5 Appendices

5.1 List of Abbreviations

Chemicals:

DTT	Dithiothreitol
NAD^+	Nicotinamide adenine dinucleotide
PCMBS	p-chloromercuribenzene sulfonic acid
PEG	Polyethyleneglycol
PIP	di-µ-iodobis(ethylenediamine)diplatinum(II)nitrate
TCEP	tris-(2-carboxyethyl)phosphine

Datasets:

Nat-2	native protein
Malmn1	native protein with maltose, NAD^+, Mn^{2+}
Ag12	C25S/C174S double mutant with maltose, NAD^+, Mn^{2+}
Ag18	native protein with maltose, Co^{2+}

Crystallographic terms:

Å	Angstrom, 0.1nm
BESSY II	Berliner Elektronenspeicherring - Gesellschaft für Synchrotronstrahlung m.b.H. II
B-factor	temperature factor
CAZY	Carbohydrate Active Enzyme Database
CSW	PDB abbreviation for cysteine-sulfinic acid

Cys-SOH	Cysteine-sulfenic acid
Cys-SO_2H	Cysteine-sulfinic acid
DESY	Deutsches Elektronen Synchrotron
EMBL	European Molecular Biology Laboratory
GH4	Glycosyl Hydrolase family 4
MAD	Multiple wavelength Anomalous Dispersion
MIR	Multiple Isomorphous Replacement
MIRAS	Multiple Isomorphous Replacement with Anomolous Scattering
PDB	Protein Data Bank
rmsd	root mean square deviation

5.2 Theory of X-ray Crystallography

5.2.1 Diffraction of Waves: Bragg´s Law

An X-ray is an electromagnetic wave with a wavelength of around 1 Å. As with all waves, it may be characterized by its wavelength, amplitude and phase (Figure 5.1, A). Diffraction of a wave occurs when the wave is scattered into directions other than that of the original beam without changing the wavelength. When an X-ray hits an electron, it sets the electron oscillating with the X-ray frequency, which may then radiate an X-ray photon of the same wavelength in a random direction as it returns to its unexcited state.

According to Bragg, the diffraction of X-rays by a crystal lattice can be described as reflection from a series of lattice planes. When an incoming wave (incident beam) hits a lattice plane under the angle θ, it will be reflected (diffracted) to produce an outgoing wave (diffracted beam) that will have the same wavelength as the incident beam, however, the phase may differ. This diffraction is shown for beam A in Figure 5.1, B. The same diffraction will occur for all planes in the series, however the extra distance traveled, described by $2d\sin\theta$, will result in a phase shift of the diffracted beam. Positive interference will only occur if the extra distance the second beam travels (d) is equal to an integer number of wavelengths ($n\lambda$), as then the scattered beam is in phase with the first. This is described by Bragg´s Law:

$$n\lambda = 2d\sin\theta$$

A set of parallel planes in a lattice is described by the indices h, k, l. These represent the fractional distance along each of the unit cell axes a, b, c which separate each plane in the series (Figure 5.1, C). Therefore each X-ray beam reflected from a set of planes can also be described by a set of indices h,k,l.

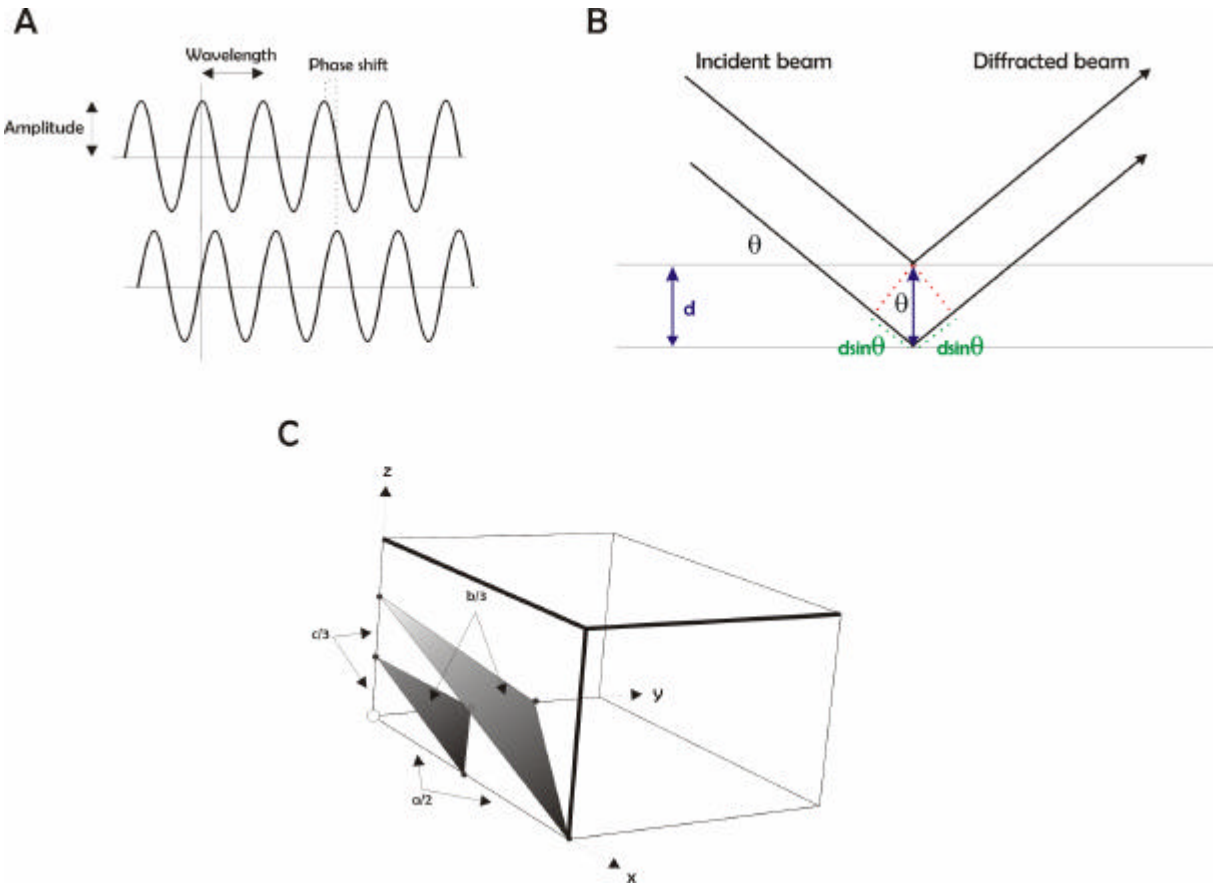

Figure 5.1 Diffraction of waves from planes

A. Wavelength and amplitude of a sinusoidal wave. Also shown is a second sinusoidal wave of the same wavelength and amplitude, but with a phase shift compared to the top wave. **B.** Bragg´s Law. Shown are two planes of a crystal lattice with a distance of d. The difference between the distance traveled by the upper beam and the lower beam is shown as a green dotted line. **C.** Three-dimensional view of a unit cell, showing an example of the planes characterised by the h, k, l indices 2, 3, 3.

5.2.2 Reciprocal Space and the Ewald Sphere

In 1921, Ewald introduced a geometric construction which illustrates the relationship between the orientation of the crystal in the beam and the resulting X-ray reflections (Figure 5.2). A sphere is drawn around the crystal in position X. The X-ray beam enters this sphere in point I, and hits the crystal under the angle ?. The reciprocal lattice is a lattice with the same symmetry elements as the real crystal lattice, but with unit cell dimensions a^*, b^*, c^*, which have the following relationship to the unit cell dimensions a, b, c of the real lattice:

$$a^* = \frac{1}{a}, \quad b^* = \frac{1}{b}, \quad c^* = \frac{1}{c}$$

At the point where the X-ray beam leaves the sphere, the origin (O) of a reciprocal lattice is placed. Based on this construction, for a given angle ?, only those lattice points of the reciprocal lattice fulfill Bragg´s law which fall onto the circumference of the sphere (here point P). These result in a reflection. For a complete diffraction data set the crystal has to be rotated in such a way that each lattice point of the reciprocal lattice passes through the Ewalds sphere at least once. For more details, see (Drenth, 1994; McRee, 1999).

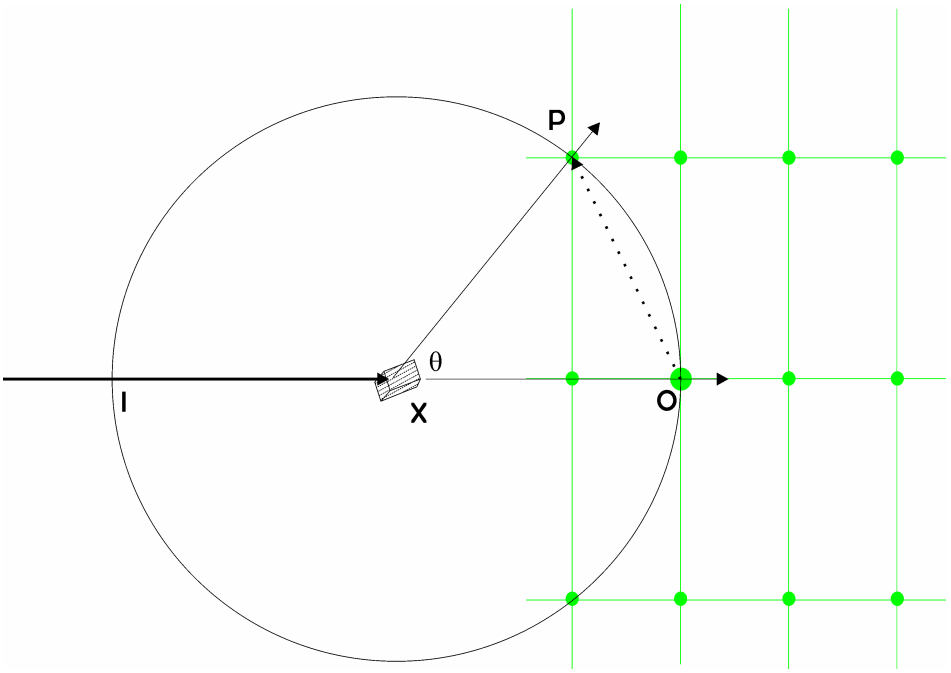

Figure 5.2 Ewald sphere and the reciprocal lattice

5.2.3 The Phase Problem

The diffracted beam is scattered by the electrons in the protein. Therefore the diffraction data in principle contain the information on the location of the electrons in the unit cell. The electron density (?) at a certain position in the unit cell (x, y, z) is a function of the amplitude (|F|) and the phase (α) of all diffracted beams:

$$r(xyz) = \frac{1}{V}\sum_{hkl} |F(hkl)| \exp\left[-2pi(hx + ky + lz) + ia(hkl)\right]$$

The amplitudes or structure factors |F| are proportional to the intensities of the reflections and can thus be determined experimentally. However, the phase information is lost during the data collection. This is called the phase problem of crystallography.

5.2.4 Solving the Phase Problem by Isomorphous Replacement

The method of multiple isomorphous replacement (MIR) was first used by Max Perutz in 1953. It relies on the observation that a heavy atom, having a much higher concentration of electrons in a small volume, will scatter X-rays strongly. If a heavy atom can be introduced into a protein crystal via soaking or co-crystallisation, without otherwise changing the protein or the crystal lattice (i.e. the crystal remains isomorphous), the differences in the scattering intensities from this crystal to the native crystal can be exploited to indirectly determine the missing phase angles.

The structure factors from the native crystal ($|F_P|$) and those of the heavy atom-derivatised crystal ($|F_{PH}|$) are scaled together to remove any variation caused by crystal quality. $|F_P|$ is then subtracted from $|F_{PH}|$, which gives the scattering intensities for only the heavy atoms ($|F_H|$). This information is then used for the Patterson synthesis, a fourier synthesis in which the phases are assumed to be zero, the amplitudes are squared and a center of symmetry is added (Drenth, 1994). This will produce a map which gives the vectors between the heavy atoms in real space, however not their actual position. Using these vectors it is possible to determine the positions of the heavy atoms in the unit cell. With these sites, and the structure factors, the Fourier transformation can now be used to determine the unknown phases for these heavy atoms. From the heavy atom phases, the phases of the protein can be derived, and the initial electron density map of the protein can be calculated.

At this step there are still two possible solutions for every phase angle, only one of which is correct. This is the major limition to structure determination via Single Isomorphous Replacment (SIR). The ambiguity can only be resolved by adding further information. In the Multiple Isomorphous Replacement (MIR) method, this is provided by further derivatives. Another method is to additionally use the anomalous signal produced by heavy atoms incorporated into the crystal (Single or Multiple Isomorphous Replacement with Anomolous Signal, SIRAS / MIRAS). Alternatively, data from a single crystal with one or more heavy atoms incorporated can be collected at several wavelengths (Multiple Wavelength Anomolous Dispersion, MAD). This has the advantage that non-isomorphism between different crystals is avoided.

5.2.5 Crystallographic Refinement

The nature of the data used to build the protein model, from data collection to phase solution to map calculation, is inherently full of error. To correct the model for this error, a crystallographic refinement is carried out.

In this, small shifts in the coordinates (bonds and angles) are introduced and the resulting energy values of the molecule are calculated. The aim is to identify the conformation with the lowest energy and the best fit to the experimental data (McRee, 1999). This is equivalent to finding the minimum of a target function E with:

$$E = E_{chem} + w_{xray} E_{xray}$$

where E_{chem} is the sum of the energy for all covalent and non-covalent interactions, including bond lengths and angles, torsion angles, non-bonded contacts, planar groups and chiral volumes. E_{xray} describes the differences between the observed structure factors, F_o, and those which can be calculated from the model, F_c:

$$E_{xray} = \sum_{hkl} \left[\left| F_o(h) \right| - k \left| F_c(h) \right| \right]^2$$

where h is comprised of all the h,k,l indices of all reflections. w_{xray} is a weighting term which allows fine tuning of the influence of the X-ray term on the overall target function. A critical aspect of the refinement procedure is that the number of refined parameters (for each protein atom the position x, y, z and the B-factor) may not exceed the number of observations, that is the number of reflections collected from a crystal. The number of parameters can be reduced indirectly by restraining the stereochemistry to values derived from small molecule crystallography (Engh and Huber, 1991). The outcome of the refinement procedure therefore is an atomic model of the protein with optimised stereochemistry and a minimised overall energy which best fits the X-ray data collected.

The progress of the refinement is judged by the crystallographic R-factor, R_{cryst}, which describes the deviation between the measured structure factors F_o and calculated F_c.

$$R_{cryst} = \frac{\sum_{hkl} \left| F_o(h) \right| - k \left| F_c(h) \right|}{\sum_{hkl} \left| F_o(h) \right|}$$

As this factor is similar to the X-ray term, it may be artificially lowered when too many parameters are refined at the same time (over-fitting). A reference factor, termed R_{free}, was introduced by Brünger, where a subset of the reflections (usually 5-10% of the data collected) are excluded from refinement (test set) (Brünger, 1993). The R_{free} is then calculated as for R_{cryst}, but only against this test set.

5.3 Theory of SAXS

5.3.1 SAXS in an Ideal Solution

SAXS of biological macromolecules in solution follows the same rules that govern the scattering of X-rays by electrons in solid matter. As the distances involved in these proteins are much greater than interatomic distances, the scattering particles are described with a continuous electron density. In the following discussion the electron density at a single point, r, is represented by p(r) and the characteristic vector of the incident beam as s_0.

An ideal solution is one in which the particles within do not interact with each other and whose thermodynamis quantities follow a linear dependence on concentration. This is approximated in SAXS by the use of extremely dilute solutions and the measurement of a range of concentrations. The monodispersity of a solution describes the identity of the molecules within and holds true for particles with a defined shape such as native proteins. In several cases, such as denatured proteins, the molecules may be monodisperse in size, but show a large degree of polydispersity in shape. It is also possible to have polydispersity in size but monodispersity in shape with particles such as micelles, and polydispersity for both can be observed for aggregations. In an ideal monodisperse solution all particles scatter independently and the measured intensity is merely the sum of the intensities from each individual molecule:

$$I(s) = N i_1(s)$$

$I(s)$ = experimental scattering intensity
N = number of particles in sample
$i_1(s)$ = scattered intensity of the particle

It is possible to directly determine the size of the particle by knowing the intensity at the zero angle ($s = 0$) with the formula:

$$I(0)/c = \frac{N_A M}{\mu^2} (1-\rho_0\psi)^2$$

$I(0)$ = experimental scattering intensity at angle 0
c = concentration
N_A = Avagadro´s constant
M = molecular weight
μ = M/m, ratio of M to the number of electrons
ρ_0 = electron density distribution of particle at angle 0
ψ = electronic partial specific volume of particle

Therefore the intensity at zero angle, scaled by the concentration, is proportional to the molecular weight and if ρ_0, ψ and the intensity of the incident beam are known, the intensity at the origin will give the molecular weight.

5.3.2 Radius of Gyration

The movement of a particle within a solution follows Brownian motion. An ideal suspension of many particles would therefore assume all particles to be at all positions and in all orientations. This isotropic intensity distribution is proportional to the scattering for a single particle averaged over all orientations. Close to the origin (angle = 0), the scattering pattern of any single particle can be approximated by a Gaussian, the width of which is proportional to the square of the radius of gyration of the particle: the Guinier relation (Guinier and Fournet, 1955).

$$R_g^2 = \frac{\int_{V_r} \Delta\rho(\mathbf{r}) r^2 \, dV_r}{\int_{V_r} \Delta\rho(\mathbf{r}) \, r \, dV_r}$$

Here, the origin of the vector is taken to be the centre of mass of the particle, and the radius of gyration (R_g) describes its sphericity. The relation of the radius of gyration to the scattering intensity close to the origin is described by:

$$I(s) \cong I(0) \exp\left(-\frac{4\pi^2}{3} R_g^2 s^2\right)$$

By graphing the natural log of I(s) against s^2, a linearised representation is created, as Guinear plot (Figure 5.3). From this plot, linear regression yields the radius of gyration from the slope, together with the intensity at the origin. However, a Guinear approximation does rely on an ideal monodisperse solution.

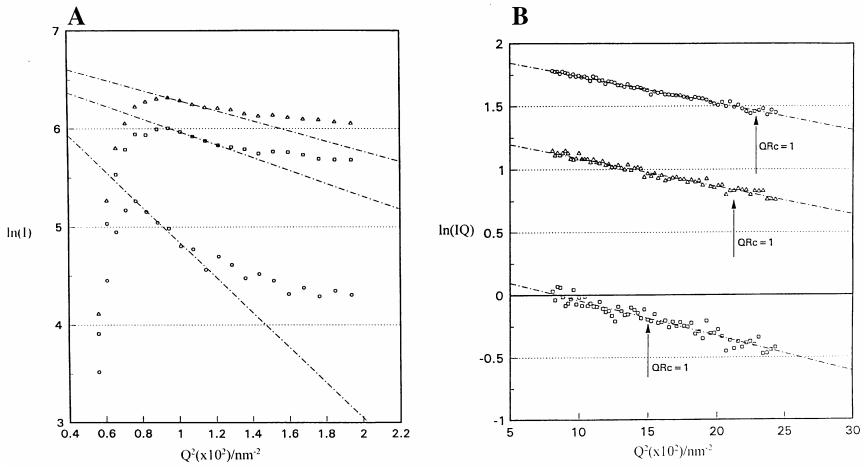

Figure 5.3 Plots of the protein K-gliadin

A. Guinier (R_g) plots for K-gliadin in 70% (v/v) aq. ethanol at concentrations of 9.3 (a), 4.7 (*), 2.3 (E) and 1.3 (O) mg/ml. Limits for the Guinier approximation to be valid (QRg 61) are shown. **B.** Cross-section (R_c) plots at concentrations of 9.3 (a), 4.7 (O), 2.3 (E) and 1.3 (*) mg/ml. Limits for the cross-section approximation to be valid (QRg 61) are shown. Taken from (Thomson et al., 1999).

5.3.3 SAXS Data Processing

The scattering intensity from a single particle in an ideal solution cannot be measured directly, rather a set of intensities ($I_{exp}(s_i)$ where I = 1...N) in a restricted angular range $s_{min} < s < s_{max}$ is measured. This set will contain statistical errors and inaccuracies produced by the beam divergence, detector resolution etc, and so the task of data processing is restoring the $I(s)$ from $I_{exp}(s_i)$. For an ideal solution, $I(s)$ is related to the pair distribution function of the particle p(**r**) by the Fourier transform

$$I(s) = 4\pi \int_0^{D_{max}} \rho(\mathbf{r}) \frac{\sin 2\pi sr}{2\pi sr} dr$$

D_{max} = maximum particle diameter

The function ?(**r**) contains the same information as $I(s)$ and the data processing can be done indirectly by restoring the ?(**r**). By providing the ?(**r**), the function $I(s)$ is automatically extrapolated beyond the measured range of scattering vectors back to the origin. By using the entire scattering curve, the system is much less sensitive to small amounts of interactions or aggregations than the Guinier relation, and, like the latter, the effects of these interactions are restricted to the innermost part of the scattering pattern.

The program GNOM performs this indirect transform, identifying the multiplier which compensates between the quality of fit to the data and the smoothness of the p(**r**) function. The program visually compares solutions for different multiplier values automatically, finding the optimal solution or indicating that there may be problems in the assumptions (such as particle size, D_{max}) used.

5.3.4 Shape Determination

The analysis of a protein with known structure (ie. from crystallography or NMR results) is performed by the program CRYSOL (Svergun et al., 1995), which, as well as predicting the scattering curve produced by the protein, also takes into account the surrounding hydration shell and its effect on scattering.

The *ab initio* shape determination may take two approaches: the first is applied by the program SASHA and employs spherical harmonics to describe the particle shape, thereby making it possible to directly restore the shape from the scattering data. (Kozin et al., 1997). The second approach utilised by the programs DAMMIN and GASBOR, is based on a trial and error analysis of scattering patterns produced by models filled with spheres (dummy atoms) or, in the latter case, dummy residues. Examples of this approach are shown in Figure 5.4.

5.3.5 Instrumentation

The X-ray beam originating from a synchrotron source passes through a monochromator which produces a monochromatic beam. The vertical beam displacement reduces background radiation from the synchrotron, and reflection at a glancing angle from the toroidal mirror eliminates harmonics and focuses in the two directions normal to the beam. The pairs of slits eliminate parasitic scattering around the monochromatic beam, essential to reach the small angles required by SAXS. The camera contains another three pairs of slits, also to remove parasitic scattering, and defines the minimum observable scattering angle. The sample is fixed in a temperature controlled quartz capillary, but the sample-to-detector distance can be adjusted by moving the detector within the range of 0.5 and 10m.

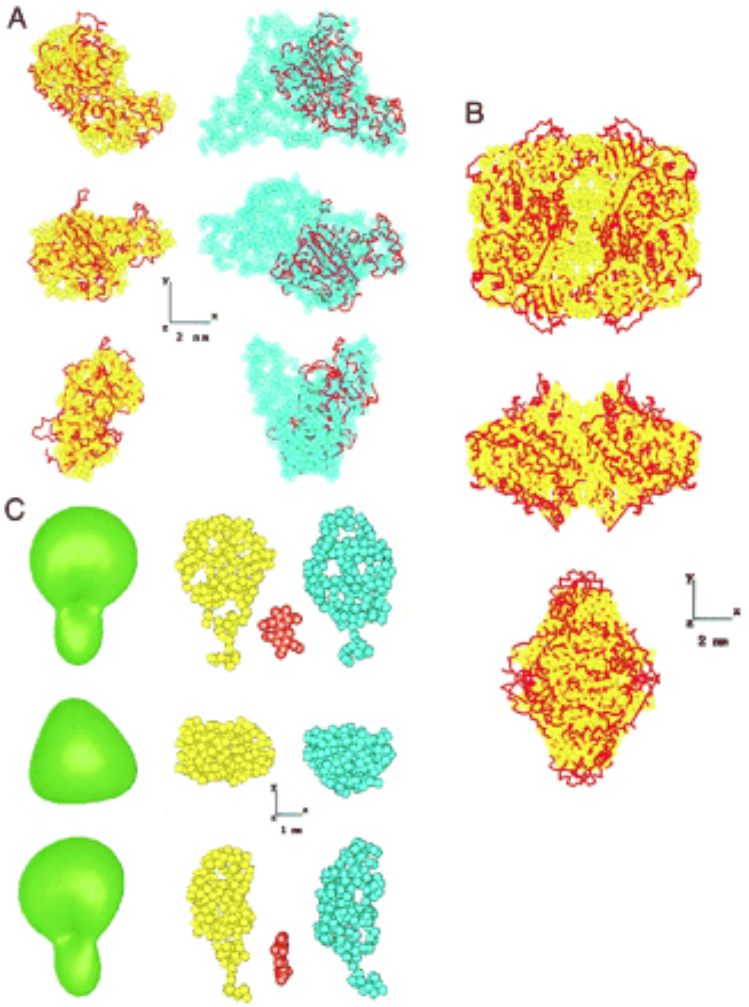

Figure 5.4 *Ab initio* Shape Determination with SAXS

A. Comparison of the atomic model of monomeric hexokinase (*red*) and the dummy residue (DR) models (*yellow*) of the monomeric (left) and dimeric (right) hexokinase. **B.** Atomic model of tetrameric yeast pyruvate decarboxylase superimposed with the DR model. **C.** Ab initio models of chitin binding protein: left, low resolution envelope produced by SASHA, middle and right, two most different DR models. The α-chitin substrate is shown in *red*. From (Svergun et al., 2001).

5.4 CNS Files for Maltose

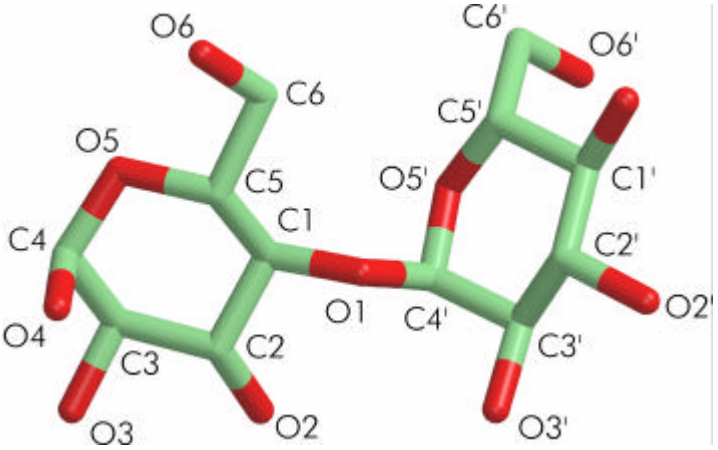

Figure 5.5 Schematic diagram of maltose

Apostrophes in atom labels follow standard PDB nomenclature.

5.4.1 CNS Parameter File for Maltose

```
BOND CM1   CM2     1000.0   1.520
BOND CM1   OM7     1000.0   1.412
BOND CM1   OM11    1000.0   1.401
BOND CM2   CM3     1000.0   1.516
BOND CM2   OM8     1000.0   1.421
BOND CM3   CM4     1000.0   1.529
BOND CM3   OM9     1000.0   1.431
BOND CM4   CM5     1000.0   1.534
BOND CM4   OM10    1000.0   1.443
BOND CM5   CM6     1000.0   1.537
BOND CM5   OM11    1000.0   1.437
BOND CM6   OM12    1000.0   1.423
BOND OM7   CM16    1000.0   1.438
BOND CM13  CM14    1000.0   1.507
BOND CM13  OM22    1000.0   1.416
BOND CM14  CM15    1000.0   1.536
BOND CM14  OM20    1000.0   1.404
BOND CM15  CM16    1000.0   1.512
BOND CM15  OM21    1000.0   1.433
BOND CM16  CM17    1000.0   1.526
BOND CM17  CM18    1000.0   1.539
BOND CM17  OM22    1000.0   1.458
BOND CM18  OM23    1000.0   1.440
BOND CM13  OM19    1000.0   1.412

ANGLe CM2   CM1   OM7      500.0    109.86
ANGLe CM2   CM1   OM11     500.0    113.67
ANGLe OM7   CM1   OM11     500.0    113.33
ANGLe CM1   CM2   CM3      500.0    109.99
ANGLe CM1   CM2   OM8      500.0    104.37
ANGLe CM3   CM2   OM8      500.0    111.18
ANGLe CM2   CM3   CM4      500.0    106.56
ANGLe CM2   CM3   OM9      500.0    109.50
ANGLe CM4   CM3   OM9      500.0    114.13
ANGLe CM3   CM4   CM5      500.0    110.45
ANGLe CM3   CM4   OM10     500.0    108.75
ANGLe CM5   CM4   OM10     500.0    107.04
ANGLe CM4   CM5   CM6      500.0    114.21
```

```
ANGLe CM4  CM5  OM11    500.0   114.12
ANGLe CM6  CM5  OM11    500.0   105.27
ANGLe CM5  CM6  OM12    500.0   110.12
ANGLe CM1  OM7  CM16    500.0   117.47
ANGLe CM1  OM11 CM5     500.0   118.76
ANGLe CM14 CM13 OM22    500.0   113.07
ANGLe CM13 CM14 CM15    500.0   115.82
ANGLe CM13 CM14 OM20    500.0   108.95
ANGLe CM15 CM14 OM20    500.0   100.76
ANGLe CM14 CM15 CM16    500.0   105.81
ANGLe CM14 CM15 OM21    500.0   110.66
ANGLe CM16 CM15 OM21    500.0   115.61
ANGLe OM7  CM16 CM15    500.0   107.00
ANGLe OM7  CM16 CM17    500.0   101.43
ANGLe CM15 CM16 CM17    500.0   108.96
ANGLe CM16 CM17 CM18    500.0   112.12
ANGLe CM16 CM17 OM22    500.0   116.29
ANGLe CM18 CM17 OM22    500.0   102.64
ANGLe CM17 CM18 OM23    500.0    97.97
ANGLe CM13 OM22 CM17    500.0   111.15
ANGLe CM14 CM13 OM19    500.0   104.00
ANGLe OM22 CM13 OM19    500.0   111.00

DIHEdral OM11 CM1  CM2  CM3     750.0 0    60.00
DIHEdral OM11 CM1  CM2  OM8     750.0 0   180.00
DIHEdral OM7  CM1  OM11 CM5     750.0 0    90.00
DIHEdral CM1  CM2  CM3  CM4     750.0 0   -60.00
DIHEdral CM1  CM2  CM3  OM9     750.0 0   180.00
DIHEdral OM8  CM2  CM3  CM4     750.0 0   180.00
DIHEdral OM8  CM2  CM3  OM9     750.0 0    60.00
DIHEdral CM2  CM3  CM4  CM5     750.0 0    60.00
DIHEdral CM2  CM3  CM4  OM10    750.0 0   180.00
DIHEdral OM9  CM3  CM4  CM5     750.0 0   180.00
DIHEdral OM9  CM3  CM4  OM10    750.0 0   -60.00
DIHEdral OM22 CM13 CM14 CM15    750.0 0    60.00
DIHEdral CM13 CM14 CM15 CM16    750.0 0   -60.00
DIHEdral CM13 CM14 CM15 OM21    750.0 0   180.00
DIHEdral OM20 CM14 CM15 CM16    750.0 0   180.00
DIHEdral OM20 CM14 CM15 OM21    750.0 0    60.00
DIHEdral CM14 CM15 CM16 CM17    750.0 0    60.00
DIHEdral OM21 CM15 CM16 CM17    750.0 0   180.00
DIHEdral OM7  CM16 CM17 OM22    750.0 0   180.00
DIHEdral CM15 CM16 CM17 CM18    750.0 0   180.00
DIHEdral CM15 CM16 CM17 OM22    750.0 0   -60.00
DIHEdral CM16 CM17 CM18 OM23    750.0 0   180.00
DIHEdral OM22 CM17 CM18 OM23    750.0 0    60.00
DIHEdral CM16 CM17 OM22 CM13    750.0 0    60.00
DIHEdral CM18 CM17 OM22 CM13    750.0 0   180.00
DIHEdral OM19 CM13 OM22 CM15    750.0 0   180.00
DIHEdral OM19 CM13 CM14 OM20    750.0 0   180.00

IMPRoper CM1  CM2  OM7  OM11    750.0 0  -35.000
IMPRoper CM2  CM1  CM3  OM8     750.0 0  -35.000
IMPRoper CM3  CM2  CM4  OM9     750.0 0   35.000
IMPRoper CM4  CM3  CM5  OM10    750.0 0  -35.000
IMPRoper CM5  CM4  CM6  OM11    750.0 0  -35.000
IMPRoper CM14 CM13 CM15 OM20    750.0 0  -35.000
IMPRoper CM15 CM14 CM16 OM21    750.0 0   35.000
IMPRoper CM16 OM7  CM15 CM17    750.0 0  -35.000
IMPRoper CM17 CM16 CM18 OM22    750.0 0  -35.000
IMPRoper CM13 CM14 OM19 OM22    750.0 0  -35.000

NONBonded CM1  0.1200  3.7418    0.1000  3.3854
NONBonded CM2  0.1200  3.7418    0.1000  3.3854
NONBonded CM3  0.1200  3.7418    0.1000  3.3854
NONBonded CM4  0.1200  3.7418    0.1000  3.3854
NONBonded CM5  0.1200  3.7418    0.1000  3.3854
NONBonded CM6  0.1200  3.7418    0.1000  3.3854
```

```
NONBonded OM7  0.1591  2.8509    0.1591  2.8509
NONBonded OM8  0.1591  2.8509    0.1591  2.8509
NONBonded OM9  0.1591  2.8509    0.1591  2.8509
NONBonded OM10 0.1591  2.8509    0.1591  2.8509
NONBonded OM11 0.1591  2.8509    0.1591  2.8509
NONBonded OM12 0.1591  2.8509    0.1591  2.8509
NONBonded CM13 0.1200  3.7418    0.1000  3.3854
NONBonded CM14 0.1200  3.7418    0.1000  3.3854
NONBonded CM15 0.1200  3.7418    0.1000  3.3854
NONBonded CM16 0.1200  3.7418    0.1000  3.3854
NONBonded CM17 0.1200  3.7418    0.1000  3.3854
NONBonded CM18 0.1200  3.7418    0.1000  3.3854
NONBonded OM19 0.1591  2.8509    0.1591  2.8509
NONBonded OM20 0.1591  2.8509    0.1591  2.8509
NONBonded OM21 0.1591  2.8509    0.1591  2.8509
NONBonded OM22 0.1591  2.8509    0.1591  2.8509
NONBonded OM23 0.1591  2.8509    0.1591  2.8509
```

5.4.2 CNS Topology File for Maltose

```
MASS CM1    13.01900
MASS CM2    13.01900
MASS CM3    13.01900
MASS CM4    13.01900
MASS CM5    13.01900
MASS CM6    14.02700
MASS OM7    15.99900
MASS OM8    17.00700
MASS OM9    17.00700
MASS OM10   17.00700
MASS OM11   15.99900
MASS OM12   17.00700
MASS CM13   14.02700
MASS CM14   13.01900
MASS CM15   13.01900
MASS CM16   13.01900
MASS CM17   13.01900
MASS CM18   14.02700
MASS OM19   17.00700
MASS OM20   17.00700
MASS OM21   15.99900
MASS OM22   17.00700

ATOM  C1   TYPE CM1   CHARge  0.0  END
ATOM  C2   TYPE CM2   CHARge  0.0  END
ATOM  C3   TYPE CM3   CHARge  0.0  END
ATOM  C4   TYPE CM4   CHARge  0.0  END
ATOM  C5   TYPE CM5   CHARge  0.0  END
ATOM  C6   TYPE CM6   CHARge  0.0  END
ATOM  O1   TYPE OM7   CHARge  0.0  END
ATOM  O2   TYPE OM8   CHARge  0.0  END
ATOM  O3   TYPE OM9   CHARge  0.0  END
ATOM  O4   TYPE OM10  CHARge  0.0  END
ATOM  O5   TYPE OM11  CHARge  0.0  END
ATOM  O6   TYPE OM12  CHARge  0.0  END
ATOM  C1'  TYPE CM13  CHARge  0.0  END
ATOM  C2'  TYPE CM14  CHARge  0.0  END
ATOM  C3'  TYPE CM15  CHARge  0.0  END
ATOM  C4'  TYPE CM16  CHARge  0.0  END
ATOM  C5'  TYPE CM17  CHARge  0.0  END
ATOM  C6'  TYPE CM18  CHARge  0.0  END
ATOM  O1'  TYPE OM19  CHARge  0.0  END
ATOM  O2'  TYPE OM20  CHARge  0.0  END
ATOM  O3'  TYPE OM21  CHARge  0.0  END
ATOM  O5'  TYPE OM22  CHARge  0.0  END
```

```
ATOM  O6'  TYPE OM23  CHARge  0.0  END

BOND  C1   C2       BOND  C1   O1       BOND  C1   O5       BOND  C2   C3
BOND  C2   O2       BOND  C3   C4       BOND  C3   O3       BOND  C4   C5
BOND  C4   O4       BOND  C5   C6       BOND  C5   O5       BOND  C6   O6
BOND  O1   C4'      BOND  C1'  C2'      BOND  C1'  O5'      BOND  C2'  C3'
BOND  C2'  O2'      BOND  C3'  C4'      BOND  C3'  O3'      BOND  C4'  C5'
BOND  C5'  C6'      BOND  C5'  O5'      BOND  C6'  O6'      BOND  C1'  O1'

DIHEdral  O5   C1   C2   O2
DIHEdral  C1   C2   C3   O3
DIHEdral  O2   C2   C3   C4
DIHEdral  C2   C3   C4   O4
DIHEdral  O3   C3   C4   C5
DIHEdral  C1'  C2'  C3'  O3'
DIHEdral  O2'  C2'  C3'  C4'
DIHEdral  O3'  C3'  C4'  C5'
DIHEdral  O1   C4'  C5'  O5'
DIHEdral  C3'  C4'  C5'  C6'
DIHEdral  C4'  C5'  C6'  O6'
DIHEdral  C6'  C5'  O5'  C1'

IMPRoper  C1   C2   O1   O5
IMPRoper  C2   C1   C3   O2
IMPRoper  C3   C2   C4   O3
IMPRoper  C4   C3   C5   O4
IMPRoper  C5   C4   C6   O5
IMPRoper  C1'  C2'  O1'  O5'
IMPRoper  C2'  C1'  C3'  O2'
IMPRoper  C3'  C2'  C4'  O3'
IMPRoper  C4'  O1   C3'  C5'
IMPRoper  C5'  C4'  C6'  O5'

ACCEptor  O1   C1
ACCEptor  O2   C2
ACCEptor  O3   C3
ACCEptor  O4   C4
ACCEptor  O5   C1
ACCEptor  O6   C6
ACCEptor  O2'  C2'
ACCEptor  O3'  C3'
ACCEptor  O5'  C1'
ACCEptor  O6'  C6'
ACCEptor  O1'  C1'
```

5.5 CNS files for NAD^+

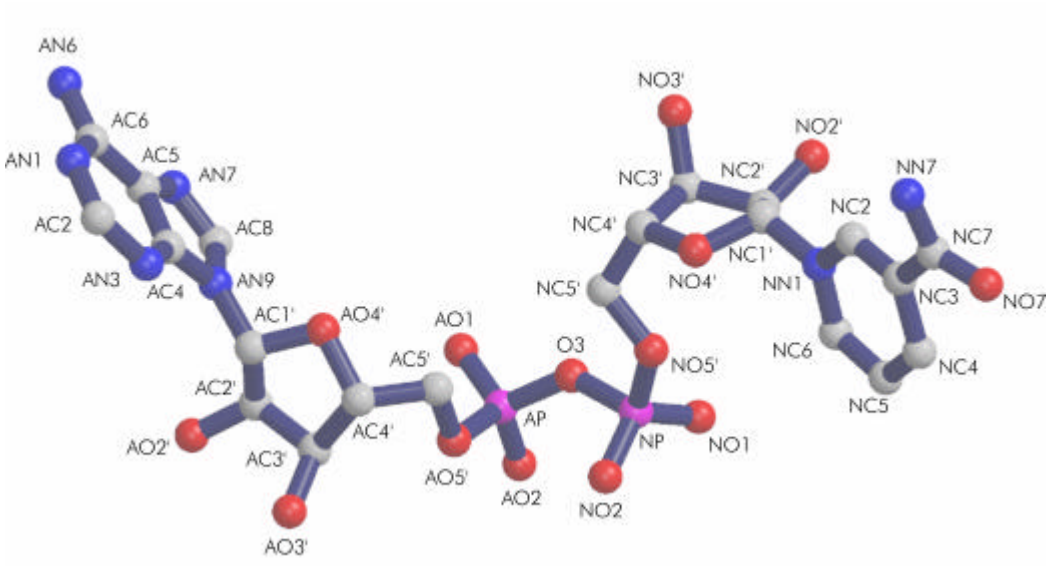

Figure 5.6 Nomenclature of NAD^+.

Atoms are labelled according to standard nomenclature. The apostrophe on the sugar ring atom labels are replaced by an asterix in CNS and PDB files, according to CNS and PDB convention.

5.5.1 CNS Parameter File for NAD^+

```
BOND PQ1  OQ2    1000.0  1.482
BOND PQ1  OQ3    1000.0  1.495
BOND PQ1  OQ4    1000.0  1.590
BOND PQ1  OQ23   1000.0  1.594
BOND OQ4  CQ5    1000.0  1.464
BOND CQ5  CQ6    1000.0  1.513
BOND CQ6  OQ7    1000.0  1.482
BOND CQ6  CQ8    1000.0  1.571
BOND OQ7  CQ12   1000.0  1.412
BOND CQ8  OQ9    1000.0  1.458
BOND CQ8  CQ10   1000.0  1.521
BOND CQ10 OQ11   1000.0  1.390
BOND CQ10 CQ12   1000.0  1.540
BOND CQ12 NQ13   1000.0  1.449
BOND NQ13 CQ14   1000.0  1.386
BOND NQ13 CQ22   1000.0  1.402
BOND CQ14 NQ15   1000.0  1.322
BOND NQ15 CQ16   1000.0  1.343
BOND CQ16 CQ17   1000.0  1.459
BOND CQ16 CQ22   1000.0  1.405
BOND CQ17 NQ18   1000.0  1.318
BOND CQ17 NQ19   1000.0  1.369
BOND NQ19 CQ20   1000.0  1.331
BOND CQ20 NQ21   1000.0  1.386
BOND NQ21 CQ22   1000.0  1.382
BOND OQ23 PQ24   1000.0  1.641
BOND PQ24 OQ25   1000.0  1.476
BOND PQ24 OQ26   1000.0  1.487
BOND PQ24 OQ27   1000.0  1.622
BOND OQ27 CQ28   1000.0  1.461
```

```
BOND CQ28 CQ29   1000.0  1.479
BOND CQ29 OQ30   1000.0  1.486
BOND CQ29 CQ31   1000.0  1.574
BOND OQ30 CQ35   1000.0  1.437
BOND CQ31 OQ32   1000.0  1.437
BOND CQ31 CQ33   1000.0  1.544
BOND CQ33 OQ34   1000.0  1.401
BOND CQ33 CQ35   1000.0  1.573
BOND CQ35 NQ36   1000.0  1.493
BOND NQ36 CQ37   1000.0  1.424
BOND NQ36 CQ44   1000.0  1.409
BOND CQ37 CQ38   1000.0  1.322
BOND CQ38 CQ39   1000.0  1.509
BOND CQ38 CQ42   1000.0  1.510
BOND CQ39 OQ40   1000.0  1.249
BOND CQ39 NQ41   1000.0  1.338
BOND CQ42 CQ43   1000.0  1.498
BOND CQ43 CQ44   1000.0  1.412

ANGLe OQ2  PQ1  OQ3     500.0   109.50
ANGLe OQ2  PQ1  OQ4     500.0   109.50
ANGLe OQ2  PQ1  OQ23    500.0   109.50
ANGLe OQ3  PQ1  OQ4     500.0   109.50
ANGLe OQ3  PQ1  OQ23    500.0   109.50
ANGLe OQ4  PQ1  OQ23    500.0   109.50
ANGLe PQ1  OQ4  CQ5     500.0   109.50
ANGLe OQ4  CQ5  CQ6     500.0   107.46
ANGLe CQ5  CQ6  OQ7     500.0   105.98
ANGLe CQ5  CQ6  CQ8     500.0   115.86
ANGLe OQ7  CQ6  CQ8     500.0   104.21
ANGLe CQ6  OQ7  CQ12    500.0   106.74
ANGLe CQ6  CQ8  OQ9     500.0   110.27
ANGLe CQ6  CQ8  CQ10    500.0   106.25
ANGLe OQ9  CQ8  CQ10    500.0   113.70
ANGLe CQ8  CQ10 OQ11    500.0   117.96
ANGLe CQ8  CQ10 CQ12    500.0   100.43
ANGLe OQ11 CQ10 CQ12    500.0   114.32
ANGLe OQ7  CQ12 CQ10    500.0   106.17
ANGLe OQ7  CQ12 NQ13    500.0   105.67
ANGLe CQ10 CQ12 NQ13    500.0   109.94
ANGLe CQ12 NQ13 CQ14    500.0   132.69
ANGLe CQ12 NQ13 CQ22    500.0   124.32
ANGLe CQ14 NQ13 CQ22    500.0   102.96
ANGLe NQ13 CQ14 NQ15    500.0   113.38
ANGLe CQ14 NQ15 CQ16    500.0   107.25
ANGLe NQ15 CQ16 CQ17    500.0   132.75
ANGLe NQ15 CQ16 CQ22    500.0   108.46
ANGLe CQ17 CQ16 CQ22    500.0   118.78
ANGLe CQ16 CQ17 NQ18    500.0   123.00
ANGLe CQ16 CQ17 NQ19    500.0   114.10
ANGLe NQ18 CQ17 NQ19    500.0   122.90
ANGLe CQ17 NQ19 CQ20    500.0   125.44
ANGLe NQ19 CQ20 NQ21    500.0   123.00
ANGLe CQ20 NQ21 CQ22    500.0   114.84
ANGLe NQ13 CQ22 CQ16    500.0   107.94
ANGLe NQ13 CQ22 NQ21    500.0   128.27
ANGLe CQ16 CQ22 NQ21    500.0   123.79
ANGLe PQ1  OQ23 PQ24    500.0   133.63
ANGLe OQ23 PQ24 OQ25    500.0   108.96
ANGLe OQ23 PQ24 OQ26    500.0   108.66
ANGLe OQ23 PQ24 OQ27    500.0    97.94
ANGLe OQ25 PQ24 OQ26    500.0   121.72
ANGLe OQ25 PQ24 OQ27    500.0   106.33
ANGLe OQ26 PQ24 OQ27    500.0   110.61
ANGLe PQ24 OQ27 CQ28    500.0   119.28
ANGLe OQ27 CQ28 CQ29    500.0   110.21
ANGLe CQ28 CQ29 OQ30    500.0   109.29
ANGLe CQ28 CQ29 CQ31    500.0   114.36
```

```
ANGLe OQ30 CQ29 CQ31    500.0   104.90
ANGLe CQ29 OQ30 CQ35    500.0   109.81
ANGLe CQ29 CQ31 OQ32    500.0   110.36
ANGLe CQ29 CQ31 CQ33    500.0   103.15
ANGLe OQ32 CQ31 CQ33    500.0   107.66
ANGLe CQ31 CQ33 OQ34    500.0   111.87
ANGLe CQ31 CQ33 CQ35    500.0   100.58
ANGLe OQ34 CQ33 CQ35    500.0   112.07
ANGLe OQ30 CQ35 CQ33    500.0   102.68
ANGLe OQ30 CQ35 NQ36    500.0   109.06
ANGLe CQ33 CQ35 NQ36    500.0   115.23
ANGLe CQ35 NQ36 CQ37    500.0   115.42
ANGLe CQ35 NQ36 CQ44    500.0   120.53
ANGLe CQ37 NQ36 CQ44    500.0   120.00
ANGLe NQ36 CQ37 CQ38    500.0   120.00
ANGLe CQ37 CQ38 CQ39    500.0   120.87
ANGLe CQ37 CQ38 CQ42    500.0   120.00
ANGLe CQ39 CQ38 CQ42    500.0   117.69
ANGLe CQ38 CQ39 OQ40    500.0   118.04
ANGLe CQ38 CQ39 NQ41    500.0   119.33
ANGLe OQ40 CQ39 NQ41    500.0   120.61
ANGLe CQ38 CQ42 CQ43    500.0   120.00
ANGLe CQ42 CQ43 CQ44    500.0   120.00
ANGLe NQ36 CQ44 CQ43    500.0   120.00

DIHEdral OQ23 PQ1  OQ4  CQ5     750.0 0    90.00
DIHEdral OQ4  PQ1  OQ23 PQ24    750.0 0    90.00
DIHEdral OQ4  CQ5  CQ6  OQ7      50.0 3   180.00
DIHEdral OQ4  CQ5  CQ6  CQ8     750.0 0   -60.00
DIHEdral OQ7  CQ6  CQ8  CQ10    750.0 0     0.00
DIHEdral OQ9  CQ8  CQ10 CQ12    750.0 0    90.00
DIHEdral CQ12 NQ13 CQ14 NQ15    750.0 0   180.00
DIHEdral CQ22 NQ13 CQ14 NQ15    750.0 0     0.00
DIHEdral CQ12 NQ13 CQ22 CQ16    750.0 0   180.00
DIHEdral CQ12 NQ13 CQ22 NQ21    750.0 0     0.00
DIHEdral CQ14 NQ13 CQ22 CQ16    750.0 0     0.00
DIHEdral CQ14 NQ13 CQ22 NQ21    750.0 0   180.00
DIHEdral NQ13 CQ14 NQ15 CQ16    750.0 0     0.00
DIHEdral CQ14 NQ15 CQ16 CQ17    750.0 0   180.00
DIHEdral CQ14 NQ15 CQ16 CQ22    750.0 0     0.00
DIHEdral NQ15 CQ16 CQ17 NQ18    750.0 0     0.00
DIHEdral NQ15 CQ16 CQ17 NQ19    750.0 0   180.00
DIHEdral CQ22 CQ16 CQ17 NQ18    750.0 0   180.00
DIHEdral CQ22 CQ16 CQ17 NQ19    750.0 0     0.00
DIHEdral NQ15 CQ16 CQ22 NQ13    750.0 0     0.00
DIHEdral NQ15 CQ16 CQ22 NQ21    750.0 0   180.00
DIHEdral CQ17 CQ16 CQ22 NQ13    750.0 0   180.00
DIHEdral CQ17 CQ16 CQ22 NQ21    750.0 0     0.00
DIHEdral CQ16 CQ17 NQ19 CQ20    750.0 0     0.00
DIHEdral NQ18 CQ17 NQ19 CQ20    750.0 0   180.00
DIHEdral CQ17 NQ19 CQ20 NQ21    750.0 0     0.00
DIHEdral NQ19 CQ20 NQ21 CQ22    750.0 0     0.00
DIHEdral CQ20 NQ21 CQ22 NQ13    750.0 0   180.00
DIHEdral CQ20 NQ21 CQ22 CQ16    750.0 0     0.00
DIHEdral PQ1  OQ23 PQ24 OQ25    750.0 0    90.00
DIHEdral OQ23 PQ24 OQ27 CQ28    750.0 0    60.00
DIHEdral OQ25 PQ24 OQ27 CQ28     50.0 3   180.00
DIHEdral OQ27 CQ28 CQ29 OQ30    750.0 0   -60.00
DIHEdral OQ30 CQ29 CQ31 OQ32    750.0 0   -90.00
DIHEdral OQ32 CQ31 CQ33 CQ35    750.0 0    90.00
DIHEdral CQ35 NQ36 CQ37 CQ38    750.0 0   180.00
DIHEdral NQ36 CQ44 CQ43 CQ42    750.0 0     0.00
DIHEdral NQ36 CQ37 CQ38 CQ39    750.0 0   180.00
DIHEdral NQ36 CQ37 CQ38 CQ42    750.0 0     0.00
DIHEdral CQ43 CQ42 CQ38 CQ37    750.0 0     0.00
DIHEdral CQ37 CQ38 CQ39 NQ41    750.0 0     0.00
DIHEdral CQ37 CQ38 CQ39 NQ40    750.0 0   180.00
DIHEdral CQ38 CQ42 CQ43 CQ44    750.0 0     0.00
```

```
DIHEdral CQ42 CQ38 CQ39 NQ41    750.0 0   180.00
DIHEdral CQ42 CQ38 CQ39 OQ40    750.0 0     0.00
DIHEdral CQ37 NQ36 CQ44 CQ43    750.0 0     0.00
DIHEdral CQ39 CQ38 CQ42 CQ43    750.0 0   180.00
DIHEdral CQ35 NQ36 CQ44 CQ43    750.0 0   180.00

IMPRoper PQ1  OQ2  OQ3  OQ4     750.0 0   35.000
IMPRoper CQ6  CQ5  OQ7  CQ8     750.0 0  -35.000
IMPRoper CQ8  CQ6  OQ9  CQ10    750.0 0  -35.000
IMPRoper CQ10 CQ8  OQ11 CQ12    750.0 0  -35.000
IMPRoper CQ12 OQ7  CQ10 NQ13    750.0 0   35.000
IMPRoper NQ13 CQ12 CQ14 CQ22    750.0 0    0.000
IMPRoper CQ16 NQ15 CQ17 CQ22    750.0 0    0.000
IMPRoper CQ17 CQ16 NQ18 NQ19    750.0 0    0.000
IMPRoper CQ22 NQ13 CQ16 NQ21    750.0 0    0.000
IMPRoper PQ24 OQ23 OQ25 OQ26    750.0 0   35.000
IMPRoper CQ29 CQ28 OQ30 CQ31    750.0 0  -35.000
IMPRoper CQ31 CQ29 OQ32 CQ33    750.0 0  -35.000
IMPRoper CQ33 CQ31 OQ34 CQ35    750.0 0  -35.000
IMPRoper CQ35 OQ30 CQ33 NQ36    750.0 0   35.000
IMPRoper NQ36 CQ35 CQ37 CQ44    750.0 0    0.000
IMPRoper CQ38 CQ37 CQ39 CQ42    750.0 0    0.000
IMPRoper CQ39 CQ38 OQ40 NQ41    750.0 0    0.000

NONBonded PQ1  0.5849  3.3854    0.5849  3.3854
NONBonded OQ2  0.1591  2.8509    0.1591  2.8509
NONBonded OQ3  0.1591  2.8509    0.1591  2.8509
NONBonded OQ4  0.1591  2.8509    0.1591  2.8509
NONBonded CQ5  0.1200  3.7418    0.1000  3.3854
NONBonded CQ6  0.1200  3.7418    0.1000  3.3854
NONBonded OQ7  0.1591  2.8509    0.1591  2.8509
NONBonded CQ8  0.1200  3.7418    0.1000  3.3854
NONBonded OQ9  0.1591  2.8509    0.1591  2.8509
NONBonded CQ10 0.1200  3.7418    0.1000  3.3854
NONBonded OQ11 0.1591  2.8509    0.1591  2.8509
NONBonded CQ12 0.1200  3.7418    0.1000  3.3854
NONBonded NQ13 0.2384  2.8509    0.2384  2.8509
NONBonded CQ14 0.1200  3.7418    0.1000  3.3854
NONBonded NQ15 0.2384  2.8509    0.2384  2.8509
NONBonded CQ16 0.1200  3.7418    0.1000  3.3854
NONBonded CQ17 0.1200  3.7418    0.1000  3.3854
NONBonded NQ18 0.2384  2.8509    0.2384  2.8509
NONBonded NQ19 0.2384  2.8509    0.2384  2.8509
NONBonded CQ20 0.1200  3.7418    0.1000  3.3854
NONBonded NQ21 0.2384  2.8509    0.2384  2.8509
NONBonded CQ22 0.1200  3.7418    0.1000  3.3854
NONBonded OQ23 0.1591  2.8509    0.1591  2.8509
NONBonded PQ24 0.5849  3.3854    0.5849  3.3854
NONBonded OQ25 0.1591  2.8509    0.1591  2.8509
NONBonded OQ26 0.1591  2.8509    0.1591  2.8509
NONBonded OQ27 0.1591  2.8509    0.1591  2.8509
NONBonded CQ28 0.1200  3.7418    0.1000  3.3854
NONBonded CQ29 0.1200  3.7418    0.1000  3.3854
NONBonded OQ30 0.1591  2.8509    0.1591  2.8509
NONBonded CQ31 0.1200  3.7418    0.1000  3.3854
NONBonded OQ32 0.1591  2.8509    0.1591  2.8509
NONBonded CQ33 0.1200  3.7418    0.1000  3.3854
NONBonded OQ34 0.1591  2.8509    0.1591  2.8509
NONBonded CQ35 0.1200  3.7418    0.1000  3.3854
NONBonded NQ36 0.2384  2.8509    0.2384  2.8509
NONBonded CQ37 0.1200  3.7418    0.1000  3.3854
NONBonded CQ38 0.1200  3.7418    0.1000  3.3854
NONBonded CQ39 0.1200  3.7418    0.1000  3.3854
NONBonded OQ40 0.1591  2.8509    0.1591  2.8509
NONBonded NQ41 0.2384  2.8509    0.2384  2.8509
NONBonded CQ42 0.1200  3.7418    0.1000  3.3854
NONBonded CQ43 0.1200  3.7418    0.1000  3.3854
NONBonded CQ44 0.1200  3.7418    0.1000  3.3854
```

5.5.2 CNS Topology File for NAD^+

```
MASS PQ1    30.97400
MASS OQ2    17.00700
MASS OQ3    17.00700
MASS OQ4    15.99900
MASS CQ5    14.02700
MASS CQ6    13.01900
MASS OQ7    15.99900
MASS CQ8    13.01900
MASS OQ9    17.00700
MASS CQ10   13.01900
MASS OQ11   17.00700
MASS CQ12   13.01900
MASS NQ13   14.00700
MASS CQ14   13.01900
MASS NQ15   15.01500
MASS CQ16   12.01100
MASS CQ17   12.01100
MASS NQ18   14.00700
MASS NQ19   15.01500
MASS CQ20   13.01900
MASS NQ21   15.01500
MASS CQ22   12.01100
MASS OQ23   15.99900
MASS PQ24   30.97400
MASS OQ25   17.00700
MASS OQ26   17.00700
MASS OQ27   15.99900
MASS CQ28   13.01900
MASS CQ29   13.01900
MASS OQ30   15.99900
MASS CQ31   13.01900
MASS OQ32   17.00700
MASS CQ33   13.01900
MASS OQ34   17.00700
MASS CQ35   13.01900
MASS NQ36   14.00700
MASS CQ37   13.01900
MASS CQ38   12.01100
MASS CQ39   12.01100
MASS OQ40   15.99900
MASS NQ41   14.00700
MASS CQ42   14.02700
MASS CQ43   13.01900
MASS CQ44   13.01900

ATOM AP    TYPE PQ1   CHARge  0.0  END
ATOM AO1   TYPE OQ2   CHARge  0.0  END
ATOM AO2   TYPE OQ3   CHARge  0.0  END
ATOM AO5*  TYPE OQ4   CHARge  0.0  END
ATOM AC5*  TYPE CQ5   CHARge  0.0  END
ATOM AC4*  TYPE CQ6   CHARge  0.0  END
ATOM AO4*  TYPE OQ7   CHARge  0.0  END
ATOM AC3*  TYPE CQ8   CHARge  0.0  END
ATOM AO3*  TYPE OQ9   CHARge  0.0  END
ATOM AC2*  TYPE CQ10  CHARge  0.0  END
ATOM AO2*  TYPE OQ11  CHARge  0.0  END
ATOM AC1*  TYPE CQ12  CHARge  0.0  END
ATOM AN9   TYPE NQ13  CHARge  0.0  END
ATOM AC8   TYPE CQ14  CHARge  0.0  END
ATOM AN7   TYPE NQ15  CHARge  0.0  END
ATOM AC5   TYPE CQ16  CHARge  0.0  END
ATOM AC6   TYPE CQ17  CHARge  0.0  END
ATOM AN6   TYPE NQ18  CHARge  0.0  END
ATOM AN1   TYPE NQ19  CHARge  0.0  END
ATOM AC2   TYPE CQ20  CHARge  0.0  END
```

```
ATOM AN3   TYPE NQ21  CHARge  0.0  END
ATOM AC4   TYPE CQ22  CHARge  0.0  END
ATOM  O3   TYPE OQ23  CHARge  0.0  END
ATOM NP    TYPE PQ24  CHARge  0.0  END
ATOM NO1   TYPE OQ25  CHARge  0.0  END
ATOM NO2   TYPE OQ26  CHARge  0.0  END
ATOM NO5*  TYPE OQ27  CHARge  0.0  END
ATOM NC5*  TYPE CQ28  CHARge  0.0  END
ATOM NC4*  TYPE CQ29  CHARge  0.0  END
ATOM NO4*  TYPE OQ30  CHARge  0.0  END
ATOM NC3*  TYPE CQ31  CHARge  0.0  END
ATOM NO3*  TYPE OQ32  CHARge  0.0  END
ATOM NC2*  TYPE CQ33  CHARge  0.0  END
ATOM NO2*  TYPE OQ34  CHARge  0.0  END
ATOM NC1*  TYPE CQ35  CHARge  0.0  END
ATOM NN1   TYPE NQ36  CHARge  0.0  END
ATOM NC2   TYPE CQ37  CHARge  0.0  END
ATOM NC3   TYPE CQ38  CHARge  0.0  END
ATOM NC7   TYPE CQ39  CHARge  0.0  END
ATOM NO7   TYPE OQ40  CHARge  0.0  END
ATOM NN7   TYPE NQ41  CHARge  0.0  END
ATOM NC4   TYPE CQ42  CHARge  0.0  END
ATOM NC5   TYPE CQ43  CHARge  0.0  END
ATOM NC6   TYPE CQ44  CHARge  0.0  END

BOND AP   AO1       BOND AP   AO2       BOND AP   AO5*      BOND AP    O3
BOND AO5* AC5*      BOND AC5* AC4*      BOND AC4* AO4*      BOND AC4* AC3*
BOND AO4* AC1*      BOND AC3* AO3*      BOND AC3* AC2*      BOND AC2* AO2*
BOND AC2* AC1*      BOND AC1* AN9       BOND AN9  AC8       BOND AN9  AC4
BOND AC8  AN7       BOND AN7  AC5       BOND AC5  AC6       BOND AC5  AC4
BOND AC6  AN6       BOND AC6  AN1       BOND AN1  AC2       BOND AC2  AN3
BOND AN3  AC4       BOND  O3  NP        BOND NP   NO1       BOND NP   NO2
BOND NP   NO5*      BOND NO5* NC5*      BOND NC5* NC4*      BOND NC4* NO4*
BOND NC4* NC3*      BOND NO4* NC1*      BOND NC3* NO3*      BOND NC3* NC2*
BOND NC2* NO2*      BOND NC2* NC1*      BOND NC1* NN1       BOND NN1  NC2
BOND NN1  NC6       BOND NC2  NC3       BOND NC3  NC7       BOND NC3  NC4
BOND NC7  NO7       BOND NC7  NN7       BOND NC4  NC5       BOND NC5  NC6

DIHEdral AO5* AC5* AC4* AO4*
DIHEdral AC1* AN9  AC8  AN7
DIHEdral AC4  AN9  AC8  AN7
DIHEdral AC1* AN9  AC4  AC5
DIHEdral AC1* AN9  AC4  AN3
DIHEdral AC8  AN9  AC4  AC5
DIHEdral AC8  AN9  AC4  AN3
DIHEdral AN9  AC8  AN7  AC5
DIHEdral AC8  AN7  AC5  AC6
DIHEdral AC8  AN7  AC5  AC4
DIHEdral AN7  AC5  AC6  AN6
DIHEdral AN7  AC5  AC6  AN1
DIHEdral AC4  AC5  AC6  AN6
DIHEdral AC4  AC5  AC6  AN1
DIHEdral AN7  AC5  AC4  AN9
DIHEdral AN7  AC5  AC4  AN3
DIHEdral AC6  AC5  AC4  AN9
DIHEdral AC6  AC5  AC4  AN3
DIHEdral AC5  AC6  AN1  AC2
DIHEdral AN6  AC6  AN1  AC2
DIHEdral AC6  AN1  AC2  AN3
DIHEdral AN1  AC2  AN3  AC4
DIHEdral AC2  AN3  AC4  AN9
DIHEdral AC2  AN3  AC4  AC5
DIHEdral NO1  NP   NO5* NC5*
DIHEdral NC2  NN1  NC6  NC5
DIHEdral NN1  NC2  NC3  NC7
DIHEdral NN1  NC2  NC3  NC4
DIHEDRAL NC3  NC4  NC5  NC6
DIHEDRAL NC7  NC3  NC4  NC5
```

```
IMPRoper AP   AO1  AO2  AO5*
IMPRoper AC4* AC5* AO4* AC3*
IMPRoper AC3* AC4* AO3* AC2*
IMPRoper AC2* AC3* AO2* AC1*
IMPRoper AC1* AO4* AC2* AN9
IMPRoper AN9  AC1* AC8  AC4
IMPRoper AC5  AN7  AC6  AC4
IMPRoper AC6  AC5  AN6  AN1
IMPRoper AC4  AN9  AC5  AN3
IMPRoper NP    O3  NO1  NO2
IMPRoper NC4* NC5* NO4* NC3*
IMPRoper NC3* NC4* NO3* NC2*
IMPRoper NC2* NC3* NO2* NC1*
IMPRoper NC1* NO4* NC2* NN1
IMPRoper NN1  NC1* NC2  NC6
IMPRoper NC3  NC2  NC7  NC4
IMPRoper NC7  NC3  NO7  NN7

ACCEptor AO1  AP
ACCEptor AO2  AP
ACCEptor AO5* AP
ACCEptor AO4* AC4*
ACCEptor AO3* AC3*
ACCEptor AO2* AC2*
ACCEptor  O3  AP
ACCEptor NO1  NP
ACCEptor NO2  NP
ACCEptor NO5* NP
ACCEptor NO4* NC4*
ACCEptor NO3* NC3*
ACCEptor NO2* NC2*
ACCEptor NO7  NC7
```

5.6 CNS Files for Cysteine-Sulfinic Acid

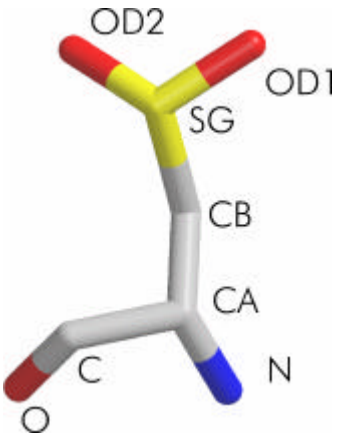

Figure 5.7 Schematic figure of cysteine-sulfinic acid (residue CSW)

5.6.1 CNS Parameter File for Cysteine-Sulfinic Acid

```
BOND NX1  CX2    1000.0  1.455
BOND CX2  CX3    1000.0  1.535
BOND CX2  CX5    1000.0  1.550
BOND CX3  SX4    1000.0  1.837
BOND SX4  OX7    1000.0  1.471
BOND SX4  OX8    1000.0  1.467
BOND CX5  OX6    1000.0  1.238

ANGLe NX1  CX2  CX3     500.0   110.41
ANGLe NX1  CX2  CX5     500.0   101.76
ANGLe CX3  CX2  CX5     500.0   111.57
ANGLe CX2  CX3  SX4     500.0   116.12
ANGLe CX3  SX4  OX7     500.0   107.89
ANGLe CX3  SX4  OX8     500.0   108.86
ANGLe OX7  SX4  OX8     500.0   113.67
ANGLe CX2  CX5  OX6     500.0   120.77

DIHEdral CX5  CX2  CX3  SX4     750.0 0   180.00
DIHEdral NX1  CX2  CX5  OX6     750.0 0     0.00
DIHEdral CX2  CX3  SX4  OX7     750.0 0   -60.00

IMPRoper CX2  NX1  CX3  CX5     750.0 0  -35.000
IMPRoper SX4  CX3  OX7  OX8     750.0 0  -35.000

NONBonded NX1  0.2384  2.8509    0.2384  2.8509
NONBonded CX2  0.1200  3.7418    0.1000  3.3854
NONBonded CX3  0.1200  3.7418    0.1000  3.3854
NONBonded SX4  0.0430  3.3676    0.0430  3.3676
NONBonded CX5  0.1200  3.7418    0.1000  3.3854
NONBonded OX6  0.1591  2.8509    0.1591  2.8509
NONBonded OX7  0.1591  2.8509    0.1591  2.8509
NONBonded OX8  0.1591  2.8509    0.1591  2.8509
```

5.6.2 CNS Topology File for Cysteine-Sulfinic Acid

```
MASS NX1    15.01500
MASS CX2    13.01900
MASS CX3    14.02700
MASS SX4    32.06600
MASS CX5    13.01900
MASS OX6    15.99900
MASS OX7    17.00700
MASS OX8    17.00700

ATOM  N    TYPE NX1   CHARge  0.0  END
ATOM  CA   TYPE CX2   CHARge  0.0  END
ATOM  CB   TYPE CX3   CHARge  0.0  END
ATOM  SG   TYPE SX4   CHARge  0.0  END
ATOM  C    TYPE CX5   CHARge  0.0  END
ATOM  O    TYPE OX6   CHARge  0.0  END
ATOM  OD1  TYPE OX7   CHARge  0.0  END
ATOM  OD2  TYPE OX8   CHARge  0.0  END

BOND  N    CA       BOND  CA   CB       BOND  CA   C        BOND  CB   SG
BOND  SG   OD1      BOND  SG   OD2      BOND  C    O

DIHEdral  C    CA   CB   SG
DIHEdral  N    CA   C    O

IMPRoper  CA   N    CB   C
IMPRoper  SG   CB   OD1  OD2

ACCEptor  O    C
ACCEptor  OD1  SG
ACCEptor  OD2  SG
```

5.7 List of GH4 enzymes

Abbreviation	Protein ID	Description	Organism
AGLA_THEMA	O33830	α-glucosidase	*Thermotoga maritima*
AGAL_BACSU	O34645	α-galactosidase	*Bacillus subtilis*
MALH_FUSMR	O06901	Maltose-6'-phosphate glucosidase	*Fusobacterium mortiferum*
AGAL_ECOLI	P06720	α-galactosidase	*Escherichia coli*
CELF_BACSU	P46320	Probable 6-phospho-β-glucosidase	*Bacillus subtilis*
CELF_ECOLI	P17411	6-phospho-β-glucosidase	*Escherichia coli*
LPLD_BACSU	P39130	LPLD protein	*Bacillus subtilis*
GLVG_BACSU	P54716	Maltose-6'-phosphate glucosidase	*Bacillus subtili.*
ORF_1Bacillus	15613475	6-phospho-β-glucosidase	*Bacillus halodurans*
ORF_2Bacillus	15612746	6-phospho-β-glucosidase	*Bacillus halodurans*
ORF_Citrobacter	12249127	α-galactosidase	*Citrobacter freundii*
ORF_1Clostridium	15023397	Maltose-6'-phosphate glucosidase	*Clostridium acetobutylicum*
ORF_2Clostridium	15026514	6-phospho-α-glucosidase	*Clostridium acetobutylicum*
ORF_3Clostridium	18143856	Maltose-6'-phosphate glucosidase	*Clostridium perfringens*
ORF_Erwinia	13517313	Putative sugar hydrolase	*Erwinia rhapontici*
ORF_1Escherichia	12515750	Phospho-β-glucosidase; cryptic	*Escherichia coli O157:H7 EDL933*
ORF_2Escherichia	12519092	α-galactosidase	*Escherichia coli O157:H7 EDL933*
ORF_3Escherichia	12518523	Putative 6-phospho-β-glucosidase	*Escherichia coli O157:H7 EDL933*
ORF_Klebsiella	12667493	6-phospho-α-glucosidase	*Klebsiella pneumoniae*
ORF_1Listeria	16799601	Similar to 6-phospho-β-glucosidase	*Listeria innocua*
ORF_2Listeria	16799615	Similar to 6-phospho-β-glucosidase	*Listeria innocua*
ORF_3Listeria	16801516	Similar to 6-phospho-β-glucosidase	*Listeria innocua*
ORF_4Listeria	16802564	Similar to 6-phospho-β-glucosidase	*Listeria monocytogenes EGD-e*
ORF_5Listeria	16802579	Similar to 6-phospho-β-glucosidase	*Listeria monocytogenes EGD-e*
ORF_1Mesorhizobium	14026141	α-galactosidase	*Mesorhizobium loti*
ORF_2Mesorhizobium	14026146	α-galactosidase	*Mesorhizobium loti*
ORF_1Salmonella	16502875	Phospho-β-glucosidase B	*Salmonella enterica*
ORF_2Salmonella	16505287	α-galactosidase	*Salmonella enterica*
ORF_3Salmonella	16419834	Phospho-β-glucosidase	*Salmonella typhimurium LT2*
ORF_1Sinorhizobium	15141450	Probable α-galactosidase	*Sinorhizobium meliloti*
ORF_2Sinorhizobium	15141176	Probable 6-phospho-β-glucosidase	*Sinorhizobium meliloti*
ORF_1Streptomyces	6855385	Putative sugar hydrolase	*Streptomyces coelicolor A3(2)*
ORF_2Streptomyces	6900960	Putative sugar hydrolase (partial)	*Streptomyces coelicolor A3(2)*
ORF_3Streptomyces	6137043	Putative α-galactosidase	*Streptomyces coelicolor A3(2)*
ORF_patent	10059896	Sequence 8 from patent US 5985622	*Erwinia rhapontici*
ORF_Vibrio	9655770	6-phospho-β-glucosidase	*Vibrio cholerae*
ORF_Yersinia	15978271	Putative glycosyl hydrolase	*Yersinia pestis*

5.8 Curriculum vitae

Full Name: Jacinta Anne Lodge

Date of Birth: 11.08.1975

Nationality: Australian

Education:

Primary School	1980-1982	Lutheran Primary School, Hamilton
	1983	Roslyn Primary School, Geelong
	1984-1987	Grovedale Primary School, Geelong
High School	1988-1993	Belmont High School, Geelong, Australia Achieved Victorian Certificate of Eductation Subjects Biology, Chemistry, Physics, Maths, Music - Solo Performance, English Score: 132/164
University	1994-1997	La Trobe University, Melbourne, Australia Bachelor of Science (Biological) with Honours Honours (Masters equivalent) Score: H1 (first class)

Career-related Employment:

01.12.96 - 31.01.97	**LaTrobe University** Department of Biochemistry Summer Studentship Trypsin inhibition by Anthracyclins
01.02.97 – 30-10.97	**La Trobe University** Department of Biochemistry Honours Student (Masters equivalent) Identification and purification of proteases from *Nicotiana tabacum*
01.05.97 - 31.02.98	**LaTrobe University** Department of Biochemistry Laboratory Technician Care of plant glasshouse and caterpillar colony. Dissection of caterpillars
01.03.98 - 31.03.99	**Hexima, Ltd** Research Assistant Agrobacterium mediated transformation of cotton plants for insect resistance. Development of ELISA-based screening technique for transformants. Production and purification of rabbit antibodies.

20.09.99 - 30.07.00	**Max-Planck Forschungsstelle "Enzymologie der Proteinfaltung"** Halle/Saale Studentship Mitochondrial import. *In vitro* protein production, radioactive assays
01.08.00 - current	**Freie Universität Berlin** Institut für Chemie - Kristallographie PhD Student Protein Crystallography: Cloning, expression and purification of proteins for crystal production. Protein crystallisation. X-ray data collection and processing using rotating anode on X-ray generators and synchrotronsources. Building and refinement of protein structures.

Seminars and Conference Participation

1997 Australian Society for Biochemistry and Molecular Biology, Melbourne, Australia
1998 Lorne Protein Conference, Lorne, Australia
1999 MPI-Halle, IMB-Jena combined meeting, Wittenberg, Germany
2001 German Biochemical Society Meeting, Bochum, Germany
Poster presentation
2002 Heart of Europe Crystallography Meeting, Goslar, Germany
Talk: A new glucosidase fold
2003 DGK Meeting, Berlin, Germany
Poster presentation

5.9 Publication of this work

Talk:

"A new glucosidase fold"

2002 Heart of Europe Crystallography Meeting, Goslar, Germany. September 25 – 27, 2002.

Poster Presentations:

Lodge, J.A, Maier, T., Raasch, C., Leibl, W. and Strater, N.

"The crystal structure of *Thermotoga maritima* α-glucosidase A defines a new clan of NAD^+-dependent glucosidases". Elfte Jahrestagung der Deutschen Gesellschaft für Kristallographie, Berlin, Germany. March 10 – 12, 2003

Lodge, J.A, Maier, T., Raasch, C., Leibl, W. and Strater, N.

"A new clan of glucosidases: The crystal structure of *Thermotoga maritima* α-glucosidase A". Seventh International School on the Crystallography of Biological Macromolecules, Como, Italy. May 10 - 14, 2003.

Publication:

Lodge, J.A, Maier, T., Raasch, C., Leibl, W. and Strater, N. (2003)

"Crystal structure of family 4 *Thermotoga maritima* AglA defines a new clan of NAD^+-dependent glucosidases". *J. Biol. Chem.* **278** (21):19151-19158.

PDB Submission:

The atomic coordinates and structure factors (code 10BB) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ. (http://www.rcsb.org/).