A comprehensive guide through the fold and stability of cytochrome c with implications for disease

Muhammad Abrar Yousaf[a], Anna Savoia[b], Annalisa Pastore[c*]

[a] *Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy*
[b] *Department of Engineering for Innovation Medicine, University of Verona, Verona, Italy*
[c] *Department of Clinical Neuroscience, King's College London, Denmark Hill Campus, London, United Kingdom*

[*]Corresponding author:
Annalisa Pastore
Department of Clinical Neuroscience, King's College London, Denmark Hill Campus, London, United Kingdom
Email: annalisa.1.pastore@kcl.ac.uk

Keywords: Cytochrome c; thrombocytopenia 4; mutations; peroxidase activity; molecular dynamics simulation

Abstract

Cytochrome c (Cyt-c), encoded by the *CYCS* gene, is crucial for electron transport, peroxidase activity, and apoptosis. Mutations in *CYCS* can cause thrombocytopenia 4 (THC4), a disorder with low platelet counts. We have, for instance, recently described six Italian families with five different heterozygous missense *CYCS* variants. These mutations likely enhance peroxidase and apoptotic activities, but the mechanisms causing reduced platelet production and increased apoptosis are unclear. This study investigates all clinically important Cyt-c variants using an integrated bioinformatics approach. Our findings reveal that all variants are at evolutionarily conserved sites, potentially disrupting Cyt-c function and contributing to disease phenotypes. Specific variants are predicted to affect phosphorylation, and ubiquitination. Cyt-c variants may influence the function of interacting genes and proteins involved in apoptosis and mitochondrial function. Molecular dynamics simulations reveal significant structural differences from the wild-type, with mutants showing reduced stability and increased flexibility, particularly in the omega loops. These changes weaken hydrogen bonds and increase distances between key residues, opening the heme active site. This open conformation may enhance accessibility to small molecules such as H_2O_2, thereby promoting peroxidase activity, which may enhance apoptosis and likely impact megakaryopoiesis and platelet homeostasis in THC4.

1 Introduction

Cytochrome C (Cyt-c) is an essential small heme-binding protein [1] that functions as an electron carrier in the electron transport chain by transferring electrons from complex III to complex IV of the cytochrome oxidase complex [2–4]. Cyt-c is a small globular protein of 105 residues in humans that contains a single covalently bound heme group attached to Cys15 and Cys18 of a CXXCH motif with two thioether bonds. His19 and Met81 are axial ligands of the six-coordinated heme iron. Heme ligation remains the same both in the ferric and ferrous forms of Cyt-c. The protein fold comprises five α-helices connected by three so-called for historic reasons Ω-loop regions, namely distal (residues 72-86), central (residues 41-58), and proximal (residues 13-29 and 28-40) Ω-loops [5–7]. These are non-regular protein structural motifs that consist of loops of six or more amino acid residues (Fetrow, 1995). Being highly conserved, Cyt-c is attached to the inner membrane of the mitochondrion to fulfill multiple important functions [8] which include a role in the intrinsic apoptotic pathway where Cyt-c initiates cell apoptosis as an antioxidant agent [9–11]. Because of its essential cellular role, it is not a surprise that Cyt-c is associated to pathology: heterozygous mutations of the *CYCS* gene on chromosome 7p15 are associated to autosomal dominant non-syndromic thrombocytopenia 4 (THC4, OMIM: 612004) [12], a disease originally described in a New Zealander family and, more recently, found also in a few other families all over the world [12–15]. Thrombocytopenia are a heterogeneous group of diseases caused by mutations in more than forty different genes [16]. Some of these diseases are rare such as it is the case for THC4. Despite the important roles of Cyt-c, it is interesting to notice that THC4 patients present only a mild phenotype that usually consists of a decreased number of platelets in circulating blood [12,13] caused by ectopic platelet release from bone marrow [12]. Platelets from THC4 patients are otherwise normal in size and morphology, and no increased bleeding tendency is usually described. While the low severity of this condition could suggest that a better understanding of the mechanisms causing it may be uninteresting, the need of being able to distinguish this condition from other more serious pathologies is essential for a correct diagnosis, which may avoid mistreatment of the patients. There have in fact been cases of THC4 that were incorrectly treated by steroids and/or splenectomy, without any significant benefit [17].

Various Cyt-c variants have been so far identified that cause THC4 [12–15,18–20]. Amongst them, G42S and Y49H have been extensively studied, and the crystal structures of some variants have been solved [21,22]. More recently, we have reported new variants and characterized the six families of the carriers for their symptoms [17]. By a clinical characterization of the 22 affected individuals, the largest series of THC4 patients ever reported, we showed that this disorder is characterized by mild-to-moderate thrombocytopenia, normal platelet size, but no additional clinical phenotypes associated with reduced platelet count. We also identified a significant correlation between the spatial region of Cyt-c affected by mutations and the extent of thrombocytopenia, which suggests different degrees of impairment of Cyt-c functions caused by different pathogenetic variants.

In the present study, we present a comprehensive bioinformatics analysis of all currently known clinical variants of Cyt-c, examining the structural features these variants induce in the Cyt-c fold through various computational tools and molecular dynamics simulations. We also explore how the structural roles of certain residues correlate with observed phenotypes affecting protein function. Hence, this study aims to propose new hypotheses regarding the impact of different Cyt-c variants on protein folding and stability, with the goal of elucidating the molecular mechanisms underlying enhanced peroxidase activity, increased apoptosis, and their effects on platelet production in THC4. These findings could offer new insights into the role of Cyt-c in the context of disease.

2 Methodology

2.1 Data collection and evolutionary conservation analysis

This study describes the functional effects of fourteen single amino acid variants (T20I, V21G, H27Y, N32H, L33V, G42S, Y49H, A52T, A52V, R92G, A97D, Y98H, L99V, and T103I) in human Cyt-c using bioinformatics approaches. These variants have been reported in patients with or without clinical characterization or identified through genetic sequencing (Table 1). The three-dimensional structure of reduced Cyt-c (PDB ID: 1j3s) was retrieved from the RCSB PDB database (https://www.rcsb.org/) [23]. Notably, the amino acid numbering in our study is offset by +1 from that in the PDB file, reflecting different sequence annotation conventions. Protein amino acid sequences of *Homo sapiens* and 91 other species were obtained from the UniProt database (https://www.uniprot.org/) [24]. A multiple sequence alignment was performed using the ClustalW tool (https://www.genome.jp/tools-bin/clustalw) [25] (Supplementary data 1) to facilitate the evolutionary conservation analysis of protein residues across different species. The evolutionary conservation of protein residues was visualized using WebLogo 3 [26].

Table 1. THC4-associated Cyt-c variants analyzed in the present study.

Protein	References (with molecular and/or clinical characterization)	References (without clinical characterization)
T20I		[27]
V21G		[27]
H27Y	[28]	
N32H		[27]
L33V		[27]
G42S	[12,17]	[20,27,29]
Y49H	[13,17]	
A52T	[17]	[27,30]
A52V	[14]	
R92G	[17]	[31]
A97D		ClinVar (ISTH-SSC Genomics)

Y98H		[32]
L99V	[17]	[20]
T103I	[29]	

2.2 Prediction of functional effect and pathogenicity of variants

The functional impacts of Cyt-c variants were assessed using SIFT (https://sift.bii.a-star.edu.sg/) [33]. SIFT predicts the effects of amino acid substitutions on protein function based on their position and nature, with scores ≤ 0.05 indicating deleterious changes. For the association of variants with the disease, the SuSPect tool (http://www.sbg.bio.ic.ac.uk/~suspect/) [34] was utilized. This tool uses sequence, structural, and systems biology features to predict the phenotypic effects of genetic variants. Trained on seventy-seven distinct features, SuSPect distinguishes between neutral and disease-causing variants, assigning scores ranging from 0 (neutral) to 100 (disease), with a cutoff score of 50 to differentiate between the two types of variants.

2.3 Prediction of molecular effects of variants

The MutPred2 platform (http://mutpred.mutdb.org/) [35] predicted molecular effects for Cyt-c variants, utilizing structural, functional and evolutionary data. This tool integrates PSI-BLAST, SIFT and Pfam for profiling, with TMHMM, MARCOIL and DisProt algorithms for predicting transmembrane helices, coiled-coil and disordered regions, respectively. This integration aids in the assessment of structural damage.

2.4 Prediction of post-translational modification

Post-translational modifications (PTMs) are crucial for predicting protein function. Methylation sites were identified using GPS-MSP (https://msp.biocuckoo.org/) [36], whereas phosphorylation sites on Serine (S), Threonine (T), and Tyrosine (Y) were predicted using NetPhos 3.1 (https://services.healthtech.dtu.dk/services/NetPhos-3.1/) [37], with scores above 0.5 indicating likely phosphorylation. PhosphoSitePlus (https://www.phosphosite.org/homeAction) [38] was used for the estimation of ubiquitination sites.

2.5 Analysis of gene and protein interactions

Cyt-c mutations may have effects on the associated genes and proteins. The *CYCS* gene interactions were analyzed using GeneMANIA (https://genemania.org/) [39] to predict and understand gene-gene relationships comprehensively. Further, the STRING database (version 12.0) (https://string-db.org/) [40] was utilized to predict protein-protein interactions for Cyt-c, assessing the impact of variants on specific interacting proteins. A high confidence threshold of 0.7 was applied in the analysis to minimize the likelihood of false positives and negatives, ensuring that only interactions with substantial supporting evidence were considered.

2.6 Prediction of protein stability

Predicting the stability of mutant proteins is crucial for understanding their activity, function, and regulation. The impact of variants on protein stability was assessed by calculating the difference

in Gibbs Free Energy ($\Delta\Delta G$, in kcal/mol) between the wild-type (ΔG_{WT}) and mutant (ΔG_{MT}) proteins, where:

$$\Delta\Delta G = \Delta G_{WT} - \Delta G_{MT}$$

Positive $\Delta\Delta G$ values indicate an increase in stability due to mutations, whereas negative values denote destabilization. For these predictions, we employed both three-dimensional structural and sequence-based methods. DynaMut2 (https://biosig.lab.uq.edu.au/dynamut2/) [41] analyzed the impact of variants on protein dynamics and stability through normal mode analysis of protein three-dimensional structure. MUpro (https://mupro.proteomics.ics.uci.edu/) [42], a sequence-based predictor, utilized a support vector machine model to assess the effects of mutations. These complementary approaches offer a comprehensive understanding of the stability of mutant proteins.

2.7 Molecular dynamics simulation

Molecular dynamics (MD) simulations were performed to investigate the impact of variants on protein stability, flexibility, and dynamics, utilizing the GROMACS 2019.2 platform [43] on the WebGRO for Macromolecular Simulations webserver (https://simlab.uams.edu/index.php). During the simulations in water, we considered only the protein without the heme group attached. Excluding the heme allowed us to focus on structural changes in the backbone and side chains. Although the heme is crucial for Cyt-c stability and function, its exclusion simplified the model to isolate mutation-specific effects on the core structure. Simulations employed the GROMOS96 43a1 force field for fully solvated systems. To ensure electroneutrality, sodium chloride was dissolved in SPC water molecules. The system pressure and temperature were maintained at 1 bar and 300 K, respectively, using the Berendsen weak coupling algorithm. The LINCS algorithm was employed to maintain all bonds at their equilibrium distances while allowing torsion and bending internal motions during the simulation [44]. The system underwent 5000 steps of energy minimization via the steepest descent algorithm to reach equilibrium, followed by a 50 ns (50000 ps) production run MD simulation. The number of frames per simulation was 1000. The trajectories were analyzed using MS Excel, and GraphPad Prism 8 for statistical analyses. To compare the significant differences in MD simulation parameters between wild-type (WT) and mutant proteins, the Kruskal-Wallis test followed by the Dunn test ($p < 0.05$) was applied due to the non-parametric nature (Shapiro-Wilks, $p < 0.05$) of the data. The results were presented as mean, median, interquartile range, or 95% confidence interval. The residue-residue percentage H-bond occupancy over 50 ns simulation was performed using VMD software [45]. The cut-off values for the distance and angle of the hydrogen bonds (donor and acceptor) were set to 3.0 Å and 20°, respectively.

2.8 Normal mode analysis

We used the protein structures obtained after the MD simulation to perform normal mode analysis using the iMODs server (http://imods.chaconlab.org/) [46], with the default parameters. This

analysis assessed the collective flexibility and mobility of the protein. The results were presented as B-factor/mobility maps, eigenvalue plots, covariance matrices, and an elastic network model.

3 Results

3.1 Cyt-c and evolutionarily conserved residues

The variants analyzed in this study are present in the proximal and central loops and the C-terminus (Fig. 1A). These mutated residues were assessed for evolutionary conservation through multiple sequence alignment and displayed as sequence logos. Amino acid positions critical for protein structure and/or function tend to be more evolutionarily conserved. Consequently, mutations at these conserved sites are usually more deleterious [47–49]. The sequence logo shows the frequency of amino acids at each position in a sequence alignment. The height of each letter in the logo is proportional to its frequency, and the total height of letters at a position represents the sequence conservation in bits. Larger letters correspond to more conserved residues, while smaller letters indicate more variability. The analysis revealed that residues T20, V21, H27, N32, L33, G42, Y49, A52, R92, A97, Y98, L99, and T103 are conserved across species from diverse taxonomic groups (Fig. 1B, Supplementary data 1). This conservation suggests that the described variants T20I, V21G, H27Y, N32H, L33V, G42S, Y49H, A52T/V, R92G, A97D, Y98H, L99V, and T103I are likely deleterious.

3.2 Prediction of the functional effect and pathogenicity of variants

Analysis of the variants and their predicted impacts on protein function and association with diseases was computed utilizing three computational prediction tools, SIFT, and SuSPect (Fig. 1C). SIFT scores (<0.05) indicate severe effects of all variants on protein function. SuSPect scores further support these findings, categorizing each variant with a 'Disease' label, where high scores (>50) suggest a potential association with the disease. This agreement across different predictive platforms underscores the robustness of the predictions.

3.3 Mapping variants on protein structure

The three-dimensional structure of Cyt-c (Fig. 2A) shows the presence of Ω-loops with the heme, thereby providing insights into the protein function in electron transport and apoptosis within cells. All Cyt-c variants were mapped onto the protein three-dimensional structure (Fig. 2B, C, D). The mutant structures of Cyt-c were generated using the SWISS-MODEL webserver (https://swissmodel.expasy.org/) [50] using the WT structure (PDB ID: 1j3s) as a template. Ramachandran plots were generated using PROCHECK tool [51] to evaluate the conformational integrity of both WT and mutant structures. These plots assess the phi and psi dihedral angles of the amino acid residues, key indicators of structural feasibility. In all mutant structures, the Ramachandran plot did not reveal any significant differences in the steric hindrance at the mutated residues compared to the WT (Supplementary data 2, Fig. S1), except for G42S (Fig. 2E). How the clinically important mutations alter the affected amino acids is shown in Supplementary Data 2 (Fig. S2).

Buried residues are crucial for the structural stability of proteins because they contribute to the hydrophobic core and are like the pillars of the three-dimensional structure. Mutations that destabilize proteins can induce unfolding or misfolding, potentially affecting the functionality of Cyt-c. Therefore, the solvent accessibility of mutation sites was calculated using the DSSP tool [52]. In this program, when the SASA score lies below 50 $Å^2$, the whole residue is considered buried. All examined amino acid residues examined in this study are buried (< 50 $Å^2$) within the protein structure. However, the degree of burial varies. Residues like V21 (5 $Å^2$) and L99 (3 $Å^2$) are deeply buried with minimal solvent exposure. L33 (9 $Å^2$), Y49 (12 $Å^2$), A52 (15 $Å^2$), and R92 (14 $Å^2$) are also highly buried, though slightly more exposed. G42 (22 $Å^2$), A97 (28 $Å^2$), T20 (29 $Å^2$), N32 (27 $Å^2$), and T103 (18 $Å^2$) have moderate burial, while H27 (43 $Å^2$) is the most exposed but still considered buried.

The T20I variant in the Cyt-c coil region (Fig. 2B), substituting a buried hydrophilic threonine with a larger hydrophobic isoleucine, likely compromises protein stability and disrupts the T20-N32 H-bonds [53]. This is due to the absence of the hydroxyl group in isoleucine, impeding the formation of the H-bond. The V21G variant in a coil region (Fig. 2B) introduces a cavity into the hydrophobic core by replacing a medium-sized valine with a smaller, more flexible glycine, potentially affecting protein stability (Fig. 2B). The critical role of H27 in a coil region (Fig. 2B) for maintaining protein stability is highlighted by the formation of a key H-bond with P45, which ensures the close association of the 20s and 40s Ω-loops [54]. Substituting histidine with the bulkier, uncharged, and hydrophobic tyrosine in the H27Y variant may induce structural alterations that significantly reduce protein stability and disrupt crucial interactions. The replacement of N32 (Fig. 2B) with a bulkier histidine in the N32H variant could impair the T20-N32 H-bonding [53] and is likely to affect the protein structural integrity, leading to instability. Additionally, the L33V variant at the turn (Fig. 2B) might create a cavity in the hydrophobic core due to valine smaller size.

G42, located in a protein loop, exhibits positive phi and psi values in the Ramachandran plot (Fig. 2E). Glycines, with their single hydrogen atom side chain, are symmetric and have a wider range of acceptable phi/psi backbone dihedral angles than other amino acids, allowing glycines to occupy also typically disallowed region of the Ramachandran plot. Replacing glycine with the bulkier serine could thus lead to steric clashes with adjacent residues due to glycine unique conformational flexibility, significantly altering the loop conformation [55] (Fig. 2E). This could be the reason that G42S mutation displaces residues 42 and 43 about 0.4 Å away from Tyr49, which is part of the core structure of the third omega loop in the experimental structure [21]. Further, the G42S variant (in a coil region, Fig. 2C), within the hydrophobic core, introduces a polar and hydrophilic serine, potentially disrupting hydrophobic interactions, destabilizing the core, and affecting the structural integrity and protein function. Y49, located in a buried hydrophobic loop, is a large residue with a bulky phenyl ring and a hydroxyl group (Fig. 2C). Replacing Y49 with the medium-sized, imidazole-ring-containing H49 could destabilize the hydrophobic core by creating a cavity. The hydrophilic and basic nature of H49 might disrupt hydrophobic interactions and introduce pH-

dependent changes, potentially impacting the protein structure and function. NMR data indicates that the Y49H variant results in a loss of resonances, suggesting increased conformational heterogeneity, potentially due to partial unfolding of the protein [22,56]. The A52T variant, replacing the nonpolar alanine located in a helix within the hydrophobic core (Fig. 2C) with the polar and hydrophilic threonine, could disturb hydrophobic interactions, affecting structure. Similarly, substitution to a larger valine in the A52V variant might affect the structure. In an experimental structure, the A52V variant of causes decreased global stability and decreased stability of the native state relative to sub-global unfolding of Ω-loop [57]. R92, a large, polar and hydrophilic residue located in the core helix near the cytochrome c active site (Fig. 2C), is crucial for interactions with the heme and other buried hydrophilic residues. Mutating R92 to the much smaller glycine could disrupt these interactions and create a cavity, compromising the protein function and structure.

The A97D variant in the helix (Fig. 2D) may disrupt the protein hydrophobic interactions due to the introduction of the hydrophilic aspartate at position 97, replacing hydrophobic alanine. The Y98H (Fig. 2D) variant could compromise the protein structural integrity by creating a cavity and disrupting hydrophobic interactions within the core helix, as it involves replacing the large hydrophobic tyrosine with a medium-sized hydrophilic histidine. Similarly, the L99V variant, located in a helix within the hydrophobic core (Fig. 2D), is likely to form a cavity due to the substitution of leucine with the smaller valine. The T103I variant in a coil region (Fig. 2D) replaces the small, polar, hydrophilic threonine with a larger, nonpolar, hydrophobic isoleucine, potentially affecting local packing and hydrophilic interactions.

Collectively, these variants could significantly influence the protein by altering its structure and functionality.

3.4 Molecular effects of the mutations

We assessed the molecular effects of the mutations of the Cyt-c variants, which could lead to pathogenicity, by the MutPred2 webserver (Supplementary data 2, Table S1). Mutations in Cyt-c can change its structure and function, leading to various molecular effects that may alter its activity. The T20I variant was predicted to affect several aspects of Cyt-c. It causes the loss of intrinsic disorder, reducing protein flexibility. It also results in altered metal binding and changes in the protein interaction surface. There is a gain of a functional site at C18, which might influence its activity, but Cyt-c is not primarily known for classical catalytic functions. The V21G leads to increased flexibility (gain of intrinsic disorder), which may destabilize the protein. It also changes metal binding and stability, both of which are essential for Cyt-c role in electron transfer. This variant results in the loss of a functional site at C18, likely affecting how the protein interacts with other molecules. The H27Y causes changes in the interaction surface and reduces the protein flexibility. It also introduces PTMs like acetylation and methylation, which may affect protein stability, although these modifications are not typical for Cyt-c regulation. The variant also impacts metal binding, which is crucial for its electron transfer function. The N32H alters the interaction surface and metal binding, both important for maintaining Cyt-c structure and function. It also

results in the loss of a functional region at K28, potentially affecting the protein interactions in the electron transport chain. The L33V leads to the gain of an allosteric site, which may introduce new regulatory features. However, it also alters the interaction surface, possibly disrupting its function in the electron transport chain.

The G42S variant causes changes in the interaction surface and solvent accessibility, which may impact how Cyt-c interacts with other molecules in its environment. The variant also affects metal binding and the presence of methylation sites, which could influence its structural dynamics. The Y49H results in changes to the interaction surface and a loss of phosphorylation at Y49. It also alters metal binding and solvent accessibility, which may impact its function in electron transport. The A52T reduces the solvent accessibility of Cyt-c, potentially impacting key functional sites. It also affects metal binding and interactions with membranes. Additionally, the variant introduces acetylation at K56, though this is not a typical regulatory feature for Cyt-c. The R92G alters the protein-coiled structure and results in the loss of an allosteric site, potentially affecting the protein shape and function. This mutation also introduces acetylation and methylation at K87, which could influence protein stability, although these modifications are not commonly linked to Cyt-c function. The A97D, Y98H, and L99V variants all affect the protein-coiled structure and disordered regions, which are important for stability. These variants also introduce changes in PTMs and solvent accessibility, which may affect Cyt-c role in electron transport and apoptosis.

3.5 Predicting the effects of PTMs

PTMs refer to the chemical modifications that proteins undergo after translation, including methylation, phosphorylation, ubiquitylation, and several others. PTMs play a pivotal role in the regulation of protein structures and functions, proving essential across various pathways, including cell signaling and protein-protein interactions, as well as numerous biological processes [58,59]. Mutations at or near the site of these modifications could impact their function, stability, localization, and interaction with other molecules. Our analysis utilized a combination of bioinformatics tools to evaluate the potential effects of Cyt-c mutations on its PTMs, which are crucial for its functionality and regulatory mechanisms. The GPS-MSP tool suggests that native Cyt-c lacks methylation sites, indicating a minimal role of methylation in its regulation. NetPhos 3.1 predicted phosphorylation sites at threonine 20 (T20), tyrosine 47 (Y47), threonine 64 (T64), and threonine 79 (T79). Cyt-c phosphorylation and acetylation regulate mitochondrial respiration, membrane potential, energy production, reactive oxygen species (ROS) levels, and apoptosis [60]. Variants such as T20I, V21G, and Y49H, located at or near these sites, could potentially modify the phosphorylation pattern, thereby affecting the protein activity or its interactions with other cellular molecules. Particularly, the T20I variant would abolish the possibility of phosphorylation at this site, potentially influencing the protein function or stability. Ubiquitination, a modification marking proteins for degradation, was predicted at lysine residues K28, K40, K54, K73, K74, K80, K87, and K100. Variants near these sites (G42S, A52T, A52V, and T103I) may affect the ubiquitination process, thus altering the protein degradation rate or its role in cellular signaling.

Collectively, Cyt-c PTMs are important for respiration, the mitochondrial membrane potential, ROS, and apoptosis [61] which could be affected by the protein variants.

3.6 Effects of variants on interactions

Gene network analysis using the GeneMANIA tool revealed that the *CYCS* gene, encoding Cyt-c, engages with twenty distinct genes (Fig. 3A). Mutations in the *CYCS* gene product could potentially disrupt these interactions by altering the structure or expression of Cyt-c, thereby weakening or preventing key protein-protein associations. Specifically, *CYCS* was found to co-express with CYC1, VDAC1, VDAC3, and EPS15L1, maintaining similar expression levels across various conditions. Disruption of these co-expression patterns could affect cellular functions such as oxidative phosphorylation and apoptosis. *CYCS* shares a domain only with CYC1, suggesting that variants impacting this shared domain may alter both structural integrity and function. Additionally, *CYCS* demonstrated co-localization with BCL2L1, VDAC1, VDAC2, and CASP9, indicating that these proteins are expressed in the same cellular locations or tissues. Variants in *CYCS* might impair its co-localization with these proteins, thereby affecting its role in apoptosis regulation and mitochondrial function. All genes within this pathway are linked to apoptosis, and any alteration in *CYCS* interaction with these proteins could disrupt the apoptotic balance within the cell. *CYCS* also displayed genetic interactions with BAX, BAK1, and DAP3, which are key regulators of apoptosis. *CYCS* variants could affect its interaction with these genes, potentially modulating apoptotic signaling and altering the cell ability to undergo programmed cell death. For instance, altered *CYCS* function could diminish its ability to trigger apoptosome formation with APAF1 and CASP9, thereby impairing the activation of downstream caspases such as CASP3, which are essential for apoptosis.

On the protein level, protein-protein interaction analysis using the STRING database revealed that Cyt-c interacts with ten other proteins, suggesting that variants in Cyt-c may disrupt these functional relationships (Fig. 3B). Notably, Cyt-c was found to interact with QCRFS1, CYC1, CASP9, APAF1, MT-CYB, BCL2L1, CASP3, BCL2, CYB5B, and CYB5A, all with high confidence scores (Supplementary data 2, Table S2). These proteins are crucial for the electron transport chain and apoptotic pathways. Variants that alter Cyt-c structure or binding affinity could impair its interaction with these proteins, potentially disrupting essential cellular processes such as energy production and programmed cell death. This analysis underscores the critical role of Cyt-c in maintaining cellular function through its interactions in both the electron transport chain and apoptotic pathways. Disruptions in these interactions due to variants could lead to imbalances in apoptosis or oxidative phosphorylation, with significant implications for cell survival and function.

3.7 Gibbs free energy-based protein stability analysis

We assessed the impacts of Cyt-c variants on protein stability by the DynaMut2 and MUpro software, using ΔΔG values which quantify the difference between the energy of the unfolded and folded states of a protein. A positive ΔΔG indicates that a mutant protein is thermodynamically more stable (requiring more energy to unfold) than the WT, whereas a negative ΔΔG suggests it is

less stable (requiring less energy to unfold). V21G, H27Y, N32H, L33V, G42S, Y49H, A52T, A52V, R92G, A97D, Y98H, L99V and T103I variants were identified as destabilizing (negative ΔΔG) by both prediction tools (Table 2). Notably, variants carrying the V21G, Y49H, R92G, Y98H and L99V mutations showed the most pronounced destabilizing effects (larger negative ΔΔG), underscoring their significant impact on protein function.

Table 2. Gibbs free energy (ΔΔG) values for Cyt-c variants computed through DynaMut2 and MUpro servers. Negative ΔΔG indicate destabilizing variants, while positive values suggest the variant as stabilizing.

Mutation	DynaMut2 (ΔΔG, kcal/mol)	MUpro (ΔΔG, kcal/mol)	Literature
T20I	0.51	–0.30	
V21G	–2.87	–2.12	
H27Y	1.12	–0.29	
N32H	–0.81	–1.25	
L33V	–2.2	–0.62	
G42S	–0.68	–0.53	Increases 41-58 Ω-loop dynamics and decrease global stability [62]
Y49H	–1.16	–1.22	Lower global stability [56]
A52T	–1.36	–0.86	
A52V	–0.76	–0.58	Destabilizes the global stability and native state [57]
R92G	–1.61	–1.10	
A97D	–0.69	–0.77	
Y98H	–1.76	–2.04	
L99V	–1.48	–1.48	
T103I	–0.17	–0.24	

3.8 Dynamics of the wildtype and mutant proteins in aqueous solution

Finally, we performed MD simulations to evaluate the conformational dynamics of Cyt-c. We observed significant alterations in the structured regions of the protein, notably affecting the α-helical and loop regions. The analysis was conducted on WT Cyt-c and on all the mutant structures over 50 nanoseconds (ns), as structural and confirmational modifications in the protein have been reported in the initial period during protein unfolding [63,64]. That is why previous Cyt-c studies conducted MD simulation analyses till or less than 30 ns [64–67].

Over the MD simulation period, WT and mutant variants exhibited distinct alterations in secondary structure, indicating structural rearrangements. Snapshots of protein structure were taken every 12.5 ns to visualize structural changes (Supplementary data 2, Fig. S3, S4, S5). The structure of WT transitioned from 41.35% alpha helix, 33.65% turn, 21.15% coil, and 3.85% extended strands (part of beta sheet) to 41.35% alpha helix, 25.96% turn, 18.27% coil, 6.73% strand, 5.77% 3-10 helix, and 1.92% isolated beta-bridges, resulting in consistent alpha helix, decrease in turns and coils, increase in strand, 3-10 helix and isolated beta-bridges (Fig. 4). During the simulation, all

the variants, except L33V and L99V, show a decrease in alpha-helical content and an increase in turn or coil conformation. Interestingly, variants Y49H, A97D, and L99V lead to the temporary development of new beta-sheets during the simulations. These beta-sheets could be part of the well-established folded units that form sequentially during the protein folding process [68]. Helix with stable H-bonds provide stability, and reduced helical conformation indicates the instability of protein structure. Conversely, increased turns or bends conformation can interfere with the proper folding of the protein, leading to misfolded proteins that are often less stable. Further, strand content varies across the mutants, while others like T20I, G42S, R92G, A97D, Y98H, L99V, and T103I lose strand structure. The 3-10 helix and bridge develop in several mutants, while pi helix is formed only in Y49H (Fig. 4).

3.9 Effect of the mutations on protein conformational stability

The root mean square deviation (RMSD) measures how much the structure deviates from a reference conformation, usually the starting structure, and is thus used as an indicator of stability. Higher RMSD values indicate reduced stability and lower values suggest greater stability [69]. WT Cyt-c exhibits the lowest RMSD, indicating stable conformation close to its initial structure (Fig. 5A). In contrast, T20I, V21G, H27Y, N32H and L33V mutant forms depict higher RMSD values, suggesting increased conformational fluctuations and reduced stability relative to the WT. Statistical analysis confirms that all variants significantly increase RMSD compared to WT, affecting Cyt-c stability (Fig. 5B). WT Cyt-c stability could be due to increased helix and coil structures and decreased extended strands. Mutant (T20I, V21G, H27Y, N32H) destabilization could be associated with conformational changes such as decreased alpha-helix. Furthermore, T20I displays a significant increase in turns; V21G increases 3-10-helices and turns; and N32H enhances 3_10-helices. L33V lower RMSD values relative to other mutants could be because the L33V mutant structure displays alpha-helical content relatively stable, but there is an increase in 3-10-helices and extended strands, indicating minor structural adjustments.

G42S, Y49H, A52T, A52V and R92G mutants (Fig. 5C, D) depict significantly higher RMSD values than WT, suggesting lower stability, with G42S and R92G being particularly destabilizing. This instability is linked to structural changes: G42S displays a decrease in alpha-helix and an increase in turn content; Y49H shows decreased alpha-helix and increased turns and pi-helix; A52T exhibits decreased alpha-helix and increased turns and 3_10-helices; and A52V shows an decrease in alpha-helix and a notable appearance of extended strands.

A97D, Y98H and T103I also show significantly higher RMSD than WT (Fig. 5E, F), indicating destabilization, especially T103I. A97D increases coils with reduced alpha-helical content; Y98H significantly increases in coils; and T103I reduces alpha-helix and increases turns while reducing alpha-helical content, suggesting significant structural alterations. L99V increases coil, indicating minor flexibility as displayed by non-significant differences with the WT ($p > 0.05$) (Fig. 5F).

3.10 Effect of the mutations on protein compactness and folding

The radius of gyration (Rg) measures protein compactness and structural stability, with lower Rg values indicating tighter packing and higher values suggesting more extended or unfolded structures. The WT protein exhibits a stable Rg, indicating consistent compactness and stability. In contrast, mutant proteins display varying Rg profiles, suggesting changes in compactness and stability due to variants (Fig. 6). Significantly higher Rg values for T20I, V21G, H27Y, N32H, L33V (Fig. 6A, B), G42S, A52V, R92G (Fig. 6C, D), A97D, Y98H, L99V, and T103I (Fig. 6E, F) mutants indicate less compact structures. A52T shows a small decrease in Rg, while Y49H has comparable compactness to the WT (Fig. 6C. D).

3.11 Effect of the mutations on protein flexibility

The root mean square fluctuation (RMSF) quantifies the flexibility of individual amino acid residues within a protein structure throughout a simulation. Higher RMSF values indicate increased residue movements, often linked to protein destabilization. Our results demonstrate that all mutated proteins show significantly higher flexibility than the WT protein (Fig. 7A, B, C). Analysis reveals the impact of variants across the Ω-loops: proximal (residues 13-40), central (residues 41-58), and distal (residues 72-86). Specifically, T20I and V21G increase flexibility in the proximal and central loops, respectively, while H27Y enhances flexibility across all Ω-loops, indicating a potential impact on protein dynamics. N32H and L33V significantly elevate RMSF in the proximal and central loops, with L33V also affecting the distal loop (Fig. 7D). The G42S variant displays heightened flexibility in all loops. Y49H shows increased RMSF in proximal and central regions, whereas A52T has a negligible effect on the proximal loop but reduces RMSF in central and distal loops. The A52V variant increases flexibility in proximal and central regions, with R92G also showing significant fluctuations across all loops and the N-terminus (Fig. 7E). The A97D, Y98H, and T103I variants significantly raise RMSF values in all loop regions, with T103I exhibiting the most pronounced fluctuations. Conversely, L99V increases flexibility solely in the proximal loop region (Fig. 7F).

3.12 H-bond network analysis suggests increased peroxidase activity

A hydrogen bond (H-bond) analysis compares the bonding patterns between the WT and mutant variants. The results reveal that the WT protein exhibits a higher number of H-bonds compared to most mutants, including T20I, V21G, N32H, G42S, A52T, R92G, A97D, Y98H, L99V, and T103I (Fig. 8A, B, C). This finding suggests that these variants potentially disrupt or alter the H-bonding network within the protein structure. Among these variants, the A52V mutant shows fewer differences in the number of H-bonds as compared to WT. Interestingly, mutants such as H27Y, L33V, and Y49H display a similar number of H-bonds as the WT protein, indicating that these variants may not significantly impact the bonding pattern (Fig. 8D, E, F).

In Cyt-c, Tyr68 forms a H-bond network with the axial ligand Met81, and conserved residues Asn53 and Thr79. Tyr68 is also involved in H-bonding interactions with the central and distal Ω-loops. This network involving Tyr68 is crucial for the conformational switch, orienting the protein

toward the apoptotic pathway [70]. Consequently, distances between these residues were calculated (Fig. 9A, Table 3).

The results indicate an overall loss of H-bonds in the T20I, V21G, N32H, G42S, A52T, A52V, R92G, A97D, Y98H, L99V, and T103I mutants of Cyt-c with increased distances between H-bonded residues (Fig. 9A, Table 3). This disruption of specific interactions may destabilize the active heme configuration by altering the positions of Met81, Asn53, and Thr79. Such structural changes could increase the accessibility of the heme catalytic site to small molecules like H_2O_2, thereby enhancing peroxidase activity and facilitating the initiation of the apoptosis pathway. In contrast, the H27Y, L33V, and Y49H Cyt-c mutants exhibit considerably stronger H-bonds involving these critical residues compared to WT, suggesting an increased catalytic activity of the heme.

Table 3. The mean, median and 95% confidence interval (CI) of distances between H-bonded residues between Tyr68-Asn53, Tyr68-Thr79 and Tyr68-Met81 of WT and mutant Cyt-c after 50 ns MD simulation. Note that amino acid numbering differs by +1 from that of the PDB file (1j3s).

Protein	Tyr68 and Asn53 (nm)			Tyr68 and Thr79 (nm)			Tyr68 and Met81 (nm)		
	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI
WT	1.1517	1.1424	1.1393-1.1407	1.0061	1.0093	1.0066-1.0121	0.9045	0.8927	0.8842-0.8990
T20I	1.2703	1.2731	1.3129-1.2766	1.0017	1.0008	0.9984-1.0038	1.0317	1.0334	1.0306-1.0370
V21G	1.3322	1.3772	1.3654-1.3874	1.2308	1.2872	1.2799-1.2936	1.3154	1.3920	1.3868-1.3973
H27Y	1.1440	1.1586	1.153-1.1639	0.9936	0.9956	0.9875-1.0022	0.8034	0.7548	0.7484-0.7650
N32H	1.1937	1.1675	1.1639-1.1725	1.3505	1.4101	1.4041-1.4166	1.0224	0.9874	0.9821-0.9923
L33V	1.1255	1.1357	1.1305-1.1396	0.9837	0.9684	0.9634-0.9735	0.9152	0.9088	0.9058-0.9125
G42S	1.2587	1.2624	1.2573-1.2676	1.0199	1.0181	1.0157-1.0201	0.9084	0.9064	0.9019-0.9109
Y49H	1.1470	1.1321	1.1259-1.1434	1.0041	1.0063	1.0038-1.0091	1.0766	1.1154	1.1069-1.1251
A52T	1.2210	1.2221	1.2153-1.2272	0.9616	0.9614	0.9583-0.966	0.9084	0.9064	0.9019-0.9109
A52V	1.2188	1.2233	1.2182-1.2481	0.9268	0.9590	0.9452-0.9703	0.9129	0.9226	0.9151-0.9274
R92G	1.2783	1.3038	1.2969-1.2838	1.4057	1.4469	1.4391-1.4587	1.2930	1.2870	1.2816-1.2935
A97D	1.2369	1.2186	1.2114-1.3395	1.0045	1.0164	1.0125-1.0197	0.8849	0.8975	0.8926-0.9006
Y98H	1.3552	1.4037	1.3957-1.3669	1.0324	1.0284	1.0256-1.032	0.9922	0.9955	0.9908-0.9994
L99V	1.2367	1.2379	1.2327-1.2299	1.0122	1.0123	1.0108-1.0145	1.0292	1.0357	1.0323-1.0404
T103I	1.1635	1.1668	1.1614-1.1556	1.0983	1.0989	1.0911-1.1034	0.9759	0.9662	0.9621-0.9710

We then analyzed the percentage of H-bond occupancy. H-bond occupancy (%) refers to the percentage of time a specific hydrogen bond is maintained during a simulation. It indicates the stability of a bond, with higher % occupancy reflecting a more consistent and stable bond, contributing significantly to the structural integrity of the protein. Lower % occupancy suggests that the bond is more transient or dynamic during the simulation. The boxplot in (Fig. 10A) illustrates the varying distribution of % H-bond occupancy across all residues. The bar chart (Fig. 10B) shows the number of H-bonds with an occupancy higher than 30%, where the WT consistently has more of these high-occupancy H-bonds compared to the mutant proteins (Fig.

10B), suggesting that WT has a stronger, more stable structure. The number of stable H-bonds varies among the mutants, with some showing a noticeable drop in H-bonds (all variants except N32H and L99V) that maintain an occupancy above 30%. This reduction in stable H-bonds in all variants suggests that variants lead to a weakening of the protein structure, as reflected in the lower H-bond occupancy, fewer stable H-bonds, and previous analyses.

3.13 Analysis of active site residues for electron transfer efficiency and apoptosis

An optimum distance between the critical electron-transfer residues around heme active conformation is important for the efficient flow of electrons [72]. Therefore, variations in the distance between electron-transfer residues, Tyr68-Cys15, Tyr68-Cys18, Tyr68-His19, Tyr68-Phe83, and Met81-Phe83 were analyzed for both WT, and mutant proteins (Fig. 9B, Table 4). In WT Cyt-c, the heme group forms covalent bonds with Cys15 and Cys18, with axial ligands His19 and Met81. The variants T20I, V21G, N32H, L33V, G42S, Y49H, A52T, A52V, Y98H, R92G, L99V, and T103I increase distances between electron-transfer residues relative to WT, while H27Y and A97D variants decrease these distances. The increased distances in certain mutants lead to an opening of the active site conformation, likely enhancing accessibility to small molecules like H_2O_2 and promoting peroxidase activity [72]. Specifically, N32H, L33V, Y49H, G42S, A52T, A52V, R92G, Y98H, L99V, and T103I variants induce significant structural changes, resulting in an open active conformation throughout the simulation. Conversely, minor distance changes in T20I, V21G, H27Y, and A97D mutants suggest that their structural integrity and heme active conformation closely resemble WT Cyt-c, enabling these mutants to retain electron transfer competency within the respiratory chain similar to WT.

Table 4. The mean, median, and 95% confidence interval (CI) of distance between key amino acid residues of the active site environment of WT and mutant Cyt-c over 50 ns MD simulation. Note that amino acid numbering differs by +1 from that of the PDB file (1j3s).

Protein	Tyr68 and Cys15 (nm)			Tyr68 and Cys18 (nm)			Tyr68 and His19 (nm)			Tyr68 and Phe83 (nm)			Met81 and Phe83 (nm)		
	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI
WT	1.0472	1.0360	1.0289-1.0216	1.7100	1.7124	1.7051-1.6451	1.6498	1.6489	1.6399-1.5854	0.8306	0.8306	0.8273-0.8344	0.6111	0.6081	0.6036-0.6339
T20I	1.8844	1.8726	1.8636-1.758	1.7244	1.7159	1.7060-1.6717	1.5720	1.5643	1.5557-1.4583	0.6821	0.6747	0.6706-0.6789	0.5890	0.5822	0.5758-0.5763
V21G	1.1460	1.1172	1.112-1.0719	1.3323	1.3205	1.315-1.3242	1.1688	1.1466	1.1420-1.1076	0.9845	0.9854	0.9832-0.9874	0.6698	0.6882	0.7012-0.6721
H27Y	1.1196	1.1145	1.1092-1.2016	1.1102	1.1066	1.0861-1.0089	1.1393	1.1365	1.1309-1.1204	0.8770	0.8676	0.8653-0.8704	0.6058	0.6037	0.6237-0.6007
N32H	1.2626	1.2578	1.2513-1.2898	1.3430	1.3897	1.3800-1.4352	1.1247	1.1517	1.1440-1.1489	0.9690	0.9467	0.9408-0.9532	0.6633	0.6705	0.6849-0.6824
L33V	1.4154	1.4067	1.4006-1.3800	1.4304	1.4026	1.3926-1.3636	1.2800	1.2618	1.2539-1.2121	1.0450	1.0322	1.022-1.0408	0.5987	0.5937	0.6393-0.6297
G42S	1.1052	1.0779	1.0716-1.0411	1.4894	1.4804	1.4748-1.5450	1.3259	1.3168	1.3119-1.3068	0.6653	0.6502	0.6461-0.6561	0.6455	0.6329	0.6272-0.6292
Y49H	1.2680	1.2512	1.2451-1.2931	1.4016	1.4004	1.3936-1.3932	1.2566	1.2562	1.2519-1.2747	1.0048	1.0039	0.9978-1.0689	0.5769	0.5642	0.5605-0.5681
A52T	1.6830	1.6992	1.7823-1.7781	1.2343	1.2395	1.2346-1.2299	1.3577	1.3671	1.3586-1.3792	0.8611	0.8606	0.8591-0.8690	0.6250	0.6246	0.6223-0.6277
A52V	1.3922	1.3920	1.5178-1.4364	1.5471	1.5458	1.5396-1.6207	1.4455	1.4610	1.4534-1.4690	0.8622	0.8569	0.8523-0.8522	0.9084	0.9064	0.9019-0.9109
R92G	1.2857	1.2789	1.2276-1.4116	1.4271	1.4374	1.4238-1.6497	1.3945	1.4154	1.4022-1.5571	0.8699	0.8654	0.8605-0.8600	0.6263	0.6245	0.6221-0.6268
A97D	1.0567	1.0363	0.9489-1.0365	0.9034	0.8576	0.8546-0.8683	1.0472	1.0334	1.0295-1.0745	0.9936	0.9872	0.9781-1.0529	0.6150	0.6165	0.6129-0.6206
Y98H	1.1987	1.1422	1.1313-1.1512	1.8396	1.7595	1.7431-1.5344	1.7805	1.7248	1.7178-1.6272	0.7533	0.7315	0.7277-0.6808	0.6438	0.6445	0.6421-0.6474
L99V	1.3752	1.3764	1.3727-1.3805	1.4160	1.4287	1.4229-1.4401	1.2510	1.2482	1.2439-1.2238	0.9696	0.9689	0.9644-1.0053	0.6142	0.6146	0.6113-0.6180
T103I	1.8069	1.8097	1.7975-1.8204	1.8378	1.7384	1.8366-1.6233	1.5584	1.4727	1.6523-1.4372	0.9081	0.8872	0.8809-0.8208	0.6355	0.6344	0.6258-0.6367

Takano et al. [73] reported that the electron transfer pathway from mitochondrial Cyt-c reductase to oxidized Cyt-c (Fe^{3+}) differs from the reduced form (Fe^{2+}) to cytochrome c oxidase. X-ray crystallography of WT Cyt-c revealed parallel orientations of Tyr75 with Trp60 and Tyr68 with Heme105. In the Cyt-c ferric state, electron transfer is proposed to begin at Tyr75, then to Trp60, with a subsequent transfer from Tyr68 to the heme, potentially via Met81 sulfur, to preserve electron neutrality. This triggers conformational adjustments, aligning Trp60 with Tyr68 to enable electron transfer between these electron-rich and electron-deficient sites. Therefore, the distances between Trp60-Tyr75 and Tyr68-Met81 are crucial for the electron transfer process (Fig. 9C). Post-simulation analysis shows that mutants T20I, V21G, H27Y, N32H, L33V, G42S, Y49H, A52T, R92G, A97D, Y98H, L99V, and T103I increase the Trp60-Tyr75 distance, while T20I, V21G, N32H, L33V, G42S, Y49H, A52T, A52V, R92G, Y98H, L99V, and T103I also show an increase in the Tyr68-Met81 distance compared to WT (Table 5). These changes suggest a potential impact on electron transfer efficiency and apoptotic activity.

Table 5. The mean, median and confidence interval (CI) of distance between Trp60 to Tyr75 and Tyr68 to Met81 residues of WT and mutant Cyt-c. Note that amino acid numbering differs by +1 from that of the PDB file (1j3s).

Protein	Trp60 and Tyr75 (nm)			Tyr68 and Met81 (nm)		
	Mean	Median	95% CI	Mean	Median	95% CI
WT	0.9552	0.9623	0.9587-0.9659	0.9045	0.8927	0.8842-0.8990
T20I	1.0681	1.0599	1.0559-1.0645	1.0317	1.0334	1.0306-1.0370
V21G	1.1015	1.0781	1.0737-1.0831	1.3154	1.3920	1.3868-1.3973
H27Y	1.0661	1.0555	1.0530-1.0590	0.8034	0.7548	0.7484-0.7650
N32H	1.0411	1.0441	1.0414-1.0476	1.0224	0.9874	0.9821-0.9923
L33V	1.0786	1.1091	1.0952-1.1206	0.9152	0.9088	0.9058-0.9125
G42S	1.1547	1.1616	1.1517-1.1702	0.9084	0.9064	0.9019-0.9109
Y49H	1.0669	1.0453	1.0407-1.0504	1.0766	1.1154	1.1069-1.1251
A52T	1.0425	1.0324	1.0283-1.0371	0.9084	0.9064	0.9019-0.9109
A52V	0.9084	0.9064	0.9019-0.9109	0.9129	0.9226	0.9151-0.9274
R92G	1.2153	1.2123	1.2077-1.2164	1.2930	1.2870	1.2816-1.2935
A97D	1.0535	1.0535	1.0450-1.0606	0.8849	0.8975	0.8926-0.9006
Y98H	1.0990	1.1075	1.0951-1.1164	0.9922	0.9955	0.9908-0.9994
L99V	1.0906	1.0589	1.0543-1.0637	1.0292	1.0357	1.0323-1.0404
T103I	1.1361	1.1355	1.1291-1.1442	0.9759	0.9662	0.9621-0.9710

3.14 Normal mode analysis

The collective functional flexibility and mobility of the WT and mutant proteins were evaluated using the iMODS server (Supplementary Data 2, Fig. S5). The deformability analysis of WT and mutant proteins reveals notable differences in flexibility across various regions of the protein structure. In each plot, the peaks indicate areas of high deformability (Supplementary data 2, Fig.

S6). The lower eigenvalues of all mutants, except for V21G and Y98H, compared to the WT suggest that the mutants exhibit higher flexibility. These lower eigenvalues indicate that less energy is required to deform the structure, making it easier for the protein to undergo conformational changes. The variants in Cyt-c protein have varying impacts on the dynamic behavior of the structure, as observed in the covariance maps (Supplementary data 2, Fig. S7). In maps, covariance matrices illustrate the correlation of motion between pairs of residues in the protein structure. In the maps, red regions represent residues that move in a correlated manner; white regions indicate uncorrelated motions; and blue regions denote anti-correlated motions. In the WT, most residues show strong correlated motions (in red), indicating coordinated movement. Variants like T20I, L33V, G42S, A52V, R92G, A97D, and T103I increase anti-correlated motions (in blue), meaning residues move more independently or in opposite directions, which suggests these variants make the structure more flexible. In contrast, A52T shows more correlated motions (red), meaning it stiffens the structure, making it more rigid. Variants H27Y, H32H, and Y49H show only minor changes compared to WT, with slightly more anti-correlated regions, indicating these mutations have a limited effect on flexibility. Y98H shows increased anti-correlated motions, suggesting it makes the structure more flexible, even though earlier analyses suggested it would increase stiffness. Overall, some mutations increase flexibility, while others make the structure stiffer, affecting the dynamic behavior of Cyt-c in different ways. The elastic network model provides insight into the stiffness or flexibility of the protein structure, with darker dots representing stronger, stiffer connections (more rigid structure) and lighter dots indicating weaker, more flexible connections (Supplementary data 2, Fig. S8). The WT has a balanced distribution of stiff (dark) and flexible (light) connections, maintaining a stable structure. Variants like T20I, V21G, L33V, G42S, R92G, A97D, Y98H, L99V, and T103I exhibit more flexible connections, indicated by lighter dots, suggesting increased flexibility in the structure. In contrast, A52T shows darker dots, reflecting increased stiffness and rigidity. Variants H27Y, H32H, and Y49H display patterns similar to WT, indicating only minor changes in flexibility. A52V shows moderate stiffness, falling between the highly flexible and rigid variants. These variants affect the dynamic properties of Cyt-c, either enhancing flexibility or increasing stiffness.

4 Discussion

The interplay between increased peroxidase activity and apoptosis in Cyt-c mutations presents a complex mechanism contributing to the pathogenesis of THC4. Cyt-c, besides the mitochondrial respiratory chain, is crucial in the intrinsic apoptosis pathway [7,74]. For apoptosis, Cyt-c moves across the outer mitochondrial membrane into the cytosol to activate the caspase cascade and initiate the apoptotic activity. This process could be initiated through several stimuli, including oxidative stress. It is believed that permeabilization of the outer mitochondrial membrane for Cyt-c release to the cytosol is induced by the peroxidase activity of Cyt-c in complex with cardiolipin, a phospholipid found in the mitochondrial inner membrane [75,76].

Post-translational modifications and mutations in Cyt-c can alter its structure and apoptotic function [77,78]. In this study, we propose that Cyt-c variants cause structural destabilization and increased flexibility, specifically through the loss of the intra-protein H-bond network and

increased distance between important residues in the heme catalytic site. These changes increase H_2O_2 access, causing oxidative stress and enhancing peroxidase activity, which stimulates apoptosis. This increase in apoptotic activity, particularly within megakaryocytes, the progenitor cells for platelets, could lead to a reduction in platelet count, manifesting as THC4 [22,28]. The precise molecular mechanisms linking these enhanced activities to the clinical phenotype of THC4 remain an area of active investigation, with current hypotheses suggesting a disruption in the balance between platelet production and apoptosis.

Mutations in Cyt-c can disrupt its interactions with mitochondrial, cytoplasmic, and nuclear partners [60,79], impacting key functions. Structural changes may alter protein folding or affect binding sites, weakening interactions with mitochondrial proteins like cytochrome c oxidase [80], which are crucial for electron transport and energy production. Mutations could also affect Cyt-c role in apoptosis by altering its interaction with Apaf-1 [22], either reducing or enhancing apoptotic signaling. Additionally, although less studied, mutations might impair Cyt-c nuclear functions, such as gene regulation and DNA repair [60], by affecting its localization or binding capabilities. In addition to the genes or proteins interacting with the Cyt-c gene or protein predicted in this study, the literature also reports HIGD1A, CHCHD2, ITPR1 [79], ANP32B, eIF2α, hnRNP C1/C2, HSPA5, SET, and YWHAE [81] interactions with Cyt-c. These genes/proteins are involved in critical cellular processes such as mitochondrial function, apoptosis, protein synthesis, calcium signaling, RNA processing, stress response, and signal transduction.

Our study shows that all variants, except L99V, display higher RMSD values during MD simulation, indicating protein destabilization. The formation of the apoptosome depends on the interaction between Cyt-c and Apaf-1. Higher RMSD values are associated with increased apoptosome activation; however, the mechanism underlying increased RMSD could differ between the pathogenic variants [22]. A previous study reporting 200 ns simulation for human Cyt-c WT, and G42S displayed similar RMSD trend with higher RMSD for G42S [82]. Mutant structures show increased flexibility in the different Ω-loop structures, which could alter the heme iron coordination, particularly by weakening the native Met81-Fe or His19-Fe bonds with increase in bond lengths [67], promoting H_2O_2 access to the heme, thus increasing peroxidase activity [22,62,78,83]. During the RMSF analysis, the increased Ω-loop flexibility, specifically central loopalso impacts Cyt-c interaction with cardiolipin facilitating Cyt-c translocation to the cytosol during apoptosis [64]. This flexibility further enhances apoptosome activation, suggesting that Cyt-c and Apaf-1 interactions may follow an "induced fit" mechanism, particularly in variants with greater conformational flexibility [22]. Structural analyses show that certain highly conserved residues are key to the structure, function, folding, and stability of Cyt-c [84,85]. Conserved Tyr68 in the 60-70 helix extends its side chain into the heme pocket and forms H-bonds with Asn53, Thr79, and Met81 [67] which are important for electron transfer and apoptosis initiation as well as heme pocket stability [64,86–90]. Disruption of Tyr68-Asn53, Tyr68-Thr79, and Tyr68-Met81 H-bonds increases caspase activation, impacting peroxidase activity and apoptosis [67,87]. Our study found that most of the variants affect H-bonds. This loss can destabilize the heme conformation

by altering the orientation of Met81, Asn53, and Thr79, leading to an open active site conformation. We further observed that the increased flexibility of the mutant proteins causes an increase in the distance between critical electron-transfer residues in the active site. These changes facilitate the access of small molecules like H_2O_2 to the heme catalytic site, promoting peroxidase activity and apoptosis [67]. The increased distance between the electron-transfer residues may also lead to the inability to transfer the electrons [91]. Our 50 ns simulation study demonstrates that Cyt-c variants increase flexibility across the protein, particularly in the loops. These results align with the 200 ns simulation findings, which show a broader distribution of conformational states in the variants compared to wild-type Cyt-c, with increased flexibility specifically observed in the Ω loops [22].

Variants H27Y, G42S, Y49H, A52T, A52V, R92G, and L99V have been previously described in the literature. H27Y, found in a Chinese family with mild THC4 and low platelet counts, can destabilize protein structure and impair function as suggested by bioinformatics analysis [28]. Normally, H27 forms a crucial H-bond with P45 for maintaining heme orientation and α-helices, affecting cardiolipin binding to Cyt-c and apoptosis [92]. Replacing histidine with tyrosine may affect electron transport in the mitochondrial respiratory chain [28]. G42S shows enhanced peroxidase activity and proapoptotic effects by altering heme electronic structure and accelerating electron exchange [12,21] G42S and Y49H variants reduce respiratory levels and increase apoptosis in yeast and mouse models [13]. A52V increases peroxidase activity by destabilizing Cyt-c native conformation. G42S, Y49H, and A52V reduce overall and local stability compared to WT Cyt-c [57]. From a clinical perspective, patients with G42S and Y49H variants have reduced mean platelet counts but normal platelet volume and function, while the A52V family shows lower mean platelet counts with a platelet function defect [18]. G42S, Y49H, A52T, R92G, and L99V variants have been described in Italian families with low platelet counts, where A52T, and R92G variants showed effects on oxidative growth and respiratory activity in a yeast model [17].

Since the open conformation of THC4 mutants increases peroxidase activity, therapeutic strategies may include developing small molecule inhibitors that block the active site or reduce H_2O_2 binding, potentially preventing excessive apoptosis. These inhibitors could help restore normal megakaryopoiesis and platelet regulation, benefiting patients with THC4-related conditions. Another approach could involve designing drugs that stabilize the wild-type conformation of THC4, reducing flexibility and structural instability caused by mutations to preserve normal protein function and prevent abnormal platelet production. This strategy is similar to current treatments for misfolded or unstable proteins in diseases like cystic fibrosis or neurological disorders [93,94]. Additionally, increasing the concentration of H_2O_2 scavengers could limit its interaction with the open active site, mitigating excessive peroxidase activity and supporting normal platelet formation. The altered structural dynamics of THC4 mutants may also serve as biomarkers for diagnosing THC4-related platelet disorders. Increased peroxidase activity and changes in megakaryopoiesis could provide measurable indicators of these conditions. Monitoring peroxidase levels, apoptosis markers, and the expression of mutant THC4 forms could aid in early

diagnosis and assess therapeutic effectiveness, as these factors are closely tied to disruptions in platelet homeostasis driven by THC4 mutations.

A limitation of this *in silico* study of Cyt-c variants is that it relies on computational models and predictions, which may not fully capture the complexity of biological processes. While these models provide useful insights, they may not accurately represent the effects of variants in a real cellular environment. Additionally, *in silico* methods often depend on the quality and accuracy of available structural and sequence data, which can limit the reliability of the predictions. Experimental validation is necessary to confirm the functional impact of these mutations. Additionally, in the present study, we analyzed the NMR Cyt-c structure (PDB ID: 1j3s), which does not include the buried water molecule. Our decision to use 1j3s was driven by our focus on studying the effects of mutations on protein stability, folding, and overall dynamics in a solvated environment, which the solution structure represents well. However, we recognize that this choice may have missed important hydrogen bond interactions involving the buried water molecule.To validate our bioinformatics findings, future studies could employ biochemical and molecular techniques to examine the effects of Cyt-c mutations on peroxidase activity and apoptosis. This process includes introducing mutations via site-directed mutagenesis, expressing and purifying the mutated Cyt-c protein, and measuring peroxidase activity using substrate oxidation assays. Apoptosis can be assessed through caspase activation assays, Cyt-c release assays, and cell viability tests. Additionally, structural changes can be investigated using X-ray crystallography or NMR spectroscopy, whereas temperature scans followed by circular dichroism and/or differential scanning calorimetry may give a quantitative measurement of the protein stability. These experiments may confirm our computational predictions and enhance our understanding of the role of Cyt-c mutations in THC4.

5 References

[1] J.M. Stevens, D.A.I. Mavridou, R. Hamer, P. Kritsiligkou, A.D. Goddard, S.J. Ferguson, Cytochrome c biogenesis System I, FEBS J. 278 (2011) 4170–4178. https://doi.org/10.1111/j.1742-4658.2011.08376.x.

[2] V.E. Kagan, V.A. Tyurin, J. Jiang, Y.Y. Tyurina, V.B. Ritov, A.A. Amoscato, A.N. Osipov, N.A. Belikova, A.A. Kapralov, V. Kini, I.I. Vlasova, Q. Zhao, M. Zou, P. Di, D.A. Svistunenko, I. V Kurnikov, G.G. Borisenko, Cytochrome c acts as a cardiolipin oxygenase required for release of proapoptotic factors, Nat. Chem. Biol. 1 (2005) 223–232. https://doi.org/10.1038/nchembio727.

[3] H.A. Kalpage, J. Wan, P.T. Morse, M.P. Zurek, A.A. Turner, A. Khobeir, N. Yazdi, L. Hakim, J. Liu, A. Vaishnav, T.H. Sanderson, M.-A. Recanati, L.I. Grossman, I. Lee, B.F.P. Edwards, M. Hüttemann, Cytochrome c phosphorylation: Control of mitochondrial electron transport chain flux and apoptosis, Int. J. Biochem. Cell Biol. 121 (2020) 105704. https://doi.org/10.1016/j.biocel.2020.105704.

[4] Y.-L.P. Ow, D.R. Green, Z. Hao, T.W. Mak, Cytochrome c: functions beyond respiration, Nat. Rev. Mol. Cell Biol. 9 (2008) 532–542. https://doi.org/10.1038/nrm2434.

[5] G.R. Moore, G.W. Pettigrew, Cytochromes c, Springer Berlin Heidelberg, Berlin, Heidelberg, 1990. https://doi.org/10.1007/978-3-642-74536-2.

[6] G.W. Bushnell, G. V. Louie, G.D. Brayer, High-resolution three-dimensional structure of horse heart cytochrome c, J. Mol. Biol. 214 (1990) 585–595. https://doi.org/10.1016/0022-2836(90)90200-6.

[7] D. Alvarez-Paggi, L. Hannibal, M.A. Castro, S. Oviedo-Rouco, V. Demicheli, V. Tórtora, F. Tomasina, R. Radi, D.H. Murgida, Multifunctional Cytochrome c : Learning New Tricks from an Old Dog, Chem. Rev. 117 (2017) 13382–13460. https://doi.org/10.1021/acs.chemrev.7b00257.

[8] R. Santucci, F. Sinibaldi, P. Cozza, F. Polticelli, L. Fiorucci, Cytochrome c: An extreme multifunctional protein with a key role in cell fate, Int. J. Biol. Macromol. 136 (2019) 1237–1246. https://doi.org/10.1016/j.ijbiomac.2019.06.180.

[9] Y. Zhao, Z.-B. Wang, J.-X. Xu, Effect of cytochrome c on the generation and elimination of O2*- and H2O2 in mitochondria, J. Biol. Chem. 278 (2003) 2356–2360. https://doi.org/10.1074/jbc.M209681200.

[10] L. Min, X. Jian-xing, Detoxifying function of cytochrome c against oxygen toxicity, Mitochondrion 7 (2007) 13–16. https://doi.org/10.1016/j.mito.2006.11.011.

[11] G.C. Brown, V. Borutaite, Regulation of apoptosis by the redox state of cytochrome c, Biochim. Biophys. Acta - Bioenerg. 1777 (2008) 877–881. https://doi.org/10.1016/j.bbabio.2008.03.024.

[12] I.M. Morison, E.M. Cramer Bordé, E.J. Cheesman, P.L. Cheong, A.J. Holyoake, S. Fichelson, R.J. Weeks, A. Lo, S.M.K. Davies, S.M. Wilbanks, R.D. Fagerlund, M.W. Ludgate, F.M. da Silva Tatley, M.S.A. Coker, N.A. Bockett, G. Hughes, D.A. Pippig, M.P. Smith, C. Capron, E.C. Ledgerwood, A mutation of human cytochrome c enhances the

intrinsic apoptotic pathway but causes only thrombocytopenia, Nat. Genet. 40 (2008) 387–389. https://doi.org/10.1038/ng.103.

[13] D. De Rocco, C. Cerqua, P. Goffrini, G. Russo, A. Pastore, F. Meloni, E. Nicchia, C.T. Moraes, A. Pecci, L. Salviati, A. Savoia, Mutations of cytochrome c identified in patients with thrombocytopenia THC4 affect both apoptosis and cellular bioenergetics, Biochim. Biophys. Acta - Mol. Basis Dis. 1842 (2014) 269–274. https://doi.org/10.1016/j.bbadis.2013.12.002.

[14] B. Johnson, G.C. Lowe, J. Futterer, M. Lordkipanidze, D. MacDonald, M.A. Simpson, I. Sanchez-Guiu, S. Drake, D. Bem, V. Leo, S.J. Fletcher, B. Dawood, J. Rivera, D. Allsup, T. Biss, P.H. Bolton-Maggs, P. Collins, N. Curry, C. Grimley, B. James, M. Makris, J. Motwani, S. Pavord, K. Talks, J. Thachil, J. Wilde, M. Williams, P. Harrison, P. Gissen, S. Mundell, A. Mumford, M.E. Daly, S.P. Watson, N. V. Morgan, Whole exome sequencing identifies genetic variants in inherited thrombocytopenia with secondary qualitative function defects, Haematologica 101 (2016) 1170–1179. https://doi.org/10.3324/haematol.2016.146316.

[15] Y. Uchiyama, K. Yanagisawa, S. Kunishima, M. Shiina, Y. Ogawa, M. Nakashima, J. Hirato, E. Imagawa, A. Fujita, K. Hamanaka, S. Miyatake, S. Mitsuhashi, A. Takata, N. Miyake, K. Ogata, H. Handa, N. Matsumoto, T. Mizuguchi, A novel CYCS mutation in the α-helix of the CYCS C-terminal domain causes non-syndromic thrombocytopenia, Clin. Genet. 94 (2018) 548–553. https://doi.org/10.1111/cge.13423.

[16] A. Pecci, C.L. Balduini, Inherited thrombocytopenias: an updated guide for clinicians, Blood Rev. 48 (2021) 100784. https://doi.org/10.1016/j.blre.2020.100784.

[17] A. Marzollo, S. Zampieri, S. Barozzi, M.A. Yousaf, J. Quartararo, D. De Rocco, M. Faleschini, C. Marconi, C.C. Berti, V. Bozzi, G. Russo, P. Giordano, P. Goffrini, S. Bresolin, A. Pastore, A. Savoia, A. Pecci, Thrombocytopenia 4 (<scp>THC4</scp>): Six novel families with mutations of the cytochrome c gene, Br. J. Haematol. (2024). https://doi.org/10.1111/bjh.19567.

[18] E.C. Ledgerwood, C. Dunstan-Harrison, L. Ong, I.M. Morison, CYCS gene variants associated with thrombocytopenia, Platelets 30 (2019) 672–674. https://doi.org/10.1080/09537104.2018.1543866.

[19] L. Salviati, Cytochrome c Defects in Human Disease, in: Mitochondrial Dis., Springer International Publishing, Cham, 2021: pp. 191–200. https://doi.org/10.1007/978-3-030-70147-5_7.

[20] E. Turro, W.J. Astle, K. Megy, S. Gräf, D. Greene, O. Shamardina, H.L. Allen, A. Sanchis-Juan, M. Frontini, C. Thys, J. Stephens, R. Mapeta, O.S. Burren, K. Downes, M. Haimel, S. Tuna, S.V. V. Deevi, T.J. Aitman, D.L. Bennett, P. Calleja, K. Carss, M.J. Caulfield, P.F. Chinnery, P.H. Dixon, D.P. Gale, R. James, A. Koziell, M.A. Laffan, A.P. Levine, E.R. Maher, H.S. Markus, J. Morales, N.W. Morrell, A.D. Mumford, E. Ormondroyd, S. Rankin, A. Rendon, S. Richardson, I. Roberts, N.B.A. Roy, M.A. Saleem, K.G.C. Smith, H. Stark, R.Y.Y. Tan, A.C. Themistocleous, A.J. Thrasher, H. Watkins, A.R. Webster, M.R. Wilkins,

C. Williamson, J. Whitworth, S. Humphray, D.R. Bentley, S. Abbs, L. Abulhoul, J. Adlard, M. Ahmed, T.J. Aitman, H. Alachkar, D.J. Allsup, J. Almeida-King, P. Ancliff, R. Antrobus, R. Armstrong, G. Arno, S. Ashford, W.J. Astle, A. Attwood, P. Aurora, C. Babbs, C. Bacchelli, T. Bakchoul, S. Banka, T. Bariana, J. Barwell, J. Batista, H.E. Baxendale, P.L. Beales, D.L. Bennett, D.R. Bentley, A. Bierzynska, T. Biss, M.A.K. Bitner-Glindzicz, G.C. Black, M. Bleda, I. Blesneac, D. Bockenhauer, H. Bogaard, C.J. Bourne, S. Boyce, J.R. Bradley, E. Bragin, G. Breen, P. Brennan, C. Brewer, M. Brown, A.C. Browning, M.J. Browning, R.J. Buchan, M.S. Buckland, T. Bueser, C.B. Diz, J. Burn, S.O. Burns, O.S. Burren, N. Burrows, P. Calleja, C. Campbell, G. Carr-White, K. Carss, R. Casey, M.J. Caulfield, J. Chambers, J. Chambers, M.M.Y. Chan, C. Cheah, F. Cheng, P.F. Chinnery, M. Chitre, M.T. Christian, C. Church, J. Clayton-Smith, M. Cleary, N.C. Brod, G. Coghlan, E. Colby, T.R.P. Cole, J. Collins, P.W. Collins, C. Colombo, C.J. Compton, R. Condliffe, S. Cook, H.T. Cook, N. Cooper, P.A. Corris, A. Furnell, F. Cunningham, N.S. Curry, A.J. Cutler, M.J. Daniels, M. Dattani, L.C. Daugherty, J. Davis, A. De Soyza, S.V. V. Deevi, T. Dent, C. Deshpande, E.F. Dewhurst, P.H. Dixon, S. Douzgou, K. Downes, A.M. Drazyk, E. Drewe, D. Duarte, T. Dutt, J.D.M. Edgar, K. Edwards, W. Egner, M.N. Ekani, P. Elliott, W.N. Erber, M. Erwood, M.C. Estiu, D.G. Evans, G. Evans, T. Everington, M. Eyries, H. Fassihi, R. Favier, J. Findhammer, D. Fletcher, F.A. Flinter, R.A. Floto, T. Fowler, J. Fox, A.J. Frary, C.E. French, K. Freson, M. Frontini, D.P. Gale, H. Gall, V. Ganesan, M. Gattens, C. Geoghegan, T.S.A. Gerighty, A.G. Gharavi, S. Ghio, H.-A. Ghofrani, J.S.R. Gibbs, K. Gibson, K.C. Gilmour, B. Girerd, N.S. Gleadall, S. Goddard, D.B. Goldstein, K. Gomez, P. Gordins, D. Gosal, S. Gräf, J. Graham, L. Grassi, D. Greene, L. Greenhalgh, A. Greinacher, P. Gresele, P. Griffiths, S. Grigoriadou, R.J. Grocock, D. Grozeva, M. Gurnell, S. Hackett, C. Hadinnapola, W.M. Hague, R. Hague, M. Haimel, M. Hall, H.L. Hanson, E. Haque, K. Harkness, A.R. Harper, C.L. Harris, D. Hart, A. Hassan, G. Hayman, A. Henderson, A. Herwadkar, J. Hoffman, S. Holden, R. Horvath, H. Houlden, A.C. Houweling, L.S. Howard, F. Hu, G. Hudson, J. Hughes, A.P. Huissoon, M. Humbert, S. Humphray, S. Hunter, M. Hurles, M. Irving, L. Izatt, R. James, S.A. Johnson, S. Jolles, J. Jolley, D. Josifova, N. Jurkute, T. Karten, J. Karten, M.A. Kasanicki, H. Kazkaz, R. Kazmi, P. Kelleher, A.M. Kelly, W. Kelsall, C. Kempster, D.G. Kiely, N. Kingston, R. Klima, N. Koelling, M. Kostadima, G. Kovacs, A. Koziell, R. Kreuzhuber, T.W. Kuijpers, A. Kumar, D. Kumararatne, M.A. Kurian, M.A. Laffan, F. Lalloo, M. Lambert, H.L. Allen, A. Lawrie, D.M. Layton, N. Lench, C. Lentaigne, T. Lester, A.P. Levine, R. Linger, H. Longhurst, L.E. Lorenzo, E. Louka, P.A. Lyons, R.D. Machado, R. V. MacKenzie Ross, B. Madan, E.R. Maher, J. Maimaris, S. Malka, S. Mangles, R. Mapeta, K.J. Marchbank, S. Marks, H.S. Markus, H.-U. Marschall, A. Marshall, J. Martin, M. Mathias, E. Matthews, H. Maxwell, P. McAlinden, M.I. McCarthy, H. McKinney, A. McMahon, S. Meacham, A.J. Mead, I.M. Castello, K. Megy, S.G. Mehta, M. Michaelides, C. Millar, S.N. Mohammed, S. Moledina, D. Montani, A.T. Moore, J. Morales, N.W. Morrell, M. Mozere, K.W. Muir, A.D. Mumford, A.H. Nemeth, W.G. Newman, M. Newnham, S. Noorani, P. Nurden, J. O'Sullivan, S. Obaji,

C. Odhams, S. Okoli, A. Olschewski, H. Olschewski, K.R. Ong, S.H. Oram, E. Ormondroyd, W.H. Ouwehand, C. Palles, S. Papadia, S.-M. Park, D. Parry, S. Patel, J. Paterson, A. Peacock, S.H. Pearce, J. Peden, K. Peerlinck, C.J. Penkett, J. Pepke-Zaba, R. Petersen, C. Pilkington, K.E.S. Poole, R. Prathalingam, B. Psaila, A. Pyle, R. Quinton, S. Rahman, S. Rankin, A. Rao, F.L. Raymond, P.J. Rayner-Matthews, C. Rees, A. Rendon, T. Renton, C.J. Rhodes, A.S.C. Rice, S. Richardson, A. Richter, L. Robert, I. Roberts, A. Rogers, S.J. Rose, R. Ross-Russell, C. Roughley, N.B.A. Roy, D.M. Ruddy, O. Sadeghi-Alavijeh, M.A. Saleem, N. Samani, C. Samarghitean, A. Sanchis-Juan, R.B. Sargur, R.N. Sarkany, S. Satchell, S. Savic, J.A. Sayer, G. Sayer, L. Scelsi, A.M. Schaefer, S. Schulman, R. Scott, M. Scully, C. Searle, W. Seeger, A. Sen, W.A.C. Sewell, D. Seyres, N. Shah, O. Shamardina, S.E. Shapiro, A.C. Shaw, P.J. Short, K. Sibson, L. Side, I. Simeoni, M.A. Simpson, M.C. Sims, S. Sivapalaratnam, D. Smedley, K.R. Smith, K.G.C. Smith, K. Snape, N. Soranzo, F. Soubrier, L. Southgate, O. Spasic-Boskovic, S. Staines, E. Staples, H. Stark, J. Stephens, C. Steward, K.E. Stirrups, A. Stuckey, J. Suntharalingam, E.M. Swietlik, P. Syrris, R.C. Tait, K. Talks, R.Y.Y. Tan, K. Tate, J.M. Taylor, J.C. Taylor, J.E. Thaventhiran, A.C. Themistocleous, E. Thomas, D. Thomas, M.J. Thomas, P. Thomas, K. Thomson, A.J. Thrasher, G. Threadgold, C. Thys, T. Tilly, M. Tischkowitz, C. Titterton, J.A. Todd, C.-H. Toh, B. Tolhuis, I.P. Tomlinson, M. Toshner, M. Traylor, C. Treacy, P. Treadaway, R. Trembath, S. Tuna, W. Turek, E. Turro, P. Twiss, T. Vale, C. Van Geet, N. van Zuydam, M. Vandekuilen, A.M. Vandersteen, M. Vazquez-Lopez, J. von Ziegenweidt, A. Vonk Noordegraaf, A. Wagner, Q. Waisfisz, S.M. Walker, N. Walker, K. Walter, J.S. Ware, H. Watkins, C. Watt, A.R. Webster, L. Wedderburn, W. Wei, S.B. Welch, J. Wessels, S.K. Westbury, J.-P. Westwood, J. Wharton, D. Whitehorn, J. Whitworth, A.O.M. Wilkie, M.R. Wilkins, C. Williamson, B.T. Wilson, E.K.S. Wong, N. Wood, Y. Wood, C.G. Woods, E.R. Woodward, S.J. Wort, A. Worth, M. Wright, K. Yates, P.F.K. Yong, T. Young, P. Yu, P. Yu-Wai-Man, E. Zlamalova, N. Kingston, N. Walker, J.R. Bradley, S. Ashford, C.J. Penkett, K. Freson, K.E. Stirrups, F.L. Raymond, W.H. Ouwehand, Whole-genome sequencing of patients with rare diseases in a national health system, Nature 583 (2020) 96–102. https://doi.org/10.1038/s41586-020-2434-2.

[21] M.D. Liptak, R.D. Fagerlund, E.C. Ledgerwood, S.M. Wilbanks, K.L. Bren, The Proapoptotic G41S Mutation to Human Cytochrome c Alters the Heme Electronic Structure and Increases the Electron Self-Exchange Rate, J. Am. Chem. Soc. 133 (2011) 1153–1155. https://doi.org/10.1021/ja106328k.

[22] M. Fellner, R. Parakra, K.O. McDonald, I. Kass, G.N.L. Jameson, S.M. Wilbanks, E.C. Ledgerwood, Altered structure and dynamics of pathogenic cytochrome c variants correlate with increased apoptotic activity, Biochem. J. 478 (2021) 669–684. https://doi.org/10.1042/BCJ20200793.

[23] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242. https://doi.org/10.1093/nar/28.1.235.

[24] A. Bateman, UniProt: A worldwide hub of protein knowledge, Nucleic Acids Res. 47 (2019) D506–D515. https://doi.org/10.1093/nar/gky1049.

[25] F. Sievers, D.G. Higgins, Clustal Omega for making accurate alignments of many protein sequences, Protein Sci. 27 (2018) 135–145. https://doi.org/10.1002/pro.3290.

[26] G.E. Crooks, G. Hon, J.-M. Chandonia, S.E. Brenner, WebLogo: A Sequence Logo Generator: Figure 1, Genome Res. 14 (2004) 1188–1190. https://doi.org/10.1101/gr.849004.

[27] K. Megy, K. Downes, M. Morel-Kopp, J.M. Bastida, S. Brooks, L. Bury, E. Leinoe, K. Gomez, N. V. Morgan, M. Othman, W.H. Ouwehand, J. Perez Botero, J. Rivera, H. Schulze, D. Trégouët, K. Freson, GoldVariants, a resource for sharing rare genetic variants detected in bleeding, thrombotic, and platelet disorders: Communication from the ISTH SSC Subcommittee on Genomics in Thrombosis and Hemostasis, J. Thromb. Haemost. 19 (2021) 2612–2617. https://doi.org/10.1111/jth.15459.

[28] F. Che, J. Zhao, Y. Zhao, Z. Wang, L. Zhang, Y. Yang, A Novel Heterozygous Pathogenic Variation in CYCS Gene Cause Autosomal Dominant Non-Syndromic Thrombocytopenia 4 in a Large Chinese Family, Front. Genet. 12 (2022). https://doi.org/10.3389/fgene.2021.783455.

[29] K. Downes, K. Megy, D. Duarte, M. Vries, J. Gebhart, S. Hofer, O. Shamardina, S.V. V. Deevi, J. Stephens, R. Mapeta, S. Tuna, N. Al Hasso, M.W. Besser, N. Cooper, L. Daugherty, N. Gleadall, D. Greene, M. Haimel, H. Martin, S. Papadia, S. Revel-Vilk, S. Sivapalaratnam, E. Symington, W. Thomas, C. Thys, A. Tolios, C.J. Penkett, W.H. Ouwehand, S. Abbs, M.A. Laffan, E. Turro, I. Simeoni, A.D. Mumford, Y.M.C. Henskens, I. Pabinger, K. Gomez, K. Freson, Diagnostic high-throughput sequencing of 2396 patients with bleeding, thrombotic, and platelet disorders, Blood 134 (2019) 2082–2091. https://doi.org/10.1182/blood.2018891192.

[30] E. Gálvez, E. Vallespín, E.G. Arias-Salgado, C. Sánchez-Valdepeñas, Y. Giménez, S. Navarro, P. Río, M. Bogliolo, R. Pujol, M. Peiró, J. Nevado, J. Zubicaray, E. Sebastián, A. Catalá, C. Beléndez, C. Díaz de Heredia, A. Galera, I. Badell, L. Madero, R. Perona, L. Sastre, J. Surrallés, J. Bueren, P. Lapunzina, J. Sevilla, Next-generation Sequencing in Bone Marrow Failure Syndromes and Isolated Cytopenias: Experience of the Spanish Network on Bone Marrow Failure Syndromes, HemaSphere 5 (2021) e539. https://doi.org/10.1097/HS9.0000000000000539.

[31] O. Gilad, O. Dgany, S. Noy-Lotan, T. Krasnov, J. Yacobovich, R. Rabinowicz, T. Goldberg, A.A. Kuperman, A. Abu-Quider, H. Miskin, N. Kapelushnik, N. Mandel-Shorer, S. Shimony, D. Harlev, T. Ben-Ami, E. Adam, C. Levin, S. Aviner, R. Elhasid, S. Berger-Achituv, L. Chaitman-Yerushalmi, Y. Kodman, N. Oniashvilli, M. Hameiri-Grosman, S. Izraeli, H. Tamary, O. Steinberg-Shemer, Syndromes predisposing to leukemia are a major cause of inherited cytopenias in children, Haematologica 107 (2022) 2081–2095. https://doi.org/10.3324/haematol.2021.280116.

[32] K. Giavi, S. Glentis, A. Bouchla, A. Apostolidou, N.M. Marinakis, A. Kattamis, E.

Katsantoni, V. Pappa, A Novel Variant of the CYCS Gene Alters Apoptosis of Megakaryocytes in a Family with Thrombocytopenia, Blood 142 (2023) 5409–5409. https://doi.org/10.1182/blood-2023-186060.

[33] P.C. Ng, S. Henikoff, SIFT: predicting amino acid changes that affect protein function, Nucleic Acids Res. 31 (2003) 3812–3814. https://doi.org/10.1093/NAR/GKG509.

[34] C.M. Yates, I. Filippis, L.A. Kelley, M.J.E. Sternberg, SuSPect: Enhanced prediction of single amino acid variant (SAV) phenotype using network features, J. Mol. Biol. 426 (2014) 2692–2701. https://doi.org/10.1016/j.jmb.2014.04.026.

[35] V. Pejaver, J. Urresti, J. Lugo-Martinez, K.A. Pagel, G.N. Lin, H.J. Nam, M. Mort, D.N. Cooper, J. Sebat, L.M. Iakoucheva, S.D. Mooney, P. Radivojac, Inferring the molecular and phenotypic impact of amino acid variants with MutPred2, Nat. Commun. 11 (2020). https://doi.org/10.1038/s41467-020-19669-x.

[36] W. Deng, Y. Wang, L. Ma, Y. Zhang, S. Ullah, Y. Xue, Computational prediction of methylation types of covalently modified lysine and arginine residues in proteins, Brief. Bioinform. (2016) bbw041. https://doi.org/10.1093/bib/bbw041.

[37] N. Blom, S. Gammeltoft, S. Brunak, Sequence and structure-based prediction of eukaryotic protein phosphorylation sites, J. Mol. Biol. 294 (1999) 1351–1362. https://doi.org/10.1006/jmbi.1999.3310.

[38] P. V. Hornbeck, J.M. Kornhauser, S. Tkachev, B. Zhang, E. Skrzypek, B. Murray, V. Latham, M. Sullivan, PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse, Nucleic Acids Res. 40 (2012) D261–D270. https://doi.org/10.1093/nar/gkr1122.

[39] D. Warde-Farley, S.L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, M. Franz, C. Grouios, F. Kazi, C.T. Lopes, A. Maitland, S. Mostafavi, J. Montojo, Q. Shao, G. Wright, G.D. Bader, Q. Morris, The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function, Nucleic Acids Res. 38 (2010) W214–W220. https://doi.org/10.1093/nar/gkq537.

[40] D. Szklarczyk, A.L. Gable, K.C. Nastou, D. Lyon, R. Kirsch, S. Pyysalo, N.T. Doncheva, M. Legeay, T. Fang, P. Bork, L.J. Jensen, C. von Mering, The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets, Nucleic Acids Res. 49 (2021) D605–D612. https://doi.org/10.1093/nar/gkaa1074.

[41] C.H.M. Rodrigues, D.E.V. Pires, D.B. Ascher, DynaMut2 : Assessing changes in stability and flexibility upon single and multiple point missense mutations, Protein Sci. 30 (2021) 60–69. https://doi.org/10.1002/pro.3942.

[42] J. Cheng, A. Randall, P. Baldi, Prediction of protein stability changes for single-site mutations using support vector machines, Proteins Struct. Funct. Bioinforma. 62 (2005) 1125–1132. https://doi.org/10.1002/prot.20810.

[43] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, Gromacs:

High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19–25. https://doi.org/10.1016/J.SOFTX.2015.06.001.

[44] B. Hess, C. Kutzner, D. Van Der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, J. Chem. Theory Comput. 4 (2008) 435–447. https://doi.org/10.1021/ct700301q.

[45] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

[46] J.R. López-Blanco, J.I. Aliaga, E.S. Quintana-Ortí, P. Chacón, IMODS: Internal coordinates normal mode analysis server, Nucleic Acids Res. 42 (2014). https://doi.org/10.1093/nar/gku339.

[47] M. Landau, I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko, N. Ben-Tal, ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures, Nucleic Acids Res. 33 (2005) W299–W302. https://doi.org/10.1093/nar/gki370.

[48] G. Celniker, G. Nimrod, H. Ashkenazy, F. Glaser, E. Martz, I. Mayrose, T. Pupko, N. Ben-Tal, ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function, Isr. J. Chem. 53 (2013) 199–206. https://doi.org/10.1002/ijch.201200096.

[49] A.M. Lesk, C. Chothia, How different amino acid sequences determine similar protein structures: The structure and evolutionary dynamics of the globins, J. Mol. Biol. 136 (1980) 225–270. https://doi.org/10.1016/0022-2836(80)90373-3.

[50] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F.T. Heer, T.A.P. De Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: Homology modelling of protein structures and complexes, Nucleic Acids Res. 46 (2018) W296–W303. https://doi.org/10.1093/nar/gky427.

[51] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291. https://doi.org/10.1107/S0021889892009944.

[52] W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637. https://doi.org/10.1002/bip.360221211.

[53] B. Nutho, S. Samsri, S. Pornsuwan, Structural Dynamics of the Precatalytic State of Human Cytochrome c upon T28C, G34C, and A50C Mutations: A Molecular Dynamics Simulation Perspective, ACS Omega 8 (2023) 15229–15238. https://doi.org/10.1021/acsomega.3c00220.

[54] F. Sinibaldi, B.D. Howes, M.C. Piro, P. Caroppi, G. Mei, F. Ascoli, G. Smulevich, R. Santucci, Insights into the role of the histidines in the structure and stability of cytochrome c, JBIC J. Biol. Inorg. Chem. 11 (2006) 52–62. https://doi.org/10.1007/s00775-005-0057-6.

[55] S. Ittisoponpisan, S.A. Islam, T. Khanna, E. Alhuzimi, A. David, M.J.E. Sternberg, Can Predicted Protein 3D Structures Provide Reliable Insights into whether Missense Variants

Are Disease Associated?, J. Mol. Biol. 431 (2019) 2197–2212. https://doi.org/10.1016/j.jmb.2019.04.009.

[56] O.M. Deacon, A.I. Karsisiotis, T. Moreno-Chicano, M.A. Hough, C. Macdonald, T.M.A. Blumenschein, M.T. Wilson, G.R. Moore, J.A.R. Worrall, Heightened Dynamics of the Oxidized Y48H Variant of Human Cytochrome c Increases Its Peroxidatic Activity, Biochemistry 56 (2017) 6111–6124. https://doi.org/10.1021/acs.biochem.7b00890.

[57] H. Lei, B.E. Bowler, Naturally Occurring A51V Variant of Human Cytochrome c Destabilizes the Native State and Enhances Peroxidase Activity, J. Phys. Chem. B 123 (2019) 8939–8953. https://doi.org/10.1021/acs.jpcb.9b05869.

[58] K. Shinwari, H.M. Rehman, N. Xiao, L. Guojun, M.A. Khan, M.A. Bolkov, I.A. Tuzankina, V.A. Chereshnev, Novel high-risk missense mutations identification in FAT4 gene causing Hennekam syndrome and Van Maldergem syndrome 2 through molecular dynamics simulation, Informatics Med. Unlocked 37 (2023) 101160. https://doi.org/10.1016/j.imu.2023.101160.

[59] Y. Shiloh, Y. Ziv, The ATM protein kinase: regulating the cellular response to genotoxic stress, and more, Nat. Rev. Mol. Cell Biol. 14 (2013) 197–210. https://doi.org/10.1038/nrm3546.

[60] P.T. Morse, T. Arroum, J. Wan, L. Pham, A. Vaishnav, J. Bell, L. Pavelich, M.H. Malek, T.H. Sanderson, B.F.P. Edwards, M. Hüttemann, Phosphorylations and Acetylations of Cytochrome c Control Mitochondrial Respiration, Mitochondrial Membrane Potential, Energy, ROS, and Apoptosis, Cells 13 (2024) 493. https://doi.org/10.3390/cells13060493.

[61] H.A. Kalpage, V. Bazylianska, M.A. Recanati, A. Fite, J. Liu, J. Wan, N. Mantena, M.H. Malek, I. Podgorski, E.I. Heath, A. Vaishnav, B.F. Edwards, L.I. Grossman, T.H. Sanderson, I. Lee, M. Huttemann, Tissue-specific regulation of cytochrome c by post-translational modifications: respiration, the mitochondrial membrane potential, ROS, and apoptosis, FASEB J. 33 (2019) 1540–1553. https://doi.org/10.1096/fj.201801417R.

[62] A.I. Karsisiotis, O.M. Deacon, M.T. Wilson, C. Macdonald, T.M.A. Blumenschein, G.R. Moore, J.A.R. Worrall, Increased dynamics in the 40–57 Ω-loop of the G41S variant of human cytochrome c promote its pro-apoptotic conformation, Sci. Rep. 6 (2016) 30447. https://doi.org/10.1038/srep30447.

[63] P. Brochette, C. Petit, M.P. Pileni, Cytochrome c in sodium bis(2-ethylhexyl) sulfosuccinate reverse micelles: structure and reactivity, J. Phys. Chem. 92 (1988) 3505–3511. https://doi.org/10.1021/j100323a037.

[64] G. Muneeswaran, S. Kartheeswaran, M. Pandiaraj, K. Muthukumar, M. Sankaralingam, S. Arunachalam, Investigation of structural dynamics of Thrombocytopenia Cargeeg mutants of human apoptotic cytochrome c : A molecular dynamics simulation approach, Biophys. Chem. 230 (2017) 117–126. https://doi.org/10.1016/j.bpc.2017.09.004.

[65] S. Abel, M. Waks, M. Marchi, Molecular dynamics simulations of cytochrome c unfolding in AOT reverse micelles: The first steps, Eur. Phys. J. E 32 (2010) 399–409. https://doi.org/10.1140/epje/i2010-10635-x.

[66] G. Muneeswaran, S. Kartheeswaran, K. Muthukumar, C.D. Dharmaraj, C. Karunakaran, Comparative structural and conformational studies on H43R and W32F mutants of copper–zinc superoxide dismutase by molecular dynamics simulation, Biophys. Chem. 185 (2014) 70–78. https://doi.org/10.1016/j.bpc.2013.11.010.

[67] G. Muneeswaran, M. Pandiaraj, S. Kartheeswaran, M. Sankaralingam, K. Muthukumar, C. Karunakaran, Molecular dynamics simulation approach to explore atomistic molecular mechanism of peroxidase activity of apoptotic cytochrome c mutants, Informatics Med. Unlocked 11 (2018) 51–60. https://doi.org/10.1016/j.imu.2018.04.003.

[68] S.W. Englander, L. Mayne, Z.-Y. Kan, W. Hu, Protein Folding—How and Why: By Hydrogen Exchange, Fragment Separation, and Mass Spectrometry, Annu. Rev. Biophys. 45 (2016) 135–152. https://doi.org/10.1146/annurev-biophys-062215-011121.

[69] A.K. Agrahari, P. Sneha, C. George Priya Doss, R. Siva, H. Zayed, A profound computational study to prioritize the disease-causing mutations in PRPS1 gene, Metab. Brain Dis. 33 (2018) 589–600. https://doi.org/10.1007/s11011-017-0121-2.

[70] Z.-P. Wang, X.-Z. Ding, J. Wang, Y.-M. Li, Double-edged sword in cells: chemical biology studies of the vital role of cytochrome c in the intrinsic pre-apoptotic mitochondria leakage pathway, RSC Adv. 5 (2015) 28258–28269. https://doi.org/10.1039/C4RA16856A.

[71] J. Zheng, S. Ye, T. Lu, T.M. Cotton, G. Chumanov, Circular dichroism and resonance Raman comparative studies of wild type cytochrome c and F82H mutant, Biopolymers 57 (2000) 77. https://doi.org/10.1002/(SICI)1097-0282(2000)57:2<77::AID-BIP4>3.3.CO;2-P.

[72] J. Zheng, S. Ye, T. Lu, T.M. Cotton, G. Chumanov, Circular dichroism and resonance Raman comparative studies of wild type cytochromec and F82H mutant, Biopolymers 57 (2000) 77–84. https://doi.org/10.1002/(SICI)1097-0282(2000)57:2<77::AID-BIP4>3.0.CO;2-Y.

[73] T. Takano, O.B. Kallai, R. Swanson, R.E. Dickerson, The Structure of Ferrocytochrome c at 2.45 A Resolution, J. Biol. Chem. 248 (1973) 5234–5255. https://doi.org/10.1016/S0021-9258(19)43594-1.

[74] M. Hüttemann, P. Pecina, M. Rainbolt, T.H. Sanderson, V.E. Kagan, L. Samavati, J.W. Doan, I. Lee, The multiple functions of cytochrome c and their regulation in life and death decisions of the mammalian cell: From respiration to apoptosis, Mitochondrion 11 (2011) 369–381. https://doi.org/10.1016/j.mito.2011.01.010.

[75] A.M. Firsov, E.A. Kotova, E.A. Korepanova, A.N. Osipov, Y.N. Antonenko, Peroxidative permeabilization of liposomes induced by cytochrome c/cardiolipin complex, Biochim. Biophys. Acta - Biomembr. 1848 (2015) 767–774. https://doi.org/10.1016/j.bbamem.2014.11.027.

[76] R. V. Chertkova, A.M. Firsov, N.A. Brazhe, E.I. Nikelshparg, Z. V. Bochkova, T. V. Bryantseva, M.A. Semenova, A.A. Baizhumanov, E.A. Kotova, M.P. Kirpichnikov, G. V. Maksimov, Y.N. Antonenko, D.A. Dolgikh, Multiple Mutations in the Non-Ordered Red Ω-Loop Enhance the Membrane-Permeabilizing and Peroxidase-like Activity of

Cytochrome c, Biomolecules 12 (2022) 665. https://doi.org/10.3390/biom12050665.
[77] A. Guerra-Castellano, I. Márquez, G. Pérez-Mejías, A. Díaz-Quintana, M.A. De la Rosa, I. Díaz-Moreno, Post-Translational Modifications of Cytochrome c in Cell Life and Disease, Int. J. Mol. Sci. 21 (2020) 8483. https://doi.org/10.3390/ijms21228483.
[78] O.M. Deacon, R.W. White, G.R. Moore, M.T. Wilson, J.A.R. Worrall, Comparison of the structural dynamic and mitochondrial electron-transfer properties of the proapoptotic human cytochrome c variants, G41S, Y48H and A51V, J. Inorg. Biochem. 203 (2020) 110924. https://doi.org/10.1016/j.jinorgbio.2019.110924.
[79] Z. Zhou, T. Arroum, X. Luo, R. Kang, Y.J. Lee, D. Tang, M. Hüttemann, X. Song, Diverse functions of cytochrome c in cell death and disease, Cell Death Differ. 31 (2024) 387–404. https://doi.org/10.1038/s41418-024-01284-8.
[80] R. V. Chertkova, N.A. Brazhe, T. V. Bryantseva, A.N. Nekrasov, D.A. Dolgikh, A.I. Yusipovich, O. Sosnovtseva, G. V. Maksimov, A.B. Rubin, M.P. Kirpichnikov, New insight into the mechanism of mitochondrial cytochrome c function, PLoS One 12 (2017) e0178280. https://doi.org/10.1371/journal.pone.0178280.
[81] J. Martínez-Fábregas, I. Díaz-Moreno, K. González-Arzola, S. Janocha, J.A. Navarro, M. Hervás, R. Bernhardt, A. Velázquez-Campoy, A. Díaz-Quintana, M.A. De la Rosa, Structural and Functional Analysis of Novel Human Cytochrome c Targets in Apoptosis, Mol. Cell. Proteomics 13 (2014) 1439–1456. https://doi.org/10.1074/mcp.M113.034322.
[82] T.C. Chin, S.M. Wilbanks, E.C. Ledgerwood, Altered conformational dynamics contribute to species-specific effects of cytochrome c mutations on caspase activation, JBIC J. Biol. Inorg. Chem. 29 (2024) 169–176. https://doi.org/10.1007/s00775-024-02044-2.
[83] S. Samsri, S. Pornsuwan, Influence of cysteine-directed mutations at the Ω-loops on peroxidase activity of human cytochrome c., Arch. Biochem. Biophys. 709 (2021) 108980. https://doi.org/10.1016/j.abb.2021.108980.
[84] S. Sperandio, I. de Belle, D.E. Bredesen, An alternative, nonapoptotic form of programmed cell death, Proc. Natl. Acad. Sci. 97 (2000) 14376–14381. https://doi.org/10.1073/pnas.97.26.14376.
[85] T. Ying, Z.-H. Wang, Y.-W. Lin, J. Xie, X. Tan, Z.-X. Huang, Tyrosine-67 in cytochrome c is a possible apoptotic trigger controlled by hydrogen bonds via a conformational transition, Chem. Commun. (2009) 4512. https://doi.org/10.1039/b904347k.
[86] T.L. Luntz, A. Schejter, E.A. Garber, E. Margoliash, Structural significance of an internal water molecule studied by site-directed mutagenesis of tyrosine-67 in rat cytochrome c., Proc. Natl. Acad. Sci. 86 (1989) 3524–3528. https://doi.org/10.1073/pnas.86.10.3524.
[87] E. Margoliash, S. Ferguson-Miller, J. Tulloss, C.H. Kang, B.A. Feinberg, D.L. Brautigan, M. Morrison, Separate Intramolecular Pathways for Reduction and Oxidation of Cytochrome c in Electron Transport Chain Reactions, Proc. Natl. Acad. Sci. 70 (1973) 3245–3249. https://doi.org/10.1073/pnas.70.11.3245.
[88] L.A. Abriata, A. Cassina, V. Tórtora, M. Marín, J.M. Souza, L. Castro, A.J. Vila, R. Radi, Nitration of Solvent-exposed Tyrosine 74 on Cytochrome c Triggers Heme Iron-Methionine

80 Bond Disruption, J. Biol. Chem. 284 (2009) 17–26. https://doi.org/10.1074/jbc.M807203200.

[89] D. Roccatano, I. Daidone, M.A. Ceruso, C. Bossa, A. Di Nola, Selective excitation of native fluctuations during thermal unfolding simulations: Horse heart cytochrome c as a case study, Biophys. J. 84 (2003) 1876–1883. https://doi.org/10.1016/S0006-3495(03)74995-9.

[90] W. Pulawski, S. Filipek, A. Zwolinska, A. Debinski, K. Krzysko, R. Garduño-Juárez, S. Viswanathan, V. Renugopalakrishnan, Low-temperature molecular dynamics simulations of horse heart cytochrome c and comparison with inelastic neutron scattering data, Eur. Biophys. J. 42 (2013) 291–300. https://doi.org/10.1007/s00249-012-0874-9.

[91] S.R. Singh, S. Prakash, V. Vasu, C. Karunakaran, Conformational flexibility decreased due to Y67F and F82H mutations in cytochrome c: Molecular dynamics simulation studies, J. Mol. Graph. Model. 28 (2009) 270–277. https://doi.org/10.1016/j.jmgm.2009.08.005.

[92] G. Balakrishnan, Y. Hu, T.G. Spiro, His26 Protonation in Cytochrome c Triggers Microsecond β-Sheet Formation and Heme Exposure: Implications for Apoptosis, J. Am. Chem. Soc. 134 (2012) 19061–19069. https://doi.org/10.1021/ja307100a.

[93] M. Eisenstein, Congruence Therapeutics: finding a fix for misfolded proteins, Nat. Biotechnol. (2023). https://doi.org/10.1038/d41587-023-00009-5.

[94] J.-C. Rochet, Novel therapeutic strategies for the treatment of protein-misfolding diseases, Expert Rev. Mol. Med. 9 (2007) 1–34. https://doi.org/10.1017/S1462399407000385.

Figures

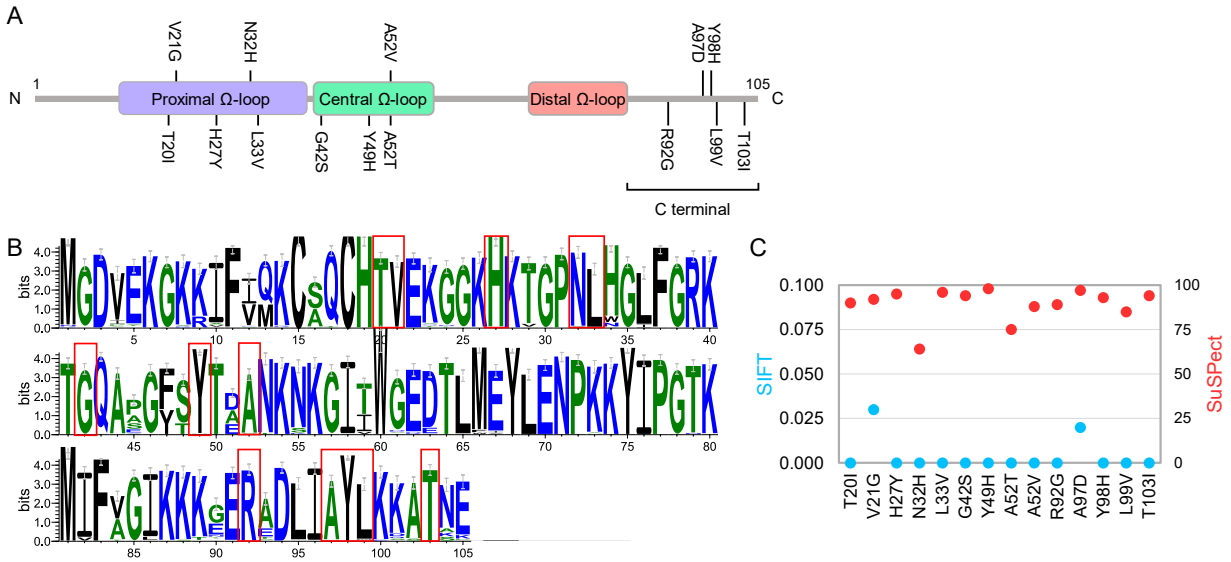

Fig. 1. Overall structure of Cyt-c and information on the position and conservation of the clinically important variants. (A) The Cyt-c variants, T20I, V21G, H27Y, N32H and L33V are located in the proximal loop, G42S, Y49H, A52T and A52V in the central loop and R92G, A97D, Y98H, L99V and T103I in the protein C-terminus. (B) Sequence logo representing the conservation of amino acids in the protein sequence. The height of each letter at a given position corresponds to the level of conservation, measured in bits, where taller letters indicate higher conservation. Amino acids are color-coded based on their chemical properties: polar residues are green, basic residues are blue, acidic residues are red, and hydrophobic residues are black. Red boxes highlight the positions of the mutated residues: T20I, V21G, H27Y, N32H, L33V, G42S, Y49H, A52T, A52V, R92G, A97D, Y98H, L99V, and T103I. These mutations occur at moderately to highly conserved positions, suggesting potential functional significance. (C) SIFT analysis shows that all variants affect protein function (score < 0.05) and SuSPect associates the variants with the disease (score > 50).

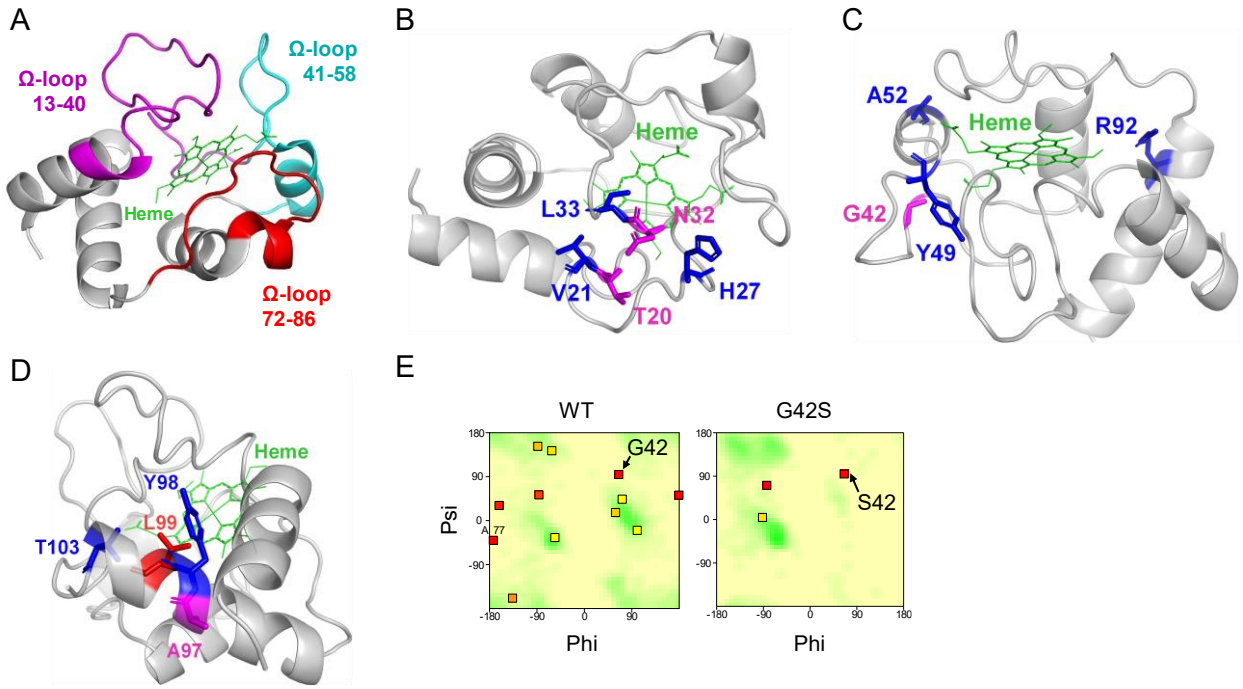

Fig. 2. Mapping the currently known clinically important variants on the structure of Cyt-c, and Ramachandran plot analysis. The protein structure is depicted as a grey ribbon, with heme in green. (A) Human Cyt-c (PDB ID: 1j3s) structure showing three Ω-loop regions: proximal (residues 13-40, magenta), central (residues 41-58, cyan) and distal (residues 72-86, red) Ω-loops, which make the large part of Cyt-c. These loops are important for both structural and functional reasons, including protein folding and conformational transition. Amino acids investigated for variants are highlighted as follows: (B) T20 (magenta), V21 (blue), H27 (blue), N32 (magenta), L33 (blue); (C) G42 (magenta), Y49 (blue), A52 (blue), R92 (blue); (D) A97 (magenta), Y98 (blue), L99 (red), T103 (blue). Note that the residue numbering is offset by +1 relative to the PDB entry 1j3s. (E) In the Ramachandran plot, G42 is an exception, positioned in a typically disallowed region due to its small size and high flexibility. The G42S variant introduces steric hindrance due to larger and less flexible serine in a disallowed region.

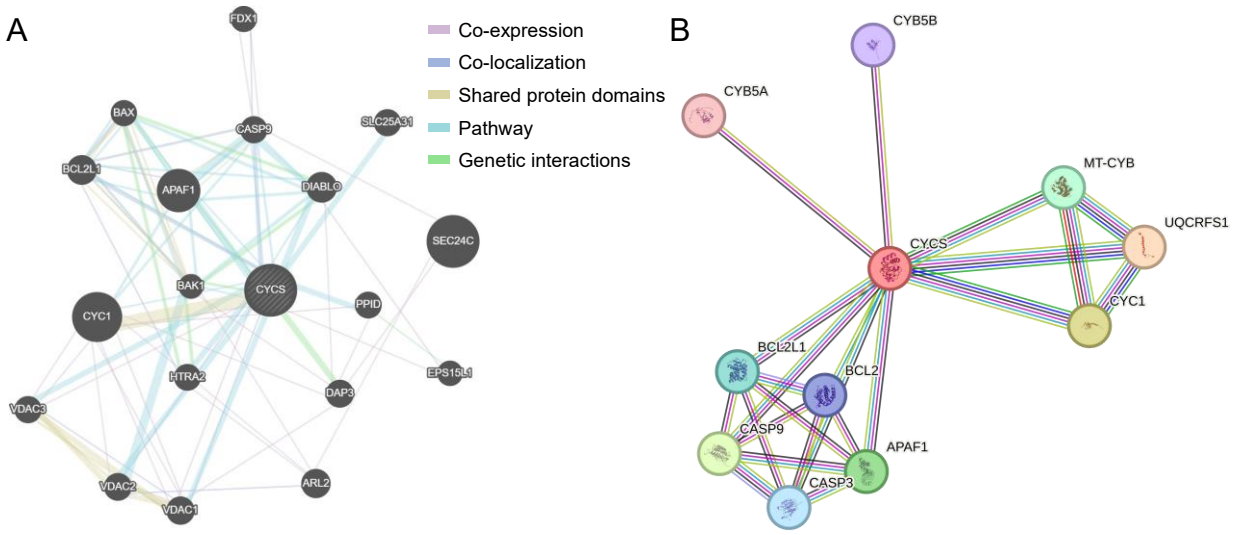

Fig. 3. Visualization of gene and protein interactions. (A) Interactions of the *CYCS* gene with other genes, analyzed using the GeneMANIA tool. (B) The protein-protein interaction network of Cyt-c analyzed using the STRING tool. Variants in these networks may affect the functions of interacting genes or proteins.

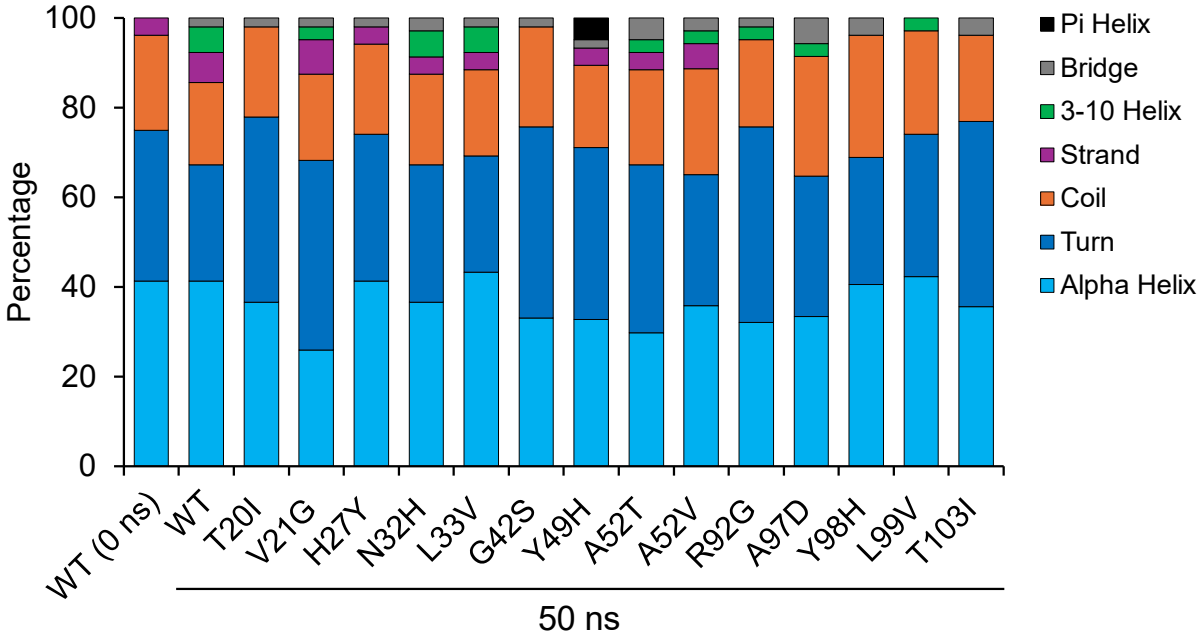

Fig. 4. The percentage distribution of secondary structure composition of the Cyt-c variants before (WT at 0 ns) and after the 50 ns simulation (50 ns). The variants display changes in the alpha helix, turn, coil, strand, 310 helix, bridge, and pi helix after simulation.

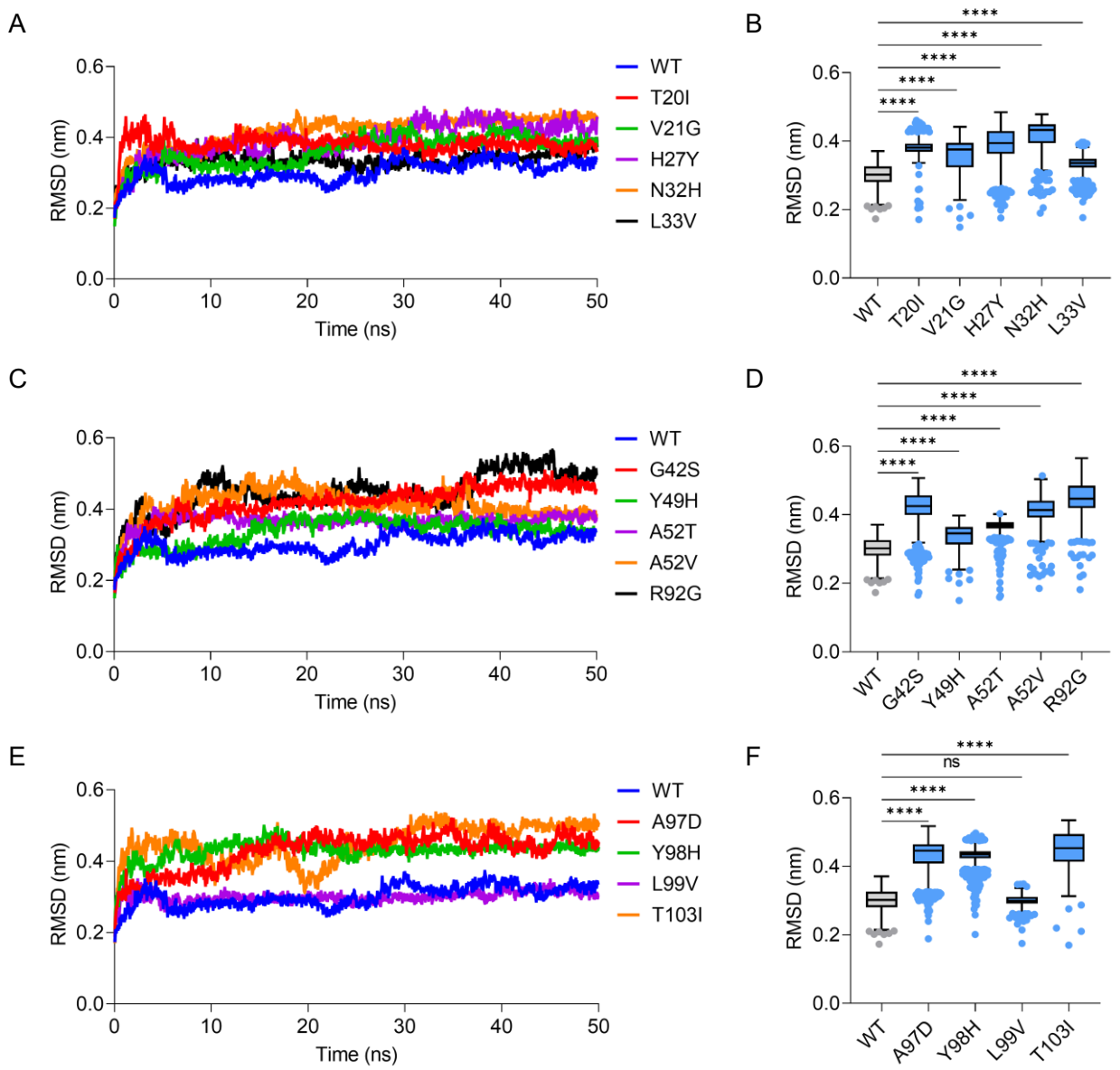

Fig. 5. Root Mean Square Deviation (RMSD) profiles of WT and mutant Cyt-c protein structures over a 50 ns (50000 ps) of MD simulation. (A) Fluctuations of RMSD values for WT and T20I, V21G, H27Y, N32H and L33V mutant proteins. (B) RMSD values of mutants T20I (median 0.381 nm, IQR 0.369-0.392 nm), V21G (median 0.375 nm, IQR 0.324-0.395 nm), H27Y (median 0.394 nm, IQR 0.363-0.429 nm), N32H (median 0.432 nm, IQR 0.394-0.448 nm), and L33V (median 0.336 nm, IQR 0.323-0.348 nm) were significantly higher than the WT (median 0.302 nm, IQR 0.281-0.325 nm) (****: $p < 0.0001$). (C) The mutant proteins showed higher RMSD values (D) which were significant (****: $p < 0.0001$) compared to WT WT (median 0.302 nm, IQR 0.281-0.325 nm). The mutant displayed median as G42S (median 0.425 nm, IQR 0.400-0.456 nm), Y49H (median 0.346 nm, IQR 0.314-0.364 nm), A52T (median 0.369 nm, IQR 0.360-0.377 nm), A52V (median 0.413 nm, IQR 0.392-0.441 nm), and R92G (median 0.447 nm, IQR 0.420-0.484 nm). (E) Variations of RMSD values over simulation where (F) A97D (median 0.446 nm, IQR 0.409-0.464 nm), Y98H (median 0.434 nm, IQR 0.423-0.444 nm), and T103I (median 0.453 nm, IQR

0.414-0.494 nm). displayed significantly higher RMSD values than the WT (median 0.302 nm, IQR 0.281-0.325 nm) (****: $p < 0.0001$). However, L99V (median 0.300 nm, IQR 0.291-0.310 nm) showed non-significant (ns) differences. In boxplots, the boxes represent the interquartile range (25th to 75th percentiles), and the lines within the boxes indicate the median values.

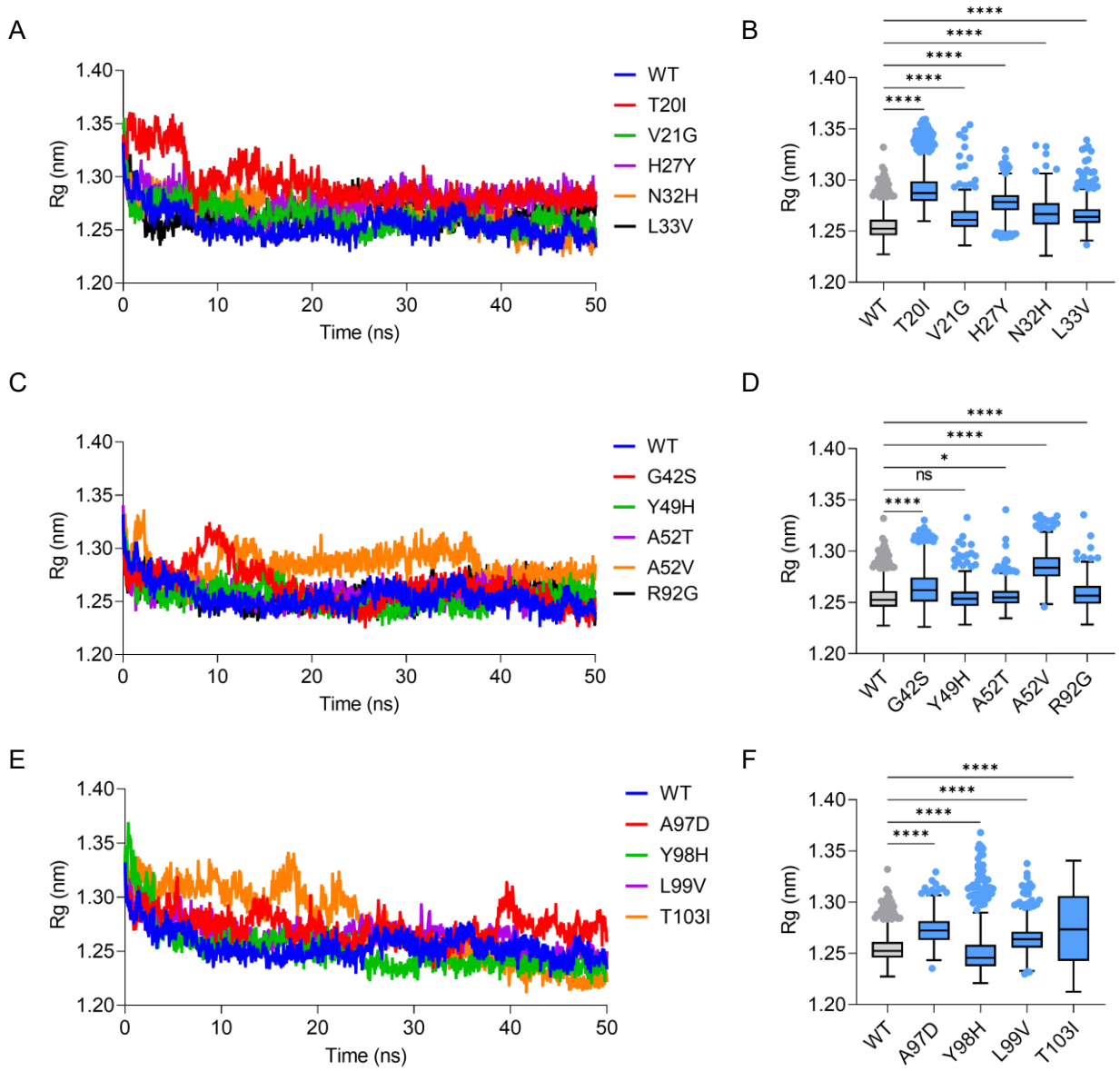

Fig. 6. Radius of gyration (Rg) analysis of WT and mutant Cyt-c protein structures over a 50 ns of MD simulation. (A) Variants of Rg values for WT and T20I, V21G, H27Y, N32H, and L33V mutant proteins. (B) Rg values of mutants T20I (median 1.287 nm, IQR 1.280-1.299 nm), V21G (median 1.261 nm, IQR 1.254-1.270 nm), H27Y (median 1.279, nm IQR 1.271-1.285 nm), N32H (median 1.267 nm, IQR 1.257-1.277 nm), and L33V (median 1.264 nm, IQR 1.258-1.271 nm) were significantly higher than the WT (median 1.252 nm, IQR 1.246-1.261 nm) (****: $p < 0.0001$). (C) Rg profiles for G42S, Y49H, A52T, A52V and R92G (D) where G42S (median 1.262 nm, IQR 1.251-1.274 nm), A52V (median 1.284 nm, IQR 1.276-1.294 nm), and R92G (median 1.257 nm, IQR 1.249-1.266 nm) were significant (****: $p < 0.0001$)

compared to WT (median 1.252 nm, IQR 1.246-1.261 nm). Y49H (median 1.254, IQR 1.247-1.261) showed non-significant results while A52T (median 1.255 nm, IQR 1.249-1.261 nm) displayed less significant differences (*: $p = 0.03$). (E) Rg variations over simulation where (F) A97D (median 1.272 nm, IQR 1.263-1.281 nm), Y98H (median 1.246 nm, IQR 1.237-1.258 nm), L99V (median 1.264 nm, IQR 1.256-1.271 nm), and T103I (median 1.273 nm, IQR 1.243-1.306 nm) displayed significantly higher Rg values than the WT (median 1.252 nm, IQR 1.246-1.261 nm) (****: $p < 0.0001$). In boxplots, the boxes represent the interquartile range (25th to 75th percentiles), and the lines within the boxes indicate the median values.

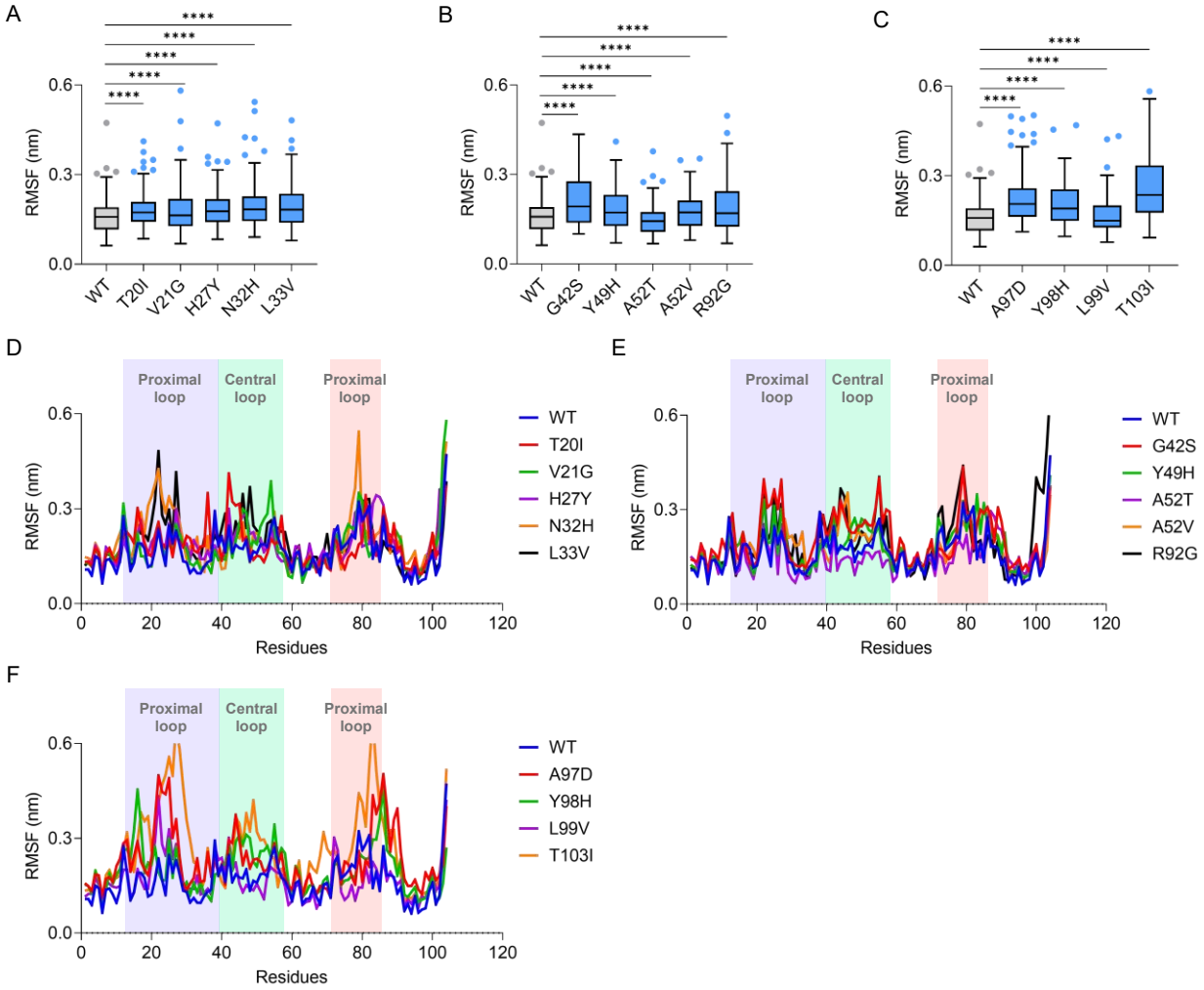

Fig. 7. Root mean square fluctuation (RMSF) profiles of Cyt-c WT and mutant protein structures. (A) Mutants T20I (median 0.173 nm, IQR 0.143-0.208 nm), V21G (median 0.164 nm, IQR 0.129-0.218 nm), H27Y (median 0.177 nm, IQR 0.142-0.217 nm), N32H (median 0.184 nm, IQR 0.146-0.225 nm), and L33V (median 0.182 nm, IQR 0.141-0.233 nm) displayed significantly higher RMSF values compared to the WT (median 0.158 nm, IQR 0.117-0.190 nm) (****: $p < 0.0001$). (B) G42S (median 0.193 nm, IQR 0.140-0.275 nm), Y49H (median 0.172 nm, IQR 0.129-0.229 nm), A52T (median 0.143 nm, IQR 0.108-0.173 nm), A52V (median 0.173 nm, IQR 0.129-0.211 nm), and R92G (median 0.170 nm, IQR 0.128-0.242 nm) exhibited significantly higher RMSF compared to the WT (****: $p < 0.0001$). (C) WT showed significantly lower RMSF values compared to A97D (median 0.205 nm, IQR 0.164-0.256 nm), Y98H (median 0.190

nm, IQR 0.150-0.253 nm), L99V (median 0.149 nm, IQR 0.127-0.199 nm), and T103I (median 0.236 nm, IQR 0.181-0.330 nm) (****: $p < 0.0001$). In boxplots, the boxes represent the interquartile range (25th to 75th percentiles), and the lines within the boxes indicate the median values. (D) Impact of variants on Ω-loop dynamics. T20I and V21G increased flexibility in proximal (residues 13-40) and central (residues 41-58) loops, respectively, while H27Y enhanced flexibility across all Ω-loops, suggesting altered protein dynamics. N32H and L33V variants notably elevated RMSF in proximal (residues 13-40) and central (residues 41-58) loops, with L33V affecting the distal (residues 72-86) loop as well. (E) G42S variant increased flexibility in all loops. Y49H showed increased RMSF in proximal and central regions; A52T minimally impacted the proximal loop but reduced RMSF in central and distal loops. A52V and R92G variants increased flexibility in proximal and central regions, with R92G exhibiting significant fluctuations across all loops and the N-terminus. (F) A97D, Y98H and T103I variants significantly increased RMSF values in all loop regions.

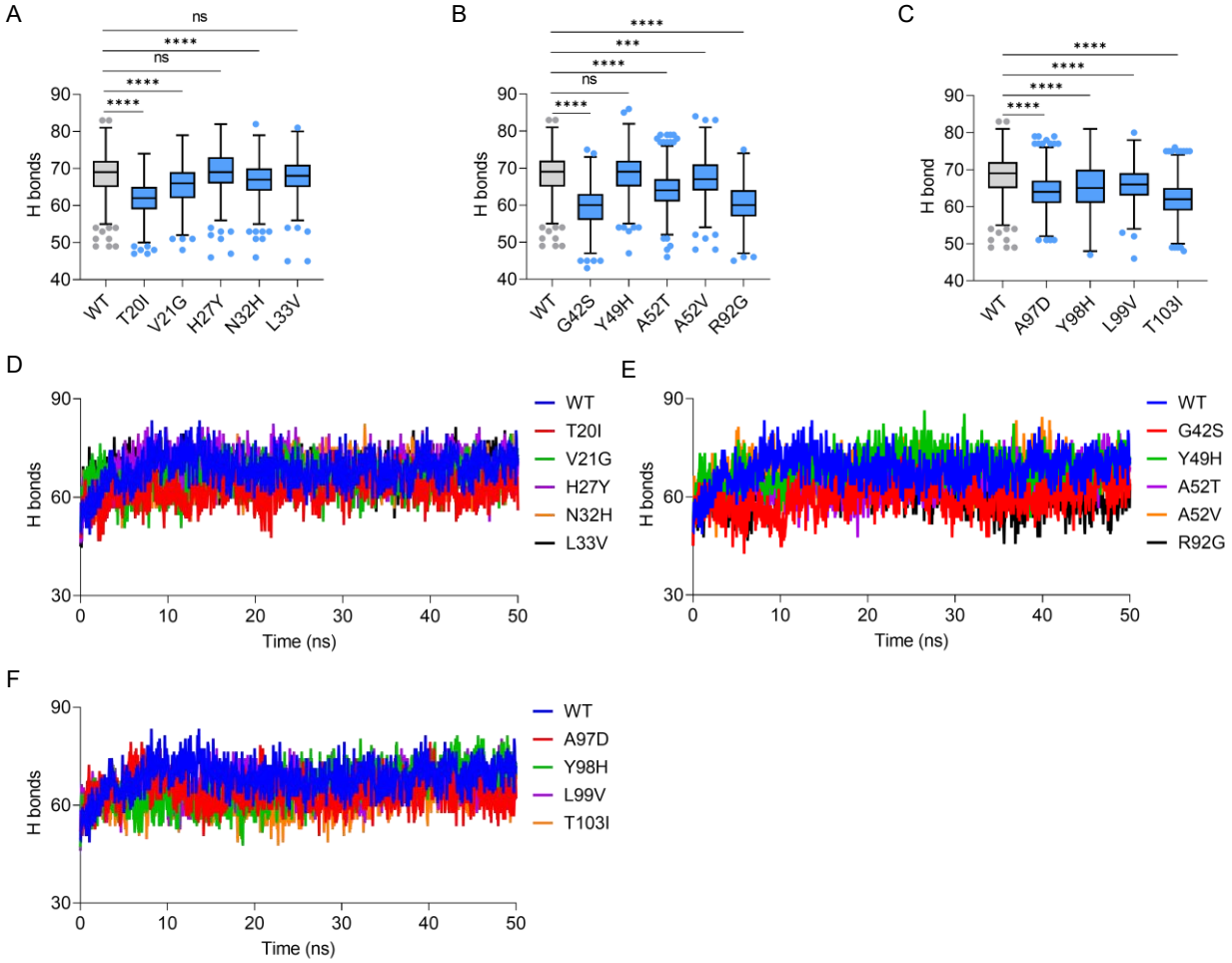

Fig. 8. Variations of H-bonds in Cyt-c WT and mutant structures. (A) WT structure (median 69.000, IQR 65.000-72.000) showed a significantly higher number of H-bonds than that of T20I (median 1.287, IQR 1.280-1.299), V21G (median 1.261, IQR 1.254-1.270), and N32H (median 1.267, IQR 1.257-1.277) (****: $p < 0.0001$). Differences of WT with H27Y (median 1.279, IQR 1.271-1.285), and L33V (median 1.264, IQR 1.258-1.271)were non-significant. (B) G42S (median 60.000, IQR 56.000-63.000), A52T (median 64.000, IQR 61.000-67.000), A52V (median 67.000, IQR 64.000-71.000), and R92G (median 60.000, IQR

57.000-64.000) showed significantly lower (****: $p < 0.0001$) compared to the WT while the difference with Y49H (median 69.000, IQR 65.000-72.000) was non-significant. (C) A97D (median 64.000, IQR 61.000-67.000), Y98H (median 65.000, IQR 61.000-70.000), L99V (median 66.000, IQR 63.000-69.000), and T103I (median 62.000, IQR 59.000-65.000) displayed significantly lesser (****: $p < 0.0001$) than the WT. In boxplots, the boxes represent the interquartile range (25th to 75th percentiles), and the lines within the boxes indicate the median values. Variation of H-bonds for WT and mutants (D) T20I, V21G, H27Y, N32H, L33V, (E) G42S, Y49H, A52T, A52V, R92G, (F) A97D, Y98H, L99V and T103I over the MD simulation of 50 ns (50000 ps).

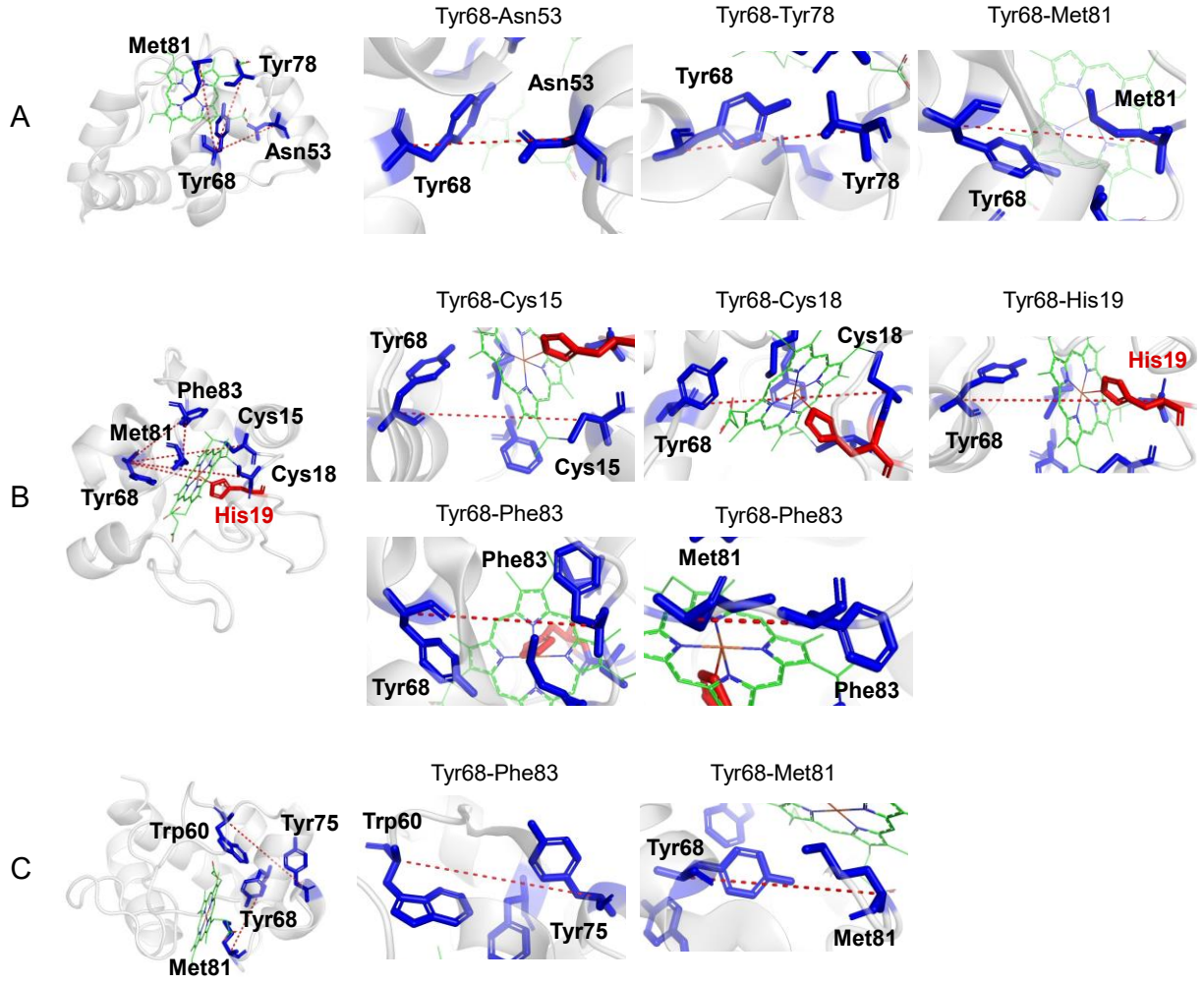

Fig. 9. Distances between H-bonded and active site residues of Cyt-c. (A) Distances between Tyr68-Asn53, Tyr68-Thr79, and Tyr68-Met81. Tyr68 forms a hydrogen bond network with the conserved residues Asn53 and Thr79, as well as the axial ligand Met81. Most variants lead to a decrease in these distances, indicating the loss of H-bonds. (B) Most variants also result in increased distances between key active site residues, including Tyr68-Cys15, Tyr68-Cys18, Tyr68-His19, Tyr68-Phe83, and Met81-Phe83. (C) Distances between Trp60-Tyr75 and Tyr68-Met81 are crucial for the electron transfer process. Variants generally cause an increase in the distance between these residues. Collectively, the increased distances between these key residues lead to a more open active site, increasing the accessibility of the heme catalytic site to small molecules such as H_2O_2, enhancing peroxidase activity and facilitating the initiation of apoptosis.

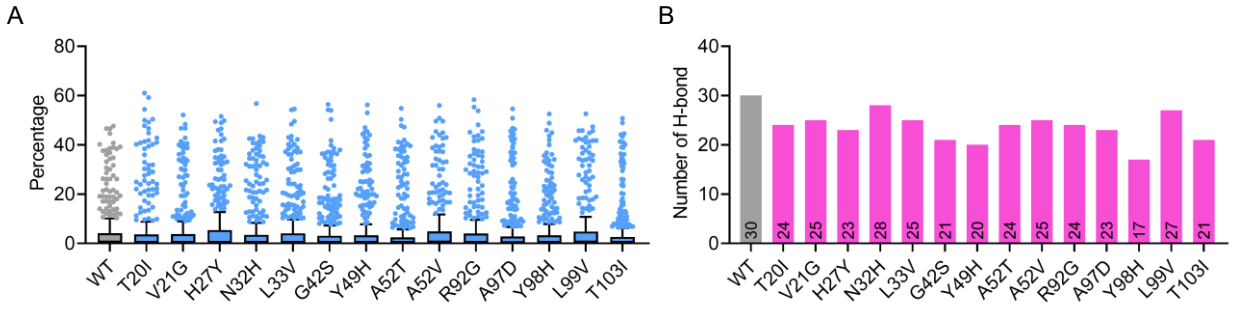

Fig. 10. Percentage H-bond occupancy for WT and mutant proteins over a 50 ns MD simulation. (A) Boxplot showing the H-bond occupancy distribution for WT and mutant proteins throughout the simulation, with each point representing the percentage of time a specific H-bond was maintained. (B) Bar chart displaying H-bonds with occupancy greater than 30% for both WT and mutant proteins, highlighting the most stable H-bonds across the variants. The comparison indicates differences in H-bond stability and occupancy between WT and mutants, which may influence the overall structural integrity and function of the proteins.