BMC Genomic Data

RESEARCH Open Access

Computational prediction of deleterious nonsynonymous SNPs in the *CTNS* gene: implications for cystinosis

Leila Adda Neggaz[1*], Amira Chahinez Dahmani[1,2], Ibtissem Derriche[3], Nawel Adda Neggaz[4] and Abdallah Boudjema[1]

Abstract

Background Cystinosis is a rare autosomal recessive lysosomal storage disorder caused by mutations in the CTNS gene, which encodes cystinosin, a lysosomal cystine transporter. These mutations disrupt cystine efflux, leading to its accumulation in lysosomes and subsequent cellular damage. While more than 140 mutations have been identified, the functional and structural impacts of many nonsynonymous single nucleotide polymorphisms (nsSNPs) remain poorly understood. Nonsynonymous SNPs are of particular interest because they can directly alter protein structure and function, potentially leading to disease. Clinically, cystinosis most often presents with renal Fanconi syndrome, photophobia and vision loss due to corneal cystine crystals, and progressive neuromuscular complications such as distal myopathy and swallowing difficulties This study aimed to identify deleterious nsSNPs in the CTNS gene and evaluate their effects on cystinosin stability, structure, and function via computational tools and molecular dynamics simulations.

Results From a dataset of 12,028 SNPs, 327 nsSNPs were identified, among which 19 were consistently classified as deleterious across multiple predictive tools, including SIFT, PolyPhen, and molecular dynamics simulations. Stability predictions revealed that most of these mutations destabilize cystinosin, with G308R and G308V located in the sixth transmembrane domain essential for transporter function having the most severe effects. Molecular dynamics simulations revealed that these mutations significantly increase local flexibility, alter hydrogen bonding patterns, and enhance solvent accessibility, resulting in structural perturbations. Notably, D305G and F142S disrupted the transmembrane domains essential for the function of cystinosin, whereas compared with the wild-type protein, G309V resulted in increased stability. Conservation analysis revealed that 16 of the 19 mutations affected highly conserved residues, indicating their crucial roles in the function of cystinosin. Additionally, protein interaction analyses suggested that mutations could impact associations with lysosomal and membrane transport proteins.

Conclusions This study identified 19 deleterious nsSNPs in the CTNS gene that impair cystinosin stability and function. These findings highlight the structural and functional importance of key residues, such as G308, D305, and F142, which play critical roles in maintaining the active conformation and transport capacity of cystinosin. These insights provide a foundation for future experimental validation and the development of targeted therapeutic strategies to mitigate the effects of pathogenic mutations in cystinosis.

*Correspondence:
Leila Adda Neggaz
addaneggaz.leila@gmail.com; leila.addaneggaz@univ-usto.dz

Full list of author information is available at the end of the article

Keywords Cystinosis, CTNS gene, Cystinosin, NsSNPs, Molecular dynamics simulations, Protein stability, Lysosomal transport, In Silico analysis

Background

Cystinosis is a rare genetic disorder classified as a lysosomal storage disease [1]. It is caused by mutations in the *CTNS* gene, which codes for the cystinosin protein responsible for transporting cystine out of lysosomes [2]. These mutations lead to the abnormal accumulation of cystine within lysosomes [3]. Cystine, the oxidized dimeric form of cysteine, accumulates as crystals within cells, gradually causing damage to various organs, particularly the kidneys, eyes, and liver [4, 5]. Cystinosis is an autosomal recessive disorder affecting approximately 1–2 individuals per 100,000 worldwide [4, 6]. It often presents early in childhood as its most severe form, nephropathic cystinosis, leading to renal Fanconi syndrome characterized by substantial loss of minerals and nutrients in the urine, which impairs growth and bone development [5, 7]. Cystinosin, a proton-coupled lysosomal cystine transporter [8], is composed of 367 amino acids [9]. It plays a crucial role in transporting cystine from the lysosome to the cytosol. Its activity relies on the proton electrochemical gradient generated by the V-ATPase, enabling cystine efflux [10]. Cystinosin belongs to the PQ-loop protein family and is characterized by conserved proline–glutamine motifs involved in various transport and nutrient signaling functions [11]. In addition to providing cysteine for the synthesis of glutathione, an essential antioxidant, cystinosin, regulates cellular responses to nutrient restriction by modulating TORC1 activity [12]. Loss of cystinosin function resulting from *CTNS* mutations disrupts cystine efflux, leading to metabolic imbalances, increased oxidative stress, and cellular damage [13]. More than 140 mutations in the *CTNS* gene have been identified to date, including nonsense mutations and deletions, which directly impact the stability and function of cystinosin [14]. Nonsense mutations, in particular, prematurely truncate protein synthesis, whereas other point mutations disrupt the conformation and functionality of cystinosin, thereby impairing cystine transport and exacerbating cystinosis pathology [15, 16]. In silico studies have revolutionized our understanding of genetic diseases such as cystinosis by enabling the analysis and prediction of mutation effects without requiring extensive and costly laboratory experiments [17]. These approaches allow researchers to identify and characterize single nucleotide polymorphisms (SNPs) with pathogenic potential and assess their impact on protein structure and stability [18, 19]. Predictive tools, such as POLYPHEN-2, SIFT, and I-Mutant, are commonly used to filter deleterious mutations [20, 21], whereas advanced methods, such as molecular dynamics simulations, help in the study of precise mutation-induced effects on the stability and interactions of cystinosin under simulated biological conditions. These methods provide a detailed view of the structural and functional alterations caused by mutations. This study relied on in silico analyses to identify genetic mutations with significant functional impact, particularly nonsynonymous SNPs (nsSNPs) likely to be deleterious. Unlike previous studies that focused solely on identifying deleterious SNPs, our approach integrates conservation analysis, stability predictions, and molecular dynamics simulations to provide a comprehensive understanding of mutation-induced dysfunctions in cystinosin. In the present study, we aimed to identify the most deleterious missense variants in the human *CTNS* gene by applying bioinformatics approaches and molecular dynamics simulations to elucidate how these mutations influence the structure, stability, and function of cystinosin, which contributes to the pathogenesis of cystinosis. This study provides valuable insights into the effects of amino acid variations on the structure, function, and ease of association of the cystinosin protein.

Materials and methods

Retrieval of the *CTNS* NsSNP dataset

All the variants of the *CTNS* gene were collected from the Ensembl variant ID ENSG00000040531. The systinosin primary sequence (UniProt accession number: O60931) was retrieved from the UniProt database [22]. The 3D structure of cystinosin used in this study was retrieved from the AlphaFold protein structure database (AF-O60931-F1-v4) [23]. The AlphaFold model showed high per-residue confidence (pLDDT >90) in the transmembrane core of the protein, while lower confidence scores were observed at the N- and C-terminal extremities. Mutant structures corresponding to the selected deleterious nsSNPs were generated using UCSF ChimeraX by substituting the corresponding residues in the AlphaFold model.

Prediction of deleterious NsSNPs

Prediction of the functional consequences of NsSNPs

The functional effects of the identified nsSNPs were evaluated via several tools, some of which were preintegrated into the Ensembl database. These tools include the following:

- SIFT (sorting intolerance from tolerance) [24]: SIFT predicts the impact of amino acid substitutions on the basis of sequence homology and physicochemical

properties. Variants with a score less than 0.05 are considered deleterious [25].

- PolyPhen (Polymorphism Phenotyping) [26] PolyPhen-2 assesses the impact of amino acid substitutions on protein structure and function. Scores between 0.96 and 1.0 indicate "probably damaging," whereas scores between 0.71 and 0.95 suggest "possibly damaging" [27].
- CADD (Combined Annotation Dependent Depletion) [28]: CADD integrates multiple annotations to assess variant deleteriousness. Higher CADD scores indicate greater potential for harmful effects [29].
- MetaLR (Meta Logistic Regression) [30]: MetaLR uses logistic regression to categorize nsSNPs as damaging (score 0.5–0.9) or tolerated (score 0–0.4). Only variants with scores ≥ 0.5 were selected for further analysis [31].
- Mutation Assessor [32]: This tool evaluates the functional impact of mutations on the basis of their frequency and position within the protein. Higher scores indicate more deleterious mutations [33].

Additionally, we used the following tools for independent analyses:

- PROVEAN [34]: PROVEAN predicts the functional impact of variants, with scores ≤ – 2.5 considered deleterious [35].
- SNPs&GO [36]: This tool predicts disease-related mutations from protein sequences, with scores > 0.5 indicating a disease-related effect [37].
- PhD-SNP () [38]: A support vector machine (SVM)-based classifier that predicts disease-associated nsSNPs with 78% accuracy [39].
- PANTHER [40]: PANTHER predicts pathogenic substitutions, with sub PSEC scores below – 3 indicating deleterious variants at highly conserved sites [41].

This combination of tools allowed for a comprehensive evaluation of the potential effects of nsSNPs on protein function.

Structural impact prediction

The potential impact of nsSNPs on protein stability was predicted via the following tools:

- I-Mutant 2 [42]: I-Mutant 2.0 predicts changes in protein stability (ΔΔG) upon mutation. Negative ΔΔG values indicate decreased stability, whereas positive values suggest increased stability [43].
- MUPro [44]: MUPro provides confidence scores for stability predictions, with scores closer to 0 indicating decreased stability [45].
- DynaMut2 [46]: DynaMut2 predicts the effect of nsSNPs on protein stability via molecular dynamics simulations, providing stability scores and vibrational entropy changes [47].
- DUET [48]: DUET combines mCSM and SDM to predict the impact of nsSNPs on protein stability, classifying mutations as stabilizing or destabilizing on the basis of changes in folding free energy (ΔΔG) [49].

These tools provided detailed insights into the structural and functional consequences of the identified nsSNPs.

Protein–protein interactions

To understand how nsSNPs may impact protein interactions, we used GeneMania and string: GeneMANIA identifies potential gene–gene and protein–protein interaction networks by integrating multiple data sources, including physical interactions, co-expression, co-localization, and shared protein domains [50]. The tool has been extensively benchmarked and is widely used for functional gene prioritization, with demonstrated reliability in studies involving various biological processes, including those related to lysosomal function [51]. STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) [52] is a database and web tool designed to predict and visualize protein–protein interactions. It integrates information from various sources, including experimental data, curated databases, and computational predictions, to provide a comprehensive view of functional associations between proteins. For this study, a high-confidence threshold of 0.700 was applied. The interaction types are represented by color-coded lines: green for activation, red for inhibition, blue for binding, pink for posttranslational modifications, and yellow for expression [53].

Prediction of evolutionary conservation

ConSurf [54]: ConSurf calculates the evolutionary conservation of each residue by analyzing homologous sequences. Residues with high conservation scores (close to 9) are typically functionally or structurally important, whereas variable positions (score close to 1) are less likely to be critical for function [55].

Molecular dynamics simulation analysis

Molecular dynamics (MD) simulations offer a powerful tool for exploring the evolution of molecular systems and predicting their properties on the basis of atomic interactions. Simulations of the predicted structures were conducted via the WebGro server to evaluate their stability and flexibility [56]. The AMBER99SB-ILDN force field,

combined with the TIP4P water model, was applied in a triclinic periodic box with NaCl at a concentration of 0.15 M to simulate a biologically relevant environment. The system was equilibrated at 300 K and a pressure of 1.0 bar. The simulation was run for 50 ns, generating 5000 frames per simulation. The Leap frog integrator was used within the NVT and NPT ensembles to control the temperature and pressure. To analyze the results, we calculated the RMSD (root mean square deviation), which measures structural stability. RMSF (root mean square fluctuation): Assesses the flexibility of individual residues. Hydrogen bonds (H-bonds) analysis: Indicates the stability of protein interactions. Radius of gyration (Rg) and solvent accessible surface area (SASA): These findings provide insights into protein compactness and the effects of mutations on protein surface properties [57].

Results data

Following the analysis of the SNP dataset, a total of 12,028 SNPs were identified. The analysis revealed the following distributions: 327 missense variants (2.72%), 163 synonymous variants (1.36%), and 9,383 intron variants (78.01%). Additionally, 265 SNPs were located in the 5' UTR (2.20%), and 1,326 were located in the 3' UTR (11.02%). These findings provide a comprehensive overview of the SNP distribution within the dataset, emphasizing regions that may hold potential functional significance for further investigation.

Prediction of deleterious NsSNPs

Prediction of the functional consequences of NsSNPs

The impact of 327 nsSNPs on the structure and function of cystinosin was investigated using five widely adopted in silico prediction algorithms: SIFT, PolyPhen-2, CADD, MetaLR, and Mutation Assessor. These tools were selected for their methodological complementarity and proven effectiveness in assessing missense variant pathogenicity. SIFT and Mutation Assessor rely on evolutionary conservation, PolyPhen-2 integrates both sequence and structural features, CADD employs a machine-learning model aggregating diverse annotations, and MetaLR combines multiple scores through logistic regression. This diverse strategy improves the reliability of consensus predictions. Among the analyzed variants, 19 were consistently identified as deleterious by all five tools, suggesting a high probability of functional disruption. These nsSNPs were classified as "deleterious" by SIFT (scores ≤ 0.05), "probably damaging" or "possibly damaging" by PolyPhen-2 (scores near 1), had CADD scores exceeding 30, were considered "likely deleterious" by MetaLR, and "damaging" by Mutation Assessor (functional impact scores ≥ 0.97) (Table 1).

The clinical significance of these mutations, as obtained from the Ensembl database, revealed a range of classifications from 'uncertain significance' to 'pathogenic'. Notably, G308R and G308 V are classified as pathogenic, which is consistent with their severe destabilizing effects on cystinosin. Similarly, S141 F and S270 F were both classified as likely pathogenic. S141 F is located in the third transmembrane domain (TM3; residues 126–150), and S270 F is in the PQ-loop 2 domain (residues 263–328). On the other hand, variants such as A137D, D161H, G296S and D305G were classified as having uncertain significance, highlighting the need for further functional studies to clarify their role in the disease.

The 19 nsSNPs identified as potentially deleterious were further evaluated via SNPs&GO, PhD-SNP, PANTHER, and PROVEAN (Table 2). These tools consistently classify most variants as "Disease" or "probably damaging," with PROVEAN scores indicating a significant potential for functional impact. This consistent prediction across multiple platforms underscores the likelihood that these variants could disrupt protein function, making them strong candidates for further investigation.

Prediction of stability for deleterious NsSNPs

The stability of the 19 deleterious nsSNPs in the cystinosin protein was analyzed via tools such as I-Mutant 2.0, MUPro, mCSM, SDM, DUET, and DynaMut2 (Table 3). The predictions from these tools generally indicate a tendency for structural destabilization induced by these mutations.

- I-Mutant 2.0 predicted destabilizing effects for 16 out of the 19 mutations, with F142S exhibiting the most negative ΔΔG value (−2.50 kcal/mol), indicating a significant loss of stability. In contrast, three mutations (A137D, S141 F, and H353D) had stabilizing effects.
- MUPro revealed that 18 mutations had destabilizing effects, with R151G having the most marked impact (ΔΔG = −2.00 kcal/mol). G308 V was the only mutation showing marginal stabilization (ΔΔG = + 0.01 kcal/mol).
- DynaMut2 predicted destabilizing effects for all the mutants, with stability changes (ΔΔG) ranging from − 0.05 kcal/mol to −2.62 kcal/mol.
- mCSM predictions indicated destabilizing effects for all the mutations, with F142S (ΔΔG = −2.813 kcal/mol) and G308R (ΔΔG = −1.13 kcal/mol) standing out as the most critical mutations.
- SDM analysis identified 11 destabilizing mutations, with G296S showing the most severe effect (ΔΔG = −4.15 kcal/mol).
- DUET confirmed these findings, identifying 18 destabilizing mutations, with the most negative values for F142S (ΔΔG = −2.897 kcal/mol) and D305G (ΔΔG = −1.415 kcal/mol). Only the P283L mutation had a stabilizing effect, with a ΔΔG of + 0.172 kcal/mol.

Table 1 Missense variants predicted as deleterious by multiple algorithms (SIFT, polyphen, CADD, MetaLR, mutation assessor)

ID of SNPs	Clin. Sig.	SUB	SIFT		PolyPhen		CADD		MetaLR		Mutation Assessor	
			Score	Pred	Score	Pred	Score	Pred	Score	Pred	Score	Pred
rs1413852868	uncertSig	A137D	0	Del	0.956	ProbDam	32	LikDel	0.963	Dam	0.954	high
rs1436441738	LikPath ~ Path	S141 F	0	Del	1	ProbDam	33	LikDel	0.975	Dam	0.95	high
rs764288123	Indet	F142S	0.02	Del	0.991	ProbDam	32	LikDel	0.971	Dam	0.954	high
rs1555563010	LikPath	R151G	0	Del	0.836	PossDam	32	LikDel	0.968	Dam	0.966	high
rs371533565	uncertSig	D161H	0	Del	0.999	ProbDam	32	LikDel	0.974	Dam	0.933	high
rs2150925336	LikPath	S270 F	0	Del	0.998	ProbDam	32	LikDel	0.963	Dam	0.932	high
rs2150925451	uncertSig/LikPath/Path	K280 T	0	Del	0.998	ProbDam	34	LikDel	0.972	Dam	0.951	high
rs2150925451	uncertSig/LikPath/Path	K280R	0	Del	0.997	ProbDam	35	LikDel	0.972	Dam	0.951	high
rs746469285	Indet	P283L	0	Del	0.999	ProbDam	31	LikDel	0.999	Dam	0.951	high
rs1238793405	Indet	Q284H	0	Del	0.999	ProbDam	36	LikDel	0.999	Dam	0.839	high
rs759363199	Indet	S293G	0	Del	0.993	ProbDam	33	LikDel	0.983	Dam	0.95	high
rs755702977	uncertSig	G296S	0	Del	1	ProbDam	32	LikDel	0.988	Dam	0.951	high
rs1332086669	Indet	S298G	0	Del	0.992	ProbDam	30	LikDel	0.986	Dam	0.946	high
rs1263951539	uncertSig	D305G	0	Del	1	ProbDam	32	LikDel	0.973	Dam	0.951	high
rs746307931	Path	G308R	0	Del	1	ProbDam	32	LikDel	0.988	Dam	0.951	high
rs908965524	Path	G308 V	0	Del	1	ProbDam	32	LikDel	0.988	Dam	0.951	high
rs373903147	Indet	G309 V	0	Del	1	ProbDam	31	LikDel	0.973	Dam	0.95	high
rs769659540	Indet	H353D	0	Del	0.999	ProbDam	31	LikDel	0.968	Dam	0.961	high
rs1177103234	Indet	Y357H	0	Del	0.998	ProbDam	31	LikDel	0.986	Dam	0.962	high

ClinSig Clinical Significance, *Del* Deleterious, *High* High Functional Impact, *Indet* Indeterminate, *LikDel* Likely Deleterious, *LikPath* Likely Pathologic, *Path* Pathologic, *PossDam* Possibly Damaging, *Pred* Prediction, *ProbDam* Probably Damaging, *SUB* Amino acid substitution, *UncertSig* Uncertain Significance

Table 2 Missense variants predicted as deleterious by SNPs&GO, PhD-SNP, PANTHER, and PROVEAN

SUB	PhD-SNP			SNPs&GO			PROVEAN		PANTHER	
	Score	RI	Pred	Score	RI	Pred	Score	Pred	Score	Pred
A137D	0.922	8	Dis	0.879	8	Dis	−5.533	Del	0.85	ProbDam
S141 F	0.963	9	Dis	0.933	9	Dis	−5.771	Del	0.89	ProbDam
F142S	0.935	9	Dis	0.855	7	Dis	−7.784	Del	0.89	ProbDam
R151G	0.816	6	Dis	0.704	4	Dis	−5.376	Del	0.5	ProbDam
D161H	0.895	8	Dis	0.790	6	Dis	−6.701	Del	0.89	ProbDam
S270 F	0.883	8	Dis	0.436	1	Neut	−6.000	Del	0.85	ProbDam
K280 T	0.886	8	Dis	0.798	6	Dis	−6.000	Del	0.89	ProbDam
K280R	0.864	7	Dis	0.714	4	Dis	−3.000	Del	0.89	ProbDam
P283L	0.934	9	Dis	0.823	6	Dis	−9.767	Del	0.89	ProbDam
Q284H	0.839	7	Dis	0.722	4	Dis	−5.000	Del	0.89	ProbDam
S293G	0.705	4	Dis	0.621	2	Dis	−4.000	Del	0.89	ProbDam
G296S	0.883	8	Dis	0.811	6	Dis	−6.000	Del	0.89	ProbDam
S298G	0.768	5	Dis	0.591	2	Dis	−4.000	Del	0.85	ProbDam
D305G	0.866	7	Dis	0.801	6	Dis	−7.000	Del	0.89	ProbDam
G308R	0.900	8	Dis	0.829	7	Dis	−8.000	Del	0.89	ProbDam
G308 V	0.869	7	Dis	0.814	6	Dis	−9.000	Del	0.89	ProbDam
G309 V	0.811	6	Dis	0.762	5	Dis	−9.000	Del	0.89	ProbDam
H353D	0.931	9	Dis	0.691	4	Dis	−9.000	Del	0.89	ProbDam
Y357H	0.916	8	Dis	0.757	5	Dis	−4.967	Del	0.89	ProbDam

Del Deleterious, *Dis* Disease, *Neut* Neutral, *Pred* Prediction, *ProbDam* Probably Damaging, *RI* Reliability Index, *SUB* Amino acid substitution

Table 3 Stability predictions for deleterious NsSNPs via I-Mutant 2.0, MUPro, DynaMut2, mCSM, SDM, and DUET

ID of SNPs	SUB	I-Mutant2.0 DDG (Pred)	Mupro DDG (Pred)	DynaMut2 DDG (Pred)	mCSM Score(Pred)	SDM Score(Pred)	DUET Score (Pred)
rs1413852868	A137D	0.29 (IS)	−0.72 (DS)	−1.47 (DS)	−1.94 (DS)	−2.53 (DS)	−2.18 (DS)
rs1436441738	S141 F	0.57 (IS)	−0.45 (DS)	−0.66 (DS)	−0.79 (DS)	2.40 (IS)	−0.01 (DS)
rs764288123	F142S	−2.50 (DS)	−1.73 (DS)	−2.62 (DS)	−2.81 (DS)	−2.23 (DS)	−2.89 (DS)
rs1555563010	R151G	−2.23 (DS)	−2.00 (DS)	−0.78 (DS)	−0.61 (DS)	−0.85 (DS)	−0.74 (DS)
rs371533565	D161H	−1.74 (DS)	−1.89 (DS)	−1.07 (DS)	−0.89 (DS)	0.48 (IS)	−0.62 (DS)
rs2150925336	S270 F	−0.32 (DS)	−0.21 (DS)	−0.91 (DS)	−1.15 (DS)	0.84 (IS)	−0.78 (DS)
rs2150925451	K280 T	−1.37 (DS)	−0.51 (DS)	−0.94 (DS)	−1.34 (DS)	−1.12 (DS)	−1.40 (DS)
rs2150925451	K280R	−0.50 (DS)	−0.14 (DS)	−0.29 (DS)	−0.71 (DS)	0.39 (IS)	−0.38 (DS)
rs746469285	P283L	−1.47 (DS)	−0.21 (DS)	−0.54 (DS)	−0.47 (DS)	1.57 (IS)	0.17 (IS)
rs1238793405	Q284H	−1.57 (DS)	−0.37 (DS)	−0.72 (DS)	−0.78 (DS)	0.61 (IS)	−1.08 (DS)
rs759363199	S293G	−1.49 (DS)	−0.96 (DS)	−0.50 (DS)	−1.10 (DS)	0.72 (IS)	−0.72 (DS)
rs755702977	G296S	−1.57 (DS)	−1.02 (DS)	−0.05 (DS)	−0.64 (DS)	−4.15 (DS)	−1.06 (DS)
rs1332086669	S298G	−1.92 (DS)	−1.17 (DS)	−0.22 (DS)	−0.58 (DS)	0.72 (IS)	−0.18 (DS)
rs1263951539	D305G	−1.76 (DS)	−1.87 (DS)	−0.75 (DS)	−1.42 (DS)	−0.46 (DS)	−1.41 (DS)
rs746307931	G308R	−1.50 (DS)	−0.27 (DS)	−0.70 (DS)	−1.13 (DS)	−1.89 (DS)	−1.11 (DS)
rs908965524	G308 V	−1.05 (DS)	0.01 (IS)	−0.91 (DS)	−0.40 (DS)	−0.14 (DS)	−0.11 (DS)
rs373903147	G309 V	−0.43 (DS)	−0.42 (DS)	−0.96 (DS)	−0.32 (DS)	−0.14 (DS)	−0.03 (DS)
rs769659540	H353D	0.37 (IS)	−0.36 (DS)	−0.59 (DS)	−0.83 (DS)	−0.81 (DS)	−0.77 (DS)
rs1177103234	Y357H	−2.05 (DS)	−1.25 (DS)	−1.19 (DS)	−1.26 (DS)	−0.45 (DS)	−1.08 (DS)

SUB Amino acid substitution, *Pred* Prediction, *DS* Decrease stability, *IS* Increase Stability

Functional analysis via MutPred2 provided detailed insights into the structural and functional alterations caused by the deleterious nsSNPs in cystinosin (Table 4). The amino acid change D305G emerged as one of the most critical variants, with the highest MutPred2 score (0.945), and was associated with the loss of a helix (p = 0.01). alteration of transmembrane protein properties (p = 1.0e-03), and modification of metal binding (p = 0.01). These changes suggest that D305G could severely disrupt both the structure and function of cystinosin. Similarly, F142S significantly altered transmembrane function (p = 4.4e-05), indicating that it could impair the protein's ability to function properly within the membrane. In addition, P283L was linked to a gain of a helix

Table 4 Predicted effects of genetic variants via MutPred2

ID of SNPs	SUB	Score	Effects	Probability	*P* value
rs1413852868	A137D	0.575	Altered Transmembrane protein	0.31	1.2e-04
rs1436441738	S141 F	0.852	Altered Transmembrane protein	0.31	1.1e-04
rs764288123	F142S	0.811	Altered Transmembrane protein	0.34	4.4e-05
rs1555563010	R151G	0.589	Gain of Strand	0.28	0.01
	R151G	0.589	Altered Transmembrane protein	0.27	6.3e-04
	R151G	0.589	Altered Stability	0.11	0.04
rs371533565	D161H	0.872	Altered Transmembrane protein	0.30	1.5e-04
	D161H	0.872	Loss of N-linked glycosylation at N166	0.07	0.02
rs2150925336	S270 F	0.917	Altered Transmembrane protein	0.27	6.5e-04
rs2150925451	K280 T	0.790	Altered Ordered interface	0.24	0.03
	K280 T	0.790	Altered Transmembrane protein	0.11	0.03
rs2150925451	K280R	0.594	Altered Transmembrane protein	0.11	0.04
rs746469285	P283L	0.819	Gain of Helix	0.33	1.7e-03
	P283L	0.819	Loss of Strand	0.29	3.6e-03
	P283L	0.819	Altered Transmembrane protein	0.14	0.02
rs1238793405	Q284H	0.700	Altered Transmembrane protein	0.13	0.02
rs759363199	S293G	0.830	Gain of Strand	0.27	0.02
	S293G	0.830	Altered Ordered interface	0.25	0.02
	S293G	0.830	Altered Transmembrane protein	0.22	2.8e-03
rs755702977	G296S	0.927	Altered Ordered interface	0.30	0.02
	G296S	0.927	Altered Transmembrane protein	0.25	1.5e-03
rs1332086669	S298G	0.895	Altered Transmembrane protein	0.24	1.9e-03
rs1263951539	D305G	0.945	Loss of Helix	0.29	0.01
	D305G	0.945	Altered Transmembrane protein	0.26	1.0e-03
	D305G	0.945	Altered Metal binding	0.23	0.01
rs746307931	G308R	0.932	Altered Metal binding	0.20	0.02
	G308R	0.932	Altered Transmembrane protein	0.13	0.02
rs908965524	G308 V	0.939	Altered Metal binding	0.20	0.02
	G308 V	0.939	Altered Transmembrane protein	0.13	0.02
rs373903147	G309 V	0.892	Altered Metal binding	0.20	0.02
	G309 V	0.892	Altered Transmembrane protein	0.13	0.02
rs769659540	H353D	0.823	Altered Metal binding	0.36	8.3e-03
	H353D	0.823	Altered Ordered interface	0.28	0.03
	H353D	0.823	Loss of Strand	0.28	0.01
	H353D	0.823	Gain of Allosteric site at Y357	0.20	0.04
rs1177103234	Y357H	0.864	Altered Ordered interface	0.35	5.4e-03
	Y357H	0.864	Gain of Intrinsic disorder	0.35	0.02
	Y357H	0.864	Altered Metal binding	0.33	0.01
	Y357H	0.864	Altered Transmembrane protein	0.10	0.04

SUB Amino acid substitution, *P* value < 0.05 (significant)

(p = 1.7e-03) and a loss of a β-strand (p = 3.6e-03), suggesting substantial structural rearrangements while also altering transmembrane protein properties (p = 0.02). R151G was predicted to cause a gain of strand (p = 0.01) and alter transmembrane protein properties (p = 6.3e-04), in addition to reducing stability (p = 0.04), which could destabilize the protein and modify its functional properties. Furthermore, G308R and G308 V were both associated with altered metal binding (p = 0.02) and transmembrane protein properties (p = 0.02), potentially disrupting critical interactions within the protein. Another notable variant, H353D, showed multiple effects, including altered metal binding (p = 8.3e-03), loss of a strand (p = 0.01), and gain of an allosteric site at Y357 (p = 0.04), suggesting significant structural and functional alterations. Finally, Y357H was predicted to alter ordered interfaces (p = 5.4e-03), increase intrinsic disorder (p = 0.02), and modify metal binding (p = 0.01), indicating potential disruptions to the protein's structural integrity and functional interactions. Overall, these findings demonstrate the diverse and significant effects of these variants on the structural and functional properties of cystinosin.

Gene and protein interaction networks involving CTNS

Gene–gene interaction network analysis revealed a functional correlation between the *CTNS* gene and other genes. Using GeneMANIA, a composite gene–gene functional interaction network was constructed, highlighting associations between *CTNS* and 20 other genes, particularly SLC66 A1 (solute carrier family 66 member 1), MPDU1 (mannose-P-dolichol utilization defect 1), and SLC66 A2 (solute carrier family 66 member 2). The interaction of *CTNS* with all possible genes is shown in Fig. 1.

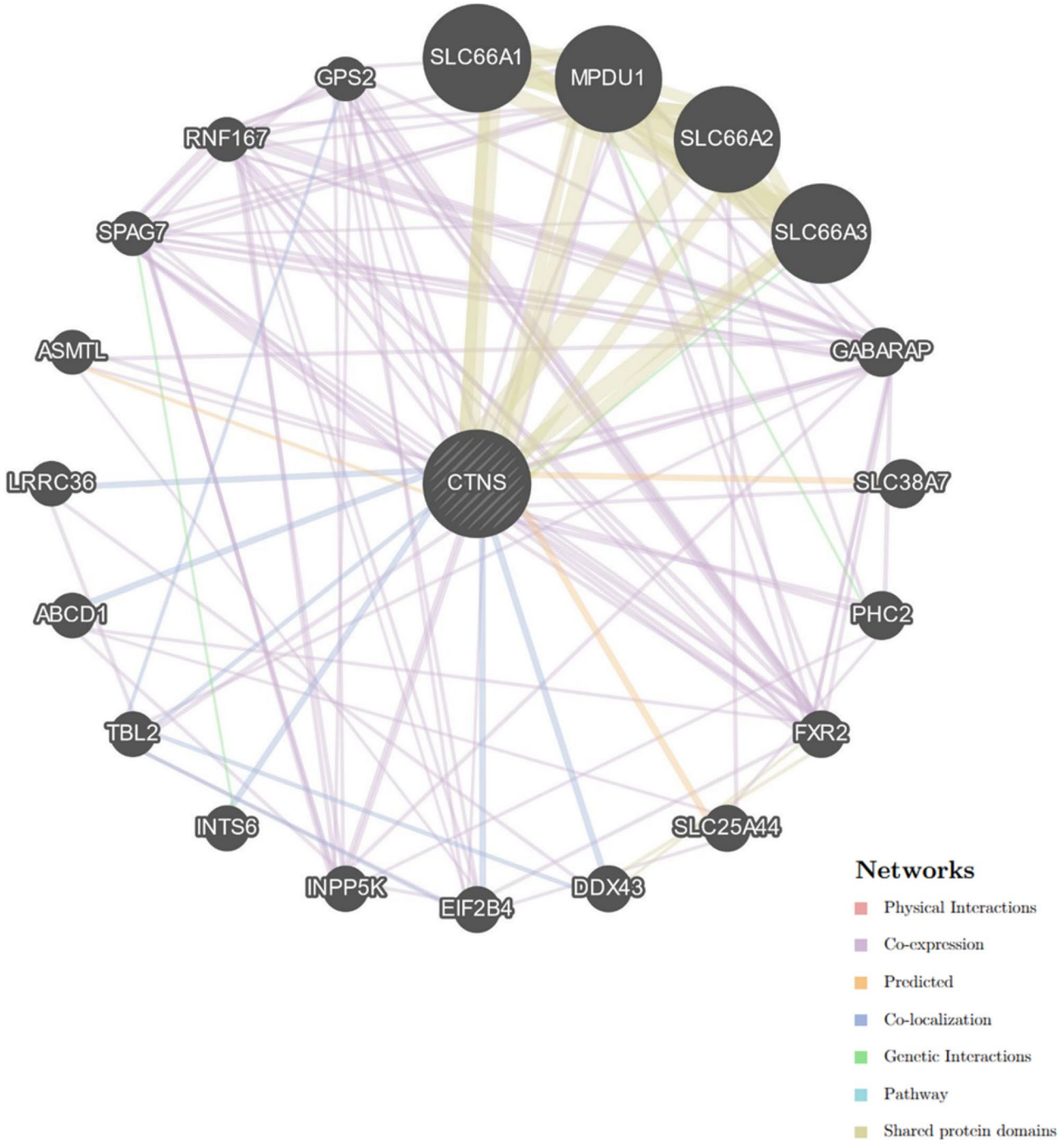

Fig. 1 Gene interaction network of CTNS predicted by GeneMANIA. The network was generated via GeneMANIA [35] and illustrates functional associations between CTNS and 20 other genes, including SLC66 A1, MPDU1, and SLC66 A2. Key interactions are highlighted, emphasizing their roles in lysosomal function and cellular metabolism

The analysis of protein interactions via the STRING platform revealed a functional network involving cystinosin and several biologically relevant partners, such as SLC66 A1, SHPK, and TRPV1. These interactions suggest an expanded role of CTNS in cellular processes related to membrane transport, metabolism, and lysosomal dynamics. The illustrated network (Fig. 2.) highlights connections on the basis of experimental data, coexpression data, and bioinformatics predictions.

Prediction of evolutionary conservation

The positions of the mutations were examined for evolutionary conservation via ConSurf (Fig. 3). Sixteen of the 19 mutations are located at highly conserved positions (score 9), highlighting that their structural and functional importance mutations, G308R, G308 V, G296S, and K280 T, affect exposed residues. In contrast, mutations D305G, G309 V, and F142S impact buried residues.

Evolutionary conservation and classification of residues as buried (b), exposed (e), functional (f), and structural (s).

Global stability (RMSD)

The global stability of cystinosin was assessed via the root mean square deviation (RMSD) (Fig. 4). The wild-type (WT) cystinosin presented an average RMSD of 0.664 ± 0.189 nm, demonstrating a stable conformation throughout the simulation. Among the mutants, G308 V exhibited the highest RMSD (0.991 ± 0.197 nm), highlighting significant deviation from the initial structure

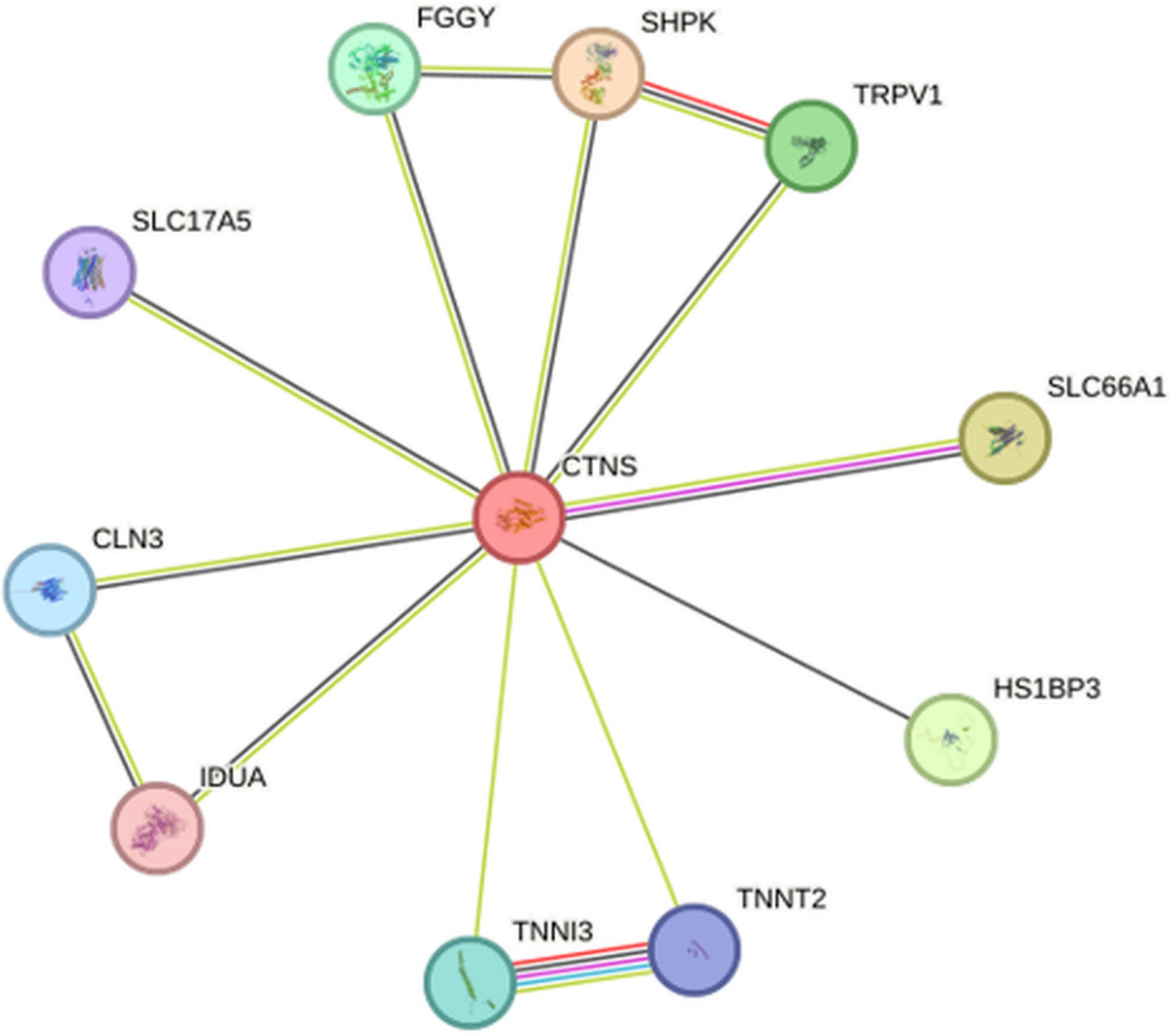

Fig. 2 Protein interaction network of CTNS analyzed via STRING. The network highlights functional associations between CTNS and key partners, including SLC66 A1, SHPK, and TRPV1, involved in membrane transport, metabolism, and lysosomal dynamics

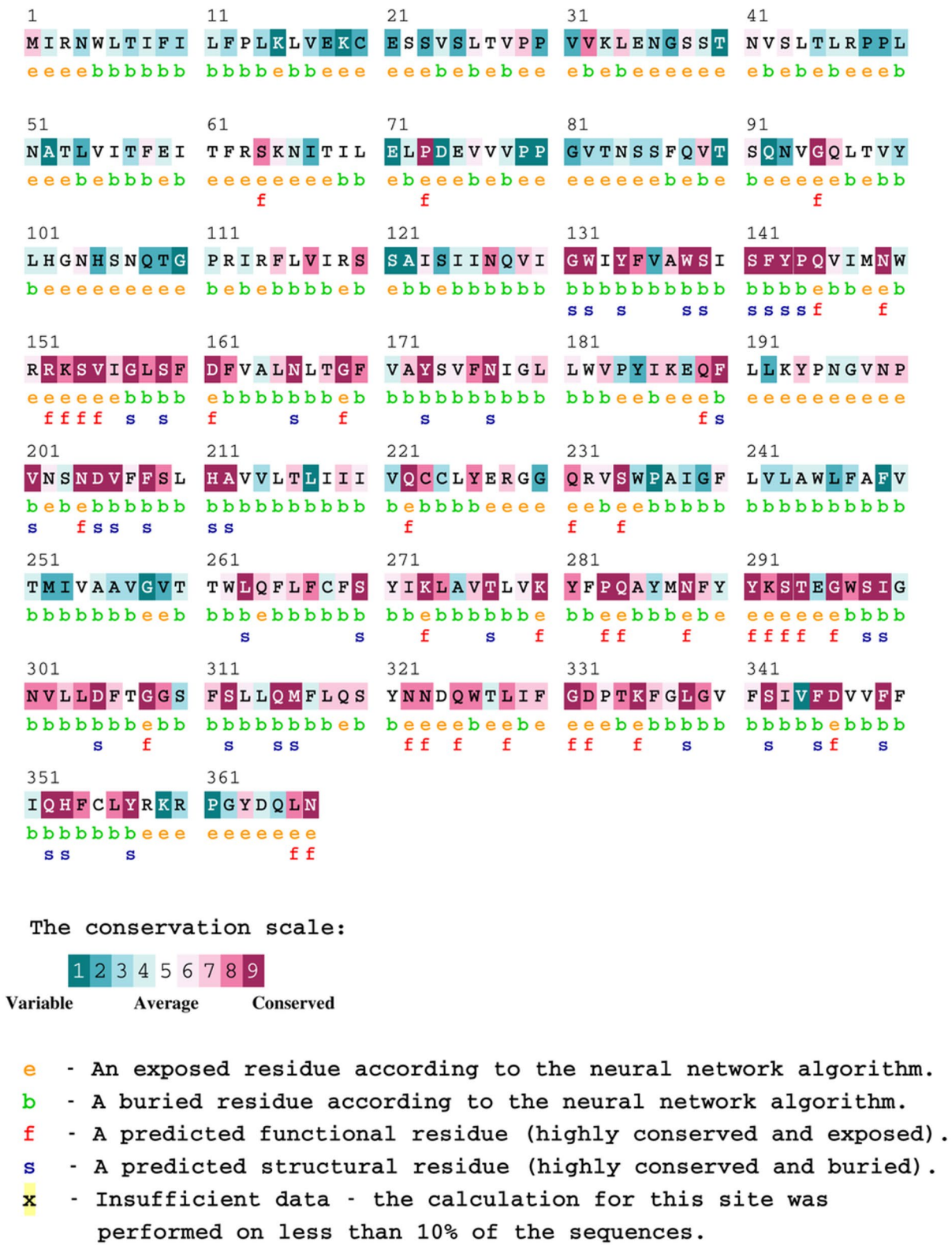

Fig. 3 Evolutionary conservation of cystinosin residues was analyzed via ConSurf

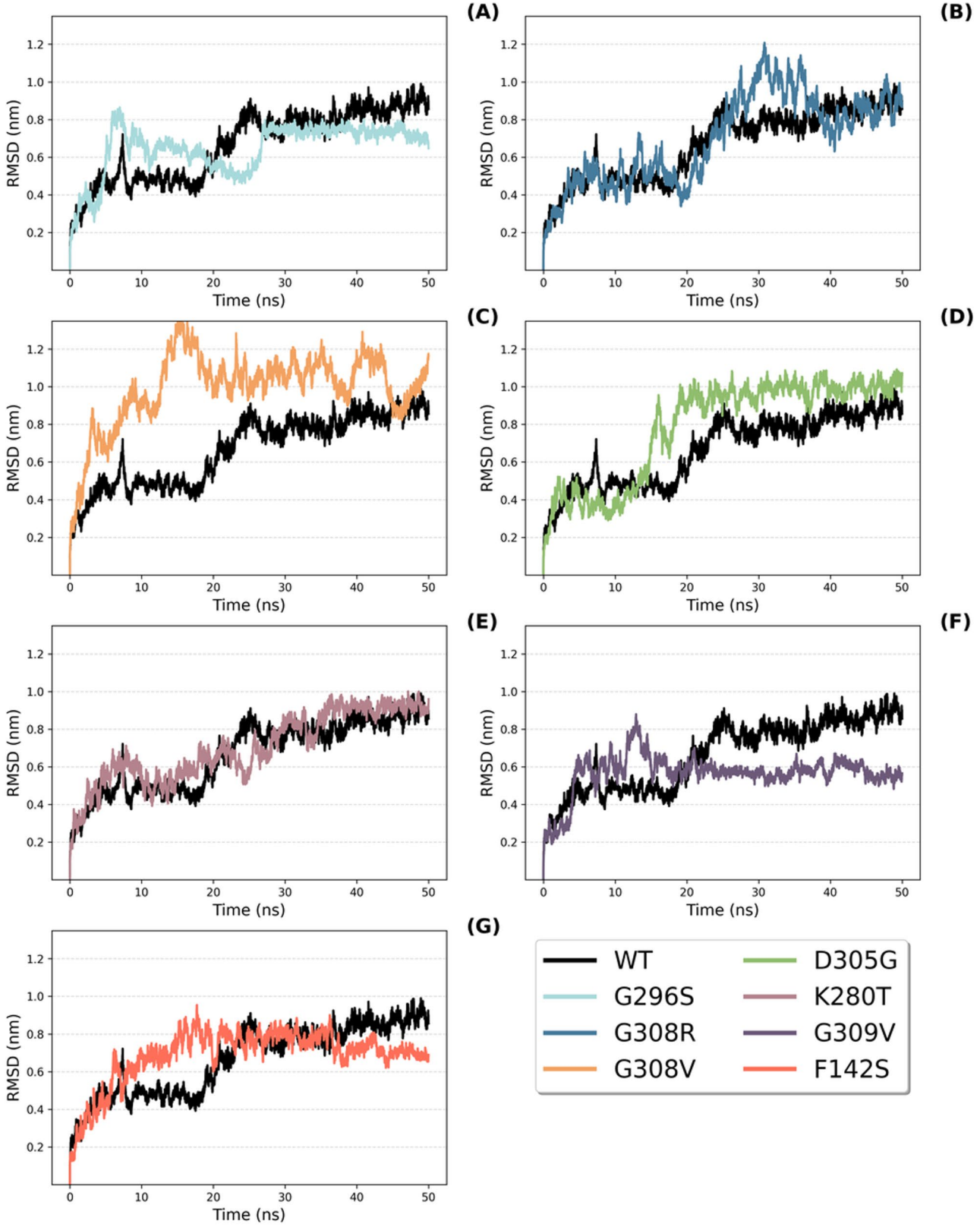

Fig. 4 RMSD analysis of wild-type CTNS and its variants. The Cα-backbone RMSD of the CTNS protein over time is shown. The y-axis represents the RMSD in nanometers (nm), and the x-axis represents the time in nanoseconds (ns). **A** WT (black) and G296S (light blue). **B** WT (black) and G308R (pastel blue). **C** WT (black) and G308 V (soft orange). **D** WT (black) and D305G (pale green). **E** WT (black) and K280 T (soft pink). **F** WT (black) and G309 V (pastel violet). **G** WT (black) and F142S (coral red)

and marked destabilization. In contrast, G309 V displayed the lowest RMSD (0.565 ± 0.105 nm), reflecting increased stability. Other mutants, such as G296S (0.647 ± 0.133 nm) and F142S (0.687 ± 0.147 nm), maintained an RMSD close to that of the WT, suggesting preserved global stability.

Local flexibility (RMSF)

Local flexibility was assessed via the root mean square fluctuation. The WT had an average RMSF of 0.275 ± 0.256 nm, reflecting balanced residue flexibility. Among the mutants, D305G (0.310 ± 0.332 nm) and G308R (0.314 ± 0.295 nm) presented the greatest fluctuations, indicating increased local instability. In contrast, G309 V (0.217 ± 0.194 nm) and G296S (0.241 ± 0.247 nm) exhibited lower fluctuations, highlighting increased rigidity in certain regions of the protein (Fig. 5).

Compactness (radius of gyration, Rg)

The compactness of cystinosin was evaluated via the radius of gyration. The WT had an average Rg of 2.793 ± 0.043 nm, corresponding to a compact and stable structure. The mutants G308 V (2.967 ± 0.060 nm) and G308R (2.911 ± 0.078 nm) presented significantly greater Rg values, indicating a loss of compactness and structural destabilization. Conversely, G296S (2.659 ± 0.063 nm) and G309 V (2.685 ± 0.047 nm) displayed a lower Rg, implying structural rigidity that could affect the flexibility required for cystine transport (Fig. 6).

Solvent accessible surface area (SASA)

The solvent-accessible surface area was analyzed to evaluate the exposure of hydrophobic and hydrophilic residues. The WT had an average SASA of 208.032 ± 4.725 nm^2. The mutants G308R (214.130 ± 5.101 nm^2) and F142S (211.930 ± 4.502 nm^2) presented increased SASA, suggesting increased residue exposure to the solvent, which could disrupt hydrophobic interactions. Conversely, G296S (201.788 ± 6.940 nm^2) and G309 V (204.475 ± 4.382 nm^2) displayed a lower SASA, reflecting reorganization of the accessible surface (Fig. 7).

Hydrogen bonds (H-bonds)

The number of hydrogen bonds was measured to assess the stability of protein interactions. The WT maintained an average of 265.766 ± 9.986 hydrogen bonds, reflecting a well-stabilized structure. The mutants G308R (258.615 ± 8.373) and G308 V (259.370 ± 9.569) presented a slight decrease in the number of H-bonds, which could contribute to their destabilization. In contrast, the number of H-bonds in G309 V (268.274 ± 7.633) was slightly greater than that in WT, indicating increased stabilization (Fig. 8).

Discussion

The study of genetic mutations in single-gene disorders, such as cystinosis, has greatly advanced owing to the use of bioinformatics tools [58]. Cystinosis is caused by mutations in the *CTNS* gene, making it essential to evaluate the functional impact of these mutations to understand their pathogenicity and potential therapeutic implications. With this in mind, our study relied on in silico analyses to identify genetic mutations with a significant functional impact, particularly nsSNPs likely to be deleterious [59, 60]. The importance of these analyses lies in their ability to predict the pathogenicity of variants and assess their impact on protein stability and function, which is essential for a better understanding of the underlying mechanisms of monogenic diseases such as cystinosis. In this study, we performed in silico structural and functional analyses to identify potentially deleterious nsSNPs. From the dataset of 12,028 SNPs, 327 (2.72%) were identified as missense variants. These mutations are particularly significant, as they can alter protein structure, stability, and function, including cystinosin-mediated cystine transport. Identifying these missense variants is a key step, as it enables us to target the mutations most likely to cause functional abnormalities, which are responsible for the pathogenicity observed in cystinosis. Among the 327 nsSNPs, 19 were predicted to be deleterious via eight predictive tools, including SIFT, PolyPhen, and the mutation assessor. These tools were selected for their proven accuracy in identifying disease-associated mutations. First, nsSNPs associated with the target gene were analyzed, identifying 19 candidate mutations for further evaluation: A137D, S141 F, F142S, R151G, D161H, S270 F, K280 T, K280R, P283L, Q284H, S293G, G296S, S298G, D305G, G308R, G308 V, G309 V, H353D and Y357H.

Stability analysis revealed crucial structural impacts of the mutations. Predictive tools such as I-Mutant 2.0 and DynaMut2 consistently showed that most mutations destabilize cystinosin. Notably, F142S and D305G had the greatest destabilizing effects, with significant ΔΔG values, indicating their strong potential to impair protein structural integrity and function. Studies have shown that there is a strong correlation between pathogenic mutations and their impact on stability (ΔΔG) or binding energy (ΔΔG binding), with a correlation coefficient of up to 0.7 [39]. Most pathogenic mutations are destabilizing, although some stabilizing mutations can also cause disease by disrupting dynamic or allosteric mechanisms [61]. ConSurf analysis assigned a high conservation score (9) to all the involved residues, highlighting their evolutionary importance. These scores suggest that these residues have essential structural or functional roles and that mutations affecting them are more likely to be pathogenic. Studies have confirmed

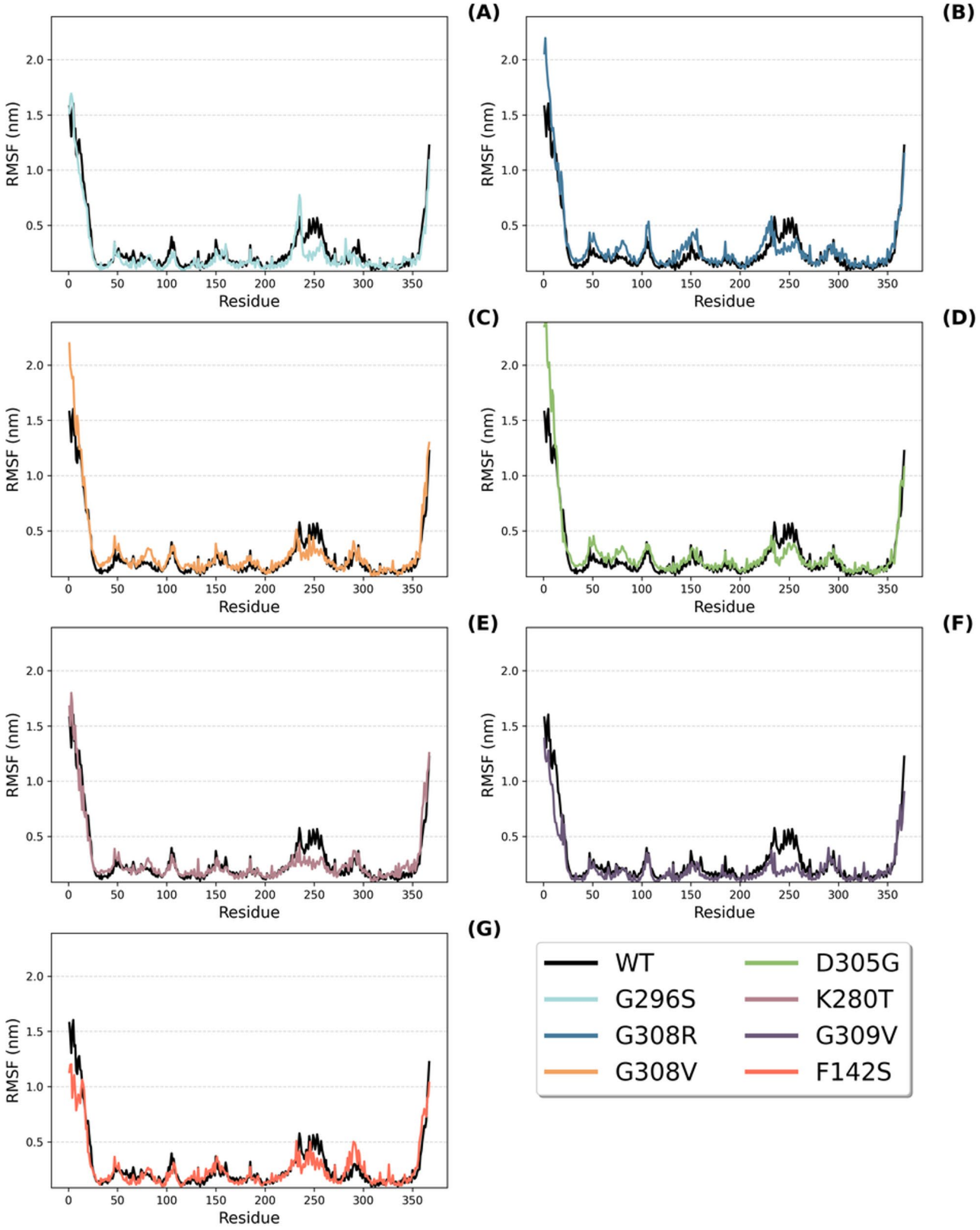

Fig. 5 The RMSF values for each residue of wild-type (WT) CTNS and its variants at 300 K are shown as a line plot. **A** WT (black) and G296S (light blue). **B** WT (black) and G308R (pastel blue). **C** WT (black) and G308 V (soft orange). **D** WT (black) and D305G (pale green). **E** WT (black) and K280 T (soft pink). **F** WT (black) and G309 V (pastel violet). **G** WT (black) and F142S (coral red)

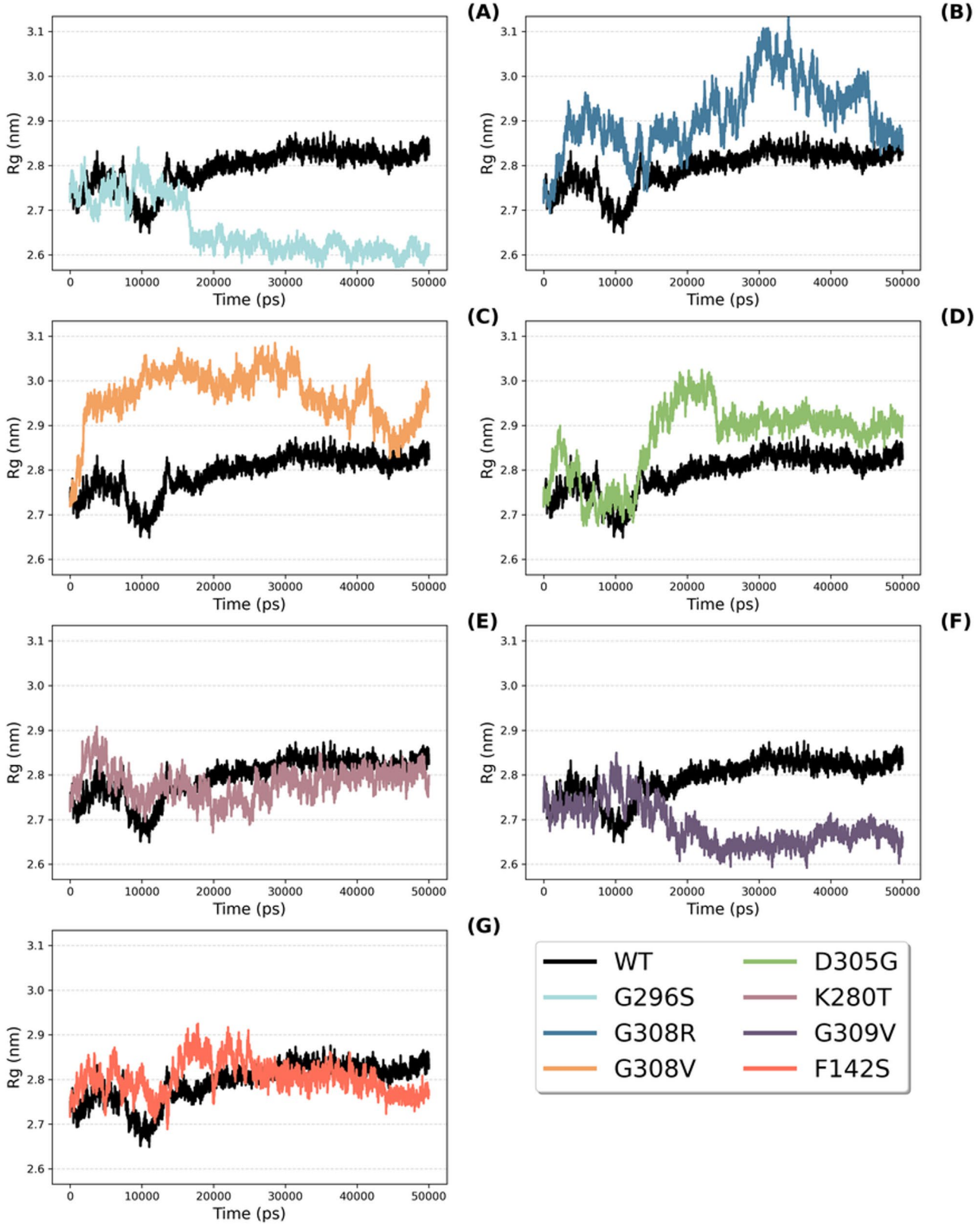

Fig. 6 (See legend on next page.)

(See figure on previous page.)

Fig. 6 The radius of gyration of the CTNS protein. The ordinate is Rg (nm), and the abscissa is time (ns). **A** Black and light blue lines show WT and G296S structures, respectively. **B** Black and pastel blue lines represent the WT and G308R structures, respectively. **C** Black and soft orange lines show WT and G308 V structures, respectively. **D** Black and pale green lines show the WT and D305G structures, respectively. **E** Black and soft pink lines show WT and K280 T structures, respectively **(F)**: black and pastel violet lines show WT and G309 V structures, respectively. **G** Black and coral red lines show the WT and F142S structures, respectively

that mutations affecting conserved residues, whether exposed or buried, have greater pathogenic potential than those targeting less conserved positions [62]. Mutations such as G308R, G308 V, G296S and K280 T affect exposed residues, potentially interfering with biological interactions or catalytic sites, whereas mutations D305G, G309 V and F142S affect buried residues, which are essential for maintaining structural stability. Previous studies have shown that the majority of disease-causing mutations occur primarily in buried positions (67%) [63]. Molecular dynamics (MD) simulations were conducted to evaluate the impact of nsSNPs on protein stability, flexibility, and compactness. Over a 50 ns trajectory, significant variations in RMSD, RMSF, Rg, and SASA were observed between the wild-type and mutant cystinosin. These changes illustrate structural perturbations induced by specific mutations, particularly D305G and F142S. The results provide valuable information on structural changes under physiological conditions. The D305G mutation significantly affects cystinosin stability and functionality, as shown by our molecular dynamics analyses and MutPred predictions. Our simulations revealed that this mutation increases local flexibility and modifies the SASA, resulting in a loss of structural rigidity. These perturbations particularly affect the helical transmembrane region (residues 299–308), which is an integral part of the PQ-loop 2 domain (residues 263–328), a key domain for protein stabilization and function. Residue D305 stabilizes the cystine-binding cavity through critical hydrogen bonding with W138, a key interaction for maintaining the transporter's active conformation. The MutPred results reinforce this analysis by assigning D305G a high pathogenic score (0.945), indicating major functional alterations. Among these predictions, MutPred suggests a loss of the transmembrane helical structure, as well as changes in the functional properties associated with D305. This residue is known for its role in cystine (L-cystine) and proton (H+) binding, as reported in previous studies. Together, these alterations severely disrupt the structural and functional interactions of cystinosin. In line with the literature, mutations such as D305G lead to a complete loss of cystine transport, highlighting the fundamental role of D305 in cystinosin function and confirming its highly pathogenic nature. Residue F142, located in the transmembrane region (residues 126–150) within PQ-loop 1 (residues 123–189), is crucial for the stability of cystinosin. Our simulations revealed that the F142S mutation destabilizes the protein by replacing the hydrophobic phenylalanine with a hydrophilic serine, increasing the SASA and exposing hydrophobic regions to the solvent. These disturbances compromise the stability of the binding cavity and could alter its affinity for cystine [64]. MutPred predictions support these observations by assigning the F142S mutation a pathogenic score of 0.811, suggesting significant functional alterations, including disruption of the protein's transmembrane structure. These results are consistent with the work of Bulut et al. [65] and Lu et al. [66], who also showed that F142 mutations affect transporter stability. The G296S mutation, located in the cytoplasmic region (residues 290–298) [67], replaces the flexible, nonpolar amino acid glycine (G) with the polar residue serine (S). Because of its flexibility, glycine plays a key role in local protein dynamics, whereas serine, with its hydrogen-bonding potential and larger volume, can induce rigidification of the affected region. This disrupts local interactions and reduces structural adaptability, which may alter the function of cystinosin. Similar effects have been observed in collagen mutations, where substitutions of glycine with larger residues, such as serine, disrupt stability and flexibility, contributing to diseases such as osteogenesis imperfecta. These observations highlight that the G296S mutation in cystinosin could affect cytoplasmic dynamics and compromise cystine transport by altering the interactions necessary for its active proton-coupled transport. Furthermore, mutations involving residues such as glycine (G) are often associated with significant functional alterations. Indeed, research into monogenic disorders shows that mutations that replace glycine with another amino acid are more likely to be pathogenic. Glycine plays a key role in flexible structural regions, and its substitution can disrupt protein dynamics and stability. Because of its flexibility, glycine enables the protein structure to adapt to local constraints, and its mutation could alter this adaptive capacity, affecting the function of the protein. G308 mutations (G308R and G308 V) further highlight the role of glycine residues in structural stability. G308 is located in the sixth transmembrane domain (TM6; residues 299–308), which is essential for transporter function. The substitution of glycine with arginine (G308R) or valine (G308 V) introduces steric and electrostatic changes, disrupting stability. G308R resulted in minor increases in the RMSD but significant increases in the Rg and SASA, indicating destabilization

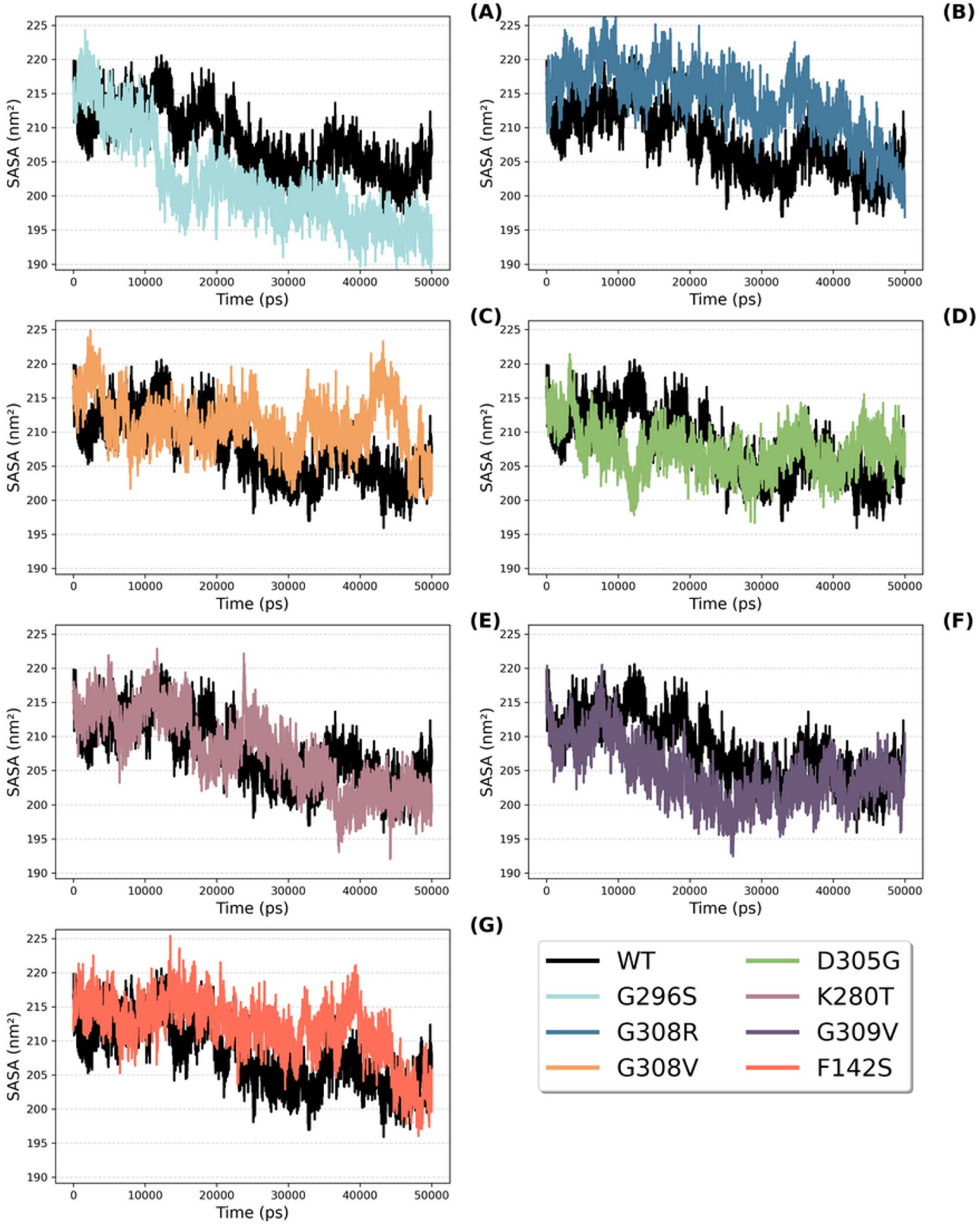

Fig. 7 (See legend on next page.)

(See figure on previous page.)

Fig. 7 The solvent-accessible surface area of the CTNS protein ordinate is the area (nm2), and the abscissa is time (ns). **A** Black and light blue lines show WT and G296S structures, respectively. **B** Black and pastel blue lines represent the WT and G308R structures, respectively. **C** Black and soft orange lines show WT and G308 V structures, respectively. **D** Black and pale green lines show the WT and D305G structures, respectively. **E** Black and soft pink lines show WT and K280 T structures, respectively (**F**): black and pastel violet lines show WT and G309 V structures, respectively. **G** Black and coral red lines show the WT and F142S structures, respectively

and greater exposure to solvents. The positive charge of arginine probably introduces electrostatic repulsion, weakening structural cohesion. Conversely, G308 V resulted in greater changes in the RMSD and Rg, with moderate increases in the SASA. Valine loading can stiffen the helix, limiting the conformational flexibility required for transport. Previous studies have demonstrated that G308R abolishes cystine transport while maintaining lysosomal localization [8]. These results highlight the role of glycine in transmembrane helices, highlighting its importance in maintaining transport function. To further understand the broader biological context of cystinosin, we explored protein–protein and gene–gene interactions. Using GeneMANIA, we constructed a composite gene–gene functional interaction network, revealing associations between *CTNS* and 20 other genes, including SLC66 A1 (solute carrier family 66 member 1), MPDU1 (mannose-P-dolichol utilization defect 1), and SLC66 A2 (solute carrier family 66 member 2) (Fig. 1). These interactions suggest that *CTNS* is part of a broader network involved in lysosomal function and cellular metabolism. Furthermore, protein interaction analysis via the The STRING platform highlighted a functional network involving cystinosin and several biologically relevant partners, such as SLC66 A1, SHPK, and TRPV1 (Fig. 2). These predicted interactions suggest that CTNS may participate in broader cellular processes related to membrane transport, metabolism, and lysosomal dynamics. However, a review of the literature indicates that, to date, there is no direct experimental evidence for physical interactions between cystinosin and these proteins.

For example, both cystinosin (CTNS) and SLC66 A1 (also known as LAAT1) are members of the SLC66 family of lysosomal amino acid transporters and are co-localized in the lysosomal membrane [68]. While their functions in lysosomal amino acid export and mTORC1 signaling are well established, a direct interaction between CTNS and SLC66 A1 has not been demonstrated experimentally. Similarly, although SHPK is often co-deleted with CTNS in patients with the common 57-kb deletion, recent studies show that SHPK deficiency does not impact cystinosin function or the efficacy of hematopoietic stem cell therapy in cystinosis, and no functional interaction has been reported [69]. As for TRPV1, its dysfunction in cystinosis patients is linked to the large genomic deletion encompassing the TRPV1 gene itself, rather than to a direct interaction with cystinosin [68].

Notably, the most substantiated experimental interactions for cystinosin involve components of the lysosomal mTORC1 signaling pathway, such as the Ragulator-Rag GTPase complex and the V-ATPase, which are essential for nutrient sensing and autophagy regulation [68]. These findings underscore the importance of further experimental studies to validate the predicted protein–protein interactions involving cystinosin and to clarify their potential implications in cystinosis pathology.

Conclusion

This study provides insights into the potential pathogenic impact of missense mutations in the CTNS gene, particularly in the context of cystinosis. Using a combination of predictive bioinformatics tools and molecular dynamics (MD) simulations, we identified 19 nsSNPs that may have significant effects on the structure and function of cystinosin. Mutations such as D305G and F142S appear to destabilize key structural regions, potentially disrupting the protein's ability to maintain its active conformation and compromising cystine transport. Residues such as G296 and G308, which play critical roles in stabilizing the transmembrane domains of cystinosin, may also contribute to protein dysfunction when disrupted.

The 50 ns MD simulations provided valuable insights into the structural and dynamic changes associated with the identified nsSNPs. However, extending the simulation time could help capture more subtle conformational changes and further elucidate protein dynamics under pathological conditions.

This study highlights the utility of in silico tools for predicting potentially deleterious mutations and prioritizing them for experimental validation. While computational predictions are a valuable first step, it is essential to emphasize that these results must be corroborated by experimental studies, such as site-directed mutagenesis, protein expression, and functional assays, to confirm their biological relevance. Further experimental investigations will be crucial for validating the pathogenicity of these nsSNPs and their roles in cystinosis.

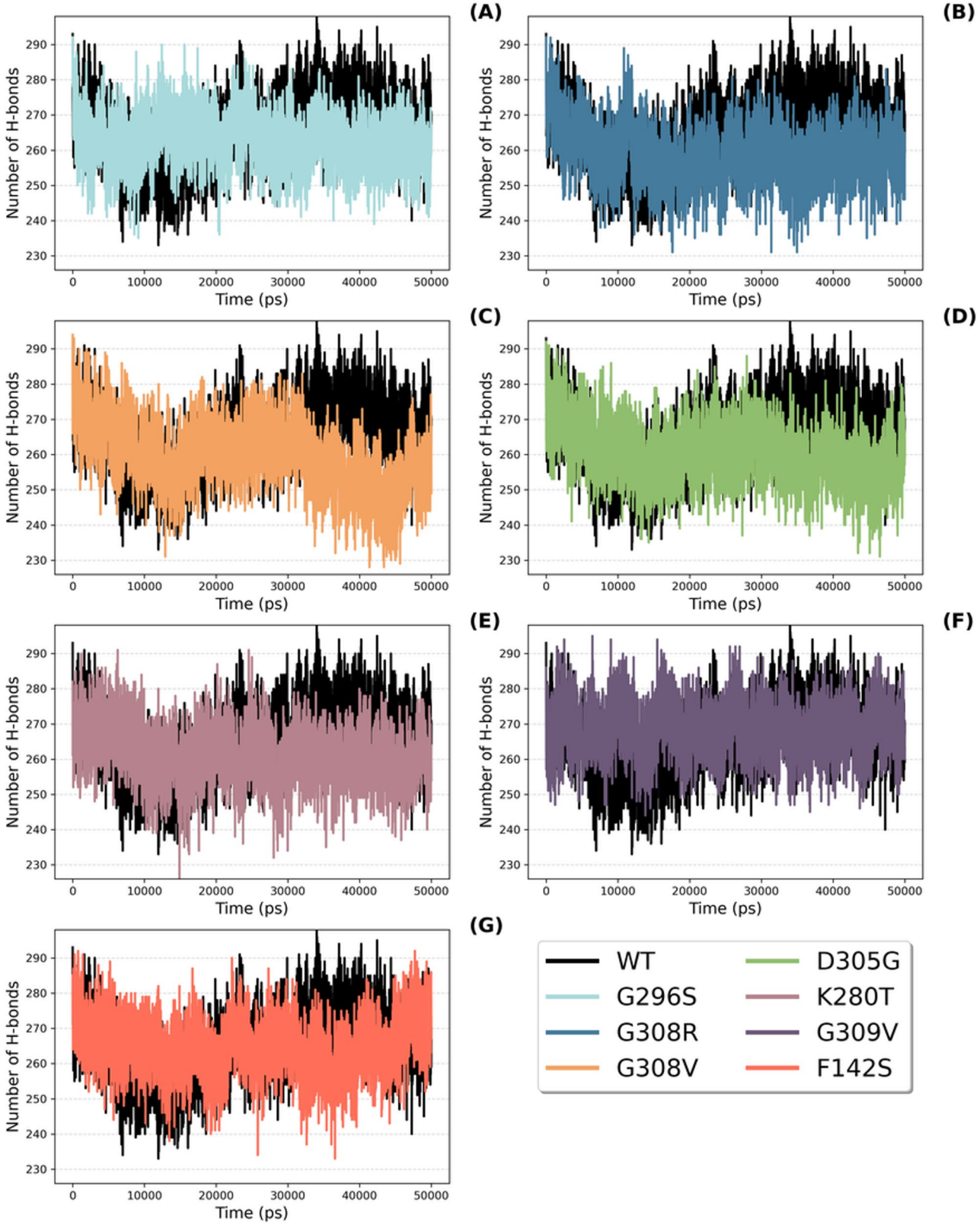

Fig. 8 Total number of hydrogen bonds in the CTNS protein for the native and mutant states. **A** WT (black) and G296S (light blue). **B** WT (black) and G308R (pastel blue). **C** WT (black) and G308 V (soft orange). **D** WT (black) and D305G (pale green). **E** WT (black) and K280 T (soft pink). **F** WT (black) and G309 V (pastel violet). **G** WT (black) and F142S (coral red)

Acknowledgements
The authors thank the developers and institutions that provided the computational tools used in this study.

Authors' contributions
L.A.N. conceived and designed the study, performed the main analyses, and wrote the first draft of the manuscript. A.C.D. contributed to the bioinformatics analyses, including the use of computational tools and data interpretation. I.D. participated in writing and revising the manuscript, improving the clarity and structure of the text. N.A.N. provided expertise in understanding the disease and reviewed the manuscript to ensure scientific accuracy. A.B. as the laboratory director, supervised the study, provided critical feedback, and validated the results. All authors read and approved the final manuscript.

Funding
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability
The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Laboratory of Molecular and Cellular Genetics (LGMC), University of Sciences and Technology of Oran Mohamed Boudiaf, Oran, Algeria
[2]Biology Department, Faculty of Natural and Life Sciences, University of Mostaganem, Mostaganem, Algeria
[3]Laboratory of Nutrition Physiology and Food Safety, University of Oran, Oran, Algeria
[4]Ophthalmology Clinic Dr. Adda Neggaz, Oran, Algeria

Received: 2 January 2025 / Accepted: 2 May 2025
Published online: 15 May 2025

References

1. Gahl WA, Bashan N, Tietze F, Bernardini I, Schulman JD. Cystine transport is defective in isolated leukocyte lysosomes from patients with cystinosis. Science. 1982;217(4566):1263–5. https://doi.org/10.1126/science.7112129.
2. Town M, Jean G, Cherqui S, et al. A novel gene encoding an integral membrane protein is mutated in nephropathic cystinosis. Nat Genet. 1998;18(4):319–24. https://doi.org/10.1038/ng0498-319.
3. Nesterova G, Gahl WA. Cystinosis: the evolution of a treatable disease. Pediatr Nephrol. 2013;28(1):51–9. https://doi.org/10.1007/s00467-012-2242-5.
4. Gahl WA, Thoene JG, Schneider JA, Cystinosis. N Engl J Med. 2002;347(2):111–21. https://doi.org/10.1056/NEJMra020552.
5. Cherqui S, Courtoy PJ. The renal Fanconi syndrome in cystinosis: pathogenic insights and therapeutic perspectives. Nat Rev Nephrol. 2017;13(2):115–31. https://doi.org/10.1038/nrneph.2016.182.
6. Levy M, Feingold J. Estimating prevalence in single-gene kidney diseases progressing to renal failure. Kidney Int. 2000;58(3):925–43. https://doi.org/10.1046/j.1523-1755.2000.00250.x.
7. Fanconi G. Die nicht diabetischen Glykosurien und hyperglykamiendes Altem kindes. Jahrbuch Kinderheilkunde. 1931;133:257–300.
8. Kalatzis V, Cherqui S, Antignac C, Gasnier B. Cystinosin, the protein defective in cystinosis, is a H(+)-driven lysosomal cystine transporter. EMBO J. 2001;20(21):5940–9. https://doi.org/10.1093/emboj/20.21.5940.
9. Cherqui S, Kalatzis V, Trugnan G, Antignac C. The targeting of cystinosin to the lysosomal membrane requires a tyrosine-based signal and a novel sorting motif. J Biol Chem. 2001;276(16):13314–21. https://doi.org/10.1074/jbc.M010562200.
10. Ruivo R, Bellenchi GC, Chen X, et al. Mechanism of proton/substrate coupling in the heptahelical lysosomal transporter cystinosin. Proc Natl Acad Sci U S A. 2012;109(5):E210–7. https://doi.org/10.1073/pnas.1115581109.
11. Saudek V, Cystinosin. MPDU1, sweets and KDELR belong to a well-defined protein family with putative function of cargo receptors involved in vesicle trafficking. PLoS ONE. 2012;7(2):e30876. https://doi.org/10.1371/journal.pone.0030876.
12. Jouandin P, Marelja Z, Shih YH, et al. Lysosomal cystine mobilization shapes the response of TORC1 and tissue growth to fasting. Science. 2022;375(6582):eabc4203. https://doi.org/10.1126/science.abc4203.
13. Festa BP, Chen Z, Berquez M, et al. Impaired autophagy bridges lysosomal storage disease and epithelial dysfunction in the kidney. Nat Commun. 2018;9(1):161. https://doi.org/10.1038/s41467-017-02536-7. Published 2018 Jan 11.
14. Jamalpoor A, Othman A, Levtchenko EN, Masereeuw R, Janssen MJ. Molecular mechanisms and treatment options of nephropathic cystinosis. Trends Mol Med. 2021;27(7):673–86. https://doi.org/10.1016/j.molmed.2021.04.004.
15. Rairikar M, Hohenfellner K, Elenberg E. Cystinosis– Pathophysiology Front Pediatr. 2021;9:643648.
16. Dries D, Berlingerio SP, Elmonem MA, Oliveira Arcolino F, Soliman N, van den Heuvel B, et al. Molecular basis of cystinosis: geographic distribution, functional consequences of mutations in the *CTNS* gene, and potential for repair. Front Pediatr. 2021;9:643648.
17. Bhat GR, Sethi I, Rah B, Kumar R, Afroze D. Innovative in Silico approaches for characterization of genes and proteins. Front Genet. 2022;13:865182. Published 2022 May 18. https://doi.org/10.3389/fgene.2022.865182
18. Uddin MM, Hossain MT, Hossain MA, et al. Unraveling the potential effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on the protein structure and function of the human *SLC30A8* gene on type 2 diabetes and colorectal cancer: an In Silico approach. Heliyon. 2024;10(17):e37280. Published 2024 Aug 31. https://doi.org/10.1016/j.heliyon.2024.e37280
19. Havranek B, Islam SM. Prediction and evaluation of deleterious and disease causing non-synonymous SNPs (nsSNPs) in human NF2 gene responsible for neurofibromatosis type 2 (NF2). J Biomol Struct Dyn. 2021;39(18):7044–55. https://doi.org/10.1080/07391102.2020.1805018.
20. Poon KS. In Silico analysis of BRCA1 and BRCA2 missense variants and the relevance in molecular genetic testing. Sci Rep. 2021;11(1):11114. https://doi.org/10.1038/s41598-021-88586-w. Published 2021 May 27.
21. Dutta AR, Roy AD. Computational analysis of deleterious single nucleotide polymorphisms (SNPs) in the human LCT gene. In: Proceedings of the 2024 IEEE International Conference on Computing, Applications and Systems; 2024; Cox's Bazar, Bangladesh. pp. 1–6.
22. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47(D1):D506–15. Available at: https://www.uniprot.org/uniprotkb/O60931/entry.
23. AlphaFold Protein Structure Database. https://alphafold.ebi.ac.uk.
24. SIFT. Sorting Intolerant From Tolerant. Singapore: Bioinformatics Institute, A*STAR. Available at: http://sift.bii.a-star.edu.sg.
25. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4(7):1073–81. https://doi.org/10.1038/nprot.2009.86.
26. PolyPhen. Polymorphism Phenotyping. Department of Biomedical Informatics, Harvard Medical School. Available at: http://genetics.bwh.harvard.edu/pph2.
27. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 2002;30(17):3894–900. https://doi.org/10.1093/nar/gkf493.
28. CADD. Combined Annotation Dependent Depletion. University of Washington. Available at: https://cadd.gs.washington.edu.
29. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. 2019;47(D1):D886–94. https://doi.org/10.1093/nar/gky1016.
30. MetaLR. Meta Logistic Regression. University of California, San Francisco. Available at: https://rbvi.ucsf.edu/netcharge/metaLR.
31. Dong C, Wei P, Jian X, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Hum Mol Genet. 2015;24(8):2125–37. https://doi.org/10.1093/hmg/ddu733.
32. Mutation Assessor. Mutation impact prediction. University of Leuven. Available at: http://mutationassessor.org.

33. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014;11(4):361–2. https://doi.org/10.1038/nmeth.2890.
34. PROVEAN. Protein Variation Effect Analyzer. JCVI. Available at: http://provean.jcvi.org.
35. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics. 2015;31(16):2745–7. https://doi.org/10.1093/bioinformatics/btv195.
36. SNPs&GO. Functional annotation of SNPs using Gene Ontology. BIOFOLD, University of Bologna. Available at: http://snps.biofold.org/snps-and-go.
37. Capriotti E, Calabrese R, Fariselli P, Martelli PL, Altman RB, Casadio R. WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation. BMC Genomics. 2013;14(Suppl 3):S6. https://doi.org/10.1186/1471-2164-14-S3-S6.
38. PhD-SNP. Predicting human deleterious SNPs. BIOFOLD, University of Bologna. Available at: http://snps.biofold.org/phd-snp.
39. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics. 2006;22(22):2729–34. https://doi.org/10.1093/bioinformatics/btl423.
40. PANTHER. Protein ANalysis THrough Evolutionary Relationships. University of Southern California. Available at: http://www.pantherdb.org.
41. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. Genet Med. 2015;17(5):405–24. https://doi.org/10.1038/gim.2015.30.
42. I-Mutant 2.0. Protein stability prediction upon mutation. BIOFOLD, University of Bologna. Available at: https://folding.biofold.org/i-mutant.
43. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res. 2005;33(Web Server issue):W306–10.
44. MUPro. Prediction of protein stability changes upon mutations. Irvine: University of California. Available at http://mupro.proteomics.ics.uci.edu.
45. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. Proteins. 2006;62(4):1125–32. https://doi.org/10.1002/prot.20810.
46. DynaMut2. Protein dynamics and stability upon mutation. University of Melbourne. Available at: http://biosig.unimelb.edu.au/dynamut.
47. Abid A, Nasir E, Awan HM. In Silico analysis of P.int mutation (c.687dupC; P.Met230Hisfs∗6) in PGAM2 gene that causes glycogen storage disease (GSD) type X. Kuwait J Sci. 2025;52(1):100344.
48. DUET. Integrated prediction of protein stability changes upon mutation. University of Melbourne. Available at: http://biosig.unimelb.edu.au/duet.
49. Pires DE, Ascher DB, Blundell TL. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. Nucleic Acids Res. 2014;42(Web Server issue):W314–9. https://doi.org/10.1093/nar/gku411.
50. GeneMANIA. University of Toronto. Available at: http://genemania.org.
51. Zuberi K, Franz M, Rodriguez H, Montojo J, Lopes CT, Bader GD, et al. GeneMANIA prediction server 2013 update. Nucleic Acids Res. 2013;41(W1):W115–22. https://doi.org/10.1093/nar/gkt533.
52. STRING. Search Tool for the Retrieval of Interacting Genes/Proteins. European Molecular Biology Laboratory (EMBL). Available at: https://string-db.org.
53. Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(D1):D638–46. https://doi.org/10.1093/nar/gkac1000.
54. ConSurf. Identification of conserved regions in proteins. Tel Aviv University. Available at: http://consurf.tau.ac.il.
55. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. 2010;38(Web Server issue):W529–33. https://doi.org/10.1093/nar/gkq399.
56. WebGRO. for macromolecular simulations. University of Arkansas for Medical Sciences. Available at: https://simlab.uams.edu.
57. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: high performance molecular simulations through multilevel parallelism from laptops to supercomputers. SoftwareX. 2015;1–2:19–25. https://doi.org/10.1016/j.softx.2015.06.001.
58. Pandey P, Panday SK, Rimal P, Ancona N, Alexov E. Predicting the effect of single mutations on protein stability and binding with respect to types of mutations. Int J Mol Sci. 2023;24(15):12073. https://doi.org/10.3390/ijms241512073. Published 2023 Jul 28.
59. Cadi AE, Krami C, Charoute AM. Prediction of the impact of deleterious non-synonymous single nucleotide polymorphisms on the human RRM2B gene: a molecular modeling study. Biomed Res Int. 2020;2020:7614634. https://doi.org/10.1155/2020/7614634.
60. Fan W, Ji HL, Kakar M, Ahmed S, Alobaid HM, Shakir Y. Computational analysis of the deleterious non-synonymous single nucleotide polymorphisms (nsSNPs) in TYR gene impacting human tyrosinase protein and the protein stability. PLoS ONE. 2024;19(11):e0308927. https://doi.org/10.1371/journal.pone.0308927. Published 2024 Nov 14.
61. Gerasimavicius L, Liu X, Marsh JA. Identification of pathogenic missense mutations using protein stability predictors. Sci Rep. 2020;10(1):153–87. https://doi.org/10.1038/s41598-020-72404-w. Published 2020 Sep 21.
62. Lee TC, Lee AS, Li KB. Incorporating the amino acid properties to predict the significance of missense mutations. Amino Acids. 2008;35(3):615–26. https://doi.org/10.1007/s00726-008-0087-9.
63. Kalatzis V, Nevo N, Cherqui S, Gasnier B, Antignac C. Molecular pathogenesis of cystinosis: effect of CTNS mutations on the transport activity and subcellular localization of cystinosin. Hum Mol Genet. 2004;13(13):1361–71.
64. Savojardo C, Manfredi M, Martelli P, Casadio R. Solvent accessibility of residues undergoing pathogenic variations in humans: from protein structures to protein sequences. Front Mol Biosci. 2021;7:626363. https://doi.org/10.3389/fmolb.2020.626363. Published 2021 Jan 7.
65. Bulut H, Moniot S, Licht A, Scheffel F, Gathmann S, Saenger W, Schneider E. Crystal structures of two solute receptors for L-cystine and L-cysteine, respectively, of the human pathogen Neisseria gonorrheae. J Mol Biol. 2012;415:560–72.
66. Lu M, Xu BY, Zhou K, et al. Structural and biochemical analyses of Microcystis aeruginosa O-acetylserine sulfhydrylases reveal a negative feedback regulation of cysteine biosynthesis. Biochim Biophys Acta. 2014;1844(2):308–15. https://doi.org/10.1016/j.bbapap.2013.11.008.
67. Guo X, Schmiege P, Assafa TE, et al. Structure and mechanism of human cystine exporter cystinosin. Cell. 2022;185(20):3739–e375218. https://doi.org/10.1016/j.cell.2022.08.020.
68. Luciani A, Devuyst O. The CTNS-MTORC1 axis couples lysosomal cystine to epithelial cell fate decisions and is a targetable pathway in cystinosis. Autophagy. 2024;20(1):202–4. https://doi.org/10.1080/15548627.2023.2250165.
69. Goodman S, Khan M, Sharma J, Li Z, Cano J, Castellanos C, Estrada MV, Gertsman I, Cherqui S. Deficiency of the sedoheptulose kinase (Shpk) does not alter the ability of hematopoietic stem cells to rescue cystinosis in the mouse model. Mol Genet Metab. 2021;134(4):309–16. Epub 2021 Nov 17. PMID: 34823997; PMCID: PMC8935660.

Publisher's Note