Crystal structure of glycosylasparaginase from *Flavobacterium meningosepticum*

JIANCHENG XUAN,¹ ANTHONY L. TARENTINO,¹ BRIAN G. GRIMWOOD,¹
THOMAS H. PLUMMER JR.,¹ TAO CUI,² CHUDI GUAN,² AND PATRICK VAN ROEY^{1,3}

¹Division of Molecular Medicine, Wadsworth Center, New York State Department of Health, Albany, New York 12201

²New England Biolabs, Beverly, Massachusetts 01915

³Department of Biomedical Sciences, School of Public Health, University at Albany, Albany, New York 12201

(RECEIVED October 8, 1997; ACCEPTED November 28, 1997)

Abstract

The crystal structure of recombinant glycosylasparaginase from *Flavobacterium meningosepticum* has been determined at 2.32 Å resolution. This enzyme is a glycoamidase that cleaves the link between the asparagine and the N-acetylglucosamine of N-linked oligosaccharides and plays a major role in the degradation of glycoproteins. The three-dimensional structure of the bacterial enzyme is very similar to that of the human enzyme, although it lacks the four disulfide bridges found in the human enzyme. The main difference is the absence of a small random coil domain at the end of the α -chain that forms part of the substrate binding cleft and that has a role in the stabilization of the tetramer of the human enzyme. The bacterial glycosylasparaginase is observed as an $(\alpha\beta)_2$ -tetramer in the crystal, despite being a dimer in solution. The study of the structure of the bacterial enzyme allows further evaluation of the effects of disease-causing mutations in the human enzyme and confirms the suitability of the bacterial enzyme as a model for functional analysis.

Keywords: aspartylglucosaminidase; crystal structure; glycoamidase

Glycosylasparaginase (aspartylglucosaminidase, EC 3.5.1.26) (Mononen et al., 1993) is a glycoamidase that cleaves the β -N-glycosidic link between asparagine and N-acetylglucosamine of asparagine-linked oligosaccharides. The enzyme converts the asparagine to an aspartic acid and releases a 1-amino-glycan that spontaneously hydrolyzes to ammonia and the intact oligosaccharide. Glycosylasparaginases have been purified and cloned from bacteria (Tarentino & Plummer, 1993; Tarentino et al., 1995), insects (Liu et al., 1996), and a range of mammals (Tollersrud & Aronson, 1992). Eukaryotic glycosylasparaginases are lysosomal proteins that play a major role in the degradation of glycoproteins (Aronson & Kuranda, 1989). Deficiency of human glycosylasparaginase leads to aspartylglycosaminuria, an inherited disease with high incidence in the Finnish population (Haila et al., 1993; Mononen et al., 1993; Mononen & Aronson, 1997).

Flavobacterium meningosepticum is the only prokaryotic organism known to express two glycoamidases that remove N-linked glycans from glycoproteins: glycosylasparaginase (Tarentino & Plummer, 1993) and peptide-N⁴-(N-acetyl- β -D-glucosaminyl)-asparagine amidase F (PNGase F) (Tarentino & Plummer, 1994). Interestingly, no sequence, structural, or mechanistic similarities

exist between PNGase F (Kuhn et al., 1994, 1995) and the glycosylasparaginases.

All glycosylasparaginases consist of an $\alpha\beta$ -heterodimer that is derived from an inactive single-chain precursor by posttranslational processing involving an autoproteolytic reaction (Guan et al., 1996). They belong to the family of N-terminal nucleophile amidohydrolases (Aronson, 1996), a group of widely diverse enzymes in which a threonine, serine, or cysteine at the amino terminal of the β -chain functions as a combined base-nucleophile catalyst (Brannigan et al., 1995). The mechanism of action of the glycosylasparaginases resembles that of the serine proteases, but with the γ -hydroxyl group of the N-terminal threonine, residue 152 and 183 in the bacterial and human enzymes, respectively, acting as the nucleophile and its amino group as the base (Tikkanen et al., 1996).

The crystal structure of human glycosylasparaginase was reported recently (Oinonen et al., 1995). The human and *F. meningosepticum* proteins are highly homologous (Tarentino et al., 1995), with about 30% sequence identity or about 60% conserved residues. The main difference from all eukaryotic glycosylasparaginases, is a 31-residue deletion at the carboxy terminus of the α -subunit of the bacterial enzyme. One other noteworthy distinction is that the human enzyme includes nine cysteines, of which eight form disulfide bonds, whereas the bacterial enzyme has five cysteines, but only one of these, Cys 232, is conserved. This difference is significant because two of the disulfides bonds in the

Reprint requests to: Patrick Van Roey, Wadsworth Center, New York State Department of Health, Empire State Plaza, P.O. Box 509, Albany, New York 12201-0509; e-mail: vanroey@wadsworth.org.

Fig. 1. Ramachandran plot of *F. meningosepticum* glycosylasparaginase produced with PROCHECK (Laskowski et al., 1993). Only residues Trp 11, in either dimer, and Ala 293 of the first dimer fall on the borderline between the generously allowed and disallowed regions. Trp 11 is conserved and also adopts a similar, less probable conformation in the structure of the human enzyme.

human enzyme, Cys 140 to Cys 156 in the α -subunit and Cys 294 to Cys 322 in the β -subunit, have been reported to be essential not only for activity, but for proper folding and activation of the enzyme in the endoplasmic reticulum (Riikonen et al., 1996).

Expression of recombinant human glycosylasparaginase in quantities suitable for functional analysis has proven to be very difficult (Mononen et al., 1995). Therefore, experiments aimed at understanding the autocatalytic activation and effects of mutations on the enzyme mechanism of action are performed using the bacterial enzyme as a model (Guan et al., 1996). We describe the crystal structure of *F. meningosepticum* glycosylasparaginase and compare its structure with that of the human enzyme.

Table 1. Final refinement statistics

<i>R</i> -value	0.188
<i>R</i> _{free}	0.270
No. of atoms	4,501
Protein	4,301
Water	200
Deviation from ideality	
Bond length (Å)	0.003
Bond angle (°)	0.895
Dihedral (°)	25.28
Average <i>B</i> -factor (Å ²)	26.5
Main-chain atoms	25.4
Side-chain atoms	27.3
Solvent	31.4

Results and discussion

F. meningosepticum glycosylasparaginase crystallizes from PEG 3350 in Tris·HCl, pH 8.5, buffer as thin plates belonging to space group P2₁. The structure was determined by a combination of molecular replacement and multiple isomorphous replacement methods. The structure of an $\alpha\beta$ -dimer of human glycosylasparaginase was used as the search model for the molecular replacement calculations. The structure was refined to an *R*-value of 0.188 and an *R*_{free} of 0.270 at 2.32 Å resolution. The asymmetric unit contains one ($\alpha\beta$)₂-tetramer. The final model includes 565 of the total 590 amino acids, residues 2–139 and 152–294 for the α and β chains of the first dimer and residues 2–141 and 152–295 for the second dimer, and 200 water molecules. Of the non-glycine and non-proline residues, 90.4% are in the most favorable region of the Ramachandran plot. Three residues, Trp 11, in both dimers, and Ala 293 in the first dimer, are located on the borderline between the generously allowed and disallowed regions. Table 1 summarizes the crystallographic refinement data. Figures 1 and 2 show the Ramachandran plot and the final ($2F_o - F_c$) electron density map in the area of Trp 11 of the first dimer, respectively. The atomic coordinates and structure factors have been submitted to the Protein Data Bank, Brookhaven National Laboratory (entry number 1ayy).

Fig. 2. Part of the ($2F_o - F_c$) map in the area of residues 10–14 of the first dimer. Map contoured at the 1.25 σ level. Figures 2, 3A, 4, 5, and 7 are prepared using SETOR (Evans, 1993).

Fig. 3. **A:** α -Carbon drawing of the structure of *F. meningosepticum* glycosylasparaginase. The α -chain is drawn in finer line than the β -chain. The amino-terminal and carboxy-terminal residues as well as every 20th residue are labeled. **B:** MOLSCRIPT (Kraulis, 1991) rendering of the three-dimensional structure of glycosylasparaginase.

The overall fold of *F. meningosepticum* glycosylasparaginase, shown in Figure 3, is very similar to that of the human enzyme: an $\alpha\beta\alpha$ -sandwich in which the two subunits are highly integrated. The core of the molecule consists of a β -sandwich containing one eight-stranded β -sheet and one four-stranded β -sheet. These sheets are flanked by five α -helices on the side of the larger sheet and three α -helices on the side of the smaller sheet. The eight-stranded β -sheet is composed of four strands from the α -subunit and four strands from the β -subunit. All strands in the second sheets are from the β -subunit. The five helices adjacent to the larger β -sheet belong to the α -subunit and those on the opposite side belong to the β -subunit.

Figure 4 shows the superposition of a dimer of the human enzyme on a dimer of the bacterial enzyme, produced by least-squares fitting of the main-chain atoms of corresponding residues in all the secondary structure elements. The RMS, average, and maximum displacements for the main-chain atoms of the 152 residues included in the fit are 0.77 Å, 0.66 Å, and 2.75 Å, respectively. These displacements are very similar to those obtained for the corresponding fit between the two dimers of the bacterial enzyme, 0.80 Å, 0.61 Å, and 2.30 Å, respectively, demonstrating the high degree of similarity between the structures of the human and bacterial proteins. The main difference in the sequences is a 31-amino acid deletion at the carboxy terminus of the α -subunit of the bacterial enzyme. This α -subunit carboxy terminus is disordered in both structures. In the bacterial enzyme, 12 residues could not be

placed for the carboxy terminus of the α -chain of the first dimer and electron density is lacking for 10 residues at the α -chain carboxy terminus for the second dimer. In the human enzyme, about 20 residues could not be modeled at the carboxy termini of either α -chain. Because it has now been well established that activation involves autoproteolytic cleavage of the link between residues 151 and 152 (bacterial sequence numbering) of the precursor, it is clear that this disordered section is released upon processing of the precursor while the amino terminus of the β -chain remains tightly bound. In the bacterial enzyme, the ordered structure ends at the carboxy terminus of helix A5, residue 139 in the first dimer and residue 141 in the second. In the human enzyme structure, an additional 21–23 residues, corresponding to about two thirds of the 31-residue insertion, are ordered and fold as an independent random coil domain. This domain is located at one end of the cleft that contains the active site and that, in the precursor form, would be occupied by the residues in the disordered α -chain terminus. The effect of the addition of the domain is an increase in the length of the cleft and an increase in the distance between the last ordered residue in the α -subunit and the amino terminus of the β -subunit. The distance between the α -carbon atoms of residues 162 and 183 in the human enzyme is 34.5 Å, compared with a distance of 23.3 Å between residues 139 and 152 of the bacterial enzyme. To compensate for the longer cleft, the human enzyme has a longer disordered carboxy terminus. The addition of the domain and the lengthening of the substrate binding cleft could possibly have an

Fig. 4. Overlap of the structure of human glycosylasparaginase (fine line) on the structure of the bacterial enzyme. Note the high similarity of the structures except for the random coil domain extension at the carboxy terminal of the α -chain, upper left, and the significant differences in the conformations of the loops connecting the smaller β -sheet and the α -helices, upper right. The latter loops surround the active site area, identified from the location of the amino terminus of the β -chain in the center of the molecule.

effect on substrate specificity as well, but no extensive comparisons of the substrate requirements have been reported. This random coil domain in the human enzyme also appears to be involved in the stabilization of the formation of the tetramer, as discussed below.

Other differences between the human and bacterial α -subunits are limited to small changes in the conformations of two loops, those connecting α -helix A2 to β -strand A2, and β -strand A2 to β -strand A3. Differences are more significant in three loops in the β -subunit. An insertion in the area of residues 242–250 of the bacterial enzyme lengthens α -helix B3, β -strand B5 and the loop connecting them. On the other hand, the turns between β -strand B6 and β -strand B7 and between β -strand B7 and β -strand B8 are shorter. Overall, these differences contribute to a more open substrate binding cleft in the bacterial enzyme compared with that of the human enzyme.

Eukaryotic glycosylasparaginases, including the human, chicken, and insect proteins, are usually observed in solution, and on sizing columns, in an $(\alpha\beta)_2$ -tetrameric form. In contrast, only a single peak, corresponding to the molecular weight of an $\alpha\beta$ -dimer, is observed on sizing gels and columns for the *F. meningosepticum* enzyme. Regardless, an $(\alpha\beta)_2$ -tetramer, very similar to that of the

human enzyme, is observed in the crystal structure, as shown in Figure 5. The two dimers are related by nearly perfect twofold symmetry. The angle between the two dimers is slightly different from that in the human enzyme so that when one dimer of the human enzyme is superimposed onto a dimer of the bacterial enzyme, the deviation between equivalent main-chain atoms of the most distant part of the other pair of dimers is about 2.5 Å.

The lower stability of the tetramer of bacterial glycosylasparaginase compared with the human enzyme can be attributed to the absence of the random coil domain at the carboxy terminus of the α -subunit. Although all other interactions between the dimers are conserved, this random coil domain makes extensive dimer–dimer contacts, primarily of hydrophobic nature. Specifically, residues 145–147 and 159–160 have interactions with loops 98–99 and 115–116 of the α -chain of the other dimer. Surface accessibility calculations, with a sphere with a radius of 1.4 Å as the probe, show that dimer–dimer interactions reduce the accessible surface for a dimer of the bacterial enzyme by only 1,905 Å², compared with 2,485 Å² for the human enzyme. In fact, the contact area of the random coil domain with the second dimer is as large as its contact area with the remainder of the same dimer. The surface areas of the random coil domain of the first dimer that are rendered

Fig. 5. Stereo diagram showing the tetramer of glycosylasparaginase. The first dimer is drawn in heavier line than the second dimer. The α - and β -subunits are shown in solid and dashed lines, respectively.

inaccessible through interaction with the rest of the same dimer and with the other dimer are $386 \text{ \AA}^2$ and $415 \text{ \AA}^2$, respectively. Corresponding inaccessible areas in the second dimer are $407 \text{ \AA}^2$ and $397 \text{ \AA}^2$. This smaller dimer–dimer contact area, and consequently weaker dimer–dimer interaction, may well explain why only a dimer is observed on sizing gels for the bacterial protein. However, crystallization as a tetramer still leads one to assume the presence of the tetramer in solution, at least at high concentrations. These observations are significant in relationship to the evidence presented by Riikonen et al. (1996), which indicates that dimerization of the precursor of the human enzyme occurs before, and is essential to, proteolytic activation.

Bacterial glycosylasparaginase contains five cysteine residues, but no disulfide bonds. This is in contrast to the human enzyme, which has nine cysteines and four disulfide bridges, two internal in each domain. Furthermore, it has been clearly demonstrated that all four disulfide bonds are important for stability of the human enzyme and two of them are essential for correct folding of the protein and for activation in the endoplasmic reticulum (Riikonen et al., 1996). Figure 6 shows the locations of each of the five cysteines in the bacterial enzyme. Residues 68 in the α -subunit and 168 in the β -subunit are located on adjacent β -strands in the center of the eight-stranded β -sheet, strands A2 and B2, and are sufficiently close to be able to form a disulfide bridge. The sulfhydryl group of Cys 68 is oriented toward Cys 168, but that of Cys 168 points away from Cys 68. The Cys 68 side chain is involved in a weak, $3.1\text{-\AA}$ hydrogen bond with Ser 184. The sulfhydryl of Cys 168 is not involved in any close contacts, but is oriented toward the substrate binding cleft. SDS-PAGE gels of the wild-type and recombinant proteins, in the absence of reducing agents, show free α and β subunits, confirming the absence of covalent linkage between the two subunits and demonstrating that the absence of the disulfide bond is not an artifact of the expression in *Escherichia coli*. Also, the fact that one of the mercury atoms in the mercury derivative binds adjacent to Cys 168 further confirms that the sulfhydryl groups are in the reduced form. Tarentino et al. (1995), in their manuscript on the cloning of *F. meningosepticum* glycosylasparaginase, reported mass spectrometry data, based on

the molecular weights of CNBr fragments, that indicated the presence of two intradomain disulfide bridges, one between residues 68 and 81 in the α -subunit and one between residues 168 and either 232 or 270 in the β -subunit. However, the crystal structure clearly shows that residues 68 and 81 are indeed on adjacent β -strands, but several residues out of register and thus not within suitable distance to form disulfide bridges. Residue 168 is close to 68, but not close to either 232 or 270. On the other hand, 232 and 270 are in reasonably close proximity, 232 on α -helix B3 and 270 on β -strand B6, but again not sufficiently close to be able to form disulfide bonds. The mass spectrometry assignments have since been reevaluated (C.R. Hauer, pers. comm.).

The active site, identified from the presence of Thr 152, is located in a funnel-shaped well in the cleft discussed above. The funnel is formed by loops that connect the two β -sheets and by loops connecting the smaller β -sheet and α -helices associated with it. Oinonen et al. (1995) published the structure of human glycosylasparaginase in free form and as a complex with aspartic acid. All residues that bind aspartic acid in the human enzyme are contributed by two loops, β -strand B2 to β -strand B3, and β -strand B4 to α -helix B1. They are conserved both in sequence and in geometry between the two structures. However, as can be seen in Figures 4 and 7, overall, the funnel is somewhat wider in the bacterial enzyme than in the human enzyme. This is primarily caused by the changes in the conformations of the loops between α -helix A2 and β -strand A2 and between β -strand B6 and β -strand B7, which are among the loops that differ most between the two enzymes, and by differences in some residues at the edge of the cleft, most notably the substitution of Gln 254 in the bacterial enzyme for Phe 278 in the human enzyme. The latter substitution is significant because Oinonen et al. (1995) suggested that Trp 11 and Phe 278 may be important for oligosaccharide binding. Although Trp 11 is conserved, the substitution of Gln for Phe 278 does not preclude this observation because both aromatic and polar–neutral residues are equally suitable for oligosaccharide binding, although with possible differences in specificities.

Oinonen et al. (1995) interpreted the structure of human glycosylasparaginase in relationship to disease-causing mutations. The

Fig. 6. Stereo diagram, prepared with MOLSCRIPT (Kraulis, 1991), showing the α -carbon tracing of an $\alpha\beta$ -dimer of glycosylasparaginase with the side chains of the cysteine residues shown in ball-and-stick mode.

Fig. 7. Stereo diagram showing the overlap of the active site areas of bacterial (bold) and human (fine) glycosylasparaginase. The main-chain atoms of all residues within close proximity of the catalytic residue Thr 152 are shown. Side chains have been included when they point into the binding cleft. All residues, except Gln 254 (Phe 278 in human), upper right, are conserved. Differences in the conformations of the loops connecting helix A2 to strand A2, lower left, and strand B6 to strand B7, right, lead to a more open active site area in the bacterial enzyme.

current study allows for additional analysis of the location and properties of the corresponding residues. Table 2 lists the most important single-site mutations that lead to aspartylglycosaminuria, and the corresponding residues in the bacterial enzyme. The disulfide bonds Cys 140–Cys 156 and Cys 263–Cys 283 of the human enzyme have been reported to be essential for activity as well as for folding and activation of the enzyme (Riikonen et al., 1996), and mutants of Cys 140 and Cys 283 are inactive. Although none of these cysteine residues are conserved in the bacterial enzyme, its molecular fold is nearly identical. Cys 140 and Cys 156 are located in the small random coil domain at the carboxy-terminal end of the α -chain of the human enzyme. Because the domain is absent in the bacterial enzyme, it remains possible that this disulfide bridge plays a role in folding that domain and in the stabilization of the tetramer. The situation for Cys 283 is more complicated. This residue is not conserved in the bacterial enzyme, but Cys 263 is conserved. Therefore, misfolding in the absence of this disulfide bond may not be a suitable explanation for the inactivity of the Cys 283 Arg mutant. The mutation Ala 78 Val is thought to cause difficulties in the packing. This residue is located in a rather large hydrophobic pocket that also contains a proline and a leucine side chain. The bacterial enzyme has a serine in position 78. Its side chain occupies an identically shaped pocket that also contains a conserved proline and a methionine side chain. Although it would seem that the pocket is occupied by larger side chains in the bacterial enzyme, suggesting some degree of adaptability, the bacterial enzyme is not a suitable model for analysis of this mutation. The residues at all other sites of disease-related mutations are either conserved or on the exterior surface of the

molecule and therefore consistent with the analysis of the human enzyme.

Materials and methods

Sample preparation

Samples of recombinant *F. meningosepticum* glycosylasparaginase used for crystallization experiments were prepared at New England Biolabs and at the Wadsworth Center. Protein samples prepared at New England Biolabs were expressed in *E. coli* as pMAL fusion proteins and purified according to published procedures (Guan et al., 1996). Samples prepared at the Wadsworth Center were made using a slightly modified procedure. In summary, the protein was expressed as a fusion protein (MBP-glycosylasparaginase) in the periplasm of *E. coli* TB-1 cells. The fusion protein was extracted from the periplasm in high yield using the bacterial permeabilization reagent Polymyxin B sulfate (MacLeod & Gyles, 1990; Schupp et al., 1995). The fusion protein was purified by chromatography on Q-sepharose, digested with Factor Xa, and rechromatographed on Q-sepharose.

Crystallization

Crystals were grown by hanging drop vapor diffusion methods at 10 °C using drops containing equal volumes of protein solution (5.8 mg/mL in 10 mM Tris·HCl, pH 7.5, 5 mM NaCl, 1 mM EDTA) and crystallization buffer, 19% PEG 3350 in 100 mM

Table 2. Single-residue mutations in human glycosylasparaginase, their location, and assumed effect compared with the corresponding bacterial residue

Mutation	Location	Effect	Bacterial	Comparison
Gly 37 Asp	Helix A2, interior	Disturbed packing	Gly 38	Conserved—consistent
Ala 78 Val	Strand A3, interior	Disturbed packing	Ser 78	Similar pocket—inconclusive
Arg 138 Gln	Helix A5, exterior	Unknown	Thr 138	Consistent
Cys 140 Ser	Random coil domain	Disulfide bond/folding	None	Role in tetramer stability
Gly 279 Arg	Strand B5, interior	Disturbed packing	Val 255	Consistent
Cys 283 Arg	Strand B5, interior	Disturbed packing	Ala 259	Tolerated

Table 3. Data collection statistics for native and derivative data

	Native	Native high resolution	Me ₃ PbAc	MeHgCl
Resolution (Å)	20.0–2.32	2.4–2.32	20.0–2.6	20.0–3.0
Number of unique data	22,627	2,190	15,272	10,631
Redundancy	2.90	2.73	2.70	3.35
Completeness (%)	95.7	94.7	90.7	95.6
$\langle I/\sigma(I) \rangle$	7.0	3.0	7.0	6.3
$I/\sigma(I) > 1$ (%)	85.4	72.2	81.5	87.3
R_{sym}	0.075	0.170	0.075	0.083
R_{mer} with native	—	—	0.122	0.139

Tris·HCl, pH 8.5. Plate-shaped crystals grow to a size of about $0.05 \times 0.1 \times 0.15$ mm in one to two weeks. Larger crystals, up to 0.3 mm in two dimensions, can be grown, but are twinned. The crystals belong to the monoclinic space group $P2_1$ with cell dimensions of $a = 46.2$ Å, $b = 115.6$ Å, $c = 52.4$ Å, $\beta = 107.2^\circ$. With two $\alpha\beta$ -dimers ($M_r = 32,135$) in the asymmetric unit, the $V_M = 2.1$ Å³/Da, corresponding to a solvent content of about 41% (Matthews, 1968). Heavy-atom derivatives were prepared by transferring crystals to wells containing 100–300 µL of the crystallization buffer enriched with heavy-atom compounds.

Data collection and processing

X-ray diffraction data for native and derivative crystals were measured at approximately 130 K (nozzle temperature 123 K) using a Rigaku R-axis IIC detector mounted on a Rigaku RU-200 rotating anode equipped with focusing mirrors (Molecular Structure Corporation) and liquid-nitrogen low-temperature device (Molecular Structure Corporation). The crystals were soaked briefly in the mother liquor enriched with 20% PEG 400 prior to flashcooling in the cold nitrogen gas stream. Table 3 lists the data collection statistics for the native and two derivative data sets. Diffraction data were processed with DENZO and scaled using SCALEPACK (Otwinowski & Minor, 1997). Native data are 95.7% complete to 2.32 Å resolution.

Structure determination and refinement

Initially, the PHASES package (Furey & Swaminathan, 1990) was used in attempts to determine the structure by multiple isomorphous replacement methods. The isomorphous difference Patterson maps for two derivatives, trimethyllead acetate (20 mM, 6 days) and methylmercury chloride (2 mM, 12 days), revealed two and three sites, respectively. However, the phasing power of these derivatives was inadequate to yield an interpretable map, presumably because of the relative weakness of the intensity data, owing to the small crystal size.

The structure was then determined by molecular replacement methods, using the program AMoRe (Navaza, 1994) as implemented in the CCP4 package (CCP4, 1994). An $\alpha\beta$ -dimer of human glycosylasparaginase (Protein Data Bank entry 1APY) was used as the search model. Two solutions for the rotation function were found consistently to have correlation coefficients that were

1.3–1.7 times background. Independently, these top two rotation function solutions yielded solutions for the translation function (x and z coordinates) that were significant above the noise level and that had correlation coefficients and R -values of 0.137 and 0.530 and of 0.104 and 0.532, respectively. The validity of the solution for the first molecule was confirmed by a cross difference Fourier calculation with the lead derivative, which showed positions for the lead atoms consistent with the solution of the isomorphous difference Patterson. Crystal packing analysis was used to determine the relative y -coordinate of the second molecule. With the first molecule held fixed, the second molecule was translated along y in fractional steps of 0.05. Only one position had no intermolecular contacts between α -carbon atoms of less than 1.0 Å.

After the positions of both molecules were determined, phase combination, including the contribution of the phasing information of the two heavy-atom derivatives, was used to improve the quality of the map and to reduce model bias. This map suggested significant differences in the structure of the second half of the β -chain. A partial model, consisting of complete α -subunits, the N-terminal half of the β -subunits, and polyaniline chain for the remaining residues, was subjected to simulated-annealing refinement, using X-PLOR (Brünger et al., 1987), reducing the R -value from 0.48 to 0.27 for the data between 8 and 2.5 Å resolution and with a $2\sigma(F)$ cut-off. One additional round of phase combination was required to allow correction of several loop segments in the β -subunit. Isolated peaks in the difference and $2F_o - F_c$ maps were included as water molecules if they were within hydrogen bonding distance from potential hydrogen bonding groups, using the waterpick option in X-PLOR with a cut-off of 3.8 Å, but were eliminated if the thermal parameter, upon further refinement, exceeded 55 Å². Further refinement consisted of alternating cycles of positional and thermal parameters using all diffraction data with $F > 0.0$ and between 10.0 and 2.32 Å resolution. The model building was performed using the graphics program CHAIN (Sack, 1988). Additional programs, such as the program SURFACE, used for surface accessibility calculations, are implemented in the program package CCP4 (1994).

Acknowledgments

This research was supported in part by USPHS grants GM-50431 (P.V.R.) and GM-34071 (T.H.P., A.L.T.) from the National Institute of General Medical Sciences, National Institutes of Health, and used the Macromolecular Crystallography Core Facilities of the Wadsworth Center.

References

- Aronson NN Jr. 1996. Lysosomal glycosylasparaginase: A member of a family of amidases that employ a processed N-terminal threonine, serine or cysteine as a combined base-nucleophile catalyst. *Glycobiology* 6:669–675.
- Aronson NN Jr, Kuranda MJ. 1989. Lysosomal degradation of Asn-linked glycoproteins. *FASEB J* 3:2615–2622.
- Brannigan JA, Dodson G, Duggleby HJ, Moody PCE, Smith JL, Tomchick DR, Murzin AG. 1995. A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* 378:416–419.
- Brünger AT, Kuriyan J, Karplus M. 1987. Crystallographic R-value refinement by molecular dynamics. *Science* 235:458–460.
- CCP4. 1994. Collaborative Computing Project No. 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
- Evans SV. 1993. SETOR: Hardware-lighted three-dimensional solid model representation of macromolecules. *J Mol Graphics* 11:134–138.
- Furey W, Swaminathan S. 1990. "PHASES"—A program package for the processing and analysis of diffraction data for macromolecules. *American Crystallographic Association Meeting Abstracts*. PA33.
- Guan C, Cui T, Rao V, Liao W, Benner J, Lin CL, Comb D. 1996. Activation of glycosylasparaginase. Formation of active N-terminal threonine by intramolecular autoprolysis. *J Biol Chem* 271:1732–1737.
- Halila R, Ikonen E, Töllersrud O, Syvanen AC, Enomaa N, Peltonen L. 1993. Aspartylglucosaminuria (AGU): Protein and gene structure of normal and mutated aspartylglucosaminidase. *Biochem Med Metabol Biol* 50:1–8.
- Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946–950.
- Kuhn P, Tarentino AL, Plummer TH Jr, Van Roey P. 1994. Crystal structure of peptide-N⁴-(N-acetyl- β -D-glucosaminyl)asparagine amidase F at 2.2 Å resolution. *Biochemistry* 33:11699–11706.
- Kuhn P, Guan C, Cui T, Tarentino AL, Plummer TH Jr, Van Roey P. 1995. Active site and oligosaccharide recognition residues of peptide-N⁴-(N-acetyl- β -D-glucosaminyl)asparagine amidase F. *J Biol Chem* 270:29493–29497.
- Laskowski RA, McArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:282–291.
- Liu Y, Dunn GS, Aronson NN Jr. 1996. Purification, biochemistry and molecular cloning of an insect glycosylasparaginase from *Spodoptera frugiperda*. *Glycobiology* 6:527–536.
- MacLeod DL, Gyles DL. 1990. Purification and characterization of an *Escherichia coli* shiga-like toxin II variant. *Infect Immun* 58:1232–1239.
- Matthews BW. 1968. Solvent content of protein crystals. *J Mol Biol* 24:54–68.
- Mononen I, Aronson NN Jr. 1997. *Lysosomal storage disease: Aspartylglycosaminuria*. Austin, Texas: Landes Bioscience.
- Mononen I, Fisher KJ, Kaartinen V, Aronson NN Jr. 1993. Aspartylglycosaminuria: Protein chemistry and molecular biology of the most common lysosomal storage disorder of glycoprotein degradation. *FASEB J* 7:1247–1256.
- Mononen I, Heisterkamp N, Dunder U, Rompanen EL, Noronkoski T, Kuronen I, Groffen J. 1995. Recombinant glycosylasparaginase and in vitro correction of aspartylglycosaminuria. *FASEB J* 9:428–433.
- Navaza J. 1994. AMoRe: An automated package for molecular replacement. *Acta Crystallogr A* 50:157–163.
- Oinonen C, Tikkanen R, Rouvinen J, Peltonen L. 1995. Three-dimensional structure of human lysosomal aspartylglucosaminidase. *Nature Struct Biol* 2:1102–1108.
- Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
- Riikonen A, Rouvinen J, Tikkanen R, Julkunen I, Peltonen L, Jalanko A. 1996. Primary folding of aspartylglucosaminidase. Significance of disulfide bridges and evidence of early multimerization. *J Biol Chem* 271:21340–21344.
- Sack JS. 1988. CHAIN—A crystallographic modeling program. *J Mol Graphics* 6:244–245.
- Schupp JM, Travis SE, Price LB, Shand RF, Keim P. 1995. Rapid bacterial permeabilization reagent useful for enzyme assays. *Biotechniques* 19:18–20.
- Tarentino AL, Plummer TH Jr. 1993. The first demonstration of a procaryotic glycosylasparaginase. *Biochem Biophys Res Commun* 197:179–186.
- Tarentino AL, Plummer TH Jr. 1994. Enzymatic deglycosylation of asparagine-linked glycans: Purification, properties and specificity of oligosaccharide-cleaving enzymes from *Flavobacterium meningosepticum*. *Methods Enzymol* 230:44–57.
- Tarentino AL, Quinones G, Hauer CR, Changchien LM, Plummer TH Jr. 1995. Molecular cloning and sequence analysis of *Flavobacterium meningosepticum* glycosylasparaginase: A single gene encodes the alpha and beta subunits. *Arch Biochem Biophys* 316:399–406.
- Tikkanen R, Riikonen A, Oinonen C, Rouvinen J, Peltonen L. 1996. Functional analyses of active site residues of human lysosomal aspartylglucosaminidase: Implications for catalytic mechanism and autocatalytic activation. *EMBO J* 15:2954–2960.
- Töllersrud OK, Aronson NN Jr. 1992. Comparison of liver glycosylasparaginase from six vertebrates. *Biochem J* 282:891–897.