

Molecular dynamics simulations elucidate the misfolding mechanisms of secretion-defective pancreatic lipase variants

Gyula Hoffka^{1,2}  and András Szabó¹ 

¹ Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Debrecen, Hungary

² Department of Chemistry, Lund University, Lund, Sweden

Keywords

chronic pancreatitis; computational model; enzyme mutation; protein folding; structural analysis

Correspondence

A. Szabó, Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Egyetem tér 1, Life Science Building 3.503, Debrecen H-4032, Hungary

Tel: +36-52-416-432

E-mail: andrasszabo@med.unideb.hu

and

G. Hoffka, Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Egyetem tér 1, Life Science Building 3.006, Debrecen H-4032, Hungary

Tel: +36-52-416-432

E-mail: hoffka.gyula@med.unideb.hu

(Received 10 April 2026, revised 15 May 2026, accepted 9 June 2026)

doi:10.1111/febs.70627

Protein misfolding of pancreatic secretory enzymes is increasingly recognized as a contributor to chronic pancreatitis. Recently, missense pancreatic lipase (PNLIP) mutations that impair lipase secretion and cause intracellular aggregation and endoplasmic reticulum (ER) stress have been linked to the disease. Among these, A174P, G233E, and C254R are located in the N-terminal domain, whereas V454F is situated in the C-terminal domain of PNLIP. To elucidate how these mutations disrupt PNLIP structure, we performed detailed molecular dynamics simulations and structural modeling of human PNLIP variants. Evolutionary conservation analyses suggested that the mutation sites contribute to PNLIP structural stability. All mutations, regardless of location, altered the residue interaction network within the N-terminal PNLIP domain, consistently affecting Glu270. Detailed structural analysis revealed that the A174P substitution disrupted several helix-stabilizing hydrogen bonds, whereas the G233E mutation introduced multiple local hydrogen bonds and caused structural distortions. The C254R mutation abolished the native Cys254-Cys278 disulfide bond and instead formed an alternative salt bridge destabilizing the lid region. All three mutations altered stabilizing interactions mainly within the N-terminal domain of PNLIP. Interestingly, the V454F substitution exerted long-range allosteric effects, modulating key interactions also within the N-terminal domain. Further structural analysis showed that Phe454 in the variant became solvent-exposed, potentially leading to aggregate formation through the mutation site. These observations provide a molecular basis for PNLIP misfolding and its downstream proteotoxic effects, supporting the development of targeted therapies for chronic pancreatitis.

Introduction

Pancreatic lipase (PNLIP; EC3.1.1.3) consists of an N-terminal α/β -hydrolase domain that carries the catalytic site and a flexible surface loop, the lid that regulates substrate access to the enzyme [1,2]. It also possesses a C-terminal β -sandwich domain responsible

for binding the cofactor colipase [3]. Although bile acids in the intestine can inhibit PNLIP, colipase binding protects the enzyme from inactivation [4]. Additionally, the human pancreas expresses and secretes two pancreatic lipase-related proteins (PLRP1 and

Abbreviations

BiP, Binding Immunoglobulin Protein; ER, endoplasmic reticulum; KIF, Key Interaction Finder; PLRP1 and PLRP2, pancreatic lipase-related proteins 1 and 2; PNLIP, pancreatic lipase; RMSF, root-mean-square-fluctuation; SASA, solvent accessible surface area.

PLRP2), which are highly homologous with PNLIP [5–8].

Protein misfolding is characterized by loss of protein function, intracellular retention and aggregation. Secretory proteins carrying misfolding mutations are typically recognized and retained in the endoplasmic reticulum (ER) by a quality control system [9]. ER retention of misfolded proteins can activate a signaling cascade known as unfolded protein response (UPR), which can lead to ER stress and cell death [10,11]. Misfolded secretory proteins are involved in a variety of diseases, such as diabetes, alpha1-antitrypsin deficiency, hypothyroidism, and chronic pancreatitis [12–15].

Chronic pancreatitis is a progressive inflammatory disorder of the pancreas that can be triggered by missense mutations. In pancreatic proteases, such alterations can promote pancreatitis progression by inducing protein misfolding, intracellular retention, and often ER stress [16–21]. Recently, an international collaborative study identified multiple heterozygous PNLIP missense variants in patients with chronic pancreatitis and control individuals [22]. Functional characterization showed that the A174P, G233E, C254R, and V454F variants displayed protein misfolding, aggregation, and induction of ER stress, suggesting that PNLIP misfolding may contribute to disease pathogenesis [23]. Clarifying the molecular mechanisms underlying PNLIP misfolding and ER retention is essential for developing targeted therapies. In this study, we performed molecular dynamics simulations and protein modeling to predict, analyze, and evaluate the structural and functional consequences of disease-associated PNLIP mutations and better understand their roles in pancreatic pathophysiology.

Results

Pancreatic lipase-like enzymes are expressed in many vertebrate species. Conserved amino acid residues contribute to the structural stability of enzymes. To evaluate the structural role of the investigated mutation sites in PNLIP, we performed a complete evolutionary amino acid conservation analysis (Fig. 1). The analysis revealed that PNLIP residues Gly233, Cys254, and Val454 are highly conserved (positional frequencies 99%, 100%, and 80% respectively), whereas residue Ala174 exhibits intermediate conservation (positional frequency 22%). At position 174, the nonpolar aliphatic residues valine, isoleucine, and leucine were also observed with frequencies of 24%, 24%, and 10%, respectively. In human PLRP1 and PLRP2, the corresponding amino acid residues were identical to those in PNLIP, with the exception of Ala174, which was

replaced by Val in both enzymes. Prediction of the pathogenicity of PNLIP mutations A174P, G233E, C254R, and V454F identified in chronic pancreatitis cohorts suggested that A174P, G233E, and C254R are likely pathogenic, whereas V454F showed an ambiguous prediction (Table S1). The findings suggest that these amino acid residues contribute to structural stability, and their mutation consequently causes PNLIP misfolding.

First, we analyzed the PNLIP structure at the full-molecule level to assess variant-induced alterations. We performed radius of gyration analyses and calculated the average and maximum radius of gyration for each frame of the trajectories (Fig. 2). Variants A174P, C254R, and V454F displayed radius of gyration values comparable to the wild-type, while variant G233E showed a somewhat increased radius, suggesting an overall loosening or expansion of the protein structure. To gain information on the general mobility of individual residues, we calculated the root-mean-square-fluctuation (RMSF) values for C α with complete protein alignment of the wild-type PNLIP and G233E variant (Fig. 3A). We also performed domain-specific trajectory alignments for both the N- and C-terminal domains before RMSF calculations (Fig. 3B). Interestingly, alignment of the whole PNLIP revealed markedly increased RMSF values for the C-terminal domain of the G233E variant compared to the wild-type (Fig. 3A), whereas only subtle differences were observed following independent domain alignments (Fig. 3B). Taken together, these results suggest a relative motion between the N- and C-terminal domains of the G233E PNLIP variant. The other variants exhibited RMSF values consistent with the wild-type (Figs S1 and S2).

To examine whether the variants associated with PNLIP misfolding display increased exposure of apolar side chains on their surface, we compared the solvent accessible surface area (SASA) of wild-type PNLIP with that of its variants (Table S2). The calculations were performed for the complete residues of the entire protein, considering strictly apolar residues (Ala, Phe, Ile, Leu and Val) and a broader set that also included Met, Trp, and Tyr. The calculated SASA values revealed that the variants exposed a similar amount of apolar residues on their surface as the wild-type PNLIP. We also assessed whether the investigated mutations altered the alpha helix and beta strand content of PNLIP at 37 °C and 277 °C (Figs S3 and S4). The high temperature reduced the overall secondary structure content but did not result in substantial differences between wild-type PNLIP and the variants.

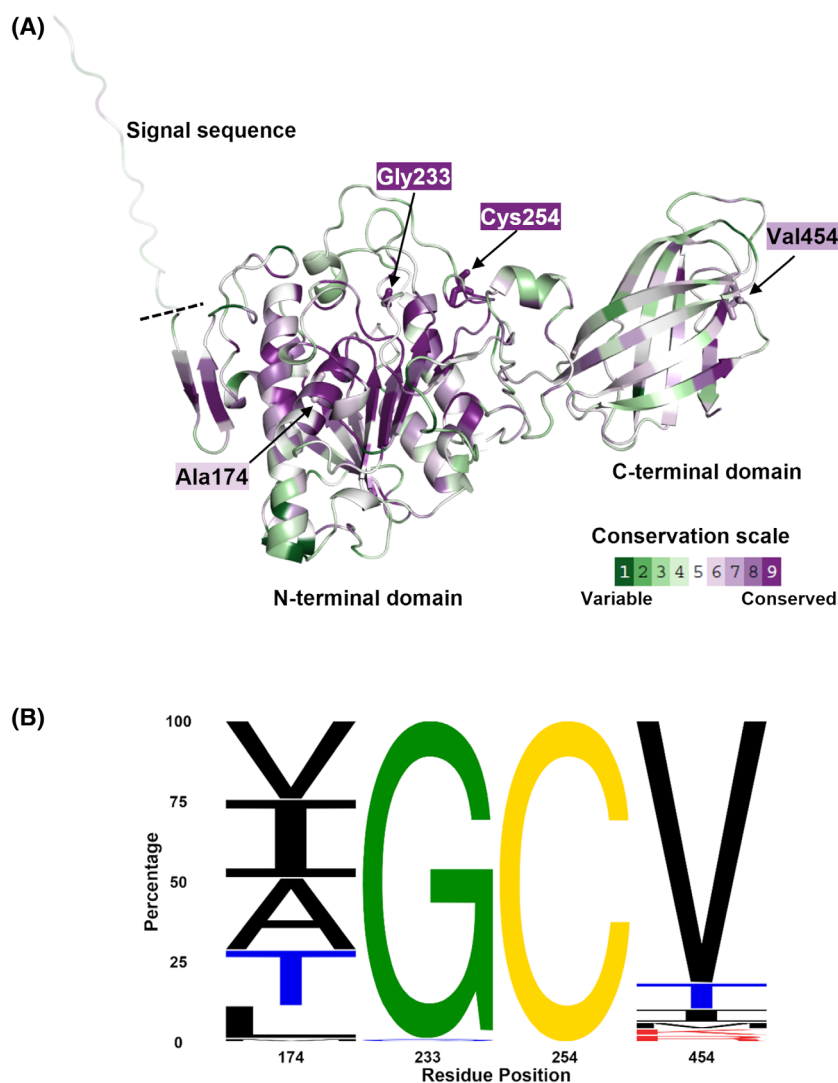


Fig. 1. Evolutionary conservation of the pancreatitis-associated mutation sites in the human PNLIP protein. (A) Conservation analysis was performed with ConSurf [78], using multiple sequence alignment of PNLIP with 149 homologous sequences from invertebrates to vertebrates and an associated phylogenetic tree. Highly similar (> 95%) and highly divergent (< 35%) sequences were excluded from the analysis. The conservation level of each residue was mapped onto the PNLIP structure using a 1–9 scale, where 9 indicates highly conserved and 1 highly variable residues. The mutation sites were colored according to their conservation grade. The structure was visualized using PyMOL. (B) Sequence logo illustrating amino acid conservation at the mutation sites. Data were obtained from the ConSurf analysis, and the logo was generated using RSTUDIO. The height of each symbol within a stack indicates the relative frequency of the corresponding amino acid, while the colors represent the chemical properties of the amino acid side chains.

To investigate the alterations in residue–residue interaction patterns between wild-type PNLIP and its variants, we evaluated and compared the interactions of wild-type PNLIP and its variants using the Key Interaction Finder (KIF) program. Residues with the highest KIF scores (> 0.7), indicating substantial differences between the wild-type and variant interaction networks, are shown in Fig. 4. Relative to wild-type PNLIP, interaction network alterations in the A174P,

G233E, and C254R variants were observed mainly in the N-terminal domain of the enzyme, in the vicinity of the mutation sites. These network changes particularly affected the active site of PNLIP. Notably, one of these residues, His280, is part of the catalytic triad of the lipase enzyme. Although the V454F mutation is located in the C-terminal domain, substantial residue–residue interaction changes were detected only in the N-terminal domain, mostly in the same region

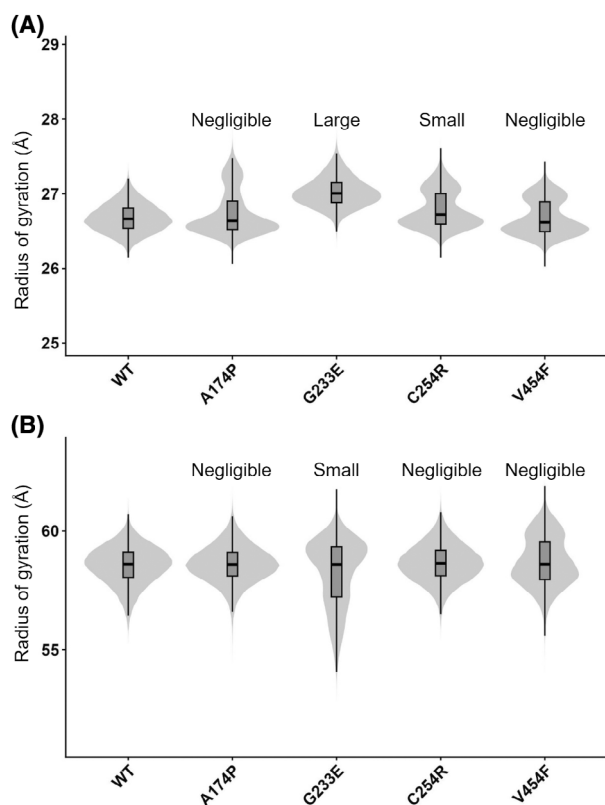


Fig. 2. Radius of gyration of wild-type PNLIP and its variants. Average (A) and maximum (B) radius of gyration values, calculated from individual frames over the complete trajectories ($n = 3$), are shown. Labels negligible (< 0.2), small ($0.2–0.5$), medium ($0.5–0.8$), and large (> 0.8) indicate the magnitude of difference according to Cohen's d , as described previously [74].

described above. This suggests that the V454F mutation may influence distant regions via allosteric interactions. Interestingly, the interaction network of Glu270 was altered in all four variants compared with the wild-type. Overall, the average per-residue KIF score was highest for the A174P and V454F variants. To identify residues with the most perturbed interaction networks, we determined the minimum, maximum, and average values of the per-residue KIF scores among the four variants for each residue (Fig. S5). The analysis revealed that residues 260–300 are highly affected by the mutations.

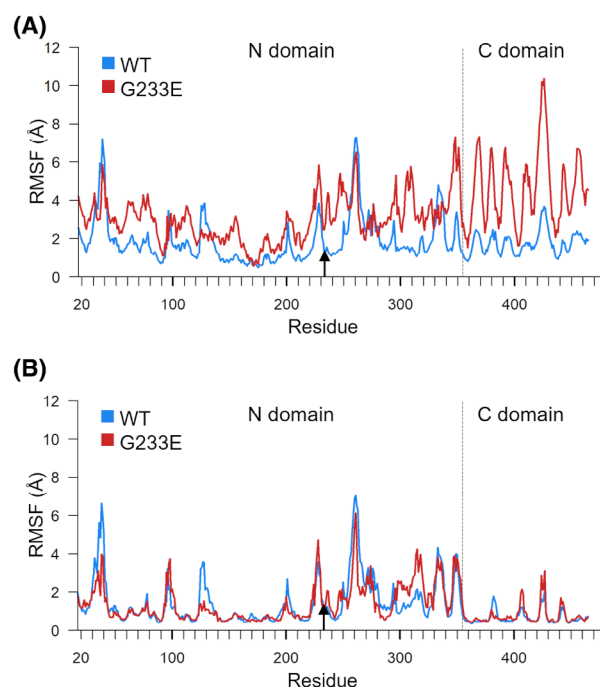
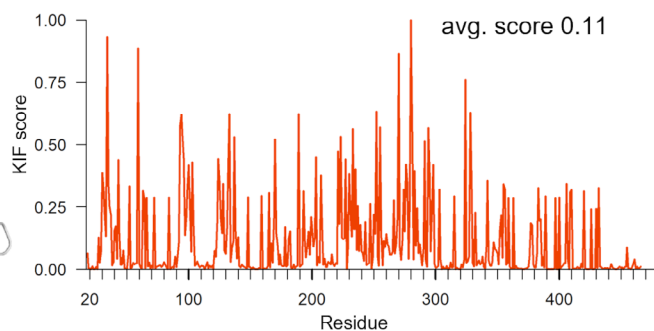
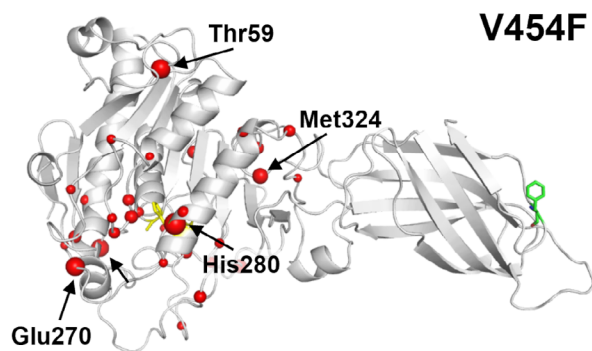
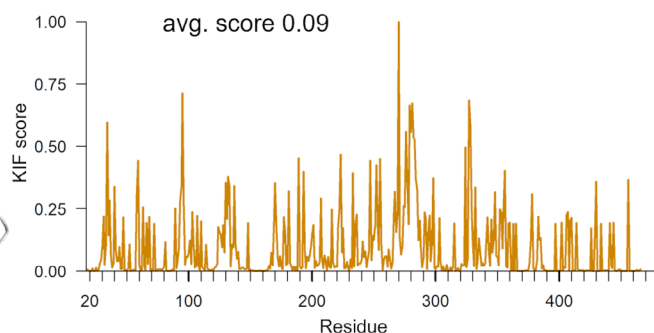
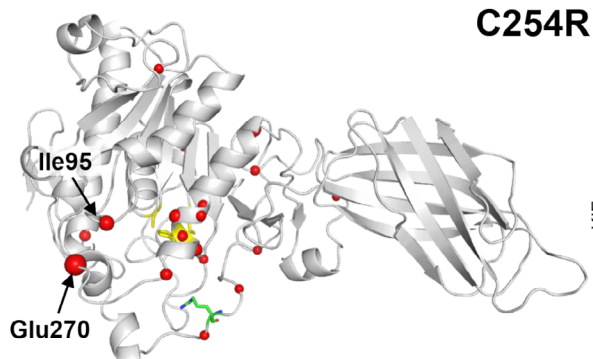
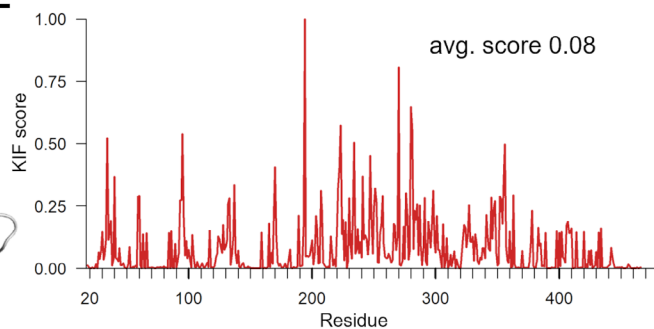
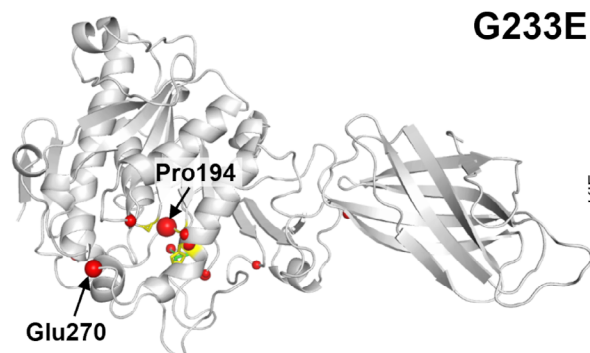
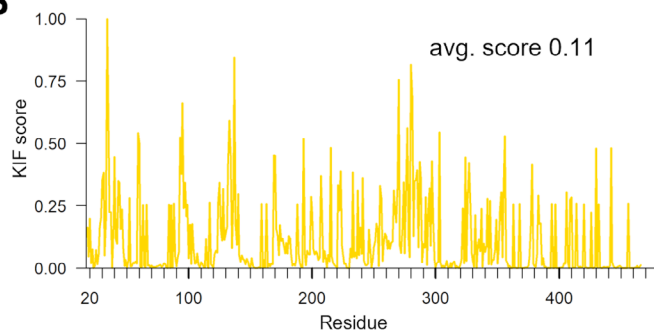
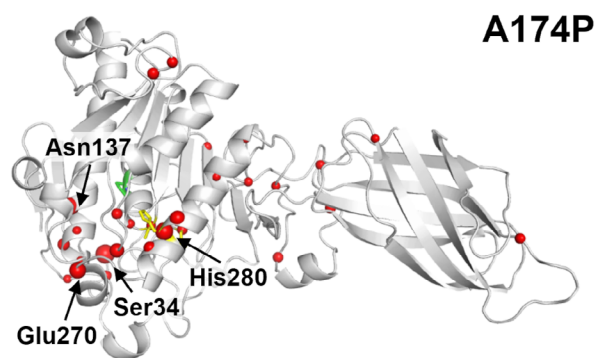


Fig. 3. Root-Mean-Square-Fluctuation (RMSF) values calculated for the $C\alpha$ atoms of the wild-type PNLIP and G233E variant. RMSF values were calculated from the combined trajectories of $n = 3$ independent replicate simulations after fitting to all $C\alpha$ atoms (A) and $C\alpha$ atoms of respective domains (B). PNLIP N-terminal and C-terminal domains span residues 17–352 and 355–465, respectively. The values for residues 353–354 were derived following C-terminal domain fitting. The arrows indicate the mutation site.

To gain more insight into the effects of individual mutations, we performed an in-depth structural and interaction analysis of the PNLIP variants in comparison with the wild-type enzyme. The A174P mutation is located in one of the alpha helices near the active site of the N-terminal PNLIP domain (Fig. 5A). Our analysis of hydrogen bond prevalence during the simulations revealed important differences. Ala174 formed multiple hydrogen bonds with nearby residues of the same helix in wild-type PNLIP (Fig. 5B). The backbone carbonyl of Ala174 formed a hydrogen bond with Ala178, which was present in 43% of the trajectory. This interaction was also observed in the A174P

Fig. 4. Comparison of interaction networks between wild-type PNLIP and its variants. Changes in residue interactions were analyzed using the Key Interactions Finder (KIF) software [75] from the combined trajectories of $n = 3$ independent replicate simulations. Red spheres represent amino acid $C\alpha$ atoms, with sphere size corresponding to the computed KIF score. Residues with KIF scores higher than 0.7 are labeled. Mutation sites are highlighted in green and the catalytic triad in yellow. The individual per-residue KIF scores are depicted as a function of residue number on the right side for each mutant. The average KIF score is also indicated. The structures were visualized using PyMOL.



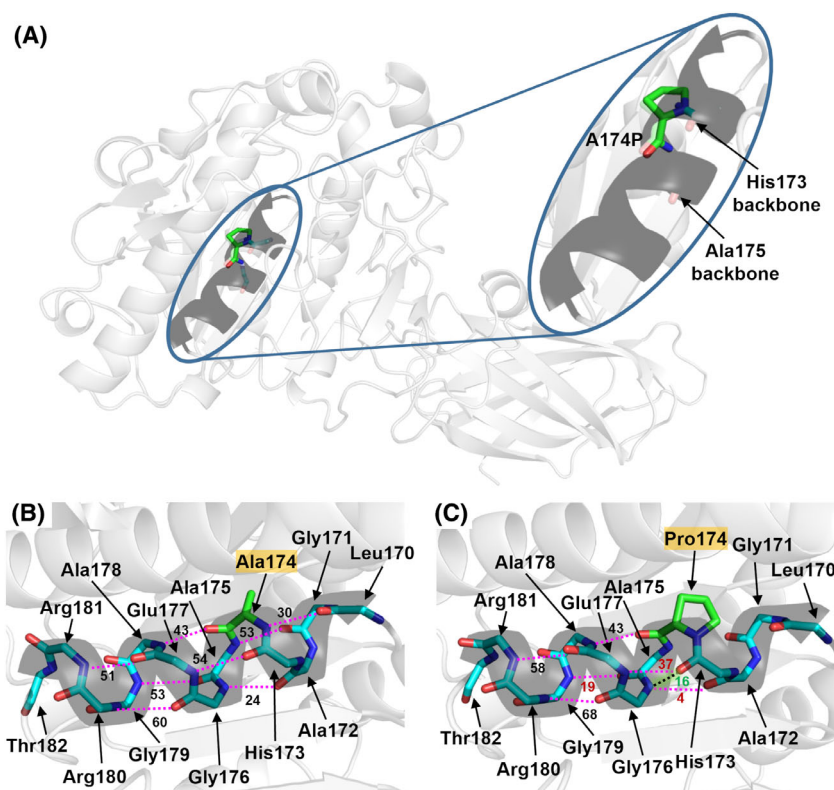


Fig. 5. Structural effect of the A174P mutation within an alpha helix of the PNLIP N-terminal domain. Representative PNLIP structures were obtained from the combined trajectories of $n = 3$ independent replicate simulations using cpptraj [68], as described in the Methods. (A) The A174P mutation site is shown in green, and the alpha helix is colored dark gray. Representative hydrogen bonds around the Ala174/Pro174 residue are shown with dashed lines in the wild-type PNLIP (B) and A174P variant (C). Hydrogen bonds are represented by dashed lines. Black, red, and green numbers represent percentages of unchanged, decreased, and increased hydrogen bond prevalence, respectively. The structures were visualized using PyMOL.

variant with the same frequency (Fig. 5C). The backbone amine of Ala174 formed a hydrogen bond with the backbone carbonyl of Leu170 in 30% of the trajectory (Fig. 5B). However, this interaction could not be observed in the A174P variant, as the proline backbone lacks a donatable hydrogen (Fig. 5C). Notably, in the A174P variant, the prevalence of several helix-stabilizing hydrogen bonds was markedly decreased in the trajectory, except for a newly formed hydrogen bond (16% of the trajectory) between His173 and Gly176, which was not present in the wild-type (Fig. 5B,C).

We also compared the differences in the hydrogen bonds formed by Gly233 or Glu233 in wild-type PNLIP and the G233E variant, respectively (Fig. 6A, B). In wild-type PNLIP, the backbone amide of Gly233 formed a single hydrogen bond with the backbone carbonyl of Asp222, which was present in 23% of the trajectory (Fig. 6A). In contrast, the charged side chain of Glu233 of the variant formed multiple

hydrogen bonds with several surrounding residues (Fig. 6B). Some of these interactions were present nearly in the entire trajectory, indicating a strong local effect of the mutation.

Comparison of wild-type PNLIP with the C254R variant suggested prominent structural alterations (Fig. 6C,D). In the wild-type enzyme, Cys254 formed a disulfide bond with Cys278, stabilizing the lid region within PNLIP N-terminal domain (Fig. 6C). In the C254R PNLIP variant, on the other hand, the Arg254 side chain contributed to a salt bridge with Glu250 instead of a disulfide bond, as demonstrated in a representative structure shown in Fig. 6D. This alternative interaction likely contributed substantially to the destabilization of the original conformer. Interestingly, distance measurements between the Arg254 side chain and Glu250 (Fig. 6E) and between Arg254 and Cys278 (Fig. 6F) indicated that the salt bridge was present in one trajectory for most of the simulation time, while Arg254 consistently remained farther from Cys278.

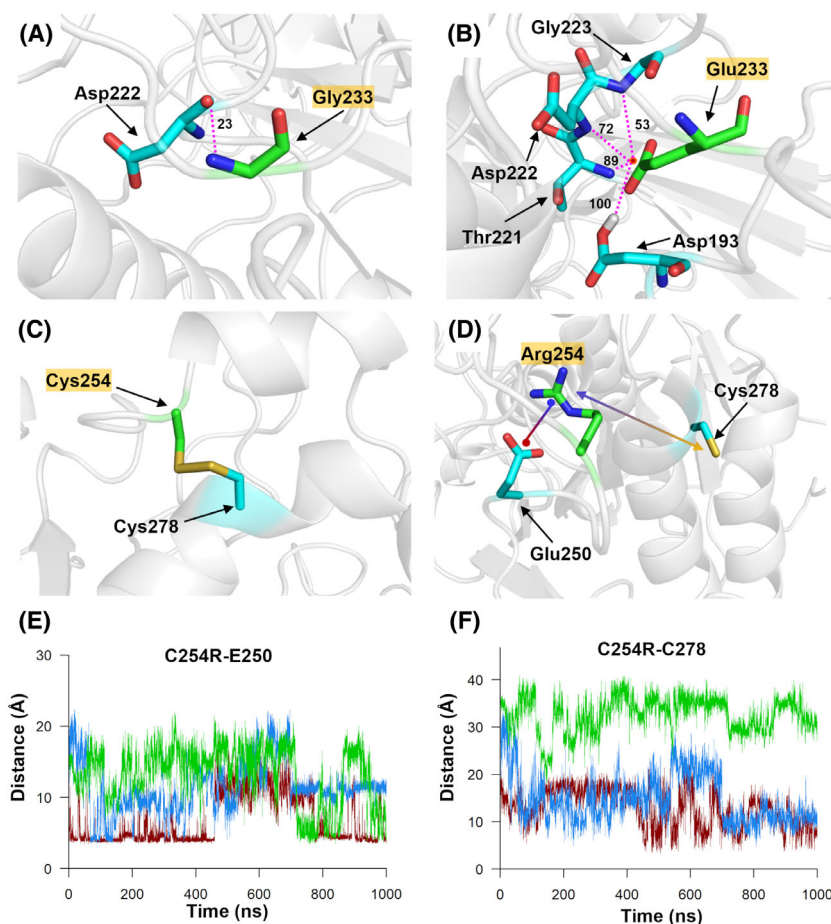


Fig. 6. Effect of the G233E and C254R mutations on the interaction network of the PNLIP N-terminal domain. Representative PNLIP structures were obtained from the combined trajectories of $n = 3$ independent replicate simulations using cpptraj [68], as described in the Methods. Interaction network around the Gly233/Glu233 residue is shown for the wild-type PNLIP (A) and G233E variant (B). Dashed lines label representative hydrogen bonds, and the accompanying numbers denote their prevalence (percentage), as obtained with cpptraj. Composite hydrogen bonds to the two terminal oxygen atoms of Glu233 are highlighted. Interaction network around the Cys254/Arg254 residue is illustrated for the wild-type PNLIP (C) and C254R mutant (D). The colored line marks a salt bridge between Arg254 and Glu250, the double arrow denotes the Arg254-Cys278 distance. Distance of the terminal carbons of Glu250 and Arg254 sidechains over the trajectories (E), and distance of the terminal carbon of Arg254 and the sulfur atom of Cys278 over the trajectories (F) are indicated. Distance measurements were carried out as described in Methods separately for the $n = 3$ independent replicate simulations. Different colors denote the individual trajectories. The structures were visualized using PyMOL.

Uniquely, the Val454 mutation site is located in the C-terminal domain of PNLIP and is positioned between two beta strands in the hydrophobic core (Fig. 7A). Distance measurements between the C α atoms of adjacent beta strands in the wild-type PNLIP and the V454F variant showed negligible differences (Fig. 7B,C). SASA values calculated for the side chain atoms of Val454 residue and all other valine residues in the wild-type PNLIP were 13.9 Å² and 8.1 Å², respectively. However, the SASA values calculated for Phe454 in the variant and all other phenylalanine residues in the wild-type were 71.6 Å² and 13.8 Å², respectively. These results suggested that

the Phe454 residue is exposed to the solvent, rather than participating in the hydrophobic core (Fig. 7D). The exposed Phe residue in the variant may promote PNLIP aggregation with π -stacking; therefore, we performed AlphaFold3 predictions of six-member multimers. The V454F PNLIP variant showed a pronounced complex formation promoted by Phe-Phe interactions introduced by the mutation (Fig. 8A,B). In contrast, the Val454 residue in the wild-type PNLIP did not contribute to oligomerization (Fig. 8C,D). This observation suggests a more direct role for the V454F mutation in promoting PNLIP aggregation.

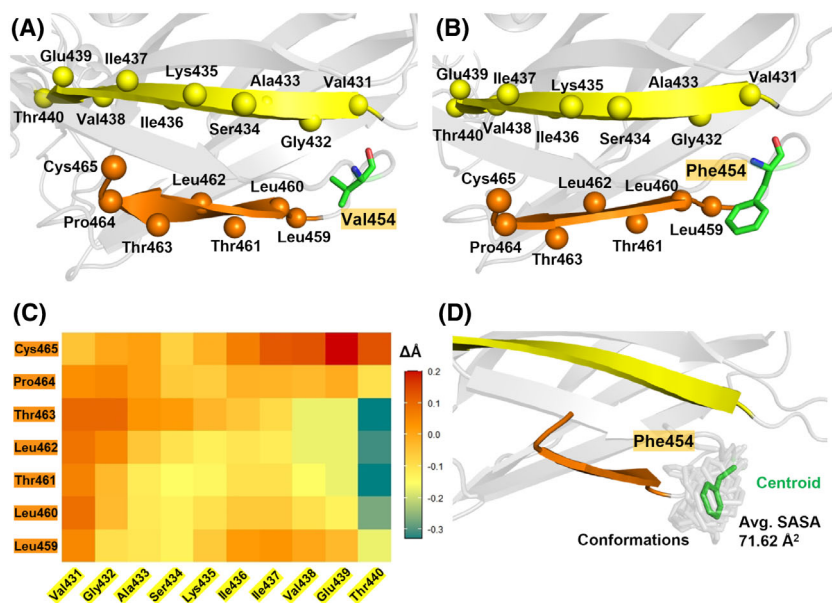


Fig. 7. Effect of the V454F mutation on the stability of two parallel beta strands in the PNLIP C-terminal domain. Representative PNLIP structures were obtained from the combined trajectories of $n = 3$ independent replicate simulations using cpptraj [68], as described in Methods. The spacing between beta strands near residue Val454/Phe454 is illustrated for wild-type PNLIP (A) and the V454F mutant (B). α atoms of the beta strand residues are represented as spheres. (C) Difference in residue-residue spacing across the two beta strands is shown. Plot colors indicate the differences in Å, and residue labels are colored to match the beta strand colors in panels A and B. (D) Conformations of Phe454. The main cluster centroid is represented with green sticks; simulation snapshots taken at each 25 ns of the simulation trajectories are represented with transparent gray sticks. The average sidechain SASA value is given in Å². Distance and SASA analyses were performed using the combined trajectories of $n = 3$ independent replicate simulations. The structures were visualized using PyMOL.

Discussion

The first reported missense mutation (T221M) in PNLIP was identified in two brothers who were homozygous for the variant and exhibited lipase deficiency [24]. Further investigation indicated that the mutation may impair PNLIP folding, resulting in the intracellular retention of insoluble protein aggregates and increased ER stress [25]. In addition, the T221M variant induced spontaneous chronic pancreatitis in a mouse model [26]. Recently, rare heterozygous pancreatic lipase (PNLIP) variants have been identified in European patients with chronic pancreatitis and healthy individuals [22]. Most of these variants were secreted at levels similar to wild-type PNLIP. However, some exhibited increased susceptibility to degradation by pancreatic proteases, a protease-sensitive phenotype that correlated with disease occurrence. The same publication reported additional rare variants, such as A174P, G233E, C254R, and V454F, of which G233E and C254R were identified only in chronic pancreatitis patients, whereas A174P and V454F were found only in controls [22]. Due to the low frequency of these mutations, genetic association with

pancreatitis could not be determined. However, *in vitro* experiments demonstrated that these PNLIP variants exhibit impaired secretion and accumulate in cells as insoluble aggregates, triggering ER stress and cell death [23]. The findings suggest incomplete penetrance of PNLIP variants consistent with similar observations for other pancreatic protein variants [27–30].

In this study, we explored the structural changes of the A174P, G233E, C254R, and V454F misfolded PNLIP variants at the atomistic level using molecular dynamics simulations. We used the inactive conformer of PNLIP (PDB ID: 1N8S) as the basis for our models because the triglyceride–water interfaces required for lipase activation are absent in pancreatic acinar cells [31]. A relevant consideration for this approach is that our molecular dynamics simulations relied on the structure of the fully folded enzyme, as structural data for misfolded or partially folded PNLIP variants are currently unavailable and challenging to study in general.

The A174P mutation is located within one of the major alpha helices of the PNLIP N-terminal domain, which is critical for maintaining the structural integrity of the enzyme. Proline is well known to have a very

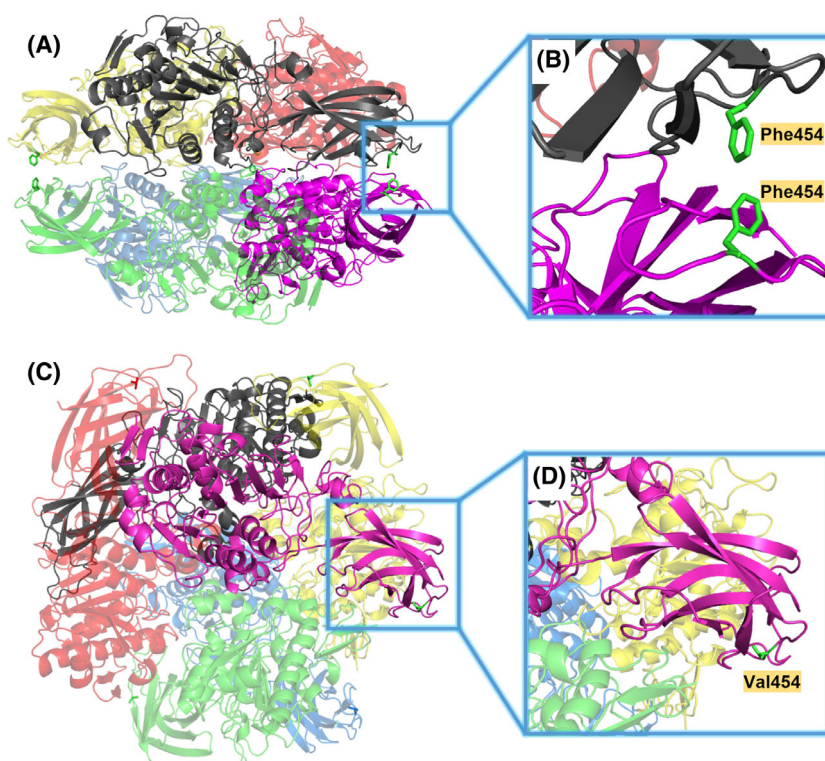


Fig. 8. Aggregated complex of the V454F PNLIP variant. (A) The hexameric aggregate of the variant is shown with individual monomers colored differently. (B) Close-up of the V454F mutant residue, highlighting interactions between mutated residues. (C) The hexameric aggregate of wild-type PNLIP is shown, with (D) a close-up of the Val454 residue in one of the monomers. The aggregate structure was predicted using AlphaFold3 software [76], as described in Methods. The highest-ranking model out of $n = 5$ prediction runs is displayed. The structures were visualized using PyMOL.

low helix-forming propensity in globular proteins [32]. In our simulations, this alpha helix was stabilized by several hydrogen bonds in wild-type PNLIP; however, the Ala174-to-Pro mutation disrupted this hydrogen bond network. This is because the rigid ring structure of proline restricts the backbone flexibility and, importantly, lacks an amide hydrogen required for helix stabilization. In addition, the mutation altered residue–residue interaction patterns throughout the N-terminal domain.

The G233E mutation site is located near the active site of PNLIP N-terminal domain. In the wild-type enzyme, we revealed that the backbone amide group of Gly233 contributed to a weak hydrogen bond with nearby Asp222. In contrast, the mutant Glu233 residue formed multiple, stable hydrogen bonds with the surrounding amino acid residues. These local changes in the interaction network within the N-terminal domain altered the relative positioning of the two domains, as indicated by the changes in the radius of gyration and RMSF values calculated after fitting to all C α atoms or to the C α atoms of the individual domains.

The C254R mutation prevents the formation of the Cys254–Cys278 disulfide bond, which stabilizes the lid [33]. The lid is a mobile surface loop that regulates substrate access to the active site of PNLIP [4]. It is generally accepted that the enzyme alternates between active and inactive conformational states, in which structural changes are observed mainly in the lid region. Our results indicated that, instead of a disulfide bond, a salt bridge was established in the C254R variant between Arg254 and Glu250 residues. The new salt bridge could alter the structure of the lid and the active site region within the N-terminal domain of PNLIP. Notably, in our simulations, the mutations modified the interaction pattern of Glu270 and His280 within the N-terminal domain, highlighting their potential role in stabilizing the PNLIP structure. Previous studies have indicated that missense mutations affecting disulfide-bond-forming cysteines in secretory proteins are relatively rare. In PRSS1, CPA1, and CTSC, cysteine substitutions disrupt proper protein folding and reduce secretion, increasing susceptibility to pancreatitis [16,17,34]. Similar cysteine mutations in

proinsulin are also strongly associated with type 1 diabetes [35,36].

The Val454 mutation site is located between two parallel beta strands within the C-terminal domain of PNLIP. However, our analysis demonstrated that the V454F mutation affected the residue interaction network in the N-terminal domain, likely through an allosteric mechanism. Similar long-range effects have been reported for certain SARS-CoV-2 spike protein variants and for the sickle-cell hemoglobin, where single amino acid substitutions induce structural changes at substantial distances across the protein [37–42]. We also observed that while the Val454 residue participates in stabilizing the hydrophobic core of the C-terminal domain, the Phe454 residue in the variant becomes exposed to the solvent. This feature of the V454F variant may promote the formation of intracellular protein aggregates through the mutation sites. Consistent with this notion, *in vitro* cell culture experiments have shown that misfolded PNLIP variants are prone to forming insoluble intracellular aggregates and increase the expression of BiP, an ER resident chaperone [23,25]. This chaperone binds to certain protein aggregates with high affinity, contributing to the activation of ER stress signaling [43]. Therefore, the formation of intracellular protein aggregates may play a significant role in the progression of chronic pancreatitis.

Molecular dynamics simulations and other computational approaches have previously been used to facilitate the development of pharmacological chaperones targeting disease-causing misfolded proteins. During molecular dynamics simulations, transient pockets may appear on the protein surface that remain undetectable with conventional methods. Such transient pockets can be specifically targeted by hotspot-based pharmacological prediction tools [44]. Molecular dynamics-based transient pocket analysis was applied to identify a binding site on a p53 variant and facilitated the development of pharmacological chaperones [45,46]. Similarly, transient pockets have been described for HIV-1 protease and homogentisate 1,2-dioxygenase variants [47,48]. Molecular dynamics approaches can also facilitate drug design targeting previously identified binding sites. For instance, pharmacological chaperones have been shown to improve the catalytic activity of a Gaucher disease-associated beta-glucocerebrosidase variant [49]. Similarly, the identification of allosteric hotspots facilitated the development of pharmacological chaperones against a mutant alpha-galactosidase associated with Fabry disease [50]. Additionally, a study on prion protein misfolding revealed that docking and molecular dynamics simulations can assist the identification of pharmacological chaperones that

stabilize native conformations and suppress pathogenic aggregation [51].

In summary, we found that missense mutations may induce notable structural changes in PNLIP leading to misfolding. These results may facilitate the development of pharmacological chaperones that promote the correct folding and secretion of PNLIP variants, thereby reducing their proteotoxic effects and supporting personalized therapeutic approaches for chronic pancreatitis.

Methods

System preparation

The coordinate file of human PNLIP in complex with procolipase (PDB ID: 1N8S [2]) was downloaded from the Protein Data Bank [52]. Procolipase was removed from the structure, as previous experimental studies on PNLIP were carried out in its absence. PNLIP variants were generated with Chimera [53] using the Dynaomics [54] rotamer library. The Propka 3.1 [55,56] software was used to determine the pKa values of titratable amino acid sidechains, and the corresponding protonation states were set accordingly. The protonation state of neutral histidine residues was determined with Chimera [53].

PNLIP and its variants were prepared for simulations with tleap as implemented in Amber22 [57]. PNLIP was described using the Amber ff14SB [58] force field, and the missing hydrogen atoms were added automatically. Disulfide bonds between the correct cysteine residues were added into the tleap input file. The system was solvated in a truncated octahedral box of TIP3P [59] water with a minimum solute-box distance of at least 12.0 Å. Based on the number of added water molecules and the charge of the protein, the amount of sodium and chloride ions required to achieve a salt concentration of ~0.15 M and to neutralize the system charge was determined with the Split Charge protocol [60]. The original systems were solvated again with the same setup, by replacing added water molecules with the calculated amount of ions.

Molecular dynamics simulation protocol

Molecular dynamics simulations were carried out using the Amber22 [57] software package. The GPU accelerated version [61–63] of the pmemd algorithm of Amber22 was applied with a Single-Precision Floating-Point (pmedd.cuda_SPFP [63]). The system was minimized with the steepest descent minimization algorithm for 2500 steps, followed by 7500 steps of conjugate gradient minimization, using a 10 kcal mol⁻¹ Å⁻² positional restraint on nonhydrogen atoms of the protein. After that, the system was heated to 100 K (−173 °C) for 250 ps of simulation time in an NVT

ensemble, and then to 310.15 K (37 °C) for 250 ps of simulation time in an NPT ensemble, keeping the same positional restraint. Subsequently, the same minimization, NVT and NPT heating steps were repeated, with a 5 kcal mol⁻¹ Å⁻² positional restraint on the C α atoms of the protein. This was followed by six equilibrations of 100 ps, in an NPT ensemble, each with increasingly smaller positional restraints on C α atoms, starting with 5 and ending with 0 kcal mol⁻¹ Å⁻² value. Subsequently, six simulated annealing cycles were applied. First, the systems were heated to 400 K (127 °C) over a 2 ns simulation and equilibrated at that temperature for 50 ns, followed by cooling to 273.15 K (0 °C) in a 15 ns simulation. In the consequent five cycles, the systems were reheated to 400 K in a 10 ns simulation, equilibrated at 400 K for 50 ns and cooled to 273.15 K in a 15 ns simulation. Finally, the systems were heated to 310.15 K (37 °C) in a 5 ns simulation. All molecular dynamics simulations up to this point were performed using a 1 fs time step, and long-range electrostatics were described using the Particle Mesh Ewald (PME) [64] approach, with a 10 Å cut-off for nonbonded interactions. During heating and equilibration, the temperature was controlled using Langevin dynamics [65], and all bonds involving a hydrogen atom were constrained using the SHAKE algorithm [66]. Finally, a production run was carried out with a 2 fs time step for a duration of 1 μ s. Trajectory sampling was executed every 500 ps. For each variant, three independent simulations were performed using the same protocol.

A further set of simulations was also executed, in which the final NPT heating step (before the six equilibrations) was carried out for a duration of 1 ns, and the final temperature was set to 550 K (277 °C). The production run in this case was carried out for a duration of 500 ns. All other settings were the same as described earlier.

Molecular dynamics simulation analysis

The analysis steps described below refer to simulations performed at 310.15 K (37 °C), unless stated otherwise. The cpptraj tool of AmberTools22 [67] was used to compute the root-mean-square-fluctuation (RMSF) values for C α atoms and complete residues (excluding hydrogen atoms) over the trajectory for each of the three replicas. Analyses of secondary structure elements during the simulations were executed with the secstruct command of cpptraj [68]. The same settings were used for simulations at 550 K (277 °C). Radius of gyration values were also calculated with cpptraj [68], considering both the average and maximum values for each frame of the trajectory. Statistical analysis of the radius of gyration calculations was carried out in R [69], using RSTUDIO [70], and the tidyverse [71] and effsize [72] packages. Cohen's d effect size test [73] using the standard algorithm was performed to compare each variant with the wild-type PNLIP. The magnitudes (negligible, small,

medium, or large) were assessed using the thresholds published previously [74]. Solvent accessible surface area (SASA) values were also calculated with cpptraj.

The effect of mutations on the protein interaction network was analyzed with the Key Interactions Finder (KIF) [75] software. Following the standard KIF protocol, all interaction types were considered, and interaction strengths were calculated for each frame of the trajectories. The residue identifier at the mutation site of each PNLIP variant was renamed to match that of the wild-type for comparison. The per-residue KIF scores were calculated accordingly.

Hydrogen bond prevalence was calculated with cpptraj [68]. Each frame of the trajectories was taken into account. Prevalence is defined as the percentage of frames for which cpptraj detected a certain hydrogen bond.

Distance measurements were carried out with cpptraj [68]. In case of the C254R variant, the distance between the Arg254 and Glu250 was measured between C ζ of Arg254 and the C δ of Glu250. The distance between Arg254 and Cys278 was measured between C ζ of Arg254 and the sulfur atom of Cys278. For the V454F variant and the wild-type used for comparison, distances between residues Leu459-Cys465 and Val431-Thr440 were calculated between the C α atoms.

Clustering of the V454F variant was performed in cpptraj using the hierarchical agglomerative approach and the average linkage method, applying an epsilon value of 5.0 Å based on C α atoms.

Molecular structures were visualized with PyMOL and graphical plots were generated with RSTUDIO.

AlphaFold3 prediction

To model the aggregation of the V454F variant, we have applied the AlphaFold3 software [76]. The sequence of PNLIP was obtained from the UniProt database [77] (identifier: P16233). The signal sequence (residues 1 to 16) was removed, and the V454F mutation was introduced by renaming the mutated residue in the sequence. The final sequence was uploaded into the AlphaFold3 prediction webserver [76]. The prediction of the variant was modeled in a hexamer state. Five independent standard models were generated, and the final model was selected based on the highest Ranking Score.

Author contributions

AS conceived and directed the study. GH performed system preparation, molecular dynamics simulations, AlphaFold predictions, and data analysis. AS performed the evolutionary analysis. GH and AS interpreted the results, prepared the figures, and wrote the manuscript.

Acknowledgement

We acknowledge KIFÜ for awarding us access to resources based in Hungary. Gergely Nagy is gratefully acknowledged for critical reading of the manuscript and helpful suggestions. This work was supported by the University of Debrecen Program for Scientific Publication.

Conflict of interest

The authors declare no conflicts of interest.

Data availability statement

The initial structures, preparation input files and molecular dynamics protocols are openly available in a public repository that issues datasets with DOI (Zenodo database). The molecular dynamics simulation trajectories are available upon request from the authors. Molecular dynamics starting structures, preparation scripts and MD scripts are available at the following Zenodo link: <https://doi.org/10.5281/zenodo.18886432>. The AlphaFold3 models of the wild-type and V454F PNLIP hexamers are available in ModelArchive at <https://www.modelarchive.org/doi/10.5452/ma-7ycd9> link.

References

- Winkler FK, D'Arcy A & Hunziker W (1990) Structure of human pancreatic lipase. *Nature* **343**, 771–774.
- van Tilbeurgh H, Sarda L, Verger R & Cambillau C (1992) Structure of the pancreatic lipase-procolipase complex. *Nature* **359**, 159–162.
- Bourne Y, Martinez C, Kerfelec B, Lombardo D, Chapus C & Cambillau C (1994) Horse pancreatic lipase. The crystal structure refined at 2.3 Å resolution. *J Mol Biol* **238**, 709–732.
- Lowe ME (2002) The triglyceride lipases of the pancreas. *J Lipid Res* **43**, 2007–2016.
- Zhu G, Fang Q, Zhu F, Huang D & Yang C (2021) Structure and function of pancreatic lipase-related protein 2 and its relationship with pathological states. *Front Genet* **12**, 693538.
- Holmes RS & Cox LA (2012) Bioinformatics and evolution of vertebrate pancreatic lipase and related proteins and genes. *JDMGP* **3**, 1.
- Berton A, Sebban-Kreuzer C & Crenon I (2007) Role of the structural domains in the functional properties of pancreatic lipase-related protein 2. *FEBS J* **274**, 6011–6023.
- Xiao X & Lowe ME (2015) The β 5-loop and lid domain contribute to the substrate specificity of pancreatic lipase-related protein 2 (PNLIPRP2). *J Biol Chem* **290**, 28847–28856.
- Adams BM, Oster ME & Hebert DN (2019) Protein quality control in the endoplasmic reticulum. *Protein J* **38**, 317–329.
- Malhotra JD & Kaufman RJ (2007) The endoplasmic reticulum and the unfolded protein response. *Semin Cell Dev Biol* **18**, 716–731.
- Oakes SA & Papa FR (2015) The role of endoplasmic reticulum stress in human pathology. *Annu Rev Pathol* **10**, 173–194.
- Sun J, Cui J, He Q, Chen Z, Arvan P & Liu M (2015) Proinsulin misfolding and endoplasmic reticulum stress during the development and progression of diabetes. *Mol Asp Med* **42**, 105–118.
- Yamasaki M, Sendall TJ, Pearce MC, Whistock JC & Huntington JA (2011) Molecular basis of α 1-antitrypsin deficiency revealed by the structure of a domain-swapped trimer. *EMBO Rep* **12**, 1011–1017.
- Citterio CE, Kim K, Rajesh B, Pena K, Clarke OB & Arvan P (2023) Structural features of thyroglobulin linked to protein trafficking. *Protein Sci* **32**, e4784.
- Sahin-Tóth M (2017) Genetic risk in chronic pancreatitis: the misfolding-dependent pathway. *Curr Opin Gastroenterol* **33**, 390–395.
- Schnúr A, Beer S, Witt H, Hegyi P & Sahin-Tóth M (2014) Functional effects of 13 rare PRSS1 variants presumed to cause chronic pancreatitis. *Gut* **63**, 337–343.
- Witt H, Beer S, Rosendahl J, Chen J-M, Chandak GR, Masamune A, Bence M, Szmola R, Oracz G, Macek M *et al.* (2013) Variants in CPA1 are strongly associated with early onset chronic pancreatitis. *Nat Genet* **45**, 1216–1220.
- Cassidy BM, Zino S, Fjeld K, Molven A, Lowe ME & Xiao X (2020) Single nucleotide polymorphisms in CEL-HYB1 increase risk for chronic pancreatitis through proteotoxic misfolding. *Hum Mutat* **41**, 1967–1978.
- Tjora E, Gravdal A, Engjom T, Cnop M, Johansson BB, Dimcevski GG, Molven A & Fjeld K (2021) Protein misfolding in combination with other risk factors in CEL-HYB1-mediated chronic pancreatitis. *Eur J Gastroenterol Hepatol* **33**, 839–843.
- Thiel F, Reiser M & Weiss FU (2022) A rare PRSS1 p.S127C mutation is associated with chronic pancreatitis and causes misfolding-induced ER-stress. *Pancreatol* **22**, 1112–1119.
- Sándor M, Vitale DS, Nagy ZA, Ibrahim SY, Abu-El-Haija M, Lazou M, Vajda S & Sahin-Tóth M (2025) Misfolding PRSS1 variant p.Ala61Val in a case of suspected intrauterine pancreatitis. *Pancreatol* **25**, 70–81.
- Lasher D, Szabó A, Masamune A, Chen J-M, Xiao X, Whitcomb DC, Barmada MM, Ewers M, Ruffert C,

- Paliwal S *et al.* (2019) Protease-sensitive pancreatic lipase variants are associated with early onset chronic pancreatitis. *Am J Gastroenterol* **114**, 974–983.
- 23 Toldi V, Kassay N & Szabó A (2021) Missense PNLIP mutations impeding pancreatic lipase secretion cause protein misfolding and endoplasmic reticulum stress. *Pancreatol* **21**, 1317–1325.
 - 24 Behar DM, Basel-Vanagaite L, Glaser F, Kaplan M, Tzur S, Magal N, Eidlitz-Markus T, Haimi-Cohen Y, Sarig G, Bormans C *et al.* (2014) Identification of a novel mutation in the PNLIP gene in two brothers with congenital pancreatic lipase deficiency. *J Lipid Res* **55**, 307–312.
 - 25 Szabó A, Xiao X, Haughney M, Spector A, Sahin-Tóth M & Lowe ME (2015) A novel mutation in PNLIP causes pancreatic triglyceride lipase deficiency through protein misfolding. *Biochim Biophys Acta* **1852**, 1372–1379.
 - 26 Zhu G, Wilhelm SJ, George LG, Cassidy BM, Zino S, Luke CJ, Hanna M, Stone S, Phan N, Matiwala N *et al.* (2023) Preclinical mouse model of a misfolded PNLIP variant develops chronic pancreatitis. *Gut* **72**, 1340–1354.
 - 27 Németh BC & Sahin-Tóth M (2014) Human cationic trypsinogen (PRSS1) variants and chronic pancreatitis. *Am J Physiol Gastrointest Liver Physiol* **306**, G466–G473.
 - 28 Solomon S & Whitcomb DC (2012) Genetics of pancreatitis: an update for clinicians and genetic counselors. *Curr Gastroenterol Rep* **14**, 112–117.
 - 29 Strum WB & Boland CR (2023) Advances in acute and chronic pancreatitis. *World J Gastroenterol* **29**, 1194–1201.
 - 30 Girodon E, Rebours V, Chen JM, Pagin A, Levy P, Ferec C & Bienvenu T (2020) Clinical interpretation of SPINK1 and CTSC variants in pancreatitis. *Pancreatol* **20**, 1354–1367.
 - 31 Belle V, Fournel A, Woudstra M, Ranaldi S, Prieri F, Thomé V, Currault J, Verger R, Guigliarelli B & Carrière F (2007) Probing the opening of the pancreatic lipase lid using site-directed spin labeling and EPR spectroscopy. *Biochemistry* **46**, 2205–2214.
 - 32 Pace CN & Scholtz JM (1998) A helix propensity scale based on experimental studies of peptides and proteins. *Biophys J* **75**, 422–427.
 - 33 Ranaldi S, Belle V, Woudstra M, Rodriguez J, Guigliarelli B, Sturgis J, Carrière F & Fournel A (2009) Lid opening and unfolding in human pancreatic lipase at low pH revealed by site-directed spin labeling EPR and FTIR spectroscopy. *Biochemistry* **48**, 630–638.
 - 34 Beer S, Zhou J, Szabó A, Keiles S, Chandak GR, Witt H & Sahin-Tóth M (2013) Comprehensive functional analysis of chymotrypsin C (CTSC) variants reveals distinct loss-of-function mechanisms associated with pancreatitis risk. *Gut* **62**, 1616–1624.
 - 35 Colombo C, Porzio O, Liu M, Massa O, Vasta M, Salardi S, Beccaria L, Monciotti C, Toni S, Pedersen O, Hansen T, Federici L, Pesavento R, Cadario F, Federici G, Ghirri P, Arvan P, Iafusco D, Barbetti F, & Early Onset Diabetes Study Group of the Italian Society of Pediatric Endocrinology and Diabetes (SIEDP) (2008) Seven mutations in the human insulin gene linked to permanent neonatal/infancy-onset diabetes mellitus. *J Clin Invest* **118**, 2148–2156.
 - 36 Liu M, Sun J, Cui J, Chen W, Guo H, Barbetti F & Arvan P (2015) INS-gene mutations: from genetics and beta cell biology to clinical disease. *Mol Asp Med* **42**, 3–18.
 - 37 Das JK, Thakuri B, MohanKumar K, Roy S, Sljoka A, Sun G-Q & Chakraborty A (2021) Mutation-induced long-range allosteric interactions in the spike protein determine the infectivity of SARS-CoV-2 emerging variants. *ACS Omega* **6**, 31312–31327.
 - 38 Ray D, Quijano RN & Andricioaei I (2022) Point mutations in SARS-CoV-2 variants induce long-range dynamical perturbations in neutralizing antibodies. *Chem Sci* **13**, 7224–7239.
 - 39 Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, Zhang X, Muruato AE, Zou J, Fontes-Garfias CR *et al.* (2021) Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **592**, 116–121.
 - 40 Henry ER, Cellmer T, Dunkelberger EB, Metaferia B, Hofrichter J, Li Q, Ostrowski D, Ghirlando R, Louis JM, Moutereau S *et al.* (2020) Allosteric control of hemoglobin S fiber formation by oxygen and its relation to the pathophysiology of sickle cell disease. *Proc Natl Acad Sci USA* **117**, 15018–15027.
 - 41 Weinkam P & Sali A (2013) Mapping polymerization and allostery of hemoglobin S using point mutations. *J Phys Chem B* **117**, 13058–13068.
 - 42 Liu Z, Gillis TG, Raman S & Cui Q (2024) A parameterized two-domain thermodynamic model explains diverse mutational effects on protein allostery. *elife* **12**, RP92262.
 - 43 Deans EE, Kotler JLM, Wei W-S & Street TO (2022) Electrostatics drive the molecular chaperone BiP to preferentially bind oligomerized states of a client protein. *J Mol Biol* **434**, 167638.
 - 44 Grasso D, Galderisi S, Santucci A & Bernini A (2023) Pharmacological chaperones and protein conformational diseases: approaches of computational structural biology. *Int J Mol Sci* **24**, 5819.
 - 45 Wassman CD, Baronio R, Demir Ö, Wallentine BD, Chen C-K, Hall LV, Salehi F, Lin D-W, Chung BP, Hatfield GW *et al.* (2013) Computational identification of a transiently open L1/S3 pocket for reactivation of mutant p53. *Nat Commun* **4**, 1407.
 - 46 Joerger AC, Bauer MR, Wilcken R, Baud MGJ, Harbrecht H, Exner TE, Boeckler FM, Spencer J & Fersht AR (2015) Exploiting transient protein states for

- the Design of Small-Molecule Stabilizers of mutant p53. *Structure* **23**, 2246–2255.
- 47 Pietrucci F, Vargiu AV & Kranjc A (2015) HIV-1 protease dimerization dynamics reveals a transient druggable binding pocket at the Interface. *Sci Rep* **5**, 18555.
 - 48 Bernini A, Galderisi S, Spiga O, Bernardini G, Niccolai N, Manetti F & Santucci A (2017) Toward a generalized computational workflow for exploiting transient pockets as new targets for small molecule stabilizers: application to the homogentisate 1,2-dioxygenase mutants at the base of rare disease alkaptonuria. *Comput Biol Chem* **70**, 133–141.
 - 49 Nakagome I, Kato A, Yamaotsu N, Yoshida T, Ozawa S-I, Adachi I & Hirono S (2018) Design of a new α -1-C-alkyl-DAB derivative acting as a pharmacological chaperone for β -glucocerebrosidase using ligand docking and molecular dynamics simulation. *Molecules* **23**, 2683.
 - 50 Citro V, Peña-García J, den-Haan H, Pérez-Sánchez H, Del Prete R, Liguori L, Cimmaruta C, Lukas J, Cubellis MV & Andreotti G (2016) Identification of an allosteric binding site on human lysosomal alpha-galactosidase opens the way to new pharmacological chaperones for Fabry disease. *PLoS One* **11**, e0165463.
 - 51 Yamaguchi K, Kamatari YO, Ono F, Shibata H, Fuse T, Elhelaly AE, Fukuoka M, Kimura T, Hosokawa-Muto J, Ishikawa T *et al.* (2019) A designer molecular chaperone against transmissible spongiform encephalopathy slows disease progression in mice and macaques. *Nat Biomed Eng* **3**, 206–219.
 - 52 Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S *et al.* (2002) The protein data Bank. *Acta Crystallogr D Biol Crystallogr* **58**, 899–907.
 - 53 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF chimera—A visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.
 - 54 Scouras AD & Daggett V (2011) The dynamomics rotamer library: amino acid side chain conformations and dynamics from comprehensive molecular dynamics simulations in water. *Protein Sci* **20**, 341–352.
 - 55 Olsson MHM, Søndergaard CR, Rostkowski M & Jensen JH (2011) PROPKA3: consistent treatment of internal and surface residues in empirical p K_a predictions. *J Chem Theory Comput* **7**, 525.
 - 56 Søndergaard CR, Olsson MHM, Rostkowski M & Jensen JH (2011) Improved treatment of ligands and coupling effects in empirical calculation and rationalization of p K_a values. *J Chem Theory Comput* **7**, 2284–2295.
 - 57 Case DM, Aktulga HM, Belfon K, Ben-Shalom IY, Berryman JT, Brozell SR, Cerutti DS, Cheatham TE, Cisneros GA, Cruzeiro VWD *et al.* (2022) “Amber 2022.”
 - 58 Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE & Simmerling C (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput* **11**, 3696–3713.
 - 59 Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW & Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926–935.
 - 60 Machado MR & Pantano S (2020) Split the charge difference in two! A rule of thumb for adding proper amounts of ions in MD simulations. *J Chem Theory Comput* **16**, 1367–1372.
 - 61 Götz AW, Williamson MJ, Xu D, Poole D, Le Grand S & Walker RC (2012) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. *J Chem Theory Comput* **8**, 1542–1555.
 - 62 Salomon-Ferrer R, Götz AW, Poole D, Le Grand S & Walker RC (2013) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J Chem Theory Comput* **9**, 3878–3888.
 - 63 Le Grand S, Götz AW & Walker RC (2013) SPFP: speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput Phys Commun* **184**, 374–380.
 - 64 Darden T, York D & Pedersen L (1993) Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J Chem Phys* **98**, 10089.
 - 65 Langevin P (1908) Sur la theorie du mouvement brownien. *CR Acad Sci Paris* **146**, 530.
 - 66 Ryckaert J-P, Ciccotti G & Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* **23**, 327–341.
 - 67 Case DA, Aktulga HM, Belfon K, Cerutti DS, Cisneros GA, Cruzeiro VWD, Forouzes N, Giese TJ, Götz AW, Gohlke H *et al.* (2023) AmberTools. *J Chem Inf Model* **63**, 6183–6191.
 - 68 Roe DR & Cheatham TE (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput* **9**, 3084–3095.
 - 69 R Core Team (2025) “R: A Language and Environment for Statistical Computing.”.
 - 70 Posit team (2025) RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA.
 - 71 Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J *et al.* (2019) Welcome to the Tidyverse. *Joss* **4**, 1686.
 - 72 Torchiano M (2016) “Effsize - a package for efficient effect size computation.”.

- 73 Cohen J (2013) *Statistical Power Analysis for the Behavioral Sciences*. Routledge, New York
- 74 Cohen J (1992) A power primer. *Psychol Bull* **112**, 155–159.
- 75 Crean RM, Slusky JSG, Kasson PM & Kamerlin SCL (2023) KIF—key interactions finder: A program to identify the key molecular interactions that regulate protein conformational changes. *J Chem Phys* **158**, 144114.
- 76 Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J *et al.* (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500.
- 77 The UniProt Consortium, Bateman A, Martin M-J, Orchard S, Magrane M, Adesina A, Ahmad S, Bowler-Barnett EH, Bye-A-Jee H, Carpentier D *et al.* (2025) UniProt: the universal protein knowledgebase in 2025. *Nucleic Acids Res* **53**, D609–D617.
- 78 Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T & Ben-Tal N (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* **44**, W344–W350.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. RMSF values were calculated for the wild-type and mutant PNLIP enzymes using C α atoms following alignment to the complete protein.

Fig. S2. RMSF values were calculated for the wild-type and mutant PNLIP enzymes using C α atoms following alignment to the respective domains.

Fig. S3. RMSF values of all non-hydrogen atoms for wild-type and mutant PNLIP enzymes.

Fig. S4. Per-residue propensity of alpha helix formation derived from simulations.

Fig. S5. Per-residue propensity of beta strand formation derived from simulations.

Table S1. AlphaMissense pathogenicity predictions for PNLIP variants.

Table S2. Solvent-accessible surface area (SASA) values of wild-type PNLIP and variants.