Chapter 2

Methods and Experimental Strategy

2.1 Biochemical Methods

2.1.1 Sequence Analysis

To solve the structure of the BRCT region from human BRCA1, a special strategic approach had to be developed. Since no structural information from homologous domains had been available at that time, one important question was concerned with the boundaries of the BRCT domains. Sequence analysis identified the most conserved residues within the BRCT region. Many of these were hydrophobic (aliphatic or aromatic) and had been predicted to be part of the hydrophobic core of the BRCT domains [32,33]. In addition sequence and hydrophobic cluster analysis revealed regions which were potentially part of α-helices or β-sheets, and others which potentially formed loops. Despite these valuable contributions from bioinformatics, the identification of domain boundaries within protein sequences remains rather challenging. Since sequence similarity of homologous proteins and domains tend to decrease toward the termini the accuracy of prediction exceeds only rarely ± 3-5 amino acid residues.

2.1.2 Molecular Cloning

A set of BRCT constructs were generated by molecular cloning, including the isolated single BRCT domains, BRCT-n (BRCA1$^{[1646\text{-}1744]}$) and BRCT-c (BRCA1$^{[1756\text{-}1863]}$), two variants of the BRCT tandem repeat (BRCA1$^{[1646/1650\text{-}1863]}$), and two amino-terminally extended variants (BRCA1$^{[1559/1590\text{-}1863]}$). BRCT-encoding DNA fragments were generated by a two-step PCR-amplification (nested PCR) of cDNA templates. PCR-amplified DNA fragments were processed with the restriction endonucleases BamH1 and Sal1, purified, and ligated into the multiple cloning site (MCS) of the expression plasmid pGEX-4T1 (GE HEALTHCARE, formerly AMERSHAM BIOSCIENCES).

2.1.3 Proteolytic Dissection

Limited proteolysis provides an easy and very sensitive method to assess the protein quality in terms of homogeneity and stability [45]. In a first approximation, stable protein domains behave like small proteins, both of which represent autonomous folding units. In the folded state, most small proteins and stable domains form compact globular entities which are usually very resistant to proteolytic degradation because potential cleavage sites are hardly accessible. Unstructured, solvent-exposed loop or linker regions, however, may be subject to rapid proteolysis. By using limited amounts of different (site-specific) proteases, properly folded proteins/domains therefore remain largely unaffected whereas un-/misfolded polypeptides are cleaved down to more stable fragments. In practice, limited proteolysis experiments are designed such that the formation of intermediate cleavage products can be followed, depending on the time course and/or the amount of protease. The fragments most resistant to proteolytic degradation are usually identified by mass spectrometry and N-terminal sequencing. To identify the domain boundaries experimentally and to assess the stability of the BRCT region, various BRCT constructs were recombinantly expressed in *E. coli* as GST fusion proteins and subjected to limited proteolysis experiments using trypsin, chymotrypsin, elastase, thermolysin, subtilisin, and V8 protease (endo-Glu-C) (mass ratio of protease to protein 1:200 – 1:1000). This approach identified the most stable BRCT construct to be used for structure analysis. By using GST-BRCT fusion proteins, GST served as BRCT-independent control for proteolytic dissection as GST is largely resistant toward proteolytic degradation in these assays. The proteolytic fragments were separated and analyzed by SDS-PAGE.

2.1.4 Co-Expression with the Chaperonins GroEL/GroES and tRNAs

Heterologous expression of human proteins in prokaryotic cells may be disfavored by at least three major factors: (i) poor expression due to a different codon usage in the host organism and, by that, the presence of rare codons in the human/recombinant gene, (ii) improper and/or incomplete folding due to the lack of eukaryotic chaperones and/or post-translational modification systems, and (iii) the biological function of the target protein which may interfere with vital cellular functions in the host organism. To overcome these difficulties associated with protein expression and folding, *E. coli* BL21[pREP4-groESL/CodonPlusRIL] cells expressing recombinant $tRNA^{Arg}$, $tRNA^{Ile}$, $tRNA^{Leu}$ (harboring human anticodons) were used. In addition the IPTG-inducible chaperonins GroEL and GroES were co-expressed to improve the yield of properly folded target protein. The important role of the GroEL/GroES

chaperone system for efficient protein folding in the crowded environment inside living cells is described in Appendix A.2.1 ("Protein Folding *in vitro* and *in vivo*").

Cells were grown at 32°C in LB medium, supplemented with 1.3% (w/v) D-glucose, or M9 minimal medium in case of ^{13}C- and/or ^{15}N-labeling, containing 100 mg/L carbenicillin, 30 mg/L kanamycin, and 34 mg/L chloramphenicol as antibiotics. Expression of recombinant proteins and the chaperonins GroEL/GroES were induced by the addition of 0.5 mM IPTG at an OD_{600} (optical density at 600 nm) of 0.7-0.9. To avoid intracellular aggregation and the formation of inclusion bodies the growth temperature was reduced to 25°C prior to induction. After an additional 5-8 hours incubation cells were harvested and lysed with a French press.

2.1.5 Improved Protein Folding *in vitro*

During this thesis work, a protein purification protocol was established that made use of the over-expressed chaperonins GroEL/GroES after cell lysis: Chaperonin-assisted protein folding relies on the energy source ATP as well as Mg^{2+}, K^+ and Cl^- ions as co-factors. Adding these components to the lysis buffer therefore effectively creates an *in vitro* (re)folding step prior to the protein purification procedure. In fact, the system GroEL/GroES, Mg^{2+}-ATP and KCl lowers the energy barrier between misfolded intermediates and the folded state. Due to the high substrate affinity of chaperonins, GroEL/GroES very often co-purify with un-/misfolded intermediates. The Mg^{2+}-ATP driven chaperonin reaction cycle, however, assists in folding, thereby lowering the affinity of GroEL/GroES toward the target protein. By using highly selective affinity chromatography under native-like conditions, most unbound components, including GroEL/GroES, are separated from the recombinant target protein, whereas co-purifying GroEL/GroES can be considered indicative of impurities from un-/misfolded intermediates. Extended incubation times (2-3 hours, 4°C) were used to allow GroEL/GroES assisted protein (re)folding prior to protein purification. This effectively creates the prerequisites for coupled (re)folding and limited proteolysis: un-/misfolded intermediates are either trapped by the chaperonin system or hydrolyzed by host proteases present in the reaction mixture upon cell lysis. By that the chaperone-mediated (re)folding step can be monitored by analyzing the protein fraction that is not degraded in limited proteolysis assays.

2.1.6 Isolation of Recombinant Proteins

Molecular cloning techniques and the rational design of recombinant proteins have become a matter of routine over the last decade. In many cases the target protein is expressed as fusion protein that is covalently linked to polypeptides and/or protein domains with defined properties. The fusion partner may be used to maintain or improve the solubility of the target inside the host cell, as it is the case with maltose binding protein (MBP), thioredoxin (Trx) and solubility enhancement tags (SETs), or it may serve as high-affinity interaction partner for an immobilized binding target in an early chromatography step. The most important peptide and protein affinity tags are oligohistidine tags [$(His)_{6-10}$], Strep-tag (based on the strong biotin-streptavidin interaction), and glutathione S-transferase (GST). Expression, solubility and affinity purification of various BRCT constructs were tested with different affinity tags, such as $(His)_6$, MBP and GST.

Affinity Chromatography

The molecular basis for affinity chromatography is the highly specific and strong interaction between the affinity tag of the recombinant fusion protein and an immobilized binding partner. Since the interaction is exceedingly specific and strong only very few host proteins co-purify with the fusion protein under stringent conditions. Recovery of the recombinant fusion protein from the column matrix is routinely achieved by the addition of a soluble form of the immobilized binding partner or a low-molecular weight competitor. By this, the affinity chromatography step may yield the target (fusion) protein with less than 10% impurities. Usually the peptide linker between the target protein and its fusion partner is rather flexible which may be obstructive for structural studies. Therefore the linker usually contains a protease recognition motif which allows the affinity tag to be cleaved by site-specific proteases.

For most BRCT constructs, the best results in terms of protein solubility and purity upon affinity chromatography were obtained with GST fusions. Recombinantly expressed GST-BRCT fusion proteins were isolated from crude cell extract components by affinity chromatography with glutathione-sepharose 4B (GE HEALTHCARE, formerly AMERSHAM BIOSCIENCES) which is based on the highly specific interaction between GST and its natural substrate glutathione. GST fusion proteins were routinely eluted with 10 mM glutathione dissolved in an aqueous buffered solution (pH 7-8).

Proteolytic Cleavage of Fusion Proteins

Upon elution from the affinity matrix, GST fusion proteins were cleaved with thrombin which hydrolyses the peptide bond between Arg and Gly in the recognition sequence [Leu-Val-Pro-Arg↓Gly-Ser], thus yielding the protein target with an additional dipeptide (Gly-Ser) at the amino terminus. The BRCT-c construct ($BRCA1^{[1756\text{-}1863]}$) was designed such that the serine from the thrombin recognition site was already part of the native sequence (Ser1755). By that, the number of additional non-native amino acid residues could be minimized to only one single Gly residue. The cleavage reaction was stopped by addition of the serine protease inhibitor PMSF or the more stable Pefabloc® SC ([4-(2-Aminoethyl)-benzenesulfonyl-fluoride·HCl], ROCHE APPLIED SCIENCE).

Size Exclusion Chromatography (SEC)

The principle of size exclusion chromatography (SEC) is based on a fine pored resin material with well defined pore size. Large molecules do not enter the small pores and are thereby effectively excluded from parts of the column bed volume. Small molecules and solutes, however, pass through the pores and utilize basically the entire bed volume. As a result, large molecules elute first, whereas ions and other small solutes come off last. SEC may therefore be used to combine buffer exchange with a purification step.

In the case of BRCT, SEC was applied to remove small compounds (PMSF, ATP, glutathione) and to separate the BRCT constructs from uncleaved fusion proteins, GST, and high-molecular weight complexes or aggregates, but also to reduce the salt concentration for the following ion exchange (IEX) chromatography step. To reduce unspecific hydrophobic interactions and to avoid an additional buffer exchange step prior to the subsequent IEX chromatography, SEC of BRCT constructs was performed in a buffer containing 20 mM KP_i (pH 6.8), 60 mM KCl, supplemented with 3-5 mM fresh DTT to avoid oxidation of solvent exposed cysteine residues.

Ion Exchange Chromatography

Ion exchange chromatography may be considered a special case of affinity chromatography in which binding is mediated by ionic interactions between charged or strongly polarized functional groups. The matrix of an ion exchange (IEX) column is either positively or negatively charged, thus working either as an anion (AIEX) or cation exchanger (CIEX), respectively. The most prominent advantage of IEX chromatography is its very high resolution. In principle it is possible to separate almost identical polypeptides that differ in only one single charged residue. The method is therefore especially valuable for the detection of differ-

ent oligomeric states or fractions of partially degraded proteins in inhomogeneous samples, both of which may be seriously detrimental for crystallization. It is therefore desirable to purify proteins with the highest possible quality (i.e. homogeneity and stability). In fact, if the protein elutes in a *single* Gaussian-like peak from an IEX column, it may be taken as very strong indication of a highly homogeneous sample.

A Mono-Q AIEX column (GE HEALTHCARE, formerly AMERSHAM BIOSCIENCES) was used as final polishing step and quality control for various BRCT constructs. The main single peak fractions were used for initial crystallization screening and NMR experiments after dialysis. The experimental set-up of molecular cloning and isolation of recombinant BRCT constructs are described in [43].

2.2 Structural Biological Methods

In a first approach toward the structure determination of the BRCA1 BRCT region a number of different BRCT constructs were highly purified for large-scale crystallization screening experiments, including two variants of the BRCT tandem repeats, BRCA1$^{[1646-1863]}$ and BRCA1$^{[1650-1863]}$ (1646-BRCT and 1650-BRCT, respectively), and the isolated second BRCT domain BRCT-c (BRCA1$^{[1756-1863]}$). Initial crystallization screening was performed with commercially available Crystal Screens (HAMPTON RESEARCH, CA, U.S.A.) at different temperatures (usually 4°C, 20°C and 30°C) and with different protein concentrations using the methods of both hanging- and sitting-drop vapor-diffusion. In the crystallization set-ups typically 1-2 µL of buffer solution were mixed and equilibrated over a reservoir of undiluted buffer solution. To modulate the protein saturation kinetics the protein-to-buffer volume ratio was also varied. In addition the effect of ionic additives, such as Ca^{2+}, Mg^{2+} and Zn^{2+} was tested. Granular precipitate/microcrystals which formed in some droplets were used for (micro-)seeding experiments. However, as none of the various BRCT constructs formed diffraction quality crystals, solution NMR spectroscopy was used to obtain structural information about the BRCT region of BRCA1. Considering the important role of the BRCA1 C-terminus for cell cycle regulation and DNA repair, structural studies were performed on wild-type BRCT domains without any C-terminal truncations.

2.2.1 NMR Spectroscopy (I): General Approach

NMR spectroscopy imposes experimental restraints in terms of sample preparation, buffer conditions, temperature, and protein concentration, to name but a few. Since NMR data rely on the magnetic properties of protons, hydrogen exchange rates should be kept below the time scale of magnetization transfer and relaxation. Signal intensity therefore is normally highest below pH 6.0. However, pH may interfere with protein stability and solubility in the test tube. Denatured/precipitated protein fractions are associated with substantially increased rotational correlation times and weak signal intensities and therefore negatively affect the signal-to-noise ratio. In principle, there is a trade-off between solubility and signal-to-noise ratio on one hand, and denaturation and/or aggregation on the other. In practice, the sample temperature is adjusted as high as possible without loss of data quality, typically in the range of 30-40°C. In summary, experimental conditions have to be properly adjusted to maximize protein stability and solubility as well as the signal-to-noise ratio of NMR data.

Initial ^{1}H-^{15}N HSQC Screening of Candidate BRCT Constructs

The ^{1}H-^{15}N HSQC spectrum represents a very sensitive NMR application which was used for diagnostic purposes to determine the protein quality of different BRCT constructs. The ^{1}H-^{15}N HSQC spectrum correlates ^{15}N and proton (^{1}H) resonance frequencies, giving one NMR signal for every pair of covalently bound N-H atoms. Therefore, every amino acid residue (except for proline) in the protein sequence would give rise to at least one NMR signal due to its backbone amide group. Additional signals may be attributed to the side chain amido, guanidinium and indole functional groups of asparagine/glutamine, arginine and tryptophan residues, respectively. Ideally, the number and resonance frequencies of HSQC signals match perfectly with those expected for the protein sample. In addition, high dispersion of NMR signals within a broad range of resonance frequencies is strongly indicative of (i) a properly folded protein sample and (ii) a balanced ratio of helical and extended conformation. In general, this approach could be readily used as diagnostic tool to screen for protein constructs most suitable for structure analysis by NMR spectroscopy and/or X-ray crystallography.

As a starting point for the structure determination by solution NMR spectroscopy, a series of one-dimensional (1D) ^{1}H-NMR and two-dimensional (2D) ^{1}H-^{15}N HSQC spectra were acquired with the BRCT tandem repeat (1646-BRCT, 220 amino acid residues) and the isolated second BRCT domain BRCT-c (1756-BRCT, 110 amino acid residues) to optimize the experimental conditions. The ^{15}N isotopic labeling of recombinantly expressed BRCT constructs is described in the accompanying publication [43]. ^{1}H-NMR and ^{1}H-^{15}N HSQC spectra identified the isolated second BRCT domain (BRCT-c) as most promising target for NMR structure analysis in solution. This BRCT construct was highly resistant to proteolytic degradation but showed a tendency toward aggregation and/or precipitation in high-salt buffer solutions below pH 6.0. Thus, the conditions for NMR experiments were adjusted to pH 6.8 and very low ionic strength. All NMR experiments used to derive experimental restraints for structure calculations were performed at 296 K (ca. 23 °C) in phosphate buffer (pH 6.8), 5 mM KCl, and 10 mM fresh fully deuterated DTT to avoid oxidation of solvent exposed cysteine residues.

Resonance Assignment and Structure Calculation

To identify the resonance frequencies of coupling nuclei, a number of different NMR experiments were performed on ^{15}N- and ^{13}C/^{15}N-labeled BRCT-c samples. High resolution 2D and 3D ^{15}N-edited NOESY spectra were used for NOE assignment and distance class classification. Structure calculations were performed using XPLOR (version 3.1) and CNS

(version 1.0). Distance restraints based on NOE assignments were classified as strong (≤ 2.5 Å), medium (≤ 3.5 Å), weak (≤ 5.0 Å) and very weak (≤ 6.0 Å). Dihedral-angle restraints were obtained from $^{3}J_{HNHA}$ experiments and backbone chemical shifts using the program TALOS [46]. Hydrogen-bond restraints were identified from a 2D ^{1}H-^{15}N HSQC spectrum following exchange into D_2O. Details of NMR data acquisition and processing, assignment (^{1}H, ^{13}C and ^{15}N nuclei), and structure calculations are described in [42,43].

Cross-Checking of NMR Assignments with CYANA

To avoid the problem of model bias, the automated assignment algorithm of combined CANDID and DYANA (CYANA) [47] was tested as independent cross-check. As CYANA uses not more than the positions and signal intensities of NOESY cross-peaks (NOEs), and the resonance frequencies derived from the sequential assignment, the algorithm can be used as assignment cross-check for manually assigned NMR structures. The CYANA approach using the concepts of ***network-anchoring*** and ***restraint-combination*** [47] is addressed in Appendix A.3.2 ("Automated NOE Assignment and Structure Calculation").

NMR ensembles of BRCT-c calculated with CYANA were of similar precision in terms of the ensemble RMSD values, but the program identified additional NOEs that were not considered during manual assignment and CNS structure calculation. These new assignments were subsequently verified in the NOESY spectra by manual inspection and then used as additional restraints in further rounds of structure calculation. This approach turned out to be especially helpful in loop regions suffering from the low number of unambiguous NOEs.

2.2.2 NMR spectroscopy (II): Chemical Shift Perturbation Analysis

To identify BRCT interaction sites of candidate binding partners, NMR titration experiments proved to be very useful, especially for small ligands such as peptides. The general idea behind NMR titration experiments goes along with the observation that the resonance frequencies of NMR signals (chemical shifts) are highly sensitive to changes in the electronic and magnetic environment of the coupling nuclei. Therefore, conformational changes upon ligand binding may cause ***chemical shift perturbations (CSPs)*** in the NMR spectra. Although ligand binding may also induce long-range conformational changes, nuclei involved in direct interactions with the binding partner usually experience more significant changes in the electronic environment than nuclei distant to the binding region. By that, NMR signals of nuclei very close to the interaction site show up with strong CSPs in the spectra.

Interaction Mapping with DNA and the DNA-binding Domain of p53

The BRCT tandem repeat of BRCA1 was reported to bind the key transcription factor p53, and to the ends of double-stranded DNA fragments [48,49]. More precisely, the second BRCT domain was identified as sufficient for the interaction with both the DNA-binding domain of p53 ($p53^{[94-292]}$) and DNA fragments. To locate the p53-BRCT interaction site(s), ^{15}N-labeled BRCT-c was used for NMR titration experiments with $p53^{[94-292]}$. Similarly, the DNA-binding region of BRCT-c was assessed by using different 5'-phosphorylated and un-phosphorylated double-stranded DNA oligomers for titration experiments with ^{15}N-labeled BRCT-c. The DNA double strands were prepared by slow-cool annealing of the self-complementary 5'-GCGTACGC-3', 5'-pGCCTAGCTAGGC-3' and 5'-GCCTAGCTAGGC-3' DNA oligomers. The physical interaction with ^{15}N-labeled BRCT-c (0.2 mM test tube concentration) was monitored by CSPs showing up in a series of 2D ^{1}H-^{15}N HSQC spectra upon step-wise addition of the putative binding partner up to a molar excess of 1.2:1.0.

Interaction Mapping with the First BRCA1 BRCT Domain (BRCT-n)

To analyze the possibility of long-range conformational cross-talk between the two BRCA1 BRCT domains, NMR titration experiments with ^{15}N-labeled BRCT-c and unlabeled BRCT-n ($BRCA1^{[1646-1744]}$) as intramolecular binding partner were envisaged. By that the double BRCT repeat was expected to become effectively reconstituted *in vitro*.

2.2.3 Analytical Ultracentrifugation

The method of analytical ultracentrifugation (UCF) is commonly used to assess the association behavior of macromolecular particles. The underlying principle is based on the correlation between the molecular size and the sedimentation velocity in a strong gravitational field. In simple terms, the higher the molecular weight the higher the sedimentation velocity during centrifugation. It is therefore possible to analyze the association properties of the interaction partners and to estimate values for the dissociation constant (K_D).

To assess the (self-)association behavior of BRCT domains, analytical UCF with the second BRCT domain of BRCA1 (BRCT-c) was performed. Protein samples were analyzed in potassium phosphate buffer (pH 6.8) containing 50 mM KCl, supplemented with 5 mM fresh DTT. Details of both sedimentation velocity (SV) and sedimentation equilibrium (SE) experiments are described in [43]. Analytical ultracentrifugation experiments were done in collaboration with Holger Strauss from the group of Hartmut Oschkinat (FMP, Berlin).

2.2.4 Macromolecular Crystallography

Structure determination by X-ray crystallography requires highly ordered single crystals of the molecular entity that is to be analyzed. Crystallization depends on the sample quality as well as the overall experimental setup, including temperature, pressure, the time course and the concentration of the chemical components involved. As a rule of thumb, the probability of crystallization is proportional to the sample quality (purity, stability and homogeneity) and inverse proportional to the length of flexible and/or unstructured regions and the number of components present in macromolecular assemblies. Finding suitable crystallization conditions therefore is by no means straightforward, but usually requires a screening approach comprising a large number of different buffer conditions.

General Aspects of Enzyme-Substrate Interactions

Structural characterization of enzyme-ligand complexes is challenging due the transient nature of the interactions involved. Under physiological conditions, substrate turnover is orders of magnitudes faster than the time scales resolved by the methods of structural biology. To get insights into the molecular architecture of enzyme-substrate (E·S) interactions, the E·S complex therefore needs to be stabilized by kinetic or thermodynamic trapping. This may be accomplished by using catalytically impaired enzymes or uncleavable, modified substrates. If the reaction involves two or more steps, specifically designed substrates may be used to trap covalent enzyme-intermediates. In this case, the modified substrates effectively function as potent inhibitors leading to inactive enzyme-inhibitor (E·I) complexes. In addition, enzyme-product (E·P) complexes may be accessible to structural studies although the enzyme-ligand affinity decreases substantially upon substrate processing. However, if the local product concentration is high and product displacement is kinetically disfavored, E·P complexes may become stable enough to perform structural studies.

Crystallization and Data Collection

To identify key residues of enzyme-ligand interactions in different glucosyl-binding subsites, the catalytically inactive variant H(A16-M)$^{E105Q/E109Q}$ was used for co-crystallization experiments with the native β-glucan substrate Glcβ4Glcβ4Glcβ3Glcβ4Glcβ4Glc (scissile β-1,4 glycosidic bond underlined). Molecular cloning and purification of H(A16-M)$^{E105Q/E109Q}$ was done by Kirill Piotukh (Humboldt-Universität zu Berlin). The hexasaccharide substrate was generously provided by Antoni Planas (Universidad Ramon Llull, Barcelona, Spain). In H(A16-M)$^{E105Q/E109Q}$ both the nucleophile (Glu105) and the general acid/base catalyst

(Glu109) have been mutated to the isosteric glutamine residues The hexasaccharide substrate was designed to be bound in subsites –IV to +II such that the scissile β-1,4 glycosidic bond becomes positioned between subsites –I and +I, thus putting the β-1,3 bond between subsites –II and –I.

For crystallization screening of substrate-bound H(A16-M)$^{E105Q/E109Q}$ the commercially available Crystal Screens (HAMPTON RESEARCH, CA, U.S.A.), each comprising 48 (Crystal Screen I: 50) different buffer conditions, were used. Crystallization experiments were set up at different temperatures (4°C and 20°C) and with different protein concentrations using the method of hanging-drop vapor-diffusion with equal volumes of protein/substrate (molar ratio of substrate to protein: 10:1) and buffer solution (1-2 µL each). To optimize promising crystal growth conditions the precipitant concentration and the protein-to-buffer volume ratio were varied. Crystals were mounted on a nylon loop and flash-frozen in liquid nitrogen to reduce radiation damage. X-ray diffraction data were collected on a MAR Research Imaging Plate detector at beamline BW7B of the EMBL outstation at DESY (Hamburg, Germany).

Structure Solution and Refinement

The crystallographic phase problem was solved by the method of molecular replacement, using the native-like hybrid enzyme H(A16-M) as a search model. The structure was improved by several rounds of positional and *B*-factor refinement with a step-wise increase of the resolution range (phase extension). Additional ions and water molecules were located by a $|F_o|$-$|F_c|$ difference density peak search. Maxima in the $|F_o|$-$|F_c|$ electron density map were accepted as water molecules if the peak was also present in the $2|F_o|$-$|F_c|$ map and the donor-acceptor distance to potential hydrogen bonding partners was in the range of 2.4 – 3.5 Å. In addition, water molecules were expected to have *B*-values below 60 Å^2. To avoid model bias the carbohydrate ligand was not accounted for in the early stage of refinement. Considering all protein amino acid residues and firmly bound calcium ions and water molecules in the structural model both the $2|F_o|$-$|F_c|$ and the $|F_o|$-$|F_c|$ difference map clearly showed the electron density of four covalently linked glucosyl units which were identified and modeled as the natural reaction product (Glcβ4Glcβ4Glcβ3Glc). Finally alternative protein side chain conformations were included. A 7% subset of all unique reflections, randomly chosen from thin shell sections over the entire resolution range, was excluded from structure calculation and refinement, and used for the calculation of R_{free} to avoid over-refinement and model bias.