2 METHODS

2.1 Materials

All other chemicals for crystallisation were ordered from either Merck, Germany or Sigma-Aldrich, Germany and were at least analysis quality.

Crystallisation Reagents

Acarbose	Bayer AG, Germany
Ditheothreitol (DTT)	AppliChem, Germany
Maltose	Sigma-Aldrich, Germany
NAD^+	Sigma-Aldrich, Germany
Raffinose	Sigma-Aldrich, Germany
Trehalose	Sigma-Aldrich, Germany
tris-(2-carboxyethyl)phosphine, hydrochloride (TCEP)	Molecular Probes, The Netherlands

Heavy atom compounds

di-µ-iodobis(ethylenediamine)-diplatinum(II)nitrate (PIP)	Strem Chemicals, Newburyport, MA, USA
$HgCl_2$	Merck, Germany
$Pb(ac)_2$	Merck, Germany
p-chloromercuribenzene sulfonic acid (PCMBS)	Sigma-Aldrich, Germany
$SmCl_3$	Sigma-Aldrich, Germany
Ta_6Br_{14}	synthesised by Rupert Marx, Institut für Chemie, Freie Universität Berlin

Crystallisation equipment

24-well plates	Linbro, ICN Flow, Germany
96-well plates	Greiner, Germany
Baysilon silicon grease	Bayer AG, Germany
Cryo loops	Hampton Research, Laguna Niguel, CA, USA
Crystal Screen I and II	Hampton Research, Laguna Niguel, CA, USA
Glass coverslips (22mm x 22mm)	Mezel Glazer, Germany
PEG Ion Screen	Hampton Research, Laguna Niguel, CA, USA
Sealing tape	Mancon, Avon, OH, USA

2.2 Cloning and Purification of AglA

The *aglA* gene from *T. maritima*, encoding the α-glucosidase AglA, was cloned and the protein expressed and purified by the group of Wolfgang Liebl at the University of Göttingen. A summary is given here.

2.2.1 Identification of *aglA* (AglA) from a *T. maritima* Gene Library

A *T. maritima* gene library, prepared by the ligation of 3 to 6kb Sau3A fragments of chromosomal MSB8 DNA into the vector pUN121 (Ruile et al., 1997), was transformed into *E. coli* JM83 (Bibel et al., 1998). Replica plates were screened for transformants by an activity assay, whereby the plates were heated to 65°C and then overlayed with filter paper containing a 50mM $NaKPO_4$ buffer with 3mM 4-methylumbelliferyl-α-D-glucoside. After incubation at 65°C the plates were inspected for fluorescent colonies under UV light (366nm). The thus identified clones were sequenced and the *aglA* gene characterised.

2.2.2 Cloning of *aglA* (AglA) from *T. maritima*

A 2.4kb EcoRV-BamHI fragment of chromosomal DNA containing *aglA* was cloned into the SmaI site of expression vector pWLQ2, creating pWBE2.4 (Figure 2.1) (Goßlar, 1992). The *aglA* gene was then subcloned into pET21a to create pET21a-aglA (Raasch, 2001).

2.2.3 Purification of AglA

pET21a-aglA (AglA) containing cultures were induced at OD_{600} = 0.5 with 0.3mM IPTG and allowed to express for 5 hours at 37°C (Raasch et al., 2000). The cells were washed with 20mM Tris-HCl pH 8.0 and then resuspended in the same buffer at around 1ml/g cells. The cells were disrupted using a French press and centrifuged (40,000xg, 30min, 4°C) (Figure 2.2, lane 2). The supernatant was collected in 0.6ml aliquots and incubated at 85°C for 15 min before being cooled on ice and recentrifuged (40,000xg, 30min, 4°C). The AglA–enriched supernatant (Figure 2.2, lane 3) was then subjected to anion exchange chromatography on Source 15 Q media and eluted with a linear gradient from 0.1M to 0.4M NaCl. The AglA-containing fractions were confirmed with activity tests, pooled and concentrated in 50mM Tris-Cl pH 7.0 in MACROSEPTM microconcentrators to a

final concentration of 8 to 16mg/ml. The SDS-PAGE (Figure 2.2, lane 4) shows that AglA (apparent molecular weight of approx. 58 kDa, calculated molecular weight 55 kDa) is approximately 95% pure, with only one further contaminating band of unknown origin at 45 kD.

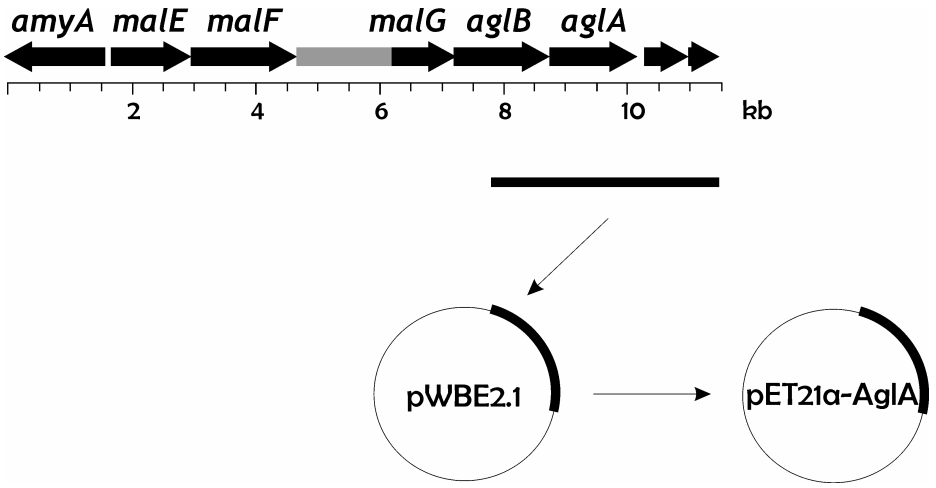

Figure 2.1 Cloning of *aglA* from *T.maritima*

Shown is a schematic figure of the gene cluster and size. Indicated by the black bar is the fragment of the genome cloned into the expression vector.

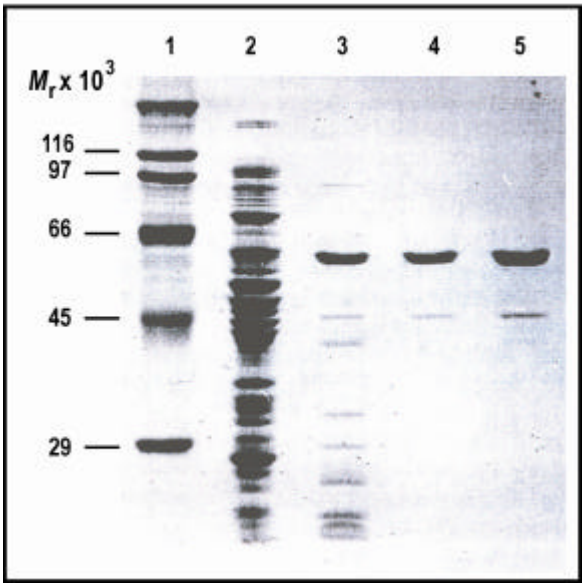

Figure 2.2 SDS-PAGE of AglA purification.

Coomassie-stained SDS-PAGE gel. **lane 1,** MW markers; **lane 2**, crude extract after french press; **lane 3**, heat denatured supernatant; lane **4**, pooled fractions after anion exchange chromatography; **lane 5**, pooled fractions after gel filtration (from (Raasch et al., 2000).

2.3 Crystallographic Investigations of AgIA

2.3.1 Theory of Protein Crystallisation

The growth of protein crystals is subject to many variables, such as the purity and homogeneity of the protein sample, the nature of the precipitant, the concentration of both the protein and the precipitant, ionic strength, temperature and pH. The underlying theory of protein crystallisation is that a supersaturated solution (i.e. one in which there is not enough solvent to maintain full hydration of the molecules present) will be thermodynamically driven towards an equilibrium state where it is partitioned between a soluble and a solid phase. Although as the proteins interact with each other they lose rotation and translational movement, thereby lowering the entropy of the system, they form new and stable bonds. This reduces the free energy of the system and provides the driving force for the ordering process (McPherson, 1982). This interaction usually results in the more kinetically favourable amorphorous precipitants, but may under certain conditions form crystals (McPherson and Weickmann, 1990). These conditions, however, cannot be suitably predicted and must instead be identified by screening approaches (Jancarik and Kim, 1991).

The basic strategy of growing protein crystals is to form a temporary supersaturated solution by altering various conditions such as salt or precipitant content. As crystal nuclei form and enlarge, the solution immediately surrounding the crystal becomes depleted of protein, thereby creating a protein concentration gradient through the solution. The outcome is illustrated by the solubility curve in Figure 2.3. The curve represents the border between supersaturated and undersaturated solution, and the curve itself constitutes the point of equilibrium between the protein being in solution and forming precipitation or crystals. The three zones described in Figure 2.3 are as follows. First, the precipitation zone, where the protein immediately precipitates to form amorphous aggregates. Second, the nucleation zone, where excess protein molecules will form crystals. The closer to the precipitation zone, the more and smaller the crystals will be. Finally, the metastable zone which is created after a supersaturated solution has nucleated and allows crystals to slowly grow larger, thereby avoiding imperfections in the lattice.

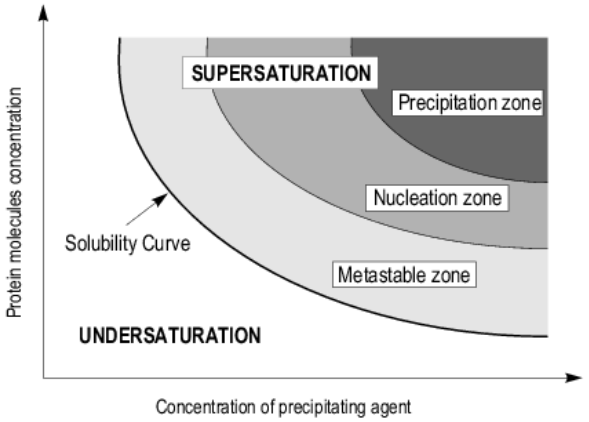

Figure 2.3 Crystallisation phase diagram

2.3.2 Vapour Diffusion Method

The most widely applied technique to grow protein crystals is the vapour diffusion method. In this, a small drop is formed by equal volumes of protein solution and reservoir solution containing the precipitant at the desired final concentration. The resulting drop is suspended over the reservoir solution as a hanging drop (Figure 2.4 A) or sitting drop (Figure 2.4 B) and the environment sealed. Due to the initial difference in precipitant concentration between the drop and the reservoir, the drop will equilibrate over time with the reservoir solution. Water will evaporate from the drop, the volume of the drop will decrease and consequently the protein concentration will increase. In the phase diagram shown in Figure 2.5, the system starts at an unsaturated state (Figure 2.5, point A, protein concentration C_{pi}) and concentrates to a supersaturated state (Figure 2.5, point B). As crystals appear and grow, the concentration will decrease until it reaches the solubility curve (Figure 2.5, point C, concentration C_{pf}).

To crystallize a protein by the vapour diffusion method, a reservoir solution containing buffer, salt and precipitant is pipetted into a well. A small drop (usually 2-5µl) of the protein solution is pipetted onto a siliconised glass coverslip and to this is added an equal volume of the reservoir solution. The cover slip is fixed over the well with silicon grease, and the reservoir and hanging drop is left to equilibrate at 18°C. During this equilibration process, which may take days or weeks, conditions may develop in which crystallisation of the protein occurs. The drop is checked regularly over the following days and weeks under the microscope.

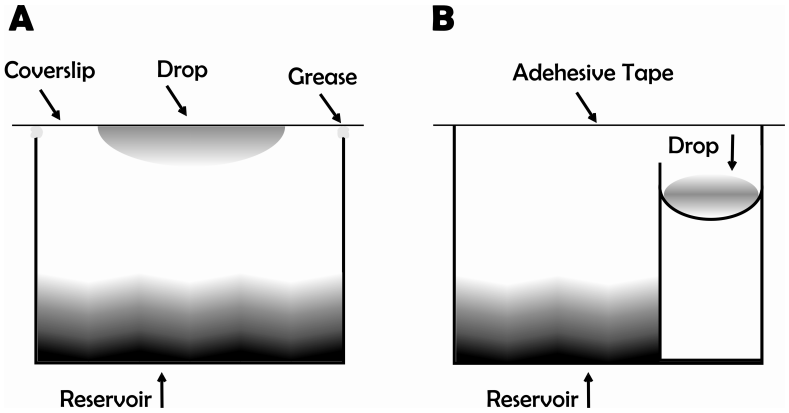

Figure 2.4 Crystallisation setups

Shown are the crystallisation setups used in this work. **A.** Hanging drop method in 24-well format. **B.** Sitting drop method in 96-well format.

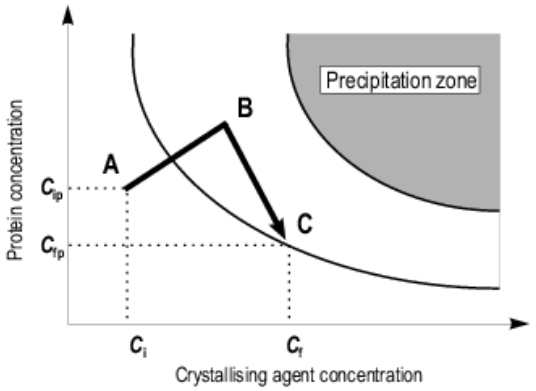

Figure 2.5 Phase diagram of a vapour diffusion experiment

2.3.3 Crystallisation Conditions

The crystallisation experiments performed in this study were carried out using three protein samples (termed AglA batch I, II and III). Initial crystal conditions were determined by screening with the commercially obtained Crystal Screens I and II and the PEG-Ion Screen from Hampton Research. Screening was performed in 24-well format trays by the hanging drop vapour diffusion method with 500µl volume reservoir solutions and 4µl drops consisting of 2µl protein solution and 2µl reservoir solution.

Fine tuning of the crystal conditions was performed in either the 24-well format mentioned above, or in 96-well Greiner plates designed for three sitting drops per reservoir. In the latter setups, the reservoir was 100µl, and each drop consisted of 0.5µl protein solution mixed with 0.5µl reservoir solution. For each batch of protein, refined

optimized conditions were determined individually. In addition, dependant on which substrate or cofactor was to be incorporated into the active site, further reagents were added to either the drop or reservoir. Details are given in Table 3-1, pg. 30.

2.3.4 Cryo-Crystallography

One of the major problems encountered while measuring crystals is the radiation damage the crystal incurs during the course of a data collection. The initial damage is caused by X-rays breaking bonds and releasing free radicals. Secondary damage occurs as these radicals diffuse through the crystal lattice and react with other parts of proteins in the crystal (Garman and Schneider, 1997). These secondary effects can be overcome by cooling the crystal during data collection to very low temperatures, thus abolishing diffusion. For this purpose, the crystal is cooled in liquid nitrogen or propane and measured at very low temperatures, which reduces the radiation damage and often enables a single crystal to survive a full data collection experiment. This technique is known as cryo-crystallography and is now employed as standard for most experiments.

The primary concern in cryo-crystallography is that the liquid surrounding the crystal should not freeze and form crystals, rather cool to a glass-like phase which will not interfere with diffraction of the protein crystal. In order for this to happen cryo-protectants must be added to the reservoir solution as otherwise water will form ice crystals which will themselves diffract. Appropriate cryo-protectants are, for example, low molecular weight PEGs, glycerol or ethylene glycol, which are added to the mother liquor.

Crystals of AglA were transferred into such a cryo-protectant buffer by the gradual stepwise addition of the cryo buffer to the drop and concurrent removal of the mother liquor. The crystal was then left to equilibrate for 10 – 15 minutes prior to mounting in a nylon loop, then flash-cooled in liquid nitrogen or a nitrogen gas stream and stored until measurement in liquid nitrogen. During the diffraction experiment, the crystal was kept at a constant temperature of 100K with the aid of a cryo system which keeps a steady flow of nitrogen flowing over the crystal. Details of the cryo buffers used are given in Table 3-1, pg. 30.

2.3.5 Heavy Atom Derivatives

In order to solve the phase problem, which is discussed in depth in Appendix 5.2, native crystals of AglA were soaked with a number of heavy atom compounds. The critical parameters of heavy atom derivatisation are (i) which compounds are to be used, (ii) soaking concentration, (iii) soaking duration and (iv) whether a "back-soaking" step (soaking against buffer without the heavy atom compound) is included after derivatisation. Alternatively, heavy atoms can be introduced by co-crystallisation or by co-translational incorporation (e.g. selenomethionine).

Heavy atom compounds used for the MIRAS structure determination of AglA are listed in Table 3-2 (pg. 31). They were dissolved in the reservoir buffer and added at a concentration of 5mM to a 10µl drop of the mother liquor containing a crystal. The crystal was incubated overnight before being transferred to the cryo buffer and frozen as described above.

2.3.6 Experimental Setup and Data Collection

The standard method for data collection is the rotation method. In this, the crystal is rotated in the X-ray beam so that all lattice points of the reciprocal lattice rotate through the Ewald sphere and are collected as reflections. This is done in small oscillation steps of 0.1° – 1° per diffraction image. A complete data set consists of a consecutive series of these images covering a total rotation of 30° to 360°, depending on the symmetry (space group) of the crystal.

The standard experimental setup is shown in Figure 2.6. The crystal is mounted in a cryo loop on a goniometer head (labelled 1) which allows rotation of the crystal around one axis. A camera (labelled 4) allows visualization of the crystal to centre it in the X-ray beam (labelled 2). The cryo stream (labelled 3) ensures the crystal is kept at 100K by a constant flow of nitrogen. In this figure, the X-ray beam comes from the left, passes through the crystal and is blocked by the beam stop (labelled 5, moved out of position in this figure) to ensure the detector (labelled 6) is not directly hit by the incident beam.

The X-ray beam is produced by either a rotating anode source (home source) or as Bremsstrahlung at a synchrotron source. For a general introduction see (Arndt, 2001). The diffracted beams (reflections) are collected with a detector, of which two commonly used

types exist. Image plate detectors collect the reflections via temporary activation of atoms, which are promoted to an excited state when they absorb an X-ray photon. These can then be "read" by scanning the plate with red light. Alternatively, charge-coupled device (CCD) detectors are used. Here, the arriving X-ray photons excite a fluorescence screen and the light is transmitted in real time via fibre optics to the charge-coupled device within the detector (Blow, 2002).

Initial analysis of crystal quality was performed by taking a first image and evaluating the quality of the diffraction. Using DENZO and XDISP of the HKL program suite (Otwinowski and Minor, 1997), this initial frame was indexed and processed and a data collection strategy was determined. Following collection of a full data set, DENZO was used to integrate the intensities from all frames. SCALEPACK, also from the HKL program suite, was then employed to scale the reflections from all frames, summing up partial reflections and refining the cell parameters and mosaicity.

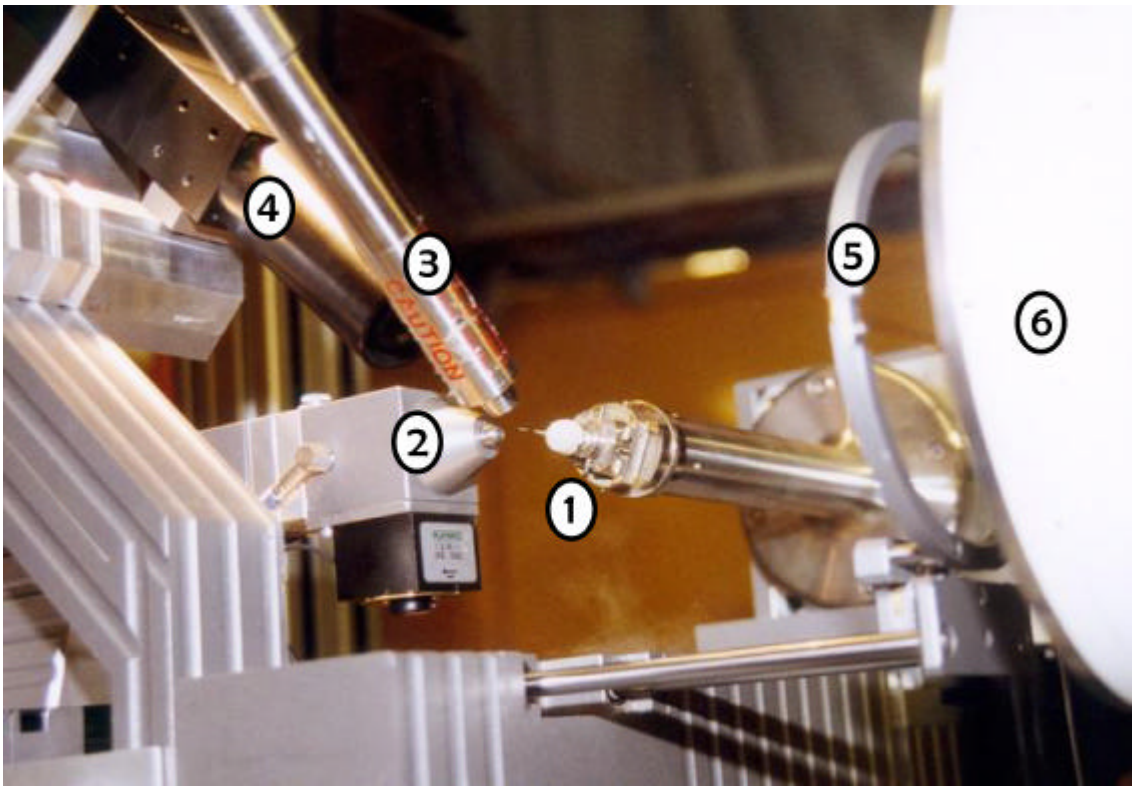

Figure 2.6 Setup of a crystallographic data collection experiment

On Beamline X-13 at DESY, Hamburg. 1. Crystal mounted in a loop on the goniometer head. 2. X-ray beam source. 3. Cryo-stream. 4. Camera to centre the crystal in the X-ray beam. 5. Beam stop. In this figure the beam stop has been moved to the right, in order to mount the crystal. 6. Detector. In this figure the detector is a MarResearch CCD detector.

2.3.7 Solving the Phase Problem

Datasets of four of the heavy atom derivatives of AglA were collected at the BAMline beamline at BESSY II, Berlin. Heavy atom sites and initial phases were determined using the program SOLVE (Terwilliger and Berendzen, 1999). Two additional data sets of $HgCl_2$ and $SmCl_3$ derivatives were collected on a home anode and the heavy atom sites were located from difference Fourier maps. Phases were calculated with SHARP (de La Fortelle and Bricogne, 1997) using the isomorphous and anomalous contributions. A more detailed introduction to the theory of protein crystallography can be found in Appendix 5.2, pg. 86.

2.3.8 Model Building and Refinement

Following solution of the phase problem, phase refinement and extension was done with the program DM (Cowtan, 1994). This resulted in readily interpretable electron density maps in which the model could be built manually using the graphics program O (Jones et al., 1991).

Refinement of the protein model was performed using the program package CNS (versions 1.0 and 1.1) (Brunger et al., 1998). Following each round of rebuilding in O, the structure was refined using CNS procedures "minimize" (restrained all-atom refinement), "bindividual" (refinement of individual B-factors) and a second "minimize". Maps were calculated with the procedure "model_map". Water molecules were automatically picked with 5 successive rounds of "water_pick" and "water_delete", and were then manually checked in O. Extra water molecules not picked automatically were checked for with the O interface xono (N.Sträter, unpublished) from the peaks of the F_o-F_c electron density maps.

The pdb, parameter and topology files for NAD^+, maltose and cysteine-sulfinic acid were obtained from the HIC-Up server (Kleywegt and Jones, 1998). The topology and parameter files for each were manually corrected and can be found in the Appendices 5.4, 5.5 and 5.6. The final structures were checked with Procheck and What Check (Hooft et al., 1996; Laskowski et al., 1993)

2.4 Small-angle X-ray scattering

Small-angle scattering (SAS) is a technique which provides structural information on particles with dimensions ranging from ten to several thousand Angstrom and is utilised for studies on such fields as metal alloys, microemulsions, liquid crystals and biological macromolecules (Vachette and Svergun, 2000). There are several fields of SAS which are defined by the source: X-ray scattering (SAXS), neutron scattering (SANS) or light scattering (LS).

SAXS is an appropriate method for acquiring structural information of proteins in solution at low resolution. As such, it is capable of supplying information on gross structural features for molecules difficult to crystallise and can also clarify structural morphology changes which may occur as artifacts of the crystallisation process. It is because of this latter ability that SAXS was used in the present work, as it provided information on the dimerisation state of AglA. A more detailed introduction to SAXS is given in Appendix 5.3, pg. 92.

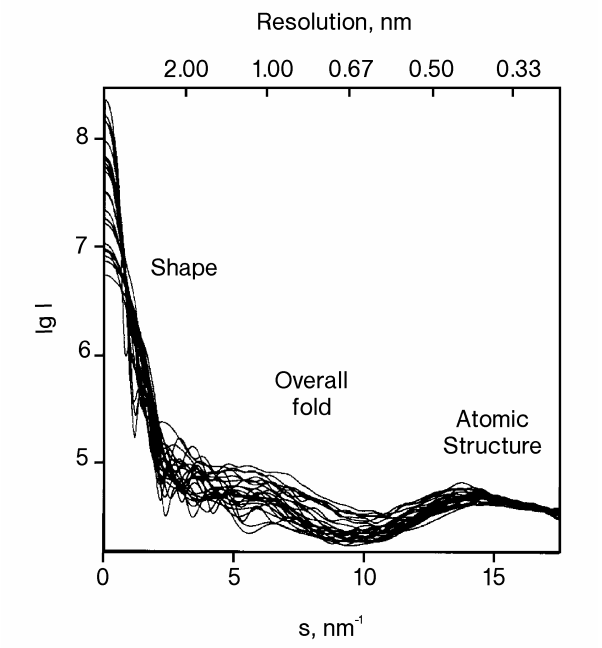

Figure 2.7 The information content of small angle scattering (SAS) data.

At low (2-3nm) resolution, the curves are very different, and the quickly decaying functions of s are primarily determined by the particle shape. At medium resolution (2-0.5nm), the differences are already less pronounced and, above 0.5nm resolution, all the curves are very similar. From (Svergun et al., 2001).

2.4.1 Experimental Details

SAXS measurements were performed on the Non Crystalline Systems (NCS) beamline X33-D1/2 at the EMBL outstation Hamburg (DESY) using multiwire proportional chambers with delay line readout. The scattering patterns were recorded at a sample-detector distance of 1.7m, equivalent to a range of momentum transfer of $0.02 < s < 0.35$ $Å^{-1}$. The protein was diluted in 0.1M Tris pH 7.0 to concentrations between 3 and 15mg/ml. Ten one-minute exposures at 15°C were performed on each protein concentration with no protein degradation observed, and a buffer blank measurement was made before and after each protein sample. The programs SAPOKO and PRIMUS (Konarev P.V., Koch, M.H.J. & Svergun, D.I., manuscript in preparation) were used to adjust for beam intensity and detector response, to average the frames and subtract the buffer background. Models of the two putative dimers and their fit to the experimental data were evaluated with the program CRYSOL (Svergun et al., 1995).

2.5 Analysis of the Crystal Structure and Preparation of Figures

Sequence alignments were performed using ClustalW (Thompson et al., 1994) and displayed with ESPript (Gouet et al., 1999), while the dendogram was prepared with Treeview (Page, 1996). Comparison of the structure to known structures was done via the DALI server (Holm and Sander, 1993). Superpositions were performed using LSQMAN from the Uppsala Software Factory (Kleywegt, 1996). Topology diagrams were prepared with TOPS (Westhead et al., 1999). Images of the co-ordination of ligands were prepared with LIGPLOT (Wallace et al., 1995). The active site tunnel surface was calculated and visualized with MAMA (Kleywegt and Jones, 1999) and VOIDOO (Kleywegt and Jones, 1994). Images of the AglA molecule were prepared with MolDraw (N.Sträter, unpublished), a graphical user interface to MolScript (Kraulis, 1991) and rendered with Raster3D (Merritt and Bacon, 1997).