THE PROTEIN SOCIETY

The T99K variant of glycosylasparaginase shows a new structural mechanism of the genetic disease aspartylglucosaminuria

Suchita Pande and Hwai-Chen Guo*

Department of Biological Sciences, University of Massachusetts Lowell, Lowell, Massachusetts, 01854

Received 24 February 2019; Accepted 21 March 2019
DOI: 10.1002/pro.3607
Published online 9 April 2019 proteinscience.org

Abstract: **Aspartylglucosaminuria (AGU) is an inherited disease caused by mutations in a lysosomal amidase called aspartylglucosaminidase (AGA) or glycosylasparaginase (GA). This disorder results in an accumulation of glycoasparagines in the lysosomes of virtually all cell types, with severe clinical symptoms affecting the central nervous system, skeletal abnormalities, and connective tissue lesions. GA is synthesized as a single-chain precursor that requires an intramolecular autoprocessing to form a mature amidase. Previously, we showed that a Canadian AGU mutation disrupts this obligatory intramolecular autoprocessing with the enzyme trapped as an inactive precursor. Here, we report biochemical and structural characterization of a model enzyme corresponding to a new American AGU allele, the T99K variant. Unlike other variants with known 3D structures, this T99K model enzyme still has autoprocessing capacity to generate a mature form. However, its amidase activity to digest glycoasparagines remains low, consistent with its association with AGU. We have determined a 1.5-Å-resolution structure of this new AGU model enzyme and built an enzyme–substrate complex to provide a structural basis to analyze the negative effects of the T99K point mutation on K_M and k_{cat} of the amidase. It appears that a "molecular clamp" capable of fixing local disorders at the dimer interface might be able to rescue the deficiency of this new AGU variant.**

Keywords: autoprocessing; autoproteolysis; glycoasparagine; lysosomal amidase; lysosomal disease

Introduction

Aspartylglucosaminuria (AGU) is a rare disease caused by the failure of lysosomes to process the protein-to-carbohydrate linkage of Asn-linked glycoproteins.[1–3] AGU affects multiple systems of the body with severe clinical symptoms including poor nerve response, connective tissue lesions, coarse facial features, vertebral and long bone changes, and osteoporosis.[3,4] It had been mapped to mutations in the gene encoding for a lysosomal enzyme called glycosylasparaginase

Abbreviations: AGA, aspartylglucosaminidase; AGU, aspartylglucosaminuria; GA, glycosylasparaginase; NAcGlc, *N*-acetylglucosamine; NAcGlc-Asn, N^4-(β-*N*-acetylglucosaminyl)-L-asparagine; rmsd, root mean square deviation/displacement.

Enzyme: glycosylasparaginase (GA) or aspartylglucosaminidase (AGA), EC # **3.5.1.26**.

Statement: This study describes structure-–function studies of a new mutant enzyme that causes a genetic disease called aspartylglucosaminuria (AGU). We report a crystal structure at 1.5 Å resolution as well as kinetics characterization of this new AGU variant. The results reveal a new disease-casusing mechanism: a local disorder at the substrate-binding site propagated through the enzyme's dimer interface. They also suggest that a "molecular clamp" for stabilizing the dimer could be a potential therapeutic.

Grant sponsor: NIH GM128152 DK075294.

**Correspondence to:* Hwai-Chen Guo, Department of Biological Sciences, University of Massachusetts Lowell, 1 University Avenue, Lowell, MA 01854. E-mail: hwaichen_guo@uml.edu

(GA) or aspartylglucosaminidase (AGA)[3] [EC # 3.5.1.26]. About 30 different AGU alleles have been reported around the world,[5–7] including independent missense mutations, frameshifts, and polypeptide truncations/extensions. Due to a founder effect, AGU incidence is concentrated in Finland, with one major allele (denoted AGU_{FIN}) found in 98% of the AGU patients. The AGU_{FIN} allele carries two concurrent substitutions R161Q + C163S, in which the C163S substitution is the causative mutation for AGU deficiency.[8] In addition to preventing the disulfide bond formation between C163 and C179, the C163S substitution also causes destabilization of a loop structure unique in the protein precursor and thus prevents a precise dimerization essential for GA autoproteolytic activation[9] (see below). A separate Finnish allele has a point mutation of nucleotide C to T that changes residue 234 from a threonine to an isoleucine, called either T234I or T257I AGU variant in the literature[6]; these numbering differences are due to omitting or including the signal peptide (Table I). In North America, different AGU mutations have also been reported. For example, a Canadian AGU allele carries a substitution of glycine by an aspartic acid at residue 203 (denoted G203D variant, Table I). Recently, a new AGU variant was discovered in two male siblings in the United States, wherein, the paternal allele carried a single C to A base exchange at position 365 of the GA coding region.[10] This new American AGU mutation results in substitution of Thr99 by a charged residue Lys (named T99K variant, Table I).

GA is involved in protein degradation by catabolizing Asn-linked glycoproteins in lysosomes. It is widely distributed in various tissues and many species of eukaryotes and prokaryotes.[11–13] GAs are conserved in protein sequences and 3D structures and share the same activation mechanism to generate a mature amidase.[14–18] This enzyme is initially synthesized as an inactive single-chain precursor in which α and β subunits are joined together via a surface loop (called precursor- or P-loop) that prevents the protein from accommodating glycoasparagine substrates.[16] An obligatory processing step is thus required to cleave off the P-loop through intramolecular autoproteolysis.[19] This autoprocessing event results in the mature and active form of the amidase with separate α and β subunits to open up the substrate site for glycoasparagines. The autoprocessing also generates the N-terminal amino group at the β subunit to serve as the base of the mature amidase.[16] During the metabolic turnover of Asn-linked glycoproteins, autocleaved GA hydrolyzes glycoasparagine N^4-(β-*N*-acetylglucosaminyl)-L-asparagine (NAcGlc-Asn) that connects a carbohydrate to the side chain of an asparagine in glycoproteins.[5]

Although autoproteolysis occurs intramolecularly, a precise dimerization of GA appeared to be a prerequisite to trigger autoproteolytic activation.[15] Through various biochemical and biophysical approaches, including size-exclusion chromatography, cross-linking experiments, and analytical ultracentrifugation (AUC) studies,[9] we demonstrated that GA forms dimers in solution. Furthermore, GA crystal structures in various crystal forms and/or from different mutants or AGU variants all form the same dimer structures.[14–16,20] Several structure-based mutations designed to disrupt the dimer interface resulted in proteins forming high molecular weight aggregates,[9] with concurrent impairment of autoproteolytic activation, and subsequent lack of amidase activity. However, this essential dimerization is not sufficient to trigger the autoproteolytic activation of GA because subtle structural changes can alter the dimerization status and thus prevent autoproteolytic activation. Data suggest that many AGU mutations remain as dimers but cannot undergo autoproteolysis and thus lack amidase activity for digesting glycoasparagines.[5,21–24] These results suggest that the precise monomer–monomer interactions of GA play essential roles in GA autoprocessing and autoactivation.

Although AGU and other GA mutations have long been known to impair autoprocessing of GA precursor and/or impair its amidase activity in lysosomes,[5,6,9,16,25] the molecular details of AGU variants have only begun to be studied carefully *in vivo* and *in vitro*.[10,22–24,26] These studies indicate that different AGU substitutions result in different impacts on GA autoactivation, hydrolysis of glycoasparagines, and/or thermal stability of the enzyme. As such, detailed structural and catalytic consequences of different AGU mutations vary widely. Model structures of two AGU variants have recently been reported. The first one corresponds to the Canadian G203D variant.[22,23] The second model structure corresponds to the Finnish T234I variant.[24] These structural studies revealed that the Canadian variant appears to cause local conformational change outside of

Table I. *Different Numbering Schemes of Three AGU Variants with Crystal Structures Available and Their Corresponding Model Enzymes*[a]

	American AGU in this study	Canadian AGU[b]	Finnish AGU[b]
variant – signal peptide	T99K	G203D	T234I
variant + signal peptide	T122K	G226D	T257I
model (– signal peptide)	**T99K**	**G172D**	**T203I**

[a]Numbering differences are due to counting signal peptide and/or different lengths of the precursor-loop (P-loop) that connects the α and β subunits in the precursors.
[b]The other two AGU models with crystal structures determined and reported elsewhere.[22–24]

the amidase substrate site, which in turn disrupts the requisite autoprocessing step, thus persisting as a metabolically nonfunctional precursor protein.[22,23] On the other hand, the T234I Finnish variant appears to have a detectible level of autoprocessing capability to generate a mature amidase.[24] However, since the Thr-to-Ile substitution is located at the rim of the substrate-binding site, this variant becomes a less efficient amidase due to negative impacts on both K_M and, to a greater extent, k_{cat} of its amidase activity.[24] Interestingly, unlike these two well-characterized AGU variants, mutation of the recently found American AGU variant is located outside of the amidase substrate site and appears to have substantial autoprocessing activity to generate a mature amidase. It thus remains as a puzzle why this variant would still cause the AGU disease. In this study, we generated an AGU model enzyme, named T99K (Table I), corresponding to the American T99K variant. This model enzyme allowed for studies of the detailed biochemical and structural consequences of the Thr-to-Lys mutation at residue 99. Results indicate that replacement of the conserved threonine with a lysine has negative impacts on both K_M and k_{cat} of amidase activity. However, in contrast to the T234I variant, the T99K substitution has a greater impact on K_M than k_{cat} of substrate hydrolysis. We also report here a 1.5-Å-resolution crystal structure to enable structure–function analyses of this novel American AGU variant.

Results

Construction of a model enzyme to study the T99K AGU variant

A HeLa culture system has been reported recently for cellular characterization of AGU variants,[10] but it remains difficult to obtain sufficient amounts of enzymes and variants for *in vitro* biochemical or structural analyses.[21] In contrast, a Flavobacterium system has become an excellent model system for these studies. The justifications to use the bacterial model include the following: (a) the Favobacterial GA shares the same αββα fold with the human enzyme[27]; (b) Favobacterial and human GAs have a high sequence homology with an expect value of 1e-48 in BLAST[28]; (c) both enzymes utilize the same autoproteolytic process to cleave single-chain precursors into mature amidases that allow glycoasparagine digestion[17,25]; (d) although different in glycosylations, these two enzymes have an essentially identical 3D structure[15,27]; (e) all the amino acids around the catalytic center, those within 4 Å of the bound substrate NAcGlc-Asn, are 100% identical between these two homologs[29]; (f) the bacterial homolog allows production of crystallographic amount and purity of AGU model enzymes to determine three-dimensional structures and study their detailed kinetic parameters.[22–24] Overall, these facts advocate suitability of using the Flavobacterium homolog to study the effects of AGU mutations on GA catalytic activities. To this end, we have generated a model bacterial enzyme to analyze consequences of the American AGU variant on the GA structure and functions. According to the structure-based sequence alignments,[15] Thr99 in Flavobacterium GA is equivalent to Thr99 in the human counterpart. Thus, an AGU model with a Thr-to-Lys substitution at residue 99 of Flavobacterium GA, called T99K model (Table I), is generated to study the American T99K variant. Results of structural and biochemical analyses of this new model variant are then compared to two previously characterized AGU model enzymes: the G172D model corresponds to the Canadian G203D variant, whereas the T203I model corresponds to the Finish T234I variant (Table I). These numbering differences are due to a shorter P-loop in the model enzymes. AGU variants and the wildtype (WT) enzyme were purified according to the previously described protocols.[30]

Autoprocessing and hydrolytic activities of the T99K model enzyme

Purified T99K model enzyme was first analyzed for its autoactivation and substrate hydrolysis activities and compared to those of the WT enzyme as well as the Finish AGU model T203I reported earlier.[24] As shown in Figure 1(A), WT GA autoproteolyzed spontaneously into the functional mature form with the α and β subunits. In contrast, the purified T203I model remained as a single-chain precursor, indicating a deficiency in its autoproteolytic processing. For the T99K model, essentially all the purified protein had been autoprocessed into the mature form during purification [Fig. 1(A)]. Also like the WT enzyme, the T99K model has a residual fraction of sample recalcitrant to autoprocessing, probably due to a low level of protein denaturation. We also compared their amidase activities using a substrate analog, aspartic acid β-(*p*-nitroanilide); hydrolysis was monitored by measuring absorbance at 405 nm for the release of *p*-nitroalinine. In general, a good correlation was found between the autoprocessing [Fig. 1(A)] and amidase activities [Fig. 1(B)]. As expected from the autoproteolysis analysis mentioned above [Fig. 1(A)], T203I has negligible hydrolysis activity since it stayed mainly as an inactive precursor. On the other hand, the T99K enzyme showed relatively high amidase activity [~75% of WT, Fig. 1(B)], correlating well with high autoprocessing activity [Fig. 1(A)].

Comparative studies of catalytic kinetics between the T99K model and the WT enzyme

With such a relatively high amidase activity (75% of WT), it is intriguing that the T99K variant still causes the inherited disease AGU. To address this question, we analyzed further detailed catalytic kinetics of the T99K variant and compared them to the WT enzyme. Using the purified and autoprocessed amidases, kinetic studies of the T99K model were conducted

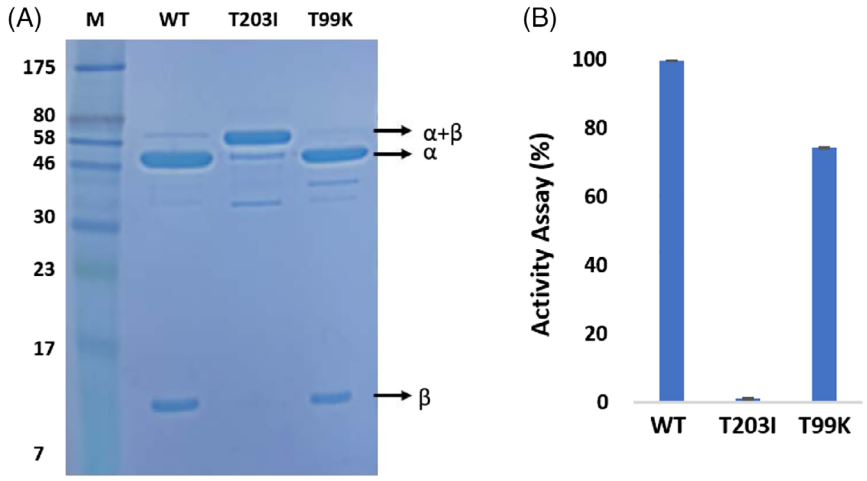

Figure 1. Autoactivation and amidase activities of AGU T99K model compared to the wildtype (WT) GA and another AGU model T203I. (A) SDS-PAGE analysis of the T99K model enzyme. Lane M represents a mixture of molecular weight markers. Lanes WT, T203I, and T99K are purified WT GA and AGU models T203I and T99K, respectively. The sizes corresponding to precursor (α + β) and autocleaved subunits (α and β) are marked. (B) Amidase activity of WT GA and AGU models T203I and T99K. The activity of WT GA is normalized to 100%. Data are the average of three repeats with standard error shown as an error bar.

and compared to the results of the WT enzyme. Kinetic parameters K_M and k_{cat} of the variant and WT GA were determined using the natural substrate NAcGlc-Asn. Typical kinetic analyses of the initial rates as a function of substrate concentration are shown in Figure 2. WT enzyme had a K_M for the natural substrate NAcGlc-Asn of 0.090 mM, and the k_{cat} was 14.18 s^{-1} (Table II). These kinetic parameters are similar to previously reported values for human and bacterial GAs.[31,32] Substitution of Thr99 by a lysine greatly reduced the enzyme activities; both K_M and k_{cat} were adversely affected (Table II). The T99K variant exhibited a bit more than twofold decrease in $k$cat. Relative to the WT enzyme, the K_M value of the T99K variant for the natural substrate was affected more than k_{cat} by about threefold. As a result, the specificity constant (k_{cat}/K_M) of the T99K variant decreased by sevenfold when compared to the WT enzyme. It is possible that the kinetic effects of T99K mutation could be more severe at the physiological concentrations of glycoasparagines in lysosomes. Nonetheless, these negative kinetic effects of the T99K variant are in sharp contrast to those of a previously characterized T203I model,[24] which has a much more dramatic (>300-fold) reduction of the specificity constant (k_{cat}/K_M) (Table II).

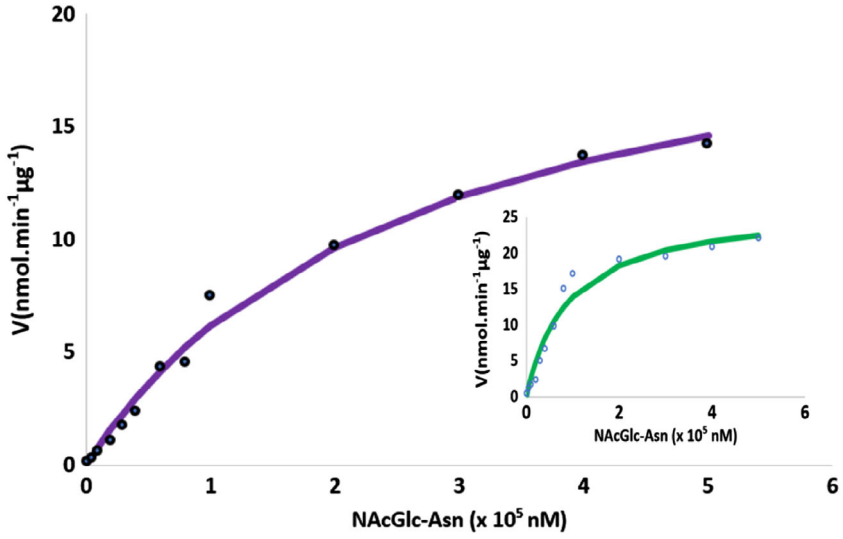

Figure 2. Kinetic analyses of the T99K mature amidase and the wildtype (WT) GA. Purified T99K and WT GA (0.04 µg) were used to determine the catalytic parameters K_M and k_{cat} of hydrolyzing NAcGlc-Asn as a substrate. Initial rates (V) were calculated based on reactions at three different time points, and are expressed in unit of nmoles.min^{-1}.μg^{-1} protein for the T99K variant (—●—) and the WT GA (inset, —○—).

Table II. *Kinetic Parameters of Two AGU Model Enzymes and the Wildtype GA*

Enzyme	Wildtype GA	AGU T99K model	AGU T203I model[a]
K_M (mM)	0.090 ± 0.018	0.260 ± 0.037	0.166 ± 0.029
k_{cat} (sec^{-1})	14.18 ± 0.99	5.93 ± 0.42	0.08 ± 0.01
k_{cat}/K_M (sec^{-1} mM^{-1})	160	23	0.48

[a]Kinetic parameters of the T203I model are taken from a previous publication[24] and are listed here for comparison.

Structural determination of the T99K variant

To provide structural basis for the effects of the Thr-to-Lys point mutation on enzyme structure and catalysis, we carried out structural determination of the T99K model by X-ray crystallography. To this end, the purified T99K protein was subject to multiple crystallization screens. Promising needle-like crystals from initial screenings were further improved by micro- and macroseeding techniques to obtain diffracting quality crystals. Before X-ray data collection, the crystals were soaked with 20% glycerol as a cryoprotectant for data collection at 100 K. The crystal structure of the T99K model has now been determined and refined to 1.5 Å resolution. The space group and cell constants of T99K crystals are as follows: $P2_1$, a = 46.2, b = 96.0, c = 61.5 Å; β = 90.1°, with an R_{free} of 0.225 and an R_{work} of 0.194. Other crystallographic statistics are summarized in Table III. Consistent with the autoproteolysis analysis [Fig. 1(A)], initial electron density maps calculated from the diffraction data clearly showed that the P-loop (residues 142–151) linking α and β subunits in the precursor had been autocleaved into the mature form in both molecules of the asymmetric unit, with a well-resolved N-terminal amino group at Thr-152 [Fig. 3(A)].

Structural comparisons between the T99K model and the WT enzyme

To study the effects of the Thr-to-Lys point mutation on enzyme structure, determined atomic coordinates of the T99K model and the WT GA enzymes (PDB code 2GAW)[15] were superimposed and compared. Overall, the structure of the T99K model is very similar to that of the WT enzyme, with the typical αββα-sandwich fold. The rmsd of all main-chain atoms of 275 residues is 0.29 Å, indicating a well-folded GA protein with no gross conformational changes or misfolding. Similar to the apostructure of the WT enzyme, the T99K variant acquires a wide opening near the substrate-binding site.[15] This suggests that T99K has its substrate-binding site fully opened through autoproteolysis and is ready to accommodate the substrate NAcGlc-Asn.

There are, however, a few significant deviations of the main-chain trace, with rmsd >1 Å, between these two structures. The most significant deviation is located around the mutated residue 99. Despite higher-resolution diffracting data being collected for the T99K structure, the calculated electron densities do not give good densities to build the main-chain backbone for residues near the mutation site, residues K99 to H101 [Fig. 3(B), right panel]. In contrast, the same segment in the WT structure has well-resolved electron density for a clearly determined conformation [Fig. 3(B), left panel]. In the T99K structure, even though we were able to build other neighboring residues, the average rms displacement at residue 98 is 1.32 Å (vs. an average of 0.29 Å). Correlating well with these large positional deviations, neighboring residues (98, 101, and 102) of the T99K variant also have much higher temperature factors B than the WT structure (with ΔB 14.3 $Å^2$ over an average ΔB of 2.6 $Å^2$), indicating a static disorder (adopting multiple conformations) and/or a dynamic disorder (staying mobile on protein surface) near the mutated site Lys99. Furthermore, the side chain of H101 also appears to become disorder, resulting in loss of specific interactions with R180 and other residues observed in the WT structure

Table III. *Crystallographic Data Collection and Refinement Statistics*

Resolution (Å)[a]	36.9–1.50 (1.53–1.50)
Space group	$P2_1$
Cell dimensions	
a, b, c (Å)	46.2, 96.0, 61.5
α, β, γ (°)	90, 90.1, 90
No. of molecules per asymmetric unit	2
I/sigma-I	9.8 (3.2)
Completeness (%)	99.0 (97.8)
R_{sym} (%)[b]	8.0 (42.0)
Structure refinementresolution (Å)	20.0–1.50
No. of reflections	80,277
R_{work}[c]	0.194
R_{free}[d]	0.225
R.m.s. deviations[e]	
Bond lengths (Å)	0.02
Bond angles (°)	2.21
Ramachandran plot	
Most favored regions (%)	97.2
Additional allowed regions (%)	2.3
Outliners (%)	0.4
B factors ($Å^2$)	
Main chain	12.5
Side chain	16.4
Water	19.1

[a]Numbers in parenthesis refer to the outermost resolution bin.

[b]$R_{sym} = \Sigma_h\Sigma_i \mid I_{hi} - I_h \mid / \Sigma_h\Sigma_i I_{hi}$ for the intensity (I) of i observation of reflection h.

[c]$R_{work} = \Sigma \mid F_{obs} - F_{calc} \mid / \Sigma \mid F_{obs} \mid$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively.

[d]R_{free} was calculated as R_{work}, but with 5% of the amplitudes chosen randomly and omitted from the start of refinement.

[e]R.m.s. deviations are deviations from ideal geometry.

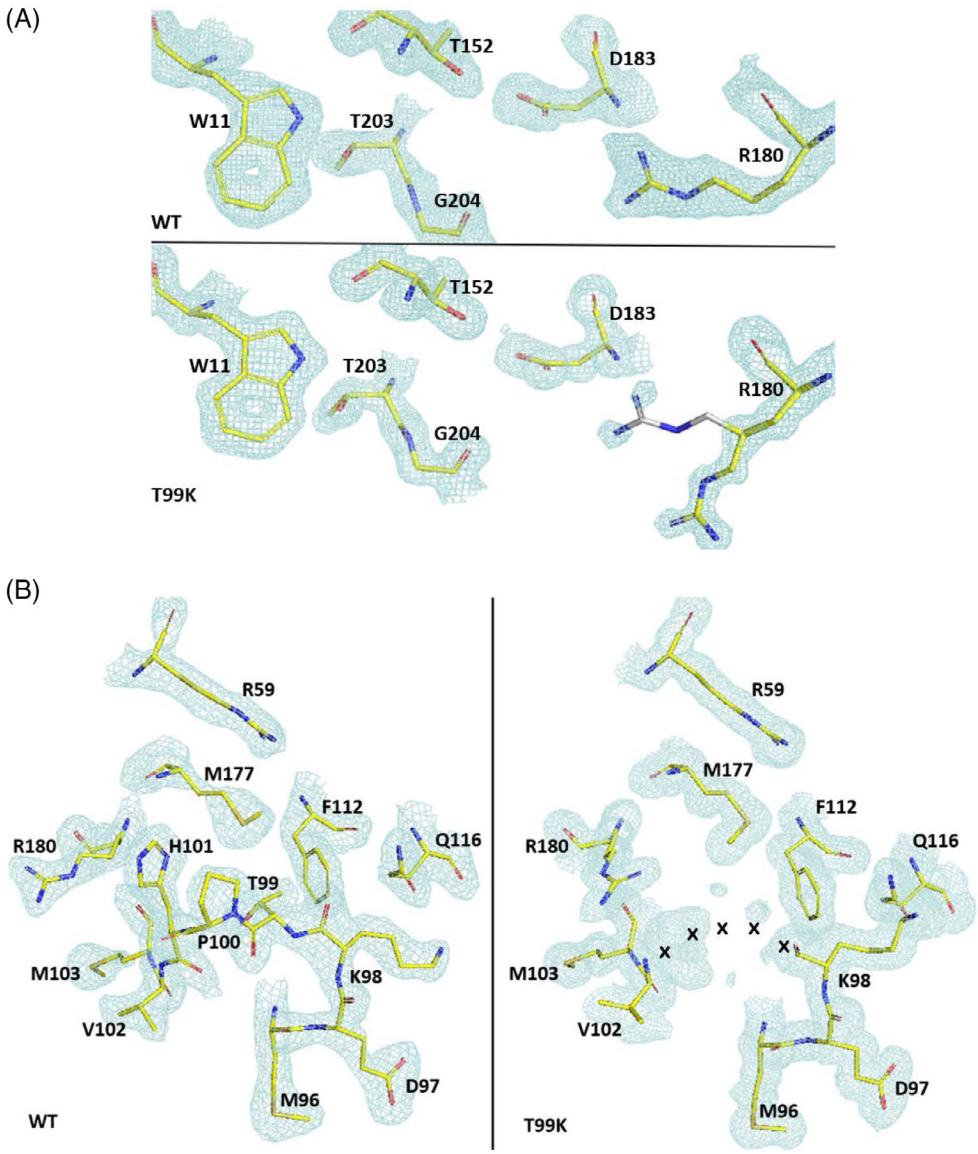

Figure 3. Electron density maps at the catalytic site of the T99K model and the wildtype (WT) enzyme. (A) Comparison of electron density maps at the amidase catalytic sites between the WT enzyme (top panel)[15] and the T99K model enzyme (bottom panel). The cyan electron density corresponds to a $2F_o$-F_c type map at 1.5 Å resolution contoured at the 1.2-σ level. The hydrolytic nucleophile is the side-chain hydroxyl group of Thr152. Side-chains of a few key active site residues are shown by atom type: Yellow for carbon atoms, blue for nitrogen atoms, red for oxygen atoms. Side chains of R180 of the T99K variant are shown with two alternative conformations (in bottom panel), with a minor form in gray carbon atoms. (B) Side-by-side comparison of electron density maps between the WT GA (left panel) and the T99K model enzyme (right panel).[15] Color schemes are the same as in (a). Note the contrast between the well-resolved and continuous density for T99, P100, and H101 in the left panel, and a lack of density for the corresponding segment in the right panel (marked by xxxxx). Also note different side-chain conformations of R180 and K98 between the WT GA (left panel) and the T99K model (right panel). In the T99K structure (right panel), lack of continuous density near the mutated site K99 indicates no defined conformation for residues K99 to H101 due to a static or dynamic disorder.

(see below). Two other significant local main-chain deviations are observed, by an average of 1.256 Å and 1.033 Å rmsd at residues T138 and P290, respectively. These latter two main-chain deviations are located about 43 and 37 Å away from the mutated residue, respectively, thus are not caused by direct contacts with the mutated K99. Instead, T138 and P290 locate near the C-terminal ends of the α and β subunits, respectively, and have been observed in other GA structures with consistently higher B factors.[15,20,22–24] Thus, these latter two deviations are more likely to be due to a static disorder and/or dynamic flexibility typical for polypeptides' C-terminal tails, along with elevated B factors (43.9 and 29.0 Å^2, respectively, compared to an overall average B of 13.2 Å^2).

On the other hand, the main-chain deviation at residues 98–101 appears to be direct results of the Thr-to-Lys mutation at residue 99. The single a.a. substitution results in main-chain and side-chain disorder of residues 98–101, that is further propagated to the

substrate-binding site through the GA dimer interface (see below). It has been shown that GA functions as a dimer[9] and crystallized with the same dimer structure in all different crystal forms and variants.[14–16,20,22–24] We have previously shown that an intact dimer structure is critical for GA functions.[9] Mutations selected to disrupt this dimer interface results in loss of GA autoprocessing and amidase activities. At this dimer interface of WT structure, the side chain of H101 from each monomer packs against a surface pocket formed by M177, R180, and E207 of the other monomer (Fig. 4), which was demonstrated by well-resolved electron density maps [Fig. 3(B)]. In contrast, in the T99K dimer structure, there is no/poor electron density to define main-chain trace of residues 99–100 [Fig. 3(B)]. Furthermore, it appears that the side chain of H101 from one monomer has become disordered, with concurrent disorders observed for side chains of R180 and E207 in the other monomer [Fig. 3(B)]; both latter residues are parts of the substrate-binding pocket.

A T99K–substrate complex model for structure–function analyses

To analyze the effects of the R180 and E207 disorder at the substrate site as a result of T99K mutation, we built a T99K–substrate complex model, based on a previously published structure of the WT GA–substrate complex (2GL9).[29] As demonstrated in SDS-PAGE mentioned above [Fig. 1(A)], the T99K model appears to be capable of autoprotolysis. However, the mature variant has a lower k_{cat} and a significantly higher K_M when compared to the WT enzyme (Table II). The T99K–substrate complex model could also allow us to analyze the structural basis for the reduction of specificity activity. To this end, a superimposition by all the secondary structure elements was done between the WT–substrate complex and the T99K apostructure. The substrate model from the WT–substrate complex was then placed into the apo- T99K model structure to build the T99K–substrate complex structure. In the WT–substrate complex structure,[29] nine substrate-binding residues are in close contact (within 4 Å distances) with the substrate molecule: W11, F13, S50, T152, R180, D183, T203, G204, and G206 [Fig. 5(A)]. For the T99K variant, all these substrate-binding residues have identical conformations as the WT enzyme, with the exceptions of R180 and nearby E207. As pointed out above, in the T99K variant, the side chain of R180 became disordered on the dimer interface as a result of the disordered segment around the mutated residue in the other monomer [Fig. 3(B)]. Similarly, the side chain of nearby E207 also appears to become flexible, indicated by bifurcated electron densities to allow alternative conformations built. Based on the WT complex structures,[29] R180 and E207 make up one end of the substrate-binding side. Thus, these concerted disorders of R180 and E207 in the T99K variant appear to have impeded the binding of the substrate, with an about threefold increase in K_M (Table II).

On the other hand, all the residues directly involved in catalysis are intact in the T99K variant, including the nucleophile at T152, the proposed oxyanion holes T203 and G204 to stabilize the negatively charged transition-state intermediate, and the base at the N-terminal amino group of the β subunit. There is, thus, only a moderate (about twofold) reduction on k_{cat} (Table II). For comparison, another AGU model T203I has a 177-fold reduction on k_{cat} (Table II) due to a direct involvement of the T203 side chain as the oxyanion hole for catalysis[24] [Fig. 5(B)]. Overall, kinetics data have corroborated structural analyses for the enzymatic roles of a few key residues. We previously showed that T203I mutation has a much more severe impact on k_{cat} (177-fold) than K_M (~twofold). In

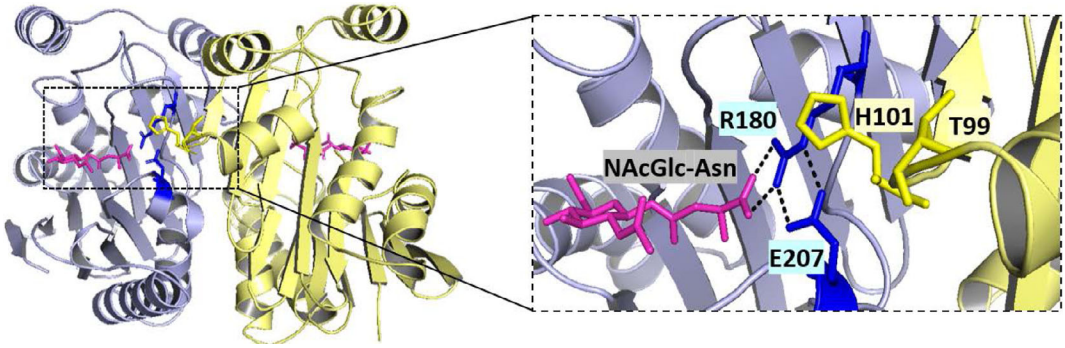

Figure 4. Dimer structures of the wildtype (WT) GA. Ribbon drawing of the dimer structure of the WT GA in complex with a substrate.[29] One molecule is in *light blue*, and the other is in *yellow*. An NAcGlc-Asn substrate molecule (colored in magenta) is modeled into each of the two amidase sites that are facing away from each other in the dimer. A closeup panel is shown on the right to highlight dimer interface contacts around the substrate-binding residue R180 in the light-blue molecule: H101 (colored in yellow, on the same surface loop as T99) from one monomer contacts R180 and E207 (colored in blue) from the other monomer. For clarity, not shown is M177 of the blue molecule that is also contacting H101 of the yellow molecule. These interacting side chains are part of the WT GA dimer interfaces, but become disorder in the AGU variant T99K structure due to the Thr-to-Lys substitution at residue 99. Selected hydrogen-bond interactions are denoted by dashed lines.

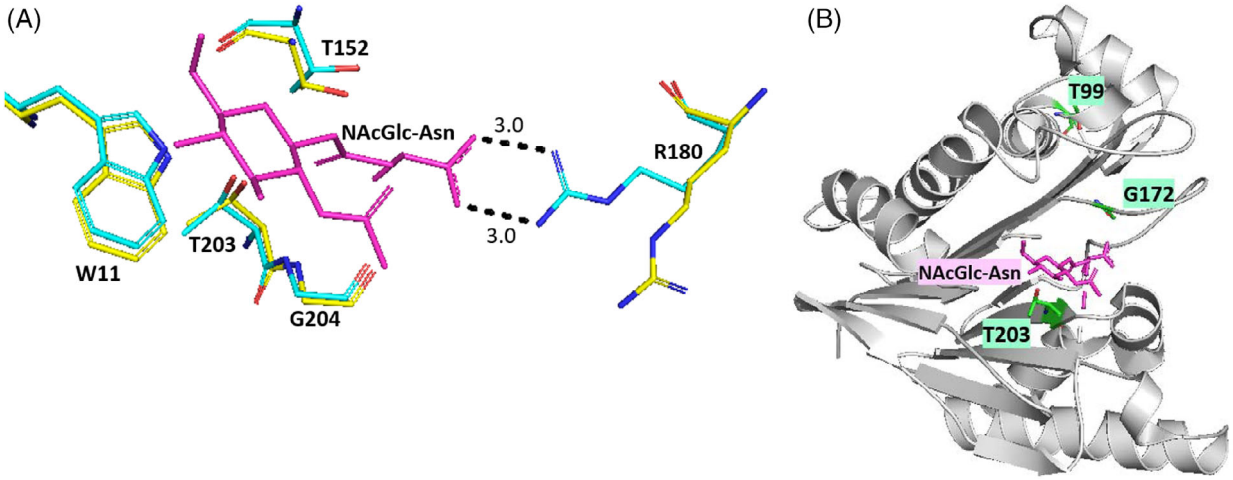

Figure 5. Comparisons of T99K–substrate and the WT–substrate complex models. (A) Superperposition of the WT–substrate complex[29] with the T99K model. Side chains of a few key active site residues are shown by atom types: Blue for nitrogen atoms, red for oxygen atoms, yellow for carbon atoms of the T99K model, and cyan for carbon atoms of the WT enzyme. Dashed lines represent hydrogen bond interactions between R180 and the substrate NAcGlc-Asn (in magenta) in the WT complex, with distances designated in angstroms (Å). (B) Ribbon drawing of the GA monomeric structure in complex with a substrate. An NAcGlc-Asn substrate molecule (colored in magenta) is placed in the amidase site. Locations of the three AGU models discussed in this study (T99K, G172D, and T203I) are indicated with green labels.

contrast, this study revealed that side-chain disorders of R180 and E207, a result of the T99K mutation, cause more reduction on K_M (~threefold) than k_{cat} (~twofold) of glycoasparagine hydrolysis.

Discussion

AGU patients, mostly in the Finnish population but also found in other parts of the world, carry various missense mutations or deletions in their GA genes. These AGU mutations produce defective GA variants that affect the amidase activity of GA either by preventing its autoprocessing and/or by reducing its ability to hydrolyze the substrate NAcGlc-Asn and other glycoasparagines. In this study, the T99K enzyme, which is a model for the novel American AGU allele,[10] was found to have autoprocessing activity comparable to the WT enzyme, and is thus purified as a mature enzyme with separated α and β subunits (Fig. 1). This is somewhat a surprise because a previous *in vivo* study suggests that the human T99K variant results in a precursor more defective in autoprocessing; in fibroblasts of AGU patients with the T99K variant, the enzyme is predominantly in the precursor form, but with increased T99K mature enzyme by coexpressing the WT GA.[10] We do not know if the discrepancies are due to *in vivo* versus *in vitro* analyses and/or protein sequence differences. Nonetheless, the T99K variant appears to represent a new class of AGU mutants (see below); the other AGU variants that had been previously characterized by structural and *in vitro* kinetic studies all showed dramatic reduction in their autoprocessing activity and thus mainly remained as inactive precursor proteins.[22–24]

In this study, we demonstrated by *in vitro* kinetic characterization and structural analyses that the T99K variant model, even capable of autoprocessing into the mature form, is still defective in its amidase activity for breaking down glycoasparagines in lysosomes (Fig. 2), consistent with its linkage to the AGU disease. Although the Thr-to-Lys substitution has negative impacts on both K_M and k_{cat} (Table II), the changes are only two- to threefold relative to the WT enzyme. This is in sharp contrast to a previously reported AGU T203I model that has a dramatic reduction of k_{cat} (~200-fold, Table II). To correlate the structure–function relationships, we have generated a T99K–substrate complex model in this study. As shown in Figure 5(A), almost all substrate-binding residues of the T99K variant (W11, F13, S50, T152, D183, T203, and G204, and G206) have identical conformations as the WT enzyme to bind and process the substrate, with the exceptions of R180. In the T99K variant, the side chain of R180 became flexible on the dimer interface with alternative conformations, caused by disordered segments around the mutated residue K99 and a conserved residue H101 from the other monomer. Similarly, the side chain of E207 in contact with R180 also became flexible to adopt alternative conformations. Contacting through the dimer interface, these concerted flexibility of R180, E207, and H101 from two different monomers in a T99K dimer appear to have impeded binding of the substrate, with an about threefold increase in K_M (Table II, Fig. 4). Significantly, the T99K substitution has a slightly larger impact on K_M than on k_{cat}, which is consistent with the structural data showing that T99K substitution disturbs only the substrate-binding residues (e.g., R180) but does not alter other residues directly involved in catalysis

(e.g., T152 or T203) [Fig. 3(A)]. In contrast, the T203I mutation results in a ~200-fold decrease in k_{cat} but only with an about twofold increase in K_M. As revealed in previous structural and kinetic studies,[24] the dramatic decrease in k_{cat} of the T203I model is likely due to the placement of the nonpolar isoleucine side chain to result in an impaired oxyanion hole that is critical for the transition state of glycoasparagine hydrolysis [Fig. 5(B)]. For the T99K variant, k_{cat} is only slightly altered since the residues directly involved in catalysis, such as T152 and T203, are all intact and adopt the same structure as the WT enzyme [Figs. 3(A) and 5(A)]. Overall, the complex structure model correlated well with the kinetic analyses using NAcGlc-Asn as the substrate.

Recently, progress has been reported on identifying small-molecule leads in hope to alleviate suffering of AGU patients and their families. There have been successful precedents of drug-like small molecules to induce protein folding and stimulate intracellular transport of proteins.[33–35] For AGU variants defective in autoprocessing, small molecule activators have also been described to enhance autoprocessing of an AGU precursor into a mature amidase[10] or to suppress nonsense AGU truncation.[36] However, this approach might not be the best strategy for the T99K variant, which appears to be already capable of autoprocessing into its mature form. Instead, the main defect of the T99K enzyme appears to arise from a floppy-binding site due to structural disorders around the key binding residue R180. This structural disorder appears to be a direct result of the Thr-to-Lys substitution at residue 99 of its dimeric partner, which is propagated through the GA dimer interface (Fig. 4). As a result, the T99K variant has a lower substrate-binding affinity with a higher K_M and a lower specificity constant (k_{cat}/K_M) (Table II). Nonetheless, rescuing the T99K variant might be more straightforward than the other AGU variants since it still has most of the catalytic center intact with a relatively high k_{cat} for digesting glycoasparagines. A "molecular clamp" capable of stabilizing the dimer interface disorders around R180 might be able to reduce K_M and increase the specificity constant (k_{cat}/K_M) of the T99K variant to a sufficient level for alleviating AGU symptoms. Thus, in principle, the T99K defect could be rescued by reagents that could stabilize such local disorders at the dimer interface. This novel type of reagents targeting at the T99K dimer interface appears to be more straightforward than those needed to rescue the T203I model (corresponding to the Finnish T234I variant), because the latter variant has an impaired oxyanion hole of catalysis.[24] It thus appears hopeful to find novel types of lead compounds to treat T99K-type AGU variants. Interestingly, the dimer interface disorders of the T99K variant, located ~18 Å away from the catalytic center at T152, is accessible through bulk solvent without a need to interfere with substrate binding (Fig. 4). Therefore, some chaperonin-like molecules might be found to stabilize this dimer interface disorder in order to enhance the substrate affinity and catalysis of the T99K variant. It is worth noting that, unlike previously proposed glycine-like autoprocessing activators that need to squeeze into the small autoprocessing site,[16,18] dimer interface chaperonins for T99K do not need to compete with substrate for binding to the same pocket, and could be large molecules such as dimer-binding proteins or customized antibodies. Consistent with this speculation, the bacterial model enzyme has a proline residue next to the mutated Lys99 ($\underline{K}^{99}$PH), whereas the human variant has a threonine at the corresponding location ($\underline{K}^{99}$TH). Thus it is plausible that the human T99K variant has an even more flexible/unstable surface loop at the dimer interface to destabilize the dimer structure, leading to lower autoprocessing activity.[9] This might explain why fibroblasts of T99K patients were found to have the T99K variant more defective in autoprocessing and mainly stays in the precursor form.[10] Interestingly, coexpression with the WT enzyme was able to enhance autoprocessing of the T99K variant,[10] possibly by forming a more stable heterodimer. Thus, the lead compounds that could mimic or amplify the stabilization effect of proline or the WT enzyme on the dimer interface of the T99K variant may be able to achieve therapeutic effects for the T99K AGU patients. Unlike glycine- or aspartate-type activator, this novel type of dimer interface fixers does not need to be limited in small size and likely will not have inhibitory effect on binding of glycoasparagine substrates.

Materials and Methods

Protein expression and purification

Overexpression and purification of WT GA and AGU T99K model enzymes, constructed on a Flavobacterial homolog, was done using previously published protocols.[30]

Amidase activity assay

Initial amidase activity was measured using aspartic acid β-*p*-nitroanilide (Asp[pNA]-OH) (Bachem), which is a GA substrate analog, by a method modified from a previously described approach.[32]

Determination of kinetic parameters (K_M and k_{cat})

Kinetic studies of *in vitro* autoprocessed samples were performed using the GA substrate N^4-(β-*N*-acetylglucosaminyl)-L-asparagine (NAcGlc-Asn) (Bachem). Enzyme reactions were set up in 13 different concentrations of NAcGlc-Asn from 0.001 to 0.5 mM. Each reaction was in 20 μL of 20 mM sodium phosphate buffer, pH 7.5, incubated with 0.04 ug of enzyme for an appropriate time at 37°C.[32,37] The reactions were stopped by adding 50 μL of 250 mM sodium borate buffer, pH 8.8, followed by boiling for 3 min. Released *N*-acetylglucosamine (NAcGlc) was assayed by the Morgan–Elson reaction using a SpectraMax-M2 spectrophotometer (Molecular Devices, San Jose, CA).[32,37]

Curve fitting to obtain K_M, V_{max}, and k_{cat} values were calculated using an online $K_M V_{max}$ Tool Kit (Gnuplot software).

Crystallization and data collection

A hanging-drop vapor diffusion technique was used in the initial crystallization screenings. Needle-like crystals of the T99K model enzyme grew in the screens using purified model enzymes. A microseeding method was used to improve the crystal quality. Optimal crystals were formed above a well solution containing 0.2 M NaCl, 0.1 M HEPES, pH 7.5, 27% PEG 3350, and a protein concentration of 3 mg/mL.

For data collection, the crystal was cryoprotected in solution containing 100 mM Tris–HCl buffer, pH 8.0, and 20% glycerol prior to flash freezing in liquid nitrogen. X-ray data were collected using a beamline 31-ID-D at LRL-CAT, Advanced Photon Source (APS) Argonne National Laboratory, IL. The data were processed with Mosflm29 and the CCP4 suite.[38] The crystal has $P2_1$ symmetry with two protein molecules in the asymmetric unit.

Structural determination and refinement

The structure of the T99K model was determined by a molecular replacement method, using the previously published structure of the GA WT enzyme (PDB code 2GAW) as the starting model.[15] To avoid model bias, the initial MR phases were calculated by omitting residues 93–107 near the T99K substitution. Molecular replacement was performed with Molrep and refinements were carried out with the Refmac from the CCP4 program suite,[38] with 5% of the total reflection data excluded from the beginning of refinement cycles and later used to calculate the free R factor (R_{free}) for monitoring refinement progress. This partial model was further subjected to rigid body and restrained refinements, with model rebuilding done in COOT[39] to obtain the final structure. Throughout the process, composite omit maps were calculated with the simulated annealing protocol to aid model building. The final model contains 278 residues (2–98, 102–138, 152–295) in molecule 1 and 278 residues (2–98, 102–138, 152–295) in molecule 2, and 163 water molecules, and is refined to crystallographic R_{work}/R_{free} of 19.4%/22.5% at 1.5 Å resolution. The X-ray data collection, processing, and structure refinement statistics are summarized in Table III. All the figures were drawn using PyMOL (DeLano Scientific, Palo Alto, CA) and labels were added using Adobe® Photoshop.

Complex model building

The T99K-NAcGlc-Asn complex model was generated by secondary structural matching of the T99K structure with the previously published T152C GA-NAcGlc-Asn complex (GA–substrate structure, PDB code 2GL9),[29] using the CCP4 program suite.[38] The coordinates of the NAcGlc-Asn substrate were then placed into the substrate binding site of current apo-T99K structure.

Protein Data Bank accession codes

The atomic coordinates and structure factors have been deposited in the Protein Data Bank under the accession number 6NQ6.

Acknowledgments

We thank Dr. Laura Morisco for assistance on data collection and William Bizilj for preparing the manuscript. This work was supported by Grants DK075294 and GM128152 from the NIH.

Author Contributions

SP and H-CG conceived the study and designed experiments; SP performed experiments and collected data; SP and H-CG analyzed data and wrote the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Abergel C, Moulard M, Moreau H, Loret E, Cambillau C, Fontecilla-Camps JC (1991) Systematic use of the incomplete factorial approach in the design of protein crystallization experiments. J Biol Chem 266:20131–20138.
2. Maury CP (1982) Aspartylglycosaminuria: an inborn error of glycoprotein catabolism. J Inherit Metab Dis 5:192–196.
3. Arvio M, Mononen I (2016) Aspartylglycosaminuria: a review. Orphanet J Rare Dis 11:162.
4. Autio S, Visakorpi JK, Jarvinen H (1973) Aspartylglycosaminuria (agu). Further aspects on its clinical picture, mode of inheritance and epidemiology based on a series of 57 patients. Ann Clin Res 5:149–155.
5. Aronson NNJ (1999) Aspartylglycosaminuria: biochemistry and molecular biology. Biochim Biophys Acta 1455: 139–154.
6. Saarela J, Laine M, Oinonen C, Schantz C, Jalanko A, Rouvinen J, Peltonen L (2001) Molecular pathogenesis of a disease: structural consequences of aspartylglucosaminuria mutations. Hum Mol Genet 10:983–995.
7. Opladen T, Ebinger F, Zschocke J, Sengupta D, Ben-Omran T, Shahbeck N, Moog U, Fischer C, Burger F, Haas D, Ruef P, Harting I, Al-Rifai H, Hoffmann GF (2014) Aspartylglucosaminuria: unusual neonatal presentation in qatari twins with a novel aspartylglucosaminidase gene mutation and 3 new cases in a turkish family. J Child Neurol 29:36–42.
8. Ikonen E, Enomaa N, Ulmanen I, Peltonen L (1991) In vitro mutagenesis helps to unravel the biological consequences of aspartylglucosaminuria mutation. Genomics 11:206–211.
9. Wang Y, Guo H-C (2003) Two-step dimerization for autoproteolysis to activate glycosylasparaginase. J Biol Chem 278:3210–3219.
10. Banning A, Gulec C, Rouvinen J, Gray SJ, Tikkanen R (2016) Identification of small molecule compounds for pharmacological chaperone therapy of aspartylglucosaminuria. Sci Rep 6:37583.
11. Tollersrud OK, Aronson NNJ (1992) Comparison of liver glycosylasparaginases from six vertebrates. Biochem J 282:891–897.

12. Liu Y, Dunn GS, Aronson NNJ (1996) Purification, biochemistry and molecular cloning of an insect glycosylasparaginase from *Spodoptera frugiperda*. Glycobiology 6:527–536.
13. Tarentino AL, Plummer THJ (1993) The first demonstration of a procaryotic glycosylasparaginase. Biochem Biophys Res Commun 197:179–186.
14. Oinonen C, Tikkanen R, Rouvinen J, Peltonen L (1995) Three-dimensional structure of human lysosomal aspartylglucosaminidase. Nat Struct Biol 2:1102–1108.
15. Guo H-C, Xu Q, Buckley D, Guan C (1998) Crystal structures of flavobacterium glycosylasparaginase: an n-terminal nucleophile hydrolase activated by intramolecular proteolysis. J Biol Chem 273:20205–20212.
16. Xu Q, Buckley D, Guan C, Guo H-C (1999) Structural insights into the mechanism of intramolecular proteolysis. Cell 98:651–661.
17. Saarela J, Oinonen C, Jalanko A, Rouvinen J, Peltonen L (2004) Autoproteolytic activation of human aspartylglucosaminidase. Biochem J 378:363–371.
18. Wang Y, Guo H-C (2010) Crystallographic snapshot of glycosylasparaginase precursor poised for autoprocessing. J Mol Biol 403:120–130.
19. 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 autoproteolysis. J Biol Chem 271:1732–1737.
20. Qian X, Guan C, Guo H-C (2003) A dual role for an aspartic acid in glycosylasparaginase autoproteolysis. Structure 11:997–1003.
21. Saarela J (2004) Characterization of aspartylglucosaminidase activation and aspartylglucosaminuria mutations. PhD Thesis. University of Helsinki, Helsinki, Finland.
22. Sui L, Lakshminarasimhan D, Pande S, Guo H-C (2014) Structural basis of a point mutation that causes the genetic disease aspartylglucosaminuria. Structure 22: 1855–1861.
23. Pande S, Lakshminarasimhan D, Guo H-C (2017) Crystal structure of a mutant glycosylasparaginase shedding light on aspartylglycosaminuria-causing mechanism as well as on hydrolysis of non-chitobiose substrate. Mol Genet Metab 121:150–156.
24. Pande S, Bizilj W, Guo HC (2018) Biochemical and structural insights into an allelic variant causing the lysosomal storage disorder—aspartylglucosaminuria. FEBS Lett 592:2550–2561.
25. Guan C, Liu Y, Shao Y, Cui T, Liao W, Ewel A, Whitaker R, Paulus H (1998) Characterization and functional analysis of the cis-autoproteolysis active center of glycosylasparaginase. J Biol Chem 273:9695–9702.
26. Banning A, Konig JF, Gray SJ, Tikkanen R (2017) Functional analysis of the ser149/thr149 variants of human aspartylglucosaminidase and optimization of the coding sequence for protein production. Intl J Mol Sci 18:706.
27. Oinonen C, Rouvinen J (2000) Structural comparison of ntn-hydrolases. Protein Sci 9:2329–2337.
28. Tatusova TA, Madden TL (1999) Blast 2 sequences, a new tool for comparing protein and nucleotide sequences. FEMS Microbiol Lett 174:247–250.
29. Wang Y, Guo H-C (2007) Crystallographic snapshot of a productive glycosylasparaginase-substrate complex. J Mol Biol 366:82–92.
30. Cui T, Liao P-H, Guan C, Guo H-C (1999) Purification and crystallization of precursors and autoprocessed enzymes of flavobacterium glycosylasparaginase: an n-terminal nucleophile hydrolase. Acta Cryst D55: 1961–1964.
31. Kaartinen V, Williams JC, Tomich J, JRr Y, Hood LE, Mononen I (1991) Glycosaparaginase from human leukocytes: inactivation and covalent modification with diazooxonorvaline. J Biol Chem 266:5860–5869.
32. Liu Y, Guan C, Aronson NNJ (1998) Site-directed mutagenesis of essential residues involved in the mechanism of bacterial glycosylasparaginase. J Biol Chem 273: 9688–9694.
33. Petrassi HM, Johnson SM, Purkey HE, Chiang KP, Walkup T, Jiang X, Powers ET, Kelly JW (2005) Potent and selective structure-based dibenzofuran inhibitors of transthyretin amyloidogenesis: kinetic stabilization of the native state. J Am Chem Soc 127:6662–6671.
34. Ray SS, Nowak RJ, Brown RHJ, Lansbury PTJ (2005) Small-molecule-mediated stabilization of familial amyotrophic lateral sclerosis-linked superoxide dismutase mutants against unfolding and aggregation. Proc Natl Acad Sci U S A 102:3639–3644.
35. Rivera VM, Wang X, Wardwell S, Courage NL, Volchuk A, Keenan T, Holt DA, Gilman M, Orci L, Cerasoli FJ, Rothman JE, Clackson T (2000) Regulation of protein secretion through controlled aggregation in the endoplasmic reticulum. Science 287:826–830.
36. Banning A, Schiff M, Tikkanen R (2018) Amlexanox provides a potential therapy for nonsense mutations in the lysosomal storage disorder aspartylglucosaminuria. Biochim Biophys Acta 1864:668–675.
37. Levvy GA, McAllan A (1959) The n-acetylation and estimation of hexosamines. Biochem J 73:127–132.
38. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the ccp4 suite and current developments. Acta Cryst D67:235–242.
39. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D60:2126–2132.