors, a phenomenon, which leads to events, which are not related to the actual scattering pattern. A time stamp TDC readout was developed that is able to detect simultaneous events in the detector. The system was used to study the time response of a detector to a synchrotron radiation bunch and it was shown that a time stamp TDC can facilitate time-resolved measurements with nanosecond-resolution. The construction of a

high count rate time stamp TDC would have been outside the scope of this thesis, especially as such developments are taking place elsewhere [Gebauer, 2001].

In the second part of this thesis gas detectors were used in a contribution to an important topic in biophysics: the study of intermolecular interactions in protein solutions. Many properties of biochemical systems depend on the interactions between biological macromolecules. Processes in which interactions between proteins play a role are, for instance, assembly and non-specific aggregation and more practically solubility and crystallisation. Protein interactions occur in various ways and are thus difficult to predict. At this stage, the physical interactions of proteins in aqueous solution can at best only be explained semi-quantitatively using simplified semi-empirical models [Tardieu, 1999].
The X-ray scattering pattern measured in small angle scattering (SAXS) is a product of form factor and structure factor. Whereas the form factor provides information about the shape of the solute scattering the radiation, the structure factor contains information about the structure of the solution (i.e. the distribution of the solute particles). Attractive interactions, for instance, are responsible for an increase of the structure factor at low angles, repulsive interactions lead to the opposite behaviour. SAXS thus provides a way of measuring the structure factor of a protein solution that is related to the pair distribution function, which can be calculated from the pair potentials.
In the 1940s a theory of interactions, which explains some of the phenomena occurring in such systems, was developed independently by B.V. Derjaguin & L.D. Landau and E.J.H. Verwey & J.TH.G Overbeek (DLVO). This theory, today referred to as DLVO theory, after the names of its authors, explains many aspects of the interactions and stability of colloids. According to the theory the main interactions between spherical particles are the hard-sphere interactions, a short-range attraction due to surface-surface forces, and a long range repulsion caused by the fact that the particles are charged [Verwey, 1948]. Even though proteins are polypeptide chains in physiological conditions at least the globular proteins behave as rigid molecules and many of their interactions are similar to those described in the DLVO theory. They carry a charge, which depends on the pH and phenomena such as crystallisation and aggregation prove that attractive interactions also play a role.
Several measurements of the proteins glucose oxidase and lysozyme were made under various conditions in order to investigate the interactions of these proteins in the presence of co-solutes such as salts, urea, TMAO, glycine or at different temperatures and protein concentrations. The choice of these co-solutes was made on the basis of their practical use in crystallisation or their physiological importance (urea, TMAO, glycine, salts). A computer program was written, which calculates the structure factors from the pair potentials and was shown to reproduce some results in the literature as well as all the main observations in our experiments.

This thesis is divided into eight chapters, which contain the following:
Chapter 2, **Scattering**: Introduction to scattering experiments and description of the X33 beamline at HASYLAB.
Chapter 3, **Gas Detectors**: Introduction to position-sensitive gas proportional detectors.
Chapter 4, **Bunch Structure and point-spread function**: Description of the time structure of the DORIS storage ring. Description of the point-spread function as a method to characterise the detector operation.
Chapter 5, **Interactions between biological macromolecules**: Introduction to the basics of protein interactions.
Chapter 6, **Proteins, sample preparation and theoretical background for the SAXS measurements**: Description of the proteins used for the studies and the theory of SAXS measurements and of the relation between pair potentials and structure factor on which the studies are based.
Chapter 7, **Results and Discussions**: Presentation of the results of the structure factor measurements and their discussion.
Chapter 8, **Summary and Outlook**: Summary of the results and outlook on possible future experiments.

2 Scattering

During the last centuries science has made considerable progress in investigating the microscopic building blocks of matter and their interactions leading to an understanding of diverse macroscopic physical, chemical and biochemical processes. The first step was the invention of the microscope that made it possible to observe processes not visible to the naked eye. One of the earliest microscopes was built by the Dutch clockmaker Antoni van Leeuwenhoek (1632-1723) at the end of the 17th century. In a microscope lenses create a magnified virtual image of the object. The resolution of a microscope based on refractive optics is limited by the wavelength of light. The discovery of X-rays by Wilhelm Conrad Röntgen in 1895 [Röntgen, 1895] not only opened a new wavelength range and consequently made even smaller dimensions accessible but more immediately also gave a new insight into the human body by X-ray transmission radiography (Figure 2.1). For a long time only diffractive lenses (e.g. zone plates) were used [Ebashi, 1991], today the first refractive X-ray lenses are available [Snigirev, 1996], which suffice to build an X-ray microscope that can take images with a resolution comparable to the images obtained with visible light.

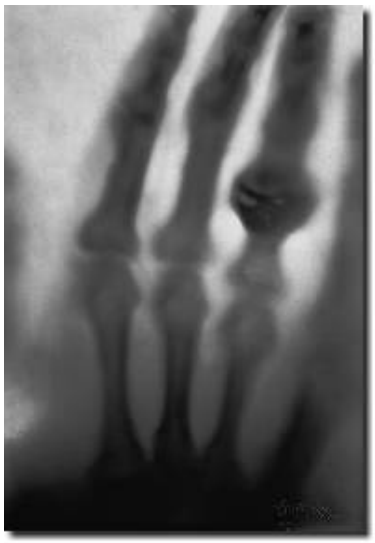


Figure 2.1: Picture of Mrs. Röntgen's hand, taken on 22nd December 1885.

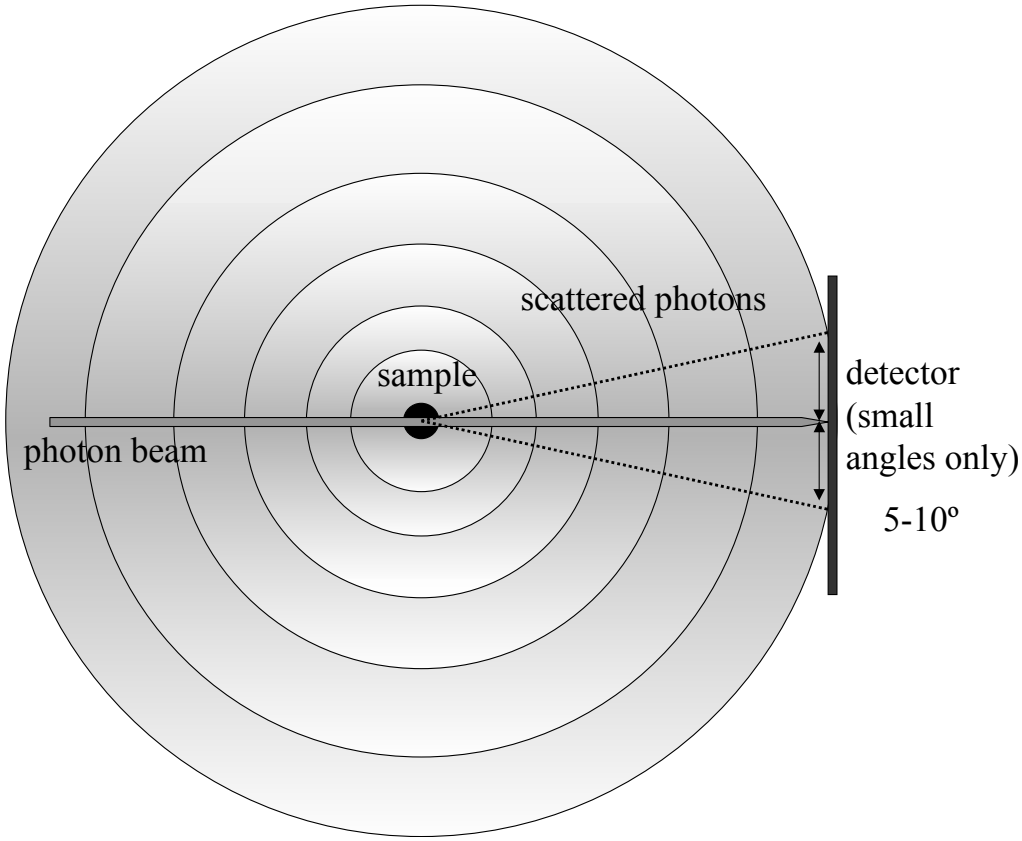


Figure 2.2: Set-up of a simple scattering experiment.

Nevertheless no real-space images with atomic resolution can be taken. The objects have to be investigated by scattering experiments (Figure 2.2). The first kind of scattering process to be recognised was discovered by Max von Laue, who was awarded the Nobel Prize for physics in 1914 for his discovery of the diffraction of X-rays by copper sulphate crystals. The diffraction condition for a lattice is described by Bragg's law:

$$n\lambda = 2d\sin\theta, \tag{2.1}$$

where n is the order of the reflection, λ the wavelength of the X-rays, d the spacing between the planes and 2θ the angle between the incident and diffracted beam. Henry and William Lawrence Bragg received the Nobel Prize in 1915. Their theory is based on the wave nature of electromagnetic radiation that was discovered in diffraction experiments with visible light. The diffraction pattern is the Fourier transform of the crystal. Its analysis makes structural information down to Ångstrom resolution accessible. The diffraction pattern of a crystal is a collection of spots. Their positions provide information about the symmetry of the lattice and the dimensions of the unit cell and their intensities are related to the content of the unit cell.

Things are less simple for the scattering of disordered systems like macromolecules in solution, which only scatter at low angles [Koch, 2003]. Here high brilliance photon sources and accurate detectors are needed and thus only with the availability of synchrotron sources and gas detectors did this technique became routinely usable for biological macromolecules during the second half of last century. The analysis is based on the theoretical findings of Peter Debye and André Guinier. Debye's formula describes the scattered intensity $I(s)$ of disordered system of molecules with N spherical atoms [Debye, 1915]:

$$I(s) = \sum_{i=1}^{N}\sum_{j=1}^{N} f_i(s) f_j(s) \frac{\sin(sr_{ij})}{sr_{ij}}, \tag{2.2}$$

where $f_i(s), f_j(s)$ are the scattering factors of the atoms and s is the scattering vector. Guinier (1939) found that in a monodisperse solution the spatial extent of the scattered intensity is directly connected to the size of the molecule and is given by the Guinier approximation:

$$I(s) \approx I(0)\exp\left(-\frac{1}{3}R_g{}^2 s^2\right). \tag{2.3}$$

R_g is the radius of gyration.

The present chapter provides an overview about the basic theory, which underlies scattering. Subsequently the current techniques used in order to perform SAXS experiments are introduced using where the X33 beam line of the EMBL at DORIS as an example.

2.1 Interactions of electromagnetic radiation with matter

Electromagnetic radiation can be used to investigate matter in several ways. For instance, radio waves that have a wavelength of approximately 1 m are used for nuclear magnetic resonance spectroscopy (NMR) and/or imaging. Microwaves (wavelength in the mm/cm-range) can be used for spectroscopy of rotation states of molecules but also for radar. Tera-hertz radiation (sub-mm) is not used for many applications yet because of the lack of efficient sources. Vibrations of molecules are investigated by infrared spectroscopy (μm-range). Visible light (400-750 nm) is used in conventional microscopy and it can excite electrons in the outer shell of atoms and molecules. UV light absorption for instance can be used to determine the protein concentrations due to the specific absorption of the peptide bound of the aromatic amino acids tyrosine and tryptophan at 280 nm. For shorter

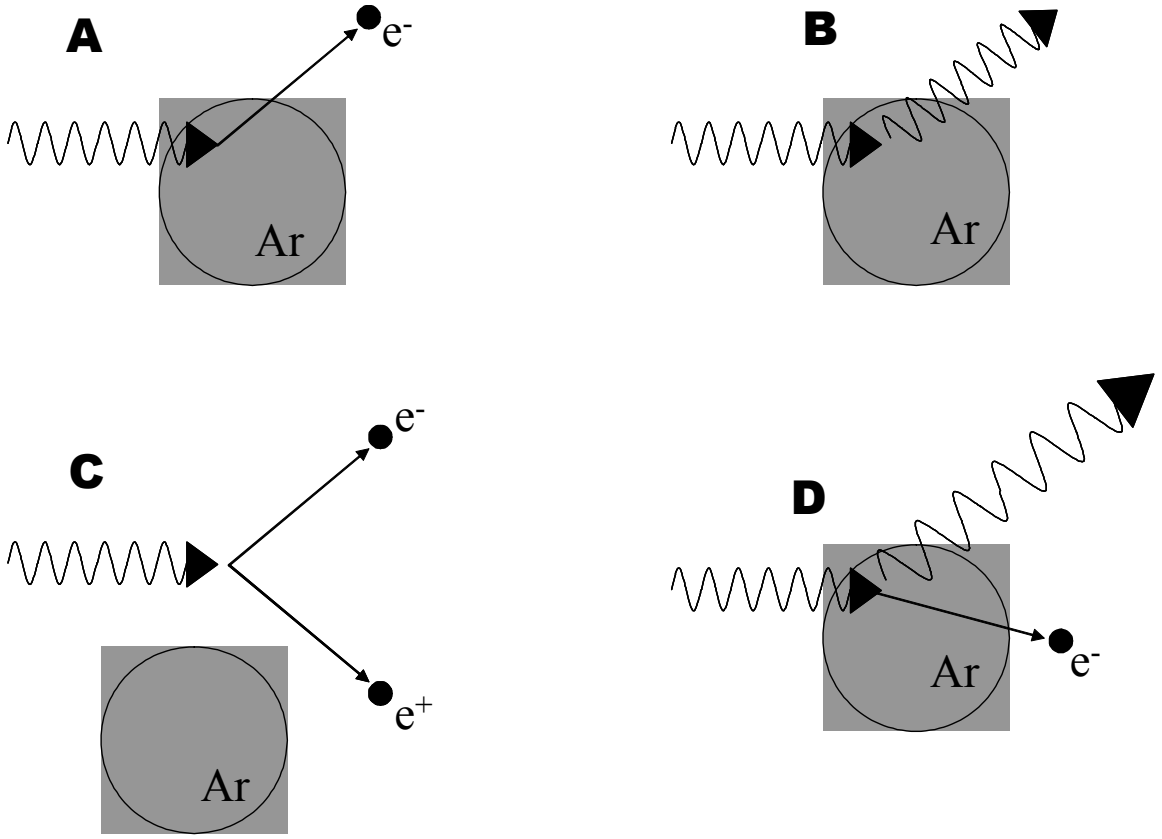


Figure 2.3: Interactions between X-rays and matter. (A) Photoelectric effect. (B) Elastic scattering. (C) Pair production. (D) Inelastic (Compton) scattering.

wavelength like X-rays the light interacts mainly with the electrons of the inner atomic shells.
There are basically four different ways by which X-rays and gamma radiation can interact with matter (Figure 2.3):

- photoelectric effect
- elastic scattering
- inelastic (Compton) scattering
- pair production

Pair production and the photoelectric effect are absorption processes. Pair production can only occur for photon energies above 1.02 MeV (the rest energy of the generated electron-positron pair) and thus it is of no interest for research with synchrotron radiation. For the photoelectric effect an atom absorbs the energy of the incident photon and is ionised. In contrast during a scattering process the incident photon interacts with an electron and changes direction. In the elastic case the electron is excited by the photon but immediately emits a photon of same energy. In the inelastic case (Compton scattering) part of the photon energy is transferred to the electron and consequently Compton scattering is always incoherent [Lengeler, 2001]. This results in a lower energy (and frequency) of the scattered photon.
The choice of wavelength is crucial in the numerous applications of electromagnetic radiation to the study of the structure of matter.
Visible light for instance is not able to penetrate even thin layers of most materials but it is appropriate to do scattering experiments of transparent aqueous solutions like in dynamic light scattering. When the wavelength of the radiation is very short (X-ray range) its

capabilities to probe the material improve but it is still necessary to find a balance between absorption and scattering.
Scattering and absorption can be understood using the classical wave theory of light. The basic idea is that an incoming electromagnetic wave (Eq. 2.4) excites a bound electron. The equation of motion of the electron (Eq. 2.5) is that of a damped oscillator driven by the incident wave where k is the damping factor and ω_0 the natural circular frequency of the electron (mass m).

$$\vec{E}(t) = \vec{E}_0 e^{i\omega t}, \tag{2.4}$$

$$\ddot{x} + k\dot{x} + \omega_0 = \frac{eE_0}{m} \cdot e^{i\omega t}, \tag{2.5}$$

The resulting displacement of the electron is described by

$$\vec{x} = \frac{e\vec{E}_0}{m} \cdot \frac{e^{i\omega t}}{\omega_o{}^2 - \omega^2 + ik\omega}, \tag{2.6}$$

The oscillating dipole itself emits a wave with the same frequency.
This semi-classical model, where the electron is represented by a driven oscillator with an electromagnetic wave as source of excitation, can properly describe the scattering process of a photon by an electron. The oscillation electron emits an electromagnetic wave with the same frequency. This dipolar radiation, can be represented by its electric field, where Ψ is the angle between the direction of polarisation and the observer's line of sight:

$$E(\vec{r},t) = \frac{e^2}{mc^2} \frac{\sin\Psi}{r} \frac{\omega^2}{\omega_0{}^2 - \omega^2} E_0 e^{i\omega t} \tag{2.7}$$

For visible light ($\lambda \approx 500\text{nm}$) ω is much smaller than ω_0.

$$\Rightarrow \frac{\omega^2}{\omega_0{}^2 - \omega^2} \approx \frac{\omega^2}{\omega_0{}^2} \tag{2.8}$$

The amplitude of the scattered light is proportional to ω^2. Blue light is more strongly scattered than red light. This is Rayleigh scattering, which explains why the sky appears blue. In the case of X-rays ($\lambda \approx 0.1\text{nm}$) ω is much larger than ω_0.

$$\Rightarrow \frac{\omega^2}{\omega_0{}^2 - \omega^2} \approx -1 \tag{2.9}$$

The scattered amplitude is in this case independent of the frequency of the incident radiation and the equation for $E(\vec{r},t)$ simplifies to

$$E(\vec{r},t) = -\frac{e^2}{mc^2} \frac{\sin\Psi}{r} E_0 e^{i\omega t} \tag{2.10}$$

This regime of scattering is called Thomson scattering.
Since for small scattering angles Ψ is close to $\pi/2$ the SAXS pattern does not depend on the polarisation of the incident radiation.
In general the amplitude of the wave at unit distance in the equatorial plane is

$$A = \frac{e^2}{mc^2} \cdot \frac{\omega^2 E_0}{\omega_0{}^2 - \omega^2 + ik\omega} \tag{2.11}$$

The scattering factor is defined as the ratio of the scattered amplitude A to that scattered by a free electron A_e in the same conditions. This is obtained by solving the same equation for $\omega_0 = k = 0$.

$$A_e = -\frac{e^2}{mc^2} \cdot E_0 \tag{2.12}$$

The dipole scattering factor is given by

$$f(\omega) = \frac{A}{A_e} = \frac{\omega^2}{\omega^2 - {\omega_0}^2 - ik\omega} = \mathrm{Re}\big(f(\omega)\big) + i\,\mathrm{Im}\big(f(\omega)\big) \tag{2.13}$$

It is composed of the real part $f_0 + f'(\omega)$ and the imaginary term $f''(\omega)$. If a medium is composed of N similar dipoles per unit volume, it can be shown that the refractive index n is also complex, which indicates that the medium is absorbs radiation.

$$n = 1 - \frac{2\pi N e^2}{m\omega^2} \cdot f \tag{2.14}$$

The linear absorption coefficient can also be expressed as a function of the scattering factor f.

$$\mu(\omega) = \frac{4\pi N e^2}{mc\omega} \cdot \mathrm{Im}\big(f(\omega)\big) \tag{2.15}$$

Taking into account all electrons of an atom the atomic scattering factor can be calculated, where f_0 is the atomic scattering factor remote from the resonant energy levels, f' is the real part of the anomalous-scattering factor, and f'' is the imaginary part of the anomalous-scattering factor.

$$f = f_0 + f' + if'' \tag{2.16}$$

X-rays thus interact by scattering and absorption with matter. In order to set up a small angle scattering experiment one has to take both phenomena into account when choosing the sample thickness d. The scattered intensity I_{sc} is proportional to the sample thickness but it is attenuated by absorption.

$$I_{sc}(d) \propto d \cdot \exp(-\mu d) \tag{2.17}$$

The optimal thickness can be determined by calculating the maximum of the scattered intensity.

$$I'_{sc}(d) \propto \exp(-\mu d) - \mu d \cdot \exp(-\mu d) \equiv 0 \tag{2.18}$$

$$\Rightarrow (1 - \mu d)\exp(-\mu d) = 0$$

$$\Rightarrow 1 - \mu d = 0$$

$$\Rightarrow d_{\mathrm{opt}} = \frac{1}{\mu} \tag{2.19}$$

Since $\mu = 10.1\frac{1}{\mathrm{cm}}$ for water and 8 keV X-rays the optimal thickness for solution scattering in aqueous solution is $d_{\mathrm{opt}} = 1\,\mathrm{mm}$ (Table 2.1).

2.2 Sources of X-ray photons

In the first X-ray experiments the radiation was produced in an evacuated tube with an

Material	**Density ρ [g/cm^3]**	**E *[keV]***	**μ [1/cm]**	**d_{opt}=1/μ [mm]**
Water	1	8	10.1	0.99
KCl solution 1M	1.05	8	18.9	0.53
Lead	11.3	8	2500	0.004
Lead	11.3	30	323	0.031

Table 2.1: Optimal sample thickness d for different substances.

anode and a cathode with several thousand volts of potential difference. A hot cathode emits electrons, which are accelerated by the field. At the time of their arrival at the anode they are rapidly decelerated and thereupon they emit electromagnetic radiation. The photons produced in this process can be separated into two categories, characteristic emission and bremsstrahlung. The characteristic emission results from fluorescence of the atoms of the anode material after an inner shell ionisation. Bremsstahlung consists of photons, which are directly produced by the electrons when they are decelerated by their electromagnetic interactions with the atoms of the anode material.
The maximal photon flux that is achievable by such a device is limited by the lifetime of the anode material. Due to the large amount of energy that is deposited at the anode it heats up quickly and must thus be cooled. In modern devices the anode rotates to decrease the energy density and facilitate cooling. The brilliance is also limited because the radiation is emitted in all directions and efficiency of X-ray optics is limited.
Nowadays electron or positron storage rings are the most X-ray sources which have the highest brilliance (photons/s/mm^2/mrad^2/0.1%bandpass). For historical reasons the X-rays emitted by these devices are called *synchrotron radiation* and storage rings are often incorrectly referred to as a synchrotron. A synchrotron is an accelerator for charged particles, in which the particles follow always the same trajectory, independently of their energy. This is facilitated by increasing the field of the bending magnets proportionally to the particle energy. The following relation between electron energy *E,* the magnetic field *B* and the magnet curvature radius *r* is valid for ultra-relativistic electrons:

$$\frac{1}{r}[\mathrm{m}^{-1}] = \frac{0.2998B[\mathrm{T}]}{E[\mathrm{GeV}]} \tag{2.20}$$

In storage rings the average energy of the particles remains constant. Only a small accelerating section in the RF cavities compensates the loss of energy due to the emission of photons in the magnetic structures.
The story of the discovery of synchrotron radiation is quite interesting. The fact that a magnetic field deflects the trajectory of a moving charge was well-known already in the nineteenth century. Also it was soon realised that an accelerated charge in such a scenario would emit electromagnetic radiation. Already in 1898, when the electron was not yet discovered, Alfred Liénard published a formula for the rate of loss of energy by an "electric mass" travelling on a circular path [Liénard, 1898]. After the discovery of the electron the problem became more attractive and was further investigated. George A. Schott derived an expression for the frequency distribution but his equation involves Bessel functions of very high order that were intractable for physicists at that time [Schott, 1912]. When in the 1930s and 1940s the first betatrons were in operation most scientists working with them were not aware of the early works anymore. In Russia the theory was rediscovered by Pomeranchuk and Ivanenko [Ivanenko, 1944]. They calculated an equation for the energy loss of electrons in a betatron. A group working at a betatron for General Electrics was measuring this radiation loss in form of an orbit shrinkage. They also tried to measure the emitted radiation but the spectrum was still unknown and they measured in the MHz range whereas their 100 MeV betatron was emitting mainly in the visible range.
Finally on April 24th 1947 the GE group first observed the electromagnetic radiation emitted by electrons travelling on a circular path at relativistic speed. This was found at the world's second operating synchrotron [Blewett, 1998]. This device was working at 70 MeV and its vacuum chamber was transparent. Thus during its first operation a bluish-white spot appeared at the side of the chamber where the beam was approaching the observer.
During the 1950s and 60s the nature of synchrotron radiation was investigated and it turned out that it could be well described by Schwinger's equations [Schwinger, 1949]. In 1961 a

serious research program started at the National Bureau of Standards that made use of synchrotron radiation as a probe for the investigation of other substances or processes. In the following years scientist went to machines with higher electron energies. These were able to produce radiation with even shorter wavelengths. At this time the most advanced electron accelerators were used in high-energy physics. Some projects were started in which the unavoidably produced synchrotron radiation was used for other experiments. One example is the use of 6 GeV [Holmes, 1998] electrons at DESY which led to wavelengths down to 0.1 Å. Nowadays most synchrotron sources are dedicated to synchrotron radiation experiments.

In the following section some details about synchrotron radiation generating devices are summarised. There are three kinds of devices for this purpose in a storage ring: Bending magnets, wigglers and undulators.
The bending magnet was historically the first device used for the generation of synchrotron radiation. Bending magnets are an essential component of any circular accelerator or storage ring because they are needed to force the electrons/positrons on a circular track. If charged particles are guided through a curved track they have to be accelerated perpendicularly to the curve. The bending power of the magnet and the ratio of mass and velocity of the charge determine the spectrum and angular distribution of the radiation.
A quantitative description of the synchrotron radiation spectrum is given by Schwinger's formula (taken from [Ebashi, 1991]):

$$I(\lambda,\Psi)=\frac{27}{32\pi^3}\frac{e^2c}{R^3}\left(\frac{\lambda_c}{\lambda}\right)^4\gamma^8\left[1+(\gamma\Psi)^2\right]^2\left[K_{2/3}{}^2(\xi)+\frac{(\gamma\Psi)^2}{1+(\gamma\Psi)^2}K_{1/3}{}^2(\xi)\right], \tag{2.21}$$

where $K(\xi)$ are modified Bessel functions of the second kind, Ψ is the vertical angle between the orbital plane and the direction of the beam, $\gamma=\frac{E}{E_0}$ is the energy of the electrons in units of its rest energy, and $\lambda_c=\frac{4}{3}\frac{R}{\gamma^3}$ is the critical wavelength, below which the intensity decreases exponentially. The term for the spectral brightness on-axis is somewhat simpler:

$$I(\lambda,\Psi=0)=1.325\times10^{10}\left(\frac{\lambda_c}{\lambda}\right)^2K_{2/3}\left(\frac{\lambda_c}{2\lambda}\right) \tag{2.22}$$

2.2.1 Wigglers and undulators

Wigglers and undulators are periodic magnetic structures placed in straight sections of the storage ring that force the electrons to move on a sinusoidal track. One can imagine such a device as a chain of bending magnets. In a wiggler the intensities of the radiation, which is emitted from curved sections, add up. An important parameter characterising such devices is the wiggler parameter $K=0.934\lambda_0[\mathrm{cm}]B_0[\mathrm{T}]$, where λ_0 is the length of a period and B_0 is the peak value of the magnetic field. The maximum angular deflection of the electron beam in a wiggler/undulator is $\delta=\frac{K}{\gamma}$. For $K<1$ radiation from the individual deflections interferes strongly and the device is called an undulator [Ebashi, 1991]. If $K>>1$ the device is called wiggler and its spectral distribution is similar to a bending magnet whereas the spectrum of an undulator consists of several peaks with wavelength

$$\lambda_i=\frac{\lambda_0}{2i\gamma^2}\left(1+\frac{1}{2}K^2+\gamma^2\theta^2\right),\qquad i=1,2,3,\ldots. \tag{2.23}$$

2.3 A small angle scattering beam line: X33 at DORIS

The basic requirement for a small angle scattering experiment is a collimated beam impinging on the sample in transmission mode. A very small fraction of the incident X-ray photons are scattered by the sample. The intensity distribution of the coherent scattering as a function of the angle of observation contains structural information. For a disordered system (e.g. a solution of macromolecules) the small angle part contains most of the information. A photon detector is needed to record the scattering pattern. In this section the X33 beam line of the EMBL (Figure 2.4) at DORIS (HASYLAB/DESY) will serve as an example of an experimental set-up for SAXS.
This beam line, which is dedicated to small angle scattering from non-crystalline systems, was build in 1982 [Koch & Bordas, 1983] and its optics has not changed since. The X-ray optics consists of a monochromator and a X-ray mirror, which are located in the optics hutch. In the experimental hutch there is the ionisation chamber, the sample holder, the camera (a vacuum tube with the beam stop at its end) and the detector(s). The beam transmitted through the sample can be measured with a photodiode.
The compressing monochromator is a Si 111 crystal. According to Bragg's law under a certain angle only one particular wavelength (and its higher harmonics) of the synchrotron radiation spectrum is diffracted. The monochromator can also be bent to focus the beam horizontally (Figure 2.4).
Because of the large divergence of the beam a very long quartz mirror (2m) is needed. It is based on total reflection. It is divided into eight segments arranged in a bicycle-chain type set-up (Figure 2.4). The mirror is used to focus vertically and to eliminate most of the higher harmonics (the second harmonic corresponds to a forbidden reflection of the monochromator).
Alignment of the beam consists in obtaining as small as possible a focal spot, by overlapping the reflections from the mirror segments and bending the monochromator. The different pairs of slits have to be carefully adjusted to minimise the background in the camera.

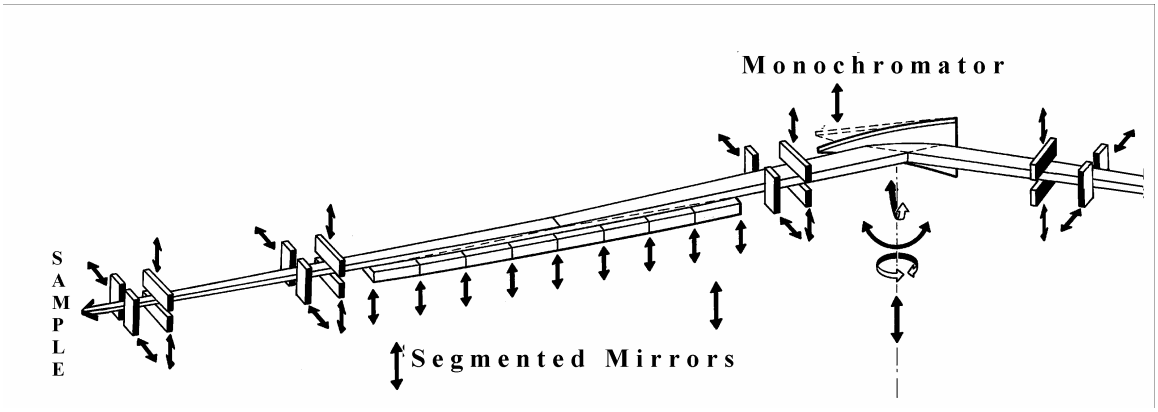


Figure 2.4: X33 beam line of the EMBL at DORIS (HASYLAB/DESY).

The purpose of the ionisation chamber is to measure the intensity of the incident monochromatic beam. Since the positrons in the DORIS storage ring have a finite lifetime the intensity decays during a measurement and the results have to be scaled in order to make measurements comparable (Figure 2.5). The ion chamber is also helpful for the alignment of monochromator and mirror and monitors the performance of the beam optics and the positional stability of the beam. The ionisation chamber is just a high voltage capacitor with air as a dielectric. The X-ray beam passes between the electrodes and

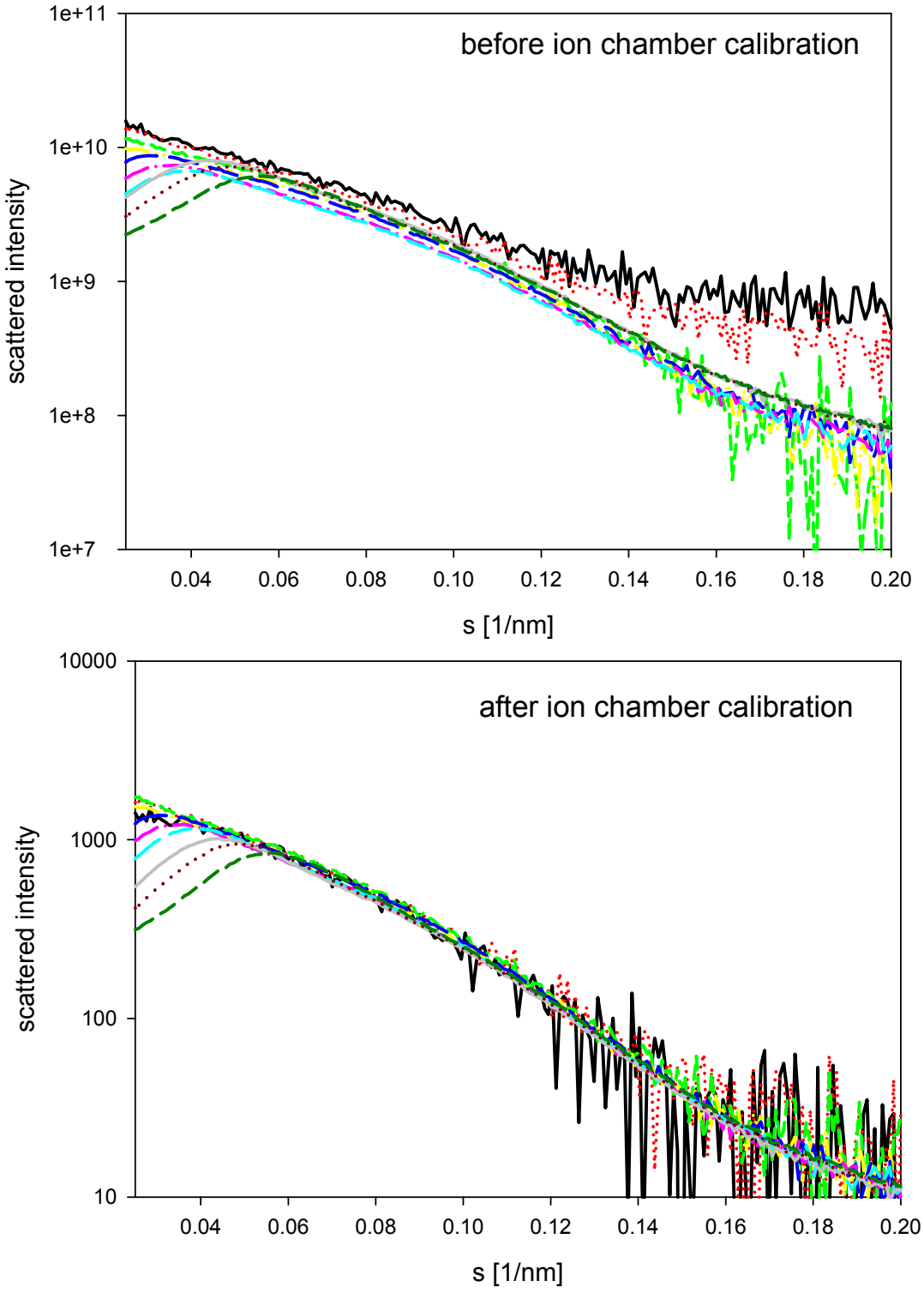


Figure 2.5: Scattering curves of glucose oxidase in water at different concentrations – with and without ion chamber calibration. The ionisation chamber normalises with the incident beam.

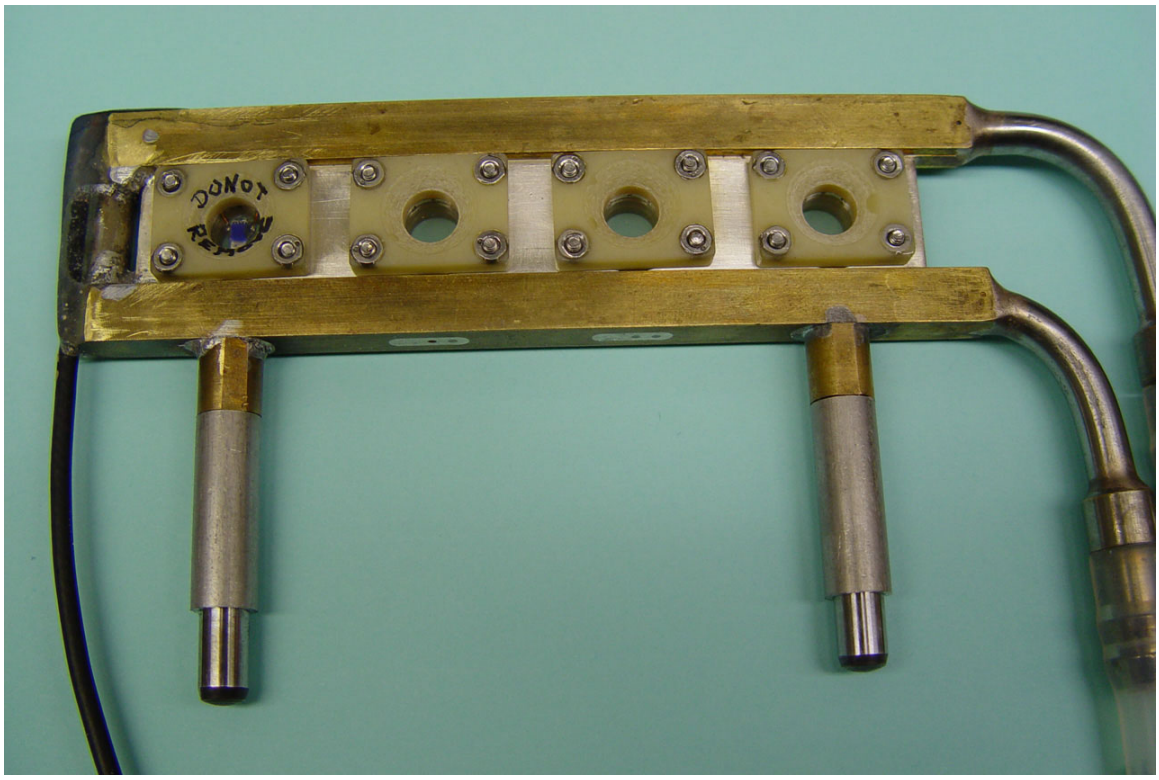


Figure 2.6: Image of a sample holder with three sample cells. The leftmost cell contains a temperature sensor. The tubing on the right can be connected to a water bath.

ionises part of the air molecules which leads to a current that is proportional to the intensity of the beam. A current amplifier (Keithley 427) delivers a voltage proportional to this current, which is recorded together with the detector data of the experiment.
Immediately after the ionisation chamber the beam passes through the sample cell (Figure 2.6). The sample cell consists of two mica windows with a cavity in between. The frame is made of silver and can be water cooled. The thickness $d = \frac{1}{\mu}$ of the cavity between the windows depends on the system, which should be investigated, as discussed in the previous section.
For SAXS experiments on proteins in aqueous solution the optimal thickness is approx. 1 mm. Behind the sample holder the beam (including the photons scattered at small angles) enters the camera that is a large evacuated metal tube with a (small) mica entrance window and a (large) mylar exit window. The presence of the tube is necessary because the used wavelength of 1.5 A is strongly absorbed and scattered by air (10% for 10 cm). The entrance window is of the same size as the windows of the sample holder whereas the diameter of the exit window (200 mm) fits to that of the detector. Another important component of the camera is the beam stop. If the sample has the optimal thickness the absorption of the sample is $1/e$ and the transmitted direct beam is much stronger than the scattered signal. Exposing the detector to the direct beam would seriously disturb the measurement and possibly damage the detector. The beam stop is made of a material (lead) with a short absorption length for X-rays. It absorbs the direct beam and the strong scattering at very low angles (due e.g. to the slits, optical elements etc.) and thus sets a lower limit for the minimal observable scattering angle. The length of the camera (i.e. the distance between sample and detector) is one of the most important parameters of a SAXS experiment as it defines the order to order resolution as well as the s-range covered. Placing a second detector at wider angles can enlarge the s-range. The minimal scattering

angle is limited by the parasitic scattering from the beam optics and it is defined by the size of the last aperture (slit). The beamstop has to be large enough to prevent the parasitic scattering from reaching the detector.
If l is the camera length, b the radius of the beam stop and a the diameter of the active area of the detector the minimal (maximal) scattering angle that can be measured (Figure 2.7) is

$$\sin(\theta_{\min}) = \theta_{\min} \approx \frac{b}{l}; \quad \sin(\theta_{\max}) = \theta_{\max} \approx \frac{a}{l}. \tag{2.24}$$

An X-ray-sensitive photodiode is built into the beam stop to facilitate measurements of the absorption of the sample, which is, for instance, needed to scale measurements of the same protein in differently absorbing buffers (e.g. high and low potassium concentration) (Figure 2.8).
The last part of the experimental set-up is the detector. At this beam line linear, area or quadrant position-sensitive gas proportional detectors are used [Gabriel, 1977] which can count single photons and thus have an excellent signal-to-noise ratio.
The response of the detector to X-ray photons is not strictly constant over its entire active

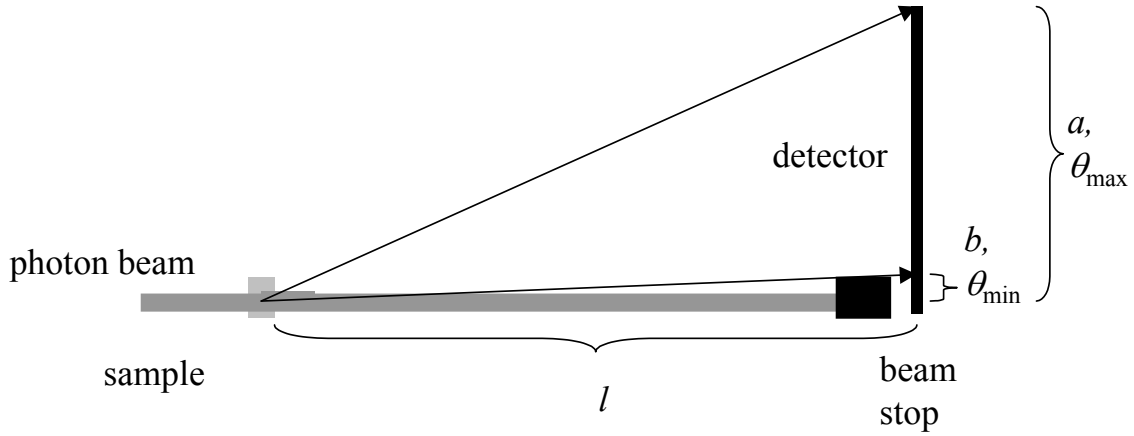


Figure 2.7: The camera length (l), the active area of the detector and the size of the beam stop, which in turn is determined by the size of the last aperture, define the maximum and minimum angles.

area and it is essential to measure its response DR(n), which can be done in two different ways. One is to use the homogenous fluorescence of an iron foil and the other one is a radioactive iron source (^{55}Fe). The fluorescence of the iron foil, which is irradiated with 8 keV synchrotron radiation, gives a flat signal that does not depend on the angle. During data processing the experimental data are divided by the detector response, to correct for the efficiency fluctuations along the detector (Figure 2.9).
Before starting the measurement the correlation between the channels of the detector and s-values must be established. This is done by measuring a standard sample with a well-known scattering pattern (e.g. collagen, tripalmitin, Ag-behenate) which is used to calibrate the scattering angle axis along the detector. In principle this could be calculated from the wavelength of the radiation and the distance between sample and detector but a calibration is more accurate and convenient.

For small angles collagen fibres are used for the calibration. The repeat of collagen is d=64.5 nm (calcified turkey tendon collagen). The scattering vector is defined as $s = \frac{1}{d} = \frac{2\sin(\theta)}{\lambda}$. ($2\theta$ is the scattering angle, λ is the wavelength of the radiation, 0.15 nm in this case.) (Figure 2.10). After calibration the measured data become comparable to data measured with other set-ups, at other beam lines and even at other wavelength (as long as the wavelength are far from any absorption edge). Alternatively one uses Ag-

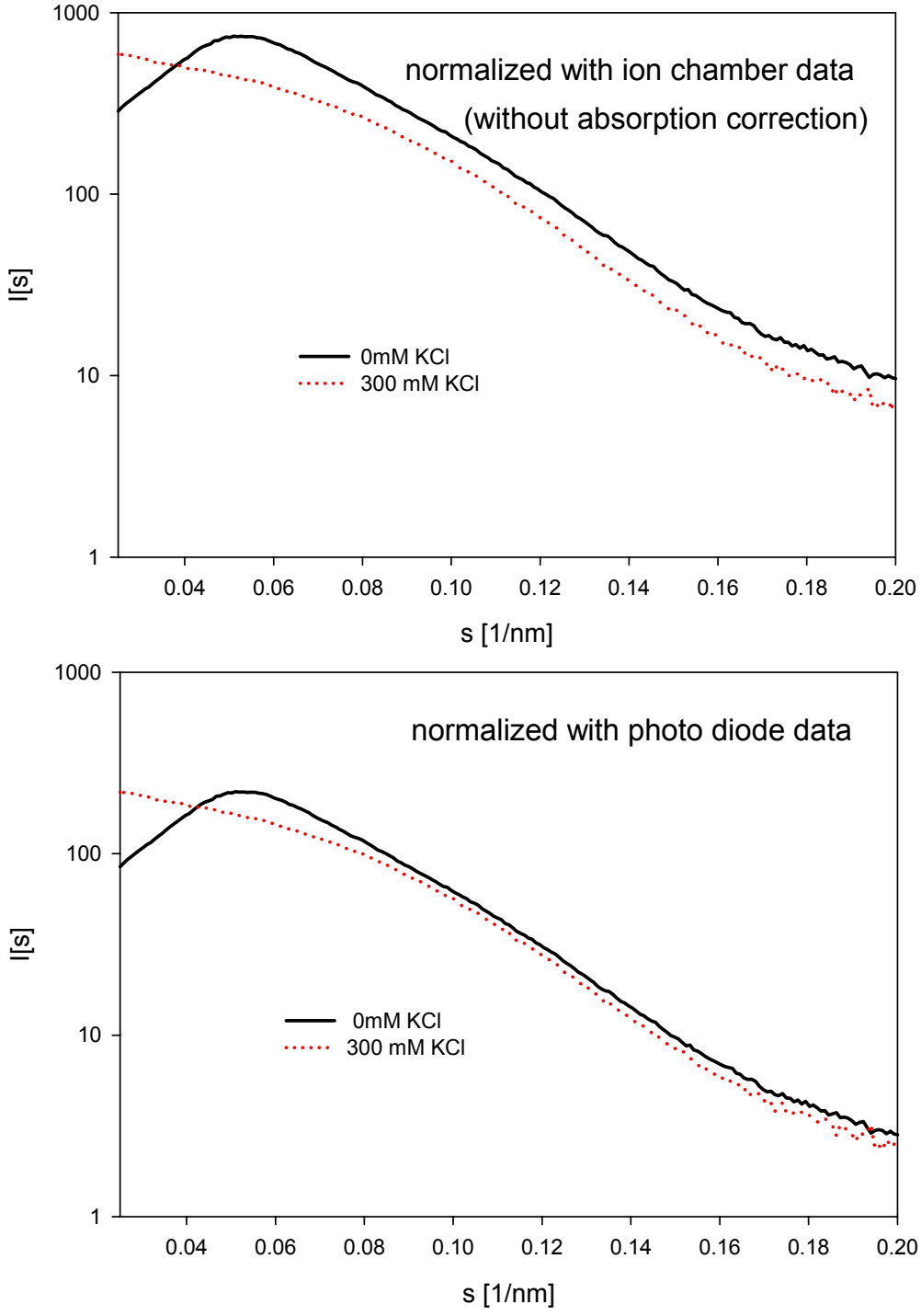


Figure 2.8: Comparison of the normalisation of differently absorbing buffers. The photo-diode normalises on the transmitted beam, the ionisation chamber on

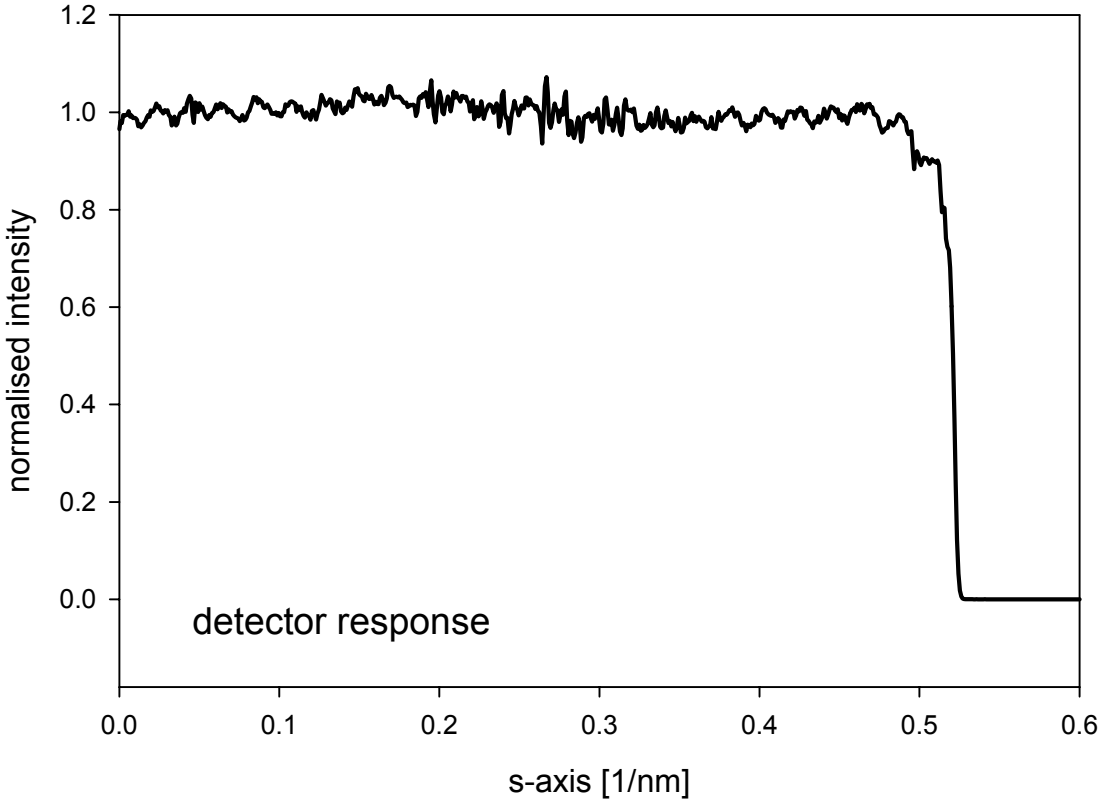


Figure 2.9: Detector response DR(*n*) measured with fluorescence from an iron foil. The detector data was calibrated with collagen. Outside the active area of the detector the response is zero.

behenate (0.5838 nm) and tripalmitin (4.06 nm) as SAXS standards.
When the sample is a macromolecule in solution the scattering due to the excess electron density between molecule and solvent is effectively measured. Thus the scattering of the buffer has to be recorded and subtracted from that of the solution to obtain the scattering of the solute. To control the cleanliness of the cells and the stability of the system one normally measures the scattering of the buffer before and after each sample. A additional method for the verification of the results is to slice the measurement in time frames collecting for example 15 frames of 1 minute each rather than one frame of 15 minutes. During data processing the frames are compared and influences like movement of the beam, radiation damage and bubbles are recognised.
Small angle scattering experiments demand particularly accurate X-ray detectors. One typical characteristic of a small angle scattering pattern of a macromolecule in solution is the large dynamic range in scattering intensity. The resolution is less of an issue. Position-sensitive gas proportional detectors (PSGPD) are extensively used in X-ray and neutron scattering experiments. Some of their characteristics - single event detection, dynamic range of 4-5 orders of magnitude, low noise, short dead-time and fast readout allowing rapid time framing - make them eminently suitable for small-angle X-ray scattering (SAXS) experiments.
Devices that detect radiation based on its interactions with a gas were used as early as the late nineteenth and early twentieth century by Marie and Pierre Curie who used ionisation chambers for their experiments.

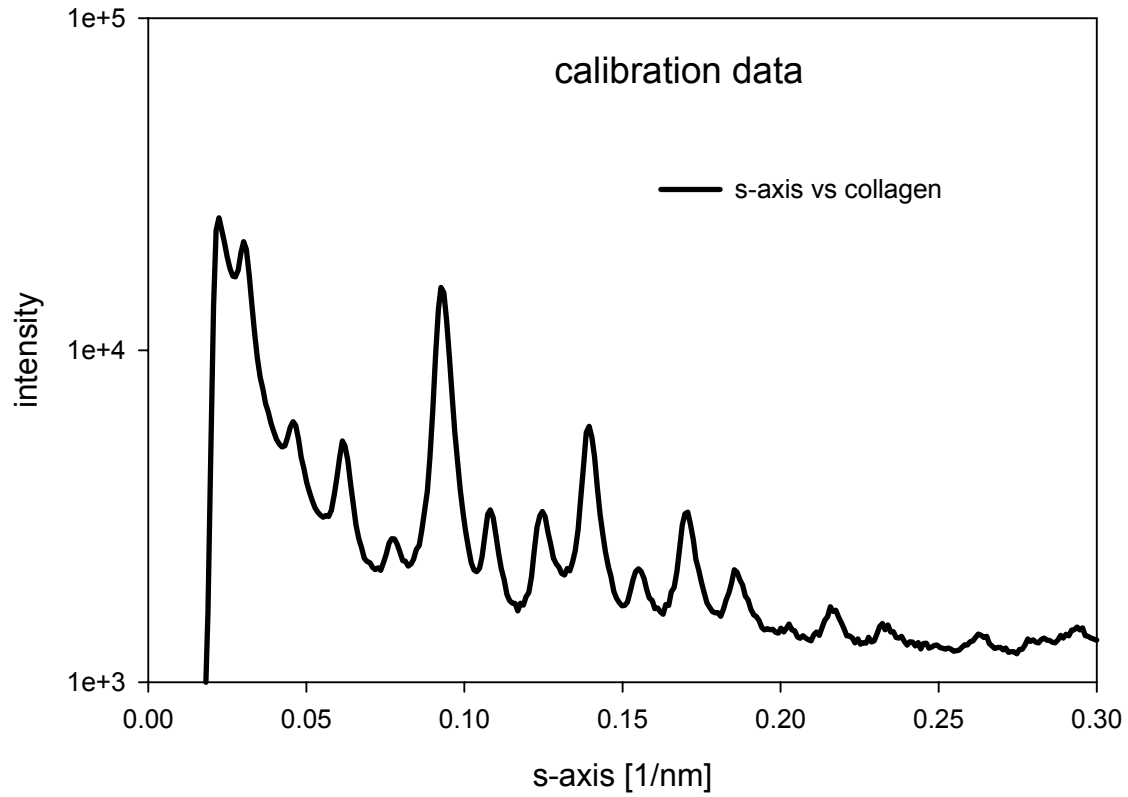


Figure 2.10: Scattering pattern of collagen as it is used for the calibration of the s-axis for SAXS measurements.

The first counting gas detector was the Geiger-counter invented by Hans Geiger and Ernest Rutherford in 1908 for the experiments which established the existence of a heavy positive nucleus at the centre of each atom.
In the second half of the twentieth century gas proportional counters were used and improved in particle physics. In 1992 Georges Charpak received the Nobel price for his development in 1968 of a new wire chamber for high energy physics which used modern electronics and was directly connected to a computer.
Soon after the successful application in particle physics the small angle scattering community started to use fast position sensitive gas detectors for their experiments. [Gabriel, 1977]

3 Gas Detectors

This section deals with the basics of the detection of X-rays in a gas detector. All values below are calculated for 8 keV X-rays, as used at the X33 beamline, and for a linear LC-delay line readout gas detector [Gabriel, 1977] (Figure 3.1) .

3.1 Processes in the gas

All gas detectors are based on the interactions between the incident X-ray photons and the atoms of the gas. The photons enter the detector through a thin beryllium window (0.3 mm, ~5.5% absorption), which is as transparent for X-rays as a glass window for visible light (4% reflected at the surface). The detection process itself is based on photoelectric absorption. The detector is filled with a mixture of argon and CO_2 at 2.3 bar. In most cases a K-shell electron (binding energy 3.2 keV) of an argon atom absorbs the energy of the X-ray photon and a photoelectron (5.8 keV) is emitted. The excited argon ion can decay in two different ways: emission of an Auger electron (89%) or emission of a fluorescence photon (11%) [Krause, 1979]. Fluorescence has some side effects because the photon transports some of the energy and deposits it at a different location. The primary electron and eventually an Auger electron will not go further than a few micrometer. They loose their energy in collisions with other atoms and thereby ionise further gas atoms and produce secondary electrons. On average 26.4 eV are needed to create one electron-ion

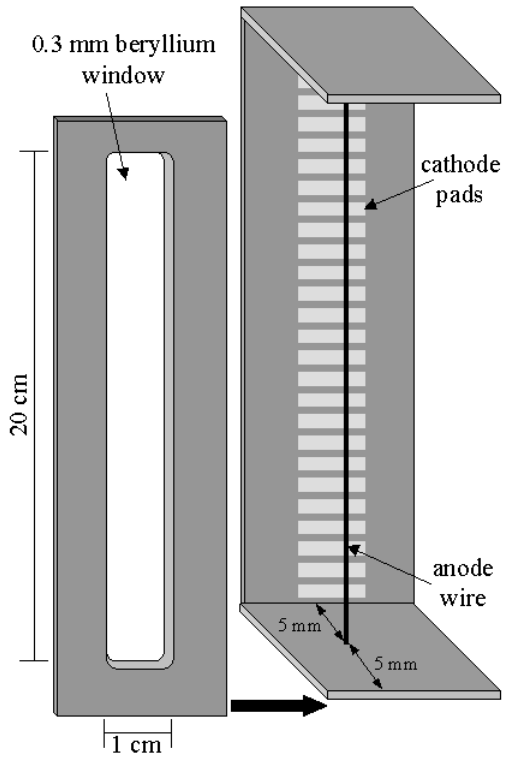


Figure 3.1: Gabriel detector with delay line cathode.

pair in argon [Blum, 1993]. Thus the 8 keV X-rays used at the X33 beam line could ideally generate ca. 300 secondary electrons in pure argon. If ca. 3 keV are carried away by fluorescence only 220 electrons will be produced.

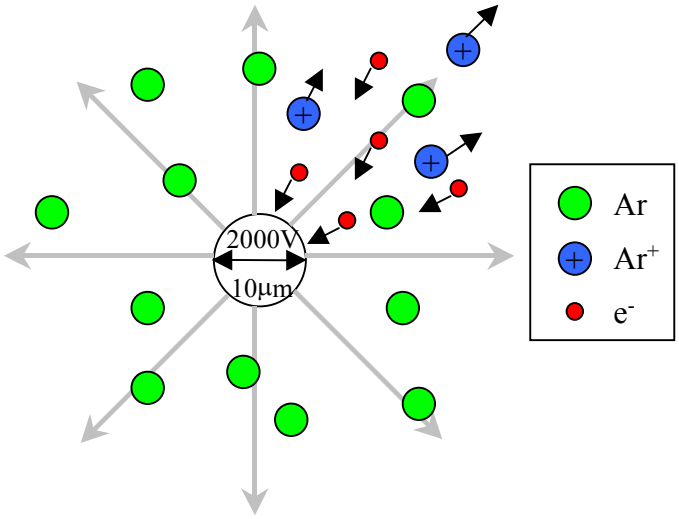


Figure 3.2: Free charges are created in the vicinity of the wire in an avalanche process.

The number of these electrons is still too small to generate a measurable signal and gas detectors of single photons would not be possible without gas amplification, which occurs when a sufficiently high electrostatic field is applied across the gas volume. A thin anode wire (diameter 10 µm) in the centre of the detector is set to a voltage of approximately 2.3 kV. The gap between anode and cathode is about 5 mm. The electrostatic field accelerates the free electrons in the direction of the anode wire. The electrons drift to the wire; during their drift their number N remains constant. In the vicinity of the wire the electrostatic field is so intense that the electrons gain enough energy between electron-atom collisions to ionise further atoms (Figure 3.2) and an avalanche starts which produces an increasing number of electron-ion pairs. In a Gabriel detector the number of secondary electrons is multiplied by a factor of $10^5 - 10^6$. Finally, this avalanche of free charges induces a signal at the electrodes that is sufficiently large to be detected by the readout electronics. In a classical proportional counter the integral of the pulses is proportional to the energy of the incident photon. The gas gain G is the ratio between the number of electrons N after the avalanche divided by the number N_0 of primary and secondary electrons. It depends on several parameters: the gas mixture, the electric field (which itself depends on the shape and layout of anode and cathode) and the gas pressure. The electric field in the avalanche region can be approximated by that in a cylindrical capacitor. The distance between the electrodes is in this case the distance between wire and detector window (b = 5 mm). The radius of the inner electrode is equal to the radius of the wire (a = 5 µm).

$$E(r) = \frac{V}{\ln\left(\frac{a}{b}\right)} \cdot \frac{1}{r} \tag{3.1}$$

Another useful parameter is the Townsend coefficient α that is a measure for the increase of the number of charges along the path s to the wire.

$$\mathrm{d}N = N\alpha\,\mathrm{d}s \tag{3.2}$$

It depends on the electric field E and the density ρ of the gas. s is the distance from the centre of the wire. This leads to an expression for the gain G:

$$G = \frac{N}{N_0} = \exp\int_{s_{\min}}^{a} \alpha(s)\mathrm{d}s = \exp\int_{E_{\min}}^{E(a)} \frac{\alpha(E)}{\mathrm{d}E/\mathrm{d}s}\mathrm{d}E \tag{3.3}$$

N is the number of electrons in the avalanche and N_0 is the initial number of electrons. $\frac{\mathrm{d}E}{\mathrm{d}s}$ is the electric field gradient. s_{min} is the position where the electrons have a sufficiently high energy to ionise other atoms.
Gas amplification is described by Diethorn's formula, which is valid under the assumption that α is proportional to E (which is reasonable for heavy noble gases and field-pressure ratios between 10^2 and $10^3 \frac{\mathrm{V}}{\mathrm{cm\,Pa}}$):

$$\ln G = \frac{\ln 2}{\ln(b/a)} \frac{V}{\Delta V} \ln \frac{V}{\ln(b/a) a E_{\min}(\rho_0)(\rho/\rho_0)} \tag{3.4}$$

For an 8 keV photon $N_0 = \frac{8\mathrm{keV}}{e\Delta V}$. V is the voltage. $E = e\Delta V$ is the amount of energy needed to ionise a single atom. ρ_0 is the normal density.
By charge separation free electrons and ions are produced but the anode and cathode signals are not generated upon absorption of electrons and ions at the electrodes. In order to understand the signal generation it is useful to think of the detector as a capacitor. If a charge travels between two points under the influence of an electric field E the electric energy of the capacitor will change by an amount ε, which is proportional to the charge and the potential difference between the two points.

$$\Delta\varepsilon = \int_{p_1}^{p_2} qE\,dr = q(\phi_1 - \phi_2) \tag{3.5}$$

This change generates the signal. Since the potential in the avalanche region is nearly at the same level as the wire the electrons do not contribute much to the signal. Most of the signal arises from the movement of the ions and especially from the beginning of their drift because of the intense field in the avalanche region.
The careful tuning of the gas gain is an important requirement for the use of a gas counter. The voltage should not be too low because in this case the gain does not suffice to detect the photons. If the voltage is too high the detector starts sparking or the avalanches become self-sustained.
The choice of the gas is another important aspect: one component of the most frequently used gas mixtures is a heavy noble gas (e.g. 70% argon). Noble gases are resistant against ageing processes, which is induced by the radiation, because their ions usually do not form chemical bonds with other materials.
The second component needed in the detector gas is a quencher for UV radiation. The processes in the avalanche region produce UV radiation, which is not adequately absorbed by argon. These photons would spread through the gas and ionise atoms outside the avalanche region. A continuous discharge could be the consequence. Thus CO_2 or methane, which strongly absorbs in the UV are added as minor component (30%) of the gas filling.

3.2 Readout electronics

The avalanche of charges induces a signal both at the cathode and at the anode. To facilitate the localisation of the avalanche the cathode of the detector is connected to a delay line (Figure 3.3). For technical reasons the delay lines used have typically ten pads per 2.54 cm but the number of pads does not determine the resolution. The signal is distributed over several pads and the resolution is limited by the precision with which it is

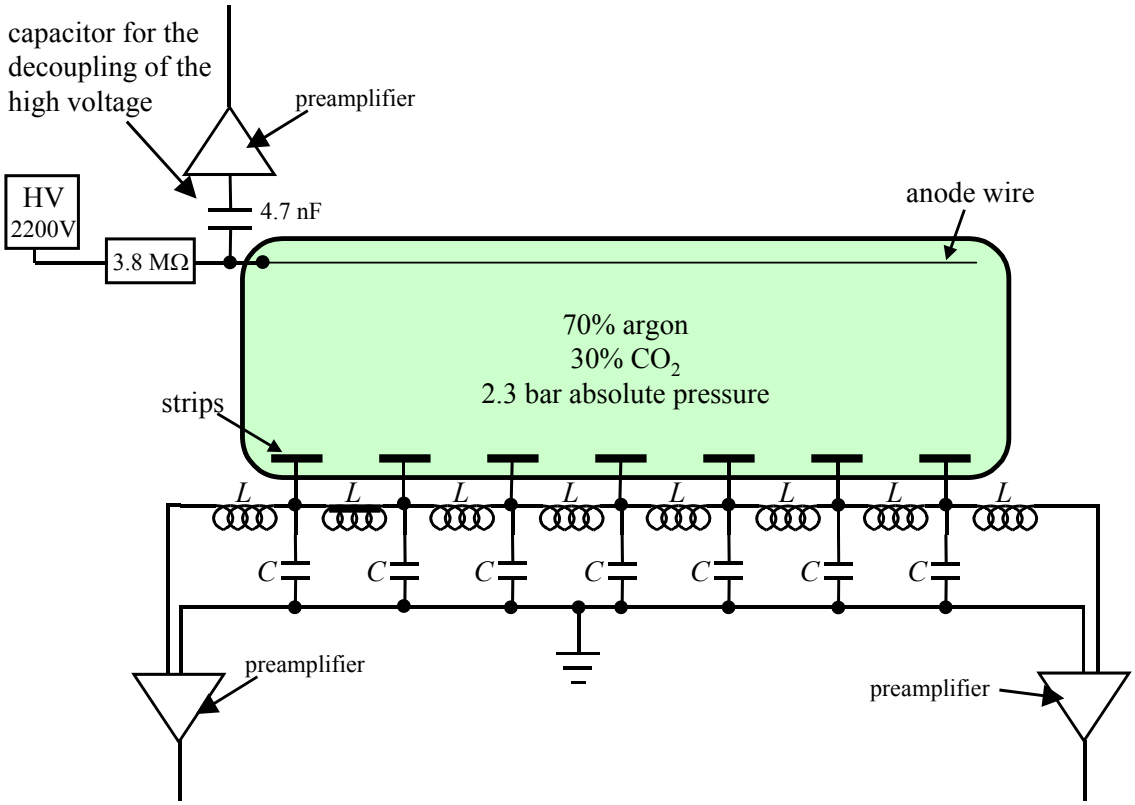


Figure 3.3: LC delay line readout of a Gabriel detector. For simplification the number of pads and LC elements is reduced.

possible to locate the centre of the distribution (i.e. in practice the rise time of the signals at the end of the delay line).
Below, the set of prompt anode signal and delayed cathode signals is referred to as an event. The position of the initial perturbation in the detector can be determined from the temporal correlation between the anode and/or cathode signals.
Inductor-capacitor (LC) ladder network delay lines provide an efficient and cost effective readout for PSGPD [Gabriel, 1977]. In this approach the pads of the cathode are directly connected to the elements of a delay line with total delay *T*. The perturbation induced by the avalanche spreads in opposite directions along the line as described by the telegraph equation and can be detected after suitable amplification and discrimination at both terminals.
Another important parameter of analog electronics is the characteristic impedance z_0. The characteristic impedance is important when different circuits are connected. If both circuits have the same z_0 then a wave originating from the first circuit can enter the second one without reflection. If the characteristic impedance of two circuits is not matched, reflections occur. This effect is comparable to a transition of a light wave from a medium with refractive index n_1 to one with n_2. A fraction of the light wave will be reflected when $n_1 \neq n_2$. For the transition from air to window glass 4% of the incident light is reflected. Therefore the delay line has to be terminated with a resistor $R_0 = \mathrm{Re}(z_0) = \sqrt{\frac{L}{C}}$ and the characteristic impedance of the preamplifiers, which amplify the cathode signals, has to be matched to the delay line. The anode signal has to be inverted after the preamplifier to be accepted by the discriminator (Phillips 715). Because the timing of the signals is essential for an accurate localisation constant fraction discriminators are used for anode and

cathode. Discriminators have a threshold which input signals have to exceed in order to produce an output pulse. In contrast to leading edge discriminator, which produces an output pulse at the time when the input pulse crosses the threshold voltage, the constant fraction discriminator fires after a constant fraction *f* of the input pulse. This circumvents timing problems that occur in leading edge discriminators if the amplitude of the pulses varies. If the shape of the pulses varies or two pulses overlap even the timing of a constant fraction discriminator becomes imprecise. The NIM signals of the discriminators are used to reconstruct the position of the avalanche.
The most common approach is to use only the two cathode signals for the reconstruction [Gabriel, 1977] as illustrated in (Figure 3.4.1). The time difference between these signals is proportional to the position of the avalanche. One cathode signal is used to start the time measurement, the other to stop it. An additional delay can be used to avoid the measurement of negative and zero times.
Alternatively, one of the cathode signals and the anode signal can be used (Figure 3.4.2). This will only affect the range of measured time differences.
In the third method only one cathode terminal is used [Gabriel, 1977] and the second end of the delay line is left open. The signal will be reflected to the end that is connected (Figure 3.4.3). This method does not improve the quality of the measurement but is the least expensive one as only one preamplifier/discriminator is needed.
These three approaches have in common that they reduce the information about an event to a single value, the time difference between two signals. They work perfectly well as long as there is only one event within the delay time *T,* but as soon as two or more events occur within the interval *T* they lead to incorrect localisation.

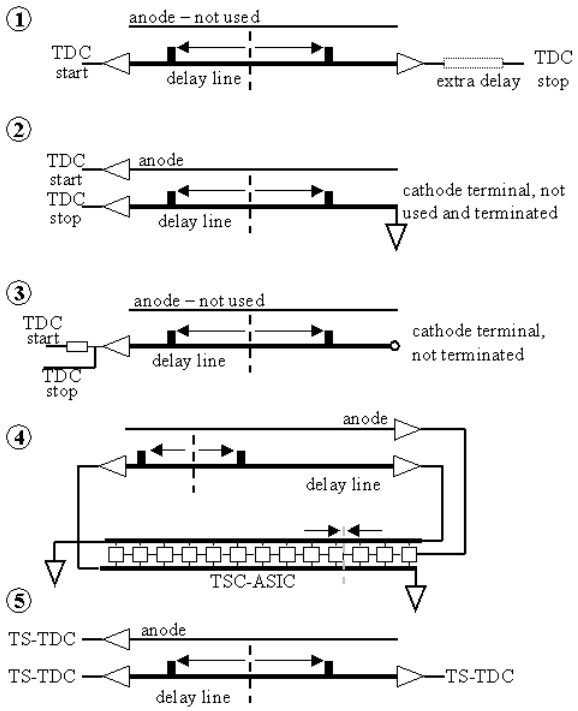


Figure 3.4: Different readout schemes. Two-cathode readout with extra delay (1), anode-cathode readout (2), one-cathode readout (3), time-to-space converter (4), time-stamp TDC (5).

The problem is circumvented in the time-to-space converter (TSC) [de Raad Iseli, 2001], which is based on two back to back delay lines (Figure 3.4.4) with equal delay *T*. The two cathode pulses belonging to the same event propagate in opposite directions on the lines and cross at a fixed time $\Delta t_{TSC} = T$ after their generation. The anode pulse delayed by *T* is used to trigger the readout of the crossing.
The TDCs, which are usually used for the readout schemes above, are combined with an

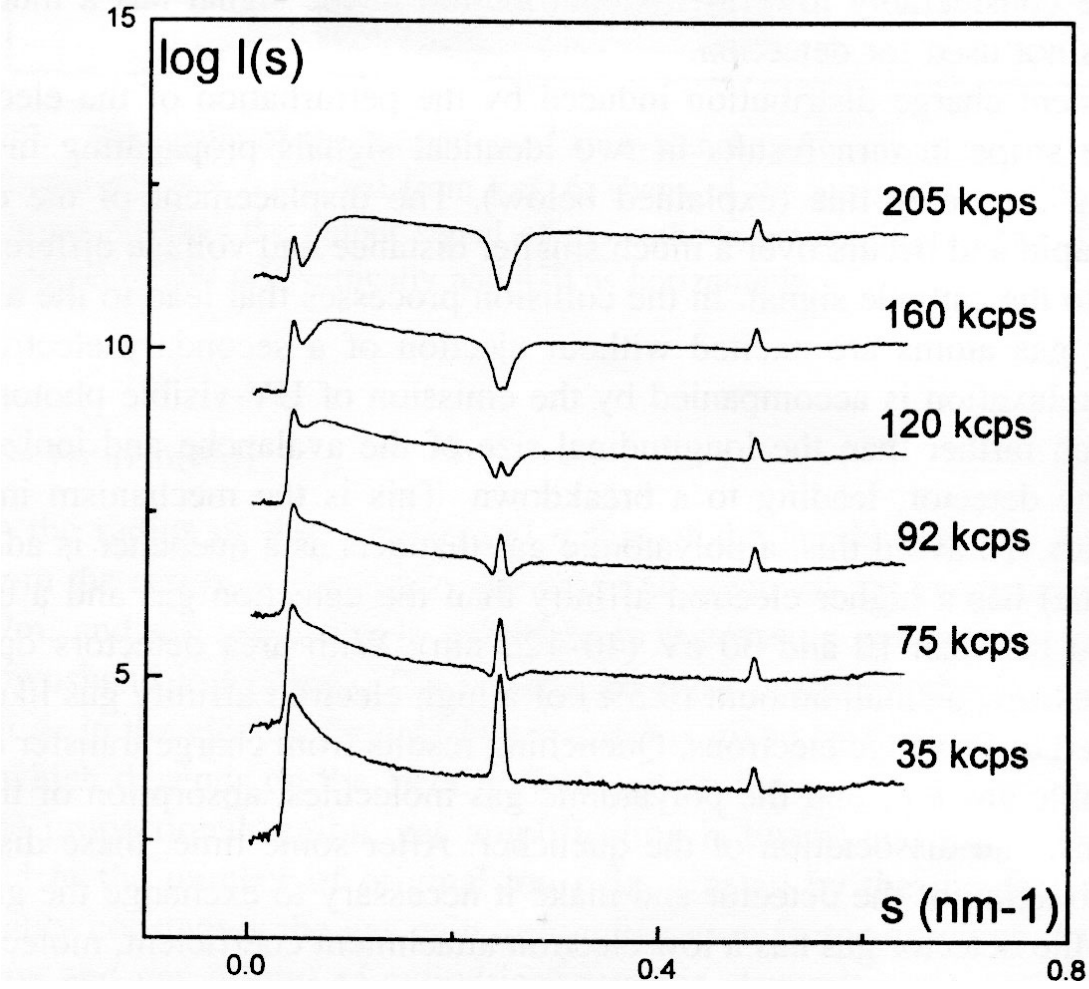


Figure 3.5: Scattering pattern (Log(intensity [a.u.])) versus scattering vector $s = 2\frac{\sin\theta}{\lambda}$ of tripalmitin measured at increasing count rate, illustrating the effect of space charges (Ar/CH_4 70:30, 6 atm). The curves have been displaced along the ordinate for better visualisation.

online histogramming unit, where the scattering patterns are built immediately by direct memory increment.

3.3 Advantages and disadvantages

Like any detection method for X-ray photons gas detectors have advantages and disadvantages. As mentioned before some of their characteristics - single event detection,

dynamic range of 4-5 orders of magnitude, low noise, short dead-time and fast readout allowing rapid time framing - make them eminently suitable for small-angle X-ray scattering (SAXS) experiments.
One of the less relevant but annoying problems of gas detectors are aging effects. During operation some deposits can accumulate on the wire. In the worst case this can disturb the field so that the detector starts sparking at this position. In this case the wire has to be exchanged.
A major drawback is the so-called space-charge effect. Whereas the electrons need only a few hundred nanoseconds to propagate through the entire detector (0.5 cm) the ions are much slower. Their mass is several thousand times large than that of an electron and therefore the same electrostatic force results in an much weaker acceleration. Also the average velocity is lowered by the fact that their mean free path between two collisions is only a fourth of that for electrons [Sauli, 1977]. They make up a cloud of positive charges that screens the potential between cathode and anode and hence lowers the electrostatic field. As a consequence the gas amplification and thus the induced signals become smaller. Some can not be detected any more. The higher the local rate of avalanches the higher the probability that some photons will not be detected. Consequently the scattering patterns are disturbed in regions of high intensity like peaks. As an extreme case peaks can become troughs due to the space charge effect (Figure 3.5). This effect reduces the applicability of gas detectors for experiments like crystallography on storage rings where the intensity of strong peaks is measured.
Another problem of gas detectors is the small absorption cross section per volume of gas. Therefore Gabriel linear gas detectors are usually operated at a pressure of approximately 2 bar. Under these conditions the counter can reach an efficiency of 30 % (8 keV photons). Increased pressure may improve the efficiency but at the same time the space charge problem becomes more severe since the drift velocity of the ions decreases.
Raising the voltage can not cure space charge effects because the drift velocity of electrons and ions runs into saturation.
Another approach to increase efficiency is to use a drift space in front of the anode that enlarges the active volume of the detector. Electrons are also produced in this drift space. The electrostatic field is not sufficiently high to result in appreciable gas amplification and the electrons are just transported to the anode.
As mentioned in the section about readout schemes a disadvantage of the common delay line readout is the fact that the TDC cannot distinguish between correlated (i.e. belonging to the same event) and non-correlated (i.e. belonging to different initial events or parasitic pulses) START and STOP signals. If two events occur at the same time at different positions in a gas detector they will both generate a START and a STOP pulse on the cathode. A measurement is started by a START and stopped by the next STOP signal. It is not possible to decide whether the pulses belong to the same initial event. In the worst case the electronics will register two false events.
As they do not allow to check whether or not the signals belong to the same event the methods cannot handle double events and a pile-up rejection mechanism must be implemented to avoid errors. This mechanism verifies that one and only one START and one and only one STOP occur within a certain time interval, usually 2*T*. With an LC-delay line readout one can establish the rejection mechanism easily by using the anode signal (Figure 3.6). This rejection mechanism is efficient and losses associated with it can easily be calculated. The rejection time can be assumed to be the dead-time of the detector system and the data acquisition. Information on how to calculate the losses due to dead-time for the cases of paralysable and non-paralysable response can be found in the literature [Knoll, 1977]. The commonly used formulae have, however, to be modified because the detection

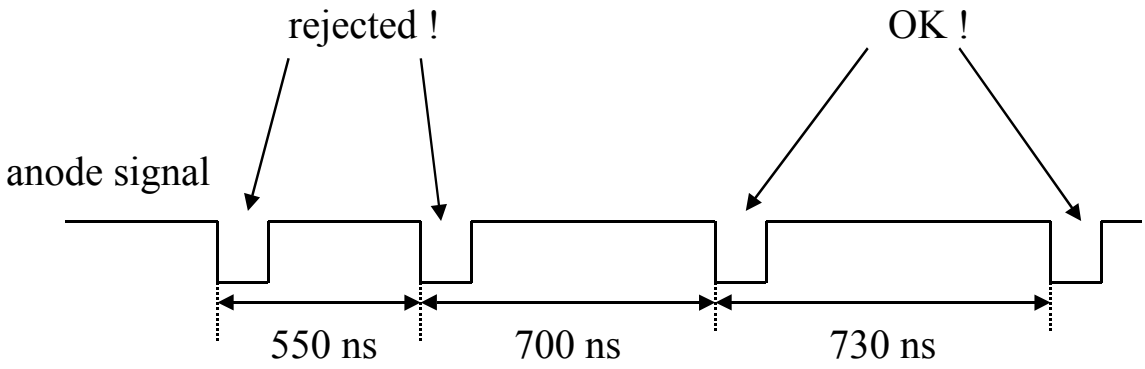


Figure 3.6: Rejection based on the anode signal. If two anode signals are separated by less than 2*T*=600 ns they are rejected.

system is an intermediate case. For a nonparalysable model the true interaction rate *n* would be

$$n = \frac{m}{1 - m\tau}. \tag{3.6}$$

m is the recorded count rate and τ is the system dead-time. For a paralysable model

$$m = ne^{-n\tau} \tag{3.7}$$

is valid.
This formula can be derived by taking into account that the probability of waiting times between random events is given by an exponential distribution:

$$p(t) = ne^{-nt} \tag{3.8}$$

n is the average rate of events.

$$P(\tau) = \int_{\tau}^{\infty} ne^{-nt}\,dt = e^{-n\tau} \tag{3.9}$$

gives the probability that the waiting time between two events is larger than τ. But the situation for readout electronics with rejection is slightly different: An event is only recorded if the waiting time between the event and the next one and the waiting time between the event and the previous are greater than τ.

$$P_R(\tau) = P(\tau) \cdot P(\tau) = e^{-2n\tau} \tag{3.10}$$

Calculations based on Poisson statistics [Knoll, 1977], for a random source, indicate that for a time interval of 300 ns, typical for LC delay lines, and a rate of $10^5\,s^{-1}$ two events will only occur in 3 % of the cases (for $10^4\,s^{-1}$ only 0.3%). At higher rates this limitation is more relevant.
For typical gas detectors used at moderate count rates the problem due to statistical double events are less important. This fact is also documented by the use of standard start-stop readout systems without rejection mechanism in many present synchrotron radiation experiments [Bateman, 2000].

3.4 Gas electron multiplier (GEM) detectors

Today the development of other types of gas detectors for synchrotron radiation applications continues. The 'thin-wire' concept is not the only one that leads to gas amplification. Another detector type is the gas electron multiplier (GEM). The working principle of a GEM detector is a thin insulator plate with small holes (diameter 70 μm, separated by 140 μm, Figure 3.7) in it. A voltage (400V) is applied to the unconnected

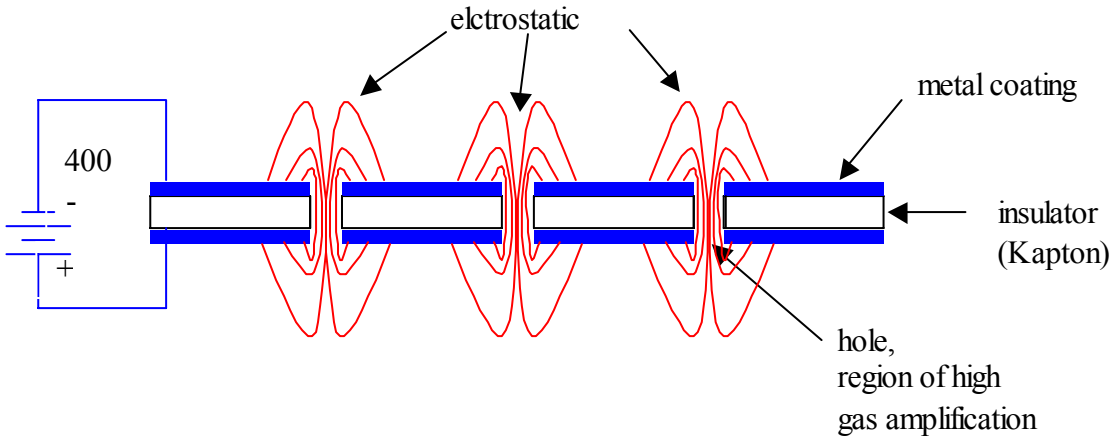


Figure 3.7: Working principle of a GEM detector.

metal coatings on both sides of the insulator. Consequently the very intense field in the tiny holes leads to high gas amplification. A GEM detector is an efficient preamplifier for single electrons.

The first GEM detector for the use in synchrotron radiation experiments was developed at the ESRF [Kocsis, 2001]. It is based on a two-dimensional delay line readout with one delay line for the x and one for the y-axis. Small pads (Figure 3.8) cover the anode plane. One half of the pads are connected to the x delay line and the other half to y.

There are three GEM layers (Figure 3.9) in the detector with a large drift space in front of the first layer. In this drift space the X-rays produce photoelectrons which drift to the holes where they induce an avalanche. Drift spaces are used in many two dimensional gas detectors for X-rays to compensate for the fact that because of their large, thin windows they can only run at atmospheric pressure, which limits the efficiency.

The ESRF GEM detector was tested in two short beam time periods at the DUBBLE beam line in May 2001 and April 2002.

During the first tests in 2001 it was not possible to acquire the full image of the detector due to limitations of the readout electronics. The efficiency (ratio of acquired counts and counts detected at the last GEM foil) was low (<15%). Possible reasons for this were very long delay lines and readout problems. The main aim of the 2002 measurements was to acquire the full image and to determine the efficiency of the detector-system data acquisition.

The detector was operated at –2786V with a mixture of argon (90%), CO_2 (9.67%) and R14 (4800.49 ppm) at atmospheric pressure. This time there was no permanent flow of gas through the detector. The detector was equipped with a new, shorter delay line: 200 LC segments with a total delay of 170 ns for x and 160 ns for the y-direction. The delay lines were made of SMD elements (inductor and capacitance) with 20% tolerance. In 2001 a printed circuit delay line was used.

The time measurements were made with the V110 TDC module [Herve, 2002]. This was connected to a VISTA histogramming unit. The data were transfer from VISTA to the data acquisition computer via a LAN interface.

At DUBBLE the X-ray energy was set to 8 keV. Images were taken at camera lengths of 3.6 m and

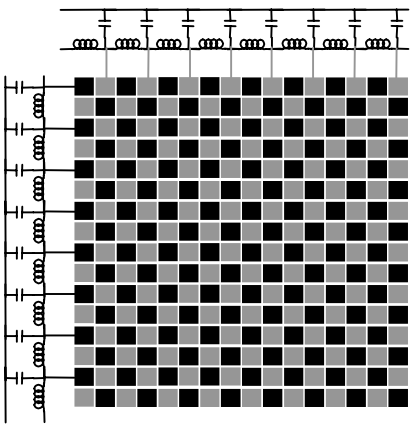


Figure 3.8: Anode of the ESRF GEM detector.

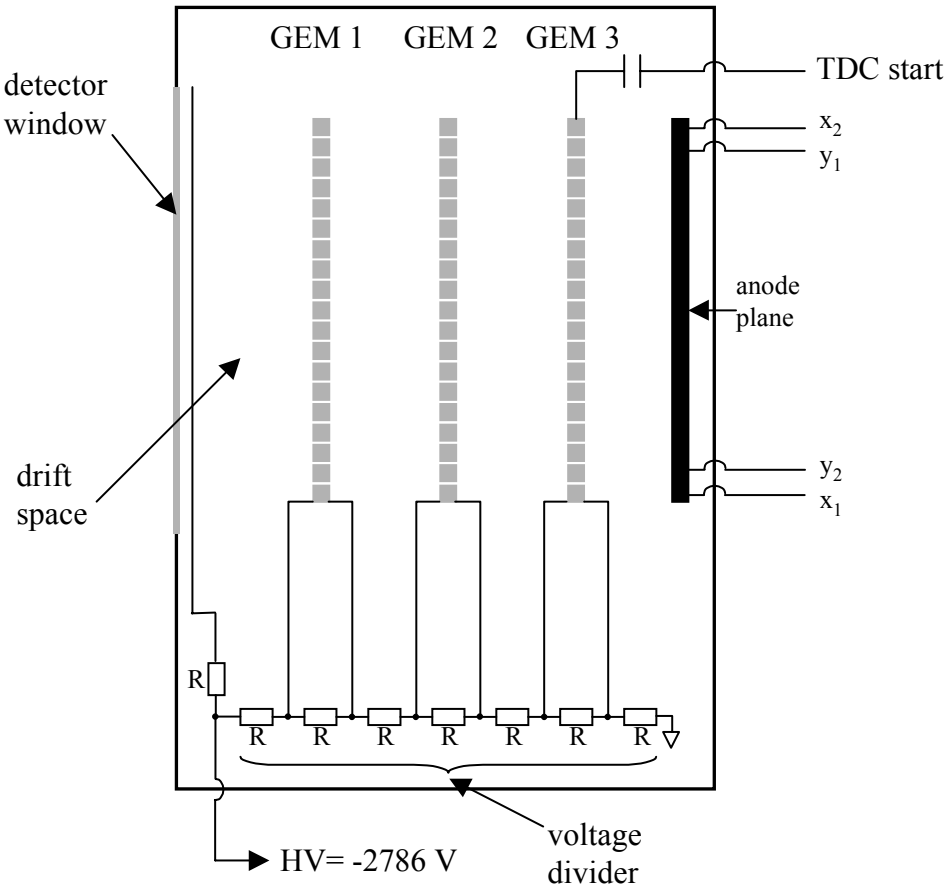


Figure 3.9: Wiring diagram of the ESRF GEM with delay line readout.

1.6 m. The images of several standard samples were collected during the two days of beam time: collagen, silver behenate, tripalmitin, ETOC, PCL, LPE, glassy carbon and an iron foil. During one measurement (iron foil) the detector was covered with a hole mask. In a further series of measurements the influence of the count rate on the pattern was investigated by inserting increasing numbers of aluminium foils in front of the detector as attenuators. In addition the absorption of the material used for the detector window at 8 keV was measured.

During the measurements there were problems to get the data from the histogramming memory to the DAQ computer, especially with the BIT3 interface. Some areas of the image were relocated during the transfer making the data unusable. The LAN interface had occasional timeouts during data transfer resulting in the loss of the data set being transferred and the need to reboot the CPU in the VME crate.

During the second beam time period it was possible to read out the full image of the GEM detector. Also the readout efficiency (ratio of acquired counts and counts at the last GEM foil) was much better (between 86 and 99.9% depending on the count rate). The detector window (0.37 mm carbon and 0.07 mm Mylar) absorbs approx. 27% of the X-rays at 8 keV. The gas (90% argon, 10% CO_2, atmospheric pressure) between the detector window and the first GEM (3.2 mm) absorbs 5.5%. Thus the total efficiency of the GEM is

$$5.5\% \times 73\% \times (99.9\% \text{ to } 86\%) = 4.0\% \text{ to } 3.4\%. \qquad (3.11)$$

Even if one assumes that ionisation behind the first GEM foil gives rise to further counts, the total efficiency should not exceed 5-10%.

Since the detector is operated at atmospheric pressure a beryllium window, like it is used in linear detectors, can replace the carbon-Mylar window. Therefore its absorption would drop from 27% to 4%.
The present images are perturbed at many positions (Figure 3.10). There are as expected some horizontal lines of empty channels resulting from dead regions where two GEM foils are glued together. In addition some vertical lines were also observed which are most probably due to a readout problem.. There is also a diagonal cross structure on the image and it seems that some events are shifted to the outer regions of the image. This was not observed in a series of tests done by M. Kocsis and F. Sever with an iron source in the lab afterwards, but some measurements at the DUBBLE beam line with the standard two dimensional gas detector appear to display a similar cross structure [Kocsis, 2002].

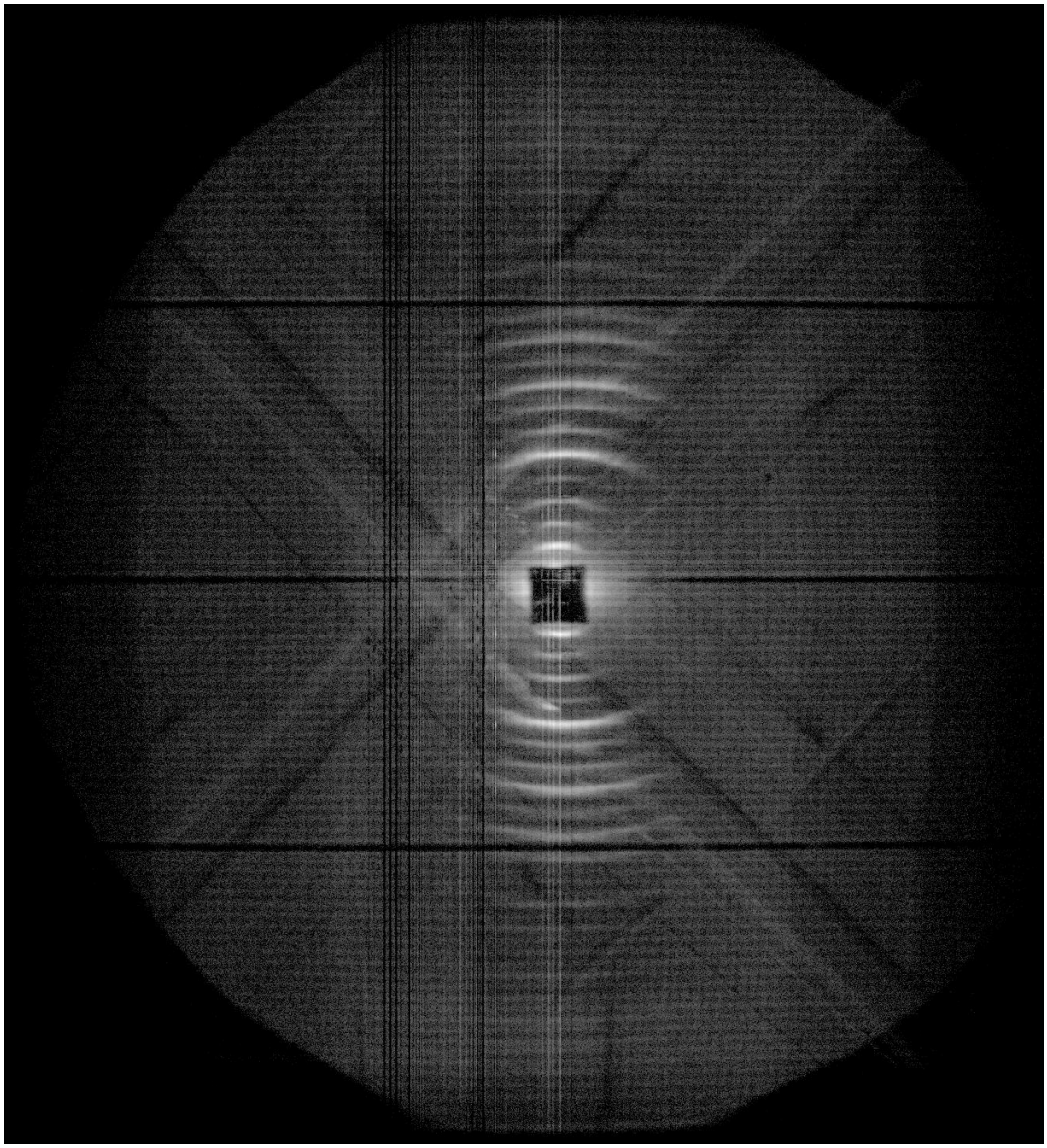


Figure 3.10: Scattering pattern of collagen (camera length 1.6 m, 2048x2048 channels). The black horizontal lines are due to blind areas where the insulator plates were glued together. The vertical lines and the diagonal cross structures are most probably formed by problems of the readout electronics.

Possible reasons could be the discriminators or the long cables between the discriminators and the TDC.
Analysis of the hole mask measurement (5 mm distance between the holes) illustrates that the images are locally distorted (Figure 3.11). This was not observed in 2001 when a printed circuit delay line was used and is due to the 20% tolerance of the delay line elements. If these distortions cannot be corrected offline SMD elements with lower tolerance or printed circuit delay lines will have to be used.
In the X-direction one readout channel is equivalent to 160 μm and in Y-direction to 140 μm. The FWHM of one hole is approximately eight channels.
The series of measurements with aluminium in front of the detector illustrates that the scattering pattern does not depend on the count rate (data not shown).
In another analysis the scattering pattern of collagen as a linear detector would record it, was extracted and corrected with the detector response (fluorescence of iron foil). The collagen peaks are well separated even at a camera length of 1.6 m. The detector response has some regular sinusoidal structure but the number of counts in this measurement is insufficient to be sure that this is caused by the detector (Figure 3.12). Further investigation is necessary.
The results of the 2002 beam time indicate that the detector is working but that a number of technical improvements to the detector and the readout electronics are necessary to make this a user-friendly and efficient system for experiments.

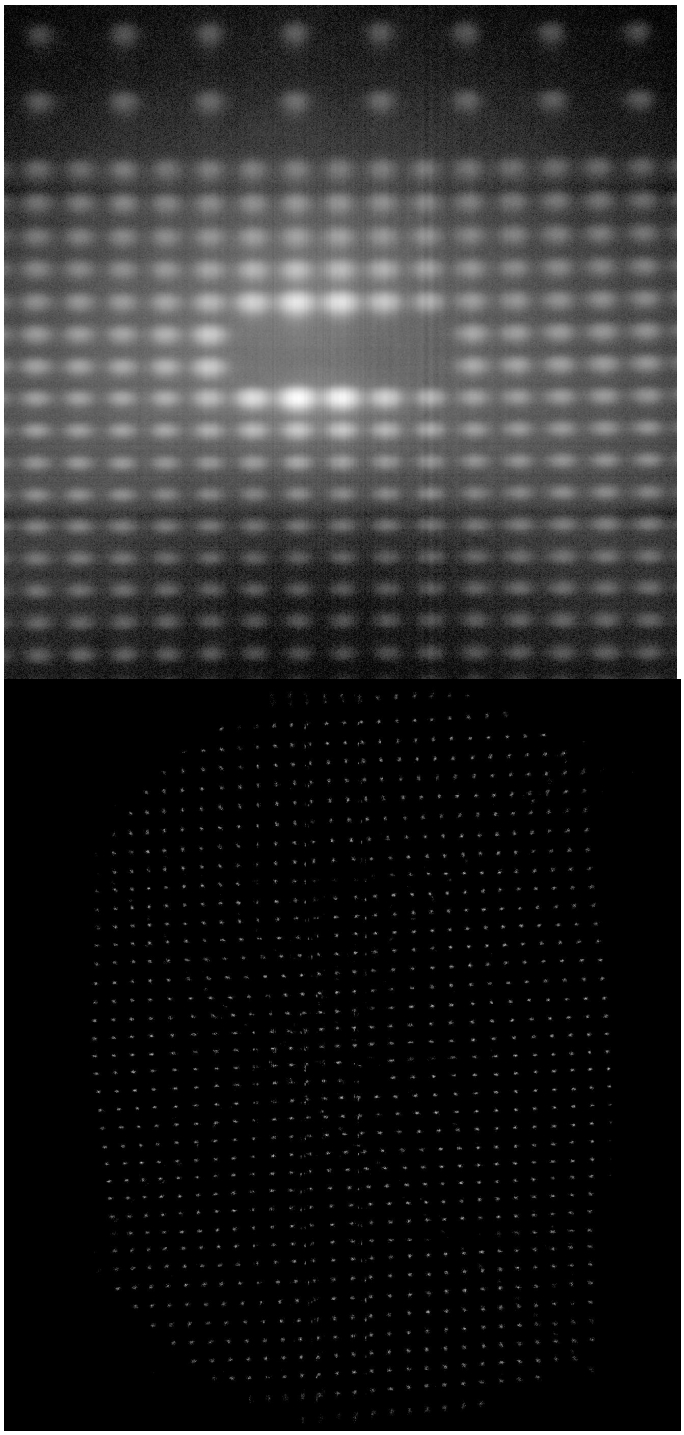


Figure 3.11: Hole mask measurements of 2001 (top) and 2002 (bottom). The images represent different areas and were taken with different masks and exposure times but whereas the lower one is clearly distorted the upper one (taken in 2001) is more symmetrical.

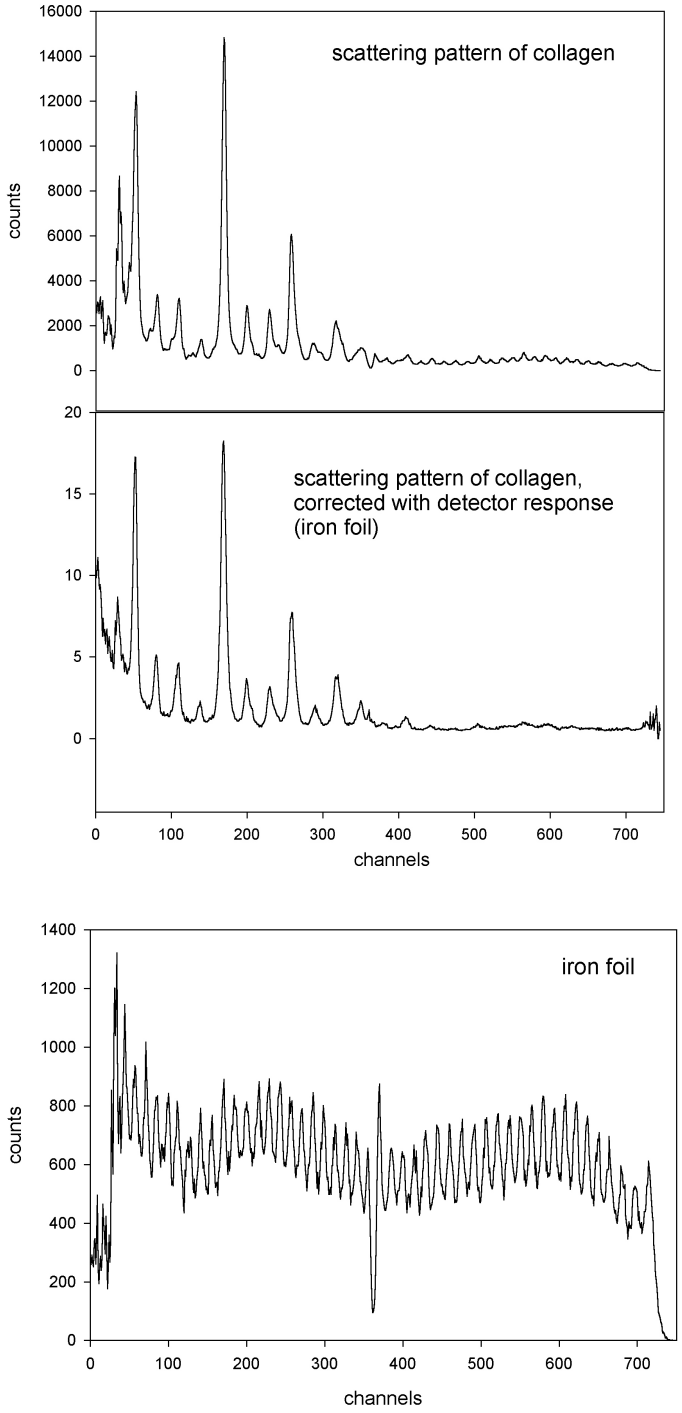


Figure 3.12: Collagen pattern and detector response (iron foil) measured by the ESRF GEM in 2002.

4 Bunch structure and point-spread function

In this chapter the pulsed time structure of synchrotron radiation sources and of the DORIS storage ring in particular are described since this provides the basis for measuring the drift time distribution in a gas proportional detector. The results of actual measurements are compared to theoretical predictions.

4.1 The bunch structure of DORIS

Synchrotron radiation produced with storage rings is never continuous but consists of regularly spaced flashes (bunches). This is a consequence of the acceleration of electrons/positrons in radiofrequency (RF) cavities. The wavelength λ of the RF and the circumference D determines the number n_b of buckets in the ring.

$$n_b = \frac{D}{\lambda} \tag{4.1}$$

A bucket is a continuously moving allowed position of the electron/positron bunches in the ring, corresponding to a potential minimum of the wave. As the time structure of the electron/positron fill in a machine determines also the structure of synchrotron radiation, in the case of DORIS (D=289.2 m; $\lambda = 0.6$ m; $\nu = 500\,\mathrm{MHz}$) where there are 482 buckets the interval between two neighbouring bunches would be only 2 ns if they were all filled. This is adequate for many techniques where a continuous source can be used but for time-

Positron energy	4.45 GeV
Initial positron beam current (5 bunches)	120 mA
Circumference	289.2 m
Number of buckets	482
Number of bunches	1 (for tests), 2 and 5
Bunch separation (minimum)	964 ns (for tests), 480 ns and 192 ns
Horizontal positron beam emittance	404 π nmrad
Vertical positron beam emittance	12 π nmrad
Positron beam energy spread (rms)	0.11%
Curvature radius of bending magnets	12.1849 m
Magnetic field of bending magnets	1.2182 T
Critical photon energy from bending magnets	16.04 keV

Table 4.1: DORIS beam parameters.

resolved measurements, like fluorescence decay experiments, it is necessary to have longer intervals between bunches. Therefore, usually, not all buckets are filled, and at DORIS two- and five-bunch modes with minimum separation times of 480 and 192 ns are used (Table 4.1) corresponding to bunch frequencies of 2.08 MHz and 5.21 MHz respectively. For a typical count rate of a PSGPD at the X33 small angle scattering beam line of 40 kHz one can easily estimate that only one out of 125 bunches gives rise to a scattered photon arriving at the detector.
The positrons in a bunch are spread over a distance of 20 mm (FWHM), and since their velocity is close to the speed of light (~30 cm/ns) this corresponds to 67 ps.
The control room of the storage ring provides an electronic signal, the bunch clock, synchronized with the revolution of the positrons in the ring which corresponds to the time of arrival of a bunch of synchrotron radiation (Figure 4.1). The electronic jitter of the signal arriving at the experimental station is 100 ps [Kaul, 2004].
The bunch structure also has an effect on the performance of the PSGPD. In the previous chapter Poisson statistics were applied. A prerequisite for the validity of this approach is that the probability to detect an event would be distributed uniformly over time. Since the positrons are bunched it is not guaranteed that Poisson statistics leads to correct predictions [Bateman, 2000]. It can, however, be assumed that the detector signals do not have the same time structure as the synchrotron radiation because during formation of the avalanche, which precedes detection, the sharp bunches are smeared out by the variable drift times of the electrons in the detector. The drift time distribution can be measured using the bunch clock and the anode signal in the detector. This allows measuring the bunch structure of the synchrotron radiation as a gas detector sees it and the validity of Poisson statistics in this case.
The bunch clock and the anode signal of the detector facilitate a measurement of the drift time distribution of the detector. The signal of the bunch clock indicates the arrival of the X-ray bunch and thus the time of primary ionisation of the gas in the detector. Before the anode signal occurs upon the formation of the avalanche the electrons have to drift to the wire. The signal recorded by the detector is the anode signal and not the true arrival of the photons. The effect of the time of flight of the photons within the detector on the drift time distribution, is negligible because photons travel approximately 3 cm in 100 ps. The drift

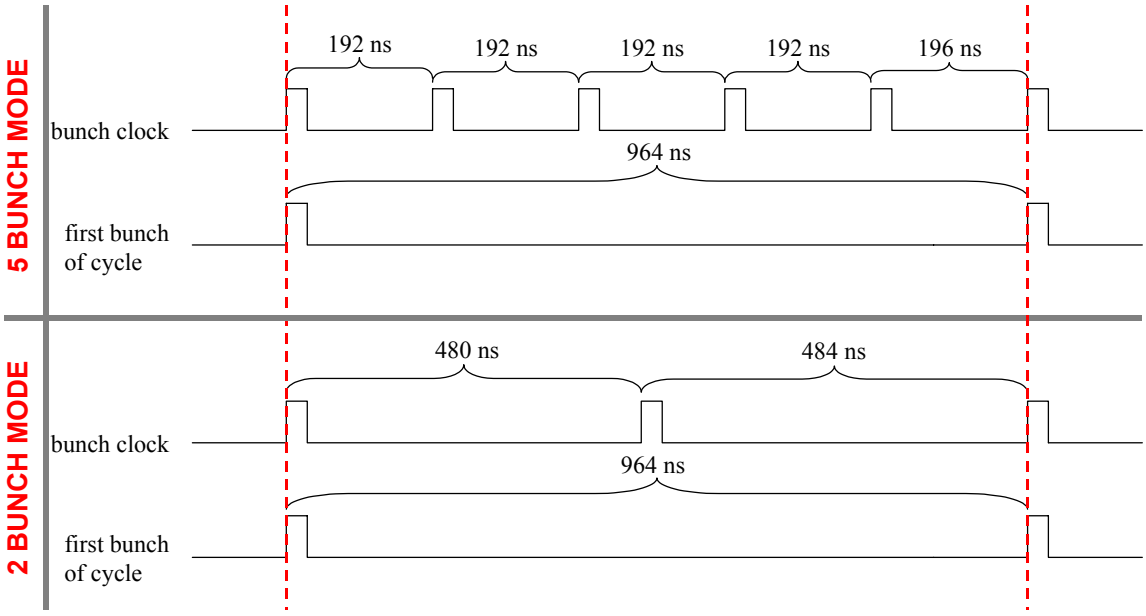


Figure 4.1: Bunch clock signal. Note that the gaps between the bunches in two- and five-bunch mode are asymmetric.

time of the electrons in the detector can thus be directly determined by measuring the interval between the initial ionisation and the avalanche defined by the bunch clock and the anode signal, respectively.
Drift times of electrons in gas detectors are very often used for the detection of ionising particles. In high-energy physics they are used to improve the accuracy of the localisation. Particles like muons at high energies (e.g. 200 GeV) can easily penetrate even thick metal walls even though they are constantly interacting with and ionising matter. Unlike X-ray photons, which always only lead to a single primary ionisation they can thus be detected by several detectors at the same time. One of these detectors can be a fast scintillator, which gives precise timing information, a second could be an array of drift tubes ("straws", gas filled kapton tubes with a central anode wire). The muon will leave a trace of electron-ion pairs in the gas. The free electrons drift to the wire and generate a signal allowing to measure their drift time and hence to determine the distance between wire and trace and thus a more accurate localisation.
As mentioned before it is more difficult to use drift-time based detection schemes for X-rays because, as mentioned above, they can only be detected once. Consequently one has to use the bunch clock for the start-signal of the drift time measurement as in the concept of the drift chamber requiring single bunch mode [Hendrix, 1980].

4.2 Drift time distribution

Before presenting the experimental set-up and the results of our measurements the basics

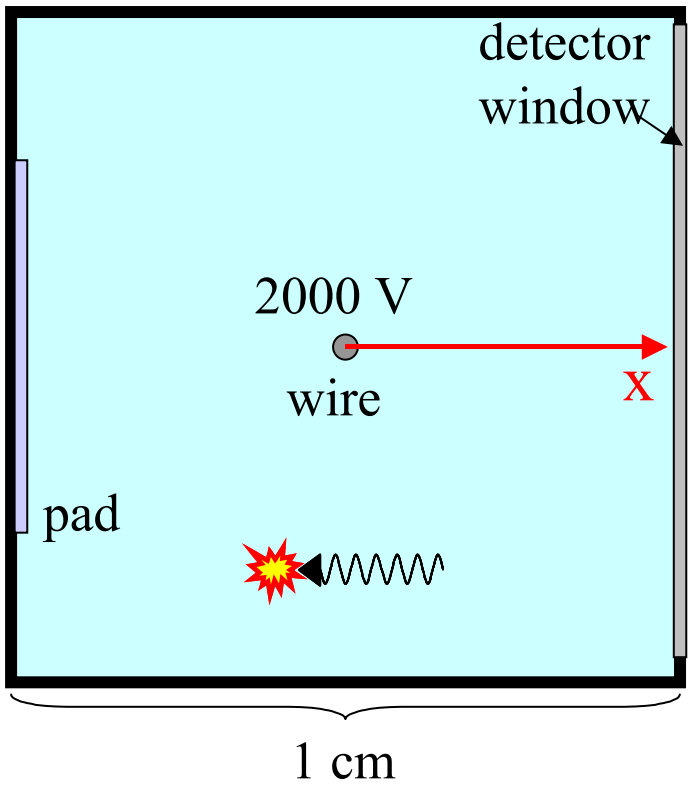


Figure 4.2: Principle of drift time measurements using the bunch structure of DORIS. The photon enters the detector through its window at the time given by the bunch clock. The electrons drift to the wire and gas amplification leads to an avalanche, which generates an anode signal. The time between bunch clock signal and anode signal is equal to the drift time.

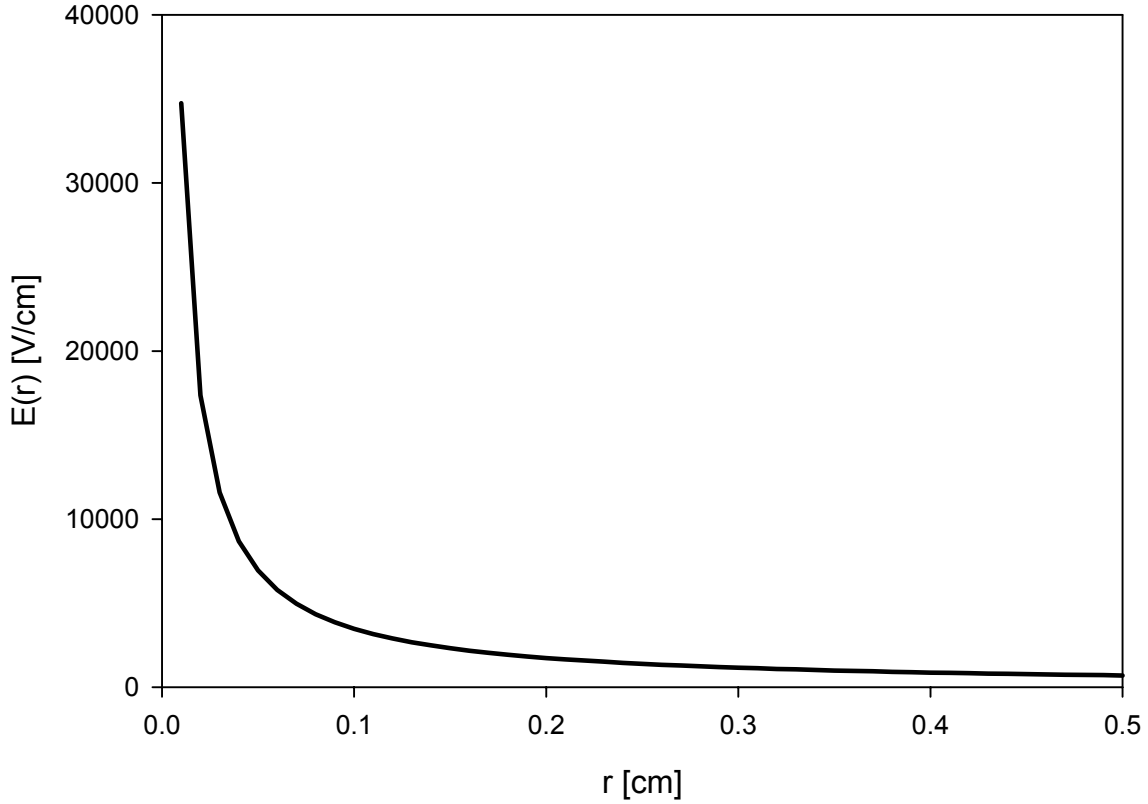


Figure 4.3: : Electric field in the linear Gabriel detector calculated assuming that the detector is a cylindrical capacitor.

of electron drift and theoretical expectations of the drift time distribution will be discussed. The simplest way to set up a model for the drift time distribution is to make a numerical simulation. Due to the symmetry of the detector along the wire the problem can be treated in two dimensions (Figure 4.2). In first approximation one can assume that the primary ionisation sites are equally distributed over the whole area. This is not entirely correct since the absorption of the gas in the detector makes ionisation 30% more likely close to the window than at the back of the detector.

The electric field in the detector can be treated like the field of a cylindrical capacitor with a thin central anode. The distance between the electrodes is in this case the distance between the anode wire (radius a = 5 μm) and the detector window (b = 5 mm). In our measurements the voltage U_{anode} was 2400V (Figure 4.3) and the gas pressure p in the detector 2.3 bar. To simplify later calculations a parameter U_0 is introduced.

$$E(r) = \frac{U_{\text{anode}}}{\ln\frac{a}{b}} \cdot \frac{1}{r} = 347.4\text{V} \cdot \frac{1}{r} = \frac{U_0}{r} \tag{4.2}$$

$$\Rightarrow \frac{E(r)}{p} = \frac{347.4\text{V}}{2.3\text{bar}} \cdot \frac{1}{r}$$

Since the detector is rectangular rather than cylindrical this is only an approximation and the larger the distance to the wire, the larger also the deviation between the approximation and the real field. This approximation suffices, however, to get a correct description of the onset of the drift time distribution and the main features of its later part of it can be discussed qualitatively.

The actual drift time can be calculated if the drift velocity υ as function of r is known. Since the electrostatic field is known, this relationship can be calculated from $\upsilon(E)$. This function depends only on the type of gas and its pressure. On the microscopic scale electrons are scattered by the gas atoms while they are accelerated by the electric field. This leads to an average drift velocity in the direction of the electric field (i.e. towards the anode) that is well below the instantaneous velocity between collisions. In the absence of field is υ=0. With increasing field the drift velocity increases but saturates above certain field strength.
A large amount of data on the drift properties of electrons and ions exists in the literature [Peisert, 1984] and gas discharges have been a major research field in the last century. Among others Heinz Raether, Professor at the Institut of Angewandte Physik at the

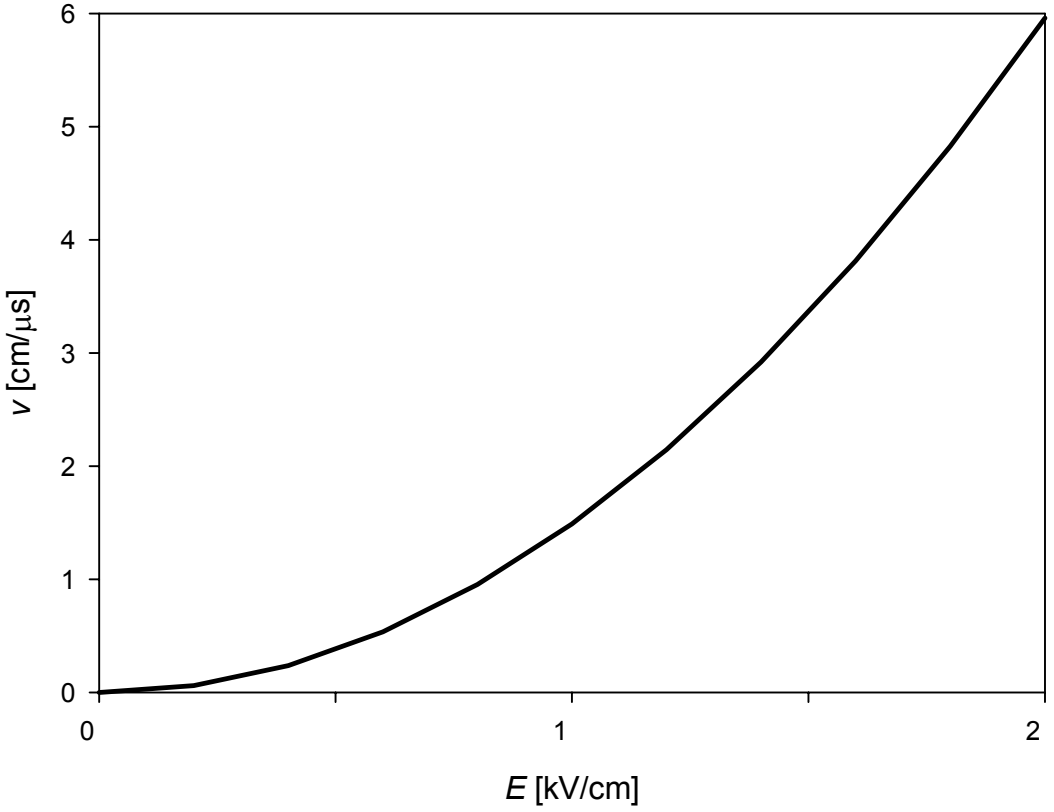


Figure 4.4: Model for the drift velocity υ as a function of the electric field before the drift velocity runs into saturation. For low field strength ν increases proportionally to the square of the electrostatic field.

University of Hamburg from 1951 to 1978, did some work on sparking in gas discharges (Raether condition). Since argon is a commonly used gas for drift chambers many different argon mixtures have been measured. [Ma, 1982] determined the drift velocities *υ(E)* in argon/CO_2 mixtures (90:10, 80:20, 70:30, 60:40 at 4 and 1 bar). For all pressures and mixtures *υ(E)* can be approximated by a parabola. When CO_2 is added to the gas the drift velocity for low field strength increases whereas the saturation velocity diminishes. At higher pressures the electrons drift more slowly for all field strengths. Based on this data *υ(E)* is estimated to be

$$\upsilon(E) = \tilde{\upsilon} E^2 , \tag{4.3}$$

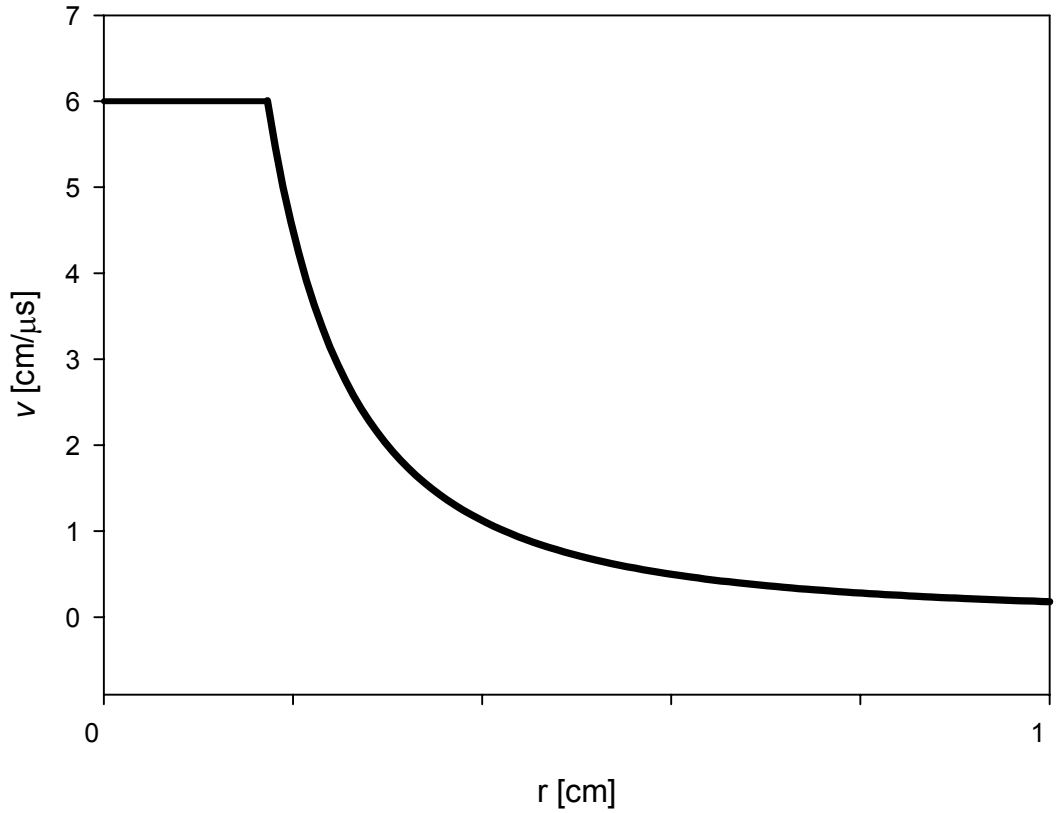


Figure 4.5: The model of the drift velocity field along *r* is used in the simulations to calculate the drift times. It was calculated from the models for the drift velocity as a function of *E* and for the electric field in the detector.

with $\tilde{\upsilon} = 1.49 \frac{\text{cm}^3}{\mu\text{s}\,\text{kV}^2}$ (Figure 4.4) and the saturation velocity is assumed to be $\upsilon_{\max} = 6 \frac{\text{cm}}{\mu\text{s}}$. Consequently the drift velocity as a function of the ionisation site is

$$\upsilon(r) = \begin{cases} \tilde{\upsilon} \dfrac{{U_0}^2}{r^2} & r \geq \sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}} {U_0}^2} \\ \upsilon_{\max} & r < \sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}} {U_0}^2} \end{cases}. \tag{4.4}$$

This velocity field (Figure 4.5) is sufficient to set-up a Monte-Carlo simulation. The wire is placed at the origin of the coordinate system. As discussed above the *x* and *y* coordinates are uniformly distributed random variables, from which radius and thus the velocity can easily be calculated. A simple iterative procedure (with ns-resolution) is then used to calculate the next position until the wire is reached. This was done for 2×10^5 X-ray photons and the results were used to obtain the drift time distribution. With the set of parameters U_{anode}=2400V, $\tilde{\upsilon} = 1.9 \frac{\text{cm}^3}{\mu\text{s}\,\text{kV}^2}$ and $\upsilon_{\max} = 6 \frac{\text{cm}}{\mu\text{s}}$ a distribution (Figure 4.6) is obtained, which increases very rapidly in the beginning before reaching a maximum around 20 ns (part 1). The distribution decreases in the second part but the rate of change

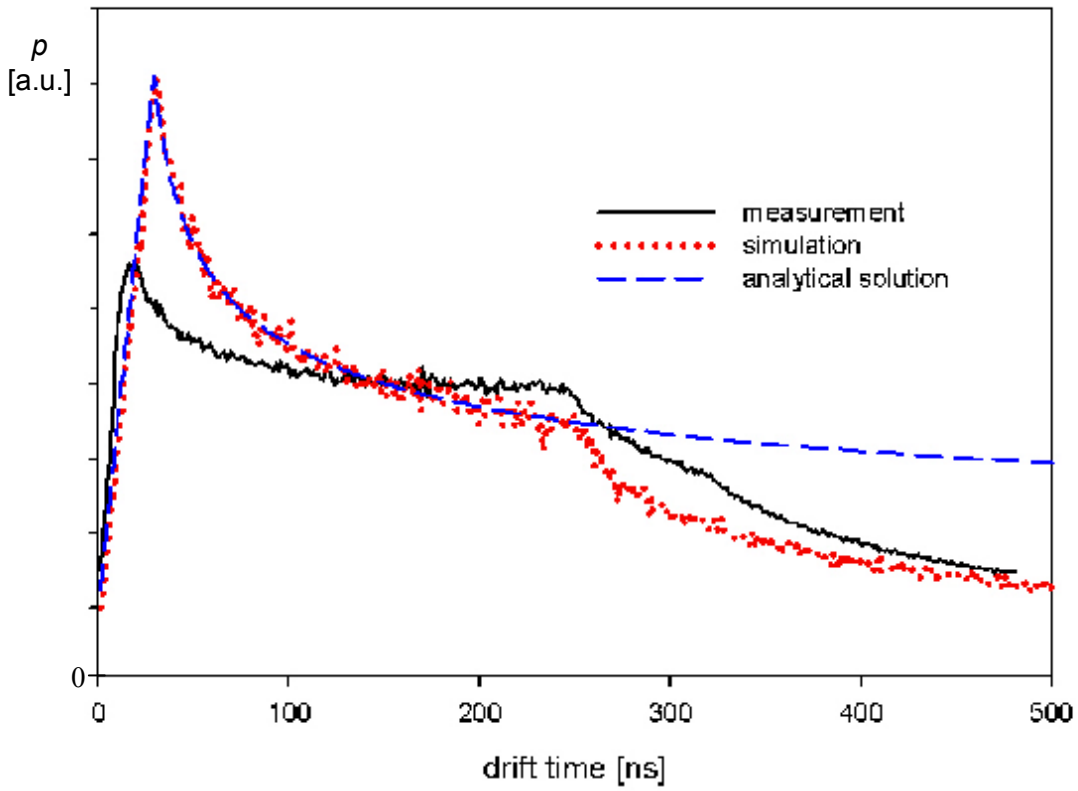


Figure 4.6: Comparison of measurement, simulation and analytical solution of *p* the probability (simulation) or number of counts (experimental) as a function of drift time.

diminishes for larger waiting times. At a certain point the distribution has a discontinuity after which it drops faster and finally reaches 0 (part 3).
The onset of the distribution is caused by the electrons, which are generated in the region of constant drift velocity (saturation). The decrease in part 2 comes from the fact that the drift velocity increases as $\frac{1}{r^2}$ further from the wire. The larger the difference in drift velocity between neighbouring ionisation sites, the wider the spread of arrival times of the corresponding electrons at the wire. The discontinuity between part 2 and 3 is due to the geometry of the detector. Whereas the effective area from where the electrons originate for a fixed distance to the wire increases proportionally to *r,* due to the cylindrical geometry in the edges of the detector this effective area decreases. This leads to a more rapid decrease of the drift time distribution, which finally drops to zero.
The simulated drift time distribution can now be compared to data acquired with a Gabriel type detector at the X33 beam line at 2.3 bar absolute pressure with an anode voltage of 2400V.
One way to measure the drift time distribution is to connect the bunch clock to the start and the anode signal to the stop of the TDC (method A). This has the disadvantage that each bunch starts a measurement but there is only one anode signal for approximately 125 bunches. Thus very often a measurement is started but it is not stopped before the internal time-out of the TDC. A better way to measure the drift time distribution is to connect the detector signal to the start and the bunch clock to the stop (method B). This set-up is used, for instance, in fluorescence decay measurements. For the measurement of the drift time distribution there is no real advantage because the number of photons is so large that the

required number of counts is in both cases reached after a short time. Thus Method A was used because the time axis always has the conventional direction.
The asymmetry of the time structure of the storage ring (Figure 4.1) together with the delay of the bunch clock cables must also be taken into account. This complication can be avoided by using the "first bunch of cycle" signal instead of the signal that gives a pulse for every bunch. It is a signal with a period of 964 ns which thus allows to measure drift

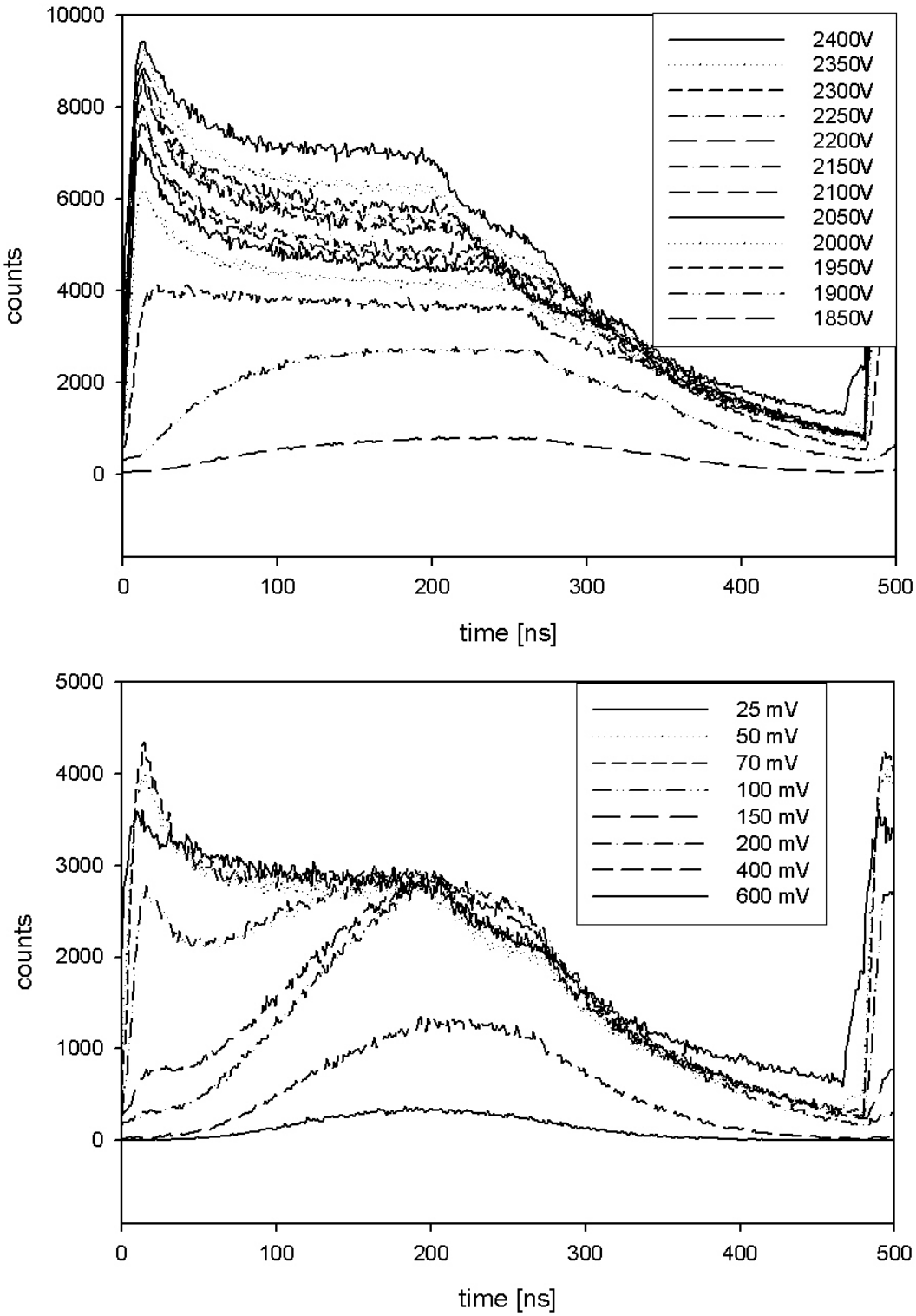


Figure 4.7: Measured drift time distribution as function of the anode voltage at a fixed discriminator threshold of 25 mV and of the discriminator threshold (lower graph) at a fixed anode voltage of 2400V.

times between 0 and 964 ns. These correspond to the drift time distribution generated by five successive synchrotron radiation bunches.
An extra delay must be used to compensate the delay of the cables and synchronise the bunch signal to the actual arrival of the bunches, to avoid distortions of the distribution.
If the value for the extra delay is not exactly known one can set it to an arbitrary value. A drift time distribution of five successive bunches is still measured but the origin of the time axis does not correspond to onset of the first bunch as it would with the exact delay value. This measurement could either be used to extract the proper extra delay or a complete drift time distribution for one bunch because the onset of the distribution is characterised by a sharp increase and its length is also known (192/196 ns).
All measurements presented below were made in the two bunch mode. In five bunch mode the drift time distributions of successive bunches overlap due to the spread of drift times. In two bunch mode the overlap is small and all predicted features of the simulation can be identified in the experimental drift time distribution (Figure 4.6). The simulation thus takes into account all important effects and facilitates the qualitative prediction of the drift time distribution. Nevertheless there are differences. In the onset of the distribution, part 1, the simulated distribution increases more slowly and the maximal value is also higher. The simulated maximum is also larger relative to the rest of the distribution. Note that, for reasons explained below, the distributions are not normalised, but they were scaled to facilitate comparison. The interpretation of part A is relatively simple. It can be attributed to the electrons coming from the region of constant saturation drift velocity. The relation between r and υ is very simple in this case and one would expect that from the slope of the distribution it should ideally be possible to determine $\upsilon_{\max}$. This requires the distribution to be normalised since otherwise the slope is arbitrary. All attempts to determine $\upsilon_{\max}$ led to too high values above $10\frac{\mathrm{cm}}{\mu\mathrm{s}}$. A series of measurements of the drift time distribution with increasing thresholds of the constant fraction discriminator at the anode revealed the cause of the problem (Figure 4.7). For the measurement, which was compared to the simulation, the threshold was set to the minimum (25 mV). If the threshold is raised the shape of the distribution changes as the pulses with short drift times seem on average to be smaller than those for the rest of the distribution. The changes in the threshold series between 25 and 50 mV suggest that there should also be changes between 0 and 25 mV. Some photons, which cause ionisation close to the wire, cannot be detected. This leads to a loss of pulses at short drift times. The assumption is in agreement with observations made by others [Bednarek, 1980], which can be explained by the so-called self-induced space charge effect. The number of charges created by the X-rays should be the same for all ionisation sites but charges, which are created close to the wire, have a shorter drift distance and time than those generated in the outer regions. During the drift the charge clouds diffuse but a dense charge cloud leads to self-induced space charge effects and consequently to a weaker signal. Figure 4.7 illustrates also that the slope can increase when signals are missing as illustrated by the curves for 25 and 50 mV discriminator settings.
In part 2 where the drift velocity is proportional to $\frac{1}{r^2}$ the decrease of the simulated distribution is faster than in reality. One reason could again be that some signals in the onset are missed. Also, the transition from constant drift velocity to $\frac{1}{r^2}$-decrease is not exactly taken into account in the model. The time from part 2 to part 3, which depends mainly on $\tilde{\upsilon}$, can be reproduced with the parameters chosen for the simulation. The exact shape of the end of the distribution cannot be reproduced, but this was expected since it

depends on the field in the corners of the detector, which can not be described by the simple cylindrical model.
The main difference between the simulation and measurement is in part 2. Changing the parameters of the simulation (U_{anode}, $\tilde{\upsilon}$ and $\upsilon_{\max}$) manually did not improve matters. In order to fit the parameters to the measurement an analytical expression for the distribution was set up based on Equation 4.4. This yields an expression for the drift time t as a function of drift distance r.

$$t(r) = \int_0^r dr' \frac{1}{v(r')} = \begin{cases} \dfrac{r}{\upsilon_{\max}} & r \le \sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}}} U_0 \\ \dfrac{2}{3}\sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}^3}} U_0 + \dfrac{1}{3\tilde{\upsilon} U_0^2} r^3 & r \ge \sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}}} U_0 \end{cases} \tag{4.5}$$

The inverse function *r(t)* is given by

$$r(t) = \begin{cases} \upsilon_{\max} \cdot t & t \le \sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}^3}} U_0 \\ \sqrt[3]{3\tilde{\upsilon} U_0^2 \left(t - \dfrac{2}{3}\sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}^3}} U_0 \right)} & t \ge \sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}^3}} U_0 \end{cases} . \tag{4.6}$$

The probability of ionisation is uniform over the entire detector area. The probability of drift distances for the cylindrical symmetry is thus proportional to r and given by

$$p_r(r) = \rho_0 r . \tag{4.7}$$

The box shape of the detector leads to intractable terms in $p_r(r)$. Since the drift behaviour in the corners cannot be calculated the analytical solution is only given here for a cylinder within the detector and will not reproduce part 3 of the distribution. The distribution p_r of the random variable r has to be converted to p_t of the random variable t with $t(r)$ given by Equation 4.5. In general

$$p_t(t) = p_r\big(r(t)\big) \cdot \frac{dr}{dt}(t) \tag{4.8}$$

is valid.
With

$$\frac{dr}{dt}(t) = v(t) = \begin{cases} \upsilon_{\max} & t \le \sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}^3}} U_0 \\ \dfrac{\sqrt[3]{3\tilde{\upsilon} U_0^2}}{3} \left(t - \dfrac{2}{3}\sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}^3}} U_0 \right)^{-\frac{2}{3}} & t \ge \sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}^3}} U_0 \end{cases} . \tag{4.9}$$

one obtains

$$p_t(t) = \rho_o \begin{cases} \upsilon_{\max}^2 \, t & t \le \sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}^3}} U_0 \\ \dfrac{\left(\sqrt[3]{3\tilde{\upsilon} U_0^2}\right)^2}{3} \left(t - \dfrac{2}{3}\sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}^3}} U_0 \right)^{-\frac{1}{3}} & t \ge \sqrt{\dfrac{\tilde{\upsilon}}{\upsilon_{\max}^3}} U_0 \end{cases} . \tag{4.10}$$

The problem of the proper choice of the normalisation factor ρ_0 is circumvented by choosing it such that the average probabilities ($\hat{\rho}$) between 50 and 200 ns of the analytical solution and the data ($\hat{\rho}$ = 6283.7) are equal.

$$\hat{\rho} = \frac{1}{0.150\mu s}\int_{0.05\mu s}^{0.2\mu s} p_t(t)dt = \frac{\rho_0}{0.15\mu s}\cdot\frac{\left(\sqrt[3]{3\tilde{\upsilon}U_0^2}\right)^2}{2}\left[\left(0.2\mu s - \frac{2}{3}\sqrt{\frac{\tilde{\upsilon}}{\upsilon_{\max}^3}}U_0\right)^{\frac{2}{3}} - \left(0.05\mu s - \frac{2}{3}\sqrt{\frac{\tilde{\upsilon}}{\upsilon_{\max}^3}}U_0\right)^{\frac{2}{3}}\right] \tag{4.11}$$

This normalisation of course does not normalise the integral of the probability distribution to unity but facilitates the comparison of the shape of the distribution of the experimental data and the analytical solution, especially in part 2. The value of $\tilde{\upsilon}$ is set to very low values (which is only possible because of our normalisation) and even this solution cannot reproduce the decrease in part 2 quantitatively.
Another point that is not represented in the simulation and might have an effect is the readout electronics (described in chapter 3). Even though a constant fraction discriminator is supposed to process pulses of different height identically, there may be differences if the pulse shape (rise time) changes. One reason for such changes could be the self-induced space charge effects mentioned above. Their consequences for the timing of the discriminator output signals are difficult to predict and can not be calculated in the present case.
The anode voltage was varied in a series of measurements (Figure 4.7). A lower voltage has two effects on the drift time distribution: it reduces the gas amplification and, more importantly in this context, also lowers the drift velocity of the electrons. Consequently, the transition between part 2 and 3 shifts to larger drift times when the voltage is lowered.

4.3 Measurements with a point-like source

A further series of measurements with a point like X-ray beam was set up to test the drift time model. In order to have a thin and collimated beam a hole-mask with three columns and four rows of holes (rows are orthogonal to the anode wire, and columns parallel) was fabricated (Figure 4.8). The mask, which is made of a 2 mm thick aluminium plate (ratio of transmitted and incident beam $\frac{I_T}{I_0} < 10^{-11}$ for 8 keV photons), has the same dimensions as the front panel of the detector. The rows can be used to irradiate the detector with synchrotron radiation beams of 500 μm in diameter at different distances from the anode wire. The row of 11 holes (0.5 mm diameter, 0.4 mm spacing) at the top of the mask was used for this purpose. All other rows and columns were covered. In each of the measurements only one hole (0.5 mm diameter) of this row remains uncovered allowing a narrow photon beam to penetrate the detector window. Along the track of the beam argon atoms are ionised and the resulting electrons drift to the wire where they give rise to an avalanche (approx. 28% of the photons at 8 keV are absorbed by the gas). The drift time distributions are presented in Figure 4.9. In contrast to all previous drift time measurements these were made with the storage ring operated in five-bunch mode. For the outer holes the onset and the mean value of the distributions move away from the bunch signal. The average drift times become longer and the distribution broader. The wire is located close to hole 6, which corresponds to the shortest drift times and the sharpest onset of the distribution (Figure 4.9). For a given hole the onset of the distribution corresponds to the arrival at the wire of electrons, which are generated closest to the wire. If one

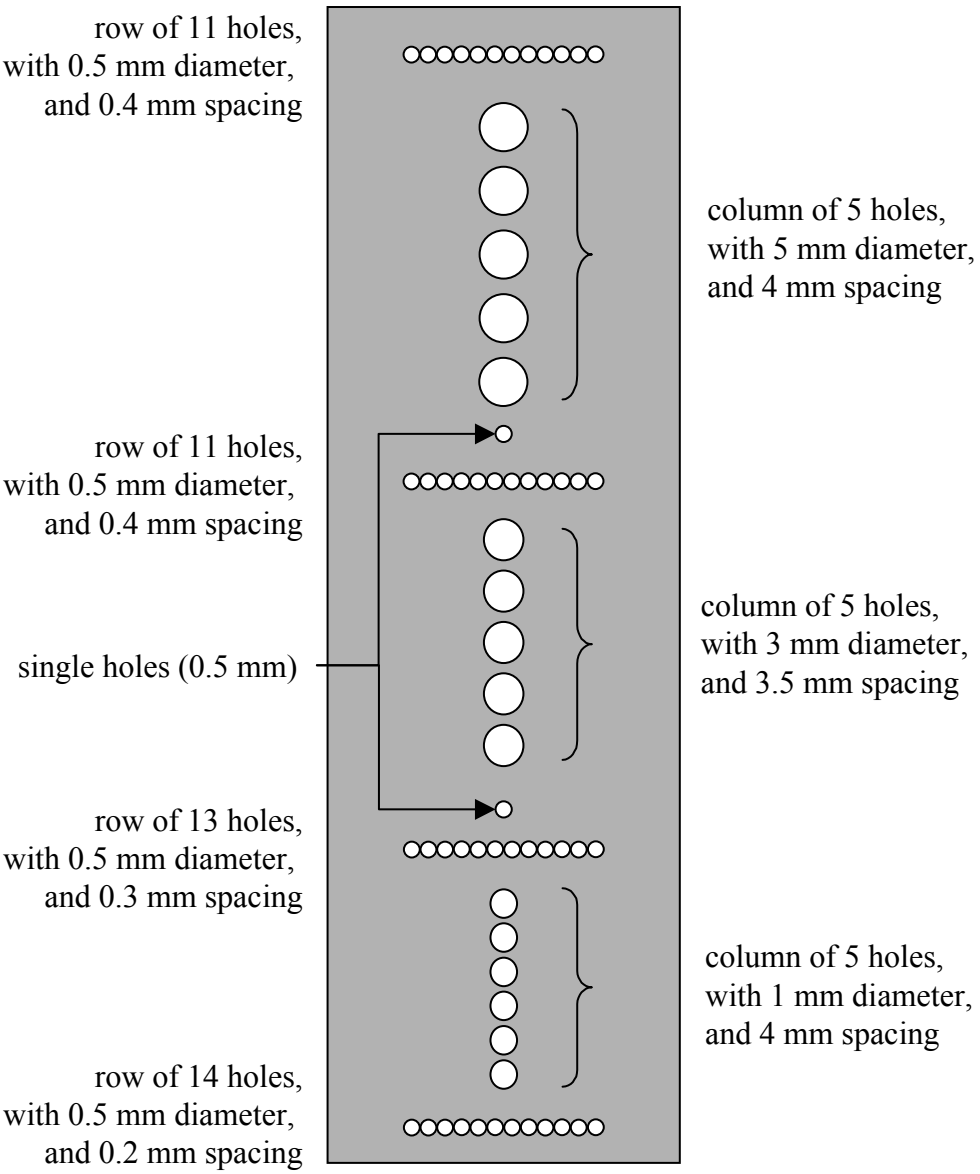


Figure 4.8: Hole-mask. The pattern on the mask consists of four rows of holes, three columns of holes and two single holes. The mask has the same dimensions as the front panel of the detector.

determines the drift time difference Δt between two holes located on the same side of the wire, it is possible to calculate the average drift time between the two holes.

$$v_{\text{drift}} = \frac{0.9\text{mm}}{\Delta t} \tag{4.12}$$

The distance between the corresponding edges of two neighbouring holes is 0.9 mm. For hole 7/8 (one finds $\Delta t = 24.0\text{ns} \Rightarrow v_{\text{drift}} = 3.8\frac{\text{cm}}{\mu\text{s}}$) and for hole 5/4 ($\Delta t = 36.1\text{ns} \Rightarrow v_{\text{drift}} = 2.5\frac{\text{cm}}{\mu\text{s}}$). Both velocities were measured in a region within 2 mm from the wire. Within the experimental error they are different and below the saturation velocity (5.0 cm/μs) given in the literature [Peisert, 1984]. The position of the centre of the hole relative to the wire is not known with the needed accuracy of a fraction of a mm.

The uncertainty on the drift velocities is large (3.8 cm/µs± 1.1 cm/µs, 2.5 cm/µs± 0.75 cm/µs). It is rather difficult to extract the drift times from the distributions because the onset is very broad, especially for the outer holes.
Other interesting information can be extracted from these measurements. For the holes 5, 6 and 7 one can clearly recognize to which synchrotron radiation bunch the avalanche belongs. For the other holes, which are closer to the edges of the detector, this is not possible. Consequently with a slit with a maximum width of 2.3 mm on the detector it is possible to decide from which bunch each photon comes even in five bunch mode of DORIS (separation of 192 ns).
The results of the drift time measurements with a pinhole at different distances from the

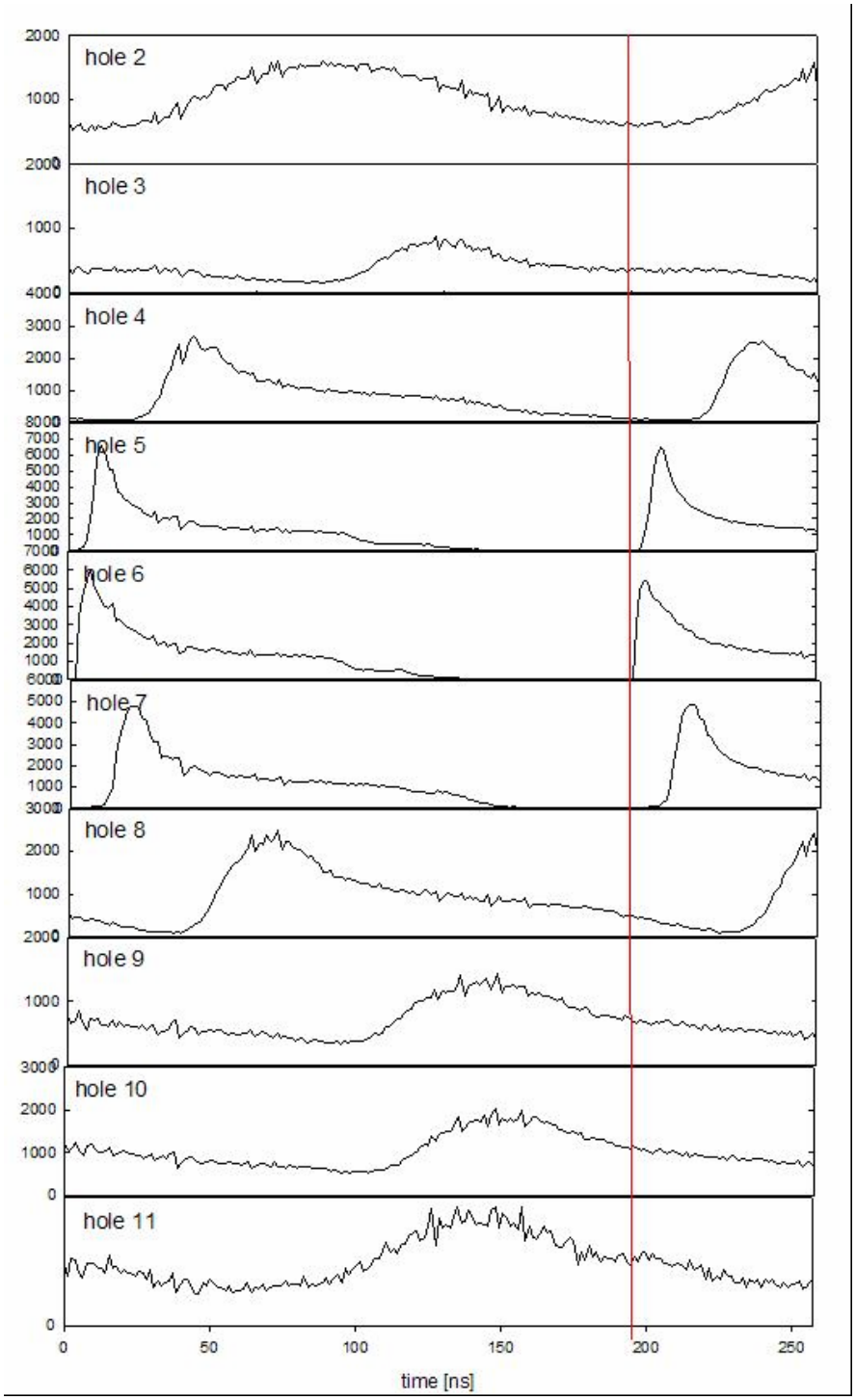


Figure 4.9: Drift time distribution in a detector irradiated through a pinhole at different positions.

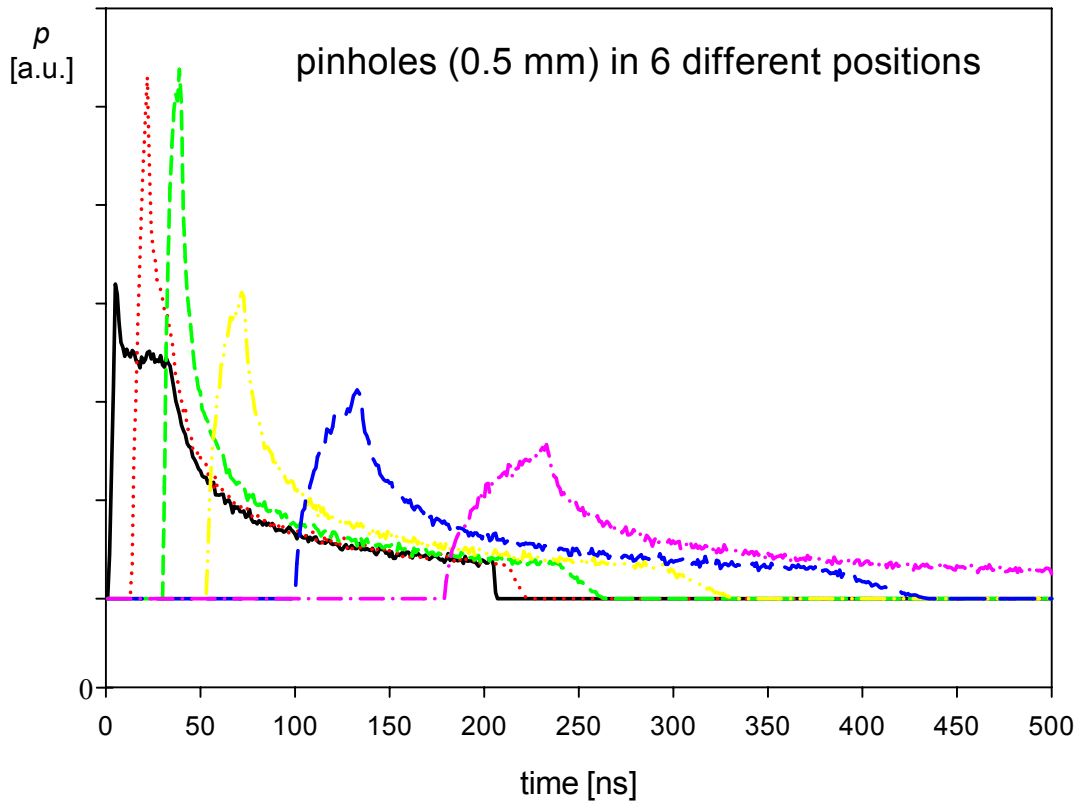


Figure 4.10: Simulation of the drift time distribution in a detector irradiated through a pinhole at different positions (probability vs. time).

wire can be reproduced qualitatively with the simulation (Figure 4.10).

4.4 Waiting time distribution

In the previous chapter we have assumed the validity of Poisson statistics. This assumption was tested by measuring the waiting time distribution (e.g. the distribution between successive events in the detector).
All calculations of double event rates are based on Poisson statistics. It was assumed that the detector behaves ideally. It is obvious that the bunch structure of the radiation and the drift time distribution of the detector influence the waiting time distribution. The time difference between two events can be measured with a start-stop TDC connecting the anode signal to both, start and stop, inputs. At the start input a short delay has to be inserted. All anode pulses thus arrive first at the stop input and then at the start input (Figure 4.11). If the pulse of the first events arrives it is ignored at the stop input since no measurement has been started before and afterwards it starts the measurement at the delayed start. The result of the measurements is thus the waiting time between the first and second events minus the delay at the start input.
The experimental waiting-time distributions between successive anode pulses made with 8 keV synchrotron radiation from the DORIS III operated in five-bunch mode and with a ^{55}Fe radioactive source indicate that the behaviour of real detectors is by no means ideal as illustrated in Figure 4.12. Although Poisson statistics predicts a somewhat counterintuitive exponentially decaying waiting time distribution

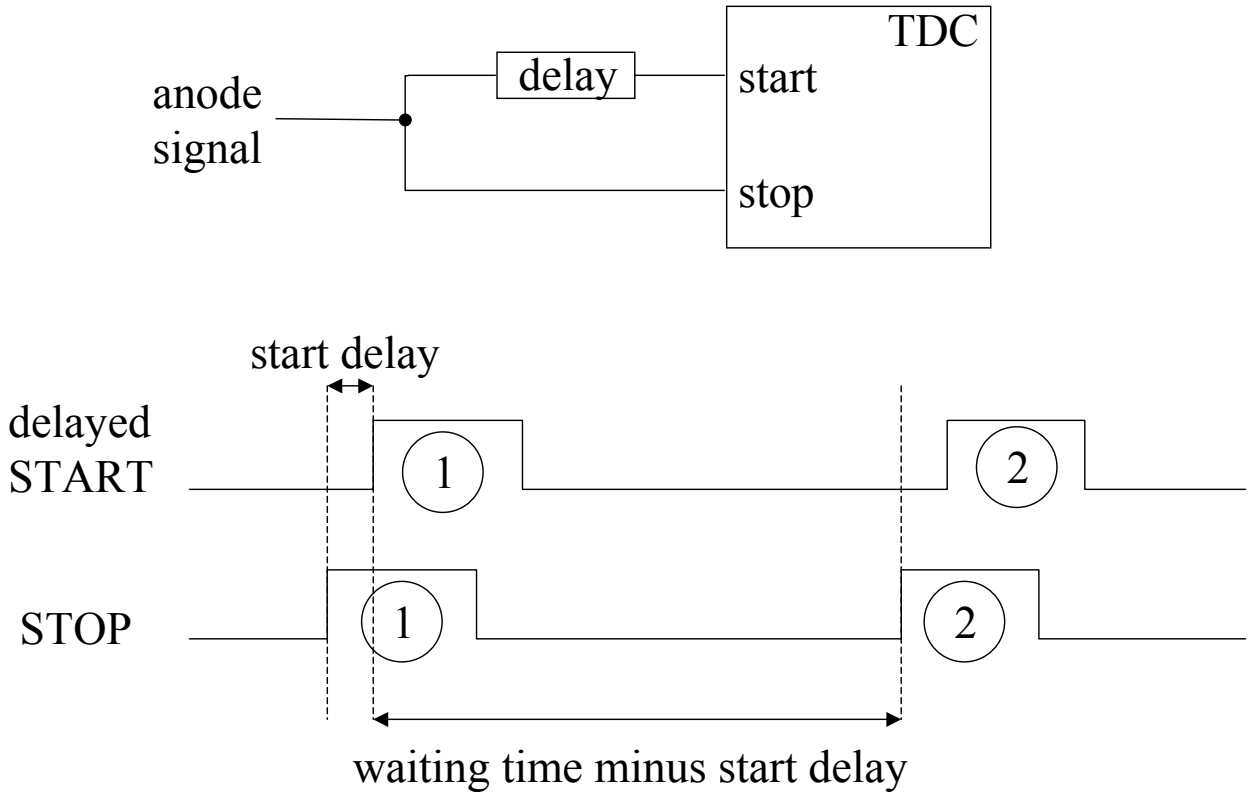


Figure 4.11: Experimental set-up for measurements of the waiting time measurement.

$$p(t) = n^2 \tau \exp(-nt) \tag{4.13}$$

the count rates during the measurements were so low ($1.1 \times 10^4\,\mathrm{s}^{-1}$) that, a flat distribution would have been expected within the time interval represented in Figure 4.12 ($n = 1.2 \times 10^4\ \mathrm{s}^{-1} \Rightarrow p(1\mu s) = n^2 \tau \exp(-n \cdot 1\mu s) = 0.988 n^2 \tau$). Clearly, the frequency of short waiting times is much higher than expected.

The temporal structure in the storage ring also influences the waiting-time distribution but the drift time of the electrons in the detector results in an apparent broadening of the well-defined bunch structure. The repeat of the observed difference in the distribution for synchrotron radiation and iron source is between 190 and 200 ns and corresponds thus to the time structure of the storage ring. The small peaks are a consequence of the sharp bunches broadened by the drift times of the electrons. Comparison with the ^{55}Fe-data indicates that the excess frequency at short waiting times is not related to the bunch structure. A more detailed analysis requires a readout system that is able to detect two events within times shorter than 300ns (i.e. the transit time of the delay line in the detector). The waiting time distribution was measured under various conditions of anode voltage, discriminator settings, and the unexpected events with short waiting times were always observed.

None of the conventional readout systems for delay lines introduced in Chapter 3 is able to detect two events in such a short time. If the signals of two events travel simultaneously on the delay line the reconstruction may fail. Time stamp TDCs (TS-TDC) with multiple-hit capability and short dead times (20 ns), which are used, for instance, in high-energy physics [Fischer, 2002] and recently for delay line readout of neutron detectors [Gebauer, 2001], make use of all signals (Figure 4.13). These devices register the time of arrival of all signals at their inputs and record their time stamps, relative to an internal ring oscillator, together with the number of the input channel. Three input channels which simultaneously

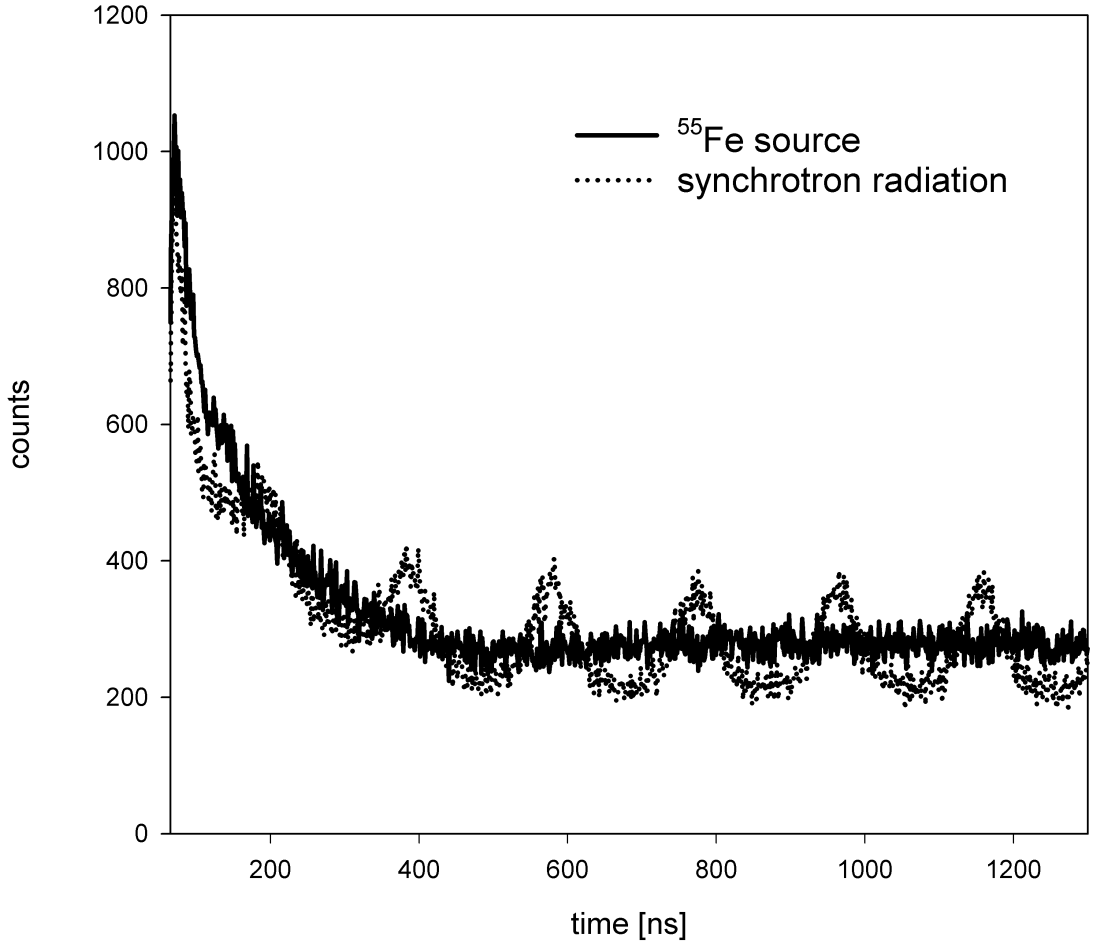


Figure 4.12: : Comparison of waiting time distributions of the anode signals measured with a ^{55}Fe source and synchrotron radiation from DORIS III operated in the 5-bunch mode.

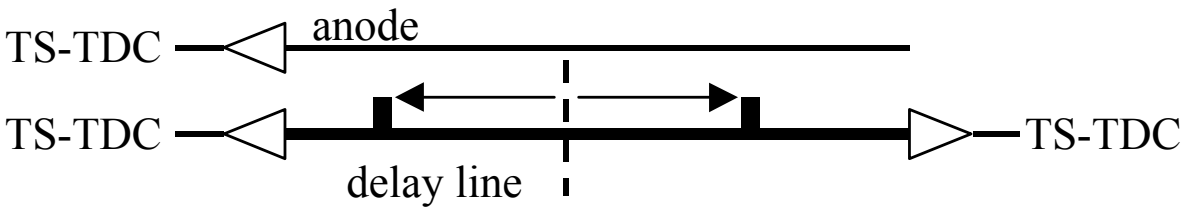


Figure 4.13: Working principle of a time-stamp TDC. The time stamps of both cathodes and the anode are constantly recorded (see also Figure 3.5).

record the time stamps for both cathode signals and the anode signal suffice for delay line readout of a linear PSGPD. Events are reconstructed offline by finding three time stamps, one at the anode (t_a) and one at each cathode (t_{c1}, t_{c2}) satisfying the check-sum in Eq. 4.14, which is valid for all events.

$$t_{c1} - t_a + t_{c2} - t_a = T \tag{4.14}$$

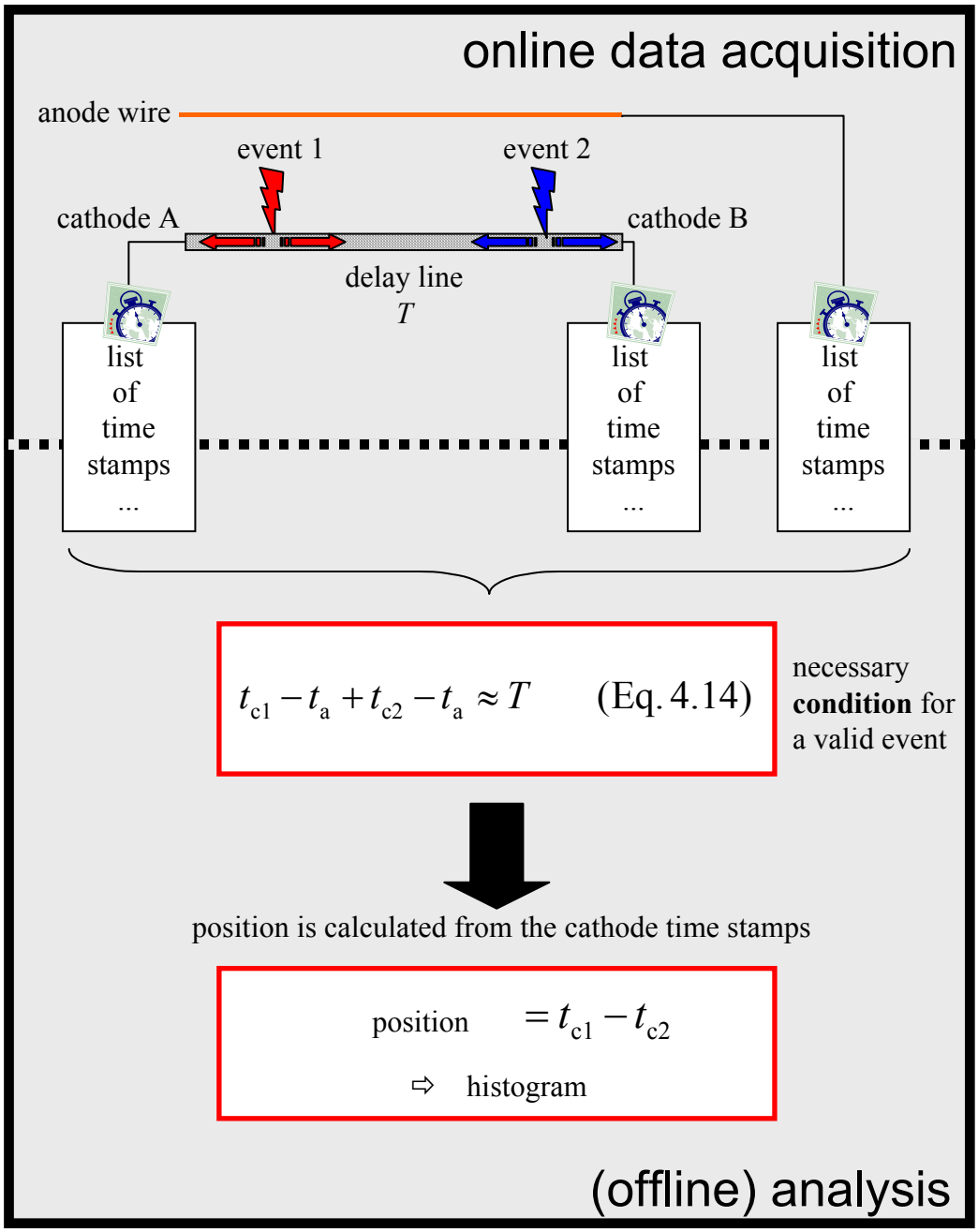


Figure 4.14: Working principle of a time-stamp delay-line readout. The TDC measures the time-stamps for each input. During analysis all time-stamps falling into a given time window relative to the anode time-stamp are checked for the condition in Eq. 4.14. The position of the corresponding event is calculated from the difference $(t_{c1} - t_{c2})$ and the histogram is incremented at the corresponding position. During analysis all triplets of time stamps (t_a, t_{c1}, t_{c2}) falling into a predetermined time window relative to the anode time stamp (t_a) satisfying the condition in Eq. 4.14 are selected.

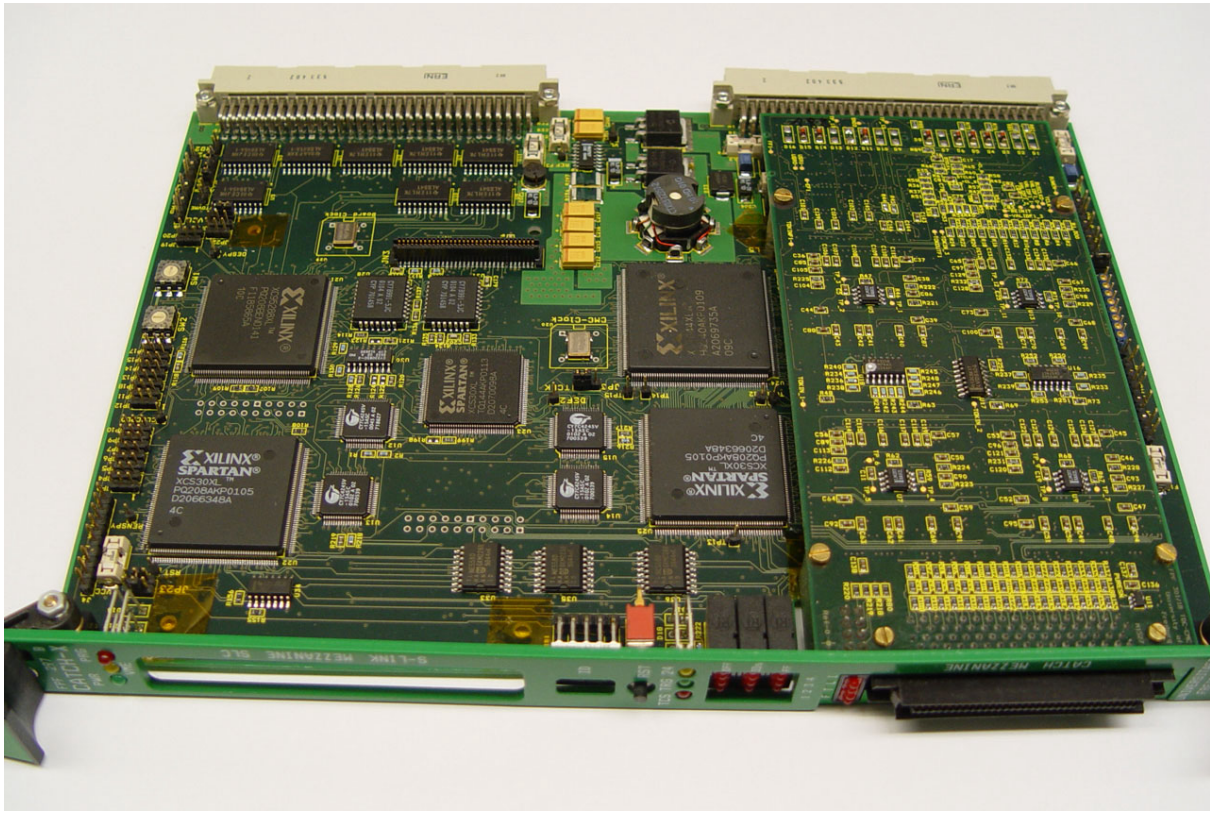


Figure 4.15: The 6U-VME TS-TDC based on the Catch-X with a F1-TDC mezzanine card (right slot).

Triplets, which do not fulfil the condition, do not belong to the same event.
With such a readout system it is also possible to reconstruct double events, which can even occur simultaneously as in the case of X-ray fluorescence on the delay line by scanning the lists for appropriate combinations of time stamps (Figure 4.14).

4.5 TS-TDC Prototype

For our measurements we used a prototype of a 6U-VME TS-TDC (Figure 4.15) based on the F1-TDC (Figure 4.16) and designed for high-energy physics (HEP) applications [Fischer, 2002] adapted to the local VME environment. The limited maximum count rate of this prototype ($\sim 1.2 \times 10^4\,s^{-1}$) is well below those in standard SAXS measurements but sufficient for the present investigation. The module only records and stores the acquired time stamps of the input signals (in this case of the anode and cathode signals) and writes them to hard disk for subsequent offline processing after receiving an external trigger signal.
Whereas in HEP a sophisticated trigger logic must be used, in the present case it suffices to connect the anode signal to the trigger input. The trigger mechanism must be bypassed for normal scattering measurements but it is useful for the investigation of double events where a simple double-pulse recognition mechanism (Figure 4.17) can be made

Figure 4.16: The F1-TDC chip.

with a fan-out, an AND gate and a short delay to produce the trigger. The readout system allows compensating for trigger latencies of up to 8 μs.
Figure 4.17 gives a schematic of the electronic set-up. Prior to the start of a measurement, three single pulses must be transmitted to three initialisation inputs to initialise the TDC. For this purpose a single pulse generator was designed and built (Figure 4.18). Throughout a measurement the TDC records time stamps but these are not acquired, as long the readout of the TDC is not released by the trigger signal. The simplest way would be to use one of the three detector signals as trigger signal. However, as we are particularly interested in double events (two events within a time equal to the length of the delay line) it was decided to implement a trigger circuit based on an AND gate, which can act as a double event filter. Double events are rare and at count rates below 10 kcps the fraction of double events with short waiting times (<300 ns) is smaller than 0.5%. A readout mode that exclusively acquires double events is very useful to rapidly accumulate a significant amount of data for analysis.
A discriminator is used as fan-out for one of the cathode signals. Cathode signals are

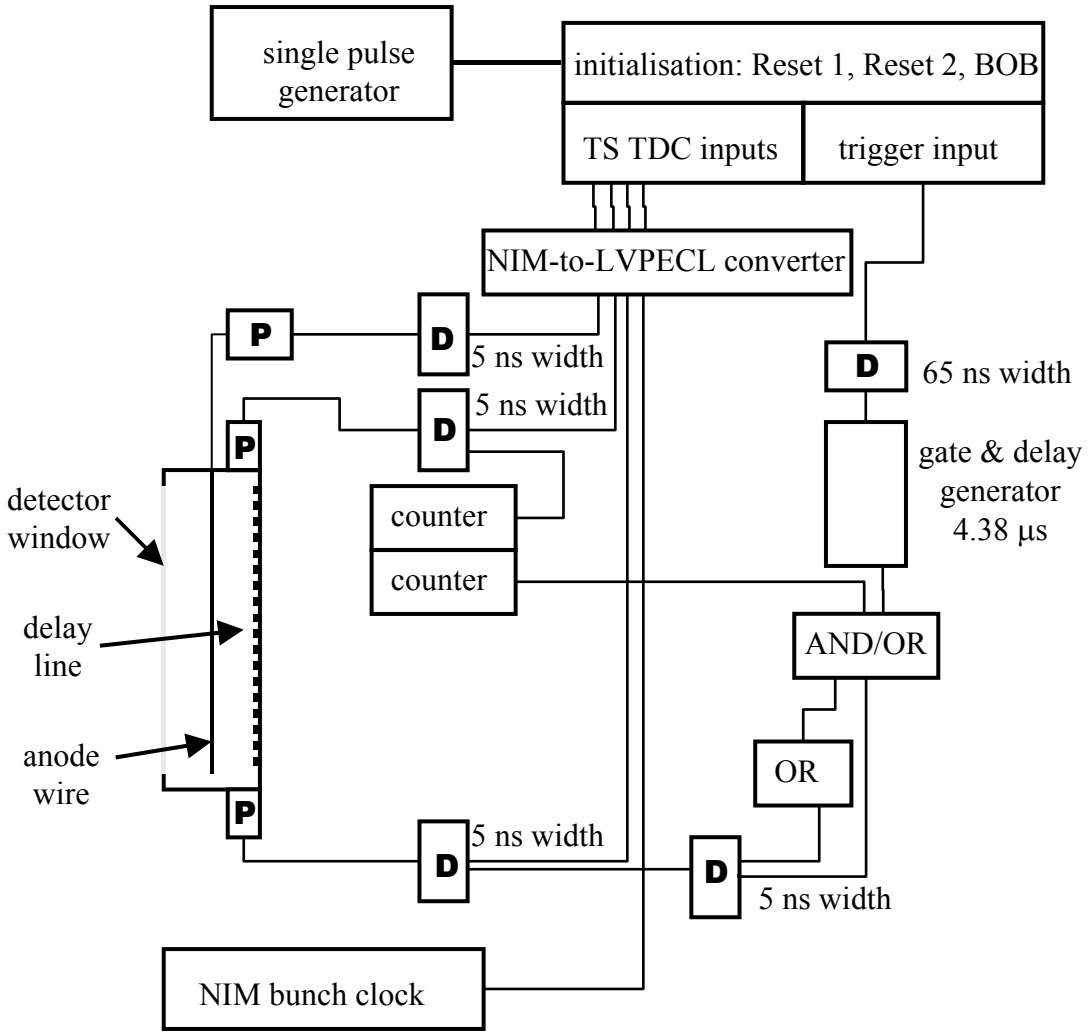


Figure 4.17: Schematic of the readout electronics for the time stamp TDC measurements. Preamplifiers (P) and discriminators (D) are connected to the detector in order to convert the detector pulses to NIM. The NIM signals are transmitted to the TDC and used to generate the trigger signal. The AND/OR symbol indicates a logic gate that can be an AND gate (double pulse filtering) or an OR gate (normal readout mode). The single pulse generator is used to initialise the TDC.

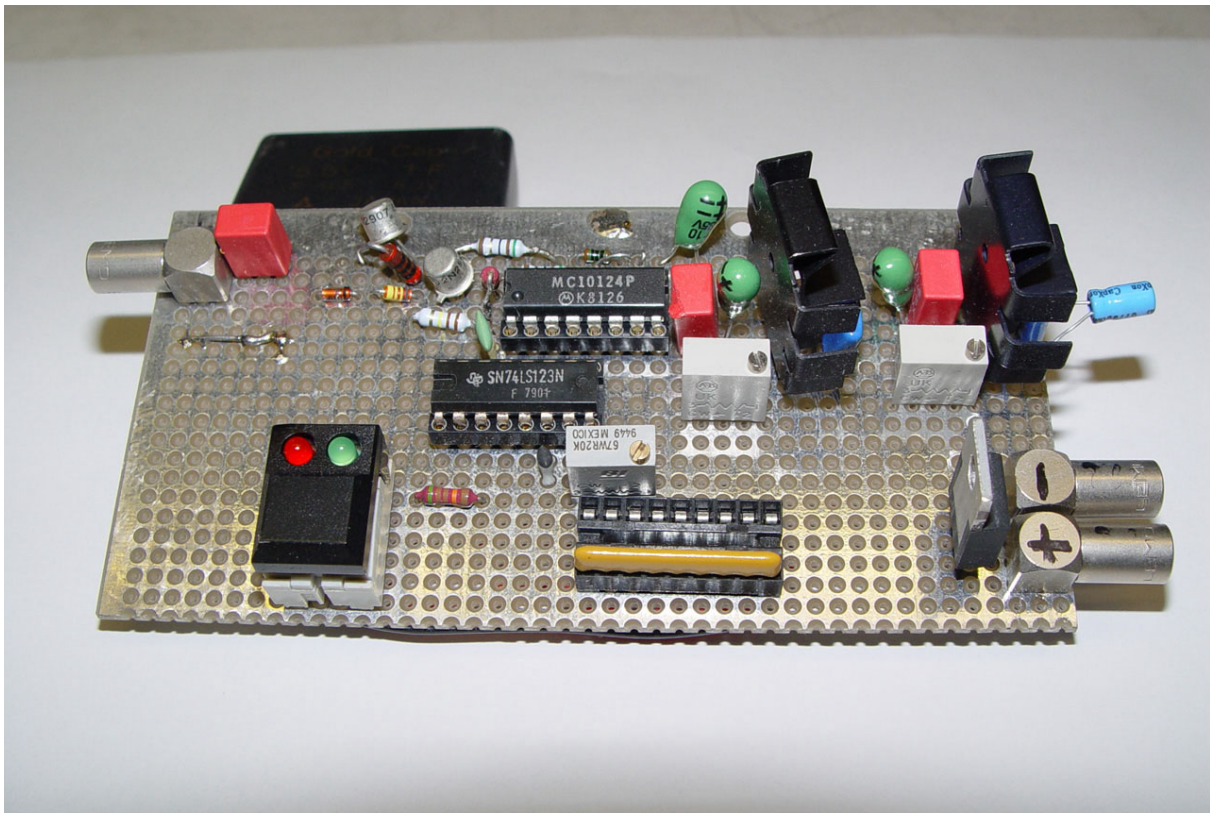


Figure 4.18: Single pulse generator. The inputs for the operating voltages on the right are labelled "+" and "-". If one presses the button a NIM pulse is generated at the output on the left.

preferable because they occur less often than anode pulses since they are on average smaller than the anode signals and thus more likely eliminated by the discriminators (threshold 25 mV). One of the discriminator output signals is converted to a 300 ns wide pulse (by the OR gate). The second discriminator output and the expanded signal are connected to the inputs of the AND gate. Because of the intrinsic delay of the OR gate the discriminator pulse and the corresponding long pulse cannot trigger the AND gate but if a second cathode pulse follows within 300 ns the AND gate gives a pulse. This pulse propagates through the gate and delay generator (GADG). Its purpose is to ensure that only one trigger signal is generated within the inherent delay of the GADG (4.38 µs). During this time the GADG does not accept further input signals. This also ensures that events are read out only once (the width of the trigger window is 4 µs). The latency of the trigger signal due to the delay of the GADG is compensated by the TDC. The F1 TDC has a build-in compensation mechanism for the trigger latency consisting of a memory that can store the time stamps for up to 4 µs.
In addition to the three detector signals the bunch clock is also connected to an input of the TDC. The distribution of time differences between the arrival of the bunch and the anode signal can be determined using the bunch clock time stamps.

4.6 First tests and the check-sum distribution

The first tests aimed at verifying that the TS-TDC can correctly acquire the SAXS pattern of a reference sample, in this case turkey tendon collagen, by comparison with the standard system based on a LeCroy 4201 TDC [Boulin, 1988]. The results in Figure 4.19 indicate that this is indeed the case, the small difference in background being due to slight changes in the geometry between measurements. Some specific features of the check-sum (Equation 4.14) distribution corresponding to such data may depend on the detector used,

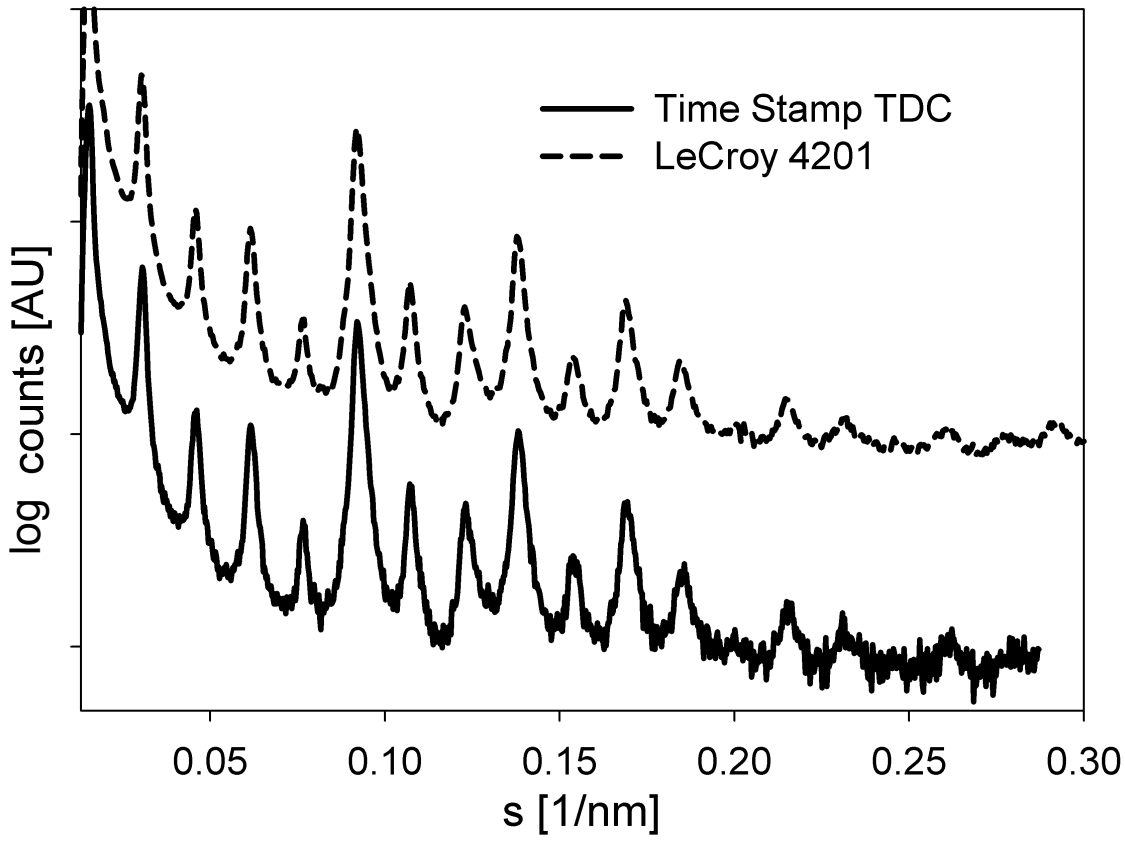


Figure 4.19: Scattering pattern of collagen acquired with a LeCroy 4201 TDC and the TS-TDC. The duration of the measurements was different. The curve for the LeCroy TDC consists of 4.0 x 10^6 events while the curve representing the time stamp TDC has only 6.7 x 10^5 events. The curves were scaled to improve comparability. The small differences are solely due to a slight change of geometry between the two measurements.

but the width of the distribution is always around 10-20 ns, which is much larger than the expected (~1 ns). The check-sum distribution depends only weakly on parameters like count rate, scattering pattern or position along the wire. It does, however, depend on the distance between the initial ionisation and the wire as indicated by a series of measurements with the mask, described earlier, using the row of eleven holes (0.4 mm spacing). During each measurement the mean distance between the primary ionisation and the anode wire was varied by leaving only one of the holes open allowing only a narrow beam to penetrate the detector at a well-defined distance from the anode wire. The check-

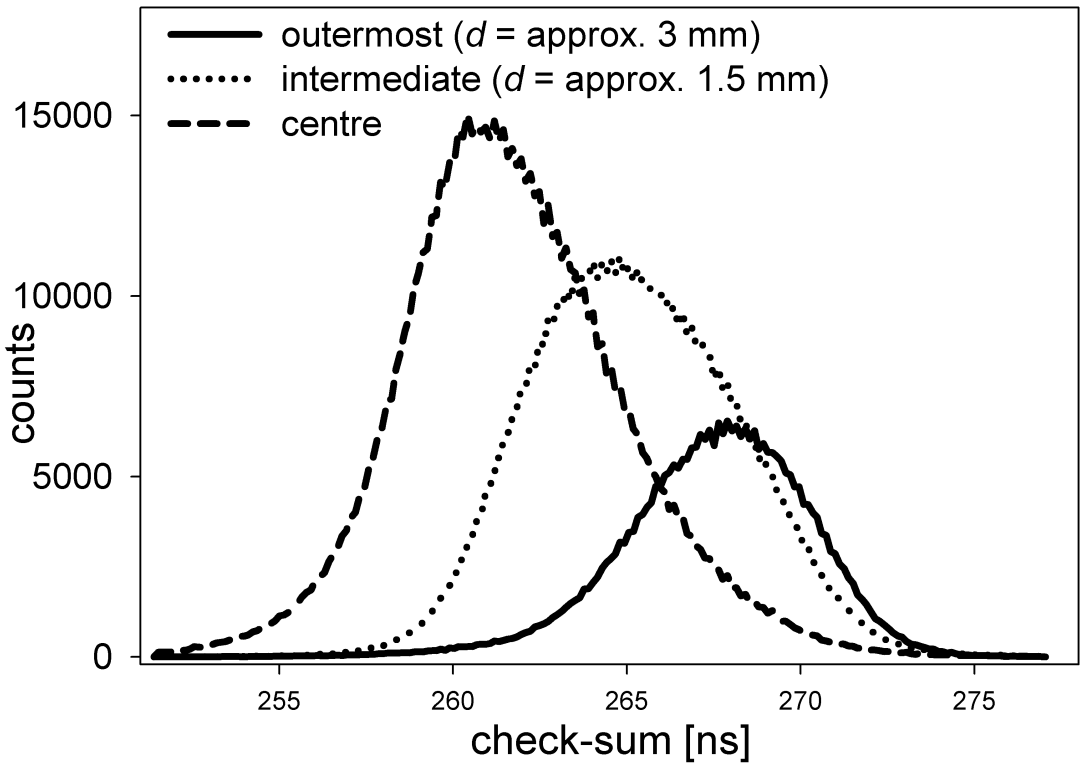


Figure 4.20: : Check-sum distributions of single events for a narrow beam. The distance between beam and wire was varied between 0 and 3 mm.

sum is broader when the ionisation is close to the wire than it is for the outer holes (Figure 4.20).
The shift of the distribution is in all likelihood due to the dependence of the pulse shape (rise time) on the drift times, which in turn depend on the distance of the ionisation site from the centre of the detector, as observed by Bednarek [Bednarek, 1984]. This effect is more pronounced when the entrance slit is further away from the centre of the detector where the field is weaker.
The width of the distribution determines the theoretical limitations of the analysis of double-events. In practice, however, no complete double events were observed for which the assignment of the signals to the events was ambiguous. Other parameters like the length of the discriminator pulses seem to be more important. Incomplete double events (missing anode or cathode signals) were also observed.

4.7 Double events and fluorescence

The aluminum hole mask shown in Figure 4.8 was fixed to the front panel of the detector and two sets of measurements (9×10^5 events each) were made, the first one recording single events only and the second one double events only.
For double events one has to differentiate between primary and secondary events. The first step to uncover the origin of the additional counts (Figure 4.12) is to analyse the waiting-time distribution of complete events (Figure 4.21). For a ^{55}Fe radioactive source the events are equally distributed over time and Poisson statistics are applicable. In the waiting-time distribution for complete events from the radioactive source, the frequency of short waiting

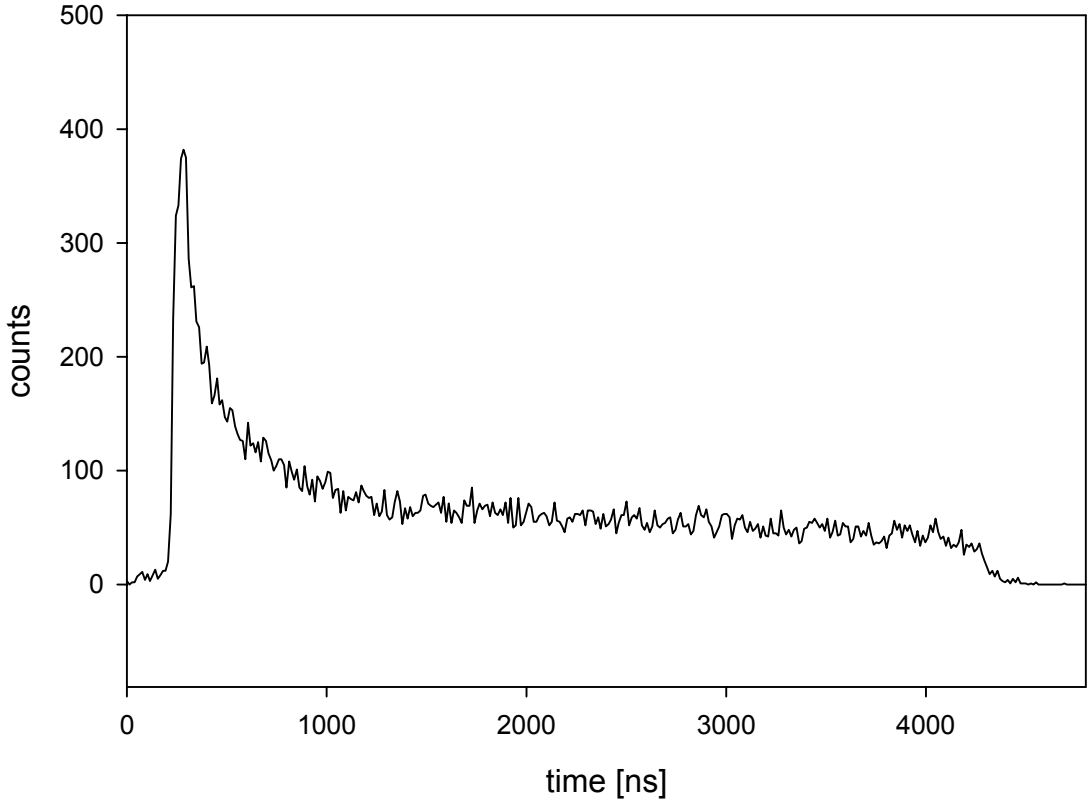


Figure 4.21: Waiting time distribution of double events. During the first 100 ns there are no counts due to the dead time of the detector system. The long dead time is probably due to the duration of the output pulses of the discriminators. After 4 μs there are no entries in the histogram because the time stamp has a 4 μs trigger window.

times is much higher than expected, as already observed for the anode signals only. The deviations from Poisson statistics are thus not an artefact (e.g. noise on the anode channel).
Comparison of the mask pattern created by the single events and by the double events gives further information about the character of the secondary events (Figure 4.22). The patterns of primary events and single events are identical but the distribution of secondary events around the holes is broader. Double-event measurements at different count rates indicate that the pattern is more smeared at low count rates (Figure 4.23).
These observations can be explained by the fluorescence of argon. The phenomenon was previously observed in measurements of the charge distribution of avalanches in gas proportional detectors [Ito, 1996] and made responsible for the existence of exponential tails in the point-spread function of position sensitive detectors [Ne, 1997].
After ionisation of an argon atom by an incident 8 keV photon there are two possibilities for the decay of the K-shell hole (Figure 4.24): the Auger effect (89%) and X-ray fluorescence (11%). Emission of an Auger electron leads to ionisation in the close vicinity of the region where the initial photoelectron and ion were formed. In contrast, fluorescence photons (3 keV) can lead to the creation of free electrons by ionisation, which may give rise to a second avalanche and thus a double event. There is a small delay between the anode signals corresponding to the initial and secondary avalanches depending on the distance between the regions where the two ionisations take place and the anode wire. Fluorescence thus explains the higher than expected number of events with short waiting times as well as the smearing of the scattering pattern of double events. This smearing depends on the count rate because there are also genuine double events – caused by two

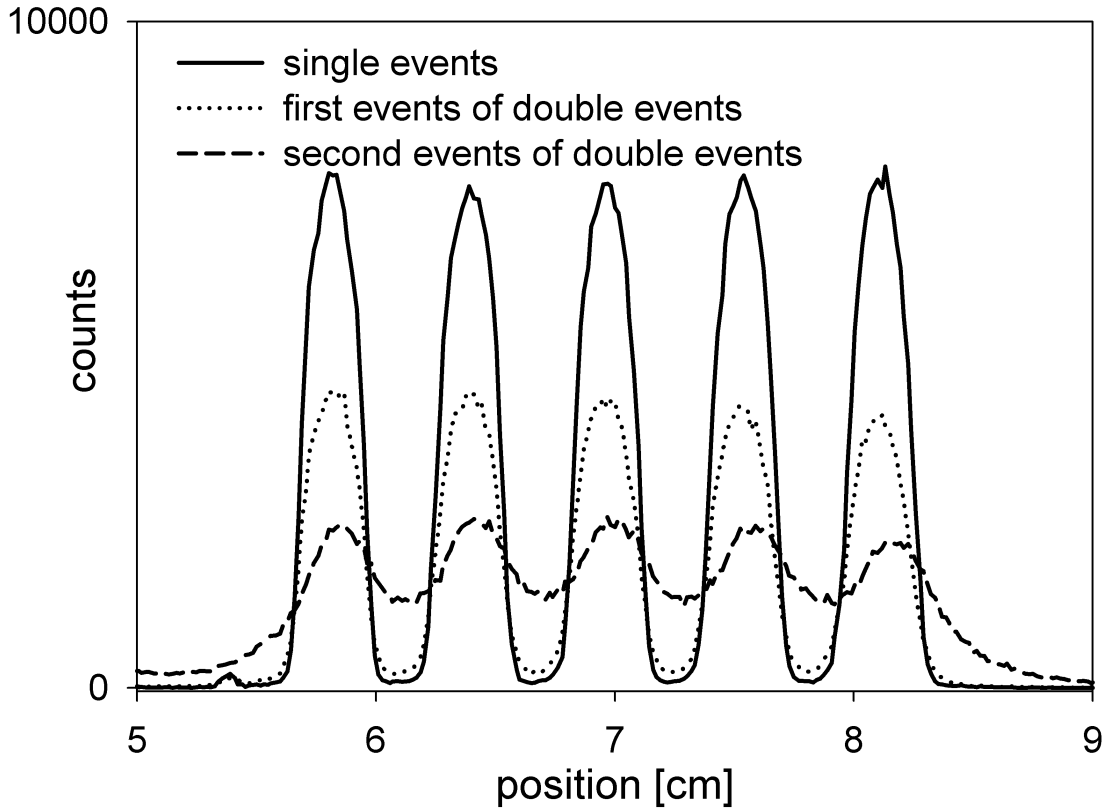


Figure 4.22: Pattern of the mask acquired with the time stamp TDC. The solid curve is the histogram of single events. The dotted curve represents primary events, the dashed one the secondary events. Note that the localisation of the secondary events is poorer than that of single and primary events.

incident 8keV-photons from the storage ring hitting the detector at the same time. The apparent improvement in resolution at higher count rates reflects the change in ratio between genuine and fluorescence-caused double events. Whereas the number of fluorescence-caused double events is always proportional to the total number of events the number of genuine double events (with a waiting time smaller than 300 ns) increases proportionally to the square of the count rate according to Poisson statistics.
In Figure 4.22 and 4.23 one does not strictly distinguish between initial and fluorescence events, only between the order of occurrence. With the present arrangement the second event in a double event will, however, in most cases correspond to the fluorescence event. As the hole mask is placed in the centre of the detector this is also where the initial ionisations take place. As fluorescence photons are emitted in all directions most of them are absorbed in the outer regions of the detector where the drift time increases rapidly. Only the small fraction that is emitted in the forward direction is detected earlier than the initial ionisation and this also explains why the pattern of primary events is slightly worse than that of single events

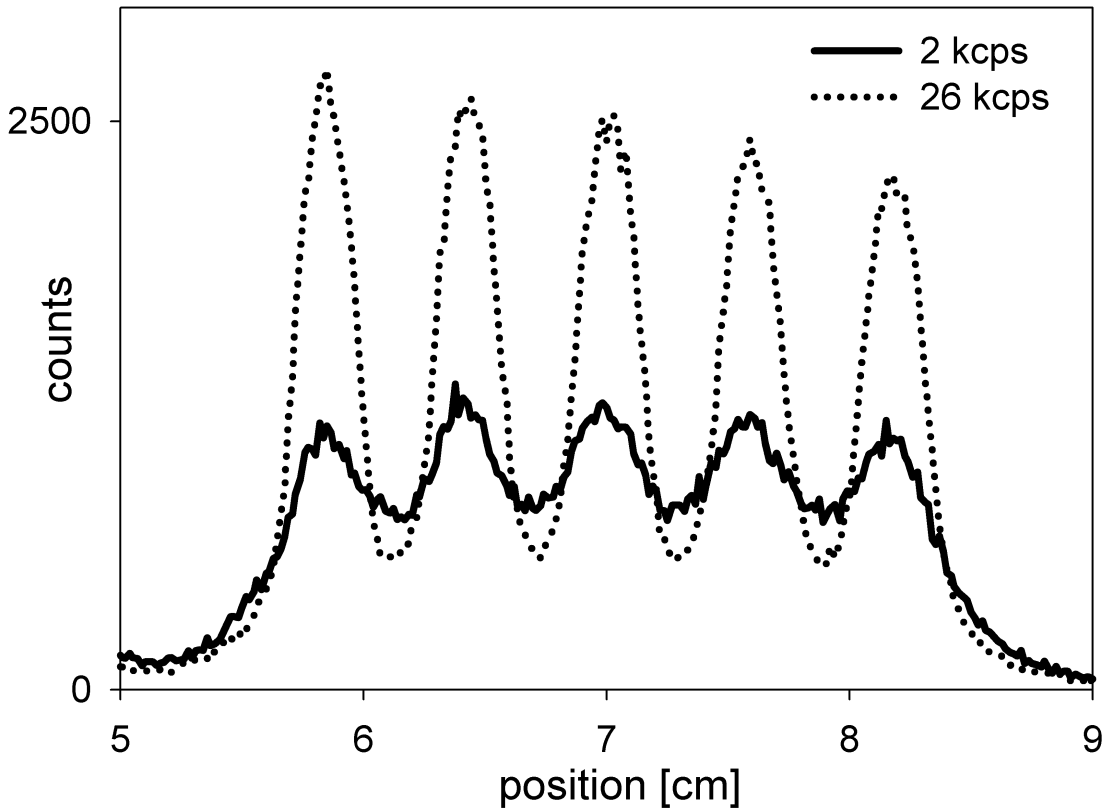


Figure 4.23: Pattern obtained with the secondary-events at different count rates. The localization of secondary events in double events improves for higher count rates. This is due to the larger fraction of real double events (two x-ray photons in the detector).

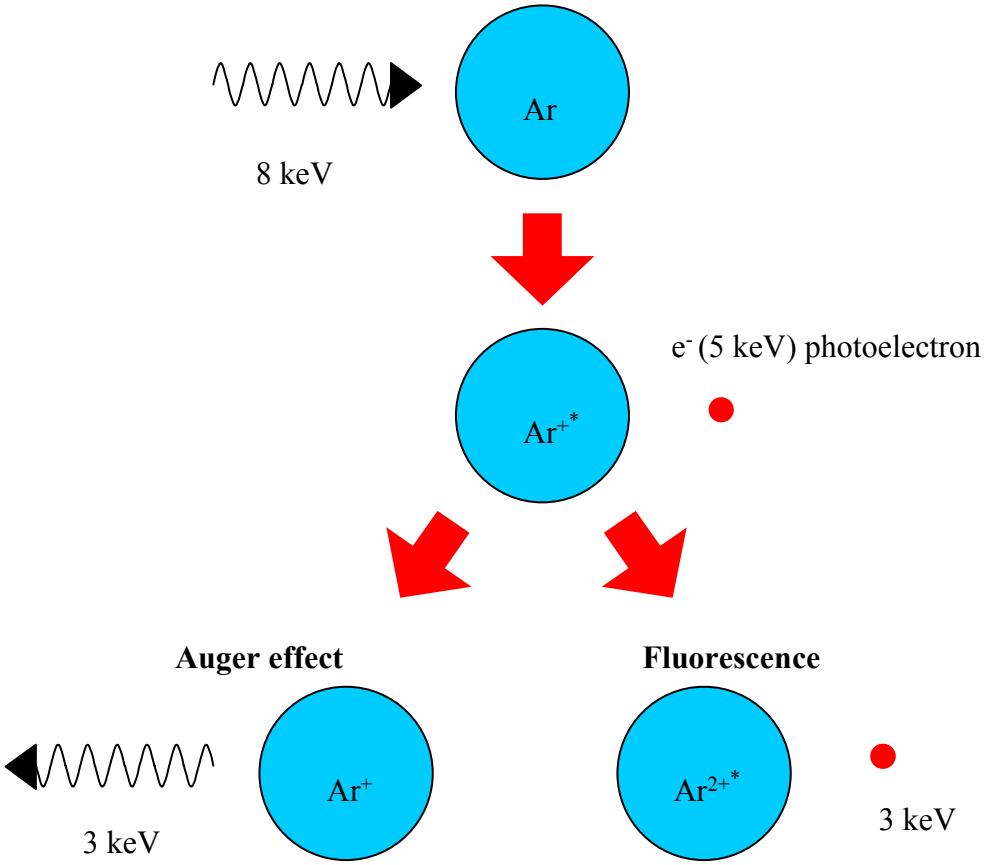


Figure 4.24: Auger effect and fluorescence are two competing modes of decay of the excited argon state that was created by the incident synchrotron radiation photon.

4.8 Point-spread function (PSF)

In SAXS experiments the scattered intensity decays over several orders of magnitude and smearing of the patterns resulting e.g. from exponential tails would reduce the accuracy of the data in PSGPD [Ne, 1997]. Similar effects have also been observed for CCD detectors with image intensifier [Pontoni, 2002]. To measure the PSF (i.e. the response to a point-like beam) of the linear detector an aluminium mask (Figure 4.8) with a 0.5 mm hole was fixed to the detector so as to irradiate the region of the anode wire. At the low count rates

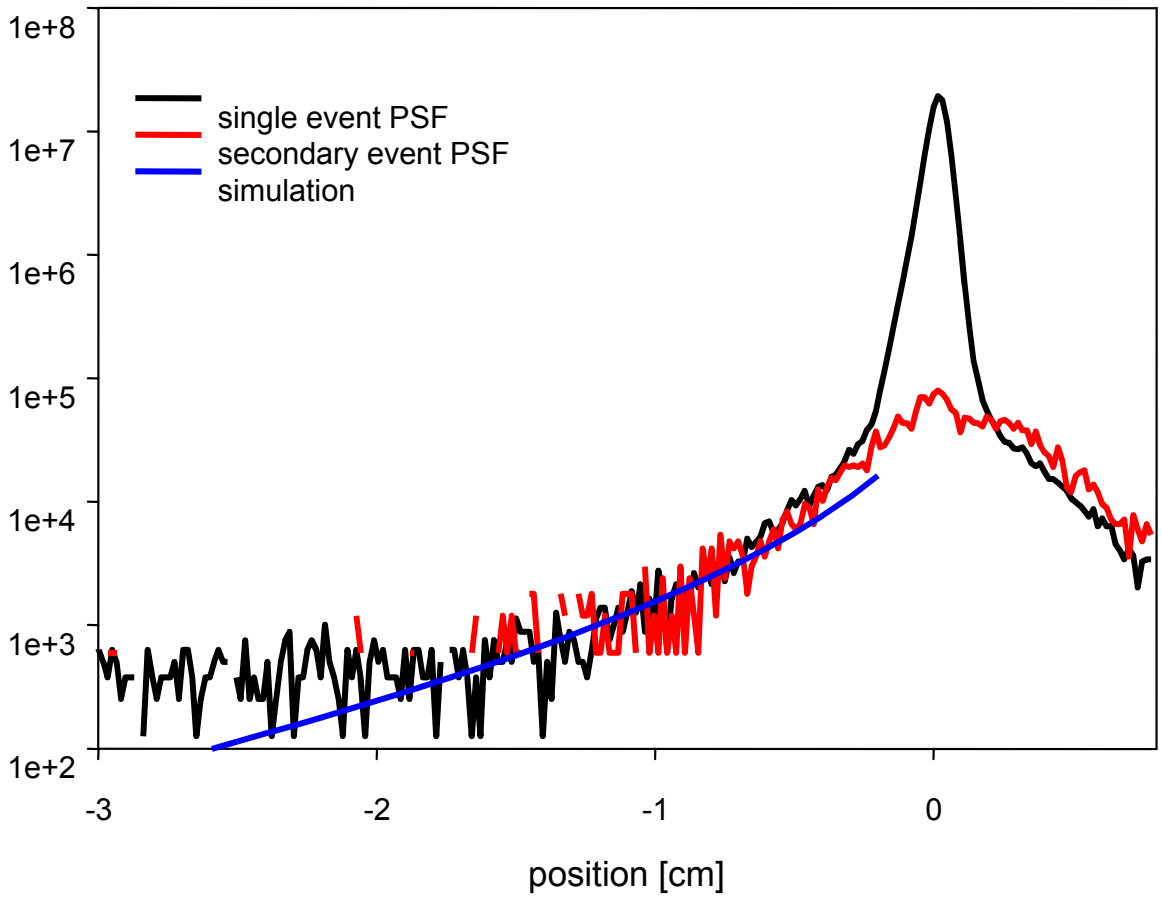


Figure 4.25: Comparison of single events, secondary events (fluorescence) and the simulation of fluorescence. The numerical calculation nicely reproduces the tails of the distributions.

used (less than 300 s^{-1}) only 65 genuine double events should occur per hour. The actual rate of observed double events was 15 double events per second (i.e. 54660/hr or 5% of the single events). This is consistent with the 11%-probability for fluorescence since not all fluorescence photons give rise to an avalanche or distinct events. The primary events give rise to a well-defined Gaussian distribution whereas the distribution of secondary events is broader and decreases exponentially as a function of the distance from the excitation region (Figure 4.25). Numerical calculations of the fluorescence can reproduce the exponential decrease. For the calculations it is assumed that a narrow beam gives rise to fluorescence along its path (Figure 4.26) whereas the fluorescence decreases according to

the Beer-Lambert's law for the fraction of the number of photons transmitted through a thickness d of an absorbing medium (calculation results in Figure 4.25).

$$N(d) = N_0 \exp\left(-\frac{d}{P(E)}\right). \tag{4.15}$$

N_0 is the initial number of photons, $P(E)$ is the inverse of the absorption coefficient μ, which is a function of photon energy in a given material. For the incident synchrotron radiation it was calculated to be P(8 keV)=0.32 cm, for argon fluorescence it is P(2.9 keV)=1.31 cm.

Pinhole measurements with different distances perpendicular to the wire provide further evidence for the hypothesis that fluorescence causes double events. Whereas for measurements with the hole in a central position the primary events are the genuine synchrotron radiation photons, for the outer holes it is more likely that the fluorescence photon is detected earlier than the original synchrotron radiation, which always occurs in the outer region. The pattern of secondary events is thus better localized than that of primary ones because the secondary events are most of the time due to synchrotron

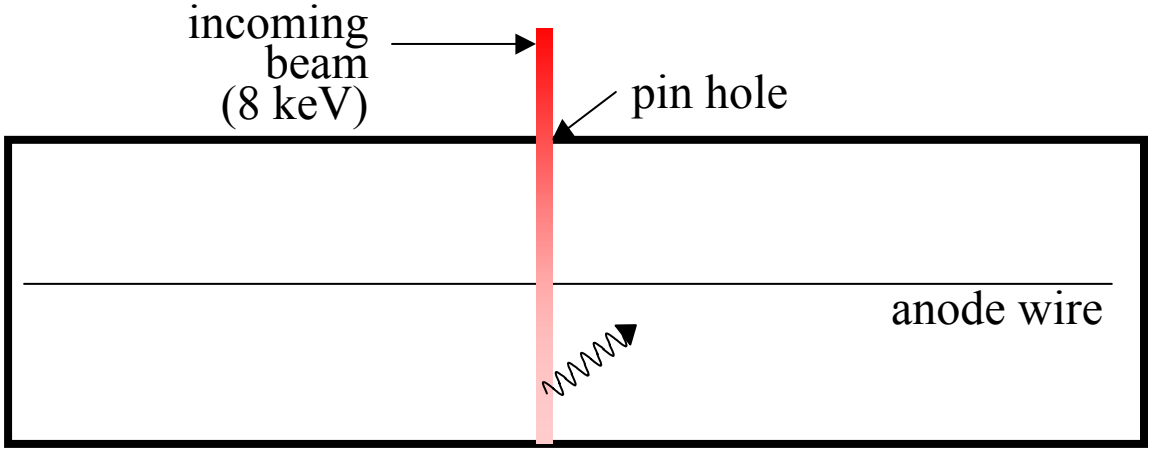


Figure 4.26: Model for the numerical calculation. A narrow beam gives rise to fluorescence. The gas absorbs direct beam and fluorescence photons.

radiation (Figure 4.27).

The PSF for the single events has exponential tails as reported in the literature [Ne, 1997]. The shape of the tails corresponds to the distribution of secondary events. The distribution of single events is also affected by fluorescence (Figure 4.25), which could cause double events that are so close to each other as to be incorrectly localized. Another explanation is that in some cases the photoelectron-induced events do not give rise to an avalanche and only the fluorescence events are detected. It is also possible that the fluorescence photon escapes to a low field region. Due to the long time needed for the electrons to drift to the wire the event is no longer recognized as a double event. Measurements in two bunch mode (500 ns separation between bunches) have shown that the drift time in the detector can be up to 500 ns.

The measurements raise the question as to whether it is possible to use the options provided by the TS-TDC to improve the PSF of the detector. Below we use two parameters: the check-sum and the drift time from the ionisation spot to the wire, which is determined from the delay between the arrival of the bunch and the anode signal. The time-stamp TDC prototype has 16 input channels and can thus also measure the time stamps for

the bunch clock signal. Consequently, it can measure the drift time and position of the events simultaneously and therefore it facilitates a correlation of both parameters. The conventional start-stop TDC could only measure drift time or position.
As discussed above, the minimum separation between successive bunches in the standard mode is 192 ns. If charges are created at the edges of the active detector volume drift times are longer than 200 ns making it impossible to discriminate, for instance, measurements of

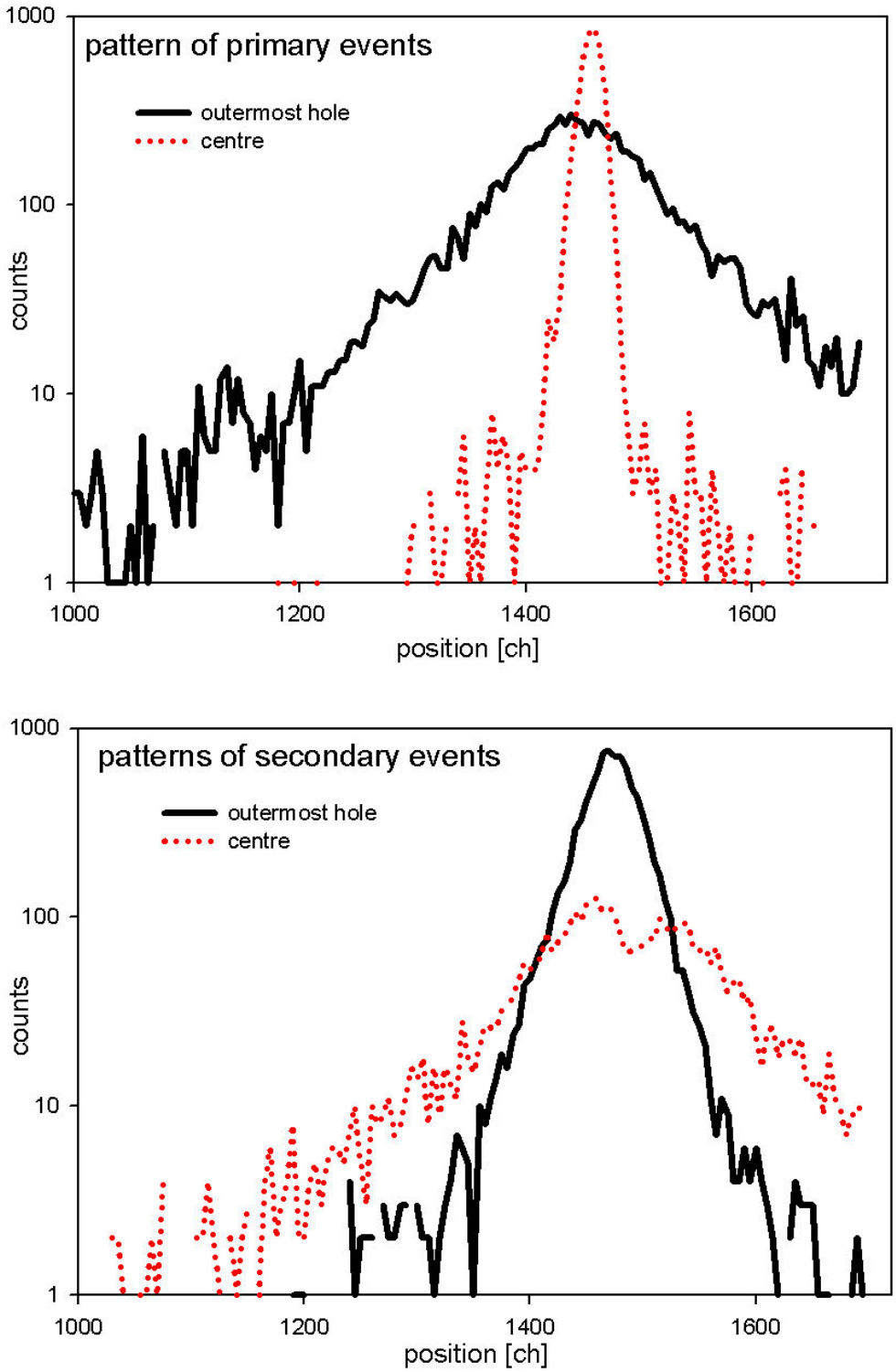


Figure 4.27: Localisation of primary events is getting worse if the ionisation takes places at the outer regions of the detector (outermost hole), in contrast the secondary events are better localised. The positions correspond to electronic channels of the TDC.

50 ns against measurements of 246 ns. The narrow beam in the centre of the detector ensures that all drift times are shorter than 200 ns and that the measurement becomes unambiguous.
In order to judge the usefulness of the parameters the check-sum was plotted against the drift time for different classes of events (Figure 4.28). The events were classified according to whether they belong to the peak region or to the tails of the PSF. The peak events are

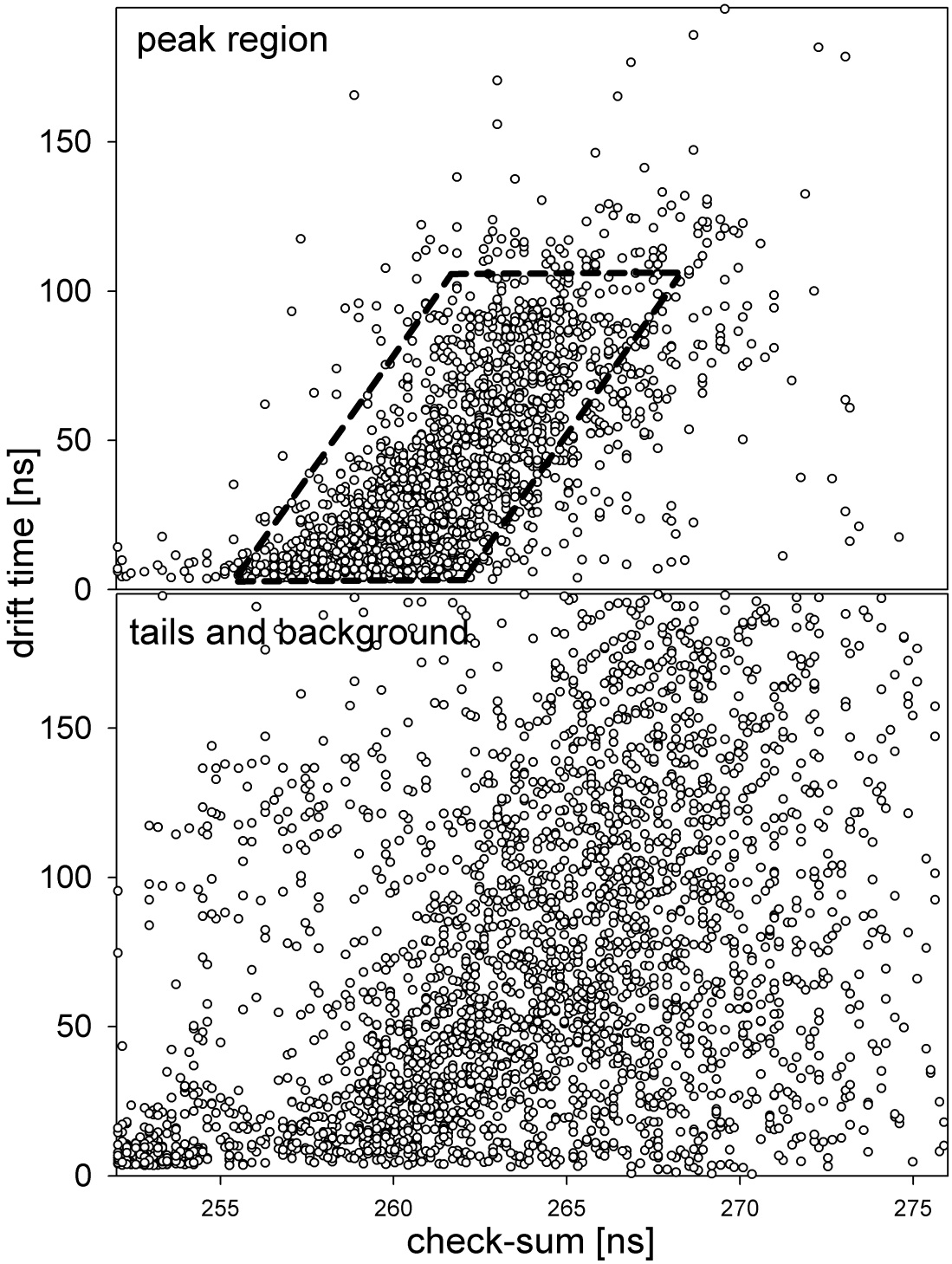


Figure 4.28: Events from the peak region and from the tails of the PSF are distributed differently if their drift time is plotted against the check-sum.

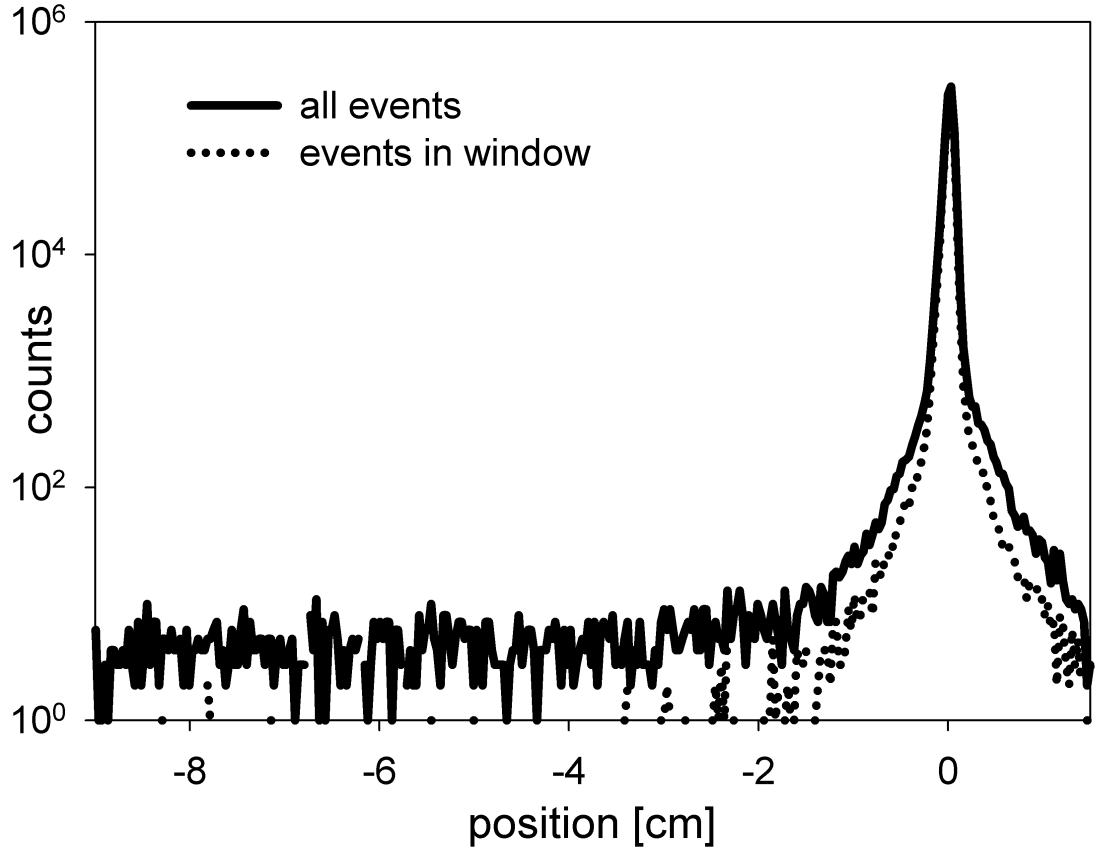


Figure 4.29: The PSF of a PSGPD can be improved if only those events are analysed that fulfil a defined relation between drift time and check-sum.

concentrated in the indicated region. Check-sum and drift time are correlated and the check-sum increases with the drift time. This is equally true for the tail events but the concentration is less pronounced.
The comparison of the PSF of all events and of these in the selected region is illustrated in Figure 4.29. The pattern of selected events contains only about 80% of the total number of events because the rest was outside the indicated check-sum drift-time region. The intensity of the tails decreases. In the pattern of all events the total number of tail events is 1.2% of the peak events; in the pattern of the selected events the tails make up only 0.6% of the Gaussian peaks. For the background the method works even better: 90% of the background noise can be rejected.

4.9 Summary

The prototype of the TS-TDC based readout electronics has been shown to be useful for detection and analysis of double events during SAXS measurements. The width of the check-sum distribution and the fluorescence of the gas clearly limit the spatial and time resolution in gas detection. The width of the check-sum, which is much larger than expected, remains to be explained. The present observations confirm the influence of argon fluorescence on the PSF and demonstrate that a rejection mechanism that analyses single events only cannot entirely eliminate these effects. The signal-to-noise ratio can, however, be improved by taking the drift time of the electrons and the checksum of the signals into account. Drift time measurements are made possible by the unique ability of the TS-TDC

to correlate the observation of scattered photons with individual radiation bunches in the storage ring. This feature could also facilitate time-resolved measurements with nanosecond-resolution.

5 Introduction: Interactions between biological macromolecules

5.1 Motivation

Many properties of biochemical systems depend on the interactions between biological macromolecules. Processes, in which interactions between proteins play a role, are for instance assembly and non-specific aggregation and more practically their solubility and crystallisation (Figure 5.1). At this stage, the physical interactions of proteins in aqueous solution can at best be explained only semi-quantitatively [Tardieu, 1999]. Although protein crystallography has become a standard method to determine the structure of proteins, it remains so far a largely empirical matter to find proper crystallisation conditions, even if some general rules have been established [McPherson, 1999]. A better understanding of the fundamental nature of interactions may contribute to turn the art of crystallisation into a reliable technique.
It would be useful to have better control over (non-specific) aggregation. This phenomenon plays an important role in several pathological conditions, such as cataract, sickle cell anaemia or Alzheimer's disease. In some cases aggregation can be put to beneficial use. A

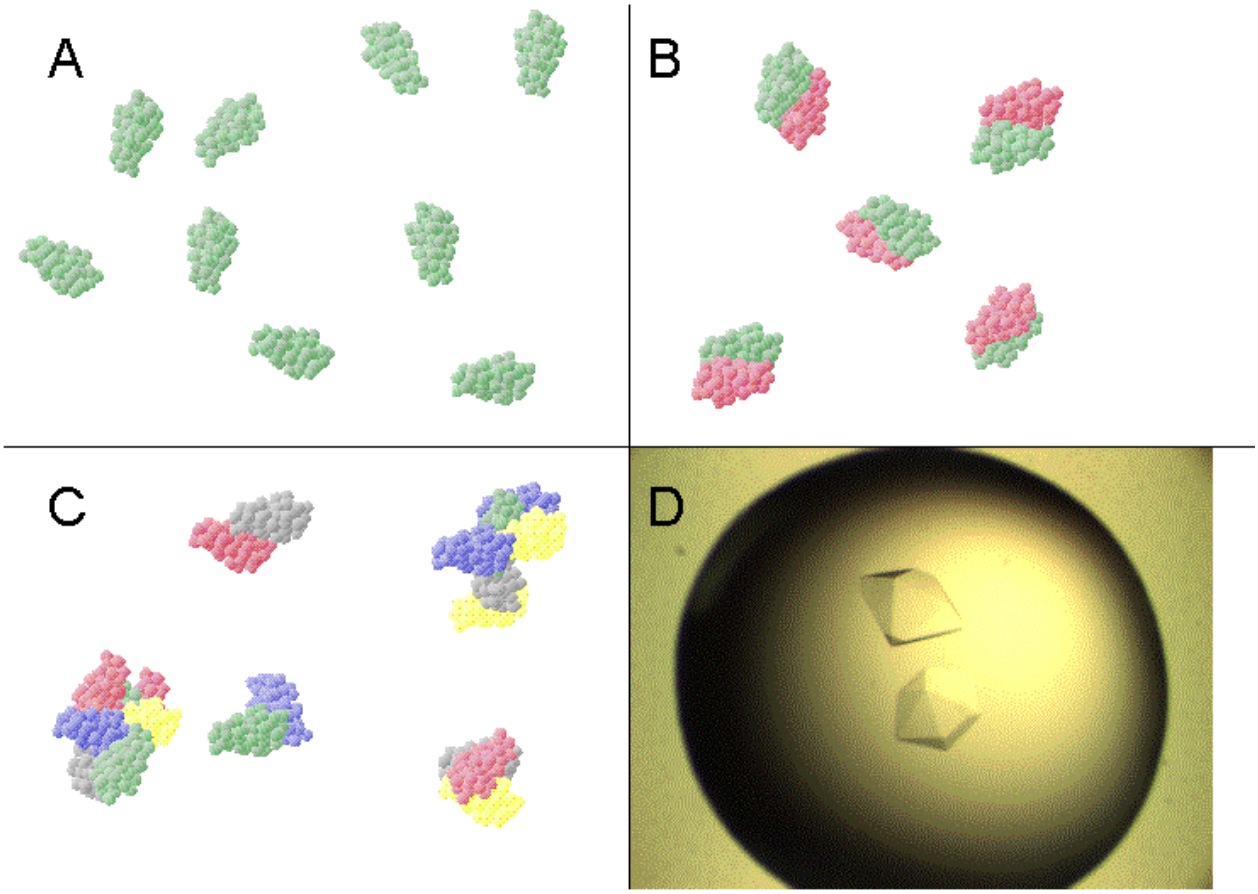


Figure 5.1: Protein interactions play a role in the transition between different states: solution of monomers (A), solution of dimers (B), unspecific precipitation (C), protein (GAPN) crystals (D, courtesy of Esben Lorentzen).

recently developed form of insulin (Insulin glargine), based on a mutation with a slightly higher pI (pH 7.0) than the native hormone (pH 5.4) forms small aggregates under physiological conditions. These are effectively reservoirs that dissolve slowly keeping the levels of insulin in the blood nearly constant over several hours [Schubert-Zsilavcz, 2001]. This represents a major improvement for diabetes patients, who can now keep stable insulin levels with a single daily injection of this new form of insulin.
Proteins interact in various ways and their interactions are thus difficult to predict. The basic physical theory of electromagnetic interactions, which essentially states that all intermolecular forces can be calculated on the basis of straightforward classical electrostatics if the spatial distribution of the electron clouds is known, is of little help in this context. In principle the electron distributions can be calculated from Schrödinger's equation but even after more than 100 years of quantum mechanics it is not yet possible to solve this equation exactly for two hydrogen atoms in vacuum. Therefore, in order to treat problems of intermolecular interactions, one uses simplified semi-empirical models. The following section provides an overview of the treatment of interactions in gases, liquids, colloids and protein solutions.

5.2 Basics

In an ideal gas at normal pressure a particle occupies on average 37000 Å^3. The volume of the molecules themselves varies, for instance, for the inert gases between 6.5 (argon) and 9.0 Å^3 (xenon). Thus the volume per molecule is several thousand times larger than the volume of the particle itself. In a pure liquid, e.g. water, an 11.5 Å^3-molecule has 29.9 Å^3 average volume per particle. This corresponds only to a factor of 2.6.
As an example of a protein solution the corresponding numbers for lysozyme at 50 mg/ml give an average solute volume per protein monomer of 474000 Å^3. The specific volume of lysozyme is approx. 0.74 cm^3/g, and consequently the volume of the particle is around 17600 Å^3. This corresponds to a factor of 27.0 for a very highly concentrated protein solution. This factor, as all values for lower protein concentrations, lies between those for gases and liquids. These numbers and the fact that the same basic interactions are involved provide evidence that it is reasonable to treat protein interactions within the same formalism, which has been useful for liquids and gases.
In order to investigate physical interactions between individual molecules it is useful to start from the properties of gases because the interactions can be described relatively simply and quantitatively. An ideal gas is a volume V filled with N molecules, which do not interact. This system serves as the basis for the derivation of the equation of state:

$$pV = Nk_BT\,. \tag{5.1}$$

p is the pressure, T the absolute temperature and k_B the Boltzmann constant. This equation, first formulated by Émile Clapeyron in 1834 combined the empirical findings of Robert Boyle $\left(pV = \text{const.}\right)$ and Jacques Charles $\left(\frac{V_1}{T_1} = \frac{V_2}{T_2}\right)$. Even for neutral molecules this approximation breaks down for high pressure and low temperatures. The equation of state of a real gas can be written as a power series in the form

$$\frac{pV}{Nk_BT} = 1 + \sum_{i=2}^{\infty} B_i(T)\rho^{i-1}\,. \tag{5.2}$$

This is the so-called virial expansion with its temperature dependent virial coefficients B_i [Hansen, 1986]. Independently of the measurement method experimental results on intermolecular interactions are often reported in the form of a truncated virial series.

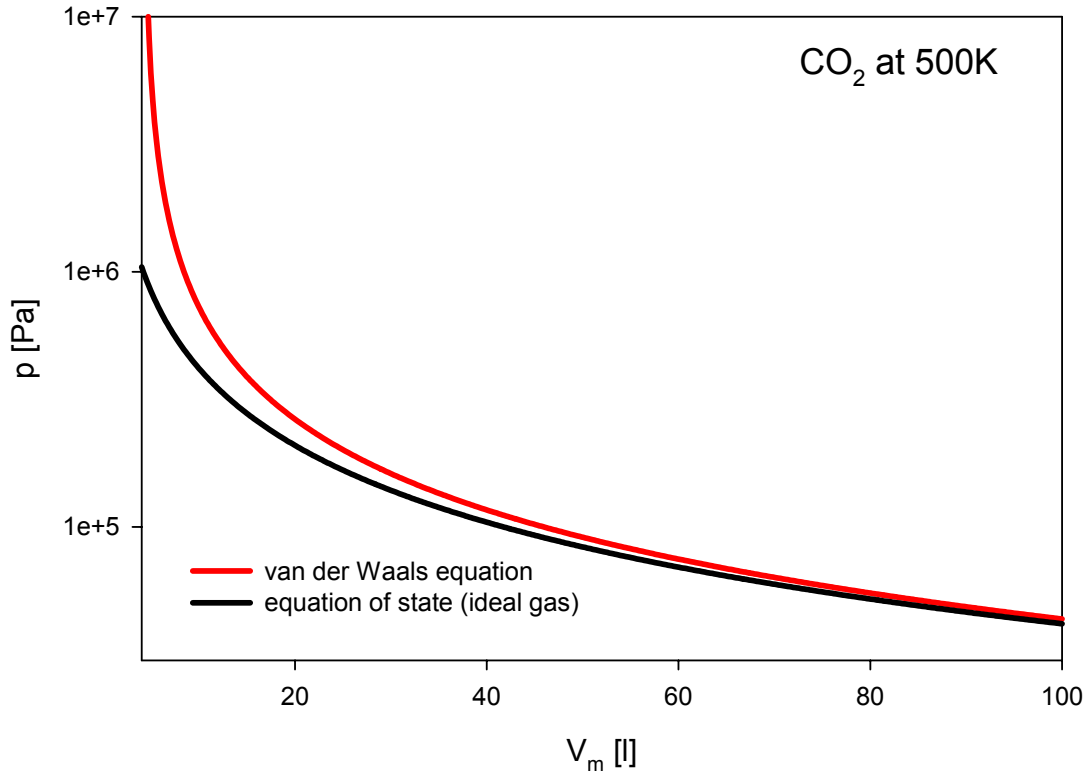


Figure 5.2: Comparison of the equation of state of an ideal gas with the van der Waals equation. The pressure of the gas CO_2 is presented as a function of the molar volume V_m. The van der Waals coefficients for CO_2 are $a = 3.688 \cdot 10^5 \frac{\text{Pa l}^2}{\text{mol}^2}$ and $b = 4.267 \cdot 10^{-2} \frac{\text{l}}{\text{mol}}$ [Atkins, 1997].

Another example of a successful approximation of the equation of state of a real gas is the so-called van der Waals equation that was proposed by J.H. van der Waals in 1873 [Atkins, 1997]:

$$p = \frac{nRT}{V - nb} - a\left(\frac{n}{V}\right)^2, \tag{5.3}$$

where n is the number of moles, $R=N_A k_B$ is the gas constant. The van der Waals coefficients a and b are characteristic of each gas but independent of temperature. b accounts for the volume of the molecules and restricts the effective volume to V-nb. a represents a weak, attractive interaction between individual molecules, which makes the pressure of the real gas smaller than that of an ideal gas. The forces responsible for the interaction are often called van der Waals forces. This model of short-range repulsion (due to the hard-sphere interactions) and long-range attraction (electromagnetic interactions such as induced dipoles) is very successful for a gas of neutral molecules (Figure 5.2).
In liquids the attractive interactions become stronger and simultaneously the short-range repulsion due to the hard-sphere interactions largely determines the structure. Whereas in a gas the outer boundaries of a system define the limits of the volume, the own volume of the molecules essentially defines the volume of the liquid.
In order to describe the interactions and structure of liquids pair potentials $u(r)$ (Figure 5.3) and distribution functions, which can also be used with gases, must be introduced. $u(r)$ is the derivative of the average, effective forces between two molecules at a distance r. It is

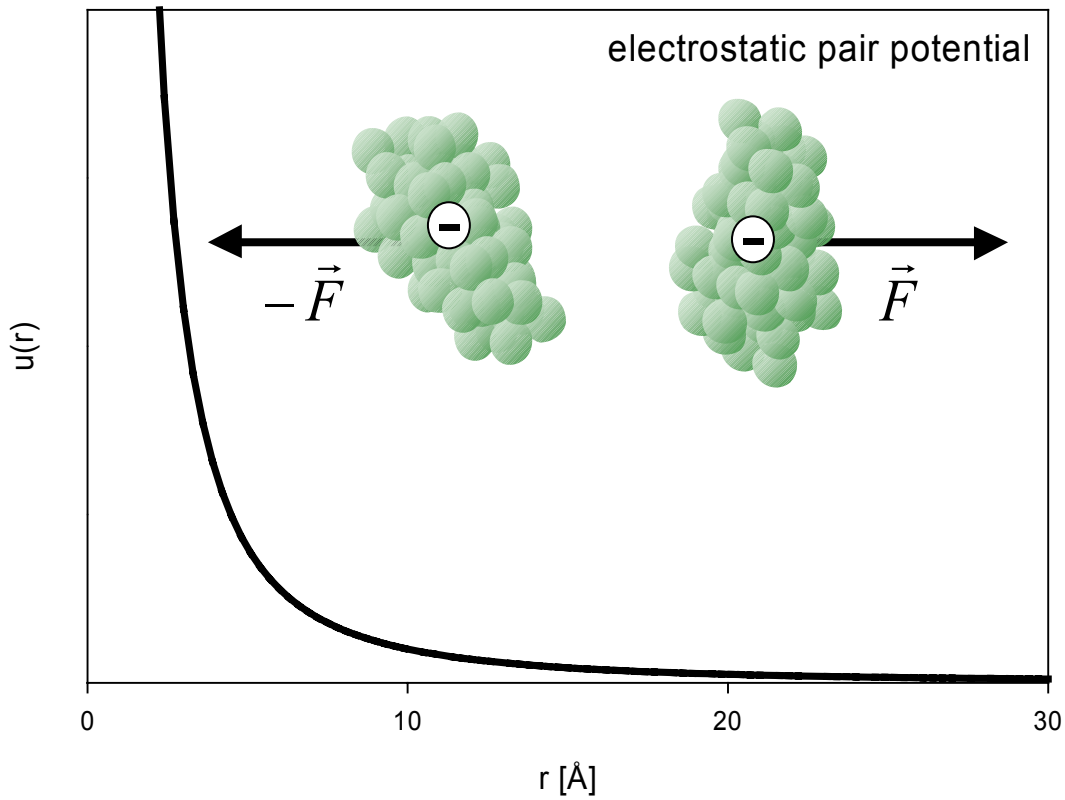


Figure 5.3: The pair potential can represent the forces between two proteins. Here the electrostatic potential between two charged molecules is taken as an example.

sometimes also called the potential of mean force. The simplest example of distribution functions is the radial pair distribution function $g(r)$, which is a measure of the probability to find another molecule with its centre at a distance r from the first one (Figure 5.4, Figure 5.5).

The most intuitive pair potential is the hard-sphere potential (Figure 5.6), which represents the fact that molecules cannot interpenetrate: Its value is zero if the two molecules do not overlap and infinite if they do. For a system of identical spherical molecules with radius R, the potential can be expressed as a function of the distance from the centre of a molecule:

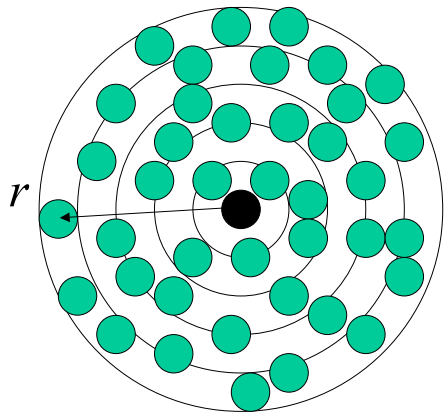


Figure 5.4: The radial pair distribution function $g(r)$ is a measure of the probability to find another molecule with its centre at a distance r from the first one.

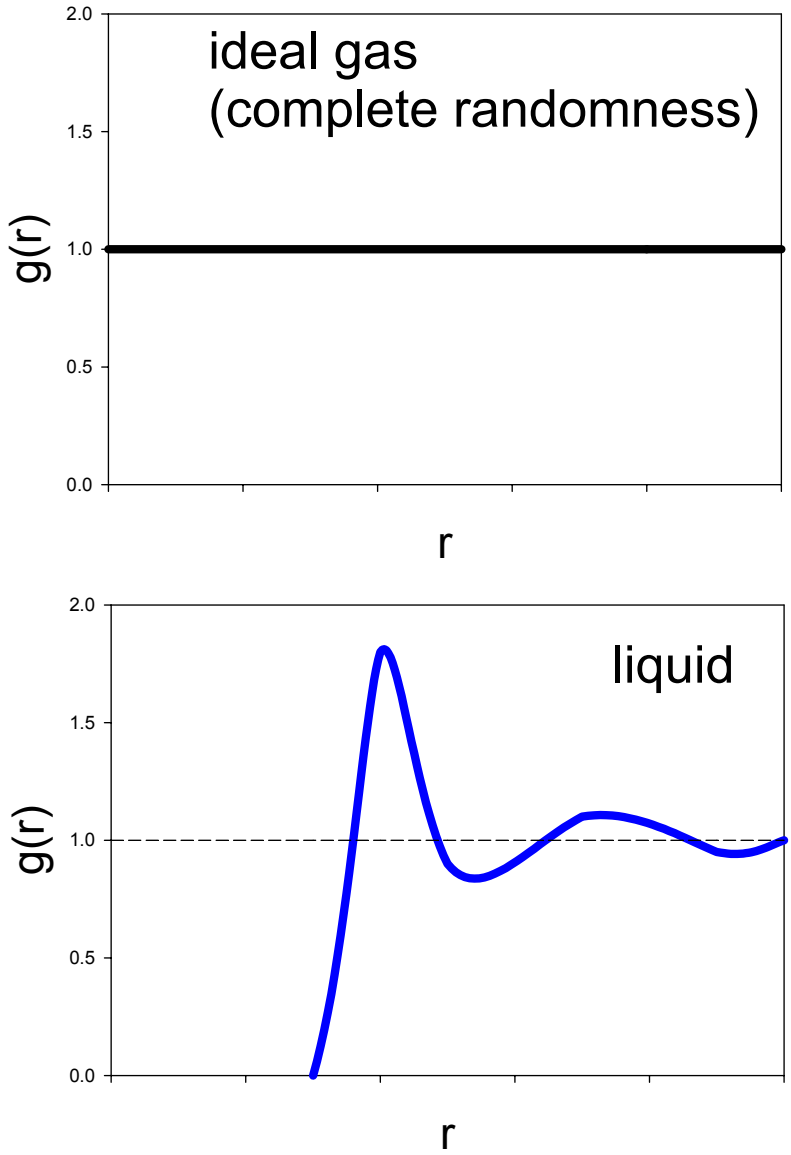


Figure 5.5: Radial pair distribution function $g(r)$ for a gas and a liquid. Whereas the distances in an ideal gas are not affected by interactions and $g(r)$ is equal to 1 in a liquid the hard sphere character and the short-range attraction play a role for the radial pair distribution function.

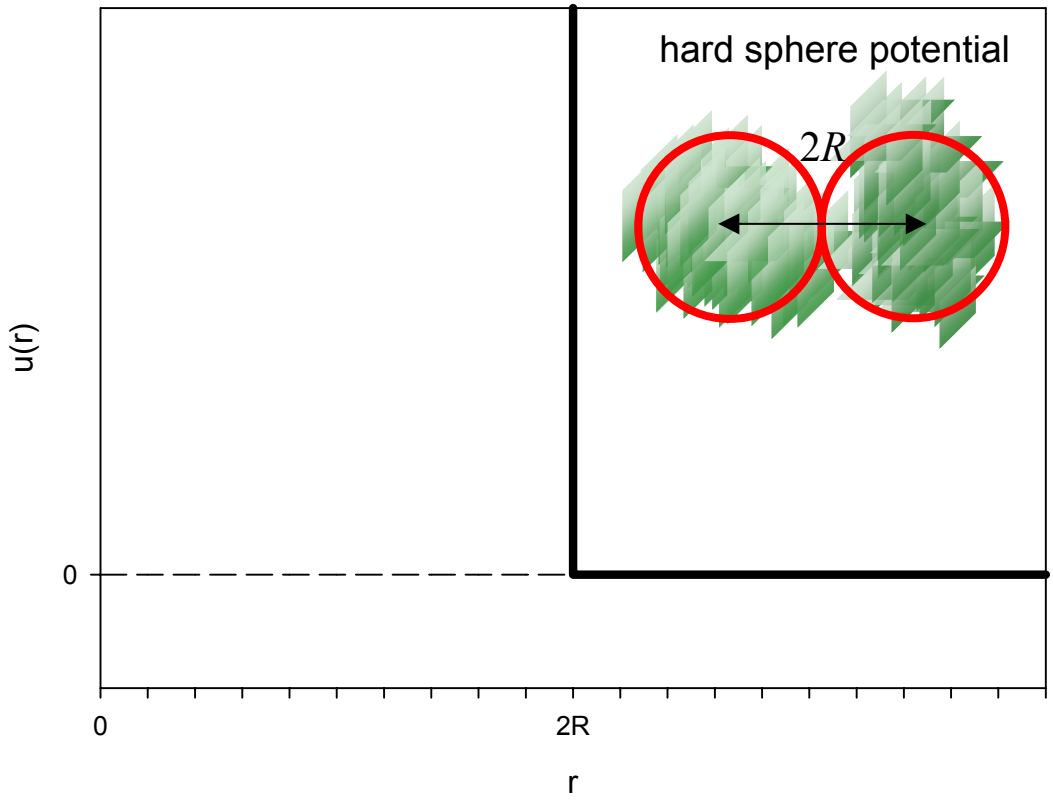


Figure 5.6: Hard sphere potential. Based on the assumption that the proteins can be approximated as spheres the potential becomes infinite if the distance drops below 2*R*.

$$u(\vec{r}) = \begin{cases} \infty, & \text{if } |\vec{r}| \le 2R \\ 0, & \text{if } |\vec{r}| > 2R \end{cases}. \tag{5.4}$$

Only in very special cases such as that of long, flexible polymer chains, such as polyethylene glycol (PEG) [Vivarès, 2002] is it appropriate to work with models of penetrable molecules.

As mentioned above all liquids have an attractive potential. Whereas attractive potentials are usually long-range in gases, in liquids short range models can successfully be applied such as the simple square-well model ([Hansen, 1986], Figure 5.7):

$$u(\vec{r}) = \begin{cases} \infty, & \text{if } |\vec{r}| \le 2R \\ -\varepsilon, & \text{if } 2\text{R} < |\vec{r}| \le 2R + \gamma \cdot (2R) \\ 0, & \text{if } |\vec{r}| > \gamma \cdot (2R) \end{cases}. \tag{5.5}$$

The parameter γ is typically 1.5. The dominant contribution for neutral molecules without permanent dipole moment to the physical origin of the attractive potential comes from the multipole dispersion interaction between the instantaneous electric moments in one molecule and those induced in the other (London forces). The moments arise because of spontaneous fluctuations in the electronic charge distribution. Over relatively small ranges of r the pair potential can be adequately represented by an exponential function of the form $\exp\left(-\frac{r}{r_0}\right)$, where r_0 is a range parameter.

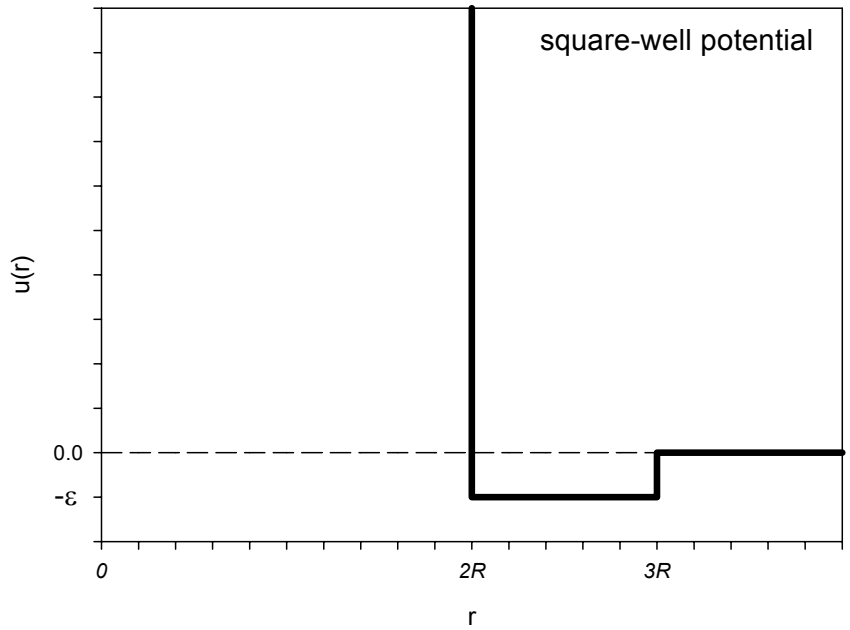


Figure 5.7: Square-well potential (γ=1.5). The simple square-well model can successfully be applied in liquids.

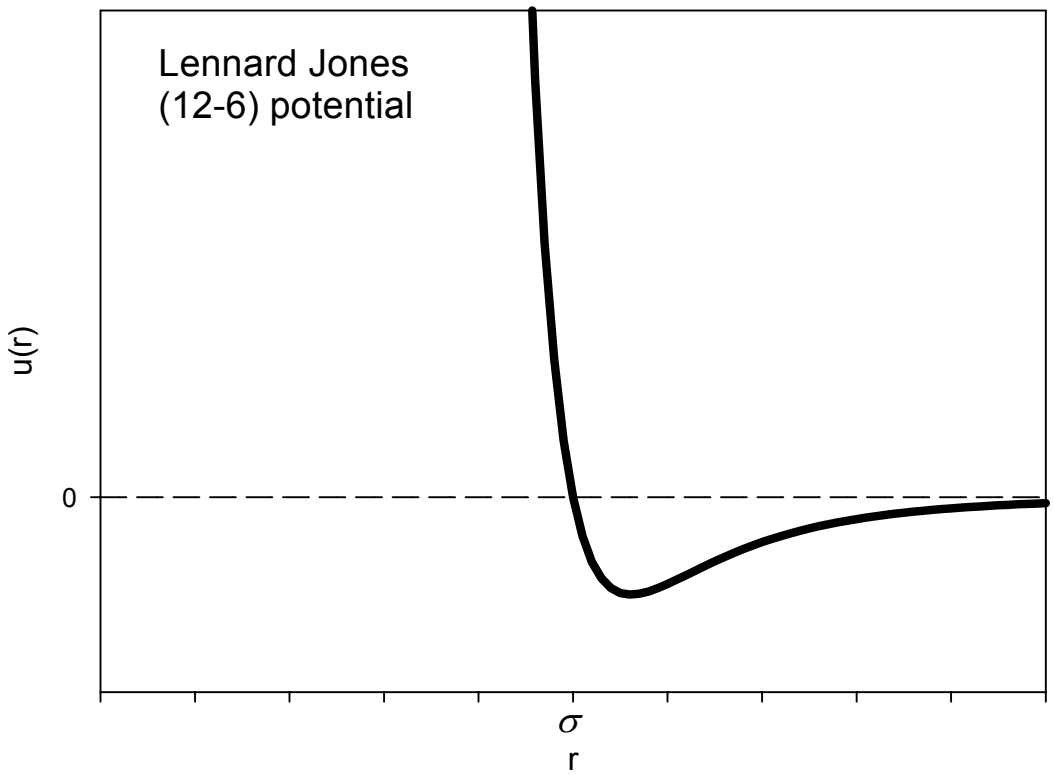


Figure 5.8: : Lennard-Jones potential.

It is often convenient to represent both, short-range repulsion and long range-attraction, by inverse power laws, i.e. r^{-n}. A famous example is the Lennard-Jones (12-6) potential (Figure 5.8):

$$u(\vec{r}) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]. \tag{5.6}$$

This provides a fair description of the interaction between pairs of rare-gas atoms and also of quasi-spherical molecules such as CH_4 and it was extensively applied to liquids, surfaces, clusters and two-dimensional systems.

The mathematical treatment of the problem can also be applied to colloids dispersed in solution. In the 1940s a theory of interactions, which can explain some of the phenomena occurring in such systems, was developed independently by B.V. Derjaguin & L.D. Landau and E.J.H. Verwey & J.TH.G Overbeek (DLVO). This theory, today referred to as DLVO theory, after the names of its authors, explains many aspects of the interactions and stability of colloids [Verwey, 1948]. According to the theory the main interaction between spherical particles are the hard-sphere interactions, a short range attraction, due to surface-surface forces, and a long range repulsion caused by the fact that the particles are charged.

The interactions in a protein solution are somewhat similar to those described by the DLVO theory for colloids. Even though proteins are polypeptide chains in physiological conditions at least the globular proteins behave as rigid molecules. Therefore it is reasonable to assume a repulsive interaction that can be represented by a hard-sphere

potential. Even though proteins, which are soluble in water (i.e. not membrane proteins), usually have a hydrophilic outer surface most of them tend to aggregate or to crystallize under appropriate conditions. It is thus reasonable to conclude that there are also attractive forces between proteins. These interactions are hydration forces and forces between multipole charge distributions in the molecules. In addition proteins carry a well-defined net charge at a certain pH. If the resulting long-range repulsion between the molecules is stronger than the attractive interactions the protein remains soluble. The potential models used in the DLVO approach for colloids take into account qualitatively similar interactions and it was therefore natural to use the potentials with proteins.
Small-angle scattering data lead to the conclusion that the interaction between small proteins like lysozyme can indeed be represented by a combination of hard-sphere, short-range attractive and Coulomb repulsive potentials [Tardieu, 1999]. The potential can be modelled as the sum of a hard-sphere and two Yukawa potentials (details are given in the next chapter), parameterised according to the DLVO theory. This quite detailed theory remains an approximation based on low-resolution information rather than on the detailed molecular structure of lysozyme. The interactions are too complex to be treated in more detail as illustrated by the list of all known interactions between molecules in solution:
As mentioned already in the beginning all interactions are basically due to the distribution of electrons and nuclei.

- The strongest interactions are the covalent bonds between the atoms. They are responsible for the primary structure of solute and solvent (Figure 5.9.A).
- The second strongest interaction is the Coulomb interaction between two point charges q_1, q_2:

$$u(\vec{r}) = \frac{q_1 q_2}{4\pi\varepsilon_o \varepsilon_r r}. \tag{5.7}$$

  The $\frac{1}{r^2}$-decay of the force in a homogenous medium in this law implies that Coulomb forces have a very long range (Figure 5.9.B).
- Not only molecules with net charges are exposed to electrostatic interactions. An asymmetry of the distribution of negative and positive charges leads to a dipole moment. A dipole will interact with charges, either negative or positive, and other dipoles. If the dipole is fixed the interaction can be attractive or repulsive depending on the relative orientation of the two molecules. If the dipole can rotate freely, such as a water molecule, the interactions are always attractive. The mean potential between two freely rotating dipoles (electric dipole moments u_1, u_2) can be formulated in the following way (Figure 5.9.C):

$$u(r) = \frac{-u_1 u_2}{3(4\pi\varepsilon_o \varepsilon_r)^2 kTr^6}. \tag{5.8}$$

- Neutral, non-dipolar molecules also are influenced by electrostatic interactions. Charges as well as dipoles induce dipole moments in non-polar molecules and this gives rise to attractive forces (Figure 5.9.D).
- For the sake of completeness the above-mentioned London forces are included in this list, too. These dispersion forces between two non-polar molecules depend on the electric polarizability α of the molecules (Figure 5.9.E).

- The solute-solvent interactions complicate especially the short-range interactions considerably. The presence of the macromolecule creates a border area region where the bulk structure of the solvent, water in the case of proteins, is disturbed. The charge distribution in the protein strengthens this effect. The protein has a hydration shell, which is formed by the first shell of water around the molecule. The protein has an effect even on water molecules beyond this layer. The total region of interactions referred to as the hydration zone. If the free energy associated with this perturbation varies with the distance between the two solute molecules, it produces an additional "solvation" or "structural" force between them (Figure 5.9.F).

A very important issue concerning interactions between proteins is that they are not only determined by the proteins themselves. The solvent also plays a major role. First the pH of the solvent determines the net charge of the proteins. At a pH that is equal to their pI proteins do not carry a net charge. At higher pH they are positively charged and at lower

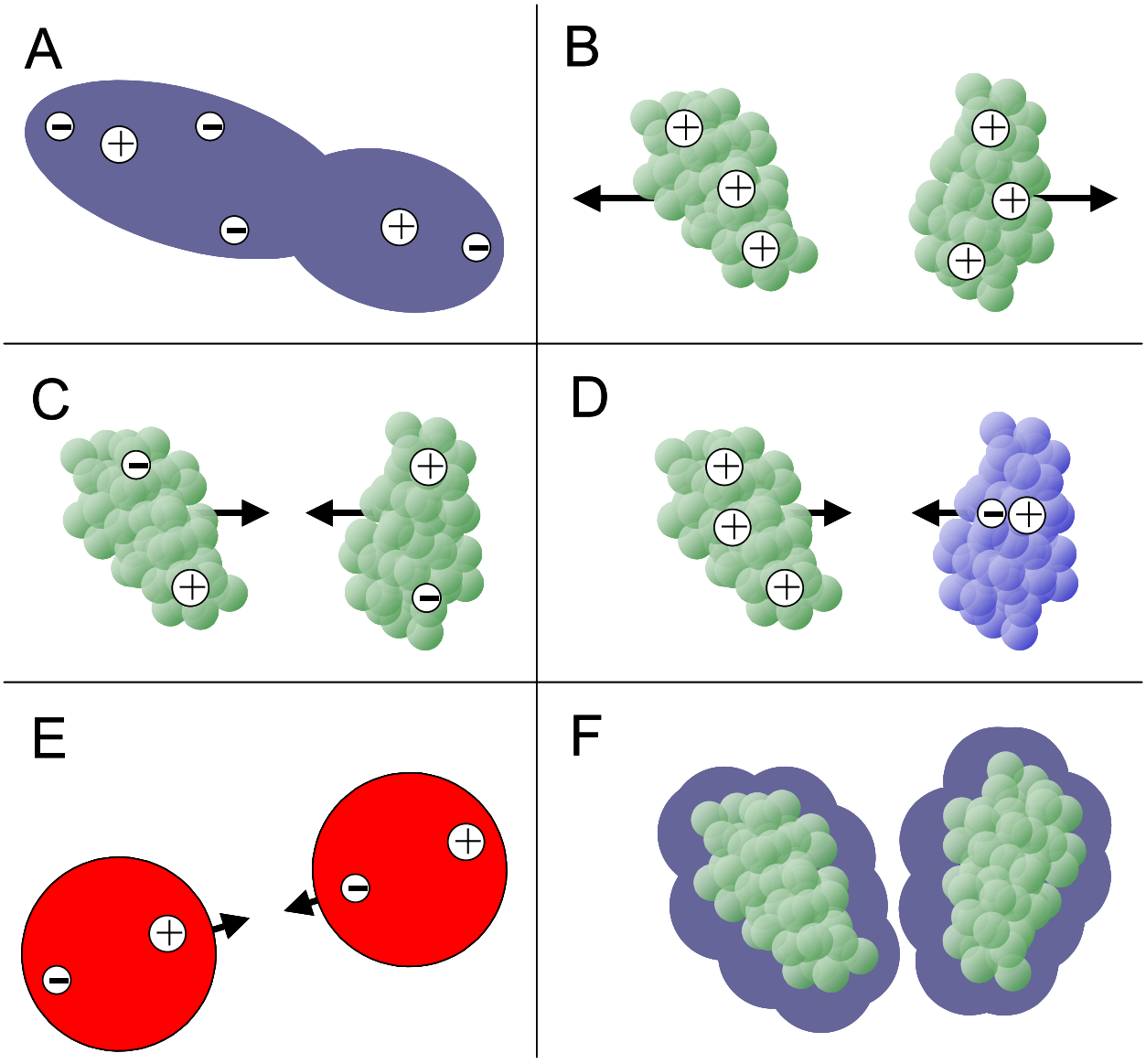


Figure 5.9: Intra- and intermolecular forces. Covalent bond (A). Electrostatic forces (B). Dipole-dipole forces (C). Forces due to induced charge distributions (D). London forces (E). Hydration forces due to the formation of a hydration layer (F).

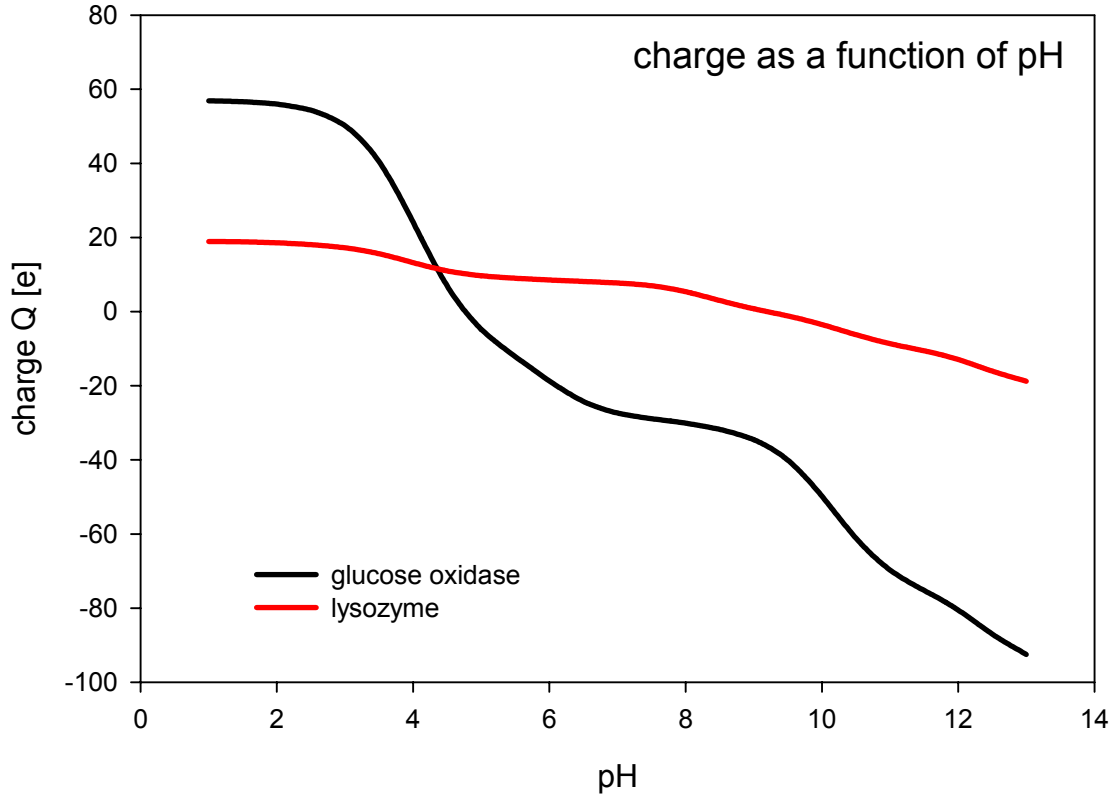


Figure 5.10: Net charge of the proteins glucose oxidase and lysozyme as a function of the pH of the solvent according to their amino acid sequence calculated by the "Compute pI/Mw" tool of ExPASy [Wilkins, 1998].

pH they have a negative net charge. Details about the charge as a function of pH can be estimated using the amino-acid sequence of the proteins (Figure 5.10).

In aqueous solution it has to be taken into account that $\varepsilon_r \approx 81$, as a result of the dipole character of the water molecules that orient in the electrostatic field and thus make the effective field weaker. The presence of ions in water is even more significant. Even in the absence of salt ions a certain fraction of water molecules dissociates into H^+ and OH^- ions.

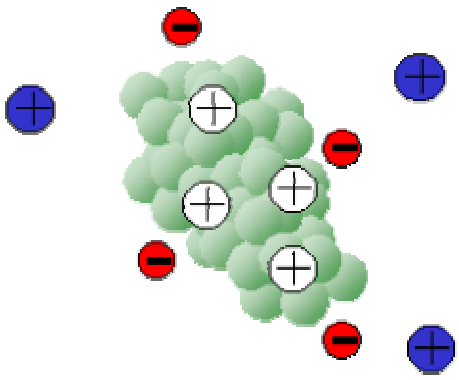


Figure 5.113: Salt screening. If charged macromolecules (in this case charged positively) are present in a solvent the counter-ions arrange in the vicinity of the macromolecule. The molecule appears neutral to other test charges further away from the molecule. The free ions screen the electrostatic interactions.

Since the pH is the negative logarithm of the H^+-ion concentration it also provides information about the number of the H^+- and OH^- -ions around the proteins and therefore the charge shielding. At pH 7 there are by definition 10^{-7} mol/l of H^+ and OH^- ions present in any aqueous solvent. That means that one in 5.55×10^8 water molecules is dissociated. At pH 4 there is thus one H^+-ion for 5.55×10^5 water molecules. The presence of ions effectively limits the range of the electrostatic interactions by charge screening (Figure 5.11). If the macromolecule is for instance positively charged, the negative counter ions will be found with a higher probability in a layer around the macromolecule. This reduces the effective charge on the protein.
The effect of charge shielding is also strengthened by the presence of any other ions and all kinds of electrostatic interactions become weaker when salts are added to the solution.

5.3 Experimental methods

Several methods exist to measure and quantify the interactions of proteins. Many of them make use of the fact that there are similarities between a protein in aqueous solution and a molecule in a gas. The volume of this "gas of macromolecules" is effectively the water volume and in analogy to the pressure of a gas one defines the osmotic pressure Π, which leads as in the cases of gases to a virial expansion:

$$\frac{\Pi}{cRT} = \frac{1}{M} + \sum_{i=2}^{\infty} A_i c^{i-1} \quad . \tag{5.9}$$

The coefficients A_i are the so-called osmotic virial coefficients and a measure for the interactions of the solute. For weak interactions and low concentrations A_2, which can be directly obtained from a measurement of the osmotic pressure (Figure 5.12) as a function of concentration provides useful information. In the case of predominantly attractive

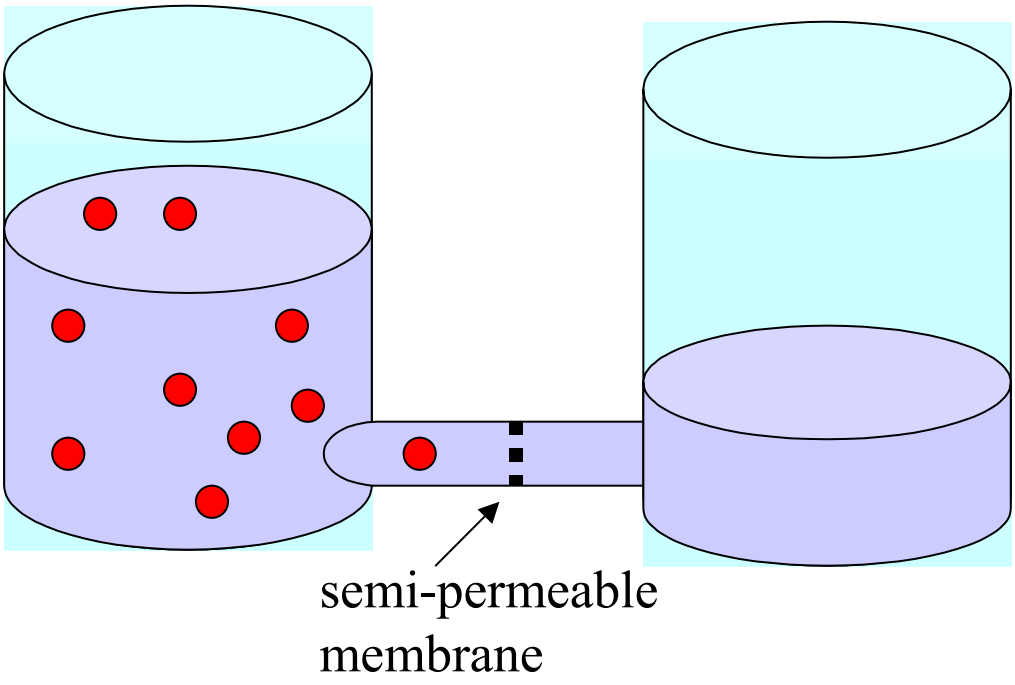


Figure 5.12: Apparatus to measure the change of osmotic pressure upon addition of a substance. The macromolecules in the left container cannot penetrate the semi-permeable membrane. Their repulsive interactions cause the rise of the solvent level in the left container.

interactions the osmotic pressure is lower than in the absence of interactions and consequently A_2 is negative. For repulsive interactions it is positive. Other methods to investigate interactions and measure A_2 are small-angle X-ray and static light scattering (Figure 5.13). Since light and X-rays are both electromagnetic waves the scattering methods are based on the same theory. At low concentrations the forward-scattering intensity can be used to determine the molecular mass of macromolecules in solution. At increasing concentrations the intermolecular interactions, namely A_2, influence the forward-scattering intensity, too. Thus it is possible to determine A_2 if the concentration and the molecular mass M_w are known. For light scattering the Zimm equation can be used to evaluate A_2. The formalism and equations used for the investigation of interactions in small-angle X-ray scattering are described in chapter 6. The advantage of small-angle X-ray scattering over light scattering is that it is not only possible to determine A_2 but also the so-called structure factor, which is a function of the scattering angle. The structure factor makes it easier to distinguish between short- and long-range forces. This information is lost for light scattering where the wavelength of the scattered radiation is too long to provide information about Å-dimensions.

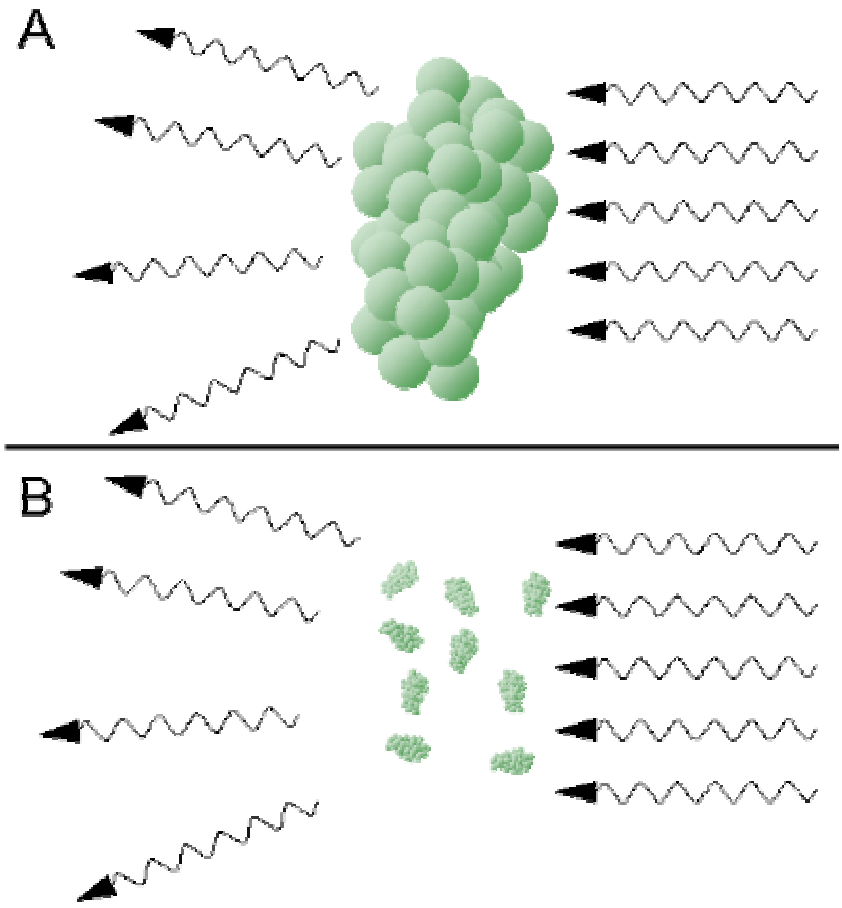


Figure 5.13: Light and X-ray scattering can be used to get information about the shape and interactions of the macromolecules. The scattering intensities contain information about the interferences within the particle (A) and information about the interactions between different particles due to interparticle interferences (B).

5.4 Statement of the problem

The proteins lysozyme (14.3 kDa) and glucose oxidase (160 kDa) were selected to investigate protein-protein interactions with SAXS as described in chapter 6. In several series of experiments both the protein concentration and the solvent composition were varied. The aim was to test the influence of variables like temperature, pH and salt concentration on the attractive and repulsive interactions and in particular the pair potentials. These parameters were varied in many previous studies and it was checked whether our studies lead to the same results. In some of the series glycine (NH_2CH_2COOH), urea (CON_2H_4) and trimethylamine *n*-oxide (TMAO) were added to the protein solutions to determine how they alter the protein-protein interactions. These substances were expected to influence the stability of proteins. Urea is know to destabilise them. It is added to protein solutions in order to unfold the protein [Bennion, 2003]. The amino acid glycine and the co-solvent TMAO however are known to have the opposite effect. They stabilise proteins and can counteract the effect of urea [Bennion, 2004].

6 Proteins, sample preparation and theoretical background for the SAXS measurements

This chapter describes the small-angle scattering methods used to investigate the particle shape and interaction potentials of the proteins hen egg white lysozyme and glucose oxidase from Aspergillus niger. It gives a brief introduction to the enzymes, their use and what is known about them. The biochemical methods for the sample preparation are described. The models for X-ray scattering by macromolecules are introduced as well as the methods for extracting structural parameters from the data. A short description of the computer software used to calculate these parameters is also given.

6.1 The proteins: Lysozyme and glucose oxidase

For the investigation of protein interactions in solution two well-known model proteins, for which the amino acid sequences and crystallographic structures are available, were chosen. The molar mass of lysozyme is 14.3 kDa and that of glucose oxidase 160 kDa (1 Da = 1 u).

6.1.1 Glucose oxidase from Aspergillus niger

The enzyme glucose oxidase (Figure 6.1) from the mould Aspergillus niger is a dimeric glycoprotein with a molar mass of approximately 160 kDa and a carbohydrate content of 16% [Hecht, 1993]. The carbohydrate moiety consists mainly of mannose (80%). The enzyme catalyses the reaction of β-D-glucose and molecular oxygen to D-glucono-1,5-lactone and hydrogen peroxide (Figure 6.2).

$$\beta\text{-D-glucose} + O_2 \xrightarrow{\text{glucose oxidase}} \delta\text{-gluconolactone} + H_2O_2 \qquad (6.1)$$

The enzymatic activity depends on its cofactor, flavin adenine dinucleotide (FAD), which is tightly but non-covalently bound to the protein and is responsible for its yellow colour. Each of the subunits of glucose oxidase contains one FAD molecule. Other mono-saccharides such as 2-deoxy- D-glucose, D-mannose, D-galactose and D-xylose are oxidized at much lower rates than β-D-glucose. The optimal enzymatic activity is found at pH 5.5. Glucose oxidase is of considerable industrial importance because it is used for the production of gluconic acid. It is also used as a food preservative since it produces hydrogen peroxide. Another interesting application is its use in diagnostics (for instance in colorimetric diagnostic kits) for the quantitative determination of sugar in samples such as blood, food, agricultural and fermentation products (biosensor). The enzyme can also be used to design biosensors that detect its inhibitors such as Ag^+, Hg^{2+} and Pb^{2+} [Wilson, 1992]. Biofuel cells (Figure 6.3) based on glucose oxidase, which generate electric energy

	lysozyme	glucose oxidase
molecular mass [kDa]	14.3	ca. 160
Configuration	monomer	dimer
PI	10.5-11	3.9-4.3
ε_{280}, literature [ml/(mg cm)]	2.66	1.6 ±0.32
amino acids (monomer)	129	583

Table 6.1: Main physico-chemical parameters of the proteins.

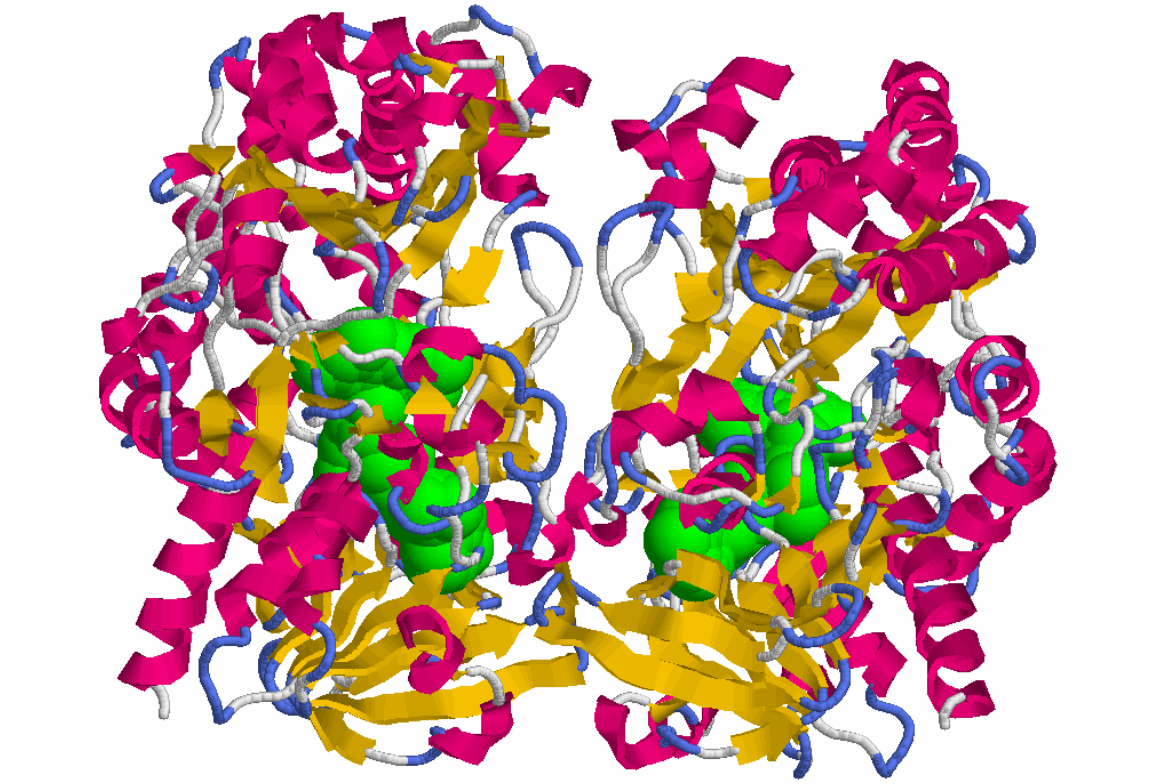


Figure 6.1: Glucose oxidase dimer [Hecht, 1993]. The green atoms represent the FAD subunit. The image was generated with the program RasWin [Sayle, 1995].

by oxidation of glucose, are being investigated [Willner, 2002]. Due to its many immediate and potential applications a considerable amount of information about glucose oxidase is available. This, together with the fact that it is commercially available, highly stable and soluble makes it an interesting example for the investigation of protein interactions.
The apo-protein consists of two identical chains each with 583 amino acids and its sequence is known [Frederick, 1989]. The translational diffusion coefficient was determined to be 4.94×10^{7} cm^2/s. Only a partial X-ray crystallographic model (Figure 6.1) is available [Hecht, 1993] because it was necessary to remove the carbohydrate moiety, which varies in size and composition from one molecule to the other, to crystallise the protein [Kalisz, 1990]. Six different forms of glucose oxidase were found with different pI in the range 3.9-4.3. From the sequence it is known that the protein has eight possible glycosylation sites [Frederick, 1989].
The size of the monomer is approx. 60Åx52Åx37Å, that of the dimer 70Åx55Åx80Å. The FAD motive has a minimum distance of 13Å to the surface of the protein (Figure 6.1).

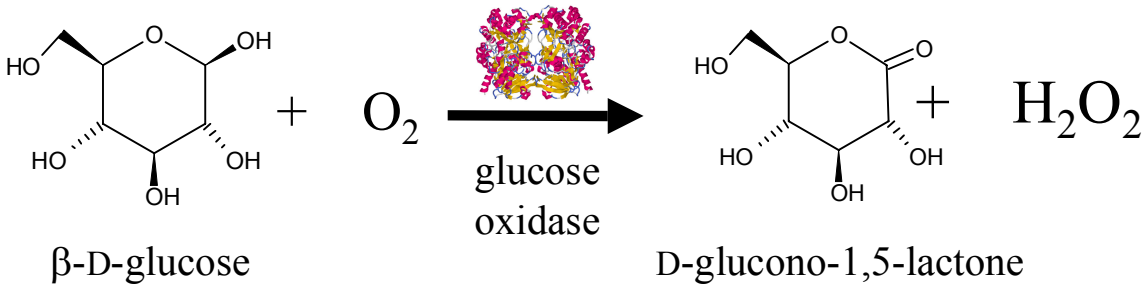


Figure 6.2: Reaction catalysed by glucose oxidase.

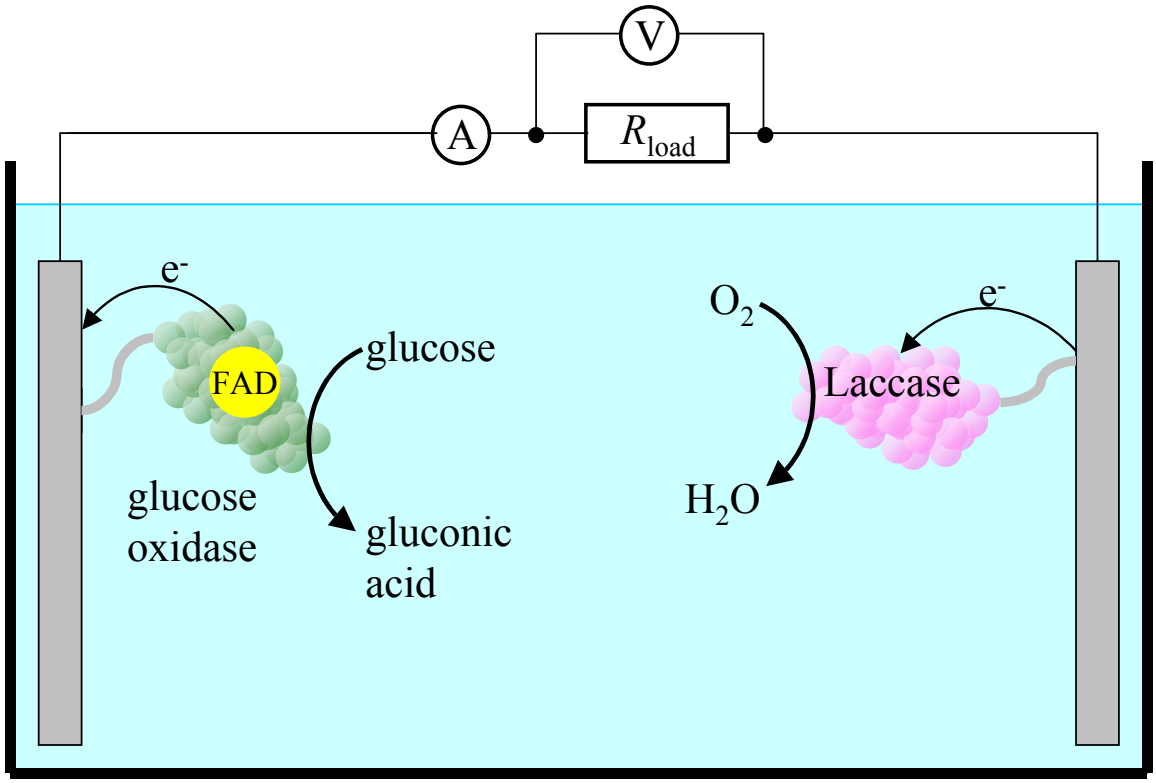


Figure 6.3: Glucose oxidase biofuel cell [Willner, 2002]. The construction of an electrically contacted glucose oxidase enzyme electrode and the complementary organisation of oxygen-reducing bioelectrocatalytic electrodes consisting of laccase creates an anode and a cathode that generate electrical power through the oxidation of glucose.

6.1.2 Hen egg white lysozyme

Lysozyme (Figure 6.4), an enzyme discovered by Alexander Fleming in 1922, is a benchmark protein in molecular biology. It catalyses the hydrolysis of specific kinds of polysaccharides comprising the cell walls of Gram positive bacteria and is part of the unspecific humoral defence system of vertebrates. It is present in most secretions (e.g. tears). In the human blood it has a concentration of 9 mg/l, in urine 0.3 mg/l [Löw, 2002]. It is used as a preservative in many food products including wine, tofu, cheese and saké. Lysozyme is also the active component in drugs against sore throat.

For the measurements presented below hen egg white lysozyme, which is active as a monomer, was used. The protein consists of a single chain of 129 amino acids with a molecular mass of 14296 Da containing four disulphide bridges. Its isoelectric point is at pH 10.5-11.

Lysozyme is one of the best known proteins corresponding to the highest number of entries from different sources in the protein data bank [Berman, 2000].

Figure 6.4: Hen egg white lysozyme [Vaney, 1996]. The image was generated with the program RasWin [Sayle, 1995].

6.2 Sample preparation

An advantage of SAXS is that structural parameters can easily be determined in different conditions (buffer, temperature etc.). Accordingly one can also investigate protein interactions as a function of temperature, salt concentration and ionic strength. Likewise it is possible to investigate interactions in the presence of cosolutes (such as PEG, glycine and urea). This chapter describes the process of sample preparation for the commercially available proteins glucose oxidase and lysozyme.
In order to keep the system simple and avoid the presence of any salt some of the measurements below were made in pure deionised water. The pH is then not accurately defined and therefore other measurements were made in buffers. These are solutions of a weak acid and a strong base or of a weak base and a strong acid in deionised water, which are used to keep the pH constant at a well-defined value. Whereas buffers for applications in molecular biology are very often filtered in order to remove any solid contaminants this procedure should be avoided in SAXS measurements. Indeed, the materials used to make the filters, may release some contaminants, which may give rise to a significant SAXS signal that disturbs the measurements. SAXS samples and buffers do not have to be strictly sterile provided they are freshly prepared. Macroscopic particles that are large enough to be filtered out would only contribute to the SAXS signal at very small angles, outside the range of observation. The last step for the preparation of the buffer is the removal of the oxygen. Traces of oxygen facilitate the formation of bubbles and of free radicals leading to

radiation damage during SAXS measurements. Oxygen can be removed by bubbling nitrogen through the buffer while stirring for 30 min to 1 h depending on the volume. Bubble formation is increasingly a problem at elevated temperatures. At room temperature or below virtually no bubbles form even in the presence of oxygen, but at temperatures above 60°C the removal of oxygen becomes essential.
Both proteins introduced in the previous section were bought as lyophilised powders, which always contain other low molecular mass substances such as salts. The powder is resuspended in a small amount of buffer taking into account the target concentration. Salts and small molecules can be removed from the powder by dialysis to obtain a well-characterised sample later on. Dialysis was performed using Pierce dialysis cassettes (Slide-A-Lyzer Dialysis Cassette; Extra Strength; 10000 MWCO). The choice of these cassettes is based on the pore-size of their semi-permeable membrane, which determines the molecular mass of the molecules that cannot diffuse through the membrane (cut-off). The ratio of solvent volumes on both sides of the membrane determines the maximum dilution that can be reached. With typical cassette volumes between 0.5 and 3 ml and dialysis buffer volumes between 200 and 1000 ml dilution factors between 60 and 2000 were achieved. The dialysis buffers in the beaker were generally exchanged twice during the process. The samples were usually dialysed overnight at room temperature and the buffer was continuously stirred. Typically the three dialysis steps lasted 1, 2 and 10 hours. Between the successive steps the samples were homogenised by vortexing.
For the choice of the initial concentration and the volume of the dialysis cassette it must be taken into account that the sample volume increases during dialysis as the high protein concentrations result in a high osmotic pressure. The protein concentration must thus be measured after dialysis. In our case the protein was most often dialysed against pure deionised water to remove all kinds of ions. Afterwards the required salts were added to the protein solution in accurately determined concentrations.
Undissolved material was removed from the sample after dialysis by centrifugation for 15 min at maximum speed. The undissolved material accumulates as a pellet at the bottom of the tube and the supernatant is transferred into a new tube.
Several protein concentration series were measured. The protein was prepared at the highest concentration and diluted with buffer with exactly the same composition as the protein buffer, ideally with the buffer of the last dialysis step.

6.3 Concentration measurements

For SAXS measurements it is mandatory to accurately determine the concentration c[mg/ml] (weight/volume) both of the calibration standard and of the sample. In practice, the errors on the concentration measurements are the most important contribution to the errors on the estimate for the molecular mass by scattering methods.
Accurate concentrations are equally important for the investigation of protein interactions since the number density of the protein contributes as a scaling factor to the relation between $g(r)$ and the structure factor SF (see the section “Interactions sensed with SAXS”).
Most methods for the determination of protein concentration are based on the absorption of ultraviolet light (UV) at 280 nm essentially by the aromatic residues (tyrosine (Tyr) and tryptophan (Trp)) in proteins (Figure 6.5). The phenylalanine (Phe) residues and the disulfide bonds also contribute to a smaller extent. Consequently the extinction coefficients ε_{280} vary greatly between different proteins depending of their amino acid composition. Note that the presence of small amounts of nucleic acids can lead to large errors since the extinction of nucleic acids is ten times greater than that of protein**.**
All protein concentrations here were determined by measuring the absorbance A at 280 nm on a Uvikon spectrophotometer 922 or a Pharmacia Biotech Ultraspec 3000.

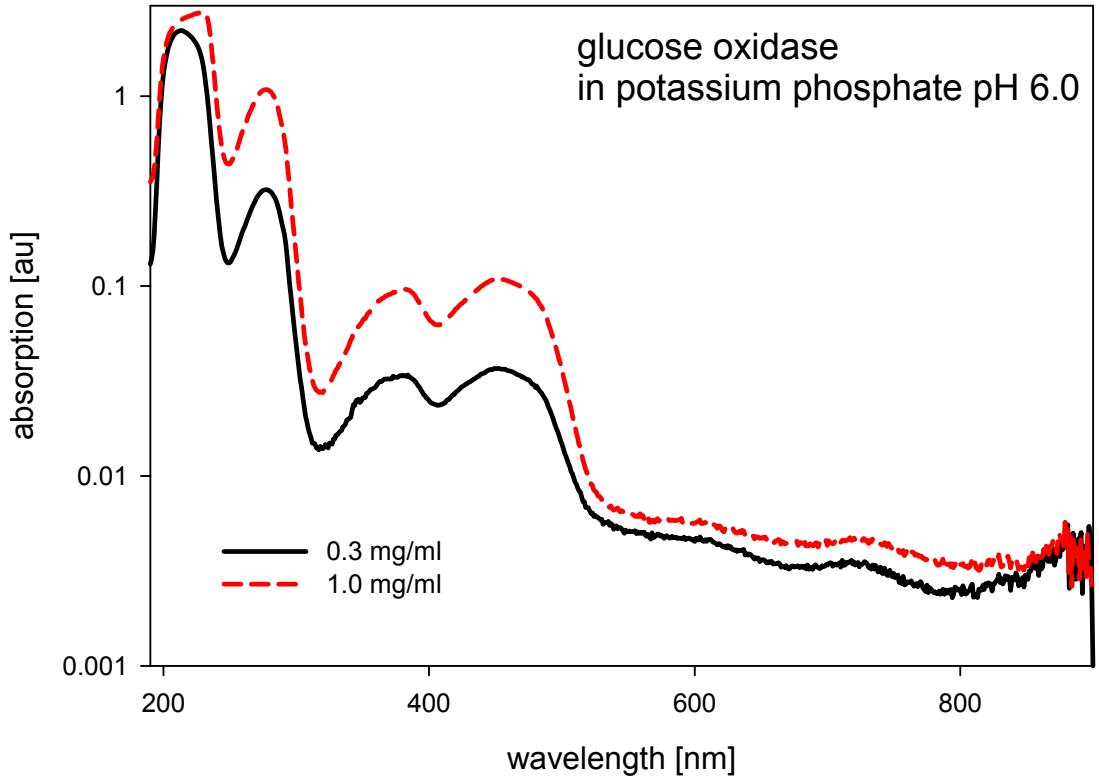


Figure 6.5: Absorption of visible light and UV of glucose oxidase.

$$A = -\log\left(\frac{I_{\text{sample}}}{I_{\text{ref}}}\right) = \varepsilon_{280} \cdot c \cdot l\,, \tag{6.2}$$

where I_{ref} is the intensity transmitted through the reference cell, i.e. a cuvette filled with buffer only (Figure 6.6). I_{sample} is the intensity after the cuvette with protein solution. The monochromatic beams impinging on the sample and buffer must of course have the same intensity. The quartz cuvettes used for absorbance measurements have a path length l of 1 cm and a volume of 700 µl and are transparent at 280 nm.

Various checks can be done to test the accuracy and precision of the instrument. For instance the measurements without cuvettes, with two matched empty cuvettes or with two matched cuvettes filled with the same buffer must result in the same absorption value A=1. If this fails it is necessary to check the device and/or clean the cuvettes more carefully. Measuring the absorbance of the same cuvette with the same sample several times gives information about the precision of the measurement. Emptying the cuvette and filling it again with the same protein solution provides information about the influence of sample

Protein	glucose oxidase	lysozyme
M [Da]	approx. 160000	14296
n_{w}	20	6
n_{y}	56	3
n_{c}	6	8
ε_{280}, calculated [ml/(mg cm)]	1.164	2.724
ε_{280}, literature [ml/(mg cm)]	1.38	2.66

Table 6.2: Extinction coefficients of glucose oxidase and lysozyme.

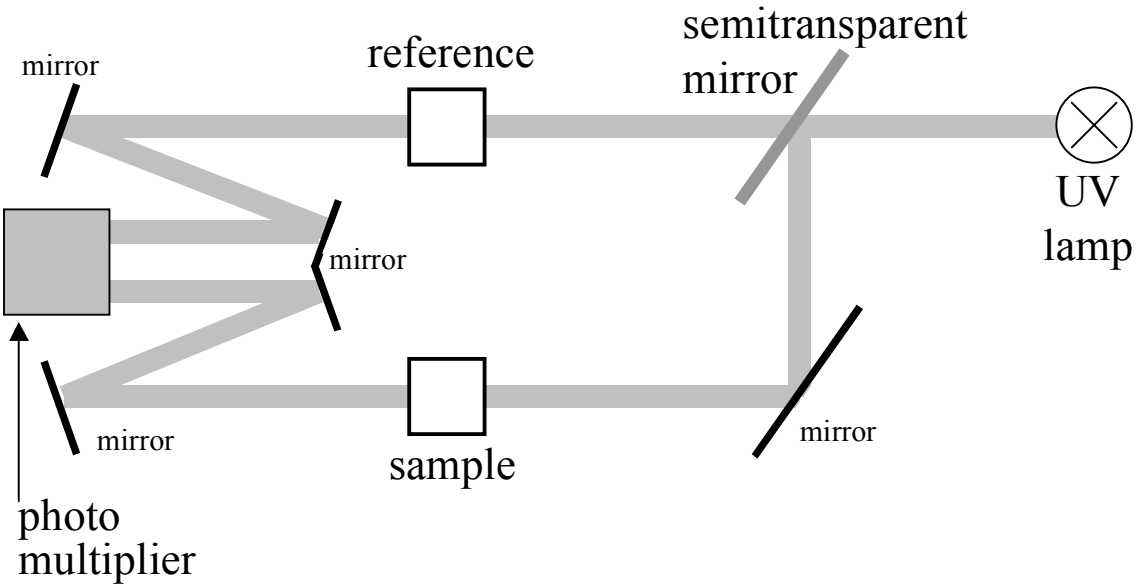


Figure 6.6: Double beam UV spectrophotometer for absorption measurements.

inhomogeneity.
For best results it is necessary to keep the absorbance of the protein solution below a value of 1. Since the protein solutions, which were prepared in order to observe interactions, are highly concentrated (up to 100 mg/ml) they must be diluted before measuring the UV absorption. The dilute protein solutions were therefore available in large quantities. It was found that washing the sample cuvettes with the dilute protein solution twice leads to more precise and reproducible results.
The extinction coefficient of a protein can be calculated approximately according to the following formula

$$\varepsilon_{280}[\frac{\text{ml}}{\text{mg}\cdot\text{cm}}] = (5690 n_w + 1280 n_y + 120 n_c)/M[\text{Da}], \tag{6.3}$$

where n_w, n_y, and n_c are the number of Trp, Tyr, and Cys residues in the protein of mass M in Da [Walker, 2002]. If tabulated values from the literature are available they should be used.

6.4 Small angle scattering by macromolecules

In chapter 2 the scattering of X-rays by single electrons was discussed. As long as the electrons can be considered to be free, i.e. not affected by any electromagnetic potential, this concept can be used to calculate the scattering pattern of any object. Molecules are modeled as an assembly of N electrons. Since the coherence length of the synchrotron radiation at X33 (longitudinal coherence $\xi_l = \frac{\lambda^2}{2\Delta\lambda} = \frac{1}{2}\cdot 10^4 \cdot 1.5\text{Å} = 7500\text{Å}$, vertical coherence length $\xi_v = \frac{\lambda a}{2\sqrt{\pi}\sigma_v} = 4000$ Å, horizontal coherence length $\xi_h = \frac{\lambda a}{2\sqrt{\pi}\sigma_h} = 8000$ Å, a is the source-to-observer distance (30m) and σ the one sigma source size ([Shang, 2003], Figure 6.7) exceeds the size of most proteins, interferences between the scattered waves of all electrons of a molecule have to be taken into account.
The separation between the scatterers results in a phase difference described by the phase factor.

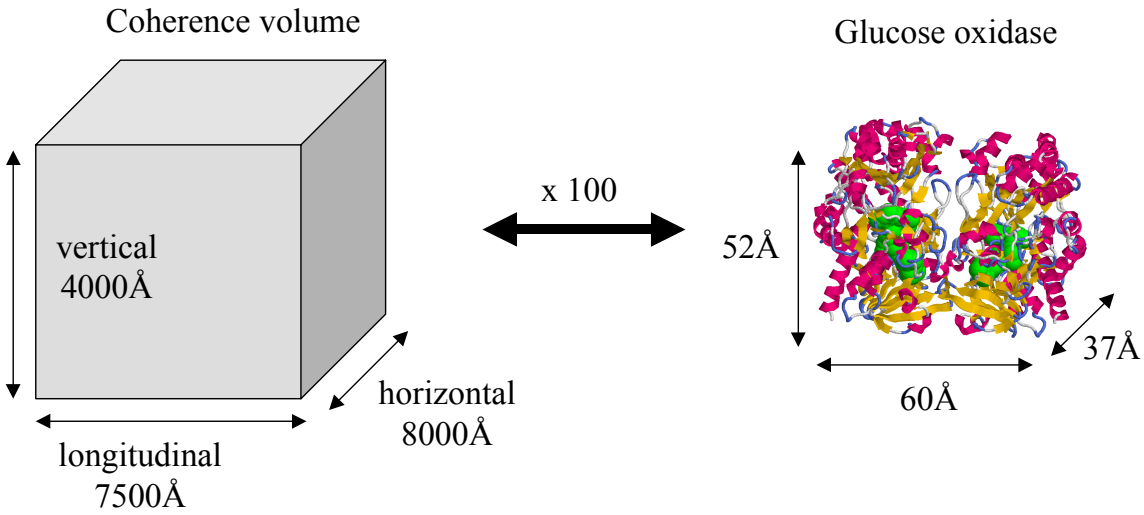


Figure 6.7: Coherence volume at the X33 beam line. The coherence length is about 100 times larger than the dimensions of the protein glucose oxidase.

$$F(\vec{s}) = \sum_{i=1}^{N} f_e \exp(i\vec{s} \cdot \vec{r}_i) \tag{6.4}$$

The momentum transfer vector $\vec{s}$ is the difference of the vectors representing the incident ($\vec{s}_i$) and the scattered wave ($\vec{s}_s$) (Figure 6.8). If the scattering angle is 2θ

$$\vec{s} = \vec{s}_s - \vec{s}_i \tag{6.5}$$

$$|\vec{s}_s| = |\vec{s}_i| = \frac{2\pi}{\lambda} \tag{6.6}$$

$$s = |\vec{s}| = |\vec{s}_s - \vec{s}_i| = \frac{4\pi \sin\theta}{\lambda} \approx \frac{4\pi\theta}{\lambda}. \tag{6.7}$$

Instead of the momentum transfer vector the scattering vector k can be used. It is parallel to s and its modulus is the inverse of the separation d between a pair of scatterers (Figure 6.9).

$$k = \frac{2\sin\theta}{\lambda} = \frac{1}{d}. \tag{6.8}$$

f_e is the scattering amplitude of a single electron. For solution scattering where the

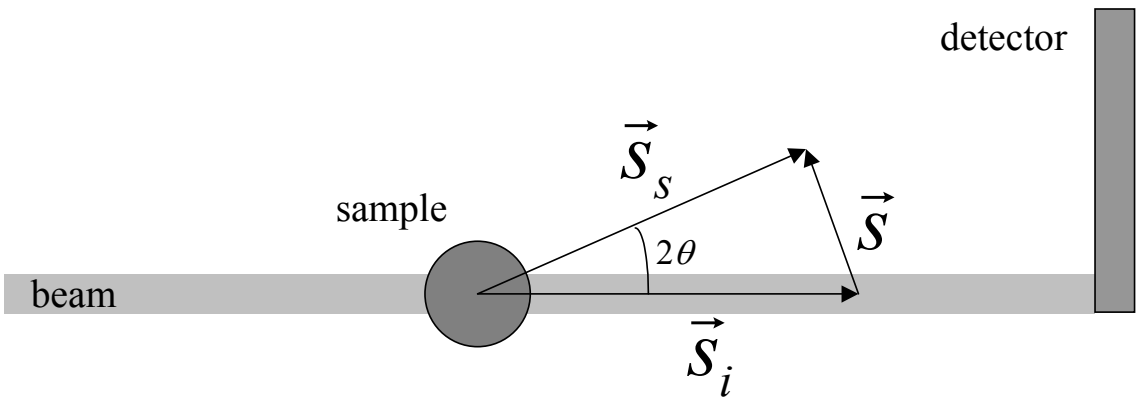


Figure 6.8: Momentum transfer vector $\vec{s}$.

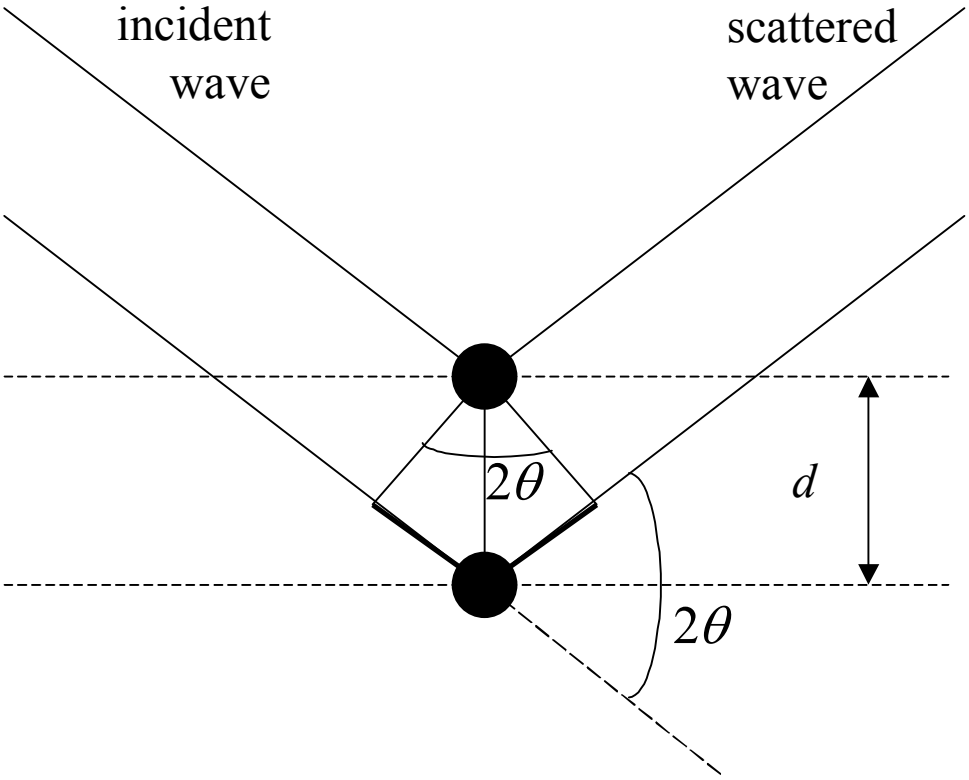


Figure 6.9: Scattering angle.

molecules can take all orientations, the phase factor can be replaced by its spherical average $\frac{\sin(sr)}{sr}$. Since for X-ray scattering at small angles $f_e = 1$ the sum over the electrons can be replaced by an integral over the function $\rho(\vec{r})$, which represents the electron density. Thus the scattering amplitude $A(\vec{s})$ of any sample (volume V) is

$$A(\vec{s}) = \int_V \rho(\vec{r}) \exp(-i\vec{s} \cdot \vec{r}) \mathrm{d}\vec{r} \,. \tag{6.9}$$

Due to the random orientation of the molecules in solution the resulting SAXS pattern is isotropic as a function of s ($\vec{s} = (s, \theta, \varphi)$).

Alternatively one can use Eq. (6.9) to calculate the scattering factors of single (spherical) atoms (for s=0 $f(0)$=Z, where Z is the number of electrons of the atom). The scattering factors of the atoms are well-known and tabulated [Wilson, 1999]. For randomly oriented molecules this leads to Debye's formula, which is very useful for calculating the SAXS pattern of crystallographic models:

$$I(s) = \sum_{i=1}^{N} \sum_{j=1}^{N} f_i(s) f_j(s) \frac{\sin(sr_{ij})}{sr_{ij}} \,. \tag{6.10}$$

Eq. (6.10) facilitates the computation of the SAXS pattern of a macromolecule, for instance from the crystallographic structure of a protein and thus the comparison of the structure of the protein in crystalline form and in solution [Svergun, 1995].

Compared to many other systems (e.g. synthetic polymers) protein solutions can be purified to obtain monodisperse or polydisperse (e.g. dimers and tetramers …) systems.

In the following section the formalism of contrast variation (for a review see [Koch, 2003]) is applied to a monodisperse solution of proteins. The scattering pattern $I_{\exp}(s)$ depends on

the difference between the average electron density of the particle ρ_p and the average electron density of the buffer ρ_b. This difference $\overline{\rho} = (\rho_p - \rho_b)$ is called the contrast.

$$\rho(\vec{r}) = (\rho_p - \rho_b)\rho_c(\vec{r}) + \rho_s(\vec{r}) = \overline{\rho}\rho_c(\vec{r}) + \rho_s(\vec{r}) \tag{6.11}$$

$\rho_s(\vec{r})$ is due to the fluctuation of the electron density inside the particle around its average value. $\rho_c(\vec{r})$ represents the shape of the particle:

$$\rho_c(\vec{r}) = \begin{cases} 0 & \text{outside the particle} \\ 1 & \text{inside the particle} \end{cases}. \tag{6.12}$$

The scattered intensity is the squared modulus of the scattering amplitude averaged over all possible particle orientations.

$$I(s) = \left\langle A(\vec{s}) \cdot A^*(\vec{s}) \right\rangle_\Omega = \overline{\rho}^2 I_c(s) + \overline{\rho} I_{cs}(s) + I_s(s)\,. \tag{6.13}$$

The indices c, s and cs on the intensity terms correspond to the indices c and s of ρ in expression (6.11) and express that I_c is due to ρ_c, I_s due to ρ_s, and I_{cs} is the mixed term.

The two most important structural parameters that can be extracted from a scattering pattern are the scattered intensity at the *s*-origin or forward scattering, *I*(0), and the radius of gyration R_g of a particle:

$$R_g^2 = \frac{\int \rho(\vec{r})(\vec{r} - \vec{r}_0)^2 d\vec{r}}{\int \rho(\vec{r}) d\vec{r}}, \tag{6.14}$$

where $\vec{r}_0$ is the position of the centre of mass. (6.14) is equivalent to

$$R_g^2 = \frac{\int\int \rho(\vec{r}_1)\rho(\vec{r}_2)(\vec{r}_1 - \vec{r}_2)^2 d\vec{r}_1 d\vec{r}_2}{2\int\int \rho(\vec{r}_1)\rho(\vec{r}_2) d\vec{r}_1 d\vec{r}_2} \tag{6.15}$$

[Feigin, 1987].

The Guinier relation links the scattered intensity to both parameters. It is based on the fact that the pattern is an even function ($I(s) = I(-s)$) and can be expanded in a series of even powers of *s*:

$$I(s) = \sum_{n=0}^{\infty} a_n s^{2n} = a_0 + a_1 s^2 + \ldots = I(0)\left[1 - \frac{1}{3}R_g^2 s^2 + O(s^4)\right] \approx I(0)\exp\left(-\frac{1}{3}R_g^2 s^2\right) \tag{6.16}$$

The last approximation is only valid for

$$\frac{1}{3}R_g^2 s^2 << 1 \quad \Leftrightarrow \quad s << \frac{\sqrt{3}}{R_g}. \tag{6.17}$$

This can be derived from

$$I(s) = \int\int \rho(\vec{r}_1)\rho(\vec{r}_2)\frac{\sin(sr_{12})}{sr_{12}} d\vec{r}_1 d\vec{r}_2\,, \tag{6.18}$$

with $r_{12} = |\vec{r}_2 - \vec{r}_1|$, if the $\frac{\sin x}{x}$-term is expanded as

$$\frac{\sin(sr_{12})}{sr_{12}} = \frac{sr_{12} - \frac{1}{6}(sr_{12})^3 + \ldots}{sr_{12}} = 1 - \frac{1}{6}(sr_{12})^2 + \ldots. \tag{6.19}$$

Then the intensity becomes

$$I(s) = \underbrace{\int\int \rho(\vec{r}_1)\rho(\vec{r}_2) d\vec{r}_1 d\vec{r}_2}_{I(0)} - \frac{s^2}{6}\int\int \rho(\vec{r}_1)\rho(\vec{r}_2)(\vec{r}_2 - \vec{r}_1)^2 d\vec{r}_1 d\vec{r}_2 + \ldots. \tag{6.20}$$

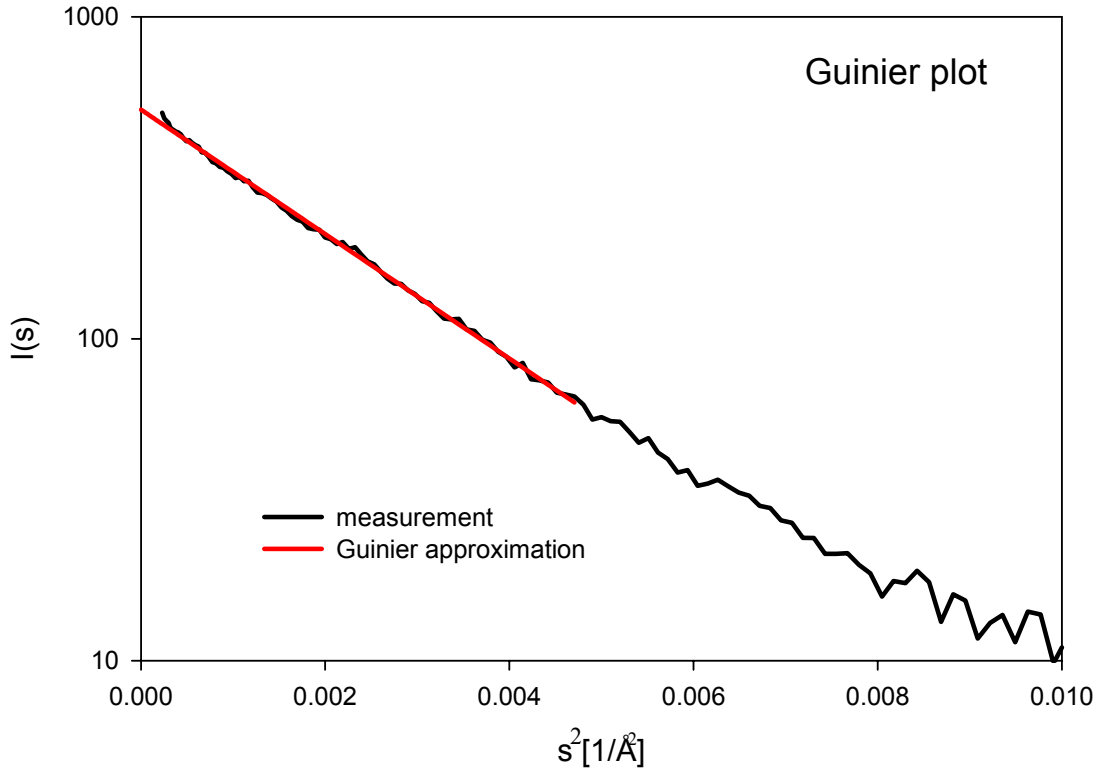


Figure 6.10: Guinier plot of glucose oxidase at pH 4.5 at 3.2 mg/ml. According to the fit the radius of gyration is 36.5 Å.

Together with Eq. (6.15) and $I(0) = \int\int \rho(\vec{r}_1)\rho(\vec{r}_2)d\vec{r}_1 d\vec{r}_2$ the Guinier relation follows.

$$I(s) = I(0)\left(1 - \frac{s^2}{3}R_g^2 + \ldots\right) \approx I(0)\exp\left(-\frac{R_g^2 s^2}{3}\right). \tag{6.21}$$

The intercept of the Guinier plot ($\log(I)$ vs. s^2) gives $I(0)$, which is directly related to the molecular mass, and its slope yields the radius of gyration (Figure 6.10).
In monodisperse solutions at low concentrations the scattered intensity of the solution is the sum of the scattering patterns $i_1(0)$ of the individual particles.

$$I(s) = N \cdot i_1(s) \tag{6.22}$$

With N molecules in the sample and

$$i_1(0) = \left\langle A(\vec{0}) \cdot A^*(\vec{0}) \right\rangle_\Omega = \int_V \int_{V'} \Delta\rho(\vec{r})\rho(\vec{r}')dr^3 dr'^3 = \overline{\rho}^2 V^2 = \Delta m^2 = \left(m - m_0\right)^2, \tag{6.23}$$

where m is the number of electrons in the particle and m_0 the number of electrons in the water, which is displaced by the molecule. From Eq. (6.22) it follows that

$$I(0) = N \cdot i_1(0) = Nm^2\left(1 - \frac{m_0}{m}\right)^2 = Nm^2\left(1 - \rho_0\psi\right)^2 \tag{6.24}$$

$$\Leftrightarrow \qquad \frac{I(0)}{c} = \frac{N_A M}{\mu^2}\left(1 - \rho_0\psi\right)^2. \tag{6.25}$$

The concentration $c = \frac{N\mu m}{N_A}$ is used to derive Eq. (6.25), where $\mu = \frac{M}{m}$ (≈ 1.87 for proteins) is the ratio of molecular mass to the number of electrons in the sample, ρ_0 is the

average electron density of the solvent, ψ the ratio of the volume of the particle to its number of electrons ($\Rightarrow \rho_0 \psi = \frac{m_0}{m}$) and N_A Avogadro's number. Note that the scattering in the *s*-origin, normalised to the sample concentration, is proportional to the molecular mass *M*.

$$\frac{I(0)}{c} \propto M \,, \tag{6.26}$$

Provided that the set-up was calibrated with a known standard it is possible to determine the molecular mass of any particle in a monodisperse solution. At X33 a fresh solution of Bovine Serum Albumin (BSA, concentration usually around c_{BSA}=5 mg/ml) is used as a standard. Provided that the proteins were measured at equal contrast and given that the molecular mass of BSA is M_{BSA}=66 kDa:

$$M = M_{BSA} \frac{I(0)}{I_{BSA}(0)} \cdot \frac{c_{BSA}}{c} \,. \tag{6.27}$$

The Fourier transform of the spherical averaged scattering intensity of the particle is the characteristic function

$$V\gamma(r) = \frac{1}{2\pi^2} \int_0^\infty s^2 I(s) \frac{\sin(sr)}{sr} \mathrm{d}s \,. \tag{6.28}$$

For a homogenous particle, $\gamma(r)$ gives the probability of finding a distance between $\vec{r}$ and $\vec{r} + d\vec{r}$ in the volume *V* of the particle (Figure 6.11). $V\gamma(r)$ is related to the histogram of

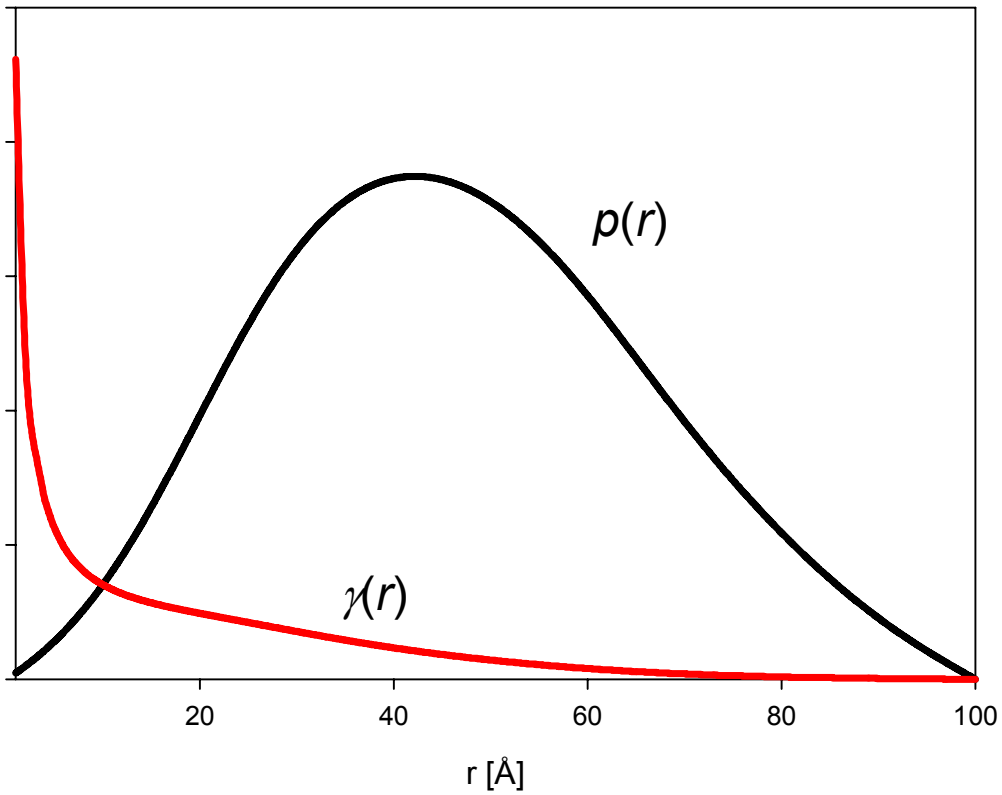


Figure 6.11: : Spherically averaged autocorrelation function *p*(*r*) (calculated with GNOM) and characteristic function $\gamma(r)$ of glucose oxidase at pH 4.5. The maximum diameter for the GNOM analysis was 100 Å, the radius of gyration was found to be 35.5Å.

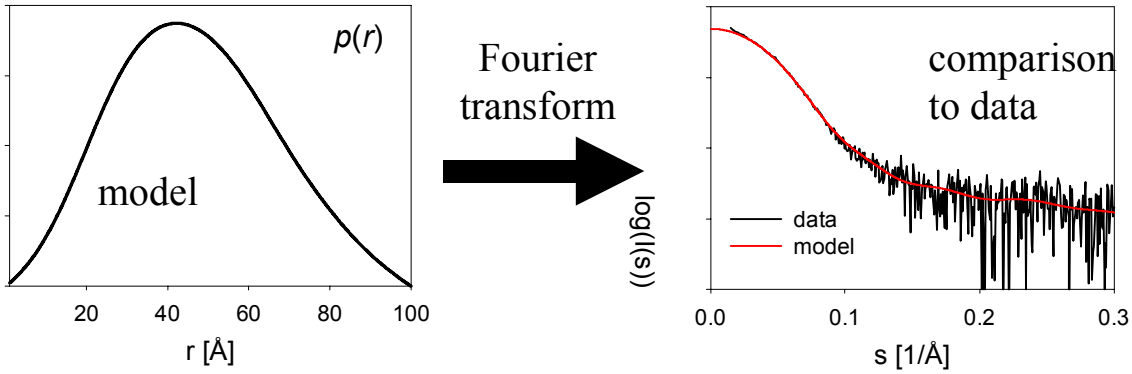


Figure 6.12: Indirect transform method. Since $I(s)$ is in practice only measured at a finite number of points and over a limited s-range with statistical and systematic errors, this method is preferable.

distances, the spherical averaged autocorrelation function $p(r)$, between the atoms of a homogenous particle (Figure 6.11):

$$p(r) = r^2 V\gamma(r) = \frac{1}{2\pi^2}\int_0^\infty srI(s)\sin(sr)\mathrm{d}s \tag{6.29}$$

$$\Leftrightarrow \qquad I(s) = 4\pi\int_0^\infty p(r)\frac{\sin(sr)}{sr}\mathrm{d}r = 4\pi\int_0^{D_{\max}} p(r)\frac{\sin(sr)}{sr}\mathrm{d}r \tag{6.30}$$

Eq. (6.30) illustrates that $p(r)$ can in principle be calculated as the Fourier transform of $I(s)$. Since $I(s)$ is in practice only measured at a finite number of points and over a limited s-range with statistical and systematic errors, the so-called indirect transform method is preferable. Here $p(r)$ is represented by a sum of orthogonal function, which is fitted to the experimental data [Svergun, 1992]. In order to minimise the effect of statistical errors the functional with the mean-square deviation is complemented by an additional term, which ensures the smoothness of $p(r)$. This term includes a parameter α, which controls whether the emphasis lies on smoothness or exact agreement between model and measurement. The fitted $p(r)$ can be used to obtain $I(0)$ from Eq. (6.30) and also R_g (Figure 6.12):

$$R_g^2 = \frac{\int r^2 p(r)\mathrm{d}r}{2\int p(r)\mathrm{d}r}. \tag{6.31}$$

Compared to the use of the Guinier relation the method using $p(r)$ is more accurate since the entire scattering curve is used and not just the onset. $p(r)$ is also the basic requirement for more sophisticated approaches to SAXS data based shape determination like 3D modeling.

The program GNOM [Svergun, 1992] was used for the determination of $p(r)$, $I(0)$ and R_g by the indirect transform method for all results presented here.

6.5 Data sampling

The readout electronics of the detector samples the scattering pattern in 1024 channels corresponding to an active detector area of 20 cm. With a camera length of 2.4 m this corresponds in practice to an s-range of approximately $0.015\,\text{Å}^{-1} \le s \le 0.4\,\text{Å}^{-1}$. An important question when it comes to 3D modelling of the particles is: What is the minimum sampling interval in the scattering curve needed to avoid any loss of information?

Fourier transforms are an important tool to answer this question as illustrated here for a time dependent function. The Fourier transform of a time-dependent function is the histogram of frequencies that characterise the process. The following example demonstrates the impact of the sampling frequency on an observation: The outcome of the observation of the movement of the minute hand of a watch clearly depends on the sampling frequency of the observation (Figure 6.13). If one samples with a interval of 5 minutes one would obtain the correct result. The hand turns clockwise with a period of one hour. If one would look only every 55 minutes on the watch one would come to the conclusion that the hand turns counter-clockwise with a period of 12 x 55 minutes. If one would sample with a period of 65 min one would end up with the correct direction of movement but with a period of 12 x 65 minutes. The reason for the incorrect observations in the last two measurements lies in the insufficient sampling rate, the true frequency of the watch hand is $\frac{1}{60} \cdot \frac{1}{\text{min}}$.

In information theory these finding are formulated more precisely and are referred to as the Nyquist-Shannon sampling theorem ([Oppenheim, 1989]). If a function $f(x)$, which is defined over a finite interval has a Fourier transform

$$\text{FT}\left[f(x)\right] = \hat{f}(k) = 0 \quad \text{for } |k| > k_{\max}, \tag{6.32}$$

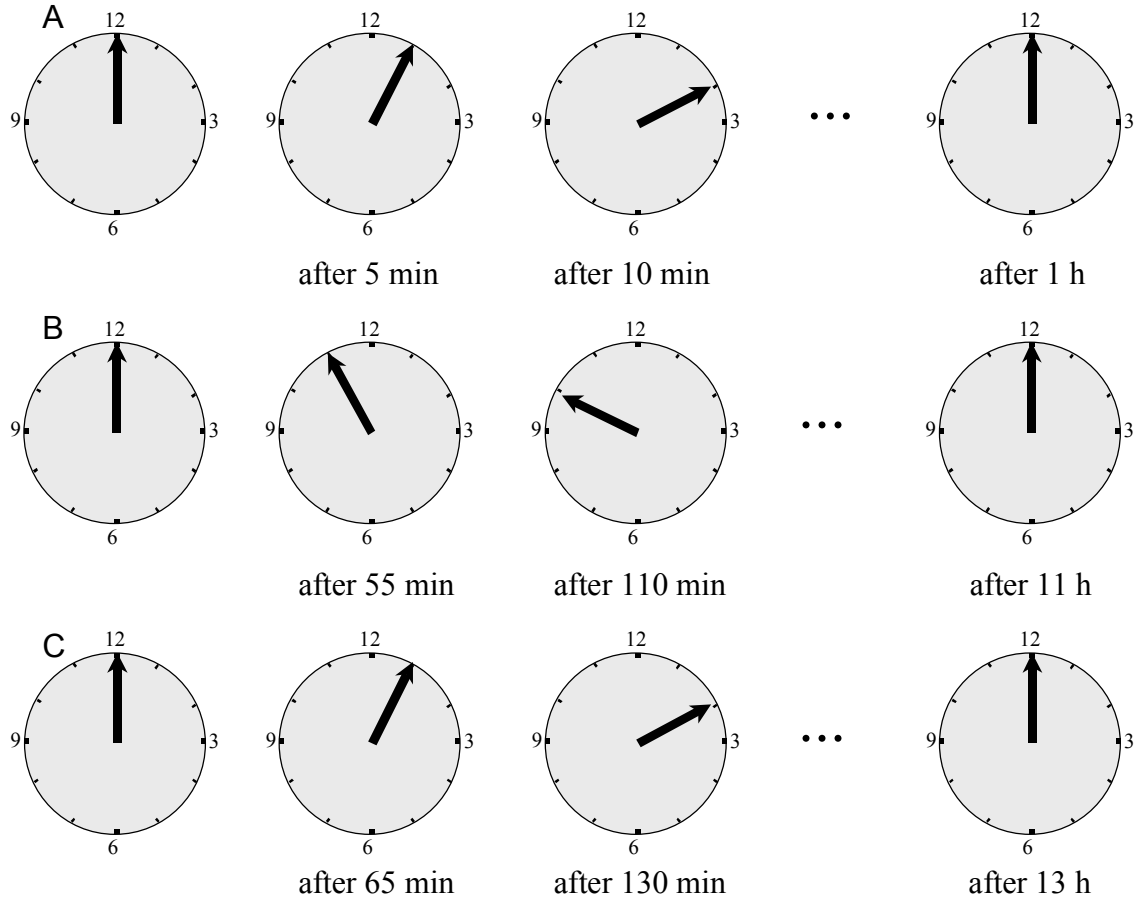


Figure 6.13: The result of the measurement depends on the sampling frequency. In series A with a sampling frequency of 5 minutes the correct result is achieved (clockwise movement with a period of 1 h). Too long sampling intervals of 55 and 65 min lead to wrong results.

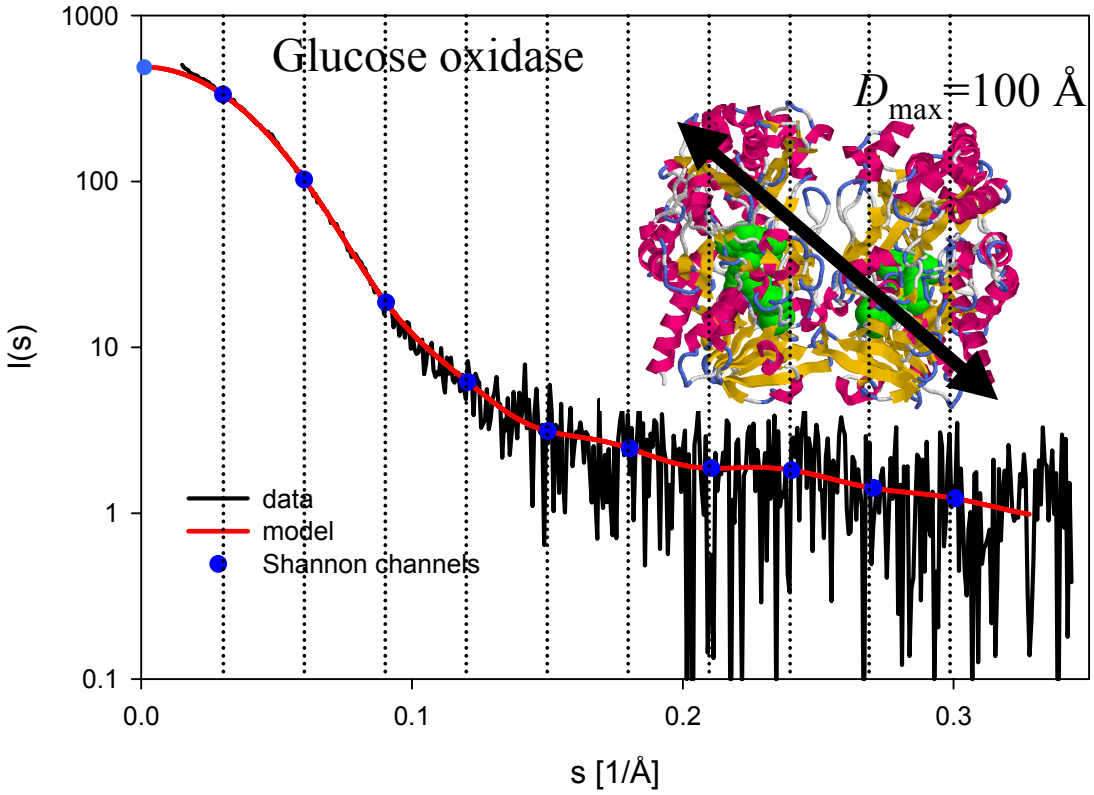


Figure 6.14: Shannon channels (spacing 0.03 Å^{-1}) for a SAXS glucose oxidase measurement. The measurement is thus oversampled.

then it is completely determined by giving the value of the function at a series of points spaced $\frac{1}{2k_{\max}}$ apart. The values $f_n = f\left(\frac{n}{2k_{\max}}\right)$ are called the samples of $f(x)$. This theorem can not only be applied to the time-frequency space but also the real-reciprocal space domain that is used in scattering (Figure 6.14). Thus the sampling rate of the scattered intensity is determined by the maximum non-zero value $D_{\max}$ (maximal distance between two atoms within the molecule) of its Fourier transform, the spherical averaged autocorrelation function $p(r)$. The continuous function I(s) can be represented by its values at a discrete set of points

$$s_k = \frac{k\pi}{D_{\max}}. \tag{6.33}$$

Consequently the number of channels N on an interval Δs is given by

$$N = \frac{D_{\max}\Delta s}{\pi}. \tag{6.34}$$

The $D_{\max}$ of the protein glucose oxidase is approximately 100 Å. The spacing between the samples is thus ca. 0.03 Å^{-1} and the number of samples on the interval $0\,\text{Å}^{-1} \le s \le 0.4\,\text{Å}^{-1}$ that corresponds to a typical range of a SAXS measurement is 13. The number of detector channels is typically around 1000. The measurements are thus oversampled. One must make sure however to measure the pattern at sufficiently small angles such that $s_{\min} < \frac{\pi}{D_{\max}}$.

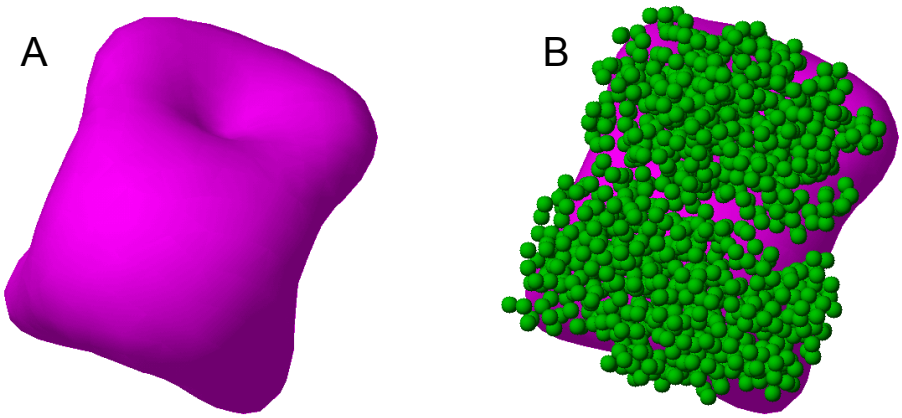


Figure 6.15: Envelope function (A) of glucose oxidase calculated with the program SASHA. In (B) the model is compared to the bead representation of the crystal structure.

6.6 Shape determination

The earliest attempts at ab initio 3D modelling were based on spherical harmonics, which are suitable because these functions form an orthogonal set [Stuhrmann, 1970]. This method leads to envelope functions of the particles (Figure 6.15). Other techniques can also be used to get models for the envelope of the molecules. The disadvantage of angular envelope functions is that they cannot represent models with complicated shapes (e.g. internal cavities). More complex models can be made with beads, unfortunately the number of parameters increases dramatically [Svergun, 1999]. By taking into account the value of $D_{\max}$, known from analysing the data with the indirect transform method, the model is restricted to a sphere with radius $\frac{D_{\max}}{2}$. For a realistic model it is also necessary to include the hydration shell of the particle. As when fitting $p(r)$ to the scattering pattern $I(s)$ a residual function is defined, which has to be minimized. Conventional minimisation algorithms are not suitable because the number of parameters is too large and the time needed would be too long. Instead one uses Monte Carlo based search algorithms like simulated annealing. The original method [Chacon, 1998], implemented in the program DALAI_GA, is based on a genetic algorithm. The more general 'dummy atoms' technique [Svergun, 1999] generates a sequence of random configurations. The current base configuration is randomly changed to obtain a new configuration. If the new configuration has a lower residual it becomes the new base configuration. Even if the residual value of the new configuration is larger there is a certain probability that it becomes the next base configuration. During the minimisation process this probability for a change to a configuration with a higher residual decreases. Consequently the movement of the base configuration through the space of configurations is similar to Brownian motion. In the beginning, at high temperatures (i.e. changing probabilities) the movement is nearly random. During the minimisation process the system cools down and the changes becomes smaller.
A general problem of many-parameter models is that the number of parameters by far exceeds the number of Shannon channels. Even if the data are neatly fitted, it is difficult to discuss the uniqueness of such a model.

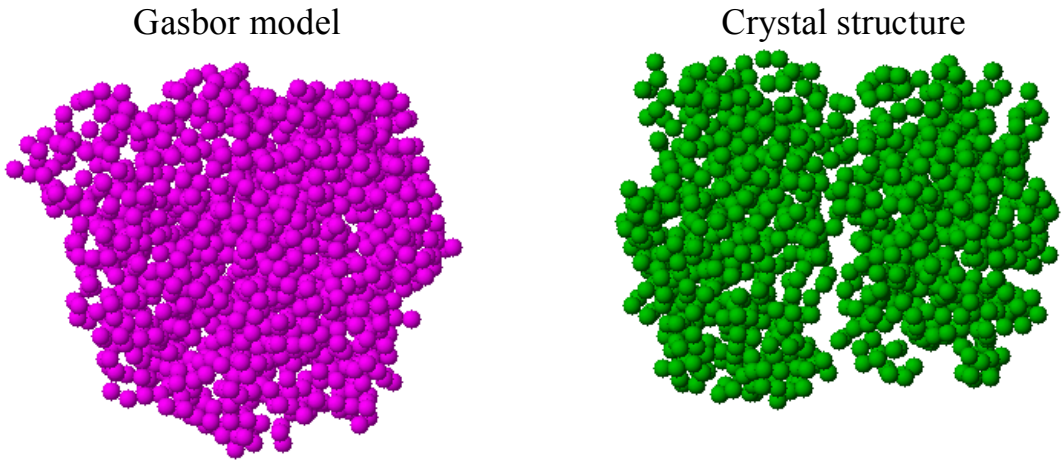


Figure 6.16: Comparison of glucose oxidase crystal structure and SAXS model calculated with GASBOR.

The above-mentioned bead model is only one example of a many-parameter system. The accuracy of the modelling can be improved by including more a priori knowledge about the particles. In the case of a protein it is known that a protein is a chain of amino acids. Since in most cases the primary structure of the protein is known the number and the average scattering of amino acid residues (dummy residues) also provide information one can make use of (Figure 6.16; [Svergun, 2001]).

6.7 Interactions sensed with SAXS

In the previous sections the small angle scattering signal of protein solutions was discussed and it was explained how it can be used to obtain information about the solute particles. One prerequisite for using these methods is that the concentration of the monodisperse protein solution should be relatively low and the scattering to be that of isolated particles. Given the dimension of the coherence volume (7500Å x 4000Å x 8000Å) at high concentrations, 5×10^5 particles are in the coherence volume (for 50 mg/ml and a 14.3 kDa protein) and the photons scattered by these particles interfere coherently (Figure 6.17). If the distribution of proteins within the sample is completely random the scattering pattern will be identical to that of the isolated particle. In reality the distribution will never be random because this would mean that there is a non-zero probability that two proteins overlap. This is never the case due to the repulsive interactions between overlapping electron orbitals (hard-sphere interactions). If the relative positions of proteins are not random due to their interactions, it means that there is a certain order resulting in interference of the scattered waves, which have an impact on the scattering pattern. If the distances between molecules remain sufficiently large and their relative orientations

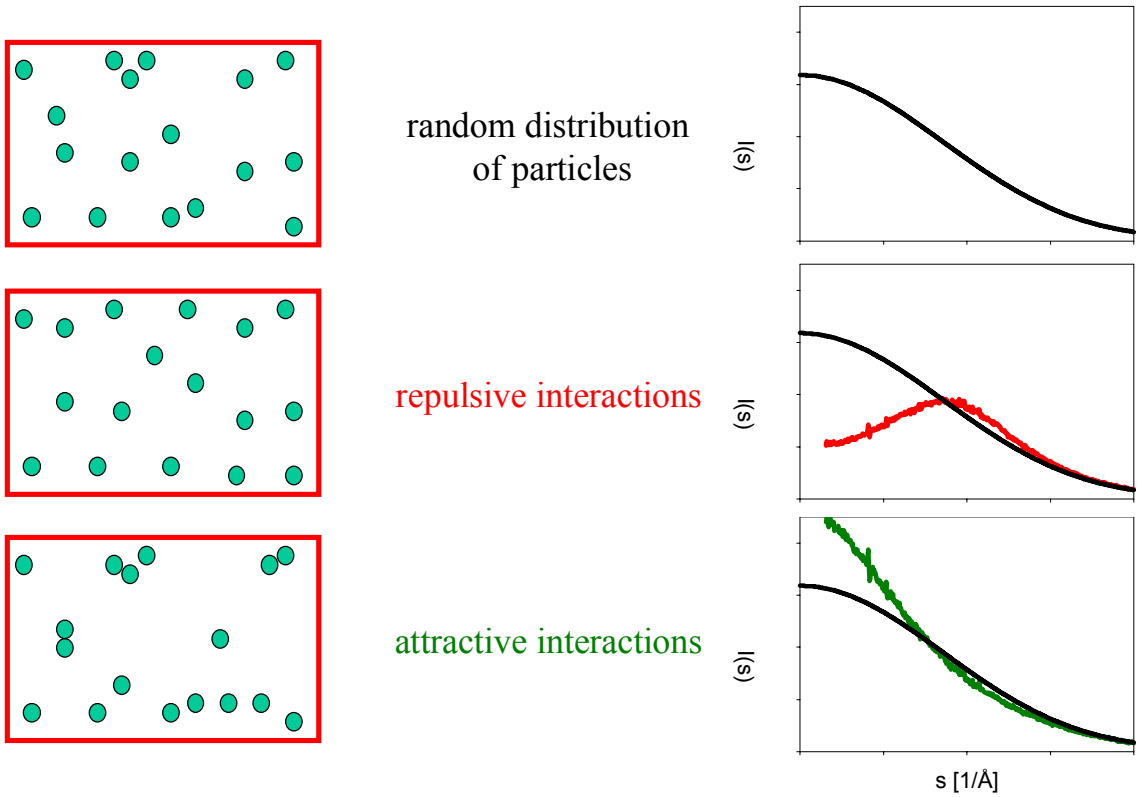


Figure 6.17: Effect of the interactions between proteins on the SAXS pattern. If the distribution of proteins within the sample is completely random the scattering pattern will be identical to that of the isolated particle.

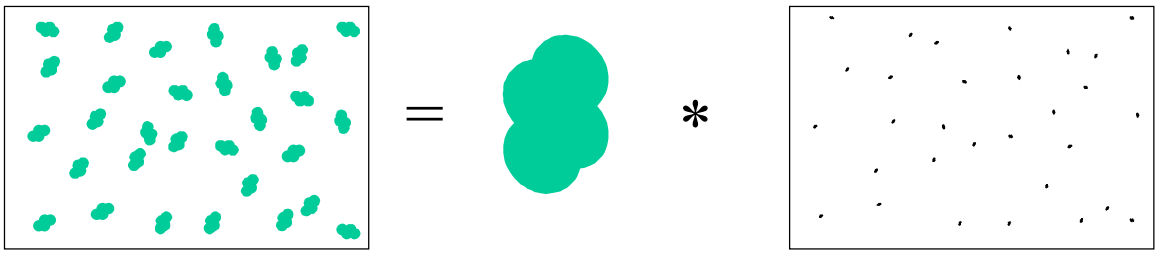
random, the scattering density of the solution can be represented as a convolution of the scattering density of a single particle averaged over all orientations and a distribution of delta functions corresponding to the centres of the individual particles (Figure 6.18).
The spherically averaged scattering of a single particle is called the form factor (FF) and corresponds to the scattering pattern of a dilute solution due solely to intra-protein interference. At high concentrations c inter-protein interference has to be taken into account and thus the scattered intensity is a product of the form factor and the so-called structure factor, which corresponds to the structure of the solution:

$$I(c,s) = \mathrm{FF}(s) \cdot \mathrm{SF}(c,s)\,. \tag{6.35}$$

All interactions between macromolecules have an impact on the structure factor. From the theory of liquids it is known that the structure factor is related to the radial pair distribution function $g(r)$ (Figure 5.4, 5.5). This function is a measure of the probability of finding another molecule with its centre in a distance r from the first one. For a hypothetical ideal gas $g(r)=1$ over the entire r-range. In a real gas, as discussed above, the molecules cannot interpenetrate and thus the probability of finding two particles of radius R within a distance of $r<2R$ is zero. In gases and liquids only short-range order exists and thus $g(r)$ differs from 1 for small r-values. In a perfect crystal there is long-range order and $g(r)$ consist of an array of delta functions at the lattice points and is equal to zero everywhere else.
In the case of a homogeneous fluid the structure factor can be written as the Fourier transform of $g(r)$.

$$\mathrm{SF}(c,s) = 1 + \rho \int_0^\infty 4\pi^2 r^2 \left(g(r) - 1\right) \frac{\sin(rs)}{rs}\,\mathrm{d}r\,, \tag{6.36}$$

where ρ is the number density of particles in the solution.



$$\rho(\vec{r}) = \rho_P(\vec{r}) * \sum \delta(\vec{r}_i)$$

$$I(c,s)=I(0,s)\mathrm{SF}(c,s)$$

Figure 6.18: The scattering density of the solution can be represented as a convolution of the scattering density of a single particle averaged over all orientations and a distribution of delta functions corresponding to the centres of the individual particles. Thus the scattered intensity is a product of the form factor and the so-called structure factor, which corresponds to the structure of the solution.

The inverse relation of Eq (6.36) facilitates the calculation of $g(r)$ from the scattering factor.

$$\rho g(r) = \frac{1}{2\pi}\int_0^\infty s^2\left(\mathrm{SF}(c,s)-1\right)\frac{\sin(rs)}{rs}\mathrm{d}s\,, \qquad (6.37)$$

A pair potential $u(r)$ gives the simplest description of interactions between two objects. A relation between the spherically averaged pair potential and the radial distribution function would thus be helpful for the investigation of protein interactions as it would facilitate the comparison of the pair potential and the scattering measurements. Unfortunately, direct analytical approximations could only be found in a few special cases. Nevertheless some procedures for numerical solutions exist.

In order to find an approach it is necessary to introduce the pair correlation or total correlation function $h(r)$ [Hansen, 1986]:

$$h(r) = g(r) - 1\,. \qquad (6.38)$$

Thus the total correlation function is zero everywhere where $g(r)$ is 1. If $h(r_0)$=0 the presence of a particle in the r-origin does not influence the probability of finding one at a distance r_0. Two particles separated by r_0 are not correlated. In a liquid, in which only short range order exists, there is always an r_{max} such that for all $r>r_{max}$ $g(r)$=1 and thus $h(r)$=0. Integrals over $h(r)$ only have to be calculated up to r_{max}. The range of the potential does not necessarily restrict the range of the correlation of particles. The example of a crystal illustrates that short-range interactions can lead to long range order. In order to find a relation between interactions and the total correlation function it is useful to introduce the direct correlation function c(r), which only represents the correlation due to the direct interactions between two particles. It is reasonable to suppose that the range of $c(r)$ is comparable to that of the pair potential $u(r)$. The differences between the functions h(r) and c(r) for a typical hard-sphere liquid are illustrated in Figure 6.19.

The total correlation function can then be written as a sum of the direct correlation and the correlation mediated through the presence of all other particles in the volume. This relation is known as the Ornstein-Zernicke (OZ) relation:

$$h(r) = c(r) + \rho \int c(|\vec{r} - \vec{r}'|) h(r') d\vec{r}' . \tag{6.39}$$

It is the fundamental equation for the pair distribution function of classical many-particle systems. On taking the Fourier transform of this equation, one obtains

$$\hat{h}(s) = \frac{\hat{c}(s)}{1 - \rho \hat{h}(s)} . \tag{6.40}$$

Also one can show that

$$SF(s) = 1 + \rho \hat{h}(k) = \frac{1}{1 - \rho \hat{c}(s)} . \tag{6.41}$$

The Ornstein-Zernicke relation contains two unknown functions and thus can not be used to calculate a solution. A exact closure relation of the OZ relation exists:

$$g(r) = \exp\left(-\beta u(r) + h(r) - c(r) + E(r)\right) . \tag{6.42}$$

$E(r)$ is an infinite sum of multicentre integrals, the so-called bridge diagrams, which are known in principle as complicated multidimensional integrals and very hard to evaluate. Instead simple approximations are used like the hypernetted-chain (HNC) approximation ($E(r)$=0). Another commonly used approximation is the Percus-Yevick approximation,

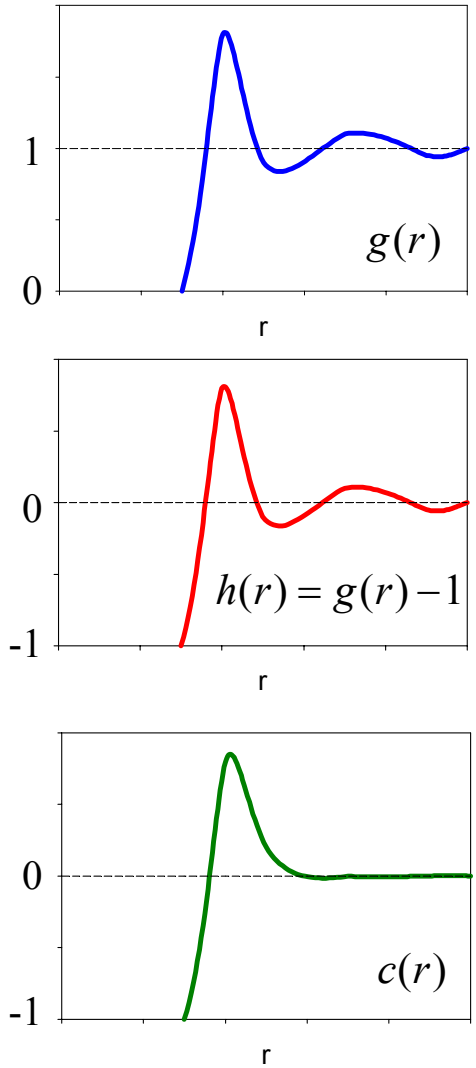


Figure 6.19: Comparison of the radial pair distribution function $g(r)$, the total correlation function $h(r)$ and the direct correlation function $c(r)$, which only represents the correlation due to the direct interactions between two particles.

which leads to good results. For short-range potentials (e.g. hard-sphere potential), an analytical solution can be obtained. The HNC approximation is superior for long-range potentials like electrostatic potentials with charge screening as used for ionic liquids [Hansen, 1986]. The HNC approximation was therefore used in previous studies of the interactions of proteins and in all calculations in the present work.
Even for the HNC approximation analytical solutions for the pair potentials are not known so, the integral equations described above must be solved by numerical methods. It is useful to introduce the function

$$\gamma(r) = h(r) - c(r)\,. \tag{6.43}$$

For the HMC approximation (6.43) then becomes

$$c(r) = h(r) - \gamma(r) = \big(g(r) - 1\big) - \gamma(r) = \exp\big(-\beta u(r) + \gamma(r)\big) - 1 - \gamma(r)\,. \tag{6.44}$$

A Fourier transform (FT) gives $\hat{c}(r)$. The OZ relation in Fourier space (6.40) is used to obtain $SF(s)$, $\hat{h}(r)$ and finally $\hat{\gamma}(r)$, which after another FT gives the next iteration of $\gamma(r)$.

6.8 Potential models used to calculate the structure factors

As mentioned in the introduction the interactions in a protein solution can best be described by a hard-sphere potential with a short-range attraction and a long-range repulsion according to the DLVO theory. For the sake of simplicity the successful

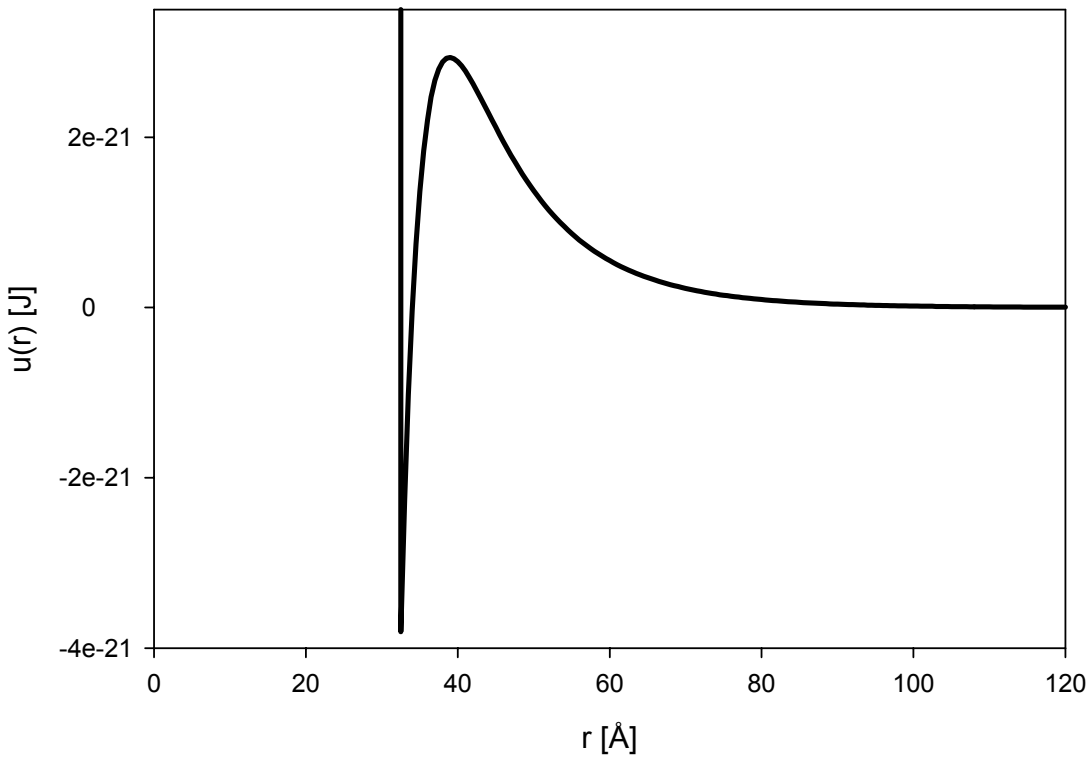


Figure 6.20: Yukawa pair potential according to DLVO theory as it was used by [Tardieu, 1999]. The parameters correspond to lysozyme at a concentration of 70 mg/ml at 15ºC. The depth of the attractive potential is 3.0 kT and its range is 3.0 Å. The effective charge was set to 6.5 and the Debye length was 13.42 Å. The hard sphere radius is 16.2 Å.

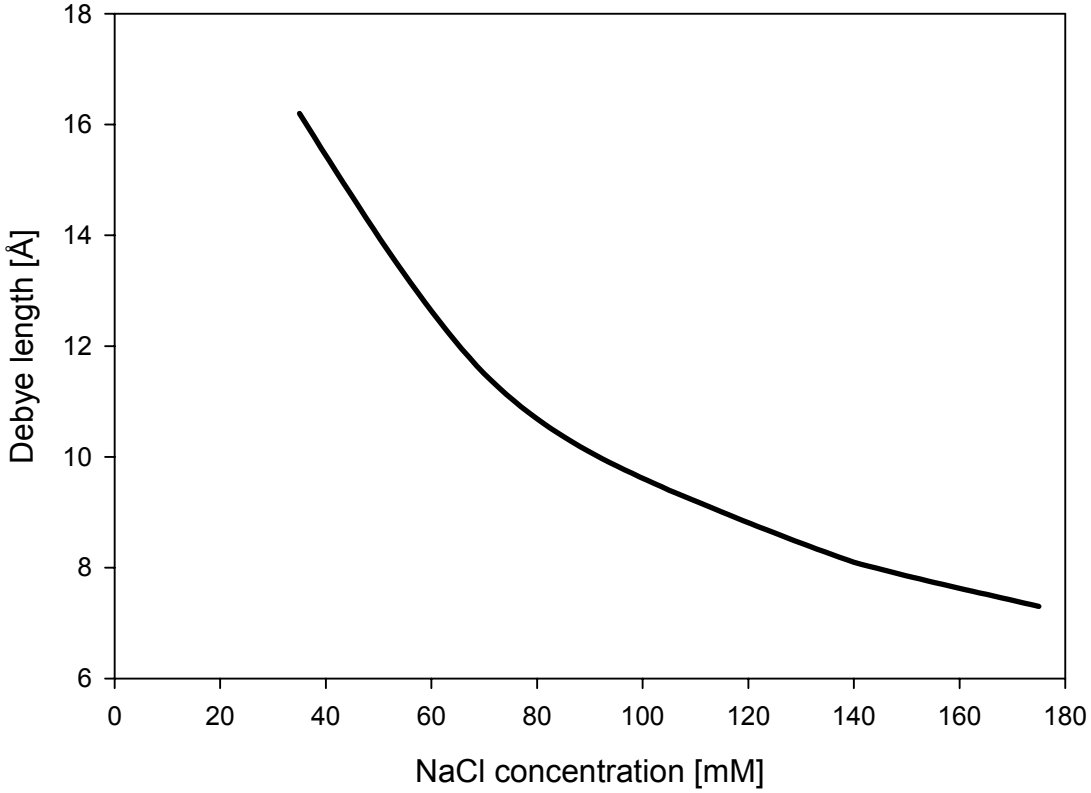


Figure 6.21: Debye length in a sodium chloride solution as a function NaCl concentration.

approach of Tardieu [1999] was used. Here the attractive as well as the repulsive potential were represented by a Yukawa potential (Figure 6.20). Thus the effective pair potential becomes

$$u(\vec{r}) = \begin{cases} \infty, & \text{if } |\vec{r}| \le 2R \\ J_a \frac{R}{r}\exp\left(-\frac{r-R}{d_a}\right) + J_r \frac{R}{r}\exp\left(-\frac{r-R}{d_r}\right), & \text{if } |\vec{r}| > 2R \end{cases} \quad (6.45)$$

J_a, d_a are the depth and range of the attractive potential and J_r, d_r are the corresponding values of the repulsive potential (charge z). According to the DLVO theory the parameters J_r, d_r can be expressed as functions of the protein charge Z_p and the Debye length λ_D, which itself is a function of the ionic strength $I = \frac{1}{2}\sum c_i Z_i^2$, where Z_i is the charge of the salt ions. The Debye length is a range parameter that becomes smaller when salt is added to the solution as a consequence of charge screening:

$$\lambda_D = \frac{1}{\sqrt{4\pi L_B \sum \rho_i Z_i^2}}, \quad (6.46)$$

where $L_B = \frac{e^2}{4\pi\varepsilon_0\varepsilon_s kT}$ is the so-called Bjerrum length of the solvent, which is 7.2 Å at 300K. According to [Israelachvili, 1992] the Debye length (Figure 6.21) in the case of monovalent ions such as NaCl, can also be written at 25ºC as

$$\lambda_D = \frac{3.04}{\sqrt{[\mathrm{NaCl}]}}\,\text{Å}. \tag{6.47}$$

The repulsive potential depth J_r is assumed to be

$$J_r = \frac{z^2}{2R}\cdot\frac{L_B}{\left(1+\frac{R}{\lambda_D}\right)^2}. \tag{6.48}$$

In the DLVO theory the attractive van der Waals potential

$$v(r) = -\frac{A}{12}\left[\frac{1}{x^2-1}+\frac{1}{x^2}+2\ln\left(1+\frac{1}{x^2}\right)\right] \tag{6.49}$$

is formulated using $x = \frac{r}{2R}$ and the Hamaker constant A, which depends on the relative permittivity $\varepsilon_p, \varepsilon_s$ and refractive index n_p, n_s of protein and solvent.

$$A = \frac{3kT}{4}\cdot\frac{\left(\varepsilon_p-\varepsilon_s\right)^2}{\left(\varepsilon_p+\varepsilon_s\right)^2}+\frac{3h\nu_e}{16\sqrt{2}}\cdot\frac{\left(n_p^2-n_s^2\right)^2}{\left(n_p^2+n_s^2\right)^{3/2}}. \tag{6.50}$$

ν_e is the main electronic absorption frequency (typically $3\times10^{15}\,\mathrm{s}^{-1}$) and h is Planck's constant.

6.9 Measurement of the SAXS pattern and processing of experimental data

The instrumentation of a SAXS beam line is described in chapter 2. It was already pointed out that it is necessary to measure and correct for the detector response DR(n) and to determine the calibration $s(n)$ of the s-axis, which assigns an s-value to each electronic detector channel n (time-bin of the delay line readout system).
In this section the procedure for measuring the small angle scattering pattern is interpreted in the formalism of contrast variation.
The scattering of protein in solution (index sol) and the scattering pattern of the buffer (index buf) only are determined in two independent measurements. The scattering pattern $I_{\mathrm{exp}}(n)$ of the protein is obtained by subtracting the scattering of the buffer from that of the solution (Figure 6.22).

$$I_{\mathrm{exp}}(n) = \frac{1}{c}\cdot\frac{1}{\mathrm{DR}(n)}\left[\frac{N_{\mathrm{sol}}(n)}{I_{0,\mathrm{sol}}}-\frac{N_{\mathrm{buf}}(n)}{I_{0,\mathrm{buf}}}\right], \tag{6.51}$$

$N(n)$ is the number of photons counted in the nth channel of a pattern, c is the protein concentration. Dividing by the intensity I_0 of the transmitted beam corrects for the decay of the primary beam intensity and the absorption of the sample.

NaCl [mM]	Debye length λ_D [Å]	NaCl [mM]	Debye length λ_D [Å]
0	-	175	7.3
35	16.2	210	6.6
70	11.5	280	5.7
105	9.4	350	5.1
140	8.1		

Table 6.3: Debye length of NaCl solutions at 25ºC (Figure 6.21).

This is particularly useful when comparing scattering patterns of a solute in two differently absorbing buffers. Whereas 1 mm of water (optimal thickness) transmits 90.4% of the X-rays at 8 keV, a 250 mM solution of NaCl transmits only 80.8% and for KCl the transmission is only 72.3%.
All the experimental data $N(n)$ were processed by the program SAPOKO to obtain the scattering patterns [Konarev, 2003].

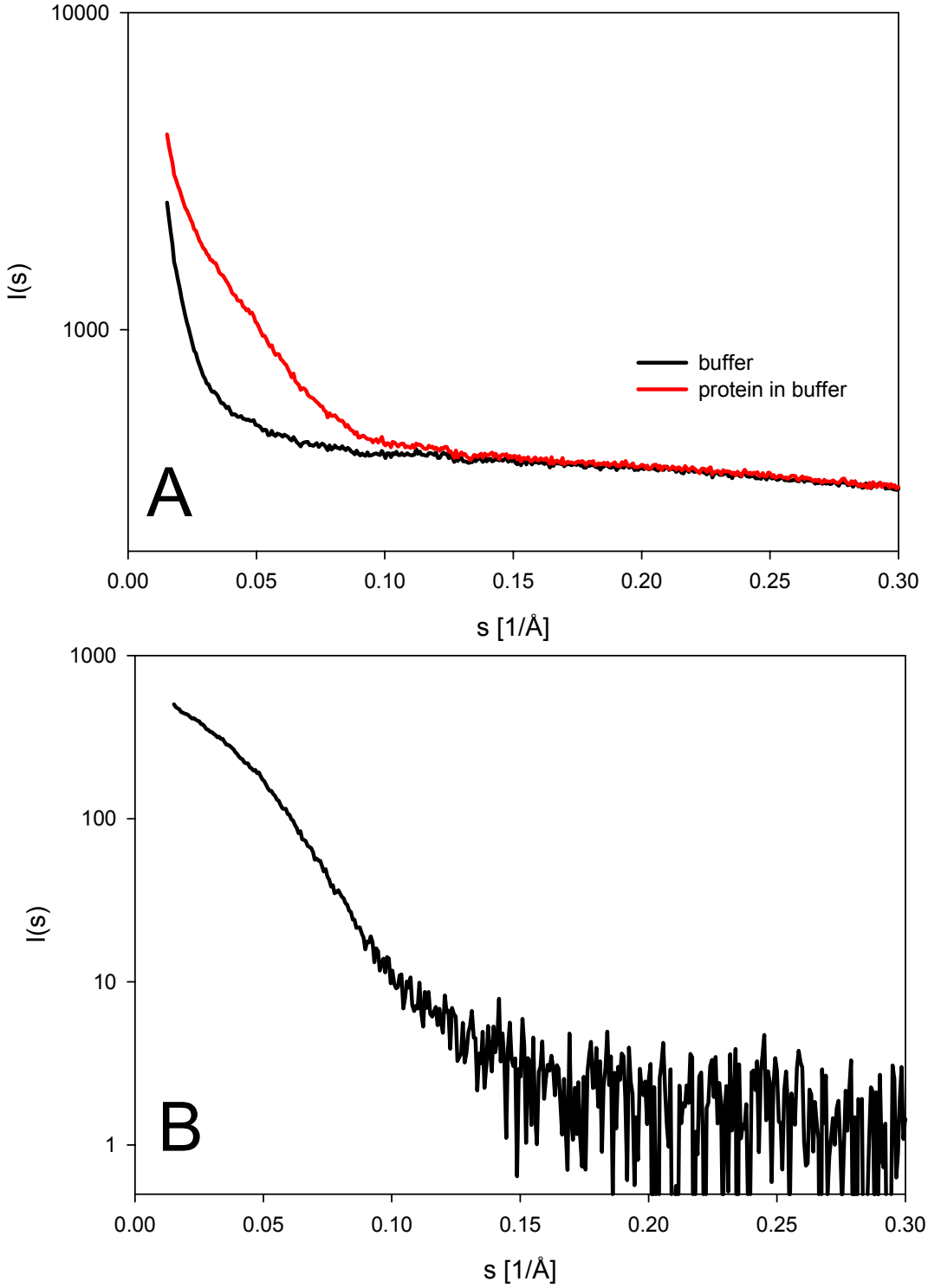


Figure 6.22: Buffer subtraction. (A) represents the scattering curves of buffer and the protein and the buffer (here glucose oxidase in 100 mM sodium phosphate at pH 4.5). (B) is the difference of the two curves calculated with Eq (6.51).

6.10 Filling and cleaning the sample cell

For a SAXS measurement it is necessary to measure buffer and sample as described in chapter 1. To obtain sufficient statistics and verify that the cell have been properly cleaned the buffer is measured before and after the sample. It is very important that the cell should be as clean as possible before injecting buffer or samples. The cell can only be accessed through three small holes for the needle of a syringe. If the cell is filled with a solution of protein A one removes it and washes it with the buffer of protein A (step 1). Afterwards the entire procedure must be repeated first with ethanol (step 2) and with the buffer of protein B (step 3), which is measured afterwards. Finally the buffer of protein B can be injected for the measurement. It is important that no bubbles remain in the cell since they would seriously affect the measurement. After the measurement the buffer must be removed. This time it is not necessary to clean the sample cell because the buffer is identical to the one surrounding the protein. It is however important to remove the buffer completely and dry the cell afterwards to avoid diluting the protein. Finally protein B solution is injected. After the measurement the protein has to be removed as described above and the buffer of protein B is measured again. In the cleaning process between two measurements of different buffers one has to replace the buffer in the first cleaning step by deionised water because this time the aim is to remove the salt ions of the buffer. This is particularly important when measuring protein interactions because the presence of even small amounts of salt (e.g. 5 mM) can significantly modify the interactions.
Under certain conditions some proteins tend to aggregate at the cell windows and even after carefully cleaning the cell they can not be removed. This can lead to differences in the buffer measurements and also make it difficult to fill the cell without bubbles. In such cases it may be necessary to exchange the cell windows to get rid of the contamination.

6.11 Statistical and systematic errors in a SAXS measurement

The most significant errors arise from the sample characterisation and the determination of the concentration. From experience it is known that 10% errors are typical for the concentration. During the process of sample preparation apart from the UV absorption measurement other sources of error come into play, the most frequent ones being pipetting errors (2-10%).
The statistical error of the detector data can be determined directly from the number of photons since the PSGPD is a single-photon counting device. The standard deviation is the square root of the number of counts. In a typical measurement at beam line X33 one measures the sample for 15 min to acquire more than 10000 counts in each of the bins (also at relatively high angles where the count rates are lower). The statistical error on each bin is thus less than 1%. The statistical errors of the detector response can also be calculated from the number of photons in the detector response measurement. Dividing the scattering pattern by the detector response is assumed to correct for any systematic error.
The statistical errors of the detector response (ΔDR), sample (ΔN_{sol}) and buffer measurements (ΔN_{buf}) together with the error in the normalisation ($\Delta I_{0,sol}$, $\Delta I_{0,buf}$) from Eq. (6.51) produce, following the rules of error propagation in the errors of the intensities (ΔI_{exp}):

$$\Delta I_{\exp}(n) = \frac{1}{c} \cdot \sqrt{\left(\frac{\Delta \mathrm{DR}(n)}{\mathrm{DR}(n)}\right)^2 + \frac{\left(\frac{N_{\mathrm{sol}}(n)}{I_{0,\mathrm{sol}}}\right)^2 \left[\left(\frac{\Delta N_{\mathrm{sol}}(n)}{N_{\mathrm{sol}}(n)}\right)^2 + \left(\frac{\Delta I_{0,\mathrm{sol}}}{I_{0,\mathrm{sol}}}\right)^2\right] + \left(\frac{N_{\mathrm{buf}}(n)}{I_{0,\mathrm{buf}}}\right)^2 \left[\left(\frac{\Delta N_{\mathrm{buf}}(n)}{N_{\mathrm{buf}}(n)}\right)^2 + \left(\frac{\Delta I_{0,\mathrm{buf}}}{I_{0,\mathrm{buf}}}\right)^2\right]}{\left(\frac{N_{\mathrm{sol}}(n)}{I_{0,\mathrm{sol}}} - \frac{N_{\mathrm{buf}}(n)}{I_{0,\mathrm{buf}}}\right)^2}}, \tag{6.52}$$

The intensity errors propagate to the errors for R_g and $I(0)$ and thus to the molecular mass and to errors in the spherically averaged autocorrelation function $p(r)$, which is determined via the inverse transformation method.
The calibration also leads to an error since usually the standard sample is fixed to the back of the sample cell (less than 5 mm from the exact sample position). On a camera length of 2400 mm this leads to a systematic error of at most 0.2 % on the calibration, which is small compared to the statistical errors.
The detectors used for the normalisation of the intensities, ion chamber and photodiode, are integrating detectors. Therefore their errors cannot simply be determined as square root of the number of photons. Their empirical variance was determined to be of the order of 0.01%.
In addition systematic errors may arise due to instabilities in the beam (in position, intensity and energy). Changes of the beam between the measurement of buffer and sample cause deviations in the difference of the two curves. Also a movement of the beam on the photo-diode leads to incorrect measurements of the intensity of the transmitted beam. The significance of these errors can vary from negligible to catastrophic. Beam stability is not correlated to the measurement so the reproducibility of the measurements is the check the beam stability.
In summary the characterisation of the samples (protein concentration) is the largest source of error (10%). The magnitude of the statistical errors is of the order of 1% and can be calculated precisely for the individual measurement. All other contributions (calibration, normalisation, beam stability) are of minor importance (<1%).

7 Results and Discussion

In this section the results of the SAXS measurements on proteins and simulations are presented and discussed. First, the structure factors calculated from our experimental data on lysozyme by the program winOZ are compared to those obtained previously by other groups under identical conditions. Secondly, using winOZ structure factors are calculated for various pair potentials. Systematic variation of individual parameters (protein concentration, hard sphere radius, effective charge, Debye length, range and depth of the attractive potential) illustrate the effect of the changes on the structure factors. Subsequently, results of the analysis of the measurements on the protein lysozyme are presented. The influence of various additives (TMAO, urea and glycine at 250 mM) was measured and special attention was paid to the differences in the short-range attractive potentials. Lysozyme was also measured at a different pH and therefore with a different net charge. Finally, the changes in the structure factor due to temperature variations (10-40ºC) and the corresponding changes of the pair potentials are presented.

7.1 Test of a program for calculating the structure factor from the pair potential

The program winOZ based on the formalism given by [Tardieu, 1999] was tested and the results were compared with those in the literature. These authors measured the scattering curve of lysozyme (c=74 mg/ml) at its pI (pH 10.5) at different temperatures, between 10ºC and 40ºC. In their analysis they assumed that only the attractive potential plays a role since the lysozyme molecules have no net charge at the pI. They found that values of $d_a = 3\,\text{Å}$ and $J_a = -2.65kT$ at T=20ºC yielded the best fit to the experimental data. A value of 16.2Å was found for the radius R of the hard sphere potential.
winOZ did not exactly reproduce the same curves with the above parameters (Figure 7.1). In order to get identical structure factors the range parameter d_a was set to 3.5Å instead of 3.0Å. The reason for this difference of 0.5 Å could be due to small differences in the numerical calculations like the sampling interval or more simply, a truncation in the original paper. The matter was not further investigated as all curves of the temperature series were correctly reproduced with the same values of the potential parameters.

HCl equivalent added	pH	λ_D [Å]	z_{eff}
0	10.5		
1	9.1	55	1
2	8.7	39	3
3	8.2	32	4.1
11	3.6	17	9
18	2.0	13	12.1

Table 7.1: Values of the Debye length and effective charge as a function of pH for a lysozyme solution (c=85 mg/ml) from [Tardieu, 1999].

In a further series [Tardieu, 1999] measured and calculated the structure factors at different pH values adjusted by addition of HCl. The parameters obtained in the temperature series were used for the attractive potential and the Debye length λ_D was calculated from the number of HCl molar equivalents added. The effective charge z_{eff} was then obtained by fitting the experimental data. Table 7.1 gives the results.
WinOZ gives similar results for exactly the same parameters (Figure 7.2). In the repulsive regime, where the long-range potentials govern the interactions, the simulations are in good agreement with the theoretical results in the literature.
The comparison of the results of winOZ and those obtained by [Tardieu, 1999] proved that despite the unexplained small difference in the range parameter d_a, the same results were obtained. The program can thus be used for the analysis of the structure factor measurements.

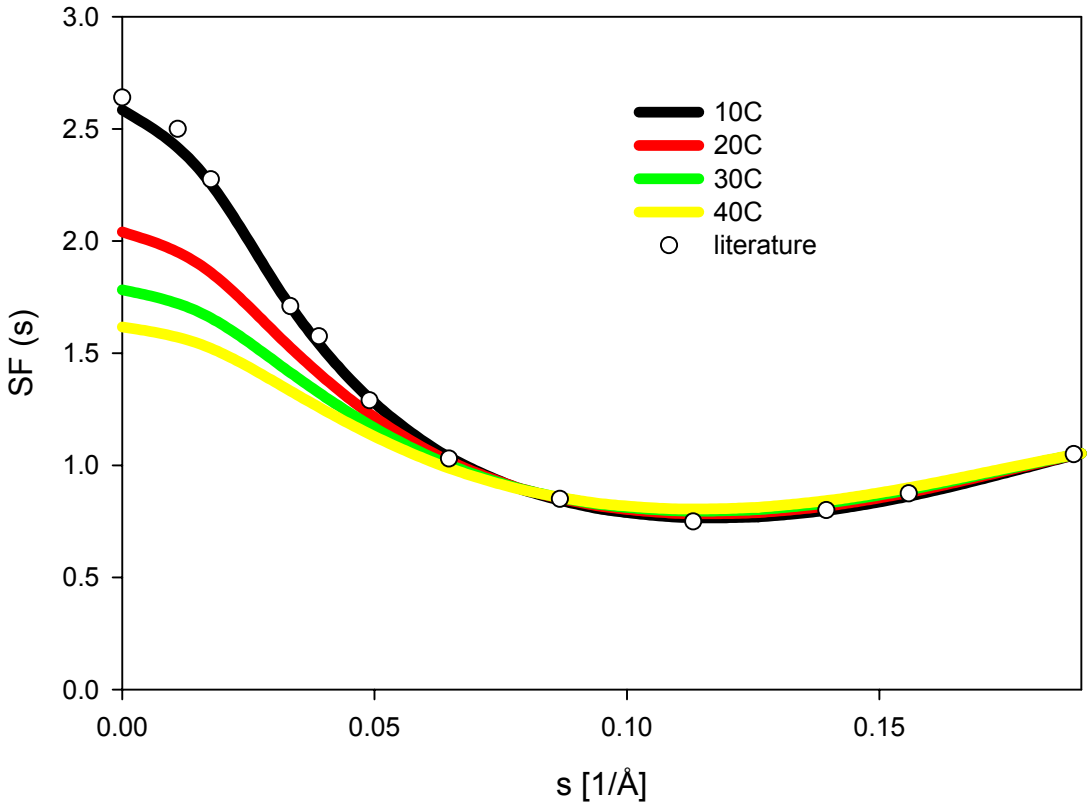


Figure 7.1: Simulation of the structure factors of lysozyme at 74 mg/ml at pI (pH 10.5, absence of repulsive electrostatic interaction) at 10, 20, 30, 40 ºC. The data points represent literature values [Tardieu, 1999] for the same conditions. In order to obtain the presented results the range of the attractive potential d_a was set to 3.5 Å whereas the value reported in the literature is 3 Å. The depth of the attractive potential was in both cases the same (-2.65 *kT* at 20ºC).

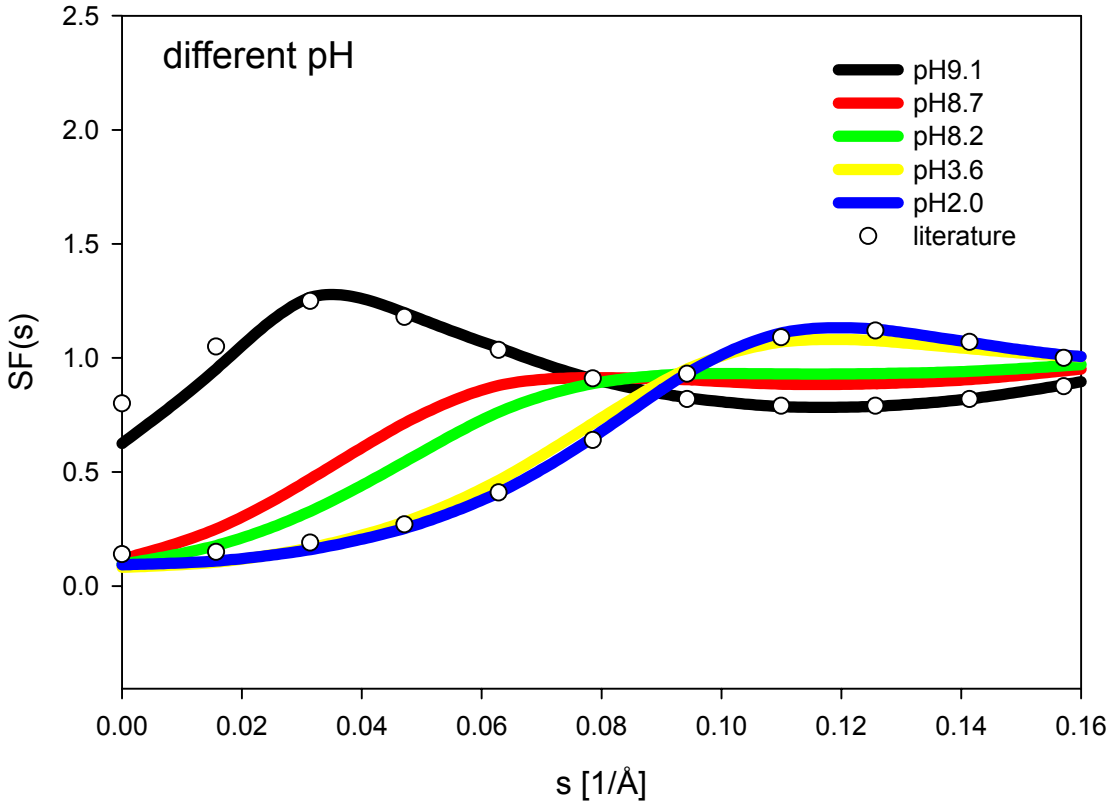


Figure 7.2: Simulated structure factors at different pH values. The data points are taken from corresponding simulations in the literature. The parameters for the simulation were identical. In the repulsive regime the simulation results are in good agreement with the literature.

7.2 Effects of the pair potential parameters on the structure factor

In order to illustrate the effects of the variation of the parameters of the pair potential the structure factors were calculated for different values of these parameters with the program winOZ. The simulation results can help to understand the complex relationship between the pair potentials and the solution scattering curves and accordingly the structure factors. To facilitate the comparison between variations of different parameters one structure factor is included in all graphs. This structure factor corresponds to a potential and protein

parameter	value for the "typical" structure factor
concentration	70 mg/ml
effective charge	6.5
Debye length	13.42 Å
depth of attr. potential	-3.0 kT (at 20ºC)
range of attr. potential	3.0 Å
hard sphere radius	16.2 Å
molecular mass	14.3 kDa (lysozyme)
temperature	15ºC

Table 7.2: Simulation parameters of the "typical" structure factor, which is shown in all plots to facilitate comparison.

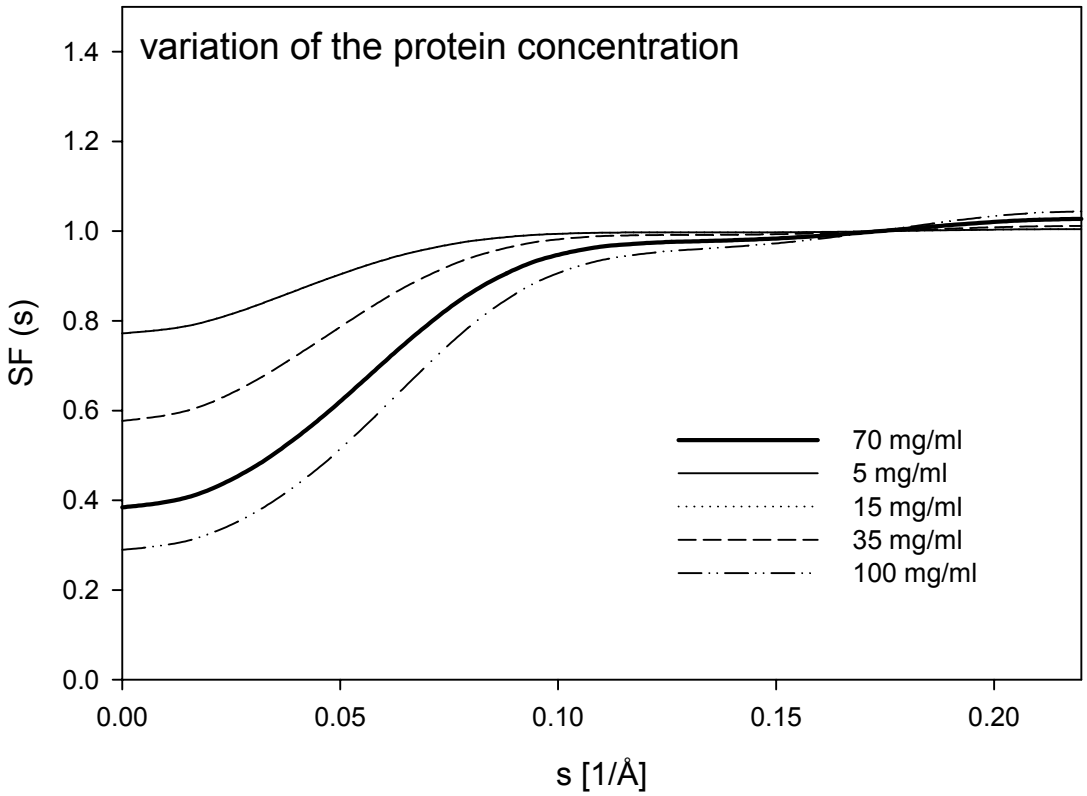


Figure 7.3: Simulation of the structure factors at different protein concentrations (5, 15, 35, 70, 100 mg/ml). The higher the concentration the larger the deviation from 1.

concentration, which are typical for the interaction of lysozyme molecules (Table 7.2).
In Figure 7.3 the concentration is varied between 5 and 100 mg/ml. As already discussed in the theory section, higher protein concentrations lead to stronger interactions and thus also to larger deviations of the structure factor from 1. For larger concentrations these interactions can also be measured more precisely.
In general an increase in the interaction potential leads to a larger deviation from 1 as is also illustrated in Figure 7.4 for an increase in effective charge. For the lowest charge (z=1.5) the repulsion is so weak that the attractive potential leads to a very large structure factor at the origin. An increase in the protein charge leads to a stronger repulsive potential and therefore to a decrease in the structure factor at the s-origin. Simultaneously the increase in charge also leads to an increase in the amplitude of the small peak in the structure factor around s=0.12 $Å^{-1}$.
The second parameter of the repulsive electrostatic contribution to the pair potential is the Debye length. It characterises the charge screening and therefore the range of the repulsive potential and depends on the counterion concentration in the solvent. The increase of the structure factor at the origin when the Debye length is decreased (Figure 7.5) is consistent with the fact that the structure factor increases when the effective charge is lowered (Figure 7.4). The evolution of the structure factor in Figure 7.5 is equivalent to the change of the structure factor in a series at fixed protein concentration and increasing salt concentration. The crossover point around s=0.08$Å^{-1}$ (for these parameters) is characteristic for such a series. A salt series is a valuable tool to vary the strength of the repulsive potential and even change the interactions from repulsive to attractive. This facilitates the observation of the effect of additives on the protein interactions in the attractive and repulsive regime without the need to change properties of the protein itself like, for instance, its charge. This

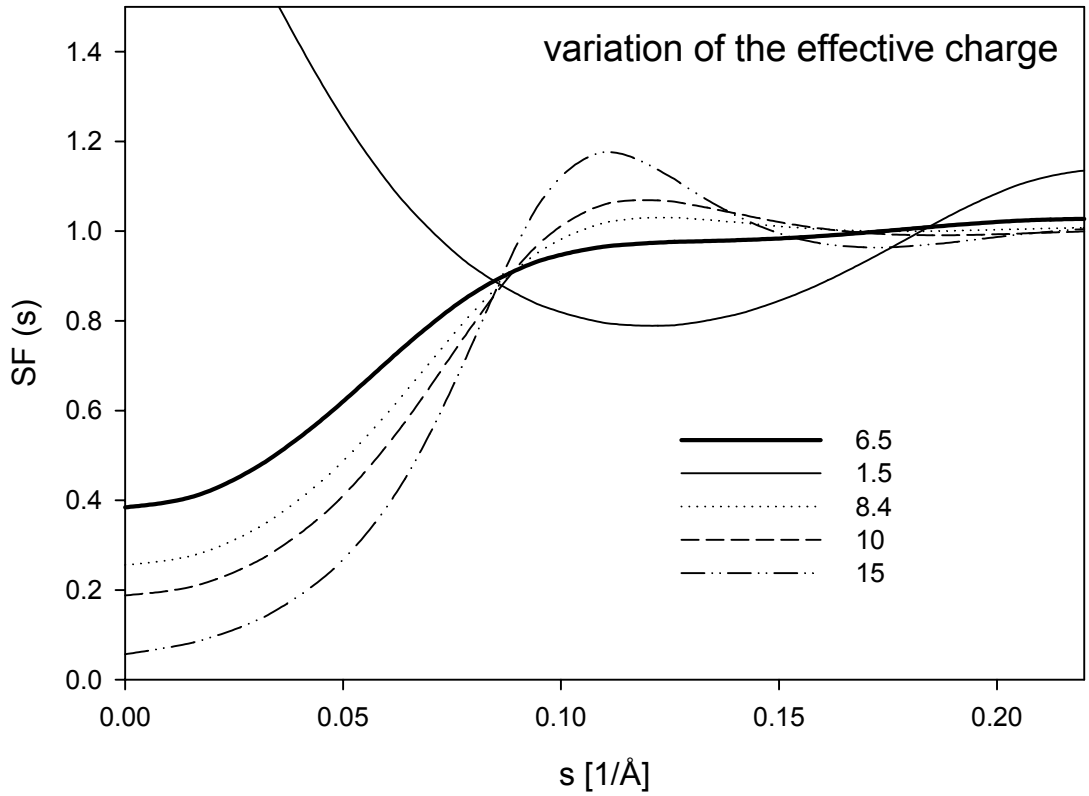


Figure 7.4: Simulation of the structure factors at different effective protein charges (z=1.5, 6.5, 8.4, 10.0, 15.0). An increase in the protein charge leads to an decrease of the structure factor at the s-origin, and to an increase of the amplitude of the small peak in the structure factor around s=0.12 Å^{-1} (for these values of the parameters). If the repulsive potential becomes very weak (z=1.5) the attraction between the proteins governs their interactions (aggregation) and leads to the strong increase of the structure factor above 1 near the s-origin.

method will be applied in the following sections to investigate the influence of urea, TMAO and glycine. These substances are expected to influence the attractive interactions, which are here characterised by a Yukawa potential with two parameters, depth and range. Both are varied independently and the simulation results are presented in Figure 7.6 for the depth of the potential and 7.7 for its range. The two parameters together control the overall strength of the attractive potential. Therefore an increase of either of them leads to an increase of the structure factor near the origin. Specific for the range of the attractive forces is that they control the separation between the maxima and minima of the structure factor. These minima and maxima can only be observed in the attractive regime and therefore in Figure 7.7 they can only be recognised by comparing the structure factors for values of the range of 5 and 10 Å.

The last parameter of the DLVO potential that remains to be varied is the hard sphere radius (Figure 7.8). The hard sphere potential is effectively a repulsive potential. If it increases, the relative contribution of the attractive potential decreases leading to an increase in the structure factor at the origin. If the molecules are in close contact the higher angle parts of the scattering pattern (s>0.1 Å^{-1}) are also influenced by the hard sphere potential.

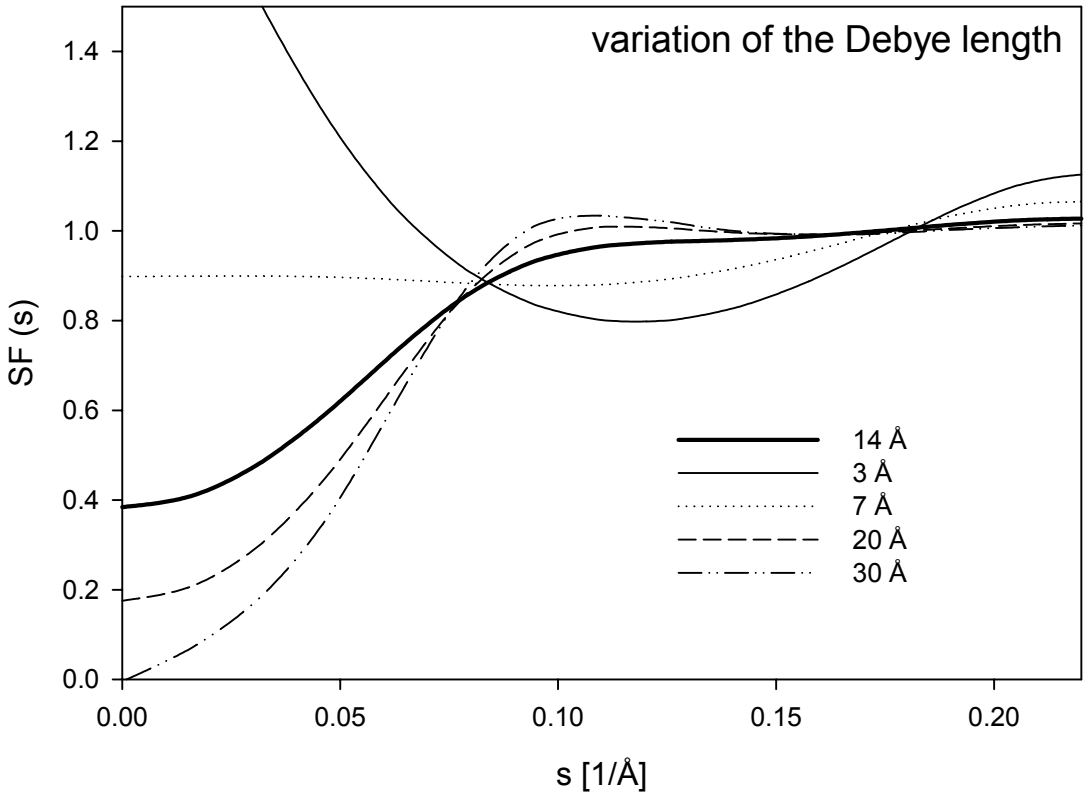


Figure 7.5: Simulation of the structure factors at different values for the Debye length λ_D of the buffer. The Debye length characterises the charge screening and therefore the range of the repulsive potential and depends on the ion concentration in the solvent. The increase of the structure factor at the origin when the Debye length decreases is in agreement with the fact that the structure factor increases when the effective charge is lowered (Figure 7.4).

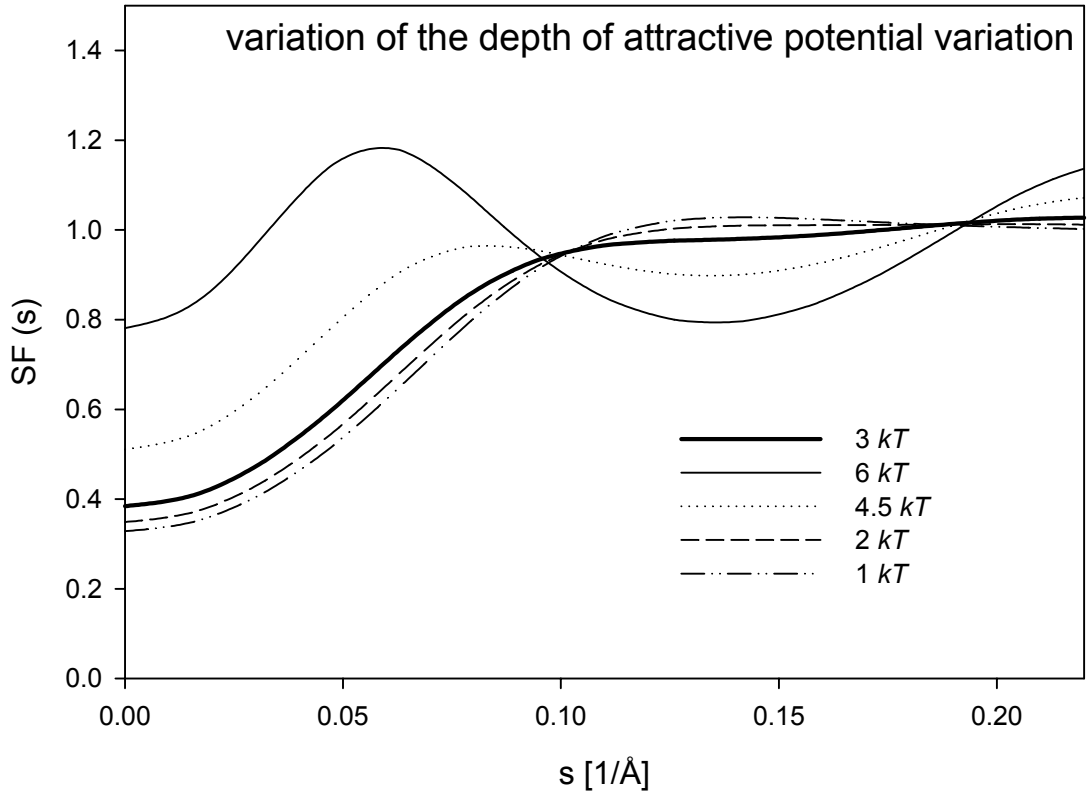


Figure 7.6: Simulation of the structure factors for different values of the depth of the attractive potential (1 to 6 *kT* at 20ºC). The increase of the depth of the attractive potential is basically equivalent to a decrease of the repulsive (electrostatic) potential (Figure 7.4).

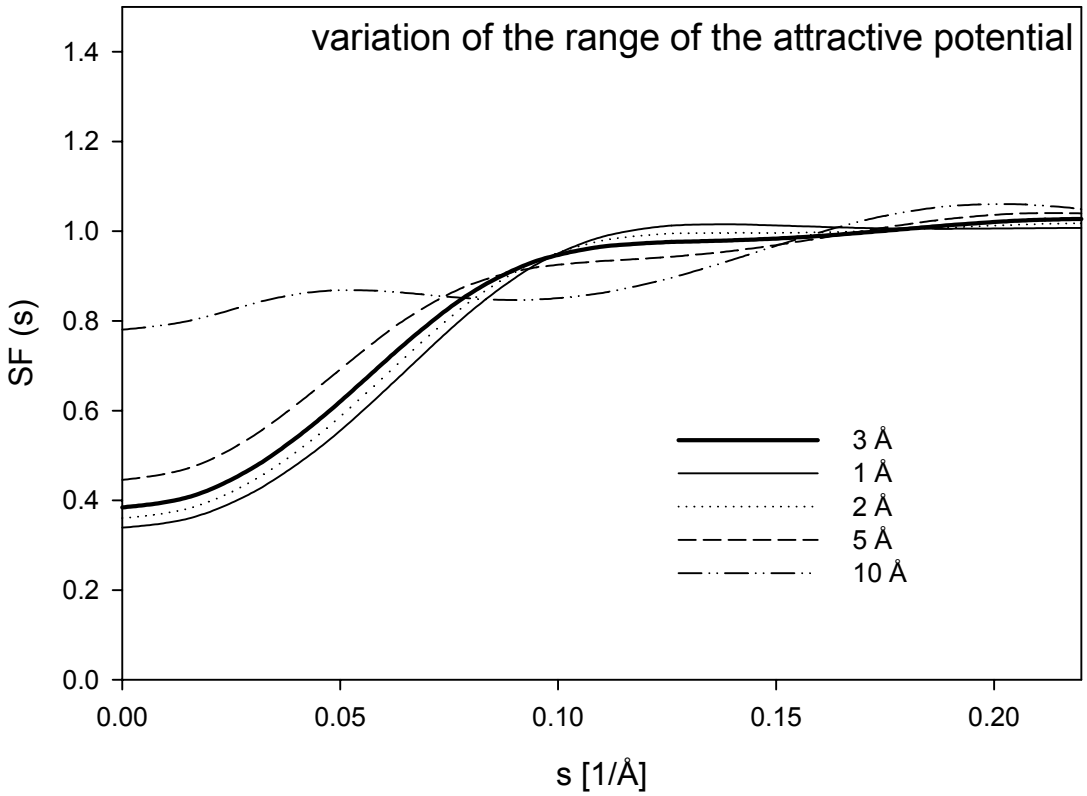


Figure 7.7: Simulation of the structure factors for different ranges of the attractive potential (1 to 10 Å). The change of the range of the attractive potential has two different effects. First at fixed depth of the attractive potential it changes the strength of the attractive potential and therefore leads to the variation of the structure factor around the origin. The second effect, which cannot be observed with dominant repulsive interactions, affects the curves in the attractive regime in such a way that it changes the distance in *s* between the maxima and minima of the structure factor (as illustrated for a range of 5 and 10 Å).

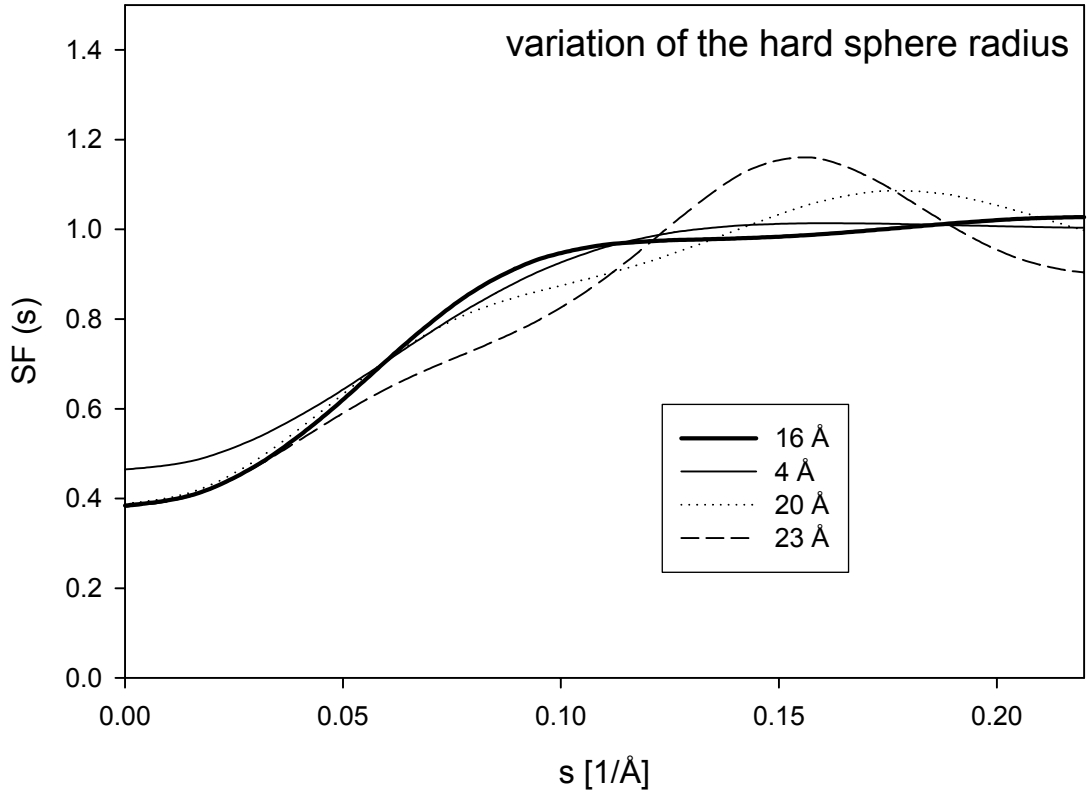


Figure 7.8: Simulation of the structure factors at different radii (4 to 23 Å) for the hard sphere potential. The hard sphere potential is equivalent to a repulsive potential and the reduction of the radius leads to an increase of the structure factor at the origin.

7.3 Effect of urea, glycine and TMAO on the short-range attractive interactions in lysozyme

Small molecules such as urea, glycine and TMAO are present in larger concentrations in the intracellular fluids of many species of all kingdoms and in particular in marine animals (see e.g. [Hochachka, 2002]). These substances affect the structure and/or function of the proteins. Urea, for instance, is known to cause chemical denaturation. Recent molecular dynamics calculations [Bennion, 2003] suggest that urea destabilises proteins by hydrogen bonding to their polar moieties, such as peptide groups, preventing the formation of intramolecular hydrogen bonds. Urea also alters the water structure and dynamics, which diminishes the hydrophobic effects and loosens the cohesion of the hydrophobic core [Bennion, 2003]. It can thus be expected that the presence of urea should influence the attractive potential of lysozyme.
The cosolvent TMAO counteracts the destabilising effects of urea most likely by increasing the effective attractive forces within the protein. TMAO is expected also to increase the intermolecular interactions.
Larger concentrations of amino acids like glycine may affect the solubility of biological macromolecules and their complexes. Glycine is known to prevent chromatin precipitation at physiological ionic strength (0-0.4 M NaCl) [Buche, 1989], and it can thus be expected to decrease attractive interactions at elevated salt concentrations.
The parameters of the attractive potential must be measured in the attractive regime. Since lysozyme has a pI of 10.5 strong electrostatic repulsion will occur in deionised water at pH

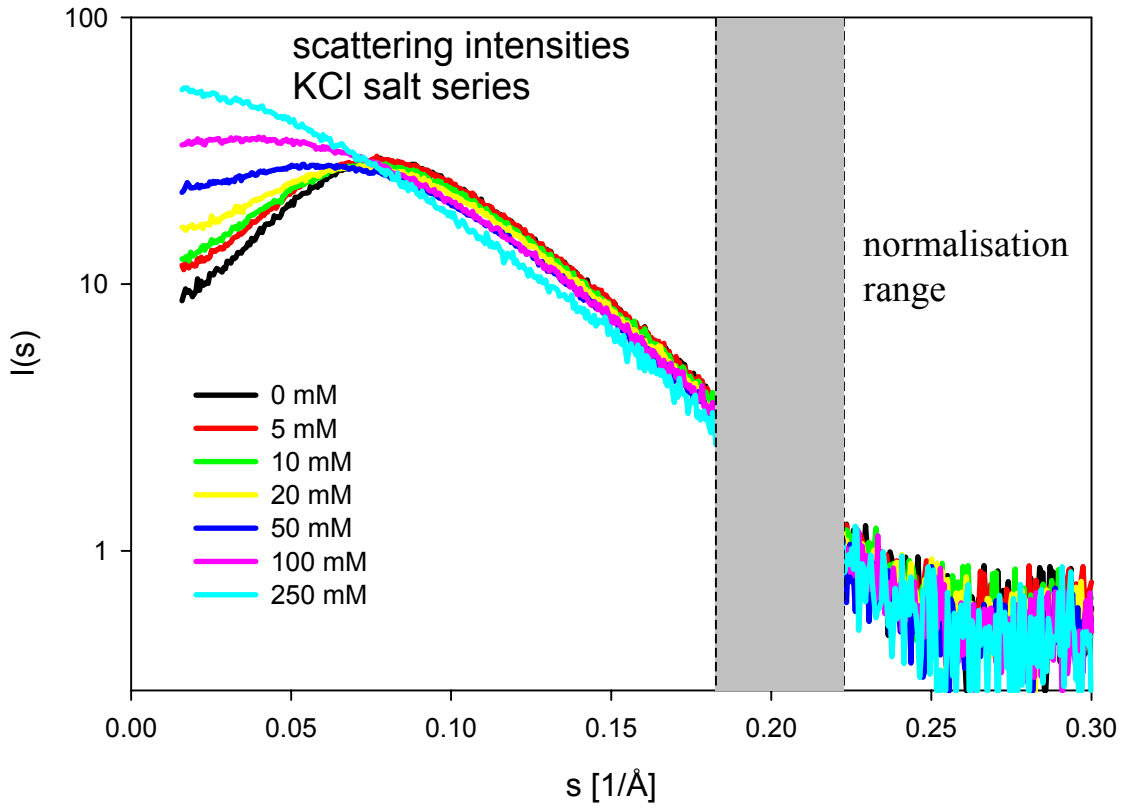


Figure 7.9: Experimental scattering intensities of the KCl salt series (0-250 mM, at a fixed lysozyme concentration of 68 mg/ml) in deionised water. Before calculating the structure factors the curves were normalised in the range 0.188 $Å^{-1}<s<0.215$ $Å^{-1}$ to correct for small differences in protein concentration.

values below 10. One possibility to reach the attractive regime is, as discussed in the previous section, to add salt to screen this repulsion. For this reason, salt concentration series were measured for all three cosolutes (urea, TMAO, glycine).
The intensity scattered at small angles is proportional to the square of the contrast between protein and solvent (i.e. the total excess scattering length density $\overline{\rho}$):

$$I(s) = \overline{\rho}^2 V^2 \exp(-\frac{1}{3} R_g^2 s^2) \qquad (7.1)$$

For X-rays $\overline{\rho}$ is the difference of the electron densities of protein and solvent. A typical electron density for protein is 0.42 Å^{-3}. The electron density of the solvent depends on the salt and co-solute concentration. For instance the electron density of water increases from 0.334 Å^{-3} in pure water to 0.337 Å^{-3} in a 250 mM KCl solution. The contrast thus drops from 0.086 Å^{-3} to 0.083 Å^{-3}, which corresponds to approximately 5 % leading to a decrease of approximately 10 % in the scattered intensities.
In order to compensate for the contrast changes and the small differences in protein concentration the scattering curves were normalised in the *s*-range between 0.188 Å^{-1} and 0.215 Å^{-1}. In this region the curves are essentially parallel and it can be assumed that the interaction effects no longer influence the structure factor in this interval. This procedure, which effectively forces the value of the structure factor in this interval to 1, was applied for all series discussed in the following section.
The form factor for the analysis was extracted from a lysozyme concentration series (data not shown). The scattering intensities were fitted with GNOM [Svergun, 1992]

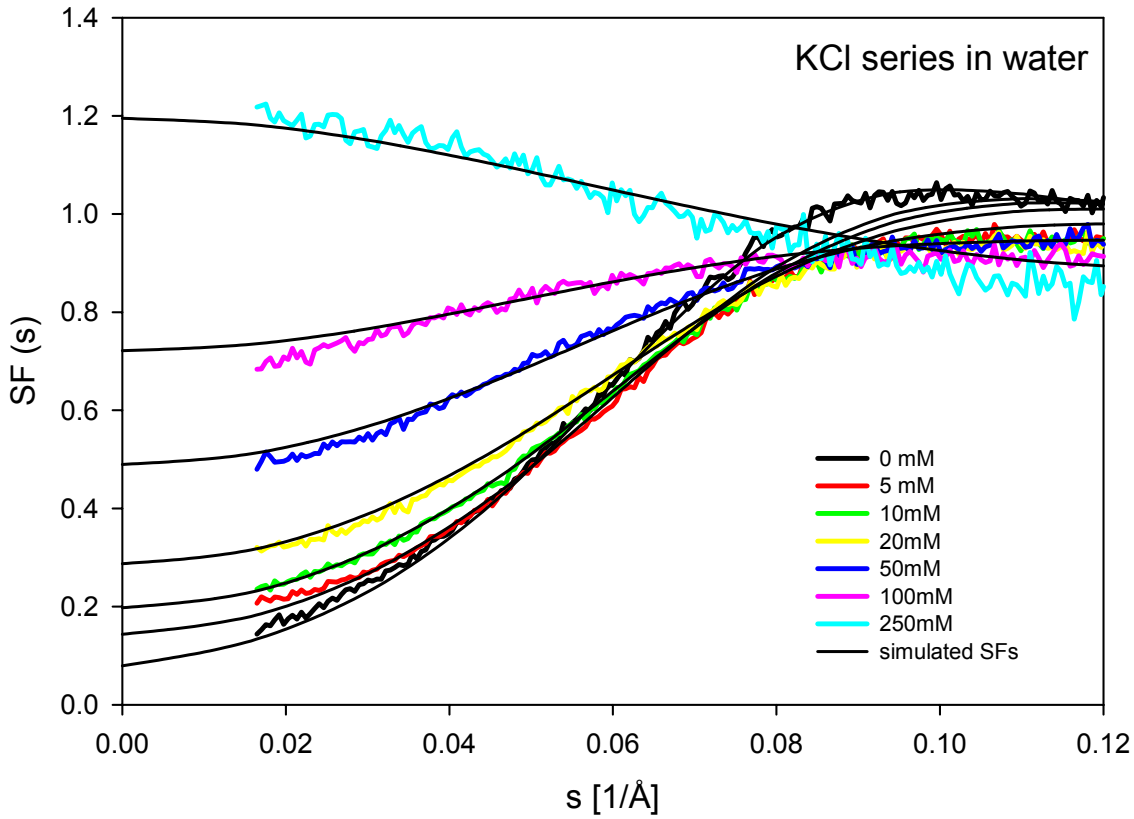


Figure 7.10: Structure factors of the KCl salt series in deionised water (lysozyme concentration 68 mg/ml). The radius of lysozyme was set to 14.2Å. An effective charge of 6.5 was found to fit the data in the repulsive regime best. The Debye length of the repulsive potential was calculated from the amount of salt added and ranged from 26.3 (0mM) to 5.8 (250 mM). The range of the attractive potential was set to 3.0Å. The depth of the attractive potential was chosen individually for each of the salt concentrations and it was found to be between -5.3 *kT* at 0 mM and -3.5 *kT* (at 20ºC) for 250 mM KCl.

(D_{max}=46Å; R_g=15.4 Å) and the GNOM fit was used as the form factor.

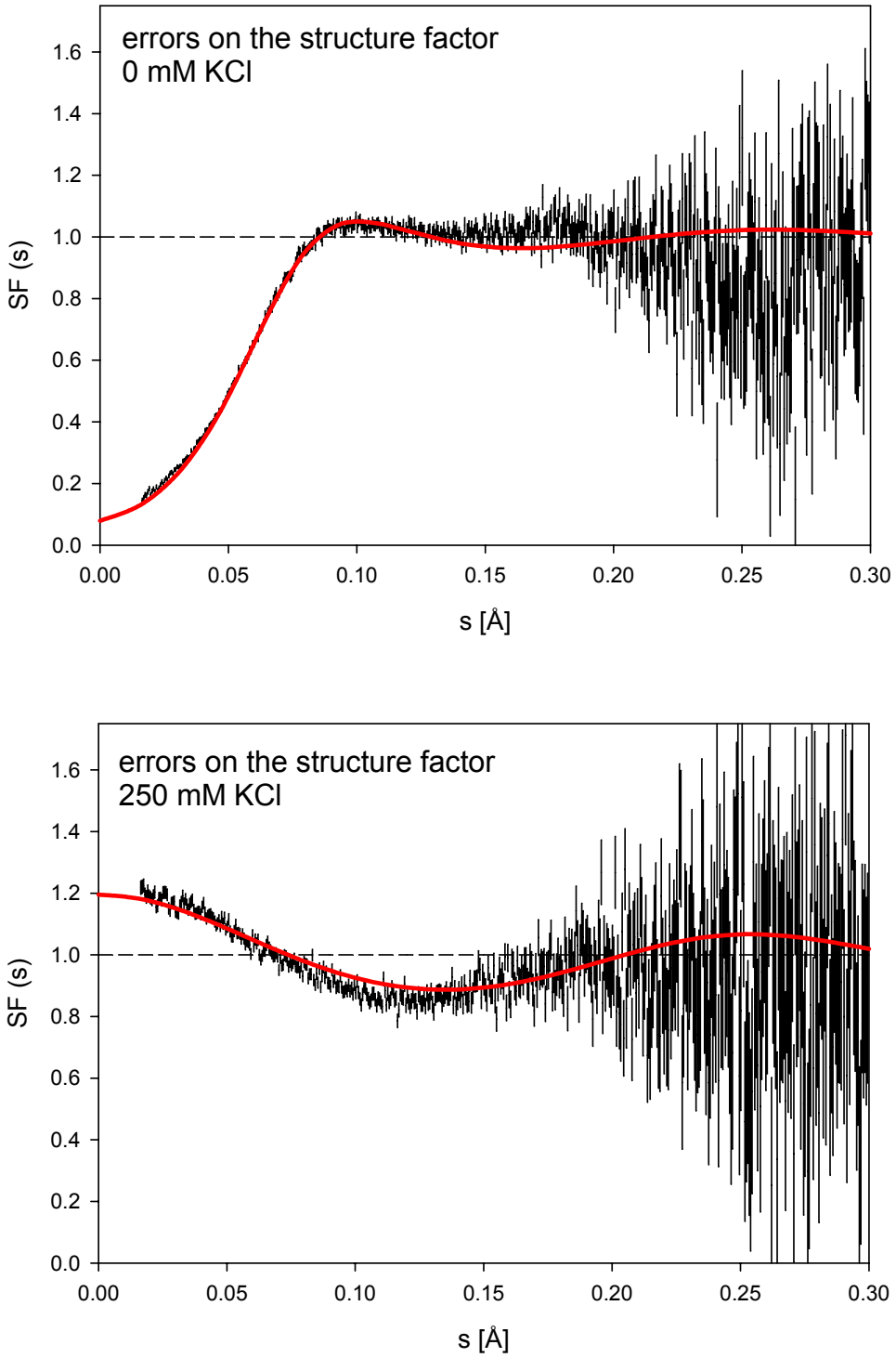


Figure 7.11: Structure factor measurements with statistical errors for the KCl series in deionised water and comparison to the model structure factors. The errors are calculated from the number of counts per channel. The error on the form factor, which was obtained with GNOM, was assumed to be 1%. Especially in the attractive regime a significant deviation is visible.

KCl [mM]	λ_D [Å]
0	26.3
5	22.4
10	19.8
20	16.5
50	12.0
100	8.9
175	6.9
250	5.8

Table 7.3: Values for the Debye length (λ_D) taking into account the inherent ionic concentration of 13 mM, for the measurements of lysozyme at different KCl concentration in deionised water and with 250 mM urea, glycine or TMAO.

First the results of the structure factor measurements in deionised water are presented. All measurements (water, urea, glycine, TMAO) were made with a lysozyme concentration of 68.0 mg/ml (Figure 7.9). The first problem arising during the selection of the fitting parameters is the choice of the Debye length λ_D. During the dialysis of lysozyme against deionised water most of the ions were removed but a small amount remains in solution. Also H_3O^+ and OH^- ions are always present. If there were no salt at all (0 mM) the DLVO theory would break down since this would result in an infinite Debye length and therefore in an infinite repulsive potential. In the present case a Debye length of 26.3 Å fits the data of the measurement in deionised water without additional KCl. This corresponds effectively to an ionic concentration of 13 mM. This inherent salt concentration has to be taken into account when the Debye length for the other KCl concentrations is calculated even if at higher salt concentrations this hardly affects the λ_D values (Table 7.3).
An effective charge of 6.5 yields the best fit to the experimental data. In the case of deionised water the value of the hard sphere radius R is not very important for the shape of the curves but measurements with stronger attractive potentials showed that 14.2 Å is a reasonable value. According to the Hamaker theory [Tardieu, 1999] a Yukawa potential with a range of 3.0 Å represents a realistic short-range attractive potential. The depth of the attractive potential was individually fitted to each of the measurements. The exact values can be found in Table 7.4. Without additional salt the interactions are, as expected, most intense (-5.3 kT at 20ºC for 0 mM KCl). Upon addition of salt (250 mM KCl) they drop to –3.5 kT at 20ºC. The calculated structure factors that fit the measurements best are shown in Figure 7.10.

KCl [mM]	deionised water	glycine 250 mM	urea 250 mM	TMAO 250 mM
0	-5.3 ± 0.5	-4.5 ± 0.5	-3.7 ± 0.5	-6.9 ± 0.5
5	-4.1 ± 0.5	-4.5 ± 0.5	-3.0 ± 0.5	-6.9 ± 0.5
10	-3.8 ± 0.4	-4.2 ± 0.4	-3.9 ± 0.4	-6.6 ± 0.4
20	-3.6 ± 0.3	-3.7 ± 0.3	-3.7 ± 0.3	-6.3 ± 0.3
50	-3.6 ± 0.2	-3.6 ± 0.2	-3.1 ± 0.2	-5.2 ± 0.2
100	-3.5 ± 0.1	-3.2 ± 0.1	-2.9 ± 0.1	-4.65 ± 0.1
175	-	-	-	-4.3 ± 0.1
250	-3.5 ± 0.1	-3.5 ± 0.1	-3.12 ± 0.1	-

Table 7.4: Values for the depth of the attractive potential (in kT at 20ºC) yielding the best fit to the experimental data.

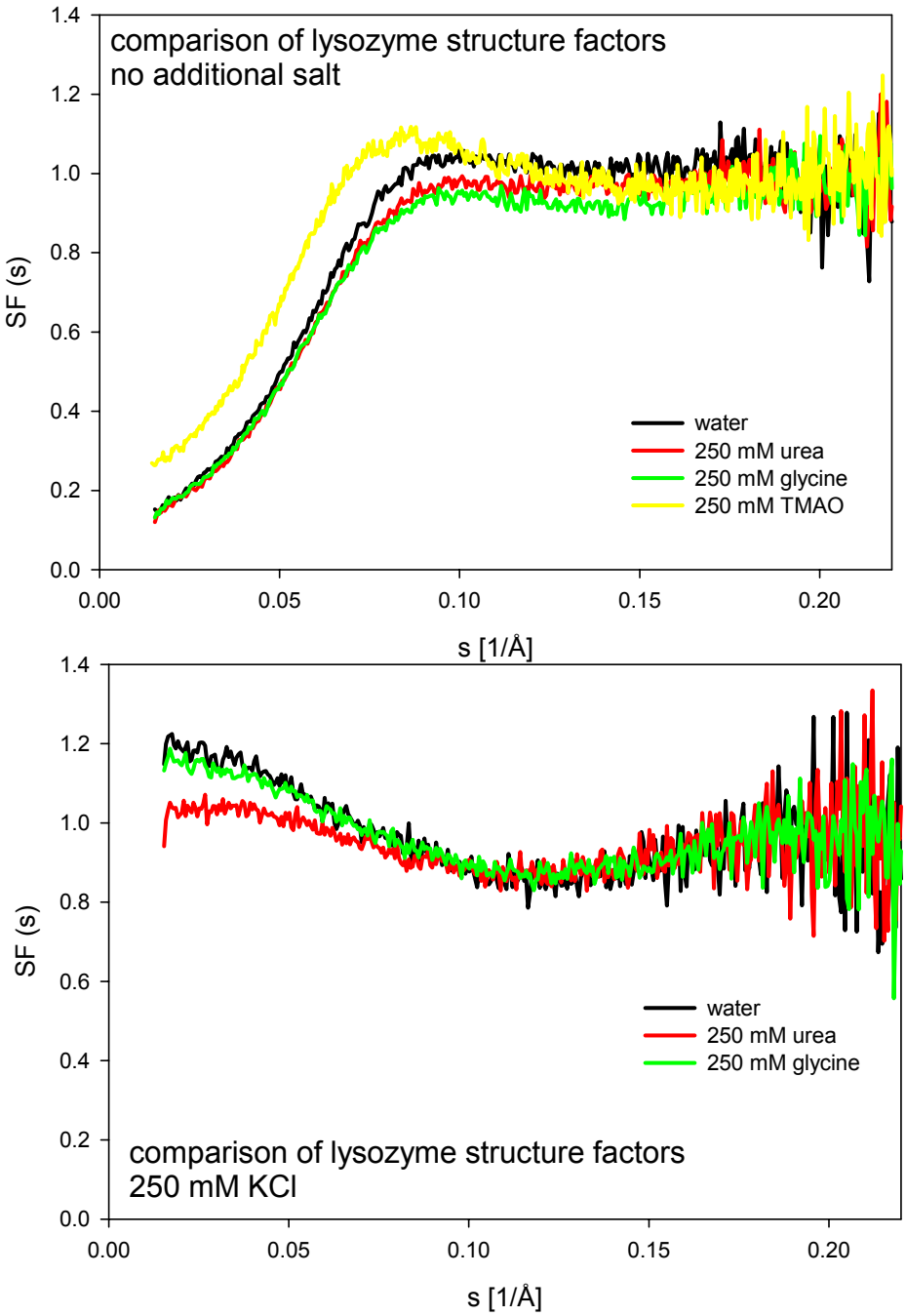


Figure 7.12: Comparison of the structure factors of lysozyme (68 mg/ml) in deionised water and in 250 mM urea, glycine and TMAO with 0 and 250 mM KCl. The measurements without salt represent the repulsive regime whereas attractive forces govern the interactions at 250 mM KCl. The interactions of lysozyme in water and in 250 mM glycine are about the same. In the attractive regime it was found that 250 mM urea decreases the attractive interactions significantly whereas even in the repulsive regime it is obvious that the cosolute TMAO significantly increases interactions. (The sample with 250 mM TMAO and 250 mM KCl was aggregated and did not yield usable structure factor.)

Figure 7.11 presents the statistical errors on the structure factors. They were calculated from the number of counts of the structure factor measurements at high concentrations. The relative error on the form factor that was calculated by the program GNOM [Svergun, 1992] was assumed to be 1% based on the fact that the errors on I(0) and R_g were of the order of 0.5%. The plot illustrates that the deviation between model and measurement can not be explained by statistical errors. There is a systematic deviation. This is not surprising given the simplifying assumption of the model (e.g. spherical shape of the protein).
It would actually have been possible to find a constant intermediate depth and range for the attractive potential that would reasonably fit all salt concentrations in deionised water. For instance an attractive potential with a range of 3.0 Å and a depth of 2.84 kT at 20ºC together with an effective charge of 6 and a hard sphere radius of 17.2 fits reasonably to all curves. The idea that it makes sense to fit the curves with individual attractive parameters was developed during the analysis of the series in TMAO. Whereas the series in glycine and urea are very similar to those in water (Figure 7.12), especially in the repulsive regime, the TMAO series is significantly different even without additional salt (0 mM KCl). It is not possible to account for the TMAO curves with a lower repulsive potential. The 0 mM KCl data can be described using a stronger attractive potential but if this value is assumed to be constant for the entire salt series in the attractive regime the structure factors would be much larger than the experimental ones. The TMAO concentration series illustrates that the structure factors change gradually with TMAO concentration (Figure 7.13).
It was checked whether the structure factors could be due to dimer formation in the sample. Even though the analysis of a low protein concentration measurement in 750 mM TMAO and the determination of the form factor under these conditions gave rise to the same

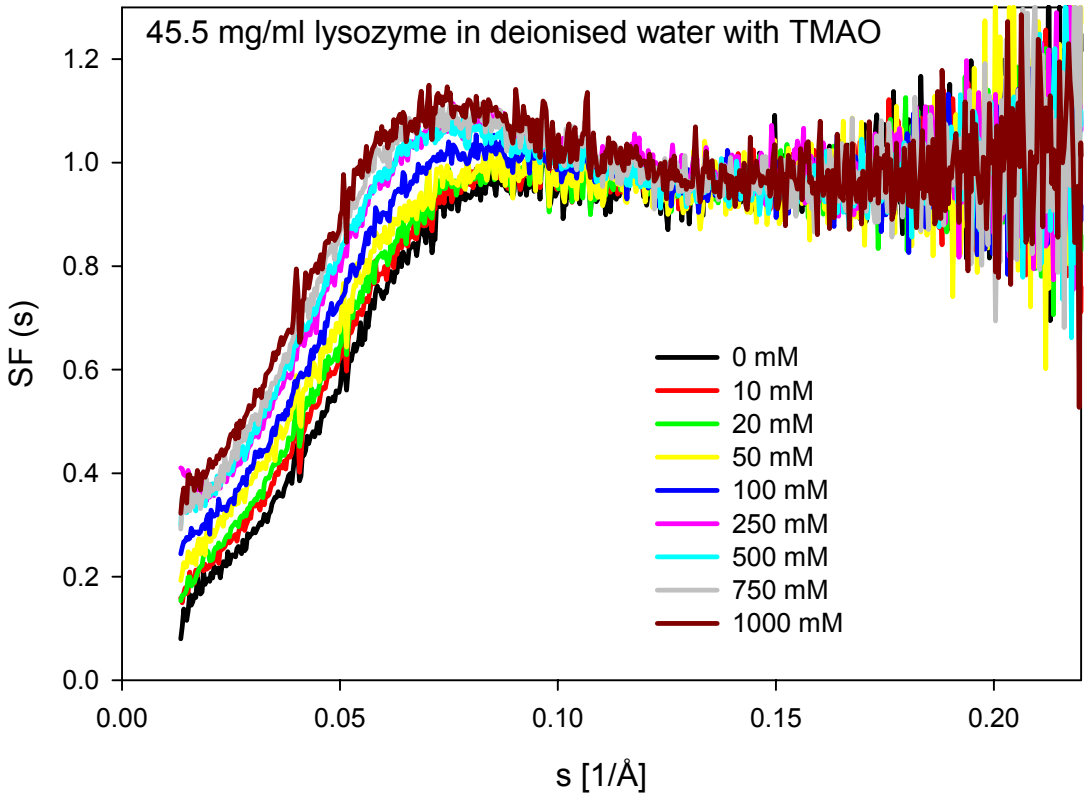


Figure 7.13: Structure factors of lysozyme (45.5 mg/ml) in deionised water at TMAO concentrations between 0 and 1M. Addition of TMAO leads to an increase of the structure factor over a wide range of *s*-values (s<0.12 Å^{-1}). The changes can not be explained with a change in the repulsive potential.

results as in water it could not be ruled out that dimers are formed at higher concentrations. To investigate the possible consequences of dimer formation the form factor of lysozyme monomers [Vaney, 1996] and three different dimers were calculated with the program CRYSOL [Svergun, 1995]. The dimers were constructed by making contacts between two monomers at positions where contacts also exist in lysozyme crystals. The form factors of the dimer have an I(0) which is four times that of the monomer. Instead of dividing the scattering intensities by the monomer form factor it was divided by those fictive dimer form factors. The expected structure factors were calculated with winOZ based on the assumption that the dimer is carrying twice the charge and has twice the volume but the

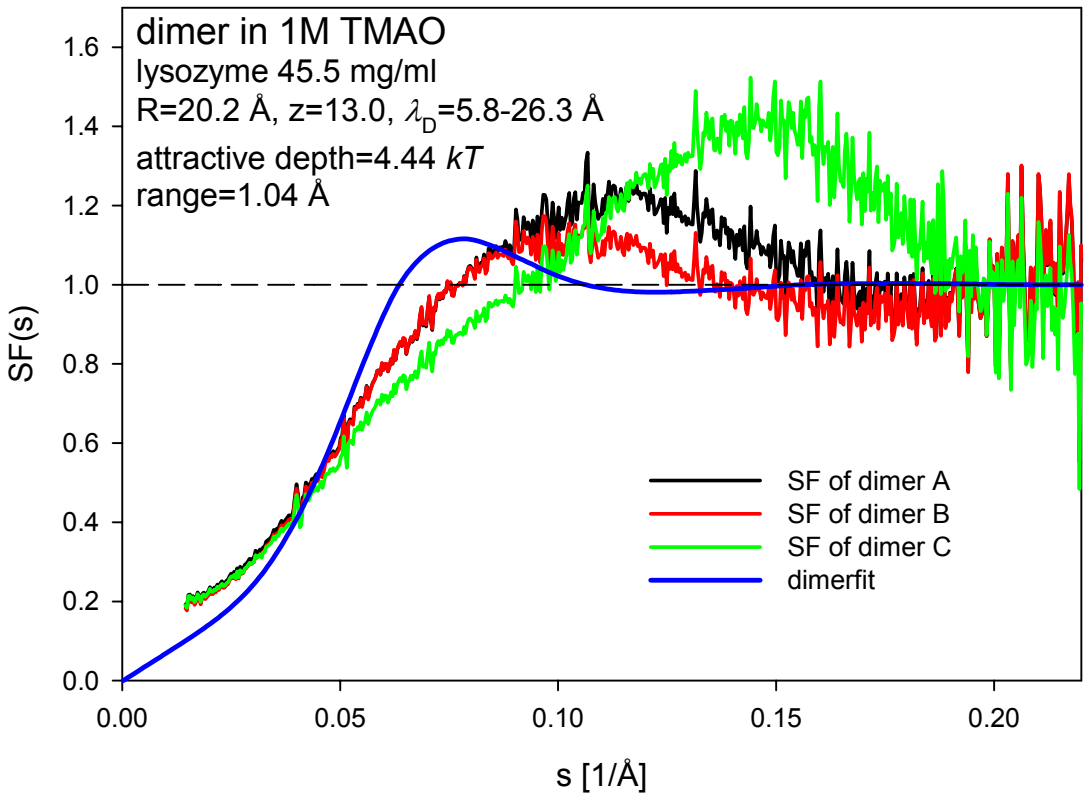


Figure 7.14: Calculation of the structure factors of lysozyme in 1M TMAO with theoretical dimer form factors. The resulting structure factors are still significantly different from the expected structure factors, which were calculated by the program winOZ with potential parameters, as expected for lysozyme dimers. The deviations of the structure factors in TMAO from those in water are clearly not due to dimer formation.

same attractive potential (Figure 7.14). None of the three structure factors (one for each dimer model) was in acceptable agreement with the calculated dimer structure factor. Therefore it was concluded that dimer formation can not lead to the observed scattering curves.

The only remaining explanation is that the attractive potential decreases with increasing KCl concentration. This assumption is reasonable since it was shown that a small polar molecule like TMAO at a concentration of 250 mM can have a significant influence on the attractive interactions. Therefore it is not surprising that 250 mM KCl would also influence attractive interactions since ions change the structure of water by aligning the dipolar water

molecules in their vicinity. Consequently the depth of the attractive potential was individually determined for all the curves and all cosolutes.
Figure 7.15 illustrates that in the attractive regime the onset of the structure factor is very sensitive to the change in the depth of the attractive potential whereas in the repulsive regime the onset does not change and only minor differences are visible in the range

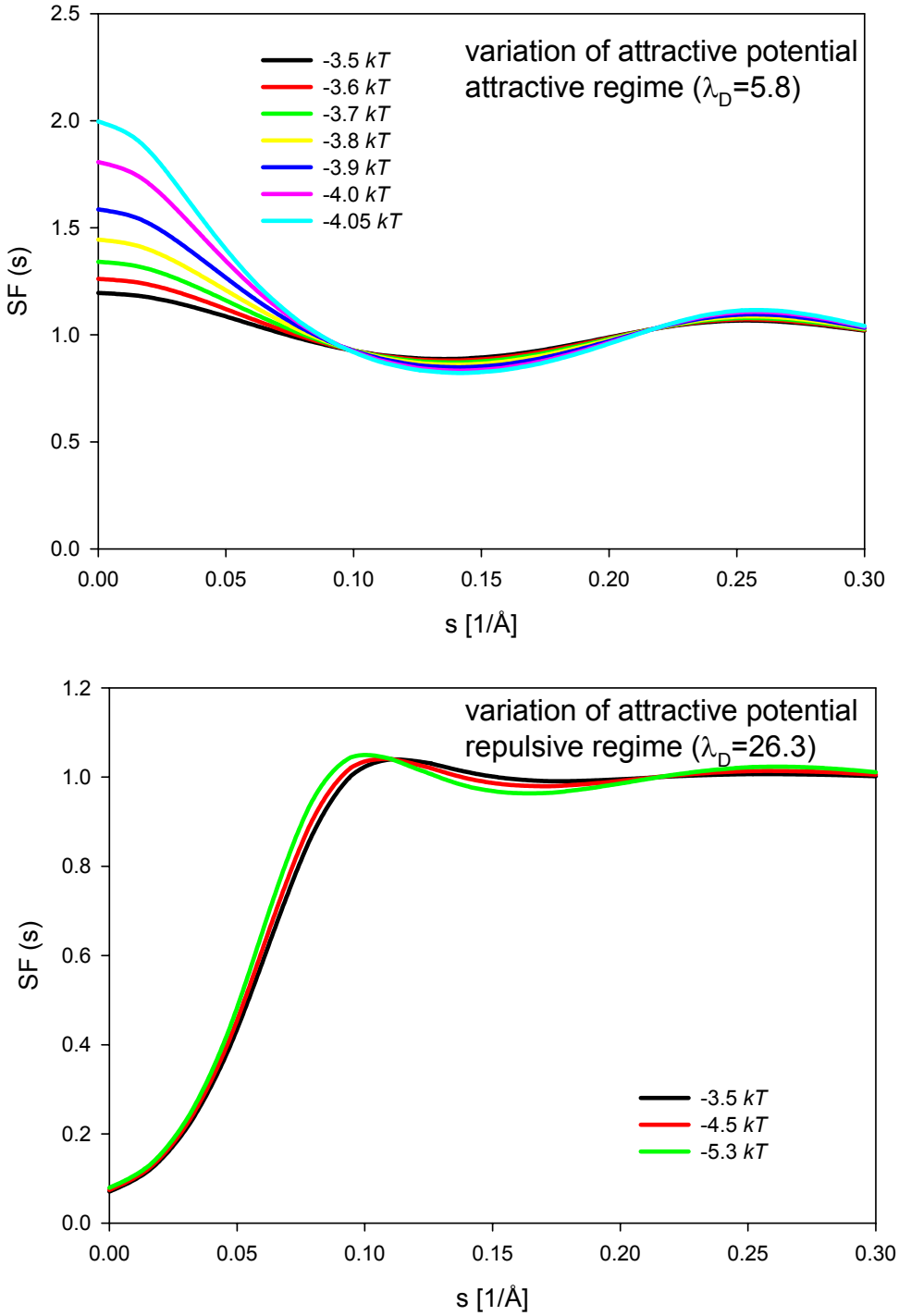


Figure 7.15: Structure factors (lysozyme 68 mg/ml) in deionised water with 0 and 250 mM KCl for different values of the depth of the attractive potential. In the attractive regime the onset of the structure factor is very sensitive to the change of the attractive potential whereas in the repulsive regime the onset does not change. Only minor differences are visible in the range $0.07 Å^{-1} < s < 0.20\ Å^{-1}$ and the uncertainty on the depth of the attractive potential is larger. This is reflected by the errors on the parameters in Table 7.4.

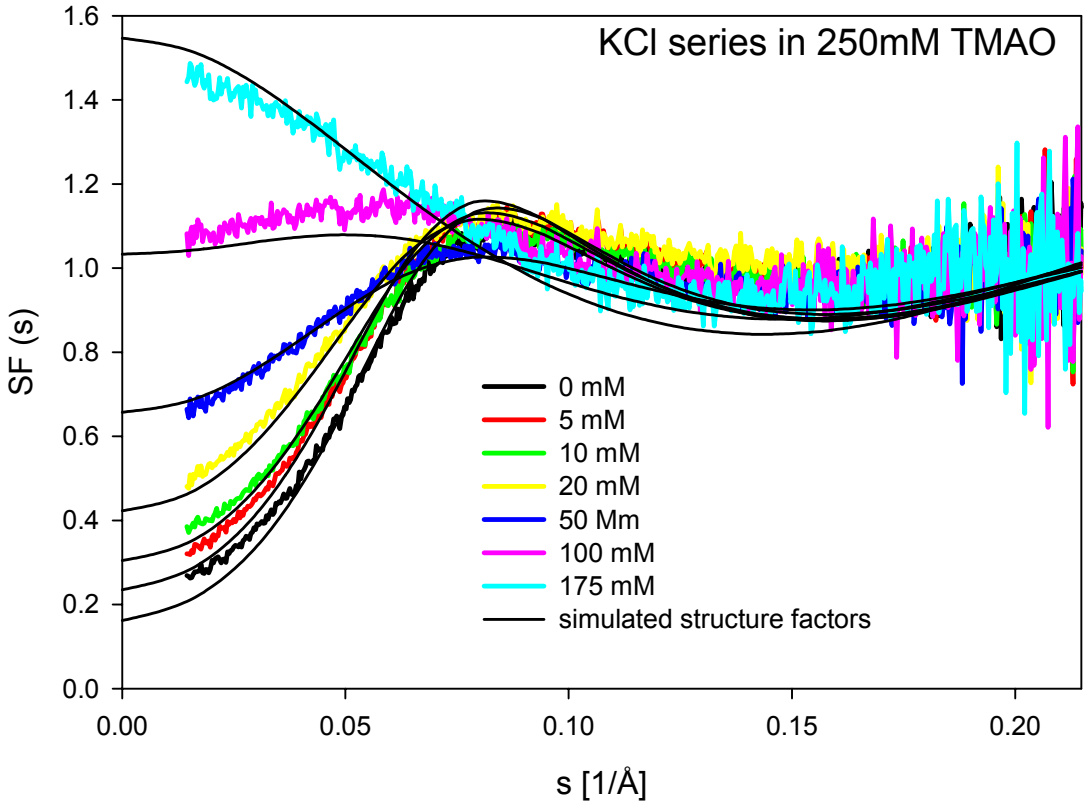


Figure 7.16: Lysozyme (68 mg/ml) structure factors of the KCl salt series in 250 mM TMAO. The radius, the effective charge, the Debye length and the range of the attractive potential were the same as for the measurements in deionised water. The depth of the attractive potential was chosen individually for each of the salt concentrations and it was found to be between -6.9 at 0 mM and -4.31 *kT* (at 20ºC) at 175 mM KCl i.e. significantly higher than in deionised water only.

$0.07Å^{-1}<s<0.20\ Å^{-1}$ and the uncertainty in the determined depth of the attractive potential is larger. This is reflected by the errors of the parameters in Table 7.4.

The measurements in deionised water (Figure 7.10), TMAO (Figure 7.16), glycine (Figure 7.17) and urea (Figure 7.18), can now be compared (Table 7.4). The same repulsive potential can be used to describe the electrostatic interactions between the lysozyme molecules with all cosolutes. Glycine as an additive at a concentration of 250 mM concentration does not have a significant effect on the structure factor. The structure factors are hardly different from those in deionised water. The differences between deionised water and the urea solutions are more pronounced. As expected a less attractive potential (depth between –3.7 and –3.12 kT compared to –5.3 to –3.5 kT at 20ºC for deionised water) fits the measured structure factors. The attractive interactions that hold the proteins together are weakened by the presence of urea.

The TMAO structure factors and therefore the strength of the attractive potential differ significantly from those in water, glycine and urea. With the same parameters for the repulsive potentials and the range of the attractive potential a depth between –6.9 and –4.3 kT at 20ºC is obtained. Since it is known that TMAO counteracts the effect of urea the increase of the attractive interactions could be expected. The change of the solvent-protein interactions due to the presence of TMAO increases the attractive interactions and stabilises proteins.

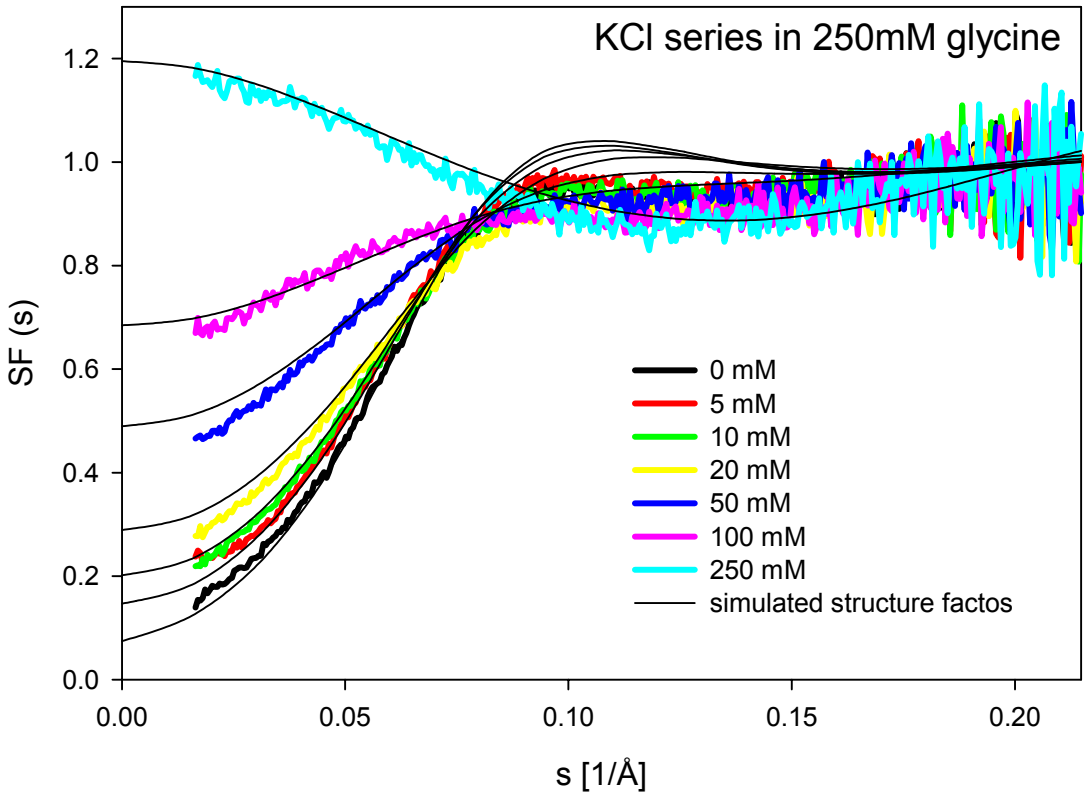


Figure 7.17: Structure factors of the KCl salt series in 250 mM glycine (lysozyme concentration 68 mg/ml). The radius, the effective charge, the Debye length and the range of the attractive potential were the same as for the measurements in deionised water. The depth of the attractive potential was chosen individually for each of the salt concentrations and it was found to be between -4.5 at 0 mM and -3.87 *kT* (at 20ºC) at 250 mM KCl and comparable to the values found for deionised water.

The measurements show that the structure factors obtained with SAXS can be used to characterise the influence of cosolutes on the attractive potential of proteins. For urea and TMAO the expected results were observed and could be quantitatively measured (Table 7.4, Figure 7. 19). Glycine had no significant impact on the attractive interactions.

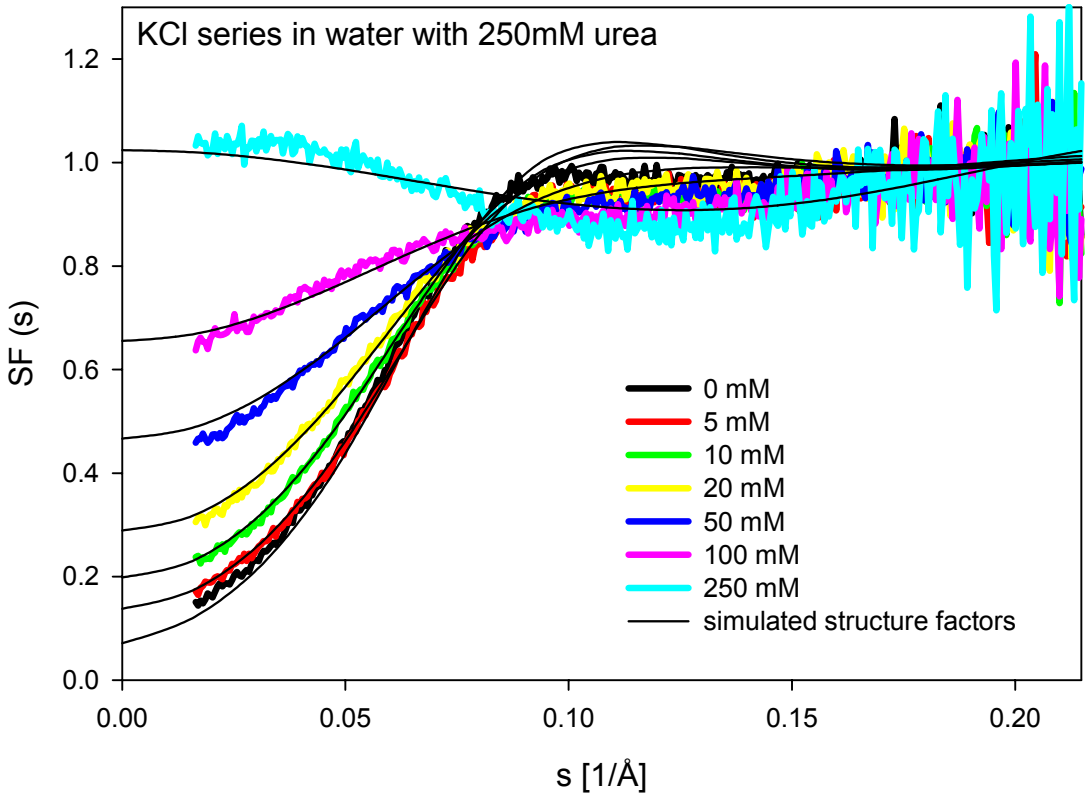


Figure 7.18: Structure factors of the KCl salt series in 250 mM urea (lysozyme concentration 68 mg/ml). The radius, the effective charge, the Debye length and the range of the attractive potential were the same as for the measurements in deionised water. The depth of the attractive potential was chosen individually for each of the salt concentrations and it was found to be between -3.7 at 0 mM and -3.12 kT (at 20ºC) for 250 mM KCl, which means that the attractive potential is significantly lower than in deionised water.

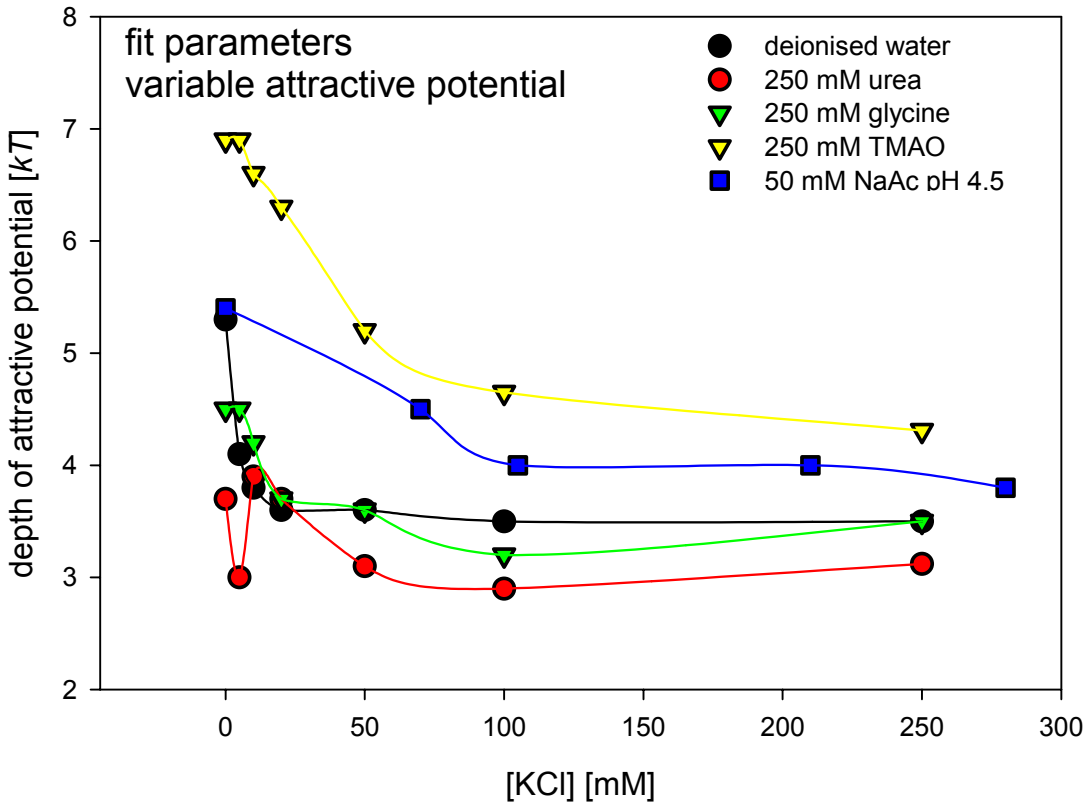


Figure 7.19: Representation of all fit parameters of the salt series in deionised water, 250 mM TMAO, 250 mM glycine and 250 mM urea. Averaged over all salt concentrations the attractive potential is about the same for deionised water and 250 mM glycine. Increased attractive interactions can be observed in 250 mM TMAO whereas 250 mM urea clearly weaken the attractive forces. The lines are only supposed to guide the eye.

7.4 Lysozyme in 50 mM NaAc at pH 4.5

Another salt series (with NaCl) was also measured in 50 mM NaAc at pH 4.5 (lysozyme concentration 100 mg/ml). For the calculation of the Debye length the buffer salts also have to be taken into account. As expected the same charge as for the potentials in deionised water could not be used here since at a lower pH (i.e. further from the pI) the charge of the protein is higher. An effective charge of 8.4 was needed to account for the

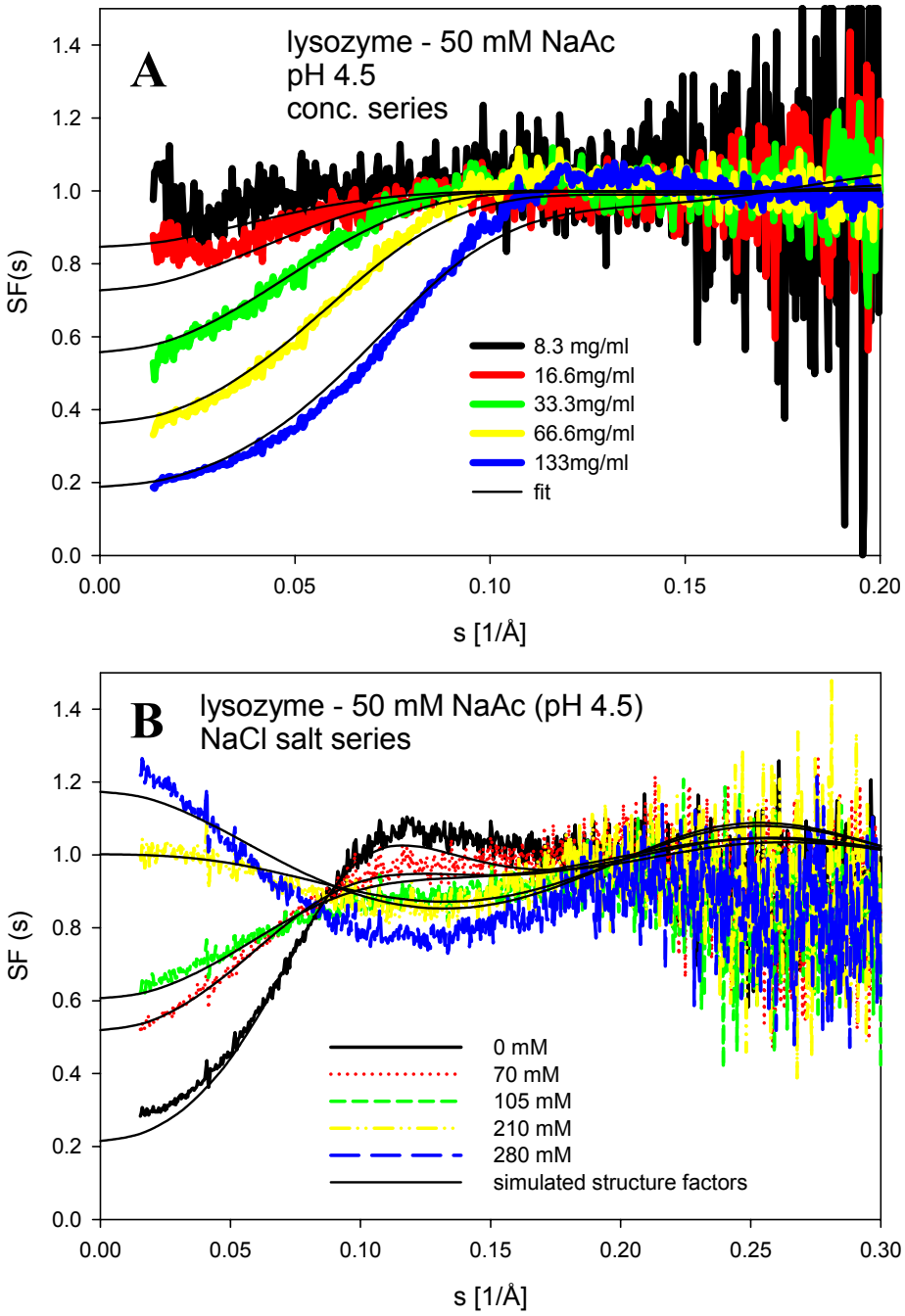


Figure 7.20: Structure factors of the protein concentration (A) and NaCl salt series (B) in 50 mM NaAc at pH 4.5 (lysozyme concentration 100 mg/ml). The radius and the range of the attractive potential were the same as for the measurements in deionised water. The effective charge was found to be 8.4. For the calculation of the Debye length the buffer ions were also taken into account.

stronger repulsive interactions. The values of the attractive potential are also significantly higher (between –5.4 and –3.8 kT at 20ºC). In the attractive regime the model does not fit the data well (280mM NaCl, Figure 7.20), but this is probably due to the fact that dimerisation is not taken into account.

7.5 Temperature dependence of lysozyme interactions

[Tardieu, 1999] published a temperature series of lysozyme in the attractive regime (pH=pI) was presented. With increasing temperature the deviation of the structure factor from 1 become smaller. This effect can be explained by the numerical calculation. If the temperature increases the average energy per particle increases as well. If one assumes that the potentials do not depend on temperature the ratio between particle energy and van der Waals and electrostatic potential increases. Therefore the motion of the proteins is less restricted by the potentials. According to the DLVO theory the deviation from 1 of the structure factor decreases, in the attractive regime as well as in the repulsive regime.
In order to check this theoretical prediction the structure factors of a solution of lysozyme (100 mg/ml) were measured in 50 mM NaAc at pH 4.5, far from the pI in the repulsive regime (Figure 7.21). Contrary to the predictions, the interactions and thus the deviation of the structure factor from 1 increased significantly. This can only be due to a change in the interaction potentials, either the attractive potential decreases or the repulsive potential increases with temperature or both.
A decrease of the attractive potential can indeed account for the changes in the region between $0.05Å^{-1}<s<0.12Å^{-1}$ as illustrated in Figure 7.21. This is unexpected because it is known from many proteins in particular those that self-assemble like virus capsid proteins, tubulin etc. that an increase in temperature leads to higher attractive interactions.

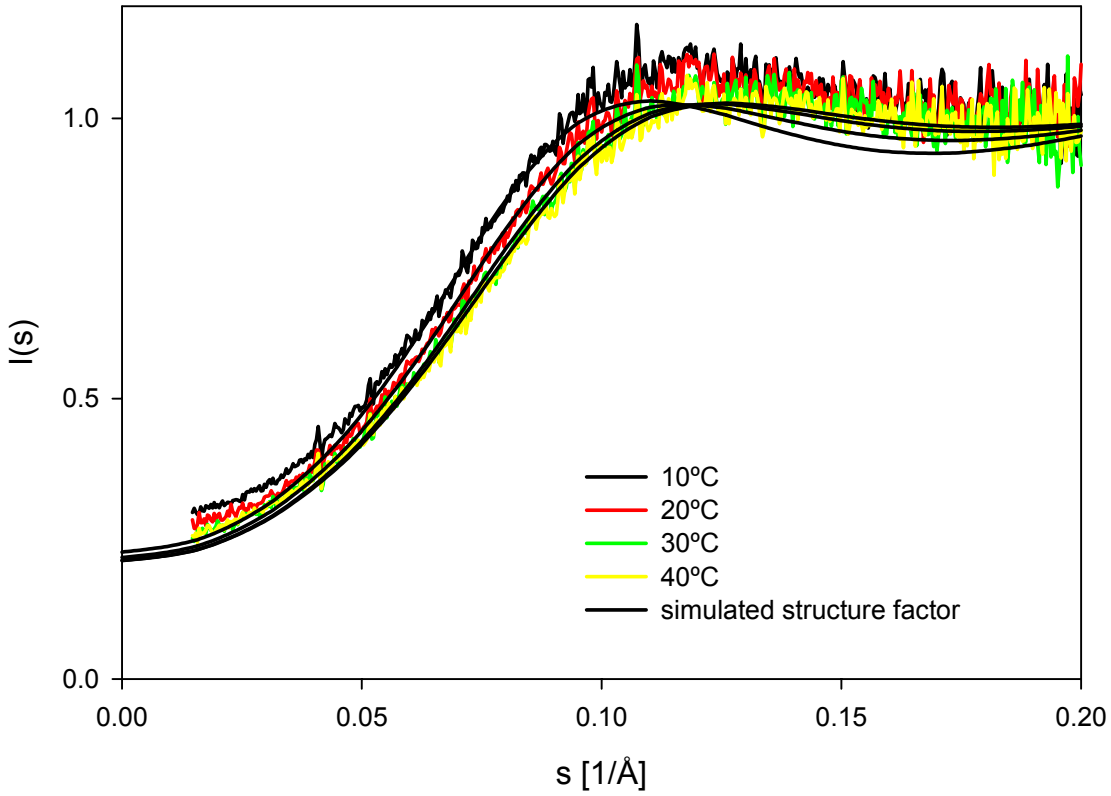


Figure 7.21: Temperature dependence of the structure factors of lysozyme (100 mg/ml) in 50 mM NaAc at pH 4.5. The differences can qualitatively be explained with an increase of the attractive potential at lower temperatures.

One possible cause is the temperature dependence of the Debye length, which was neglected in previous calculations.

$$\lambda_D(T) = \lambda_D(300K) \cdot \sqrt{\frac{T}{300}} \tag{7.2}$$

The Debye length λ_D at 300K in 50mM NaAc buffer is 13.42Å. At 10, 20, 30 and 40ºC the values are 13.04, 13.27, 13.49 and 13.71Å respectively. A larger Debye length makes the potential more repulsive but it can not overcome the effect of the higher temperature on the form factor due to the kT-term in Eq. (6.42). Introducing the temperature dependence of the Debye length changes the structure factors and makes them more repulsive at higher temperature but it can only compensate the decrease caused by the kT-term in Eq. (6.42).
Another possibility would be that the lysozyme charge increases and makes the potential more repulsive. The thermal coefficient of the pH-dependence of 50 mM NaAc buffer is only weak, typically 0.002/K. This change can not account for the differences observed in the structure factor. For the temperature change from 10ºC to 40ºC this corresponds to a change of 0.06 in pH. This change in pH would cause a corresponding small change in the charge of the protein. Whether the temperature would increase or decrease the pH is difficult to predict. It depends on whether the dissociation of NaAc is endothermic or exothermic. This has not been verified as the expected changes are of the order of the experimental precision. The charge increase could account for the change in the beginning of the curve.

7.6 Glucose oxidase

The glycoprotein glucose oxidase was studied with SAXS as a second model system. The influence of the carbohydrate moiety on the SAXS pattern was investigated and it was attempted to model the carbohydrates which are missing in the crystallographic structure of glucose oxidase.
Lysozyme has a molecular mass of 14.3 kDa whereas that of glucose oxidase is more than ten times larger (160 kDa). The main purpose of choosing glucose oxidase as a target for SAXS measurements was to check whether the models, which were successfully applied to lysozyme can also help to qualitatively and semi-quantitatively understand the interactions of glycoproteins. For studying the interactions a glucose oxidase concentration series was measured. Even though there is some aggregation of the sample in this case the results could be compared to predictions of the program winOZ. In order to investigate the attractive interactions and the changes due to the salt concentration in the buffer a KCl series was measured as for lysozyme. This facilitates the comparison between the repulsive and attractive interactions of the proteins. winOZ is again used to extract the parameter of the potential from the measurements of the structure factor.
The salt series of glucose oxidase was repeated with different salts (LiCl, NaCl, KCl) in order to measure the effect of the different cations. The protein was also measured at different pH values. The structure factors of solutions in 100 mM sodium phosphate buffer at pH 4.5, 7 and 9 were determined and compared.
Finally, the temperature dependence of interactions in the range 15-60ºC was measured in deionised water and in 50mM KCl.

7.7 Using the glucose oxidase form factor to obtain a model of the carbohydrate moiety

The first objective of the glucose oxidase measurements was to determine the form factor that is needed for all structural studies as well as the calculation of the structure factors. Because of the high salt shielding and in order to keep the influence of the repulsive interactions small a concentration series in 100 mM sodium phosphate at pH 4.5 was measured. The small angle part (s<0.87Å^{-1}) of the measurement at 3.2 mg/ml and the wide angle part (s>0.87Å^{-1}) were scaled and merged. The data were analysed with the program GNOM [Svergun, 1992] (D_{max}=98.5Å), which determined a radius of gyration of 35.0Å. The GNOM fit for the form factor was used to calculate all following structure factors. From $I(s)$ one can also see that the glucose oxidase sample was slightly aggregated. The onset of the curve had to be excluded in the GNOM analysis.
Figure 7.22 illustrates the differences between the form factors of glucose oxidase and lysozyme. Whereas the decay of the lysozyme form factor (R_g=15.0Å) is relatively slow that of the glucose oxidase form factor is more rapid as a result of the larger size of the protein. Since a crystallographic structure is available for both proteins the program CRYSOL [Svergun, 1995] was used to compare the theoretical SAXS pattern calculated from the spatial information in the PDB file and the experimental solution scattering pattern. The computed curves were fitted to the SAXS measurements. In the case of lysozyme the predicted and experimental curves are in excellent agreement (Figure 7.22) whereas for glucose oxidase there are significant differences. This is due to the absence of the carbohydrate moiety in the protein crystal. Since the glycosylated protein could not be

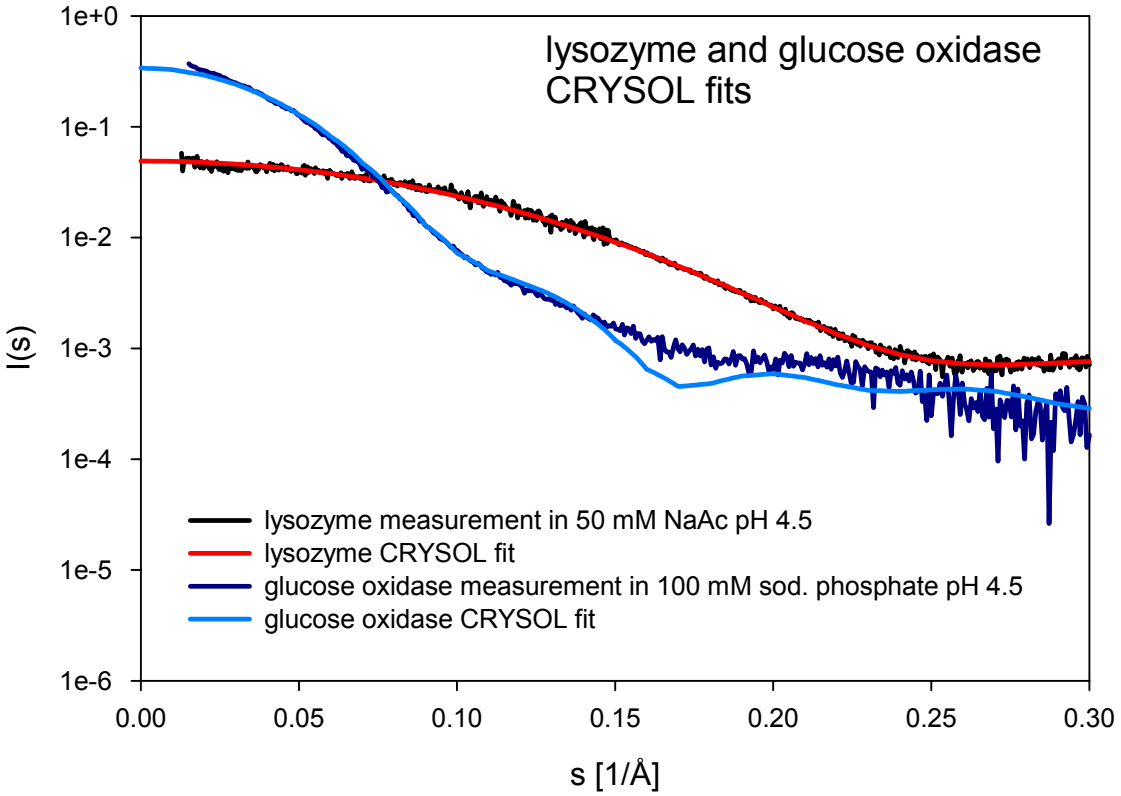


Figure 7.22: Form factors of glucose oxidase (R_g=35.0Å) and lysozyme (R_g=15.0Å). Whereas for lysozyme the measured SAXS intensities are in good agreement with the scattering curves predicted from the crystal structure with the program CRYSOL this is not the case for glucose oxidase. The absence of the carbohydrate moiety in the crystal structure (PDB entry 1CF3) gives rise to the deviation.

crystallised the sugars were cleaved and the remaining protein moiety was used to obtain the crystals [Kalisz, 1990]. In the SAXS measurements the carbohydrate moiety was present and gave rise to the differences observed in Figure 7.22.
The SAXS results can be used to generate a structural model for the protein and its carbohydrate content. A possible method is to make an ab initio shape determination with

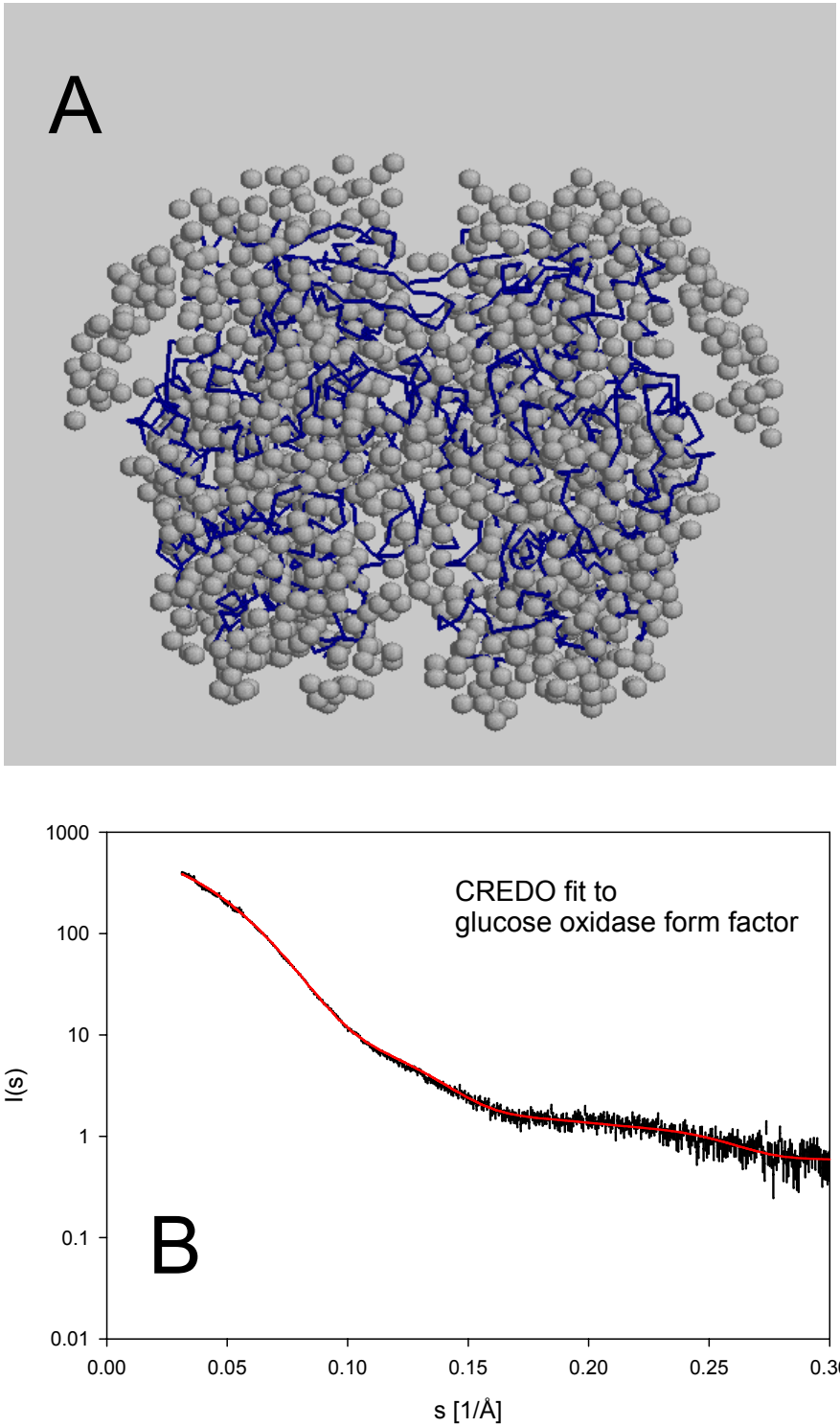


Figure 7.23: CREDO [Petoukhov, 2002] model of the carbohydrate moiety (A). The blue line is the glucose oxidase backbone. The two 'arms' at the upper end of the protein, which are not overlapping with the backbone, correspond to the carbohydrate moieties. The red line in (B) represents the scattering calculated from the CREDO model and is compared to the raw data.

the programs DAMMIN [Svergun, 1999] or GASBOR [Svergun, 2001]. In the present case one can also make use of the existing crystallographic model of the protein moiety. The program CREDO [Petoukhov, 2002] is designed to make a model of a protein complex if the structure of one part and the size (number of amino acids) of the other part is known. It will only model the unknown part, i.e. in the case of glucose oxidase, the carbohydrates. A number of amino acids with a mass equivalent to the glucose oxidase sugars (25.6 kDa) was used for the computation. The resulting model is illustrated in Figure 7.23. The two 'arms' at the upper part of the protein correspond to the carbohydrate moiety. The area where the sugars are linked to the protein is in good agreement with the area where the remaining sugar groups in the crystal structure are located.

7.8 Glucose oxidase structure factor in deionised water

Based on the knowledge of the potential parameters for lysozyme it is possible to come up with an educated guess for the glucose oxidase structure factor in the repulsive regime (deionised water, pH7). The best-fit hard-sphere radius was determined to be 16.2Å for lysozyme, which is relatively close to the radius of gyration (ca. 15Å). Therefore in the case of glucose oxidase its R_g was also chosen as the hard-sphere radius. The attractive potential is neglected since it is small compared to the repulsive potential in deionised water at pH 7. The effective charge of lysozyme was found to be 6.5. The charge of glucose oxidase can be calculated from the comparison of the charge of both proteins predicted by using the sequence information Figure 7.24 and was set to 36 for the computations.
The resulting structure factors are compared in Figure 7.25. The glucose oxidase structure factor increases much faster from the minimum close to the *s*-origin than that of lysozyme. The oscillations are also more pronounced for glucose oxidase. The period of oscillation is

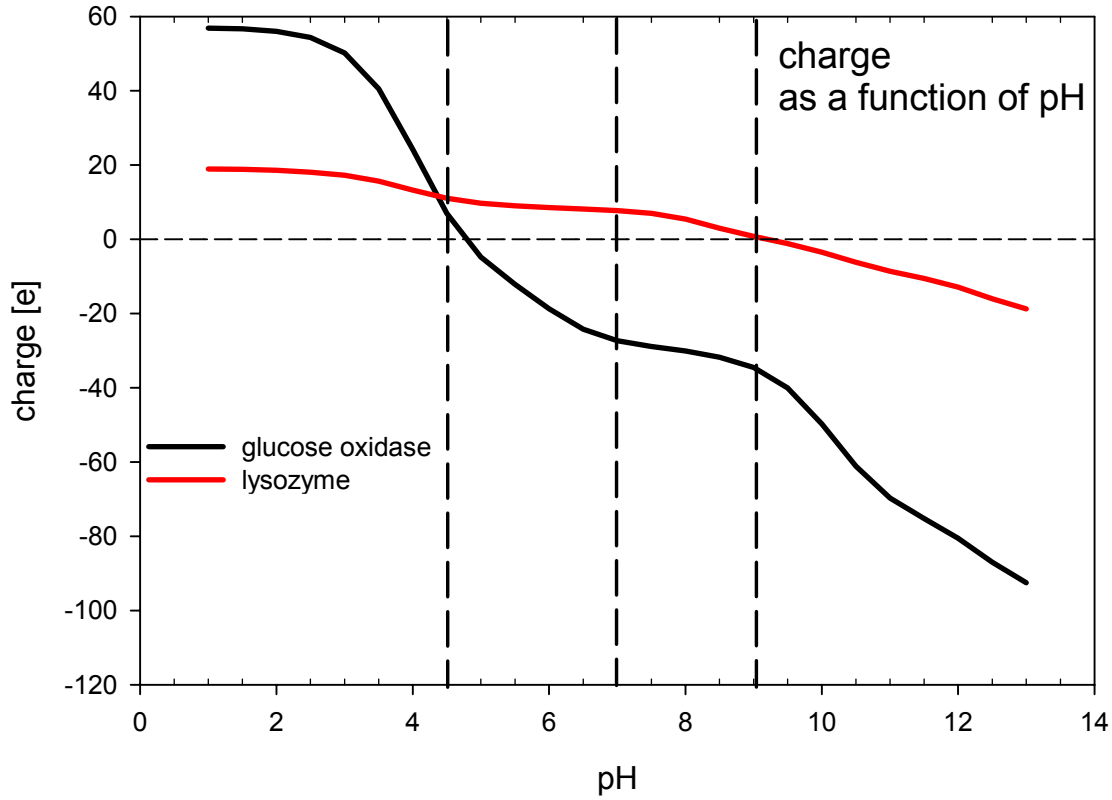


Figure 7.24: Prediction of the charge as a function of the pH for the proteins glucose oxidase and lysozyme based on their amino acid sequence.

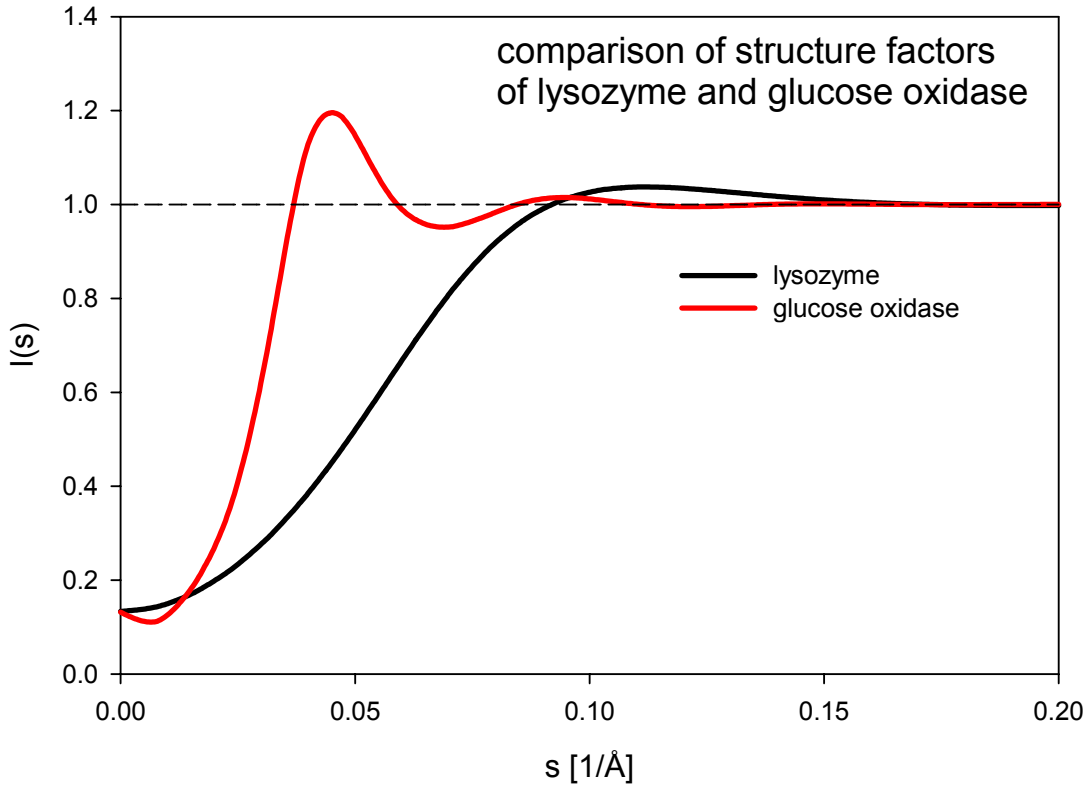


Figure 7.25: Structure factor prediction for two proteins (50mg/ml) of the size (R_g) and charge (z) of lysozyme (R_g=15Å, z=6.5) and glucose (R_g=35Å, z=36) oxidase at pH 7 in deionised water (λ_D=26.3Å). Only the hard sphere (R=R_g) and electrostatic interactions were taken into account for the simulation with the program winOZ. The charge of lysozyme was taken from the measurements of lysozyme in deionised water. The charge of glucose oxidase was calculated from the comparison of the charge of both proteins predicted by using the sequence information.

much shorter than for lysozyme because in real space the electrostatic interactions of glucose oxidase reach much further due to the larger hard-sphere radius.
In Figure 7.26 the scattered intensities of the glucose oxidase concentration series in deionised water are presented. Deionised water was chosen because due to the absence of ions the electrostatic interactions are not shielded and influence the structure factors strongly. Solutions with protein concentrations between 0.6 and 57 mg/ml were prepared. From the analysis of the low concentration curves with the program GNOM it is known that the samples are slightly aggregated. This is indicated by the fact that the onset of the $I(s)$-curves exceeds the amplitude of the form factor since the high concentration curves show that the net interactions are repulsive. Based on the assumption that the structure factor of the low concentration curve must be close to 1 it was attempted to correct the scattering curves. The scattering due to aggregates was calculated and subtracted from all curves. The corrected patterns were used to calculate the structure factors of the concentration series. The highest concentration structure factor (Figure 7.27) was used to obtain the potential parameters with the program winOZ. A surprising feature of the structure factor in deionised water is that its oscillations around 1 can still be observed at very high angles up to s=0.3Å^{-1}. With the estimated structure factor in Figure 7.25 this was not observed mainly because the attractive interactions were completely neglected in the

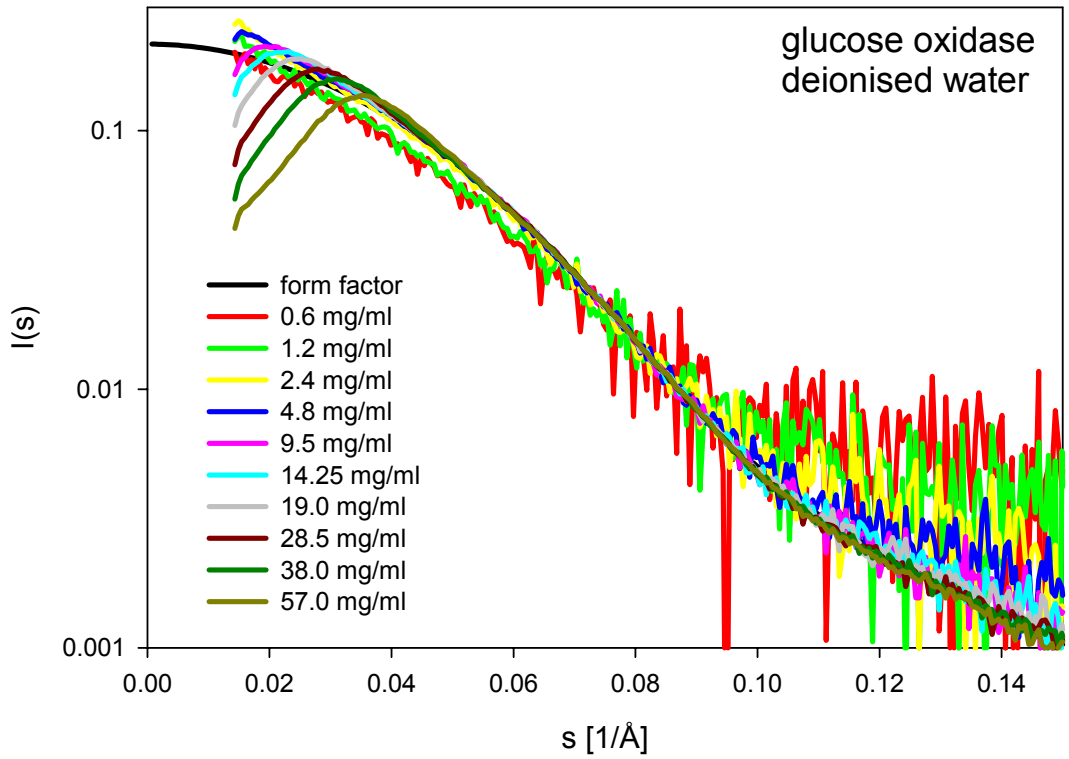


Figure 7.26: Scattering intensities of the glucose oxidase concentration series in deionised water. The samples were slightly aggregated and therefore the onset of the scattering pattern lies even for the low concentrations higher than the form factor.

computation and the hard-sphere radius was, obviously, underestimated. The parameters of the potential model that fits the experimental data best is presented in Table 7.5.
The effective charge of 22.2 is considerably lower than previously estimated. This is not entirely unexpected since the pH in deionised water is not precisely defined and can be closer to the pI of glucose oxidase (4.5) than to neutral pH (7.0). The value of the Debye length (30.3 Å) is similar to the one observed for the lysozyme solutions (26.3 Å) and is equivalent to an ion concentration of approx. 10 mM reflecting once more the fact that not all ions were removed during dialysis.
Although the absolute value of the effective charge is higher for glucose oxidase than for

parameter	best fit value
concentration	57 mg/ml
effective charge	22.2
Debye length	30.3 Å
depth of attr. potential	-8.3 *kT* (at 20ºC)
range of attr. potential	3.0 Å
hard sphere radius	50.0 Å
molecular mass	160 kDa
temperature	15ºC

Table 7.5: Simulation parameter values for the pair potential of glucose oxidase in deionised water.

lysozyme the effective charge per surface area is only 7.1x10^{-4} Å^{-2} whereas it is 2.0x10^{-3}Å^{2} for lysozyme. Thus the repulsive potential plays a larger role for lysozyme.
Based on the literature [Tardieu, 1999] the range of the attractive potential was set, as for the lysozyme simulations, to 3Å. The high value of -8.3 *kT* (at 20ºC) for the depth of the attractive potential is a surprise since according to the DLVO theory it should be the same value for all proteins. For lysozyme it was -5.3 *kT* (at 20ºC). On the other hand a weaker attractive potential cannot account for the main features of the structure factor such as the extended oscillation around 1. A hard-sphere radius of 50Å fits the experimental data best and exceeds R_g=35Å by approx. 50%. The hard-sphere radius is directly correlated with the period of the oscillation of the structure factor around 1. In contrast for lysozyme the hard-sphere radius (R=16.2Å) was approximately equal to the radius of gyration (R_g=15Å). The values for molecular mass, concentration and temperature were determined during the measurements.
Based on this pair potential model for glucose oxidase the structure factors for the entire concentration series in deionised water were calculated and compared to the measurements (Figure 7.28). The simulated curves qualitatively reproduce all features of the experimental structure factors. Remarkable is first of all the concentration dependence of the *s*-value, at which the structure factor becomes larger than 1. Whereas the structure factor at 57 mg/ml is approx. 1 at s=0.032Å^{-1}, it is only s=0.019Å^{-1} for 95 mg/ml. This feature can also be found in the simulations but here the change of *s* is smaller (below 0.01 Å^{-1} between the two measurements).

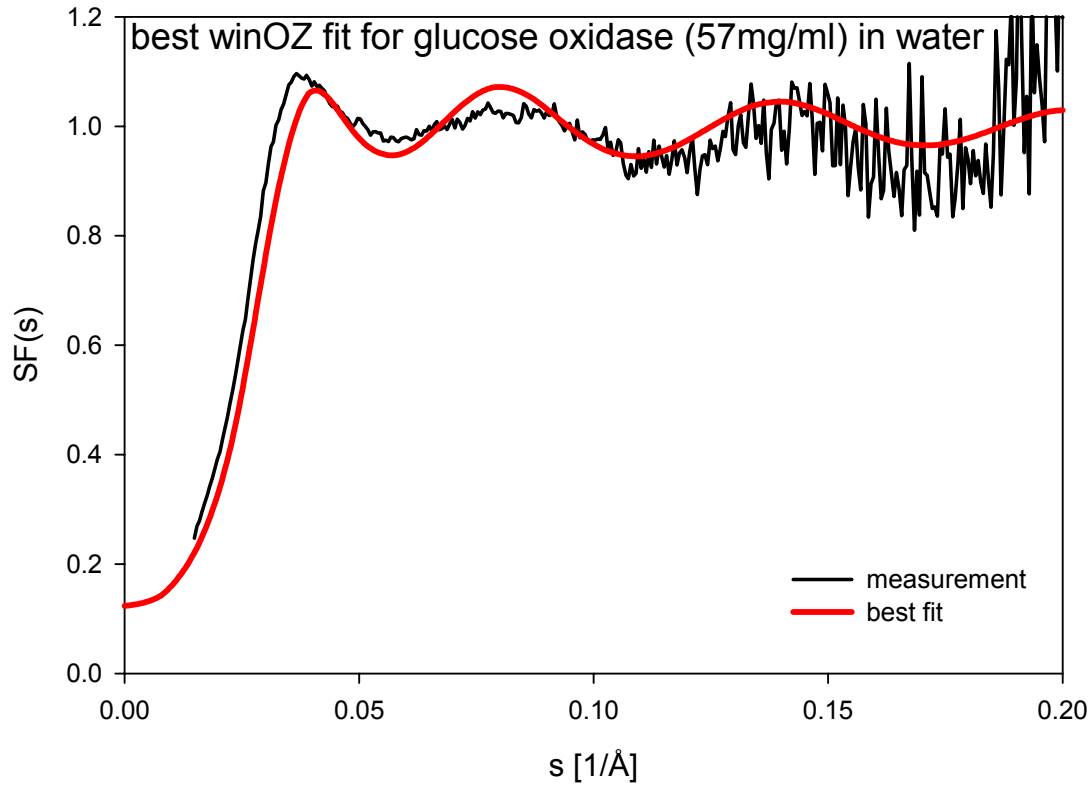


Figure 7.27: Structure factor (57 mg/ml) in deionised water and the best fit structure factor calculated with winOZ. The pair potential parameters can be found in Table 7.5.

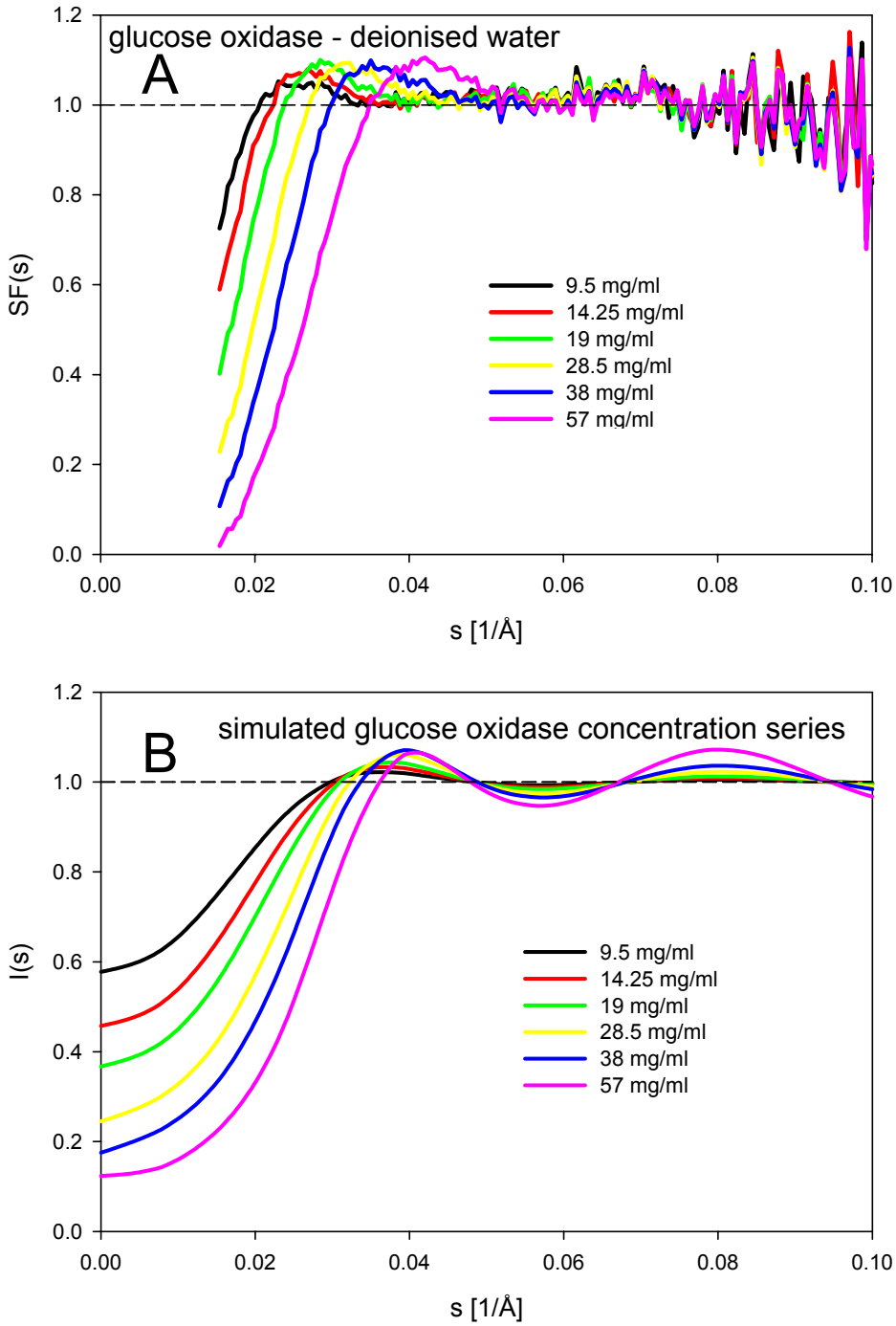


Figure 7.28: Best fit winOZ structure factor (z=22.2, R=35Å, λ_D=30.3Å) for the measurement at 57 mg/ml (B). The structure factors at lower concentrations are calculated with the same parameters. The concentration series of structure factors (A) calculated from the scattering patterns presented in Figure 7.24.

7.9 KCl Salt series of glucose oxidase

A KCl concentration series (0-300mM) was measured with glucose oxidase in deionised water in the same way as the series for lysozyme. Due to the increase of salt shielding effects with KCl concentration it was possible to turn the lysozyme interactions from repulsive to attractive. For high salt concentrations the onset of the structure factor was larger than 1. This feature was not observed with glucose oxidase. Even at a KCl concentration of 300 mM the onset of the structure factor remains below 1 (Figure 7.29). This is due to the size of the molecule. For a large molecule the repulsive interactions due to the hard-sphere core of the protein are so large that even with the same set of parameters for the attractive potential as used for lysozyme the onset of the structure factors remains below 1.
The best-fit parameters for the potential model were found with the following strategy: All parameters apart from the Debye length and the depth of the attractive potential were the same as in Table 7.5. The Debye length (Table 7.6) was calculated according to the salt concentration. Assuming that in the dialysed sample without added salt (0 mM) an inherent

KCl [mM]	λ_D [Å]
0	26.3
5	22.4
10	19.8
20	16.5
50	12.0
100	8.9
175	6.9
250	5.8

Table 7.6: Values for the Debye length (λ_D) taking into account the inherent ionic concentration of 10 mM, for the measurements of glucose oxidase at different KCl concentration in deionised water.

salt concentration of 10mM remains.
As for the lysozyme series the depth of the attractive potential had to be adapted to the individual measurements. It was found that as in the case of lysozyme the attractive potential of glucose oxidase decreases if the KCl salt concentration in the buffer is raised (Table 7.7).

KCl [mM]	depth of the attractive potential
0	-8.3
5	-6.6
20	-5.3
50	-4.8
100	-3.8
150	-3.2
300	-2.7

Table 7.7: Values for the depth of the attractive potential (in kT at 20ºC) yielding the best fit to the experimental data.

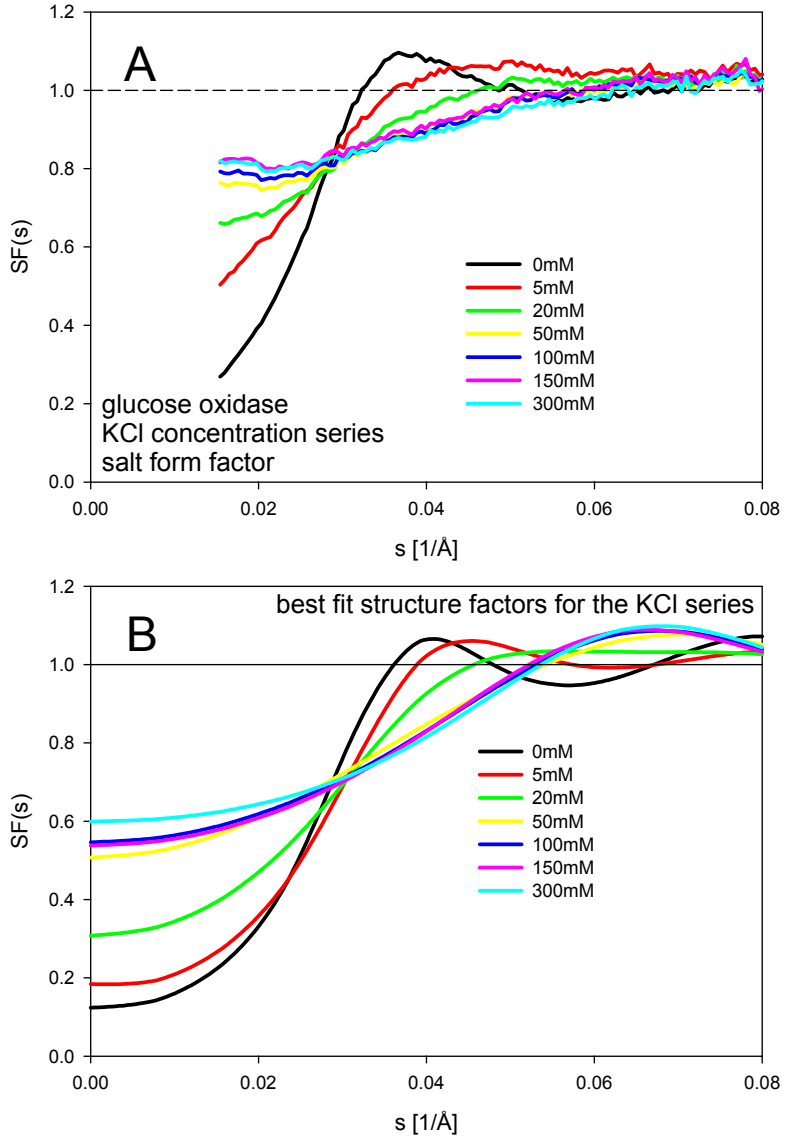


Figure 7.29: Structure factors in deionised water for the glucose oxidase (57mg/ml) KCl salt series. The simulated curves were obtained with a fixed net charge (z=22.2) and a variable attractive potential. The parameters for the depth of the attractive potential are given in Table 7.7.

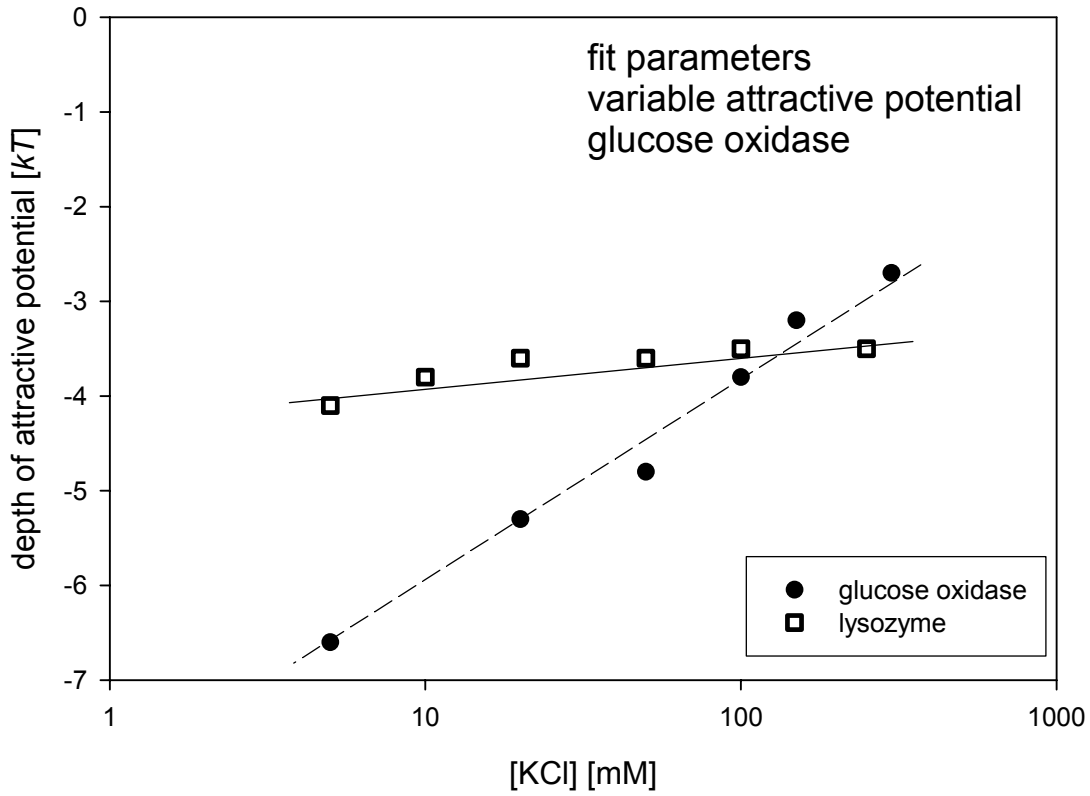


Figure 7.30: Depth of the attractive potential as a function of KCl concentration for glucose oxidase and lysozyme. The lines are only present to guide the eye.

Figure 7.30 illustrates that the attractive potential is deeper for glucose oxidase than for lysozyme at low salt concentrations. The concentration dependence is also higher and at high salt concentrations the potential is less deep than that of lysozyme.
The simulations can well reproduce the most apparent feature of the concentration series, i.e. the point in which all curves cross (around s=0.028Å^{-1}). Especially at low salt concentrations the models fit the experimental data well but with increasing salt concentration the deviation at the onset becomes larger. This could be a consequence of the inappropriate representation of the molecules as spheres. The repulsive interactions in this regime are relatively small and therefore the molecules can get in closer contact. Indeed other simulations show that with a hard-sphere radius of R=42Å it is possible to fit the experimental data at 300 mM KCl better (Figure 7.31). On the other hand with R=42Å it is impossible to reproduce the oscillation around 1 as found at high angles. This is a strong indication that the asymmetry of the particles plays a role in this context.
The potentials for the lowest and highest salt concentrations are illustrated and compared to those for lysozyme in Figure 7.32. At high salt concentration for glucose oxidase the repulsive potential is no longer significant any more and the structure factor does not change between 150 and 300 mM. All parameters for the computations are given in Table 7.8.
Since the pI of glucose oxidase lies between 4 and 5 the net charge of the protein in deionised water is negative. If ions are present in the solution the positive ions will form the first layer of counterions. In order to check whether the size of the ions effects protein interactions concentration series of glucose oxidase in LiCl, NaCl and KCl were measured and their structure factors were compared. Although there are small differences in the onset of the curves there are no reproducible and significant effects for the different cations, i.e.

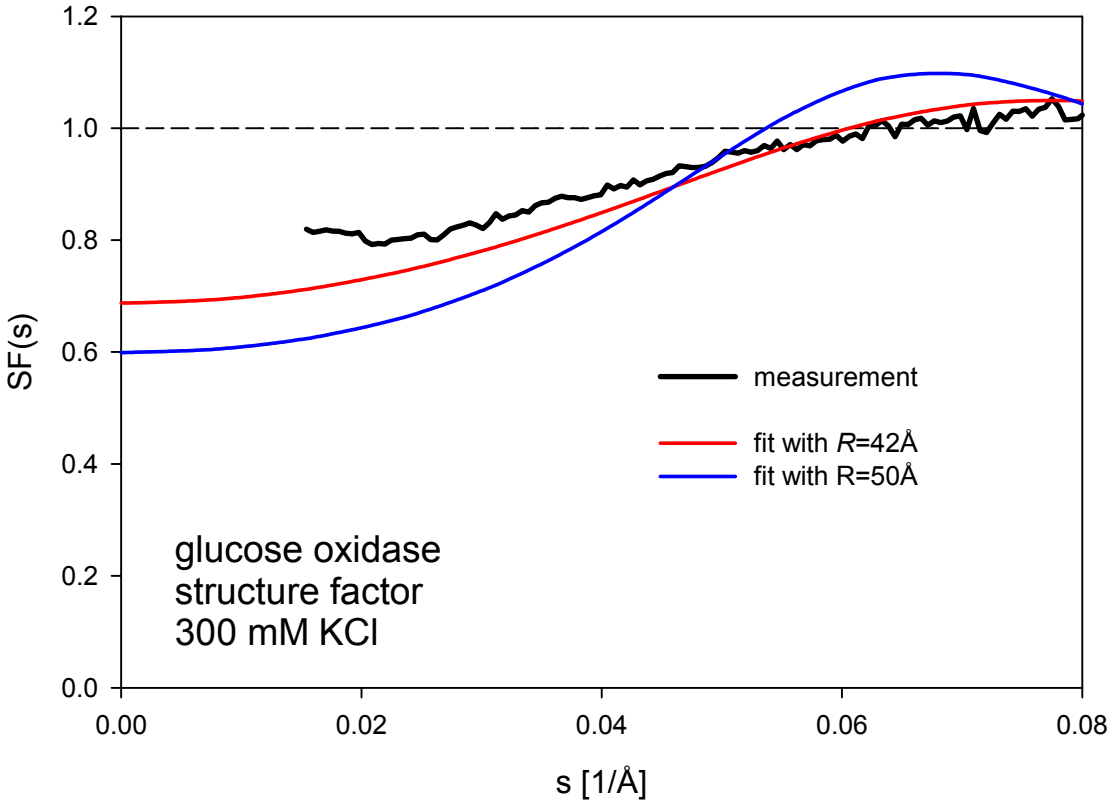


Figure 7.31: The pair potential model fits the experimental structure factors better at low angles in the attractive regime (300 mM KCl concentration) if a lower value for the hard-sphere potential (R=42Å) is used. This indicates that the restriction of the model to spherical proteins limits the precision of the approach.

the order of the curves is different for different salt concentrations. In Figure 7.33 one can see for instance that the repulsive interactions are the strongest in 8mM KCl whereas they are the weakest at 230mM. It would seem unwise at this point to discuss such small effects.

parameter	Glucose oxidase; 0 mM KCl	Glucose oxidase; 300 mM KCl	lysozyme; 0 mM KCl	lysozyme; 300 mM KCl
effective charge	-22.2	-22.2	6.5	6.5
Debye length	30.3 Å	5.5 Å	26.3 Å	5.8 Å
depth of attr. Potential	-8.3 kT	-2.7 kT	-5.3 kT	-3.5 kT
range of attr. Potential	3.0 Å	3.0 Å	3.0 Å	3.0 Å
hard sphere radius	50.0 Å	50.0 Å	16.2 Å	16.2 Å
molecular mass	160 kDa	160 kDa	14.3 kDa	14.3 kDa
temperature	15ºC	15ºC	15ºC	15ºC

Table 7.8: Pair potential parameters for the simulations, which lead to the pair potentials in Figure 7.32.

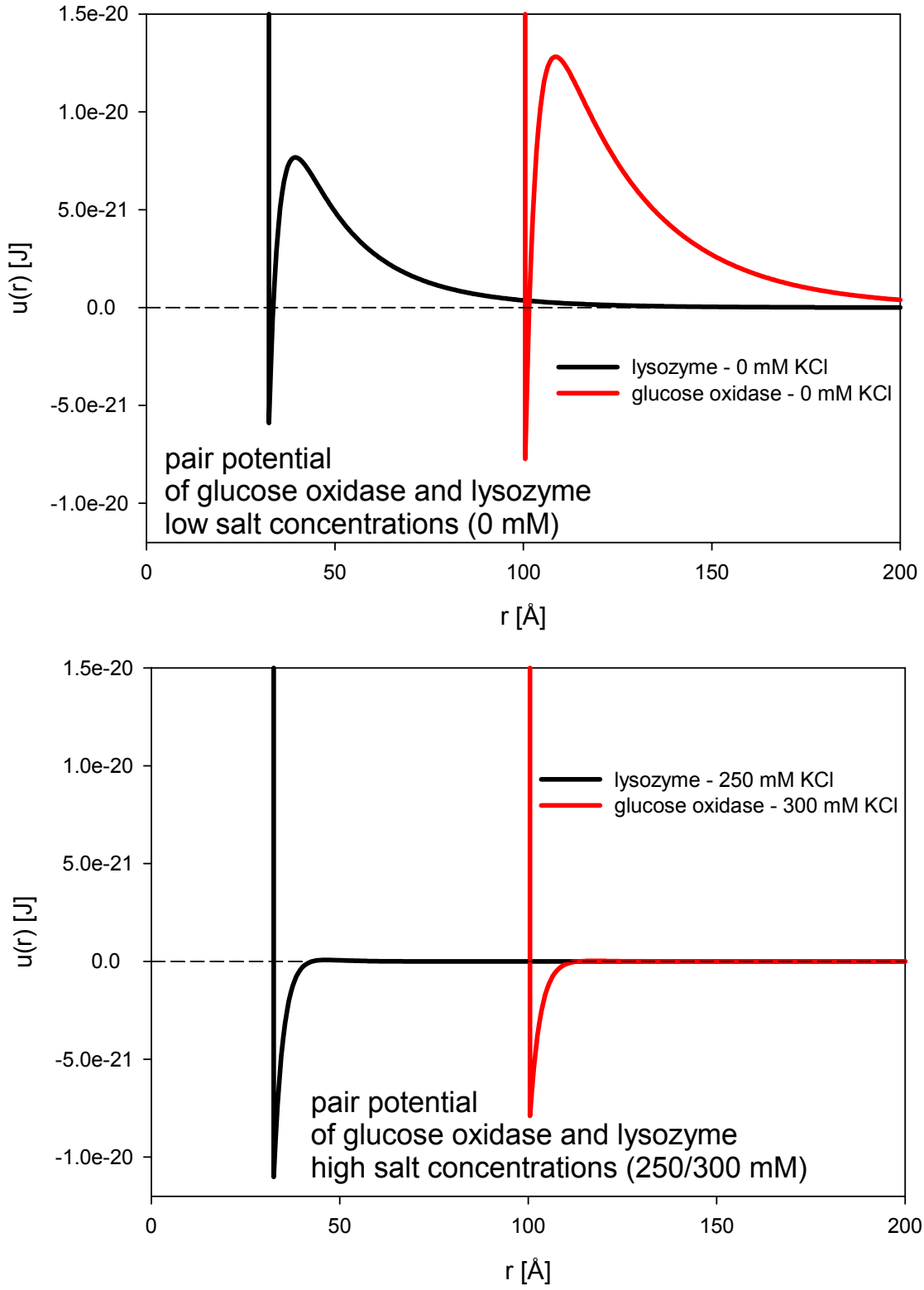


Figure 7.32: Comparison of the potentials for glucose oxidase and lysozyme at high and low salt concentrations.

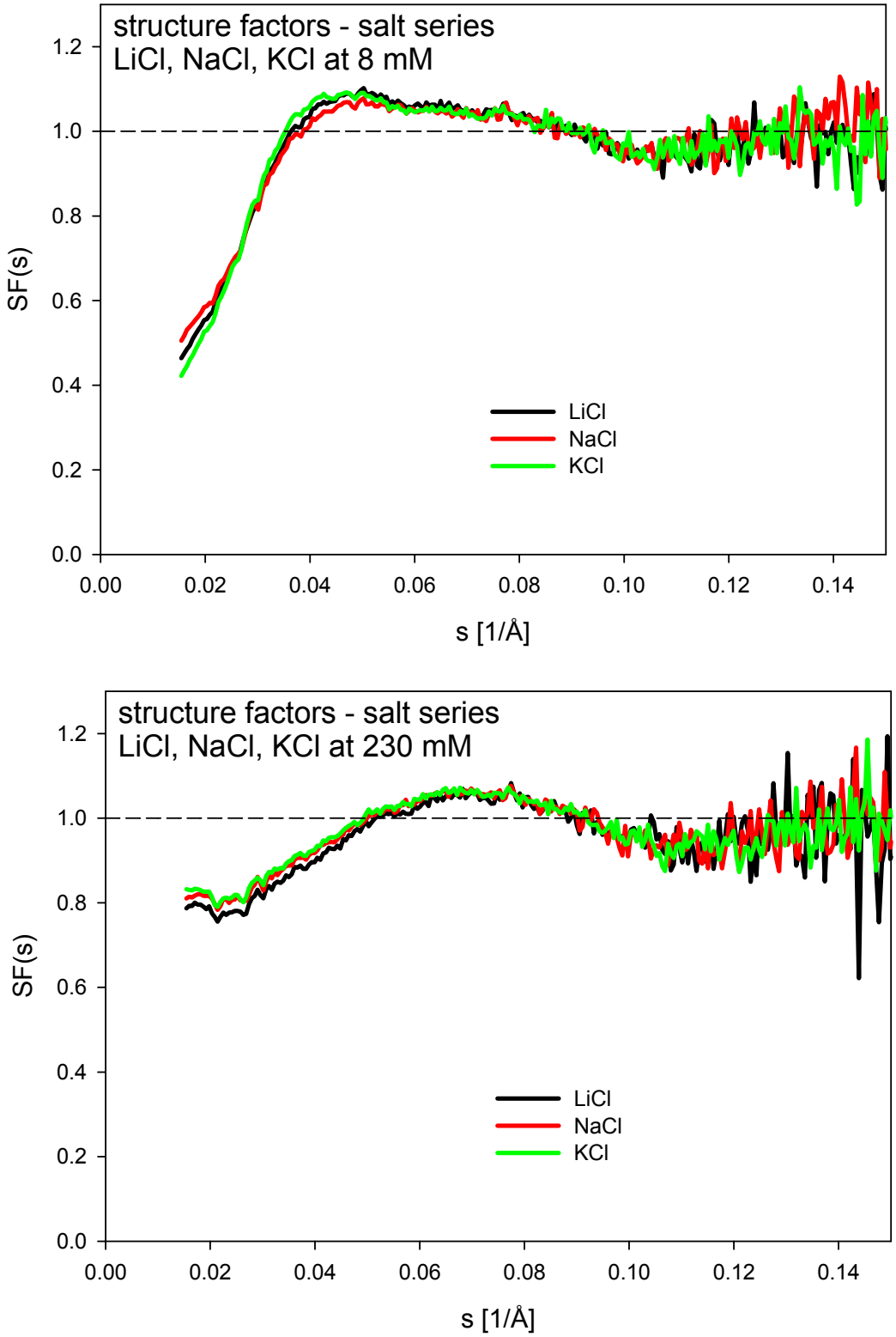


Figure 7.33: Glucose oxidase interactions in the presence of different salts (LiCl, KCl, NaCl).

7.10Glucose oxidase at pH 4.5, 7 and 9

The net charge of the protein depends on the pH. In deionised water the pH is not well defined and difficult to determine since deionised water has no buffering capacity. Therefore glucose oxidase was dialysed against 100 mM sodium phosphate at pH 4.5, 7 and 9. The final concentration of the samples after dialysis was 50.8 mg/ml. Since the isoelectric point of glucose oxidase is between 3.9 and 4.3. The repulsive interactions were expected to be rather weak at pH 4.5 and strongest at pH 9, whereas one can calculate from the protein sequence that the change of net charge from pH 7 to pH 9 is relatively small (Figure 7.34). Looking at the onset of the structure factor in Figure 7.34 the repulsion at

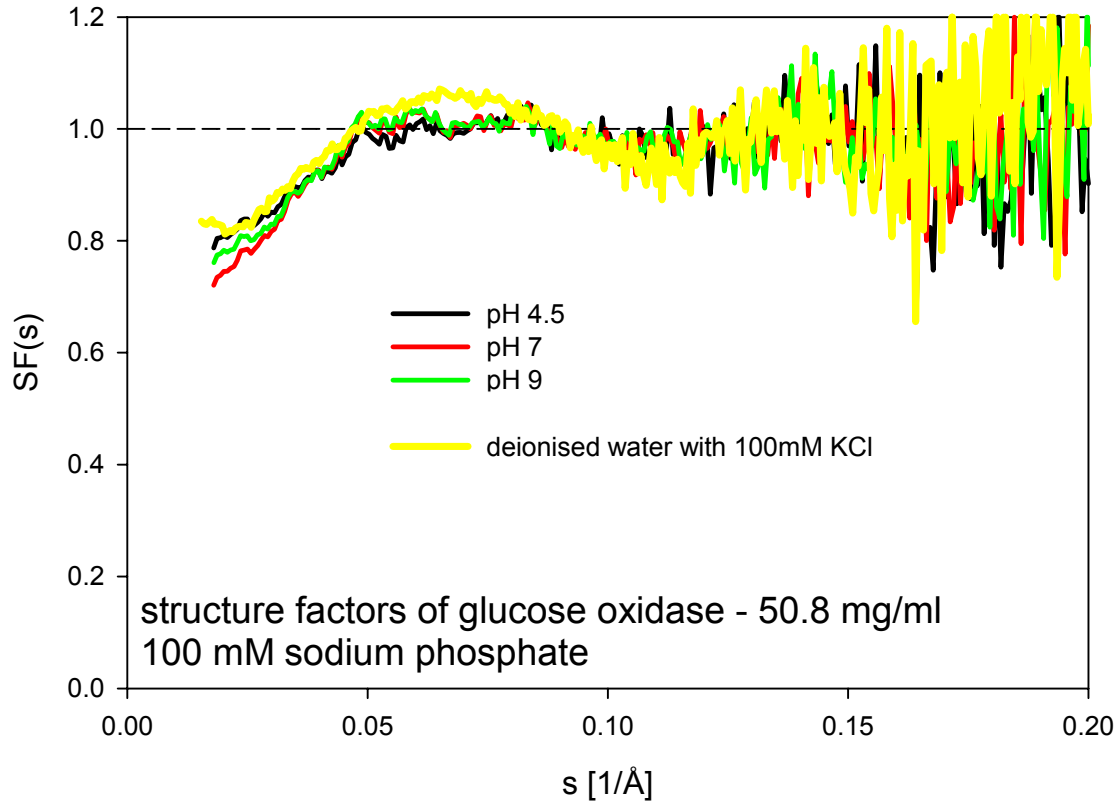


Figure 7.34: Glucose oxidase (50.8 mg/ml) structure factors at pH 4.5, 7 and 9 in 100 mM sodium phosphate.

pH 4.5 is clearly the weakest but surprisingly the onset of the structure factor at pH 9 is higher than at pH 7 and therefore also the repulsion appears to be lower. The order of the curves changes at other positions in the scattering curve (0.04 $Å^{-1}$<s<0.07 $Å^{-1}$) and it thus is reasonable to assume that minor aggregation effects affect the onset of the structure factors. Using the program winOZ it is possible to fit the effective charge of the pair potential to the entire structure factor and the charges, which it finds, fit to the prediction from the sequence data. (z=27.7 at pH 4.5, z=30.0 for pH 7 and z=30.3 for pH 9). This example demonstrates the advantage of SAXS over static light scattering because it is useful to measure the entire structure factor instead of only the second virial coefficient A_2.

7.11 Temperature dependence of glucose oxidase interactions

The temperature dependence of glucose oxidase interactions was measured in deionised water at 15, 40 and 60ºC at a protein concentration of 50 mg/ml (Figure 7.35) in the same way as for lysozyme. For lysozyme it was surprisingly found that the deviation of the structure factor from 1 increases with temperature.

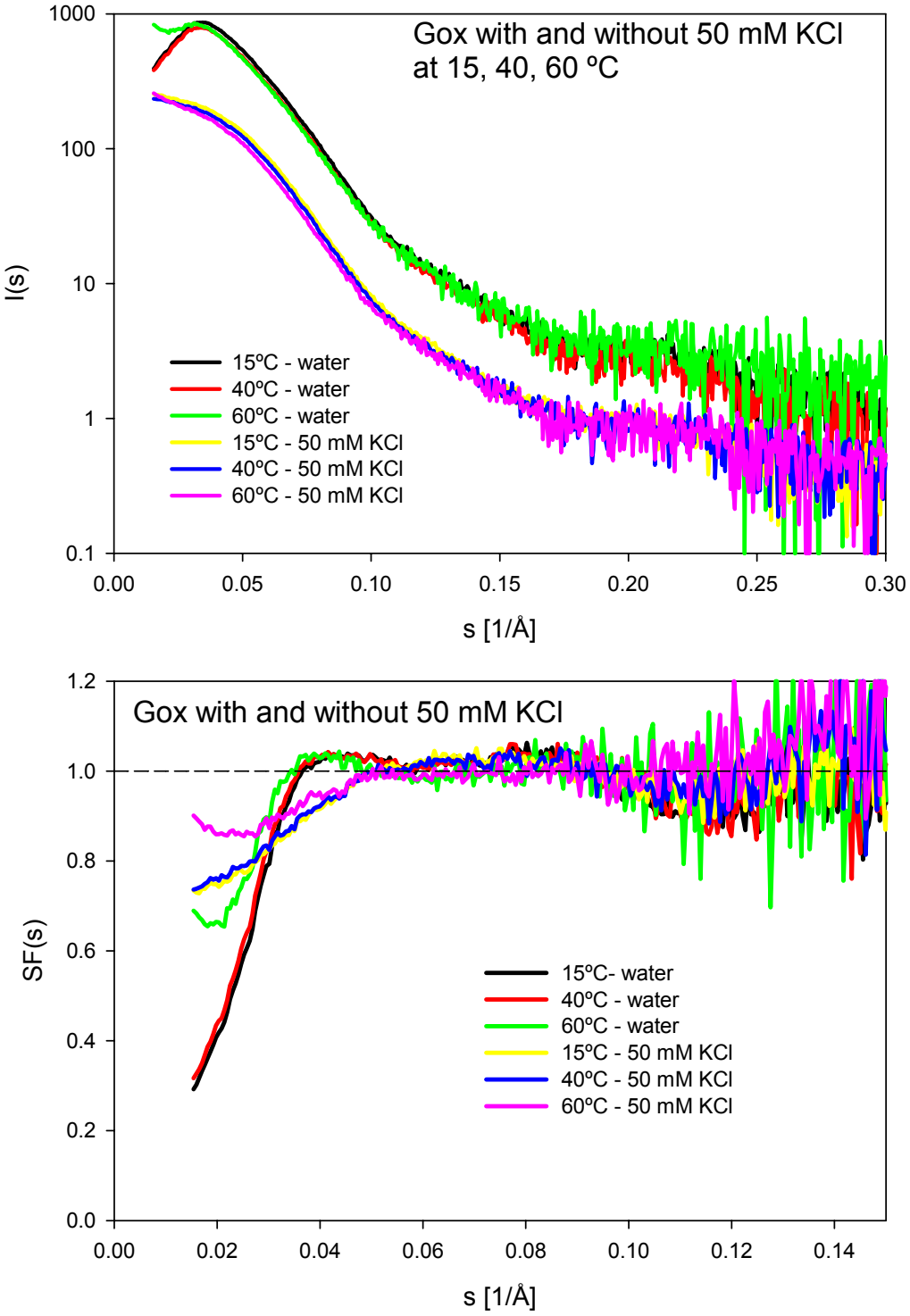


Figure 7.35: Glucose oxidase interactions (50 mg/ml) at 15, 40 and 60ºC. The samples at 60ºC are clearly aggregated. The measurements at lower temperature are in agreement with the theory.

Unfortunately glucose oxidase formed aggregates at 60ºC and does not have the good thermal stability of lysozyme. The differences between 15 and 40ºC are very small. If there is any difference between the structure factors the deviation of the structure factor from 1 is somewhat smaller for elevated temperatures. This finding would be in agreement with the DLVO theory since the temperature increase goes along with an increase of the average energy per macromolecule. If one assumes that the potentials do not depend on temperature the ratio between particle energy and van der Waals and electrostatic potential increases. Therefore the motion of the proteins is less restricted by the potentials and the deviation of the structure factor from 1 becomes smaller.

7.12 Bovine serum albumin (BSA)

As a third example a concentration series (4.3 – 52.8 mg/ml) of the protein bovine serum albumin (BSA) was studied. The protein was measured in 50 mM HEPES at pH 7.5 and at 20ºC. Its radius of gyration is ca. 30Å and the molecular mass is 67500 kDa. The structure factors of the concentration series could be reproduced with the program winOZ with the parameters in Table 7.9 where the value for the hard-sphere radius was set to that of the radius of gyration. The value for the Debye length (12.4Å) was calculated assuming a salt concentration of 50 mM. The range of the attractive potential was set to 3Å as in all previous simulations. The best fit model for the experimental structure factors has an effective charge z=25 and the depth of the attractive potential is 3.2 kT (at 20ºC). Figure 7.36 illustrates that the model describes the data of the entire concentration series well. The precision of the model parameters is limited by the fact that only one protein concentration series with fixed salt concentration was measured. If a salt concentration series were

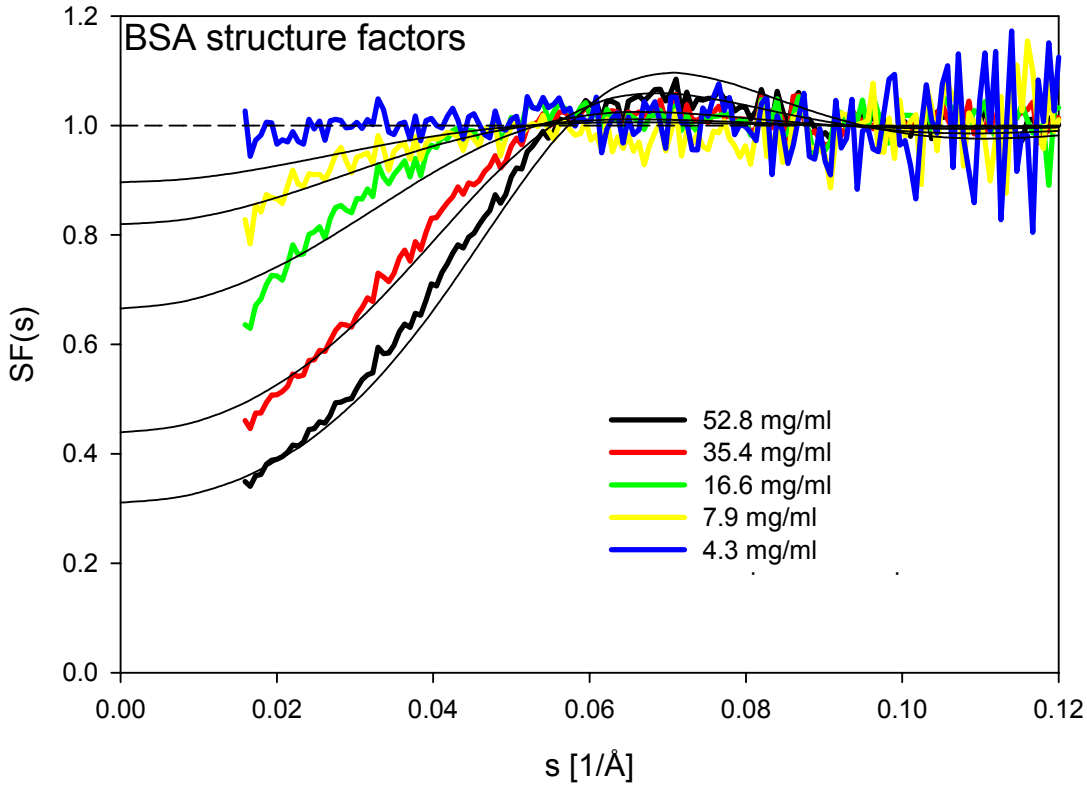


Figure 7.36: Structure factors of the BSA concentration series. The black solid lines represent the best fit model structure factors calculated with the program winOZ and the pair potential parameters given in Table 7.9.

parameter	BSA in 50 mM HEPES pH 7.5
effective charge	25.0
Debye length	12.4 Å
depth of attr. potential	-3.2 *kT*
range of attr. potential	3.0 Å
hard sphere radius	30.0 Å

Table 7.9: Pair potential parameters for the simulations of the structure factor of BSA, which lead to the structure factors in Figure 7.1.

available parameters such as the hard-sphere radius could be determined with a higher precision by measuring in the attractive regime at very high salt concentrations.
In summary the measurements show that the DLVO theory and the program winOZ lead to reasonable parameter values for globular proteins in the range of molecular mass from about 15 kDa to 200 kDa.

8 Conclusion and outlook

The most reliable structure factors for lysozyme, glucose oxidase and BSA were obtained by dividing the experimental scattering patterns at high concentrations by the form factors determined with the program GNOM [Svergun, 1992]. Especially in the case of glucose oxidase interaction effects are visible even at large angles (s<0.2 Å^{-1}).
The radius of gyration of glucose oxidase was found to be 35.0Å. The radii of gyration of BSA and lysozyme were in agreement with previous measurements [Shang, 2003]. The carbohydrate moiety of glucose oxidase significantly influences the scattered intensity and a model of the protein including the carbohydrate moiety, which fits the experimental data, was generated with the program CREDO [Petoukhov, 2002].
The simulations show that the program winOZ, written as part of this work, can be used to calculate the structure factors of protein solutions from the pair potential. The size of the proteins (in our experiments ranging from 14.3 kDa for glucose oxidase to 160 kDa for lysozyme) does not affect the usefulness of the method (DLVO theory implemented in the algorithms of the program winOZ). The main features of the structure factors observed in the protein concentration series are correctly predicted by the program. The results for lysozyme are also in good agreement with similar ones in the literature [Tardieu, 1999].
The program was further used to systematically study the effects of the variation of some of the parameters in the pair potential.
It was found that salt concentration series of structure factors could be better described if the strength of the attractive potential decreases with increasing salt concentration. In the literature [Tardieu, 1999] such concentration series were described using a variable electrostatic potential only. As our measurements of lysozyme solutions in the presence of 250 mM TMAO can only be explained using a variable attractive potential, we revised some of our earlier calculations on the salt series on lysozyme and also found a better agreement. The sensitivity of the structure factors to the attractive potentials depends on whether the measurements are made in the attractive or in the repulsive regime, with measurements in the attractive regime being more precise (±0.1 kT at 20ºC) than in the repulsive regime (±0.5 kT at 20ºC). The structure factors of glucose oxidase and lysozyme KCl concentration series were compared. Although in these cases the potentials change in a similar manner the change in the structure factor is quite different, e.g. even in the attractive regime (high salt concentration) the glucose oxidase structure factor does not become larger than one at its onset due to the large hard-sphere interactions. In the case of glucose oxidase the attractive potential is surprisingly deep at low salt concentrations. The origin of this strong attraction could be the carbohydrate moiety, which according to the model covers part of the protein surface. The value of the hard-sphere radius of glucose oxidase is about 50% larger than its radius of gyration, whereas in the case of lysozyme these values are similar. Further it was found that it does not make a significant difference whether LiCl, NaCl or KCl are used to weaken the electrostatic potential by charge screening.
Measurements in the presence of the cosolutes urea and TMAO (250 mM each) showed that it is possible to measure and quantify the well-known effects of these substances on protein stability and interactions.
For measurements at different pH-values the net charge of glucose oxidase changes in agreement with the prediction from its sequence. These results illustrate the advantages of SAXS structure factors over the use of the isolated A_2 value measured for example by

static light scattering. Small aggregation effects would have led to an artificially high A_2 value but fitting the pair potential to the entire structure factor leads to satisfactory results.
The change in the structure factor of glucose oxidase as a function of temperature is very small and its direction is, in contrast to the case of lysozyme, in agreement with DLVO theory.
A consideration of the statistical errors proves that DLVO theory can not entirely explain the structure factors and significant deviations between the model and measurements remain. In particular for glucose oxidase in the attractive regime (at high salt concentrations) the onset of the structure factor fits the experimental data better if a radius of 42Å (instead of 50Å required to explain the oscillation of the structure factor around 1 at large *s*-values) is used. This strongly suggests that the asymmetry of the particles, which is not accounted for in the DLVO theory, plays a role in this context.

8.1 Outlook

During the process of measuring, analysing and discussing our data many questions came up that remain unanswered. It would, for instance, be interesting to measure the structure factors of lysozyme at high angles with a higher precision. Better knowledge of the lysozyme structure factor at high angles would facilitate the determination of the hard-sphere radius with a higher precision. Such a measurement could be made with a quadrant detector, which effectively measures $sI(s)$ rather than $I(s)$ thereby giving better statistics at higher angles.
The origin of the surprisingly large value for the depth of the attractive potential for glucose oxidase is still unknown. One possible cause is the presence of the carbohydrate moiety at the protein surface. For the crystallisation of glucose oxidase the sugars were cleaved by an enzymatic reaction [Kalisz, 1990]. In general, it would be interesting to compare the structure factors of glucose oxidase with and without carbohydrates and compare the values for the attractive potential especially at low salt concentrations. Attempts were made to reproduce this procedure but they did not lead to immediate success. A problem is also the relative large amount of protein needed for a structure factor measurement (approx. 50 mg for a salt series).
In all salt series presented here only monovalent ions were used. Recently in various systems evidence was found that divalent salt ions can cause an attraction between equally charged macromolecules ([Netz, 2003], [Angelini, 2003]). It would be interesting to test whether this behaviour can also be detected by SAXS structure factor measurements.
The glucose oxidase KCl concentration series provided experimental evidence that the (symmetric) hard-sphere potential can not account for all features of the structure factors for one fixed *R*-value. It would be desirable to have asymmetric hard-core potentials (Figure 8.1.A) and also asymmetric charge distributions (multipole moments). For such models all quantities (such as pair distribution function, pair potential) are functions of the position (R) and orientation (Ω_1, Ω_2) of the macromolecules in a fixed frame of reference, e.g. the pair distribution function:

$$g = g(R, \Omega_1, \Omega_2) \tag{8.1}$$

It would also be interesting and highly relevant from a biochemical point of view to investigate mixtures of two proteins A and B (Figure 8.1.B). Apart from the pair potentials of the two protein species, $u_{AA}(r)$ and $u_{BB}(r)$, the mixed pair potential between different proteins $u_{AB}(r)$ would be experimentally accessible [Vivarès, 2002].
Although the theory of liquids provides some approaches, which can be used with mixtures and asymmetric potentials [Hansen, 1986] the analytical expressions become very complex.

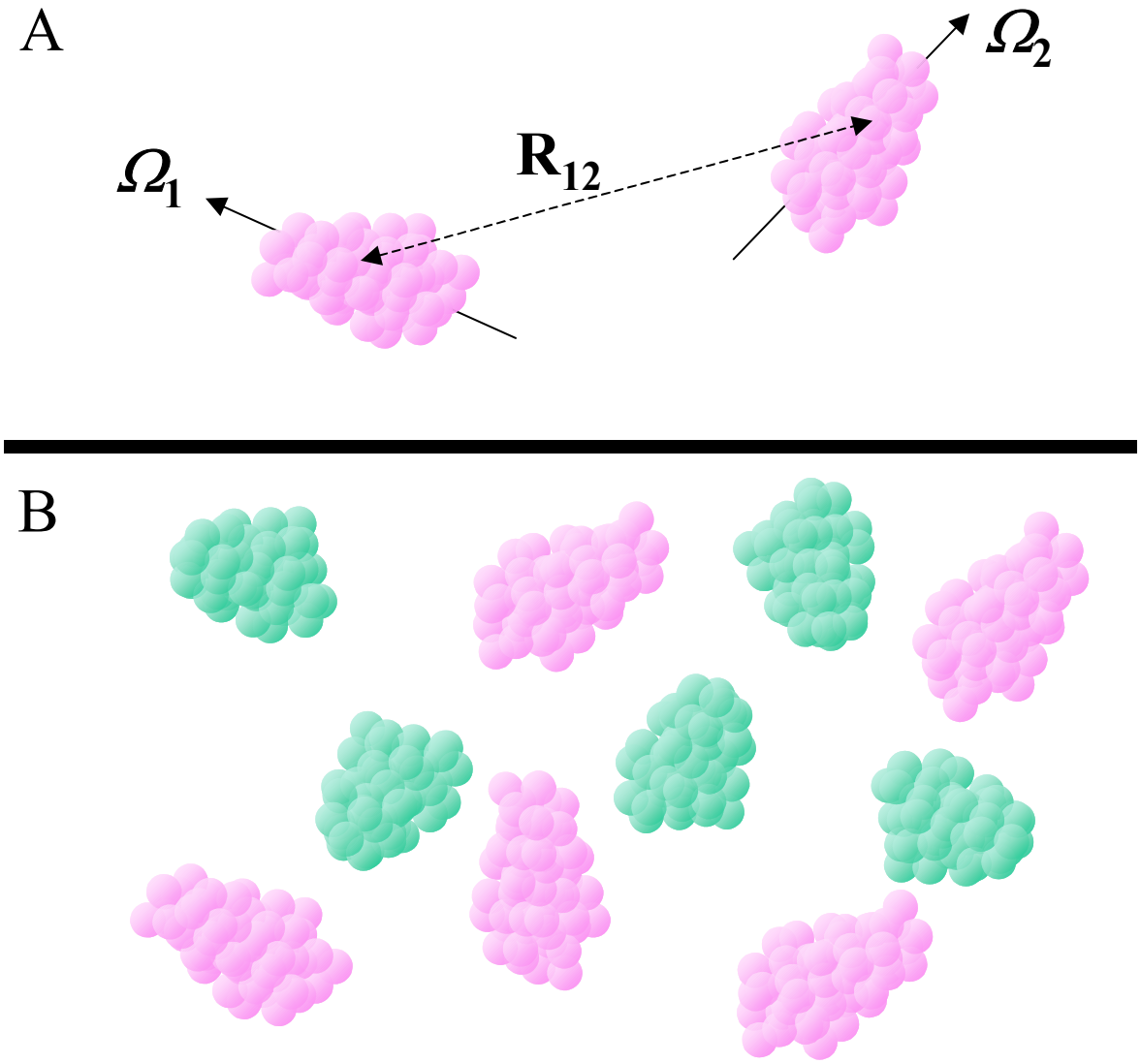


Figure 8.1: If the protein interaction models are extended to asymmetric mass and charge distributions the orientation of the molecules ($\mathbf{\Omega_1}$, $\mathbf{\Omega_2}$, Part A) has to be taken into account in all calculations and simulations. Mixtures (B) of two different proteins are interesting objects for further studies because this could provide information about pair potential between different proteins.

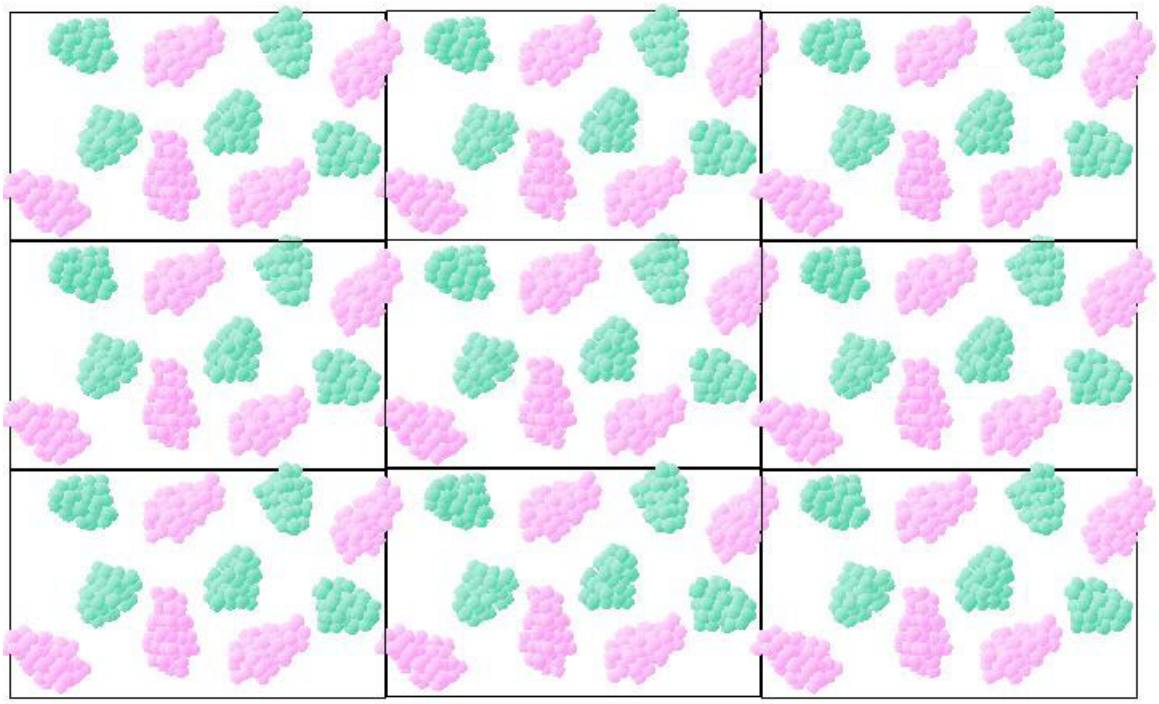


Figure 8.2: Periodic boundary condition used in molecular dynamics simulations ("computer experiments").

Alternative methods to study the implications of pair potentials on the structure of solutions are molecular dynamics simulations ("computer experiments", [Hansen, 1986]). A number of objects of interest (in this case proteins) are put in a virtual "box". Their motion is then simulated as a random walk, due to diffusion, influenced by the interactions with the environment, i.e. the other proteins in the volume. The problem of the limited volume is circumvented by introducing periodic boundary conditions (Figure 8.2). It is relatively easy to implement all kinds of potentials in these models. The large computing power needed for such molecular dynamics simulations, especially if one is interested in asymmetric, orientation-dependent potentials is a limitation.
Although the structure factors measured in SAXS experiments yield probably the most accurate results on physical protein interactions it is interesting to compare the results to other techniques, which require less experimental effort. One candidate that was already mentioned is the measurement of the second virial coefficient A_2 with static light scattering. In the literature (e.g. [George, 1994]) the values for A_2 for various proteins are documented. Knowledge of the sequence, or even the crystallographic structure, of the proteins make it easier to estimate the pair potential form our results. The potentials can be used to calculate the A_2-value and compare it to the literature. The following formula is taken from [Bonneté, 2002]:

$$A_2 = 2\pi \frac{N_a}{M^2} \int r^2 \left[1 - \exp\left(\frac{-u(r)}{k_B T} \right) \right] \mathrm{d}r \qquad (8.2.)$$

Calorimetric methods, such as isothermal titration calorimetry (ITC), provide another interesting approach. This device measures the differential heating power needed to keep sample and reference cell at the same temperature and the heat of dilution of the protein is

measured as illustrated by the result of a preliminary experiment in Figure 8.3. At the beginning of the experiment the sample cell was filled with buffer only. Then a small amount of protein is injected and thus diluted. If the heating power decreases upon dilution the interactions between the proteins are repulsive. Energy is released. For attractive interactions it is the other way round. The excess internal energy can be calculated from the pair potential and the pair distribution function [Hansen, 1986]:

$$U^{ex} = \frac{N^2}{2V}\int u(r)g(r)\mathrm{d}r \tag{8.3}$$

This should allow a direct comparison of the results of ITC dilution experiments with theoretical results.

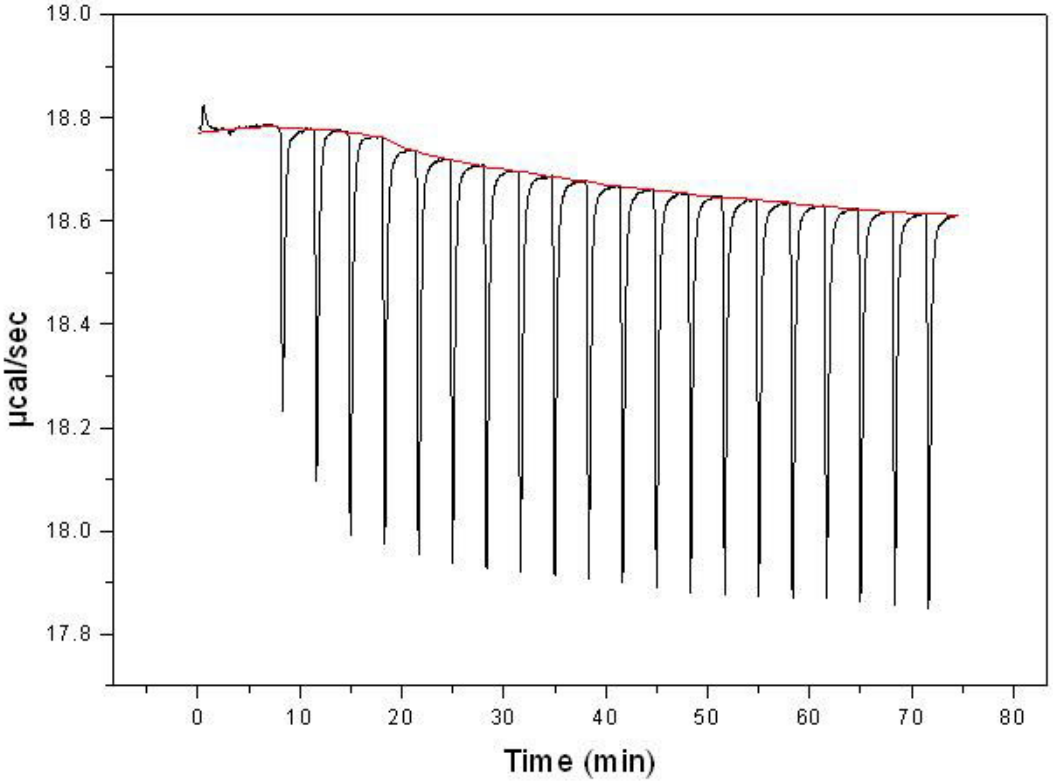


Figure 8.3: Experimental data of an ITC experiment on glucose oxidase. The differential heating power is plotted against the time elapsed since the beginning of the experiment. The peaks are due to injections of protein into the sample cell. The decay of heating power indicates the repulsive interactions between the macromolecules. Energy is released when the protein is diluted (heat of dilution).

9 Concluding remarks

The work of the last three years gave me the opportunity to learn about several aspects of the interactions of X-rays with matter and their importance for experimental studies, to become familiar with all elements of a SAXS beamline and to spend much time working with gas detector systems and their readout electronics.
The work the time-stamp TDC system proved that the detection of multiple events in a Gabriel type detector is possible and that this can be used to minimise the effects of fluorescence of the detector gas on the measurement. It also suggests that time-stamp TDCs may be useful on pulsed X-ray sources for experiments with high time resolution.
During the second part of the work, I learned everything needed to set up SAXS experiments on protein systems; from the preparation of the sample to the accomplishment of the actual measurement to the analysis and discussion of the data. The measurements lead to several findings about the physical interactions of the proteins lysozyme and glucose oxidase. An important result is that varying the attractive potentials can account for all features of the structure factors in a concentration series. Previously others had attempted, less successfully, to explain the structure factors with a variable repulsive potential. It also was possible to show that the change of attractive potential due to some physiological important cosolutes, such as urea, glycine and TMAO, can be measured and quantified by measuring the SAXS structure factors.
As seems so often the case, it appears that more questions were opened than could be answered.

10 References

[Angelini, 2003] Angelini, T.E., Liang, H., Wriggers, W. & Wong, C.L. (2003) "Like-charge attraction between polyelectrolytes induced by counterion charge density waves", PNAS **100**(15), 8634-8637.

[Atkins, 1997] Atkins, P. (1997) *Physical Chemistry* (W.H. Freeman and Company, New York).

[Bateman, 2000] Bateman, J. (2000) "The effect of beam time structure on counting detectors in SRS experiments", J. Synchrot. Radiat. **7**, 307-312.

[Bednarek, 1980] Bednarek, B. (1980) "On a new interpretation of the count-rate effects in gas-proportional detectors", Nucl. Instr. and Meth. A **178**, 173-180.

[Belloni, 1993] Belloni, L. (1993) "Inability of the hypernetted chain integral equation to exhibit a spinodal line", J. Chem. Phys. **98**(10), 8080-8095.

[Bennion, 2003] Bennion, B.J. & Daggett, V. (2003) "The molecular basis for the chemical denaturation of protein by urea", PNAS **100**, 5142-5147.

[Bennion, 2004] Bennion, B.J. & Daggett, V. (2004) "Counteraction of urea-induced protein denaturation by trimethylamine N-oxide: A chemical chaperone at atomic resolution", PNAS **101**(17), 6433-6438.

[Berman, 2000] Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. & Bourne, P.E. (2000) "The protein data bank", Nucleic Acids Research, **28**, 235-242.

[Blewett, 1998] Blewett, J. P. (1998) "Synchrotron radiation - Early history", J. Synchrotron Rad. **5**, 135-139.

[Blum, 1993] Blum, W. & Rolandi, L. (1993) *Particle detection with drift chambers* (Springer, Berlin Heidelberg New York) .

[Bonneté, 2002] Bonneté, F. & Vivarès, D. (2002) "Interest of the normalized second virial coefficient and interaction potentials for crystallizing large macromolecules", Acta Cryst. D**58**, 1571-1575.

[Boulin, 1988] Boulin, C.J., Kempf, R., Gabriel, A. & Koch, M.H.J. (1988) "Data acquisition-systems for linear and area X-ray-detectors using delay-line readout", Nucl. Instr. and Meth. A **269**, 312-320.

[Buche, 1989] Buche, A. (1989) "Glycine and other amino compounds prevent chromatin precipitation at physiological ionic strength", FEBS letters **247**(2), 367-370.

[Chacon, 1998] Chacon, P., Moran, F., Diaz, J.F., Pantos, E. & Andreu, J.M. (2000) "Low resolution structure of proteins in solution retrieved from X-ray scattering with a genetic algorithm" Biophysical J. **74**, 2760-2775.

[de Raad Iseli, 2001] de Raad Iseli, C., Reimann, T., Golding, F., Boulin, C., Epstein, A., Beloeuvre, E., Gabriel, A. & Koch M.H.J., (2001) "A data acquisition system for gas proportional detectors with delay line readout based on space-time-space conversion" Nucl. Instr. and Meth. A **467-468** 1152-1155.

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

[Fischer, 2002] Fischer, H., Franz, J., Grunemaier, A., Heinsius, F.H., Hoffmann, M., Karstens, F., Kastaun, W., Königsmann, K., Niebuhr, M., Risken, R., Schmidt, T., Schmitt, H., Schweimler, A., von Hodenberg, M. & Urban, H.J. (2002) "The COMPASS data acquisition system", IEEE Trans. Nucl. Sci. **49,** 443-447.

[Frederick, 1989] Frederick, K.R. (1989) "Glucose oxidase from aspergillus niger", J. Biological Chemisty **265**, 3793-3802.

[Gabriel, 1977] Gabriel, A. (1977) "Position sensitive X-ray detector", Rev. Sci. Instrum. **48**, 1303-1305.

[Gebauer, 2001] Gebauer, B., Schulz, C., Richter, G., Levchanovsky, F.V. & Nikiforov, A. (2001) "Development of a hybrid MSGC detector for thermal neutron imaging with a MHz data acquisition and histogramming system", Nucl. Instr. and Meth. A **471,** 249-253.

[George, 1994] George, A. & Wilson W. (1994) "Predicting Protein Crystallization from a Dilute Solution Property", Acta Cryst. D**50**, 361-365.

[Guinier, 1939] Guinier, A. (1939) "La diffraction des rayons X aux très petits angles: application à l'étude de phénomènes ultramicroscopiques", Ann. Phys. **12**, 161-237.

[Hansen, 1986] Hansen, J.-P. & McDonald, I.R. (1986) *Theory of simple liquids* (Academic Press, Amsterdam).

[Hecht, 1993] Hecht, H.J., Kalisz, H.M., Hendle, J., Schmid, R.D. & Schomburg, D. (1993) "Crystal structure of glucose oxidase form aspergillus niger refined at 2.3Å resolution", J. Mol. Biol. **229**, 153-172.

[Hendrix, 1980] Hendrix, J. & Fürst, H. (1980) "Novel position sensitive detector for X-ray-scattering experiments with storage-rings operating in the single bunch mode", IEEE Trans. Nucl. Sci. **NS-27**, 145-149.

[Herve, 2002] Herve, C. & Torki, K. (2002) "A 75 ps rms time resolution BiCMOS time to digital converter optimized for high rate imaging detectors", Nucl. Inst. and Meth. A **481** (1-3), 566-574.

[Hochachka, 2002] Hochachka, P.W. & Somero, G.N. (2002) *Biochemical Adaption* (Oxford University Press: New York).

[Holmes, 1998] Holmes, K.C. & Rosenbaum, G. (1998) "How X-ray diffraction with synchrotron radiation got started" J. Synchrotron Rad. **5**, 147-153.

[Ito, 1996] Ito, S., Tosaki, M. & Maeda, N. (1996) "Effect of self-induced space charge in a high pressure position-sensitive proportional counter", Nucl. Instr. and Meth. A **368**, 738-744.

[Israelachvili, 1992] Israelachvili, J. (1992) *Intermolecular & surface forces* (Academic Press, London).

[Ivanenko, 1944] Ivanenko, D. & Pomeranchuk, I. (1944), Akad. Nauk. Dokl. **44**, 315 (cited by [Blewett, 1998]).

[Kalisz, 1990] Kalisz, H.M., Hecht, H.J., Schomburg, D. & Schmid, R.D. (1990) "Crystallisation and preliminary X-ray diffraction studies of a deglycosylated glucose oxidase from aspergillus niger", J. Mol. Biol. **213**, 207-209.

[Kaul, 2004] Kaul, O., Private communication (2004).

[Knoll, 1977] Knoll, G.F. (1977) *Radiation detection and measurement* (John Wiley & Sons, Inc., New York) 95-103.

[Koch, 1983] Koch, M.H.J. & Bordas, J. (1983) "X-ray diffraction and scattering of disordered systems using synchrotron radiation", Nucl. Inst. Methods 208 **(1-3)**, 461-469.

[Koch, 2003] Koch, M.H.J., Vachette, P. & Svergun D.I. (2003) "Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution", Quarterly Reviews of Biophysics **36** (2), 147-227.

[Kocsis, 2001] Kocsis, M. (2001) "The status of gas-filled detectors at a third generation synchrotron (ESRF)", Nucl. Instr. and Meth. A **471**, 103-108

[Kocsis, 2002] Kocsis, M., Private communication (2002)

[Konarev, 2003] Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., Svergun, D.I. "PRIMUS: a Windows PC-based system for small-angle scattering data analysis", J. Appl. Cryst. **36**, 1277-1282.

[Krause, 1979] Krause, MO (1979) "Atomic radiative and radiationless yields for K-shells and L-shells", J. Phys. Chem. Ref. Data **8** (2), 307-327

[Lengeler, 2001] Lengeler, B. (2001) "Coherence in X-ray physics", Naturwissenschaften **88**, 249-260.

[Liénard, 1898] Liénard, A. (1898) "Champ électric et magnétique produit par une charge électtrique concentrée en un point et animée d'un mouvement quelconque", L' Éclairage Élec. **16**, 5.

[Löw, 2002] Löw, U. & Kaufmann, Stefan H. E. (2002) *Immunologie, Poster* (Springer, Berlin)

[McPherson, 1999] McPherson, A. (1999) *Crystallisation of biological macromolecules* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor).

[Ne, 1997] Ne, F., Gabriel, A., Kocsis, M. & Zemb, T. (1997) "Smearing effects introduced by the response function of position-sensitive gas detectors in SAXS experiments", J. Appl. Crystallogr. **30**, 306-311.

[Netz, 2003] Netz, R. (2003) "Gemeinsamkeiten ziehen sich an", Physik Journal **2**, 19-20.

[Oppenheim, 1989] Oppenheim, A.V. (1991) *Signale und Systeme* (Wiley-VCH, Weinheim).

[Peisert, 1984] Peisert, A. & Sauli, F. (1984) "Drift and diffusion of electrons in gases: a compilation", CERN Report **84-08**.

[Petoukhov, 2002] Petoukhov, M.V., Eady, N.A., Brown, K.A. & Svergun, D.I. (2002) "Addition of missing loops and domains to protein models by X-ray solution scattering", Biophysical J. **83**, 3113-3125.

[Pontoni, 2002] Pontoni, D., Narayanan, T. & Rennie, A.R. (2002) "High-dynamic range SAXS data acquisition with an X-ray image intensifier", J. Appl. Crystallogr. **35** 207-211.

[Ramón y Cajal, 1999] Ramón y Cajal, S. (1999) *Advice for a young scientist* (The MIT Press, Cambridge, Massachusetts).

[Röntgen, 1895] Röntgen W.C. (1895) *Über eine neue Art von Strahlen* (Stahel'schen K. Hof- und Kunsthandlung, Würzburg).

[Sauli, 1977] Sauli, F. (1977) "Principle of operation of multiwire proportional and drift chambers", CERN Report **77-09**

[Sayle, 1995] Sayle, R. & Milner-White, E.-J. (1995) "RasMol: Biomolecular graphics for all", Trends in Biochemical Sciences **20**(9), 374-376.

[Schubert-Zsilavcz, 2001] Schubert-Zsilavcz, M. & Wurglics, M. "Better blood sugar control in diabetics. Insulin glargin—a long-acting insulin analogue", Pharm. Unserer Zeit. 2001 **30**(2), 125-130.

[Schwinger, 1949] Schwinger, J. (1949) "On the classical radiation of accelerated electrons" Phys. Rev. **75**, 1912-1925

[Shang, 2003] Shang, W. (2003) *Design of a time-resolved data acquisition system for X-ray scatterng and application to proteins in solution and semicrystalline polymers,* Master thesis, University of Siegen.

[Snigirev, 1996] Snigirev, A., Kohn, V., Snigireva, I. & Lengeler, B. (1996) "A compound refractive lens for focusing high-energy X-rays", Nature **384**, 49-51

[Stuhrmann, 1970] Stuhrmann, H.B. (1970) "Ein neues Verfahren zur Bestimmung der Oberflächenform und der inneren Struktur von gelösten globularen Proteinen aus Röntgenkleinwinkelmessungen" Zeitschrift für physikalische Chemie (Neue Folge) **72**, 177-198.

[Svergun, 1992] Svergun, D.I. (1992) "Determination of the regularization parameter in indirect-transform methods using perceptual criteria", J. Appl. Cryst. **25**, 495-503.

[Svergun, 1995] Svergun, D., Barberato, C., Koch, M.H.J. (1995) "CRYSOL– a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates", J. Appl. Cryst. **28**, 768-773.

[Svergun, 1999] Svergun, D.I. (1999) "Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing", Biophysical J. **76**, 2879-2886.

[Svergun, 2001] Svergun, D.I., Petoukhov, M.V. & Koch, M.H.J. (2001) "Determination of domain structure of proteins from X-ray solution scattering", Biophysical J. **80**, 2946-2953.

[Tardieu, 1999] Tardieu, A., Le Verge, A., Malfois, M., Bonneté, F., Finet, S., Riès-Kautt, M. & Belloni, L. (1999) "Proteins in solution: from X-ray scattering intensities to interaction potentials", J. Crystal Growth **196**, 193-203.

[Vaney, 1996] Vaney, M.C., Maignan, S., Riès-Kautt, M. & Ducruix, A. (1996) "High-resolution structure (1.33Å) of a HEW lysozyme tetragonal crystal grown in the APCF apparatus. Data and structural comparison with a crystal grown under microgravity from SpaceHab-01 mission", Acta Cryst. D**52**, 505-517.

[Verwey, 1948] Verwey, E.J.W. & Overbeek, J.TH.G. (1948) Theory of the stability of lyophobic colloids (Dover publications, Inc., Mineola, New York).

[Vivarès, 2002] Vivarès, D., Belloni, L., Tardieu, A. & Bonneté, F. (2002) "Catching the PEG-induced attractive interaction between proteins", Eur. Phys. J. E **9**, 15-25

[Walker, 2002] Walker, J.M. (2002) *The protein protocols handbook* (Humana Press, Totowa, New Jersey).

[Wilkins, 1998] Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.-C., Williams, K.L., Appel, R.D. & Hochstrasser, D.F. (1998) *Protein identification and analysis tools in the ExPASy Server* in: *2-D Proteome Analysis Protocols* (Humana Press, New Jersey)

[Willner, 2002] Willner, I. (2002) "Biomaterials for Sensors, Fuel Cells, and Circuitry", Science **298**, 2407-2408.

[Wilson, 1992] Wilson, R. & Turner, A.P.F. (1992) "Glucose oxidase: an ideal enzyme", Biosensors & Bioelectronics **7**, 165-185.

[Wilson, 1999] Wilson, A.J.C & Prince, e. (1999) *International tables for crystallography, Volume C* (Kluwer Academic Publisher; Dordrecht/Boston/London).

11 Acknowledgements

The last three and a half years were marked by some hard work and fortunately some small successes. I am indebted to the many people who supported me during the process of producing my thesis.

First of all I would like to thank my supervisor Prof. Dr. Robert L. Johnson for his interest in my work and his encouragement.

Since the research presented in this thesis has been carried out at the Hamburg outstation of EMBL I want to express my special appreciation to Michel Koch, who was my supervisor at the EMBL. He gave me all freedom to develop my interests, let me take my decisions and guided me (particularly by asking critical questions) through the work. I appreciate his loyal support through all stages of the work and I truly admire his ability to read and *carefully* correct 100-page manuscripts within one day. Also I am grateful to him for waking up my interest in the history of science and for introducing me to his views on scientific and other topics during countless hours at the beam line, in our offices and in the coffee room. Although I do not always agree I would strongly suggest writing a book about it. I would buy it.

Many thanks go also to the two other members of my EMBL thesis advisory committee, Christoph Hermes and Christian Boulin, for attending my annual presentations and for the fruitful discussions afterwards.

Horst Fischer and Fritz-Herbert Heinsius have provided me with most of the hardware and knowledge to carry out the TDC work of the instrumentation part of this thesis. I thank the MVP group at DESY and in particular Olaf Hensler for making a VME CPU available and helping to adapt it to our system. Many thanks go also to Doris Jahn and Roy Kläring for their help with various technical aspects of the project, and to Bernd Robrahn who desperately attempted to teach a physicist soldering (the result of his efforts can be admired in Figure 4.18). I also would like to thank Andre Gabriel for the interesting conversations on detectors and for the pantomime presentations “How to enjoy life and drink Pastis in the south of France”, which he once gave in my office.

For the work with the GEM detector I would like to thank Menhard Kocsis, who build the detectors, and Wim Bras, who gave us the opportunity to use the DUBBLE beam line. Many thanks also go to Marc Malfois, Igor Dolbnya and Franc Sever for their support

during the measurements. I thank Sergio Funari for giving me some hours of his valuable beam time to learn something about CCD and image plate detectors.
For the protein interaction project plenty of thanks are addressed to the biochemist in the wet lab, who had to bear the presence of a physicist at their working place. Special thanks go to Anni Linden, Katja Schirwitz and Will Stanley for making my wet lab stays as successful as possible and for preventing me from poisoning myself. In this context I also thank my fellow PhD students from Heidelberg who showed me how a pipette actually works.
I would like to express my gratitude to Maxim Petoukhov, Peter Konarev and my fellow group member Weifeng Shang for their help during the data analysis and for many discussions about the theory of SAXS.
For various aspects of my projects I was supported by the staff of HASYLAB, in particular I am grateful for their moral support on the evening when I turned the interlock key *before* closing the shutter.
Special thanks, of course, go to all the people who help to make the working atmosphere so nice at EMBL.
I would like to thank my parents for their lifelong support and because I know that they are always present when there are *real* problems.
Finally, I am indebted to my girlfriend Christine Funke who certainly gave me the most valuable support during the last three and a half years.