Title:

N-glycosylation in the SERPIN Domain of the C1-Esterase Inhibitor in Hereditary Angioedema

Zhen Ren,[1] John Bao,[1] Shuangxia Zhao,[2] Nicola Pozzi,[3] H. Wedner,[1] John P. Atkinson[4]

[1] Department of Medicine, Division of Allergy and Immunology, Washington University School of Medicine, St. Louis, Missouri, USA

[2] Department of Molecular Diagnostics and Endocrinology, The Core Laboratory in Medical Center of Clinical Research, Shanghai Ninth People's Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China

[3] Department of Biochemistry and Molecular Biology, Edward A. Doisy Research Center, Saint Louis University School of Medicine, St. Louis, MO 63104, USA.

[4] Department of Medicine, Division of Rheumatology, Washington University School of Medicine, St. Louis, Missouri, USA

Correspondence:

Zhen Ren, MD, PhD

Washington University School of Medicine,

660 S Euclid Avenue,

Campus Box 8122, St. Louis, MO 63110, USA

ren.zhen@wustl.edu

Conflict-of-interest statement: H. J. Wedner reports research funding from BioMarin, KalVista, Pharvaris and Takeda; consulting fees from CSL, Takeda, Biocryst and BioMarin; and speaking fees from Genentech, GSK, CSL, Takeda and BioCryst. No other authors have conflicts of interest to report.

Abstract

Hereditary angioedema is an autosomal dominant disorder caused by defects in C1-esterase inhibitor (C1-INH), resulting in poorly controlled activation of the kallikrein-kinin system and bradykinin overproduction. C1-INH is a heavily glycosylated protein in the serine protease inhibitor (SERPIN) family, yet the role of these glycosylation sites remains unclear. To elucidate the functional impact of N-glycosylation in the SERPIN domain of C1-INH, we engineered four sets consisting of 26 variants at or near the N-linked sequon (NXS/T). Among these, six are reported in HAE patients and five are known C1-INH variants without accessible clinical histories. We systematically evaluated their expression, structure and functional activity with C1$\bar{s}$, FXIIa and kallikrein. Our findings showed that of the eleven reported variants, seven are deleterious. Deleting N at the three naturally occurring N-linked sequons (N238, N253 and N352) results in pathologic consequences. Altering these sites by substituting N to A disrupts N-linked sugar attachment but preserves protein expression or function. Further, an additional N-linked sugar generated at N272 impairs C1-INH function. These findings highlight the importance of N-linked sequons in modulating the expression and function of C1-INH. Insights gained from identifying the pathological consequences of N-glycan variants should assist in defining more tailored therapy.

Introduction

Hereditary angioedema (HAE) is an autosomal dominant (AD) disorder caused by defects in the *SERPING1* gene encoding C1-esterase inhibitor (C1-INH) (1). C1-INH is a multi-functional protein that inhibits the contact system, fibrinolytic pathway and complement activation. A dysfunctional C1-INH results in poorly controlled activation of kallikrein-kinin (PKa) system leading to bradykinin (BK) overproduction, known to be the primary mediator of recurrent and unpredictable soft tissue swelling. Laryngeal edema is associated with the risk of asphyxiation, which can be life-threatening (1).

HAE is categorized based on a low laboratory plasma C1-INH antigen level (TY I) or normal antigen but with a dysfunctional C1-INH protein (TY II) (2). This AD disorder results in haploinsufficiency which would be expected to produce 50% of the normal level of C1-INH in plasma (3). However, most patients with HAE present with plasma C1-INH antigen levels of ~ 5 to 30% of normal (4). The cause of these disproportionally decreased C1-INH plasma levels remains largely unknown (5, 6). Further, the severity of clinical symptoms do not necessarily correlate with the antigenic levels of the C1-INH (1).

Over eleven hundred variants in the *SERPING1* gene have been reported and more than one-half are missense SNVs (7). Interpreting the pathogenicity of these *SERPING1* variants can be difficult. This is in part due to the limitations of *in silico* prediction tools and of the genetic variant databases. Relative to the latter, each genetic variant database provides its own unique annotations and datasets, presenting a challenge when integrating clinical and genetic data from different resources.

C1-INH protein contains 500 aa that includes a 22 aa signal peptide, an amino-terminal domain (NTD) consisting of 112 aa and a carboxyl-terminal domain (CTD), also named the serine

protease inhibitor (SERPIN) domain, containing 366 aa (Figure 1). The SERPIN domain is highly conserved, comprising 9 α-helices, 3 β-sheets and a reactive center loop (RCL) (Figure 2) (8). The target protease binds to the RCL, forming a reversible “Michaelis-Menten complex”, and then cleaves the P1–P1′ scissile bond (R466-T467) (9). Upon the cleavage, the SERPIN-protease complex forms a covalent bond between P1 residue of the RCL and the active site serine of the protease (Figures 2-3) (9, 10). The cleaved RCL undergoes drastic conformational changes and inserts into β sheet A (SA), becoming an additional strand, S4A, in the SA domain. This process leads to the formation of an irreversible C1-INH-protease complex (9). In contrast to the well-defined structure and function of the SERPIN domain, the NTD is poorly understood. Crystal structure models of C1-INH are only available for the CTD, as the first one hundred aa from the NTD were not included in the studied structures (Figures 2 and 3) (11).

In typical glycoproteins, the glycans can constitute up to ~20% of the total weight (12). Whereas C1-INH is a heavily glycosylated protein with an estimated relative electrophoretic mobility (M_r) of 105,000, in which ~ 50% of the mass results from glycosylation (13). C1-INH contains six N-glycosylation sites, of which three sites are in the NTD while the others are in the SERPIN domain (Figure 1 and Figure 3A). The NTD also contains up to 26 O-glycosylation sites (Figure 1) (14). Even though most glycans reside in the NTD, the recombinant C1-INH without NTD appears to have preserved SERPIN function, suggesting that N-glycosylation on the SERPIN domain may play a crucial role in stabilizing C1-INH protein, assisting in protein folding, and function (10, 15, 16).

N-linked glycosylation is the predominant type of glycosylation in eukaryotic cells (17). This process is catalyzed by an oligosaccharyltransferase (OST) through recognition of an N residue in the canonical NXS/T sequon which promotes attachment of a preassembled oligosaccharide to N using an N-glycosidic bond (18, 19). The conserved NXS/T sequon requires an X be at the

+1 position, which can be any aa except for P, and either a S or a T at the +2 position of the accepting sequon (18, 19).

This study investigates the impact of *SERPING1* variants at the N-glycosylation sites in the SERPIN domain. We hypothesized that deleting the N or S/T residue in the NXS/T sequon could destabilize the protein due to it lacking glycosylation modifications. Here, we present a systematic evaluation of four sets of 26 *SERPING1* variants at or near the N-glycosylation sites in the SERPIN domain. Among these, six variants are reported in HAE patients while five variants are known human *SERPING1* variants lacking an accessible clinical history. These studies not only shed new light on the role of N-glycosylation in the SERPIN domain of C1-INH, but also validate how patient variant modeling can help identify pathological consequences and potentially point to improved clinical therapeutic strategies.

Results

This study investigates the impact of N-glycosylation on the SERPIN domain of WT C1-INH as well as the effects of known human variants through analysis of four sets of variants either at or near N-glycosylation sites (N238, N253, N352 and N272). Each set of variants includes a N to A substitution, a N deletion and a S deletion at the +2 position of the sequon. In addition, we constructed 11 *SERPING1* variants in close proximity to the N-glycosylation sites based on the published literature, the Genome Aggregation Database V4.1(gnomAD) and the Leiden Open Variation Database V3.0 (LOVD). Among these, six variants are reported in patients with HAE. The variants were expressed recombinantly in human embryonic kidney (HEK) 293T cells containing the intact glycosylation machinery. The protein expression levels of the C1-INH variants were compared to wild-type (WT) C1-INH. We then analyzed the impact of these variants on protein structure and interactions with its functional substrates, including C1s̄, FXIIa and PKa (see Table 1) (21, 22).

N-glycosylation site N238

Experimental design

The N-glycosylation sequon at N238 employs a consensus sequence, $_{236}$FV**N$_{238}$AS**RTLYSS$_{246}$ (Figure 4). To investigate the impact of impaired glycosylation at position N238 on C1-INH expression and function, we created N238A and N238del at the N-linked glycan site and S240del at the +2 position of the sequon. Additionally, we engineered A239D which is reported in gnomAD.

Genetic analysis

A239D was reported with an allele frequency (AF) of 0.0004% in the gnomAD and Mastermind databases. This variant has not been reported in other databases, including VarSome, LOVD or ClinVar. *In silico* tools predict this variant to be benign.

Structure

N238 is located on the surface of Helix E (hE), which connects strands 1 and 2 of the β-sheet A (S1A and S2A) (Figure 2). The hE plays a crucial role in facilitating the conformation change during the insertion of RCL between S3A and S5A. The deletion of N238 alters the hydrophilic surface of hE and thereby alters the packing of the hE, leading to protein misfolding (Figure 4F).

Antigenic and functional analyses

We expressed N238A, N238del, A239D and S240del separately in 293T cells. N238A demonstrates a normal protein secretory pattern while secretion of the recombinantly produced N238del and S240del are barely detectable. The secretion level of A239D is ~ 30% higher than WT (Table 1 and Figure 4A). The binding affinity of N238A and A239D to C1$\bar{s}$, FXIIa and PKa is comparable to WT (Figure 4, B-E).

Implication

The N-linked glycan site at N238 employs an NXS sequon. We hypothesize that deleting the N or S residue likely destabilizes the protein because it lacks proper posttranslational glycan structure modification. In the case of the N238 sequon, the deletion of N or S abolishes protein expression. Replacing N238A leads to normal recombinant protein expression and preserves its function (Figure 4).

A239D is at the +1 position, 'X' of the NXS sequon, so it can be any aa except P without affecting the recognition of glycan attachment. Compared to WT, A239D exhibits the same M_r and the secretory level of A239D is higher (Table 1 and Figure 4A). A239D demonstrates a normal (comparable to WT) binding affinity to FXIIa and PKa (Figure 4, B-E).

N-glycosylation site N352

Experimental design

The conserved consensus sequence at the N352 position is $_{350}$SH**N$_{352}$LS**LVILVP$_{360}$ (Figure 5A). We constructed N352A, N352del and S354del at the sequon. Based on the data from gnomAD and LOVD, we also created four additional variants: N352I, N352S, L353P and S354G.

Genetic analysis

N352S and S354G are reported in gnomAD with an AF of 0.00061% and 0.00065%, respectively. *In silico* predictions for N352S suggest that this change is benign, whereas S354G has conflicting predictions. S345G has a REVEL score of 0.845 and a CADD score of 25.33, indicating a potentially deleterious effect. Other scores, such as pholyP, Pangolin and SpliceAI, indicate that this change is benign. N352I and L353P are reported in an HAE cohort from Germany but are not included in gnomAD (23, 24).

Antigenic and Functional Analyses

Variants N352A, N352del, N352I, N352S, L353P, S354del and S354G were individually expressed in 293T cells (Table 1). The secretion of recombinantly produced N352A, N352S and S354G is comparable to WT. The secretion of N352del is markedly decreased, being ~ 10% compared to WT. The secretion of N352I is reduced to ~ 50% compared to WT (Table 1), and L353P and S354del are undetectable. N352I, N352S and S354G show a slightly lower M_r on western blot (WB) compared to WT, ~ 100,000 and 105,000, respectively. Functional analyses reveal that N352A has a normal binding to C1$\bar{s}$, PKa and FXIIa. N352S demonstrates a normal binding to C1$\bar{s}$ and FXIIa but reduced binding to PKa. Both N352I and S354G show mildly reduced binding of FXIIa and PKa, but not to C1$\bar{s}$ (Table 1 and Figure 5, B-E).

Structure

N352 is located at the end of loop in the β-sheet 2B (S2B) as it transits into S3B (Figure 5F). It is situated in the center of the breach region, where the RCL inserts into the shutter domain. Deleting N352 destabilizes the packing of S2B and S3B, which may lead to protein misfolding. The replacement of N352 to I is unfavorable. Asparagine (N) is a polar aa and prefers to be on the protein surface. Isoleucine (I) is a hydrophobic aa, which is challenging to adapt to an *α*-helical confirmation and prefers to lie within β-sheets. The substitution of N to I at N352 may change the orientation of the loop, resulting in protein misfolding and thereby leading to a decreased protein expression (Figure 5A). In the case of N352S, both N and S are polar aa. Therefore, substituting N for S is structurally tolerable but alters the glycosylation of N352. S354 is located at the tip of the S3B. Replacing a buried S354 with G would markedly reduce the side chain volume. This could cause disruption and alter the packing of the S3B region, thereby leading to C1-INH dysfunction (Figure 5F).

Implication

The N-linked glycan site at N352 employs an **NLS** sequon and the deletion of N or S in the sequon results in drastically decreased protein expression, which is similar to what we observed at the **N_{238}AS** glycosylation site (Table 2). Additionally, replacing N352A led to normal recombinant protein expression, and its functions are preserved. Replacing N352 with both S and I led to a loss of N-linked glycan attachment, resulting in a destabilized protein structure with compromised protein function (Figure 5 and Table 1).

Recombinant protein L353P is not expressed in 293T cells, which aligns with the prediction that N-glycan can attach in the NXS/T sequon, where X can be any aa except a P. Proline (P) separates the acceptor N from the hydroxyl aa S/T in the sequon and inactivates the acceptor sequence (Figure 5F), forming the structural basis for excluding P residues at the middle position (13, 25).

N-glycosylation site N253

Experimental design

The consensus sequence at the N253 position is $_{251}$LS**N_{253}NS**DANLEL$_{261}$ (Figure 6). In this sequon, we engineered N253A and N253del at the N253-glycan site and N254A and N254del at the +1 position N254. Additionally, we introduced S255del, S255G and S255T at the +2 position of the sequon.

Genetic analysis

S255G is reported in gnomAD with an AF of 0.00014% and *in silico* predictions indicate this change is benign. S255T is reported in LOVD from a HAE patient who also has a frameshift deletion in exon 7 of the *SERPING1* gene (26). This variant has an AF of 0.0027% in gnomAD and *in silico* predictions suggest this missense variant is benign (23, 27). The available

evidence is currently insufficient to determine the role of S255T in HAE; therefore, it is categorized as a Variant of Uncertain Significance (VUS).

Antigenic and functional analyses

Replacing N253 and N254 with A does not alter recombinant protein expression or secretion (Table 1). Also, deleting N253, N254 and S255 does not affect protein expression (Table 1). Except for variants N254A and S255T, the recombinantly expressed protein from all other variants, including N253A, N253del, N254del, S255del and S255G, migrate at a slightly lower M_r on WB compared to WT, likely due to loss of N-glycan attachment (Figure 6A). After deglycosylation, N254del and S255G run at the same M_r as WT (Figure 6B).

Our functional analysis indicates that N253A and N254A have normal binding affinities to C1$\bar{s}$, PKa, and FXIIa. N253del and N254del have a decrease in binding to PKa and FXIIa (Figure 6, C-F). However, their binding to C1$\bar{s}$ is normal (Table 1). The binding affinity of S255del, S255G and S255T to PKa, FXIIa and C1$\bar{s}$ is normal (Figure 6, C-F).

Structural analysis

N253 and N254 are located at the transition from the loop to hF. The replacement of N to A at N253 and N254 results in a change from hydrophilic N to hydrophobic A, which might alter the packing and orientation of the loop. S255 is located in the beginning section of the hF. Glycine (G) is the smallest aa with only one side chain of hydrogen. Due to its size and being at the transition from loop to hF, the replacement of S255 to G would not interfere with the hF backbone packing (28). In the case of S255T, S and T are both neutral and polar and the change of S to T is structurally tolerable (Table 2) (29).

Implications

The consensus sequence for the **N_{253}NS** glycosylation site displayed different variant expressions and functional profiles compared to N238 and N352. The deletion of N at the N238 or N352 sequons completely abolishes protein expression. Further, the deletion of S240 and S354 at the **N_{238}AS** and **N_{352}LS** sequon, respectively, prevents expression (Figures 4A and 5A). Whereas N253del or N254del in the **N_{253}NS** sequon does not alter the recombinant protein expression but causes a decrease in both PKa and FXIIa binding (Figure 6 and Table 1). Similarly, S255del leads to the loss of the N-glycan site but does not affect the recombinant variant protein expression or function (Figure 6, A and C-F). These changes are likely due to the secondary structure of the **N_{253}NS** sequon in C1-INH. N253 and N254 are at the loop transitioning into hE, whereas S255 resides in the hE. The deletion of N253 or N254 shortens the loop, linking S1A and hE, which leads to a decreased mobility of S1A and hE during RCL insertion. However, the deletion of S255 does not disrupt the hE packing (Figure 6G).

In the case of S255T, according to the consensus sequence for an N-glycosylation, NXS/T, the replacement of S with T theoretically should not affect the glycol attachment at the N253 site. Our data reveals that the recombinant protein of S255T exhibits the same M_r compared to the WT (Figure 6A), confirming that the replacement of T with S does not affect the N-glycan attachment at the N253 position (27).

N-glycosylation site N272

Experimental design

The sequence at the N272 site is $_{270}$T**NN**$_{272}$**KIS**RLLDS$_{280}$ (Figure 7). K273del is reported in HAE patients and noted to have a higher M_r by creating an additional N-glycosylation site at N272 (6). Given that this site contains two Ns, we constructed N271A, N271del, N272A, N272del and

S275del. Additionally, we engineered variants N271-N272del and N272D, which are reported by LOVD.

Genetic analysis

N271-N272del, N272del and K273del are reported in multiple HAE patient studies (5, 30-34). However, these three variants are absent in general population databases (gnomAD). N272D is reported in gnomAD with an AF of 0.0003% and *in silico* analyzes predict that the replacement of N with D is likely benign.

Structural analysis

K273 is in the loop connecting hF and S3A. Deletion of K273 creates an additional N-glycosylation site on N272 by changing the aa sequence N**N$_{272}$KIS** to an N-glycosylation sequon N**N$_{272}$IS**. Adding an N-glycan at the N272 position reduces the flexibility of the hF/S3A loop, which is critical for the RCL insertion during the transition from an active to a latent state (Figure 7G).

Antigenic and functional analyses

The secretion of the recombinantly produced variants N271del and N272del is reduced compared to WT (Table 1). The expression levels of variants N271A, N272A, N272D, N271-N272del, K273del and S275del are normal (Figure 7A).

The variants N271del, N272del and N271-N272del proteins exhibit a decrease in binding to C1$\bar{s}$, FXIIa, and PKa. N271A, N272A and N272D demonstrate normal binding to their substrates compared with WT (Figure 7, C-F). Deleting S275 does not change recombinant protein expression; however, the binding of S275del to PKa is markedly reduced, but not to C1$\bar{s}$ or FXIIa. The mechanism of this selective reduction in binding with PKa is under investigation.

Implications

A mass spectrometry study reported that the SERPIN domain of C1-INH carries three N-glycosylation sites, N238, N253 and N352 (14). By deleting K273, a glycosylation sequon is created at the N272 position, **$NN_{272}IS$**. The variant protein K273del's binding affinity to $C1\bar{s}$, PKa and FXIIa is reduced secondary to an additional N-linked glycan (Figure 3A) (Table 2).

Discussion

Dysfunctions in C1-INH lead to recurrent and unpredictable episodes of soft tissue swelling in HAE, which can be debilitating and life-threatening. While N-glycosylation can play a key role in stabilizing and folding of proteins, little is known about the role of N-glycosylation in the functionally relevant SERPIN domain of C1-INH. Therefore, we sought to better understand the role of N-glycosylation sites on C1-INH protein expression and function. We strategically evaluated variants at or in close proximity to the three N-linked glycan sequons in the SERPIN domain. Through this study, we determined that the loss of an N-linked sugar at N238, N253 or N352 sites is pathologic (Table 1).

The exact molecular determinants that alter N-glycan co- and posttranslational modifications are unclear (35). OST has two isoforms, STT3A and STT3B, with different roles in mediating N-linked glycosylation. The STT3A isoform is responsible for the cotranslational modification of NXS/T sequon in which the nascent polypeptide enters the lumen of the endoplasmic reticulum (ER). STT3B is mainly responsible for posttranslational N-glycosylation, modifying consensus sites that are not glycosylated by STT3A. The depletion of STT3A, but not STT3B, results in the induction of the unfolded protein response pathway. This pathway involves downregulating the transcription of secretory proteins and increasing the removal of misfolded proteins through ER-associated degradation (36-39).

In this study, we first addressed if the lack of N-glycosylation alters C1-INH protein production. The deletion of an N residue at **N_{238}AS** or **N_{352}LS** results in the abolition of recombinant C1-INH protein expression by 293T cells. Further, C1-INH variants with an S deletion at the +2 position (at these same loci), including S240del and S354del, are also not expressed. Interestingly, N238A and N352A mutants display preserved protein expression and functional levels despite lacking N-glycan attachment.

These data suggest that N-glycosylation sites at **N_{238}AS** and **N_{352}LS** sequons are likely required for cotranslational N-glycan modification. The replacement of N with A at **N_{238}** and **N_{352}** sequons is tolerable structurally and these variants could undergo folding and being exported outside of ER. However, errors in protein folding (misfolded protein) due to N and S deletion likely trigger protein cotranslational degradation (Figure 8) (40, 41).

Unlike the N-glycosylation sites at **N_{238}AS** and **N_{352}LS** sequon, the site at **N_{253}NS** contains two **N**s and exhibits different patterns. When N residues are deleted at the **N_{253}NS** site, the variants N253del and N254del show normal recombinant protein expression comparable to WT. Further S deletion at the +2 position of **N_{253}NS** sequon also results in a preserved protein expression comparable to WT. All three mutants migrate at a lower M_r on WB than WT due to a lack of N-linked glycan attachment at N253 (Figure 6).

The influences of the middle “X” residue in N-glycan sequon, NXS, have been studied. If X is a small, non-charged aa, it can be N-glycosylated in an efficient cotranslational manner. In contrast, consensus sites with bulky hydrophobic, negatively charged middle X residue, in a close spacing between NXS acceptor site or within the cysteine-rich domain, often result in a higher percentage of N-glycans being added vis posttranslational modification (27, 42, 43). We

hypothesize that the N-glycosylation at **N_{253}NS** occurs during posttranslational modification and the deletion of either N is tolerable without disrupting protein structural folding. Therefore, the recombinant protein is synthesized and secreted despite lacking N-glycan modification (Figure 6).

NXS/T sequons are highly conserved in the SERPIN domain of C1-INH. At the **N_{352}**XS/T sequon, variant L353P is not expressed in 293T cells. This is consistent with the prediction that at the X position in the NXS sequon, it can be any aa but not P as it physically prohibits N-linked sugar attachment. Additionally, the consensus NXS sequon indicates that S can be replaced by T or, less often, C. In the case of the **N_{253}NS** sequon, the variant S255T exhibits a normal protein expression and function (Figure 6), supporting the conservation of the NXS/T sequon.

N-glycosylation sites N238, N253 and N352 are located on the SERPIN domain of C1-INH and have been confirmed by mass spectrometry (14). N272 glycosylation site is noted in The Complement Factsbook (13). However, this N-glycan location site in human C1-INH has not been verified by mass spectrometry. Our understanding of the impact of glycosylation modifications on C1-INH is currently limited. The recombinant C1-INH produced in the mammary gland of transgenic rabbits is available for treating acute lesions in HAE (44, 45). Due to a lower degree of glycosylation, its efficiency suffers from an extremely short half-life of 2.4-2.7 hrs, compared to human plasma-derived C1-INH, with a half-life of 56-72 hrs (44). A question our study investigated is whether an increased number of glycosylation sites would better facilitate the C1-INH function. This was further “sparked” when we identified two unrelated HAE patients carrying heterozygous *SERPING1* variants affecting neighboring aa, N272 and K273. The patient carrying N272del presented as a TYI HAE laboratory phenotype with a low C1-INH serum level, whereas the patient with K273del demonstrated a TYII phenotype with a normal C1-INH serum level but with dysfunctional C1$\bar{s}$ binding (6). The

deletion of K273 created a new N-glycosylation site as **N_{272}IS** sequon. The recombinantly expressed K273del protein exhibited an increased M_r compared to WT, likely due to another N-linked glycan attachment (Figure 3A and Figure 7, A and B). Further functional analyses of K273del demonstrated impaired binding activity to C1$\bar{s}$, PKa and FXIIa. The presence of the additional N-linked glycan likely hindered the insertion of RCL between S3A and S5A (Figure 2 and Figure 7G). This insight, gained from the additional N-glycan site in the K273del, deepened our understanding of the impact of N-glycosylation sequon on C1-INH protein function and may be valuable for future consideration in protein modification and engineering (6, 46) of N-linked glycosylation sites.

The growing application of next-generation sequencing (NGS), exome or whole genome sequencing in investigating rare diseases has led to the identification of an increased number of variants in the *SERPING1* gene (47). To date, more than eleven hundred genetic variants are reported, among which about one-third are single nucleotide variants (SNVs) resulting in missense mutations (34, 47). Assessing the pathogenicity of these variants can be challenging as more than one-half of them are classified as a VUS (47).

In this study, we conducted strategic analyses to examine 11 reported SNVs near the three N-glycan sites in the SERPIN domain. Among these, five variants, A239D, S255G, N272D, N352S and S354G, are reported in gnomAD without an accessible clinical history. The other six variants, S255T, N352I, L353P, N271-N272del, N272del and K273del are reported in HAE patients (Tables 1 and 2). At the N352 glycosylation site, both N352I and L353P are likely pathogenic (Table 2). N352I exhibits markedly decreased protein expression and L353P is not expressed. At the N253 glycosylation site, two variants are likely benign, S255G and S255T, with normal recombinant protein expression and binding affinity to PKa, FXIIa and C1$\bar{s}$ (Figure 6C). At the N272 site, K273del has an additional N-glycan site and is dysfunctional and N272del

has low protein expression and function. N271-N272del is a rare variant reported in TYI HAE patients from a Macedonian cohort (31). Recombinant protein expression of N271-N272del is normal, but its binding affinity to C1$\bar{s}$, PKa and FXIIa is decreased. N271 and N272 are in the loop connecting hF and S5A, conserved aa in the SERPIN domains (Supplemental Figure 1). The deletion of N271-N272 likely disrupts the packing of S5A/hF and further destabilizes the SERPIN domain folding and function (Figure 7G).

It is worth noting that variants S275del and N352S show normal protein expression but selectively impair PKa binding. The mechanism of the selectively impaired SERPIN inhibition is not well understood, although it was reported in a study of twelve C1-INH P1 variants (16). The inhibitory activities of twelve R466 variants at the P1 position of RCL in the SERPIN domain were tested with C1$\bar{s}$, FXIIa, PKa and plasmin. Selectively impaired binding activity of P1R466K was observed in FXIIa, less in PKa and not in C1$\bar{s}$ or plasmin (16). Currently, the diagnosis of HAE is based on abnormal complement laboratory studies and genetic testing is not routinely performed. The functional analysis of C1INH is only commercially available to assess its binding to C1$\bar{s}$ but not to other substrates, such as PKa, FXIIa and thrombin. Various HAE therapies are available, including C1INH replacement, PKa inhibitors, FXIIa inhibitors, antifibrinolytics, and B2R antagonists, with a high average cost of $700,000 per patient-year (48). Managing the disease effectively with optimal treatment choices continues to be a persistent challenge. Thus, these findings provide valuable insights for assisting the formulation of personalized treatment by opening up the possibility of selectively choosing a medicine to directly target a specific defect that is impaired in patients with a C1-INH variant.

This investigation focuses on analyzing the impact of N-glycosylation in the SERPIN domain by examining the expression, structure and function of *SERPING1* variants. The results are most consistent with a pathologic consequence (i.e. HAE) if a missense mutation that alters/deletes

an N-linked sugar at four distinct sites in the SERPIN domain of C1-INH protein. Our findings suggest that N238 and N352 glycosylation sites undergo cotranslational modification mediated by the STT3A complex, adding oligosaccharides to the nascent protein during its insertion into the ER. This process is crucial for assisting nascent protein folding and transportation (49) (Figure 8A). The failure of attaching N-glycan to these two sites can lead to protein misfolding and trigger cotranslational protein degradation (Figure 8B) (50). Conversely, we postulate that the N253 glycosylation site is modified through posttranslational modification mediated by the STT3B complex. N253 is in close spacing between at **N_{238}AS** and **N_{352}LS** sequons that promote skipping by the STT3A complex. The posttranslational N-glycosylation at N253 is required but not necessary to acquire native protein structure. In the absence of N253-glycan attachment, the effect is tolerable compared to N238 and N352 sites (Figure 8C) (38, 50, 51). N253 site variants have preserved expression levels comparable to WT (Figure 6). In addition, our finding demonstrates that adding an additional N-glycosylation site can be deleterious by interrupting the SERPIN domain function. The scope of this study is limited to only analyzing the N-glycosylation sites on the SERPIN domain. Further research is warranted to understand the structure and function of the N-glycosylation sites on the NTD.

Methods

Sex as a biological variable. As only 293T cells were used in this study, sex was not considered as a biological variable.

Preparation and expression of variants. The *SERPING1* pcDNA3.1 expression vector (Genescript, NJ, USA) was used to create the C1-INH variants. The variants were produced using the QuikChange XL site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA). Each *SERPING1* cDNA clone was sequenced. The variants were transient transfected into 293T cells (ATCC, CRL-3216) using the Xfect reagent (Takara Bio USA, CA) where Dulbecco's

Modified Eagle Medium (DMEM) was replaced with OptiMEM® (Invitrogen, NY). Each transfection experiment was conducted in three independent biological replicates. Supernatants were collected after 48 h, concentrated 40× and then stored in aliquots at -80°C (20).

Quantification and Western blotting (WB). The quantity of each recombinant C1-INH variant protein was determined by ELISA according to the manufacturer's recommendations (Abcam, Cambridge, MA). Electrophoretic patterns were evaluated and compared to WT using transfectant supernatants that were analyzed under reducing conditions using 4-20% SDS-PAGE, transferred to nitrocellulose and then probed with 1:1,000 rabbit anti-human C1-INH mAb (Abcam, AB134918) as the primary antibody, specifically recognizing the NTD of C1-INH between amino acids 22 to 100, followed by a 1:10,000 horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (Abcam, AB205718).

*C1s̄, FXIIa and PKa binding assays***.**

C1s̄ . The binding of C1s̄ to C1-INH was measured according to the manufacturer's instruction (Quidel, San Diego, CA). In the first step, standards, controls and C1-INH variants were incubated with biotinylated C1s̄. Next, the incubation mixtures containing the C1-INH-C1s̄ complex were added to microtiter wells precoated with streptavidin. After incubation, the wells were washed 3× to remove unbound protein. Then, the goat anti-human C1-INH was added to each test well to bind with the C1-INH-C1s̄ complex captured on the surface of the streptavidin-coated microtiter wells. After washing, the HRP-conjugated goat anti-human Ab was added to each microassay well. The HRP-conjugate reacted with the C1-INH-C1s̄ complex. After adding 3, 3', 5, 5'- tetramethylbenzidine dihydrochloride (TMB) substrate, the complex generated a yellow color. The intensity of the color reaction mixture was measured spectrophotometrically at 450 nm.

493

494 *FXIIa and PKa.* Purified human FXIIa and PKa were obtained from Enzyme Research
495 Laboratories (South Bend, IN, USA). The biotinylation kit and streptavidin-coated plate were
496 obtained from Thermo Scientific (Rockford, IL, USA). FXIIa and PKa were biotinylated
497 according to the manufacturer's recommendations. In brief, FXIIa and PKa were dissolved in
498 phosphate-buffered saline (PBS) and then incubated with Sulfo-NHS-LC-biotin on ice for 2 h.
499 Excess nonreacted and hydrolyzed biotin was removed through a spin column and the biotin-
500 labeled protein concentrations were measured by NanoDrop. ELISA was utilized to assess the
501 binding of WT and variant C1-INH to FXIIa and PKa, as described previously(21). Biotinylated
502 FXIIa and PKa (25 μl of 2 μg/ml), an equivalent molar amount of C1-INH, and 50 μl of reaction
503 buffer (2% BSA in PBS-T) were added to the streptavidin-coated plate and incubated for 1 h at
504 37° C. The bound PKa-C1-INH or FXIIa-C1-INH complex was detected by 1:10,000 mouse
505 anti-C1-INH mAb (Abcam, Cambridge, MA), followed by incubation at RT for 1 h. After
506 incubation, the plates were washed 3× using PBS-Tween (300 μl/each). A 1:10,000 dilution of
507 HRP-conjugated goat anti-rabbit IgG (Abcam, Cambridge, MA) secondary Ab was added and
508 then incubated at 37° C for 1 h. The detection of bound C1-INH complex was carried out as
509 described previously (21). On at least three occasions, binding assays were performed
510 employing serially diluted samples. C1s̄, FXIIa and PKa binding assays were repeated a
511 minimum of 3× for each variant.
512
513 *Molecular modeling.* PyMOL Version 3.0 was employed to visualize and analyze the
514 protein structures.
515

Data availability. Values for all data points in graphs are reported in the Supporting Data Values file. All gel data and WBs in this study are presented in the full, unedited gel file available in the Supplemental Data with full annotations.

Statistical analyses. Statistical analyses were performed using Prism 10 (GraphPad, San Diego, USA). Comparisons between two groups were assessed using paired t-test (non-parametric). Comparisons among groups were performed using a 1-way ANOVA with Dunnett's multiple comparison test ($P < 0.05$ was considered as significant). The relative absorbance was calculated as the absorbance for each C1-INH variant divided by the absorbance of WT at protein concentrations of 125 ng/ml, 250 ng/ml, 500 ng/ml and 1 μg/ml.

Author contributions: ZR and JPA conceived and designed the experiments, analyzed the data and interpreted results. ZR and JB performed the experiments. ZR conducted the structural analysis. SZ classified variants in accordance with ACMG 2015 criteria. ZR prepared the figures and drafted the manuscript. ZR, JB, SZ, NP, JW and JPA edited the manuscript.

Acknowledgments

This publication has been supported by the Division of Allergy and Immunology in the Department of Medicine at the Washington University School of Medicine, the National Institutes of Health (NIH) under award #R35 GM136352-01 (JA) and the Washington University Institute of Clinical and Translational Sciences which is, in part, supported by the NIH/National Center for Advancing Translational Sciences (NCATS), CTSA grant #UL1TR002345. The authors thank Kathryn (Kathy) Liszewski (Washington University School of Medicine) for their helpful suggestions regarding structural analysis, experimental design and manuscript preparation. The manuscript was edited by InPrint: A Scientific Communication Network at Washington University in St. Louis.

References

1. Zuraw BL. Clinical practice. Hereditary angioedema. *N Engl J Med.* 2008;359(10):1027-36.
2. Bissler JJ, Aulak KS, Donaldson VH, Rosen FS, Cicardi M, Harrison RA, et al. Molecular defects in hereditary angioneurotic edema. *Proc Assoc Am Physicians.* 1997;109(2):164-73.
3. Bock SC, Skriver K, Nielsen E, Thogersen HC, Wiman B, Donaldson VH, et al. Human C1 inhibitor: primary structure, cDNA cloning, and chromosomal localization. *Biochemistry.* 1986;25(15):4292-301.
4. Pappalardo E, Zingale LC, and Cicardi M. C1 inhibitor gene expression in patients with hereditary angioedema: quantitative evaluation by means of real-time RT-PCR. *J Allergy Clin Immunol.* 2004;114(3):638-44.
5. Ryo LB, Haslund D, Rovsing AB, Pihl R, Sanrattana W, de Maat S, et al. Restriction of C1-inhibitor activity in hereditary angioedema by dominant-negative effects of disease-associated SERPING1 gene variants. *J Allergy Clin Immunol.* 2023;152(5):1218-36 e9.
6. Ren Z, Zhao S, Li T, Wedner HJ, and Atkinson JP. Insights into the pathogenesis of hereditary angioedema using genetic sequencing and recombinant protein expression analyses. *J Allergy Clin Immunol.* 2023;151(4):1040-9 e5.
7. Guan X, Sheng Y, Liu S, He M, Chen T, and Zhi Y. Epidemiology, economic, and humanistic burden of hereditary angioedema: a systematic review. *Orphanet J Rare Dis.* 2024;19(1):256.

8. Dijk M, Holkers J, Voskamp P, Giannetti BM, Waterreus WJ, van Veen HA, et al. How Dextran Sulfate Affects C1-inhibitor Activity: A Model for Polysaccharide Potentiation. *Structure.* 2016;24(12):2182-9.
9. Garrigues RJ, Garrison MP, and Garcia BL. The Crystal Structure of the Michaelis-Menten Complex of C1 Esterase Inhibitor and C1s Reveals Novel Insights into Complement Regulation. *J Immunol.* 2024;213(5):718-29.
10. Davis AE, 3rd, Mejia P, and Lu F. Biological activities of C1 inhibitor. *Mol Immunol.* 2008;45(16):4057-63.
11. Beinrohr L, Harmat V, Dobo J, Lorincz Z, Gal P, and Zavodszky P. C1 inhibitor serpin domain structure reveals the likely mechanism of heparin potentiation and conformational disease. *J Biol Chem.* 2007;282(29):21100-9.
12. Cummings RD. *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine* Berlin, Heidelberg: Springer; 2005:704–18.
13. Bause E. Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. *Biochem J.* 1983;209(2):331-6.
14. Stavenhagen K, Kayili HM, Holst S, Koeleman CAM, Engel R, Wouters D, et al. N- and O-glycosylation analysis of human C1-inhibitor reveals extensive mucin-type O-glycosylation. *Mol Cell Proteomics.* 2018;17(6):1225-38.
15. Helenius A, and Aebi M. Intracellular functions of N-linked glycans. *Science.* 2001;291(5512):2364-9.
16. Eldering E, Huijbregts CC, Lubbers YT, Longstaff C, and Hack CE. Characterization of recombinant C1 inhibitor P1 variants. *J Biol Chem.* 1992;267(10):7013-20.
17. Valderrama-Rincon JD, Fisher AC, Merritt JH, Fan YY, Reading CA, Chhiba K, et al. An engineered eukaryotic protein glycosylation pathway in Escherichia coli. *Nat Chem Biol.* 2012;8(5):434-6.

18. Lowenthal MS, Davis KS, Formolo T, Kilpatrick LE, and Phinney KW. Identification of novel N-glycosylation sites at noncanonical protein consensus motifs. *J Proteome Res.* 2016;15(7):2087-101.

19. Lizak C, Gerber S, Numao S, Aebi M, and Locher KP. X-ray structure of a bacterial oligosaccharyltransferase. *Nature.* 2011;474(7351):350-5.

20. Ren Z, Perkins SJ, Love-Gregory L, Atkinson JP, and Java A. Clinicopathologic Implications of Complement Genetic Variants in Kidney Transplantation. *Front Med (Lausanne).* 2021;8:775280.

21. Joseph K, Bains S, Tholanikunnel BG, Bygum A, Aabom A, Koch C, et al. A novel assay to diagnose hereditary angioedema utilizing inhibition of bradykinin-forming enzymes. *Allergy.* 2015;70(1):115-9.

22. Kajdacsi E, Jandrasics Z, Veszeli N, Mako V, Koncz A, Gulyas D, et al. Patterns of C1-Inhibitor/Plasma Serine Protease Complexes in Healthy Humans and in Hereditary Angioedema Patients. *Front Immunol.* 2020;11:794.

23. Gosswein T, Kocot A, Emmert G, Kreuz W, Martinez-Saguer I, Aygoren-Pursun E, et al. Mutational spectrum of the C1INH (SERPING1) gene in patients with hereditary angioedema. *Cytogenet Genome Res.* 2008;121(3-4):181-8.

24. Guryanova I, Suffritti C, Parolin D, Zanichelli A, Ishchanka N, Polyakova E, et al. Hereditary angioedema due to C1 inhibitor deficiency in Belarus: epidemiology, access to diagnosis and seven novel mutations in SERPING1 gene. *Clin Mol Allergy.* 2021;19(1):3.

25. Nita-Lazar M, Wacker M, Schegg B, Amber S, and Aebi M. The N-X-S/T consensus sequence is required but not sufficient for bacterial N-linked protein glycosylation. *Glycobiology.* 2005;15(4):361-7.

26. Förster T. *Natural Science*. The Bavarian Julius-Maximilians University of Wurzburg; 2006:114.

27. Shrimal S, Cherepanova NA, and Gilmore R. Cotranslational and posttranslocational N-glycosylation of proteins in the endoplasmic reticulum. *Semin Cell Dev Biol.* 2015;41:71-8.
28. Imai K, and Mitaku S. Mechanisms of secondary structure breakers in soluble proteins. *Biophysics (Nagoya-shi).* 2005;1:55-65.
29. Bordo D, and Argos P. Suggestions for "safe" residue substitutions in site-directed mutagenesis. *J Mol Biol.* 1991;217(4):721-9.
30. Bissler JJ, Cicardi M, Donaldson VH, Gatenby PA, Rosen FS, Sheffer AL, et al. A cluster of mutations within a short triplet repeat in the C1 inhibitor gene. *Proc Natl Acad Sci U S A.* 1994;91(20):9622-5.
31. Grivceva-Panovska V, Kosnik M, Korosec P, Andrejevic S, Karadza-Lapic L, and Rijavec M. Hereditary angioedema due to C1-inhibitor deficiency in Macedonia: clinical characteristics, novel SERPING1 mutations and genetic factors modifying the clinical phenotype. *Ann Med.* 2018;50(3):269-76.
32. Loules G, Zamanakou M, Parsopoulou F, Vatsiou S, Psarros F, Csuka D, et al. Targeted next-generation sequencing for the molecular diagnosis of hereditary angioedema due to C1-inhibitor deficiency. *Gene.* 2018;667:76-82.
33. Verpy E, Biasotto M, Brai M, Misiano G, Meo T, and Tosi M. Exhaustive mutation scanning by fluorescence-assisted mismatch analysis discloses new genotype-phenotype correlations in angiodema. *Am J Hum Genet.* 1996;59(2):308-19.
34. Loli-Ausejo D, Lopez-Lera A, Drouet C, Lluncor M, Phillips-Angles E, Pedrosa M, et al. In search of an association between genotype and phenotype in hereditary angioedema due to C1-INH deficiency. *Clin Rev Allergy Immunol.* 2021;61(1):1-14.
35. Malaby HL, and Kobertz WR. Molecular determinants of co- and post-translational N-glycosylation of type I transmembrane peptides. *Biochem J.* 2013;453(3):427-34.

36. Sidrauski C, and Walter P. The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. *Cell.* 1997;90(6):1031-9.
37. Malaby HL, and Kobertz WR. The middle X residue influences cotranslational N-glycosylation consensus site skipping. *Biochemistry.* 2014;53(30):4884-93.
38. Ruiz-Canada C, Kelleher DJ, and Gilmore R. Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. *Cell.* 2009;136(2):272-83.
39. Adams CJ, Kopp MC, Larburu N, Nowak PR, and Ali MMU. Structure and molecular mechanism of ER stress signaling by the unfolded protein response signal activator IRE1. *Front Mol Biosci.* 2019;6:11.
40. Molinari M. N-glycan structure dictates extension of protein folding or onset of disposal. *Nat Chem Biol.* 2007;3(6):313-20.
41. Haukedal H, and Freude KK. Implications of glycosylation in Alzheimer's disease. *Front Neurosci.* 2020;14:625348.
42. Shrimal S, and Gilmore R. Glycosylation of closely spaced acceptor sites in human glycoproteins. *J Cell Sci.* 2013;126(Pt 23):5513-23.
43. Bolt G, Kristensen C, and Steenstrup TD. Posttranslational N-glycosylation takes place during the normal processing of human coagulation factor VII. *Glycobiology.* 2005;15(5):541-7.
44. Rout PK, and Verma M. Post translational modifications of milk proteins in geographically diverse goat breeds. *Sci Rep.* 2021;11(1):5619.
45. van Doorn MB, Burggraaf J, van Dam T, Eerenberg A, Levi M, Hack CE, et al. A phase I study of recombinant human C1 inhibitor in asymptomatic patients with hereditary angioedema. *J Allergy Clin Immunol.* 2005;116(4):876-83.

46. Parad RB, Kramer J, Strunk RC, Rosen FS, and Davis AE, 3rd. Dysfunctional C1 inhibitor Ta: deletion of Lys-251 results in acquisition of an N-glycosylation site. *Proc Natl Acad Sci U S A.* 1990;87(17):6786-90.

47. Ponard D, Gaboriaud C, Charignon D, Ghannam A, Wagenaar-Bos IGA, Roem D, et al. SERPING1 mutation update: Mutation spectrum and C1 Inhibitor phenotypes. *Hum Mutat.* 2020;41(1):38-57.

48. Shah CH, Princic N, Evans KA, and Schultz BG. Real-world changes in costs over time among patients in the United States with hereditary angioedema on long-term prophylaxis with lanadelumab. *J Med Econ.* 2023;26(1):871-7.

49. Duran-Romana R, Houben B, De Vleeschouwer M, Louros N, Wilson MP, Matthijs G, et al. N-glycosylation as a eukaryotic protective mechanism against protein aggregation. *Sci Adv.* 2024;10(5):eadk8173.

50. Eisenack TJ, and Trentini DB. Ending a bad start: Triggers and mechanisms of co-translational protein degradation. *Front Mol Biosci.* 2022;9:1089825.

51. Sato T, Sako Y, Sho M, Momohara M, Suico MA, Shuto T, et al. STT3B-dependent posttranslational N-glycosylation as a surveillance system for secretory protein. *Mol Cell.* 2012;47(1):99-110.

Figure legends

Figure 1. Schematic representation of protein domains and N-glycan for C1-INH. (**A**) The N-terminal domain includes resides 23-134 (pink). The C-terminal SERPIN domain consists of residues 135-500 (blue). Signal peptide (SP) contains 22 residues and is included in the numbering. Branched symbols, N-glycosylation (n=7); black circles, verified O-glycosylation sites (n=7); white circles, potential O-glycosylation sites (n=7). N-glycosylation sites in the SERPIN domain are highlighted in red. Protease cleavage site, P1-P1′ highlighted in purple. NTD, amino-terminal domain; CTD, carboxyl-terminal domain. (Adapted from *The Complement Factsbook, Figure 23.1).* (**B**) Representation of the N-glycan structure on C1-INH. Based on a mass spectrometry study, the majority of the N-glycans on N238, N253 and N352 are biantennary, approximately 80%, with $HexNAc_4Hex_5NeuAc_2$ being the most abundant structure (14). *, the N272 glycosylation site has not been confirmed by mass spectrometry.

Figure 2. Illustration of key domains in C1-INH. The reactive center loop (RCL) is in pink. The regions involved in SERPIN function are labeled. The P15-P9 portion of the RCL, the hinge domain, is highly conserved and facilitates the insertion of RCL into the β sheets A (SA). The breach region lies on the top of SA, the initial insertion site of RCL. The shutter domain, composed of S3A and S5A, is in the center of SA and facilitates the RCL insertion. The gate region consists of strands 3 and 4 from β sheets C (S3C and S4C). N-glycosylation sites are shown as blue spheres. The P1 and P' are displayed in rainbowstick and are responsible for trapping the target protease. Strands of central SA are in cyan. The two disulfide bridges are labeled and colored in red. Protein structures used for modeling were obtained from the PDB database (PDB: 5DU3). The Figure was generated using Pymol (3.0) and serves as a model for structure analyses in Figures 4-7. NTD, amino-terminal domain; CTD, carboxyl-terminal domain.

Figure 3. Illustrations of N-glycans on the SERPIN domain in C1-INH and a C1$\bar{s}$/C1-INH complex. (**A**) Molecular model of N-glycans on the SERPIN domain in C1-INH (PDB: 5DU3). Molecular modeling was performed on the GlyCAM-Web tool, Glycoprotein Builder (https://glycam.org). The N-linked glycan structure was generated based on mass spectrometry results (14). The N-GlcNAc linkage conformation was based on the simulation generated from Glycoprotein Builder. (**B**) Structure of a C1$\bar{s}$/C1-INH complex (PDB: 8W18). The structure of active C1$\bar{s}$ is shown in a cartoon representation in rainbow. The RCL is in violet. The P1 R466 and P1′ T467 residues are displayed in rainbowstick and are responsible for trapping the C1$\bar{s}$. The disulfide bridges are labeled and colored in red. Recombinant C1$\bar{s}$ harbors a S632A mutation, making it catalytically inert (PDB: 8W18) (9). Figures were produced using PyMOL (3.0)

Figure 4. N238-glycosylation site variants. (**A**) Western blot **(**WB) analysis of supernatants from transfected wild type (WT) C1-INH and variant constructs of N238 under reducing conditions. The recombinant expression of N238A and A239D is comparable to WT (see Table 1). The secretion of N238del is markedly decreased compared with that of WT. The consensus sequence of N238- glycosylation NXS/T is highlighted in blue, orange and red. (**B**–**E**) Functional analysis of N238 variants. (**B** and **D**) Absorbance is plotted against protein concentration. (**C** and **E)** Relative absorbance (RA) is computed as the absorbance of the variant divided by the absorbance of the WT at concentrations of 1 μg/ml, 500 ng/ml, 250 ng/ml and 125 ng/ml, respectively. For PKa binding, the *P* value for the percentage differences of N238A and A239D compared to WT are 0.918 and 0.254, respectively. For FXIIa binding, the *P* value for the percentage differences of N238A and A239D compared to WT are 0.05 and 0.358, respectively. Data represent three separate experiments with bars corresponding to SEM. One-way ANOVA with Dunnett's multiple comparisons test was used. **F.** Structural

analysis of N238 (PDB: 5DU3). N238 is located on the surface of Helix E (hE). N238del results in the disruption of the consensus sequence NXS, which is required for the attachment of N-glycan. In the absence of glycosylation, N238del likely leads to protein misfolding.

Figure 5. N352-glycosylation site variants. (**A**) WB analysis of supernatants from transfected WT C1-INH and variant constructs of N352 under reducing conditions. One representative experiment of three is shown. Recombinant expression of N352A, N352S and S354G is comparable to WT, whereas N352del, L353P and S354del are barely secreted. The secretion of N352I was decreased (see Table 1). The consensus sequence of N352-glycosylation NXS/T is highlighted in blue, orange and red. (**B**–**E**) Binding analysis of N352 glycosylation site variants. (**B** and **D**) Representation of PKa and FXIIa binding of N352A, N352S, N352I and S354G compared to WT. (**C** and **E**) N352I and S354G demonstrate impaired binding to both PKa and FXIIa. N352S showed decreased binding to PKa but not to FXIIa. Data represents mean ± SEM of 3 independent experiments. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. One-way ANOVA with Dunnett's multiple comparisons test was used. (**F**) Structural analysis of N352. N352 is located in the loop connecting strands 2 and 3 in the β-sheet B (S2B and S3B), which is the hydrophobic core of C1-INH. N352del disrupts the packing of hydrophobic core and leads to protein misfolding. The substitution of L353P can cause structural disturbances by disrupting hydrogen bridges and affecting the packing of the loop between S2A and S3A, thus leading to protein misfolding. Hydrogen bond, yellow dash line.

Figure 6. N253-glycosylation site variants. (**A**) WB analysis of supernatants from transfected WT and N253 variant constructs under reducing conditions. Recombinant expression of N253 variants is comparable to WT, see Table 1. N253A, N253del, N254del and S255G disrupt the N-glycan attachment and lead to a slightly lower M_r protein compared with WT. The consensus

762 sequence of N253- glycosylation NXS/T is highlighted in blue, orange and red. It contains two
763 Ns in this sequence. (**B**) WB analysis of variants N254del and S255G before and after
764 treatment with glycosidases. Recombinant expression of N254del and S255G are comparable
765 to WT, but with a slightly lower M_r compared with WT (lane 1). After deglycosylation treatment,
766 WB demonstrates that N254del (lane 8), S255G (lane 9) and WT have the same M_r. Δ, post-
767 degylcosylation. (**C**–**F**) Functional analysis of the N253-glycosylation site variants. (**C** and **E**)
768 Absorbance is plotted against protein concentrations. (**D** and **F**) Binding affinity of N253-
769 glycosylation site variants with PKa and FXIIa compared to WT. The binding affinity of N253A,
770 N254A, S255del, S255G, S255T to PKa and FXIIa is comparable to WT. N253del and N254del
771 exhibit mildly impaired binding to PKa and FXIIa. Results shown are from 3 independent
772 experiments. Data represent mean ± SEM. Significances were calculated by 1-way ANOVA and
773 Dunnett's multiple comparisons, $*P < 0.05$, $***P < 0.001$, $****P < 0.0001$. (**G**) Structural
774 analysis of N253. The structure of active C1-INH is shown in a cartoon representation (PDB:
775 5DU3). N253 locates in the loop connecting strand 1 in the β-sheet A and Helix F.
776
777 **Figure 7. N272-glycosylation site variants.** (**A**) WB analysis of supernatants from transfected
778 WT and N272 variant constructs under reducing conditions. Recombinant expression of
779 N271del (lane 3) and N272del (lane 5) is decreased. N271A, N272A, N272D, N271-2del,
780 K273del and S275del (lanes 2, 4, 6, 7, 8 and 9) have normal secretion comparable to WT (lane
781 1), see Table 1. N272-linked glycan site has an atypical N-glycosylation consensus sequence,
782 NNKIS. The peptide sequence of N272- glycosylation NNKIS is highlighted in blue, purple,
783 orange and red. (**B**) WB analysis of variants N272del and K273del before and after treatment
784 with glycosidases. Before treatment, K273del (lane 3) has a slightly higher M_r compared with
785 N272del (lane 2). After treatment, WB demonstrates that N272del (lane 5) and K273del (lane 6)
786 have the same M_r. Human purified C1-INH is used as a positive control (lane 7, before
787 treatment; lane 8, after treatment). Adopted from *Insights into the pathogenesis of hereditary*

angioedema using genetic sequencing and recombinant protein expression analyses, by Ren et al. 2023 (6). Δ, post degylcosylation. (**C**–**F**) Functional analysis of the N272 glycosylation site variants. (**C and E**) Absorbance is plotted against protein concentrations. (**D and F**) Comparison of PKa and FXIIa binding between WT and N272 glycosylation site variants. The binding affinity of N271A, N272A, N272D to PKa and FXIIa is comparable to WT, whereas N271del, N272del, K273del and N271-N272del exhibit impaired binding activity to both substrates. Interestingly, S275del exhibits a markedly decreased binding to PKa but not to FXIIa. Data represent mean ± SEM of 3 separate experiments. ***$P < 0.001$, ****$P < 0.0001$. One-way ANOVA with Dunnett's multiple comparisons test was used. (**G**) Structural analysis of N272del and K273del. K273 is located in the loop immediately after hF. The deletion of K273 will affect the conformation of hF. K273del, previously reported, results in a new N-glycosylation site in C1-INH (37). The N271 residue is shown in pink; N272 in blue sphere and K273 in purple.

Figure 8. N-glycosylation at N238, N253 and N352 in the endoplasmic reticulum (ER). (**A**) STT3A transfers the oligosaccharide to N238 and N352 (NXS) in the nascent protein chain cotranslationally. STT3B transfers the oligosaccharide to the N253 sequon in a posttranslational manner. (**B**) Misfolded N352del triggers the cotranslational protein degradation (50). (**C**) With preserved protein structure, 253del, even without posttranslational modification, is able to pass the quality control system and be transported out of the ER.

Table legends

Table 1: N-glycosylation variants: expression and functional assessment. Wild type (WT) recombinant C1-INH concentration, 26.4 ± 3.0 μg/ml. Kallikrein (PKa), FXIIa and C1s̄ binding are measured by ELISA. *P* values are computed by using two-sided independent sample T-

814 tests and comparing the results to those for WT. [§] PKa and FXIIa binding, >75% normal; 50-
815 70% marginally decreased; <50% decreased (21). [†] C1$\bar{s}$ binding >67%, normal; 41%–67%,
816 equivocal; <41%, abnormal (6, 22). Data represent mean ± SEM of 3 separate experiments. *P
817 ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001. One-way ANOVA with Dunnett's multiple
818 comparisons test was used. [#], variants reported in HAE patients; N, no significant difference; N[‡],
819 increased expression; del, deletion; D, decreased; ND, not done; MD, marginally decreased.
820 SEM, standard error of the mean.

821 **Table 2. Clinical pathogenic implication of N-glycan variants in HAE patients**. The allele
822 frequency (AF) information is only reported for S255T, 0.0029%. ACMG, American College of
823 Medical Genetics; N, normal; D, decreased; ND, not detected; NA, not available; VUS, variant
824 of uncertain significance.
825

N-glycosylation Sequon	Variants Reported	Mutagenesis	Expression	Recombinant Secretion (µg/ml), mean ± SEM	Kallikrein (PKa) Binding§	FXIIa Binding§	C1s̄ Binding†
N238AS	A239D	N238A	N	29.2 ± 1.7	N	N	N (89.9% ± 1.7)
		N238del	D****	ND	ND	ND	ND
		A239D	N‡	34.0 ± 3.5	N‡	N	N (89.0% ± 1.3)
		S240del	D****	2.3 ± 0.5	ND	ND	D (13.3% ± 6.6)
N352LS	N352I# N352S L353P# S354G	N352A	N	21.6 ± 0.3	N	N	N (76.7 ± 2.7%)
		N352del	D****	0.32 ± 0.13	ND	ND	ND
		N352I#	D***	8.6 ± 1.6	D**	D****	N (78.0 ± 1.6%)
		N352S	N	30.7 ± 3.7	D**	N	N (82.3 ± 1.1%)
		S354del	D****	1.1 ± 0.2	ND	ND	ND
		L353P#	D****	ND	ND	ND	ND
		S354G	N	16.6 ± 1.3	D***	D****	N (86.6 ± 2.3%)
N253NS	S255G S255T #	N253A	N	24.6 ± 4.7	N	N	N (86.2 ± 0.1%)
		N253del	N‡	38.9 ± 2.8	MD*	D***	N (83.7 ± 2.7%)
		N254A	N	32.9 ± 1.8	N	N	N (86.1 ± 3.0%)
		N254del	N	30.4 ± 0.5	D****	D****	N (86.0 ± 1.4%)
		S255del	N	25.3 ± 0.4	N	N	N (88.8 ± 0.7%)
		S255G	N	22.4 ± 3.2	N	N	N (86.0 ± 3.7%)
		S255T #	N	20.0 ± 4.0	N	N	N (89.1 ± 1.1%)
NN272KIS	N272D N272del# N271-N272del# K273del#	N271A	N	22.3 ± 5.4	N	N	N (88.5 ± 5.7%)
		N271del	D***	8.3 ± 1.2	D****	D****	D (56.0 ± 7.9%)
		N272A	N	24.9 ± 4.2	N	N	N (83.0 ± 0.6%)
		N272del#	MD*	13.0 ± 1.2	D****	D****	D (51.3 ± 4.7%)
		N272D	N	24.0 ± 2.9	N	N	N (80.2 ± 2.4%)
		N271-N272del#	N	28.9 ± 1.6	D****	D****	D (29.6 ± 0.3%)
		K273del#	N	21.8 ± 3.2	D****	D****	D (36.4± 2.7%)
		S275del	N	28.1 ± 3.0	D****	MD***	N (68.4 ± 3.9%)

Table 1: N-glycosylation variants: expression and functional assessment. Wild type (WT) recombinant C1-INH concentration, 26.4 ± 3.0 µg/ml. Kallikrein (PKa), FXIIa and C1s̄ binding are measured by ELISA. *P* values are computed by using two-sided independent sample T-tests and comparing the results to those for WT. § PKa and FXIIa binding, >75% normal; 50-70% marginally decreased; <50% decreased (21). † C1s̄ binding >67%, normal; 41%–67%, equivocal; <41%, abnormal (6, 22). Data represent mean ± SEM of 3 separate experiments. *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$. One-way ANOVA with Dunnett's multiple comparisons test was used. #, variants reported in HAE patients; N, no significant difference; N‡, increased expression; del, deletion; D, decreased; ND, not done; MD, marginally decreased. SEM, standard error of the mean.

N-glycosylation Site	Variants	Assessment		Classification		Family History	Clinical Summary
		Expression	Function	ACMG Interpretation	Modified Interpretation		
N253	S255T	N	N	VUS	Likely benign	NA	**N_{253}NS** sequon likely utilizes a posttranslational glycosylation mechanism. The replacement of S with T at S255 site theoretically should not affect the glycol attachment at the N253. As S255T demonstrates, it has preserved expression and function and is likely a benign missense variant.
N352	N352I	D	N	VUS	Likely pathogenic	NA	**N_{352}LS** sequon are likely required for cotranslational N-glycan modification. The disruption at N352 glycan attachment results in a decreased protein expression and functional defects. N352I results in expression defects and is likely causing HAE. L353P demonstrates abolished expression, leading to HAE.
	L353P	ND	ND	Pathogenic	Pathogenic	Yes	
N272	N272del	D	D	Pathogenic	Pathogenic	Yes	K273del gains an additional N-glycan at the **N_{272}IS** site, which leads to a dysfunctional protein. N271 and N272 positions are conserved among SERPINs (Supplemental Figure 1). The deletion of N271, N272 or both are deleterious.
	N271_ N272del	N	D	Likely pathogenic	Likely pathogenic	Yes	
	K273del	N	D	Likely pathogenic	Pathogenic	Yes	

Table 2. Clinical pathogenic implication of N-glycan variants in HAE patients. The allele frequency (AF) information is only reported for S255T, 0.0029%. ACMG, American College of Medical Genetics; N, normal; D, decreased; ND, not detected; NA, not available; VUS, variant of uncertain significance.

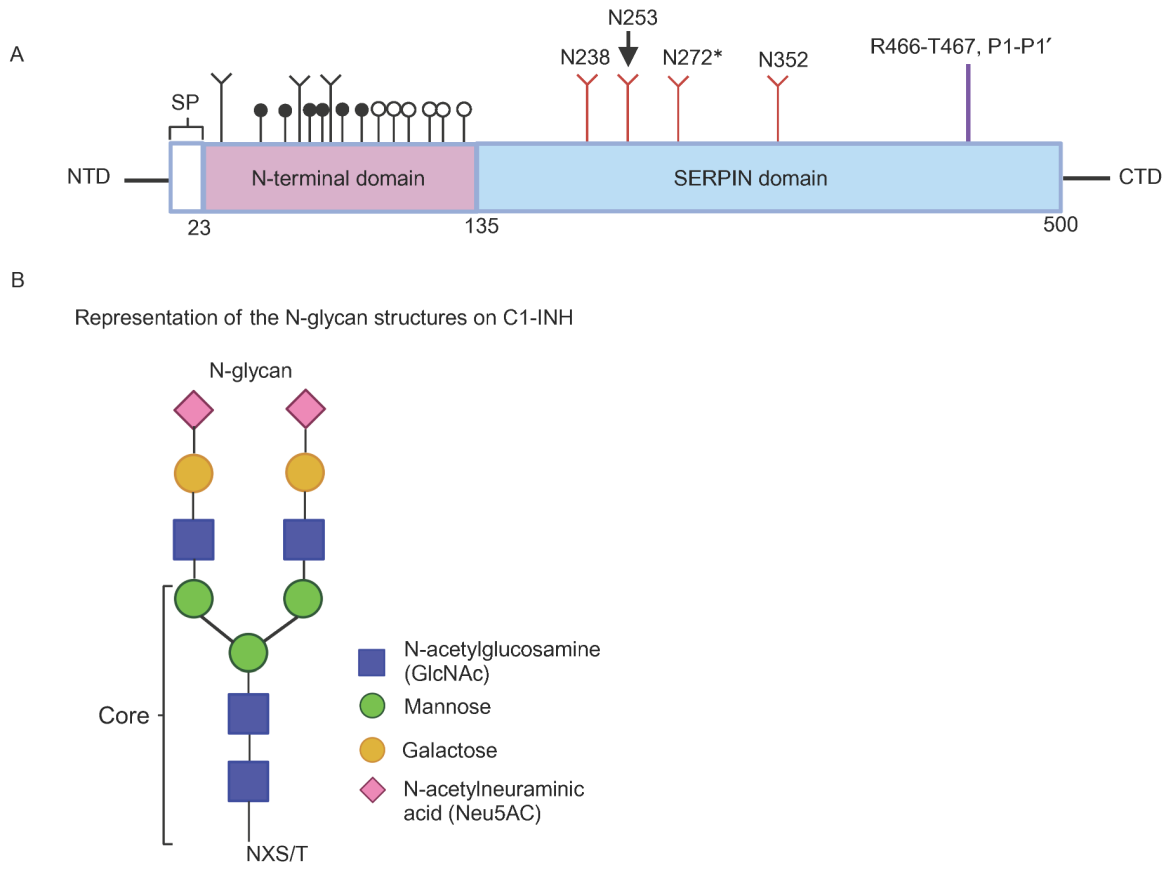


Figure 1. Schematic representation of protein domains and N-glycan for C1-INH. (**A**) The N-terminal domain includes resides 23-134 (pink). The C-terminal SERPIN domain consists of residues 135-500 (blue). Signal peptide (SP) contains 22 residues and is included in the numbering. Branched symbols, N-glycosylation (n=7); black circles, verified O-glycosylation sites (n=7); white circles, potential O-glycosylation sites (n=7). N-glycosylation sites in the SERPIN domain are highlighted in red. Protease cleavage site, P1-P1′ highlighted in purple. NTD, amino-terminal domain; CTD, carboxyl-terminal domain. (Adapted from *The Complement Factsbook, Figure 23.1).* (**B**) Representation of the N-glycan structure on C1-INH. Based on a mass spectrometry study, the majority of the N-glycans on N238, N253 and N352 are biantennary, approximately 80%, with $HexNAc_4Hex_5NeuAc_2$ being the most abundant structure (14). *, the N272 glycosylation site has not been confirmed by mass spectrometry.

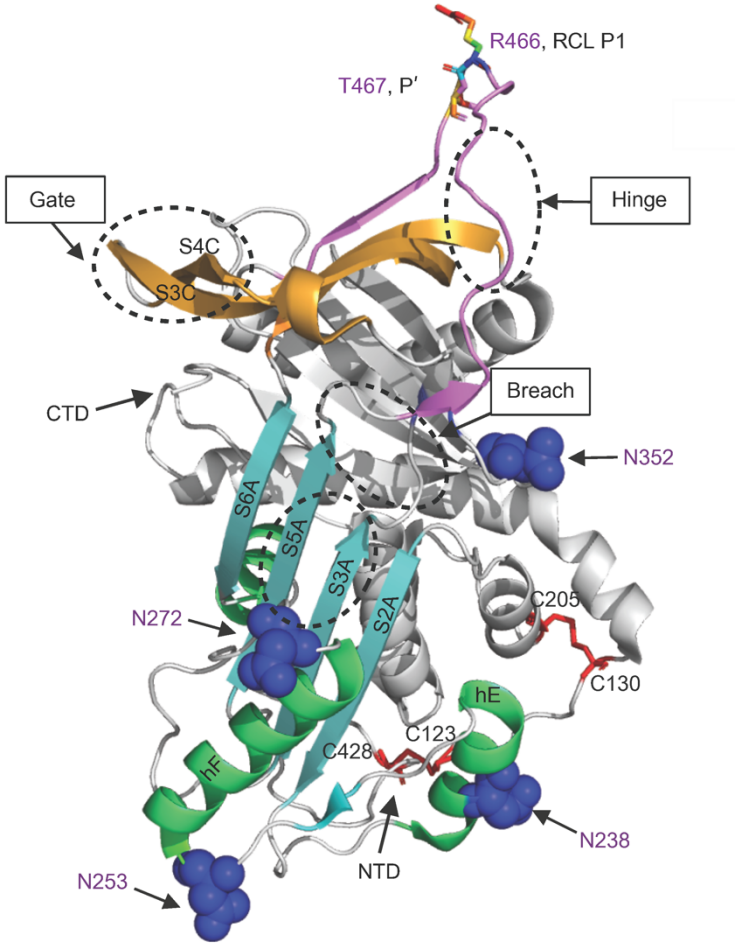
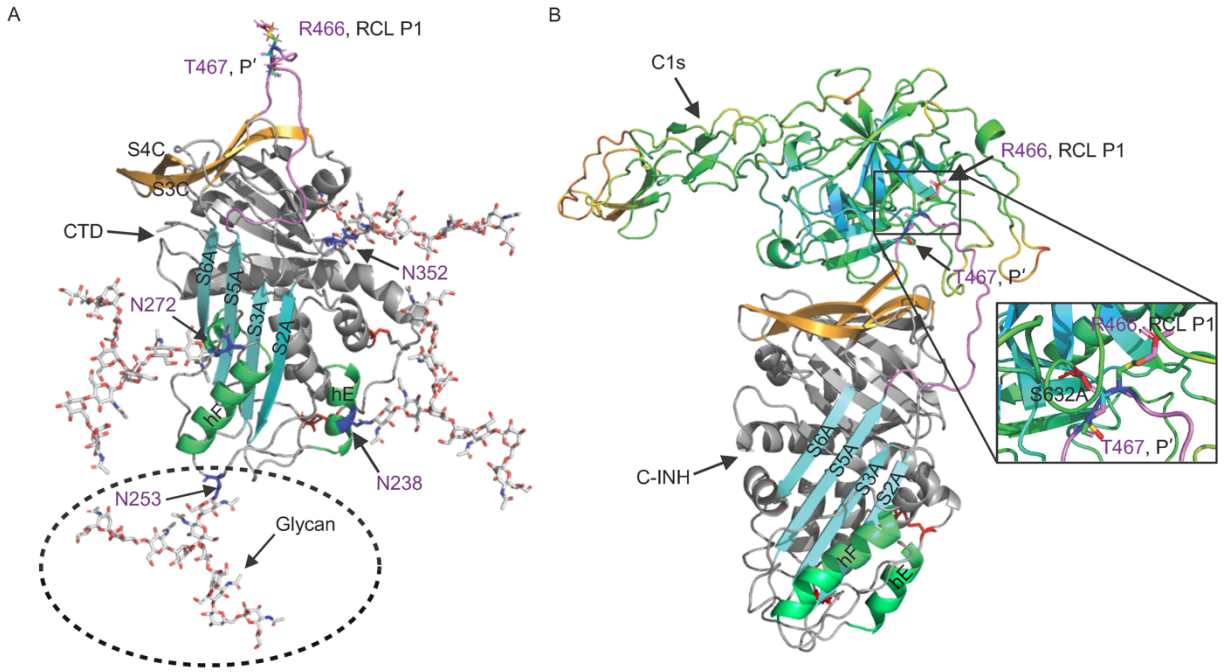


Figure 2. Illustration of key domains in C1-INH. The reactive center loop (RCL) is in pink. The regions involved in SERPIN function are labeled. The P15-P9 portion of the RCL, the hinge domain, is highly conserved and facilitates the insertion of RCL into the β sheets A (SA). The breach region lies on the top of SA, the initial insertion site of RCL. The shutter domain, composed of S3A and S5A, is in the center of SA and facilitates the RCL insertion. The gate region consists of strands 3 and 4 from β sheets C (S3C and S4C). N-glycosylation sites are shown as blue spheres. The P1 and P' are displayed in rainbowstick and are responsible for trapping the target protease. Strands of central SA are in cyan. The two disulfide bridges are labeled and colored in red. Protein structures used for modeling were obtained from the PDB database (PDB: 5DU3). The Figure was generated using Pymol (3.0) and serves as a model for structure analyses in Figures 4-7. NTD, amino-terminal domain; CTD, carboxyl-terminal domain.



879

880 **Figure 3. Illustrations of N-glycans on the SERPIN domain in C1-INH and a C1s̄/C1-INH**

881 **complex.** (**A**) Molecular model of N-glycans on the SERPIN domain in C1-INH (PDB: 5DU3).

882 Molecular modeling was performed on the GlyCAM-Web tool, Glycoprotein Builder

883 (https://glycam.org). The N-linked glycan structure was generated based on mass spectrometry

884 results (14). The N-GlcNAc linkage conformation was based on the simulation generated from

885 Glycoprotein Builder. (**B**) Structure of a C1s̄/C1-INH complex (PDB: 8W18). The structure of

886 active C1s̄ is shown in a cartoon representation in rainbow. The RCL is in violet. The P1 R466

887 and P1′ T467 residues are displayed in rainbowstick and are responsible for trapping the C1s̄.

888 The disulfide bridges are labeled and colored in red. Recombinant C1s̄ harbors a S632A

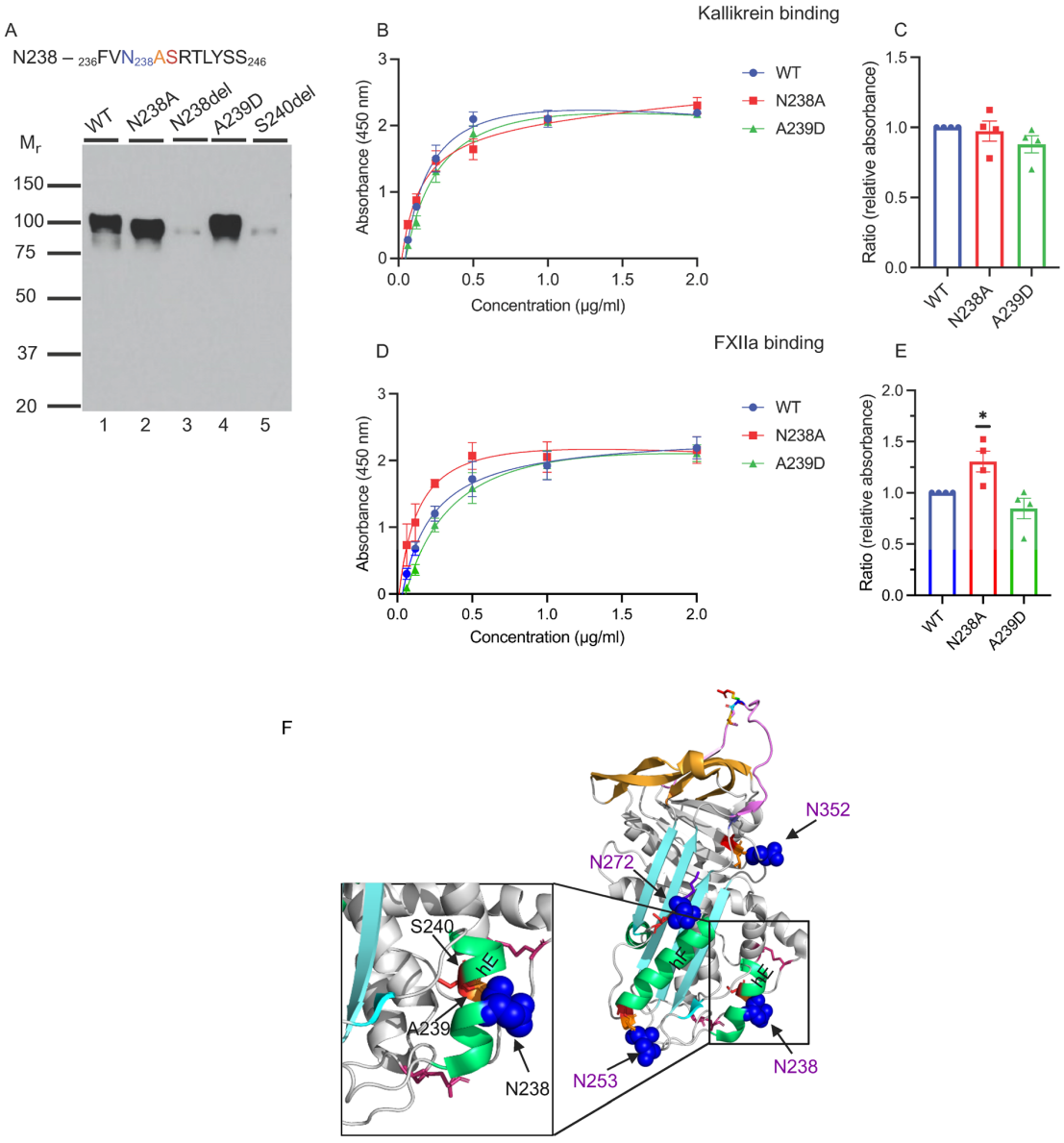
889 mutation, making it catalytically inert (PDB: 8W18) (9). Figures were produced using PyMOL

890 (3.0)

891

892

893



894

895 **Figure 4. N238-glycosylation site variants.** (**A**) Western blot **(**WB) analysis of supernatants

896 from transfected wild type (WT) C1-INH and variant constructs of N238 under reducing

897 conditions. The recombinant expression of N238A and A239D is comparable to WT (see Table

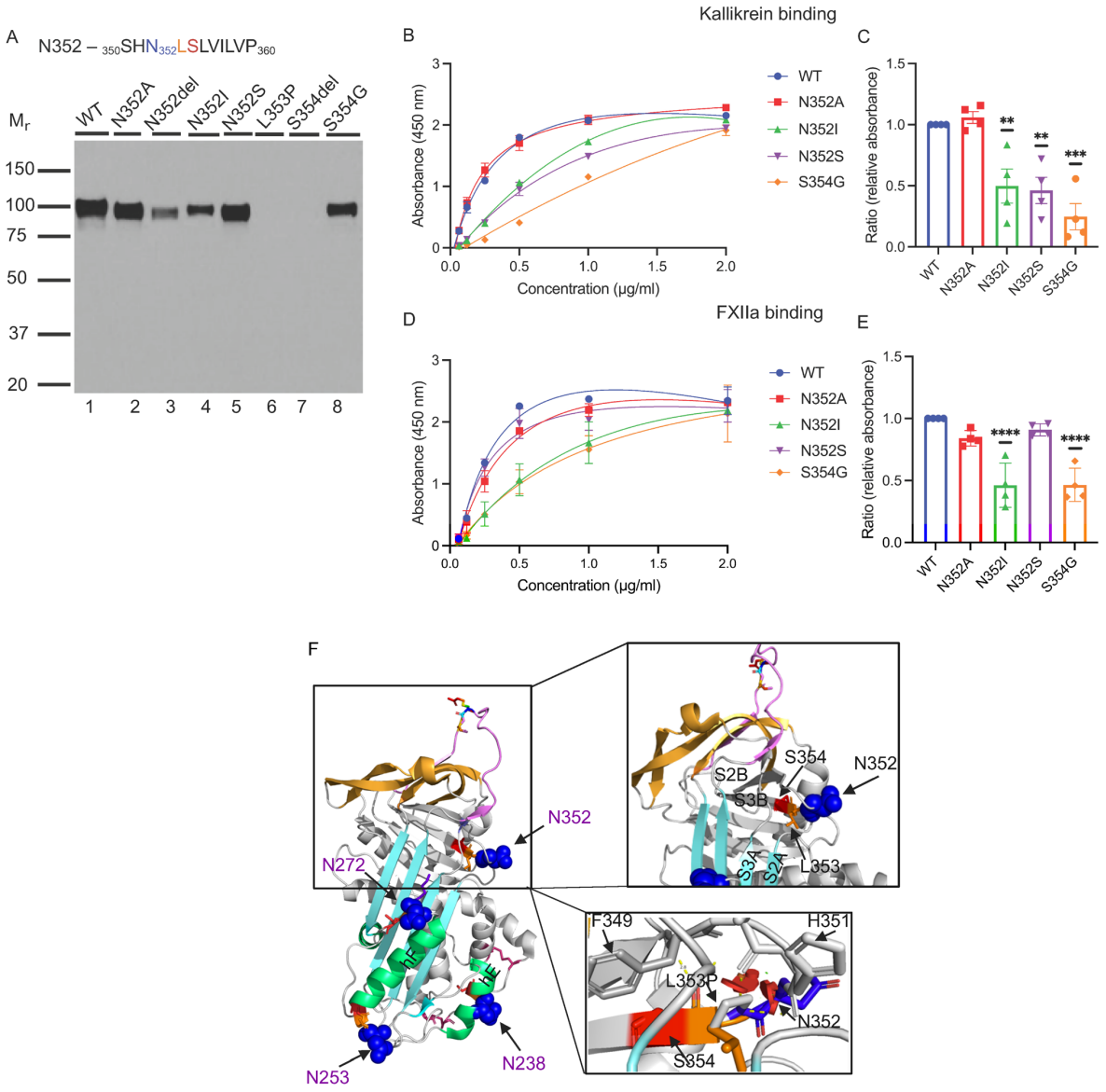
898 1). The secretion of N238del is markedly decreased compared with that of WT. The consensus

899 sequence of N238- glycosylation NXS/T is highlighted in blue, orange and red. (**B**–**E**)

900 Functional analysis of N238 variants. (**B** and **D**) Absorbance is plotted against protein

901 concentration. (**C** and **E)** Relative absorbance (RA) is computed as the absorbance of the

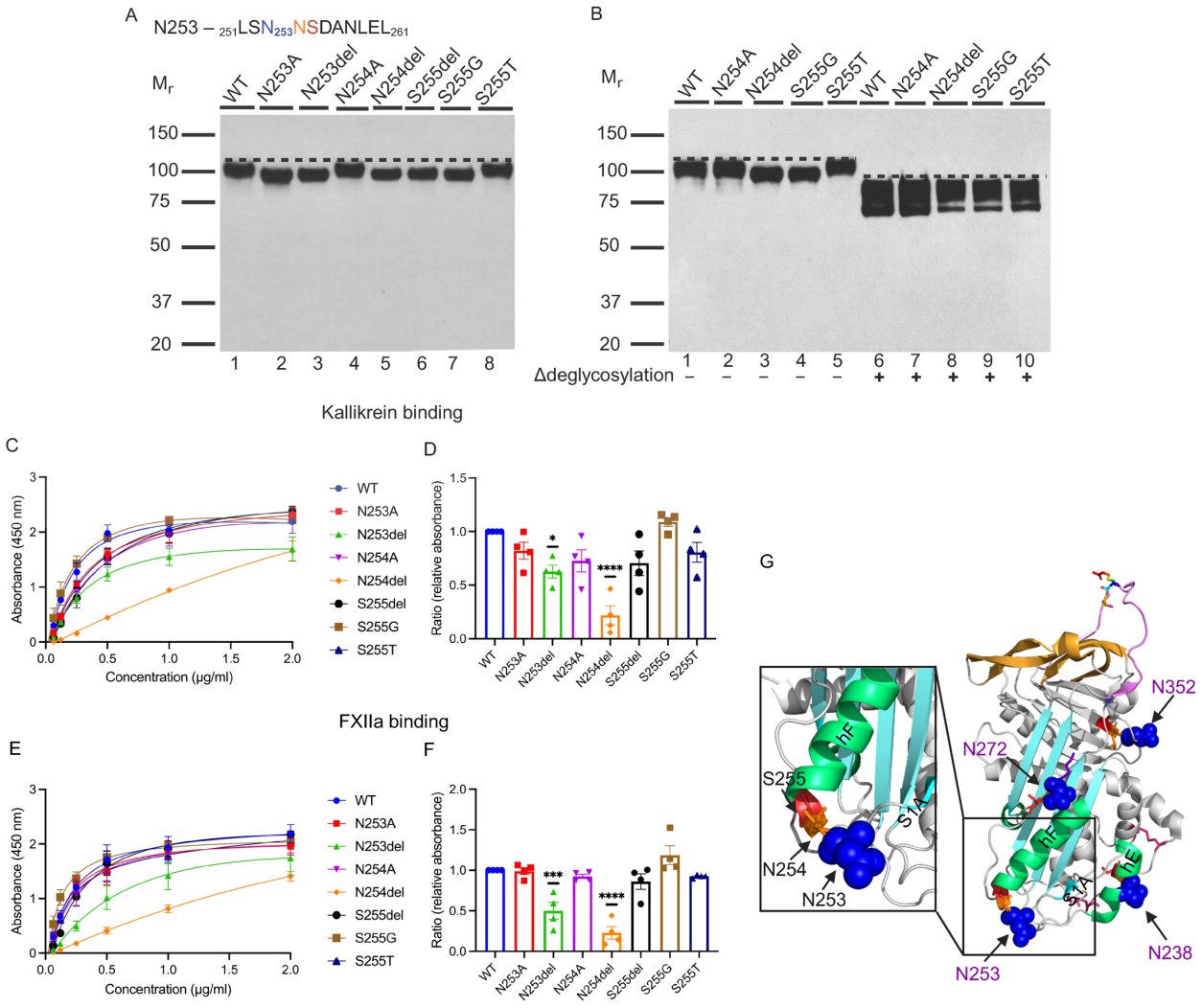
902 variant divided by the absorbance of the WT at concentrations of 1 μg/ml, 500 ng/ml, 250
903 ng/ml and 125 ng/ml, respectively. For PKa binding, the *P* value for the percentage differences
904 of N238A and A239D compared to WT are 0.918 and 0.254, respectively. For FXIIa binding, the
905 *P* value for the percentage differences of N238A and A239D compared to WT are 0.05 and
906 0.358, respectively. Data represent three separate experiments with bars corresponding to
907 SEM. One-way ANOVA with Dunnett's multiple comparisons test was used. **F.** Structural
908 analysis of N238 (PDB: 5DU3). N238 is located on the surface of Helix E (hE). N238del results
909 in the disruption of the consensus sequence NXS, which is required for the attachment of N-
910 glycan. In the absence of glycosylation, N238del likely leads to protein misfolding.
911



912

913 **Figure 5. N352-glycosylation site variants.** (**A**) WB analysis of supernatants from transfected
914 WT C1-INH and variant constructs of N352 under reducing conditions. One representative
915 experiment of three is shown. Recombinant expression of N352A, N352S and S354G is
916 comparable to WT, whereas N352del, L353P and S354del are barely secreted. The secretion of
917 N352I was decreased (see Table 1). The consensus sequence of N352-glycosylation NXS/T is
918 highlighted in blue, orange and red. (**B**–**E**) Binding analysis of N352 glycosylation site variants.
919 (**B** and **D**) Representation of PKa and FXIIa binding of N352A, N352S, N352I and S354G

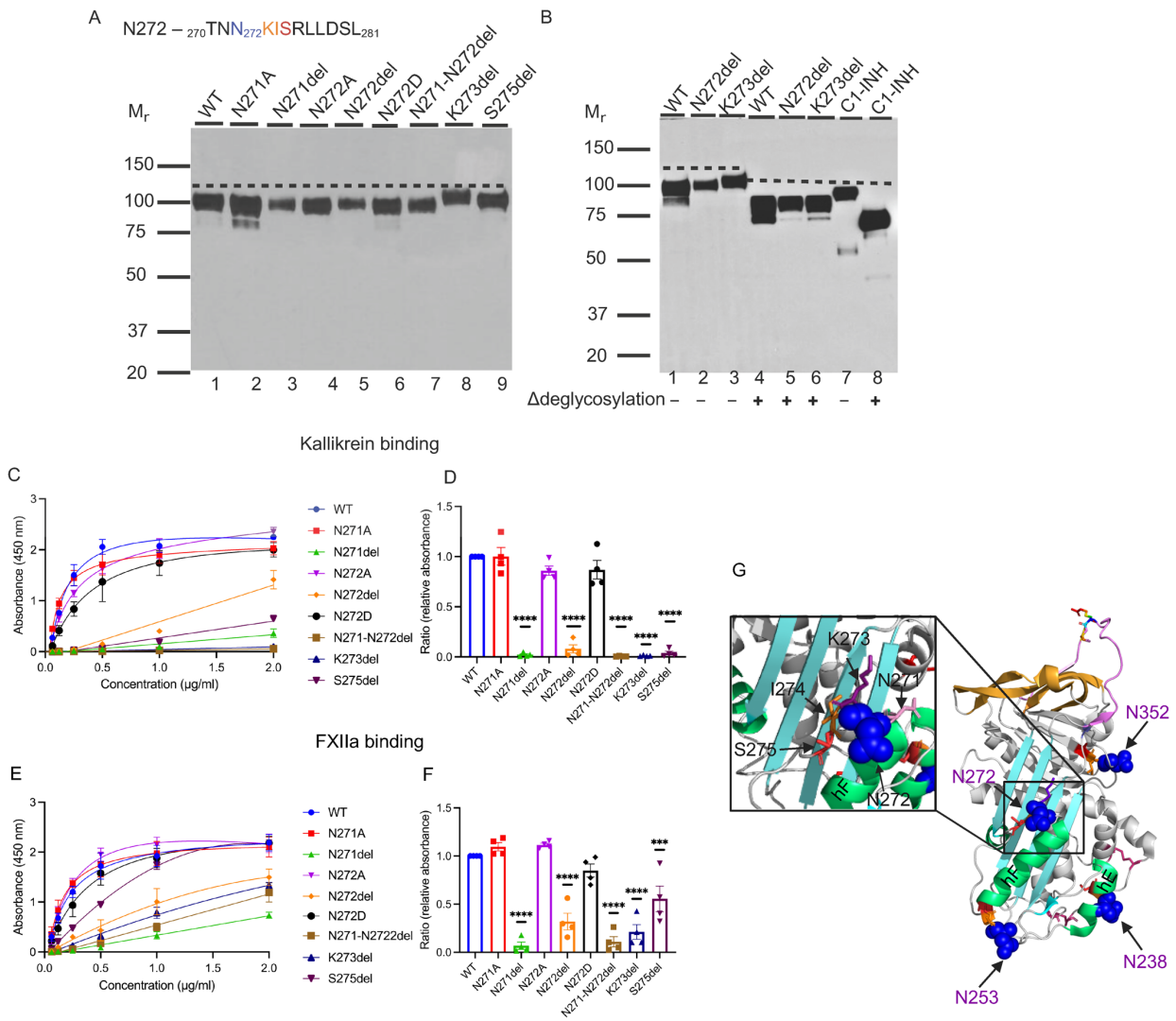
compared to WT. (**C** and **E**) N352I and S354G demonstrate impaired binding to both PKa and FXIIa. N352S showed decreased binding to PKa but not to FXIIa. Data represents mean ± SEM of 3 independent experiments. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. One-way ANOVA with Dunnett's multiple comparisons test was used. (**F**) Structural analysis of N352. N352 is located in the loop connecting strands 2 and 3 in the β-sheet B (S2B and S3B), which is the hydrophobic core of C1-INH. N352del disrupts the packing of hydrophobic core and leads to protein misfolding. The substitution of L353P can cause structural disturbances by disrupting hydrogen bridges and affecting the packing of the loop between S2A and S3A, thus leading to protein misfolding. Hydrogen bond, yellow dash line.



929

930 **Figure 6. N253-glycosylation site variants.** (**A**) WB analysis of supernatants from transfected
931 WT and N253 variant constructs under reducing conditions. Recombinant expression of N253
932 variants is comparable to WT, see Table 1. N253A, N253del, N254del and S255G disrupt the
933 N-glycan attachment and lead to a slightly lower M_r protein compared with WT. The consensus
934 sequence of N253- glycosylation NXS/T is highlighted in blue, orange and red. It contains two
935 Ns in this sequence. (**B**) WB analysis of variants N254del and S255G before and after
936 treatment with glycosidases. Recombinant expression of N254del and S255G are comparable

to WT, but with a slightly lower M_r compared with WT (lane 1). After deglycosylation treatment, WB demonstrates that N254del (lane 8), S255G (lane 9) and WT have the same M_r. Δ, post-degylcosylation. (**C**–**F**) Functional analysis of the N253-glycosylation site variants. (**C** and **E**) Absorbance is plotted against protein concentrations. (**D** and **F**) Binding affinity of N253-glycosylation site variants with PKa and FXIIa compared to WT. The binding affinity of N253A, N254A, S255del, S255G, S255T to PKa and FXIIa is comparable to WT. N253del and N254del exhibit mildly impaired binding to PKa and FXIIa. Results shown are from 3 independent experiments. Data represent mean ± SEM. Significances were calculated by 1-way ANOVA and Dunnett's multiple comparisons, $*P < 0.05$, $***P < 0.001$, $****P < 0.0001$. (**G**) Structural analysis of N253. The structure of active C1-INH is shown in a cartoon representation (PDB: 5DU3). N253 locates in the loop connecting strand 1 in the β-sheet A and Helix F.



948
949 **Figure 7. N272-glycosylation site variants.** (**A**) WB analysis of supernatants from transfected
950 WT and N272 variant constructs under reducing conditions. Recombinant expression of
951 N271del (lane 3) and N272del (lane 5) is decreased. N271A, N272A, N272D, N271-2del,
952 K273del and S275del (lanes 2, 4, 6, 7, 8 and 9) have normal secretion comparable to WT (lane
953 1), see Table 1. N272-linked glycan site has an atypical N-glycosylation consensus sequence,
954 NNKIS. The peptide sequence of N272- glycosylation NNKIS is highlighted in blue, purple,
955 orange and red. (**B**) WB analysis of variants N272del and K273del before and after treatment
956 with glycosidases. Before treatment, K273del (lane 3) has a slightly higher M_r compared with

N272del (lane 2). After treatment, WB demonstrates that N272del (lane 5) and K273del (lane 6) have the same $M_{r.}$ Human purified C1-INH is used as a positive control (lane 7, before treatment; lane 8, after treatment). Adopted from *Insights into the pathogenesis of hereditary angioedema using genetic sequencing and recombinant protein expression analyses*, by Ren et al. 2023 (6). Δ, post degylcosylation. (**C**–**F**) Functional analysis of the N272 glycosylation site variants. (**C and E**) Absorbance is plotted against protein concentrations. (**D and F**) Comparison of PKa and FXIIa binding between WT and N272 glycosylation site variants. The binding affinity of N271A, N272A, N272D to PKa and FXIIa is comparable to WT, whereas N271del, N272del, K273del and N271-N272del exhibit impaired binding activity to both substrates. Interestingly, S275del exhibits a markedly decreased binding to PKa but not to FXIIa. Data represent mean ± SEM of 3 separate experiments. ***$P < 0.001$, ****$P < 0.0001$. One-way ANOVA with Dunnett's multiple comparisons test was used. (**G**) Structural analysis of N272del and K273del. K273 is located in the loop immediately after hF. The deletion of K273 will affect the conformation of hF. K273del, previously reported, results in a new N-glycosylation site in C1-INH (37). The N271 residue is shown in pink; N272 in blue sphere and K273 in purple.

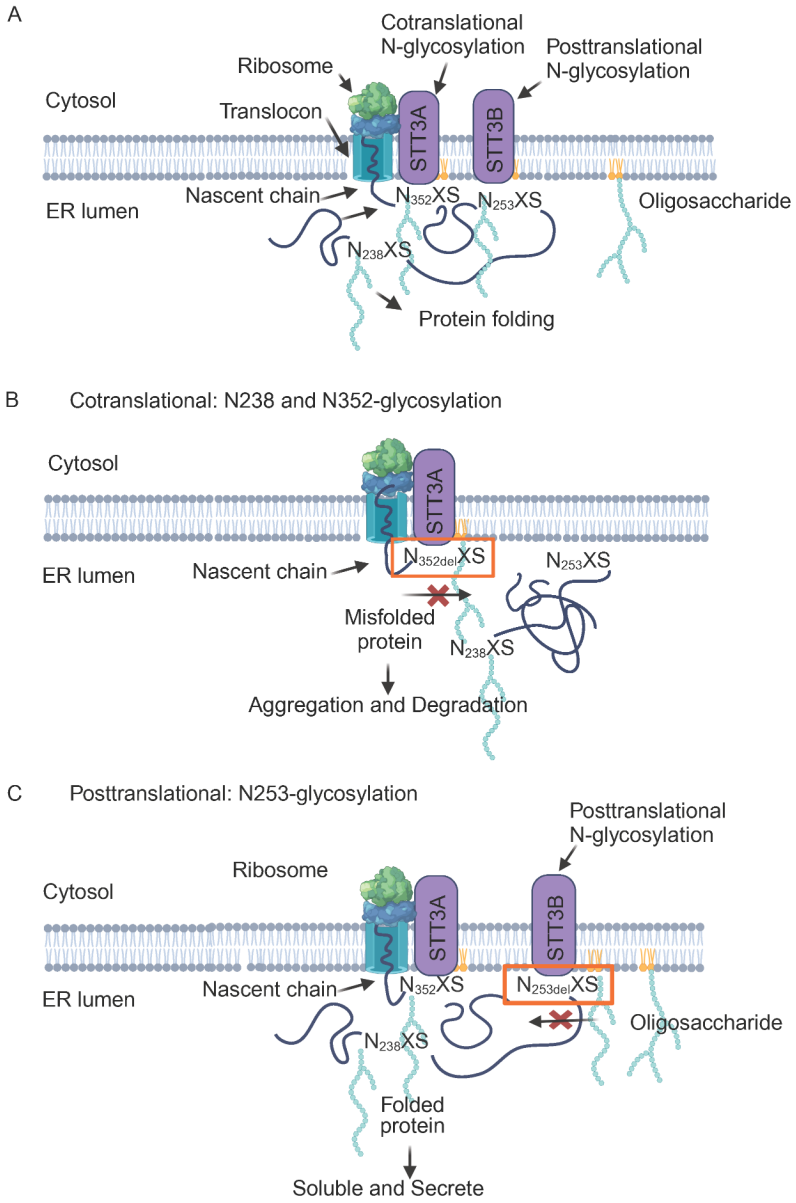


Figure 8. N-glycosylation at N238, N253 and N352 in the endoplasmic reticulum (ER). (**A**) STT3A transfers the oligosaccharide to N238 and N352 (NXS) in the nascent protein chain cotranslationally. STT3B transfers the oligosaccharide to the N253 sequon in a posttranslational manner. (**B**) Misfolded N352del triggers the cotranslational protein degradation (50). (**C**) With preserved protein structure, 253del, even without posttranslational modification, is able to pass the quality control system and be transported out of the ER.