

<https://doi.org/10.1038/s41540-025-00582-2>

Integrative structural profiling and ligand optimisation across the transthyretin mutational landscape



Ugo Lomoio^{1,6}, Valentina Carbonari^{1,6}, Federico Manuel Giorgi², Francesco Ortuso³, Pietro Lió⁴, Pierangelo Veltri⁵ ✉ & Pietro Hiram Guzzi¹

Transthyretin amyloidosis (ATTR) is a genetically diverse disorder caused by destabilising mutations in the transthyretin (TTR) protein, leading to pathological aggregation. While stabilisers like tafamidis and acoramidis are approved, their efficacy across TTR variants remains unclear. This study presents an in silico pipeline combining AlphaFold3 for structure prediction, ESM2 for sequence embeddings, DiffDock-L and AutoDock Vina for molecular docking, and DiffSBDD for ligand generation. Simulations show that binding affinities of approved ligands vary significantly among TTR variants, with some mutations (e.g., W61L, Y98F) reducing binding despite being distant from the binding site. Embedding-based clustering highlights potential benign mutations and supports scalable variant classification. Additionally, customised ligand optimisation can recover binding affinity in specific cases, though effects are mutation-dependent. These findings emphasise the need for variant-aware therapeutic strategies. This integrative approach offers a foundation for precision drug design in ATTR, enabling the development of personalised stabilisers tailored to individual mutational profiles.

Transthyretin (TTR) is a plasma protein predominantly synthesised in the liver, where it functions as a carrier for thyroxine (T4) and retinol—the latter in complex with retinol-binding protein^{1,2}. Its three-dimensional structure, illustrated in Fig. 1, comprises four identical subunits of ~14 kDa each, forming a highly stable homotetramer rich in β -sheet content^{2–10}.

The *TTR* gene (Gene ID 7276; GenBank accession NG_009490.1) is located on chromosome 18q12.1 and spans ~6.9 kb⁷. It consists of four exons separated by three introns. Exon 1 encodes a 23-residue sequence that includes a 20-residue signal peptide and the first three residues of the mature protein, while exons 2–4 encode residues 4–127. Throughout this study, residue numbering refers to the mature 127-residue monomer, excluding the signal peptide. Variants are therefore annotated using mature protein nomenclature (e.g., Val30Met or V30M), while genomic nomenclature includes the signal peptide (e.g., p.Val50Met or p.V50M).

To date, 216 point mutations have been reported within the first 127 residues of the mature TTR monomer¹¹, encompassing both benign and pathogenic variants^{12–14}. While some variants have been extensively characterised, biochemical and clinical data on rare mutations remain limited. Most reported changes are single amino acid substitutions, although notable exceptions include the Val122_{del} deletion¹⁴, duplications at Met13, Glu51,

and Ser52, and compound heterozygous mutations observed in individual patients¹⁴. In this study, the initial selection of mutations was guided primarily by their established pathogenicity, as reported in clinical and molecular literature. Among these pathogenic variants, we prioritised mutations that are predominantly associated with cardiological and neurological phenotypes—two of the most clinically relevant manifestations of ATTR. The final subset used was chosen based on a combination of disease prevalence, clinical severity, and evidence of structural destabilization. Finally, we selected the three most globally prevalent variants for molecular dynamics analysis. Figure 2 represents the analysis of 133 point mutations, primarily those classified as pathogenic or of uncertain significance.

Previous investigations have examined the spatial distribution of TTR mutations to assess potential clustering. While residue-level analysis did not reveal significant aggregation, sliding window approaches (e.g., seven-residue intervals)^{14,15} identified mutation hotspots. Notably, highly amyloidogenic and clinically aggressive variants cluster in regions 1 and 2 (residues 26–59), whereas non-amyloidogenic variants are more prevalent in regions 4 and 6 (residues 97–125), suggesting a non-random distribution across the sequence.

¹Department of Surgical and Medical Sciences, Magna Graecia University of Catanzaro, Catanzaro, Italy. ²Department of Pharmacy and Biotechnology, University of Bologna, Bologna, Italy. ³Department of Health Sciences, Magna Graecia University of Catanzaro, Catanzaro, Italy. ⁴Department of Computer Science and Technology, Cambridge University, Cambridge, United Kingdom. ⁵DIMES, University of Calabria, Rende, Italy. ⁶These authors contributed equally: Ugo Lomoio, Valentina Carbonari. ✉ e-mail: pierangelo.veltri@dimes.unical.it

Pathogenic mutations destabilise the native tetrameric structure, facilitating dissociation into monomers that misfold, aggregate, and form insoluble amyloid fibrils. This process underlies both hereditary transthyretin amyloidosis (ATTRv) and wild-type transthyretin amyloidosis (ATTRwt)^{16,17}. Over 150 pathogenic variants have been reported, with Val30Met being the most studied due to its high prevalence in endemic populations affected by familial amyloid polyneuropathy^{18,19}. Disease progression is marked by systemic fibril deposition, leading to multi-organ dysfunction and ultimately death^{1,20,21}.

Clinical manifestations vary by mutation. Destabilising variants often result in early-onset polyneuropathy or cardiomyopathy, whereas ATTRwt primarily presents later in life as cardiomyopathy^{22,23}. Progressive fibril accumulation leads to debilitating complications, including cardiac failure and peripheral neuropathy, with rapid clinical decline in the absence of treatment^{24,25}. Inflammatory responses have also been implicated in disease progression, as elevated cytokines may suppress hepatic TTR synthesis^{26,27}.

Beyond amyloidosis, altered TTR expression has been associated with other pathological conditions such as preeclampsia²⁸, highlighting its broader clinical relevance and potential utility as a biomarker. The complex relationship between TTR sequence variation, structural stability, and disease phenotype underscores the need for systematic molecular characterisation²⁹.

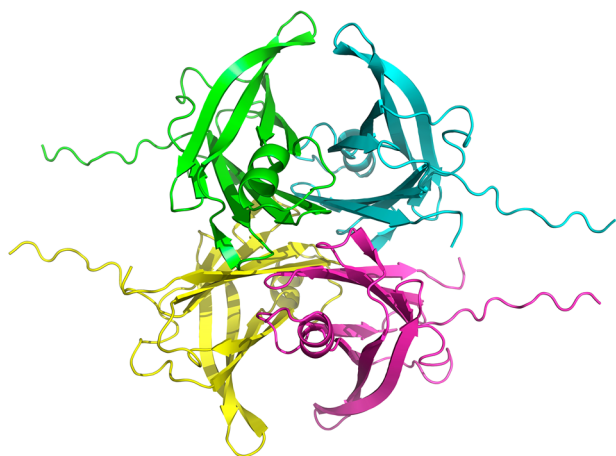


Fig. 1 | Wild-type transthyretin. PyMOL visualisation of the three-dimensional structure of wild-type transthyretin (TTR) predicted using AlphaFold3.

Under physiological conditions, TTR assembles into a kinetically stable tetramer^{17,30}. Mutations that perturb this assembly increase the pool of monomeric intermediates susceptible to aggregation^{1,18}. In contrast, stabilising mutations are associated with higher circulating TTR levels and greater longevity³¹. Therapeutic strategies aimed at stabilising the tetramer—most notably tafamidis—have shown substantial clinical benefit^{32–34}.

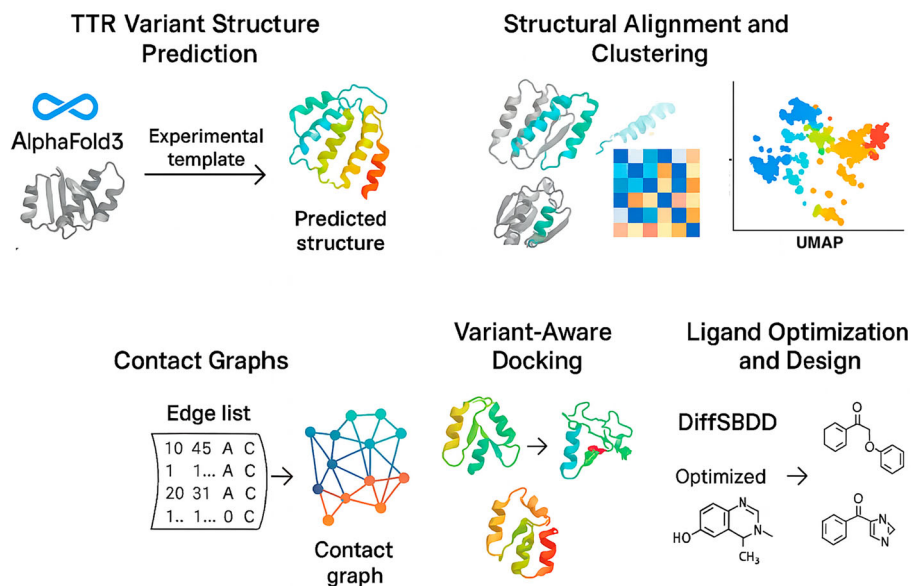
Comprehensive structural and functional profiling of TTR variants is critical for advancing precision medicine. Elucidating mutation-specific effects on folding stability, aggregation propensity, and drug binding provides key insights into disease mechanisms and therapeutic responsiveness^{35,36}. Moreover, circulating TTR concentrations and conformational changes may serve as accessible biomarkers for diagnosis and treatment monitoring^{37,38}.

This study focuses on a prioritised and curated subset of 133 TTR variants based on reported pathogenicity, with an emphasis on those associated with neurological and cardiac phenotypes—the two most clinically impactful forms of ATTR. Selection criteria included mutation prevalence, severity of clinical presentation, and evidence of structural destabilisation. The three most prevalent pathogenic variants worldwide were selected for detailed structural and dynamic characterisation via molecular simulations.

Key findings of this analysis include:

- A comprehensive mapping of mutations across the TTR sequence revealed no consistent correlation between residue position and pathogenicity.
- High-confidence structural models of both monomeric and tetrameric forms were generated for each variant using AlphaFold3.
- Predicted changes in Gibbs free energy ($\Delta\Delta G$) did not exhibit a direct relationship with pathogenic potential.
- Clustering based on structural similarity identified three major variant groups and suggested that R123S may warrant reclassification as non-pathogenic.
- Protein contact network analysis demonstrated that pathogenic mutations disproportionately impact structural regions critical for tetramer stability.
- In silico evaluations of drug binding for two clinically approved stabilisers revealed substantial mutation-dependent variability in predicted affinity.
- Structure-based modifications of existing therapeutics were proposed to improve binding efficacy in the context of destabilising variants.

Fig. 2 | Overview of the computational pipeline. Starting from the wild-type (WT) TTR sequence and known point mutations, AlphaFold3 was used to generate structural models for all the variants. Predicted structures were assessed and analysed in latent space and by means of contact networks to evaluate structural impact. Ligand docking was performed to assess mutation-dependent binding affinities. Finally, the optimisation of the existing ligand was proposed to better fit with the existing variants.



Together, these findings provide a comprehensive and structure-guided framework for understanding the molecular basis of TTR amyloidosis and informing the development of mutation-specific therapeutic strategies within a precision medicine paradigm.

Results

Overview of the results

The computational analysis of TTR variants revealed consistent structural and functional signatures associated with pathogenicity. AlphaFold3-predicted models showed high agreement with experimental structures (mean RMSD = 0.2662 Å), and TM-score clustering indicated that pathogenic mutations tend to co-localise, suggesting shared conformational effects. Embeddings from the ESM2 language model, projected via UMAP, accurately distinguished pathogenic from benign variants (ROC AUC = 0.9948), outperforming AlphaMissense, E-SNPs&GO, and VESM++. To assess functional impact, classical and AI-based docking have been combined. DiffDock-L and AutoDock Vina revealed mutation-specific binding profiles for tafamidis and acoramidis, with the latter showing higher affinity in several non-canonical contexts. Ligand optimisation via DiffSBDD improved binding to destabilised variants such as Y98F. Network analysis and MD simulations further confirmed that both ligands restore structural stability in key pathogenic variants, demonstrating the utility of AI-guided pipelines for mutation-aware drug design. A complete results overview is depicted in Fig. 3.

Structural alignment of TTR variants

Structural comparisons between AlphaFold3-predicted and experimentally resolved TTR structures revealed a high degree of agreement. The predicted tetramer achieved a TM-score of 0.9904 when aligned to the crystallographic tetramer (PDB ID: 1ICT), while the predicted monomer obtained a TM-score of 0.9958 compared to the crystallographic dimer structure (PDB ID: 3A4D), demonstrating the reliability of AlphaFold3 in reproducing native conformations.

To validate the predicted mutant structures, all available experimental mutant TTR structures were retrieved from the RCSB Protein Data Bank. For each mutation, the AlphaFold3-predicted model was aligned to its corresponding experimental structure, and the root-mean-square deviation (RMSD) was computed. A mean RMSD of 0.2662 Å (standard deviation: 0.0635) was observed for tetramer-only structures, while a slightly higher value of 0.5083 Å was recorded for tetramer-ligand complexes. These values support the structural validity of the predicted models (Table 1).

AlphaFold3-predicted models offer additional benefits: (i) they are complete, without missing residues; (ii) they include the signal peptide (residues 1–21); and (iii) they account for asymmetric interactions between monomers, whereas experimental tetramers often assume idealised global symmetries (e.g., D2 symmetry).

Pairwise structural alignments were performed for all variants, producing TM-score matrices for both monomeric and tetrameric forms. Hierarchical clustering using average linkage³⁹ revealed that pathogenic variants tended to cluster together, suggesting shared conformational perturbations, whereas benign mutations formed structurally distinct groups (Fig. 4).

The TM-score distributions for monomeric and tetrameric forms are shown in Fig. 5. Tetrameric structures exhibited narrow distributions centred around 0.96, indicating strong structural conservation. By contrast, monomeric forms showed broader variability, suggesting conformational flexibility that may underlie aggregation propensity.

Latent space analysis using ESM2 and UMAP

To explore sequence-level determinants of pathogenicity, the ESM2 transformer-based protein language model⁴⁰ has been used to compute 1280-dimensional embeddings for each TTR variant. These embeddings were projected into two dimensions using Uniform Manifold Approximation and Projection (UMAP)⁴¹, yielding an interpretable latent space representation (Fig. 6).

Clinical annotations were obtained from the UniProt Proteins API, with additional classification into amyloidogenic or non-amyloidogenic categories using the Mutations-TTR database⁴². Variants of unknown significance were often located near benign or non-amyloidogenic mutations, suggesting possible reclassification.

Subclusters were observed around residues such as L55 and V28, whose mutations (e.g., L55P, L55R, V28M) are known to cause destabilisation. Importantly, several “Unknown” variants—including S132I, M33I, and R124C—were located close to benign or non-amyloidogenic variants in embedding space. This proximity suggests that they may have similar functional profiles. For instance, R124H and R124C nearly overlapped in UMAP space; R124H is known to be benign and protective⁴³, indicating that R124C may similarly lack pathogenic effects. Conversely, S132I, while still annotated as “Unknown”, clustered with amyloidogenic variants, consistent with reports of increased aggregation propensity⁴⁴. F64L, classified as “likely pathogenic”, was also part of this cluster, consistent with its association with late-onset neuropathy and mild cardiac symptoms⁴⁵.

Taken together, the clustering patterns suggest that UMAP projections of ESM2 embeddings can reveal structural and functional similarity among variants, including those currently lacking definitive clinical annotation.

To quantify classification performance, the UMAP-based approach has been compared with respect to three state-of-the-art methods: AlphaMissense⁴⁶, E-SNPs&GO⁴⁷, and VESM++⁴⁸. As shown in Table 2, this method outperformed all baselines, achieving a ROC AUC of 0.9948, compared to 0.9219 for VESM++, and substantially higher than AlphaMissense and E-SNPs&GO. These results confirm the ability of pretrained protein language models, combined with dimensionality reduction, to capture biologically meaningful determinants of variant pathogenicity Table 3.

Network-based modelling of protein structural perturbations

Protein Contact Networks (PCNs)⁴⁹ were constructed for both wild-type and mutant TTR tetramers to investigate how single-point mutations influence residue-level centrality. For each variant, the change in centrality for every residue i relative to the wild-type has been calculated:

$$\Delta C_i = C_{i,wt} - C_{i,mut}. \quad (1)$$

Figure 7 illustrates residue-wise differences in closeness centrality across representative mutations. For visual clarity, only one chain is shown, as all four chains exhibited comparable profiles. Residues and mutations included are those for which at least five residues showed non-zero centrality changes. Complete results are available in the public repository.

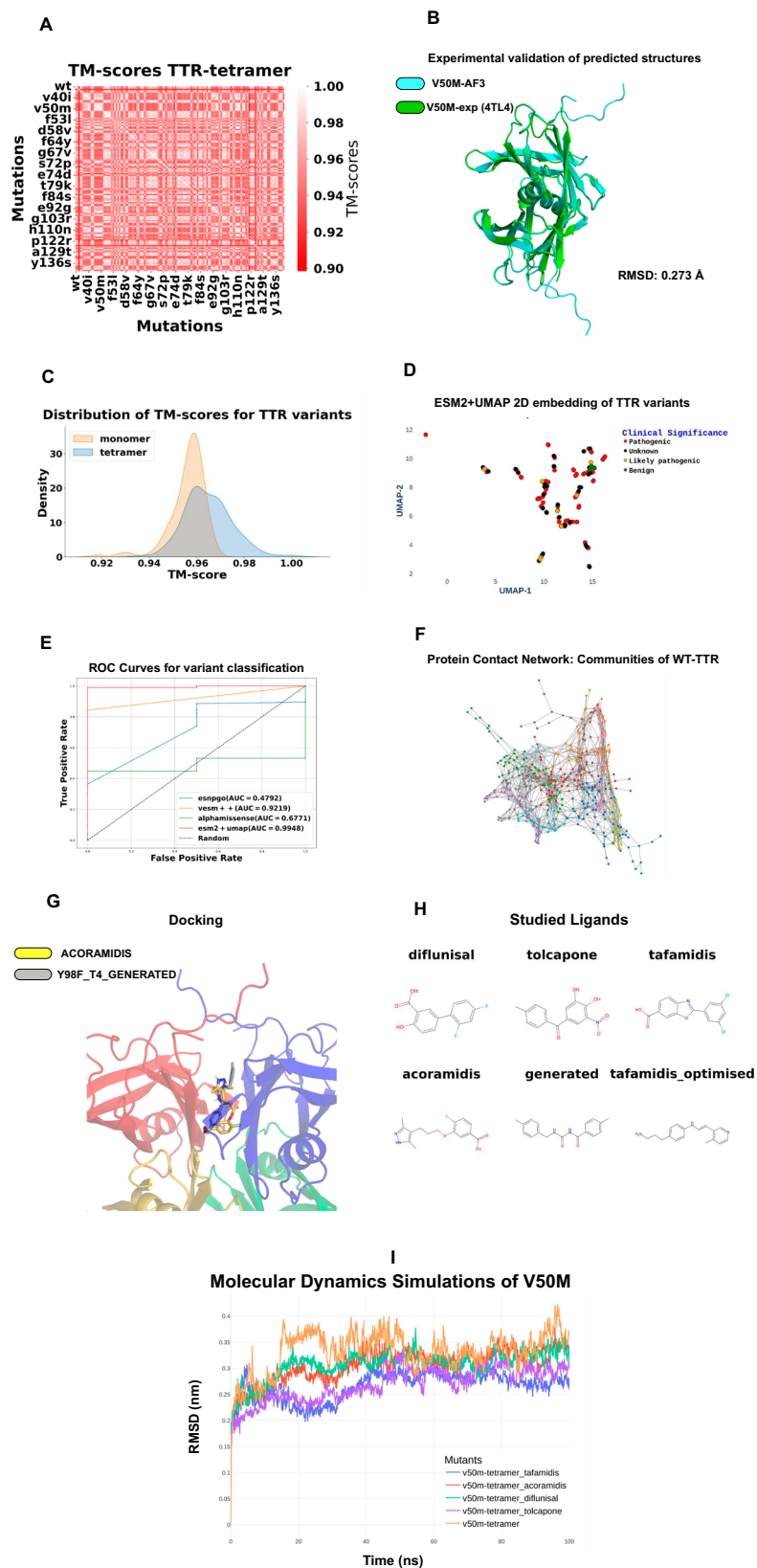
To further characterise the structural organisation of TTR, PCN-Miner has identified residue communities—i.e. densely connected subgraphs—that may correspond to cooperative structural or functional domains. The Leiden algorithm was used to partition the PCN into communities, revealing seven and ten discrete modules in the monomeric and tetrameric wild-type structures on Fig. 8). These communities provide a mesoscale representation of the protein, potentially linking mutation hotspots to collective structural dynamics.

Variant-aware molecular docking and ligand optimisation in transthyretin

Molecular coupling is a computational approach that is used to predict the binding orientation of a small molecule (ligand) to a target macromolecule, typically a protein. By estimating binding affinities and interaction modes, docking simulations offer valuable insights into the molecular basis of ligand-receptor recognition. In the context of disease-related protein variants, docking enables evaluation of how specific mutations can alter ligand binding profiles, thus informing drug efficacy, resistance mechanisms, and opportunities for drug repurposing or personalisation.

This work presents the use of two docking tools: DiffDock-L^{50,51} and AutoDock Vina⁵², to simulate interactions between approved or

Fig. 3 | Integrated computational framework for the structural and functional characterization of transthyretin (TTR) variants. A multi-level computational analysis has been performed by combining structural prediction, sequence-based representation, and dynamic simulations of TTR variants. **A** Structural alignment between the structures of variants. **B** Experimental Validation of the predicted structures. **C** Two-dimensional UMAP projection of TTR variants for classifying pathogenic and benign variants. **D** Distribution of TM-Scores for the Variants reveals that structure of the monomer are more conserved. **E** Receiver operating characteristic (ROC) curves illustrating the classification performance in distinguishing benign and pathogenic TTR variants. **F** PCN analysis identified changes at mesoscale in variants. **G** Docking pose of Acoramidis on a generatively designed Y98F variant, revealing predicted binding interactions from AI-based docking. **H** RMSD time evolution from molecular dynamics simulations of the V50M variant, supporting the structural stability of the predicted model. **I** Selected Ligands have been optimised. **J** Molecular Dynamics Simulation of the binding.



AI-generated ligands and TTR variants. DiffDock-L was used to predict 20 binding poses per variant-ligand pair, from which the pose with the highest confidence score was selected. AutoDock Vina was used to estimate binding affinities for each selected pose.

Docking with approved ligands

Tafamidis and Acoramidis are FDA-approved transthyretin stabilisers currently used in the clinical management of systemic amyloidosis¹⁴. Tafamidis⁵³ is a non-NSAID (Nonsteroidal anti-inflammatory drug)

benzoxazole compound that binds to the TTR tetramer, preventing its dissociation into monomers and subsequent amyloid formation. Approved in 2019⁵⁴, it is particularly effective in treating V50M-associated hereditary transthyretin amyloidosis.

Acoramidis⁵⁵, marketed as Attruby, gained FDA approval in 2024 for transthyretin-mediated cardiomyopathy. Designed to emulate the stabilising T139M mutation⁵⁶, Acoramidis achieves over 90% tetramer stabilisation throughout dosing intervals.

Figure 9 presents the highest confidence scores from DiffDock-L for TTR variant interactions with tafamidis and acoramidis, respectively. Two variants, A39D and G73R, exhibited consistently low confidence scores across both ligands, suggesting a common mutation-induced disruption or a docking algorithm limitation.

AutoDock Vina simulations further revealed mutation-specific differences in binding affinity. For each ligand-mutant pair, the lowest Vina score from ten poses was selected to represent optimal binding.

To ensure robustness, predicted poses were aligned with experimentally resolved ligand conformations. Alignments with the lowest RMSD were retained for further analysis using the PyMOL Align plugin.

Binding affinities relative to the V50M mutation for tafamidis and to tafamidis for acoramidis have been calculated as depicted in Fig. 10. Tafamidis showed diminished binding in several mutants, consistent with its

optimisation for V50M. In contrast, acoramidis showed broadly improved binding compared to tafamidis for all the reported mutations.

Ligand optimisation and de novo design

To improve binding against destabilising mutations, DiffSBDD⁵⁷ to optimise Tafamidis has been used. Among the ten variants with the lowest binding scores, Y98F was randomly selected for ligand redesign.

In the TTR tetramer, Y98F lies distal to the Tafamidis binding pocket (green). Despite its spatial separation, this mutation significantly reduced binding affinity, highlighting long-range structural effects⁵⁸. Optimisation via DiffSBDD improved binding to A129T and modestly benefited several others. These results indicate the need and effectiveness of a mutation-specific ligand design as reported in Fig. 11.

De novo ligand generation using the Y98F tetramer structure as reference has been employed. Residues 15, 17, 54, 106, 108, 109, 110, 117 and 119 were selected as binding pocket ref.⁵⁹. Binding affinities of the generated compound were evaluated across mutants.

As shown in Fig. 12, the generated ligand exhibited consistently strong affinities when compared with Acoramidis, showing mutation-dependent variability in docking outcomes.

A strong negative correlation (-0.4631) with statistical significance ($1.9850e^{-8}$) was found for acoramidis docking between confidence values given by DiffDock and Vina Scores predicted with AutoDock Vina. This strong, but negative, correlation may be caused by how vina and confidence scores are defined. Vina scores are always negative while confidences are always positive values between 0 and 1. This negative correlation actually indicates agreement between the two scoring methods, as lower (more negative) Vina scores represent stronger predicted binding affinities, while higher confidence values from DiffDock represent greater certainty in the predicted binding poses. No correlation found for tafamidis.

Table 1 | Validation of AlphaFold3-predicted mutant structures against experimental references

Structure type	Mean RMSD (Å)	No. of experimental structures
Tetramer	0.2662	21
Tetramer + Ligand	0.5083	12

Fig. 4 | Hierarchical clustering of TTR tetrameric variants based on TM-score similarity. Clinical classifications are colour-coded: red (Pathogenic), orange (Likely pathogenic), green (Benign), and blue (Unknown). Three main clusters are annotated (purple, black, and sea-green).

Hierarchical Clustering Dendrogram for TTR tetramer (Average Linkage)

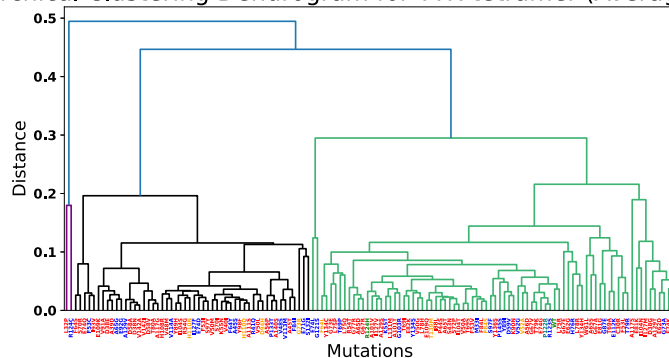
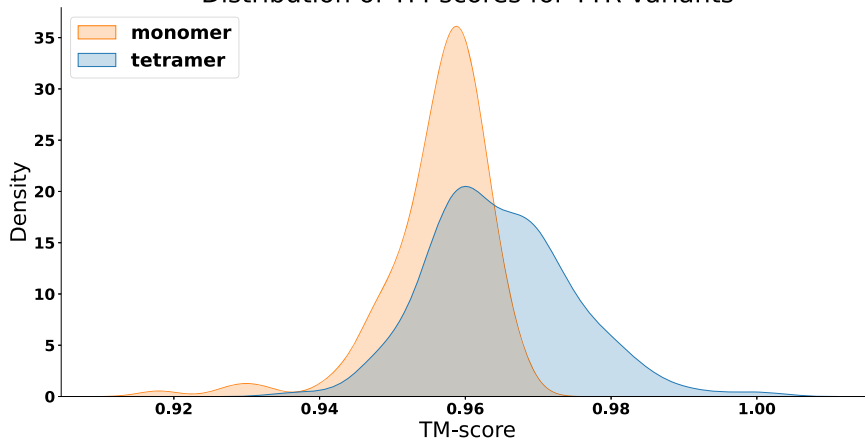


Fig. 5 | Density plot of TM-score distributions for monomeric and tetrameric TTR variants. Tetrameric structures display high structural conservation; monomeric forms show greater variability, potentially reflecting aggregation-prone states.

Distribution of TM-scores for TTR variants



UMAP projection of the ESM2 embeddings obtained for TTR variants

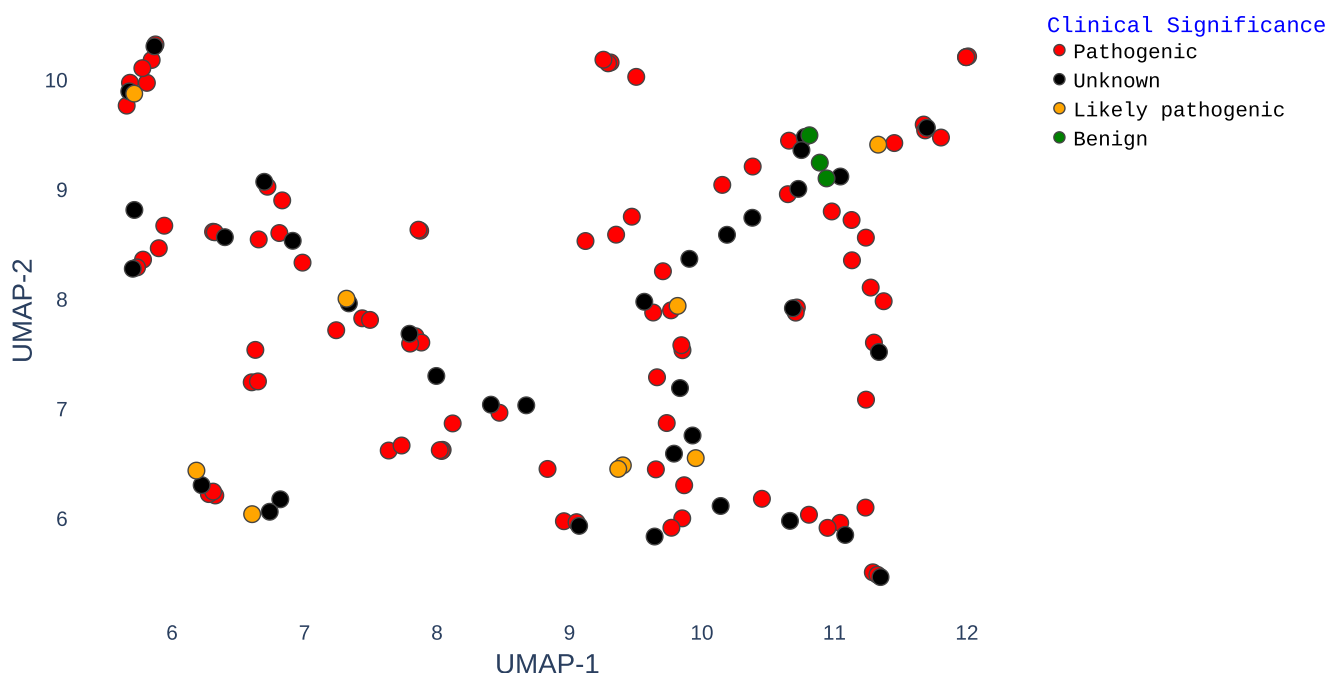


Fig. 6 | Two-dimensional UMAP projection of ESM2 embeddings for TTR variants. Each point represents a single variant, coloured by clinical annotation from UniProt (accession ID: P02766). Clusters indicate shared functional or structural properties.

Table 2 | Comparison of ROC AUC scores for classification of pathogenic vs. benign TTR variants

Method	ROC AUC
E-SNPs&GO	0.4792
AlphaMissense	0.6771
VESM++ ^a	0.9219
ESM2 + UMAP^b	0.9948

Thresholds: ^a0.2 for VESM++, ^b0.325 for ESM2+UMAP.

Values highlighted in bold represent best results.

An agreement assessment was also performed to study the alignment of DiffDock and AutoDock Vina in predicting the optimal existing ligand for each mutation. A strong agreement of 96% was found and quite always acoramidis is predicted by both methods as the optimal ligand in terms of docking.

Figure 13 depicts the 2D structures of all the ligands studied in this work, from existing ones (tafamidis, acoramidis, tolcapone and diflunisal) to the ones generated and optimised with DiffSBDD (Y98F_T4 and tafamidis_optimised_V142I).

For all the structures, a drug-likeness assessment was performed for validating the generated ligands as reported in Table 1. The optimisation of tafamidis leads to an higher QED value maintains compliance with major drug-likeness rules. The optimised tafamidis also reports enhanced membrane permeability with increased Blood-brain barrier permeability (BB Perm) and a reduction of the Topological Polar Surface Area (TPSA). Tolcapone, on the other hand, represents the classic example of the drug-likeness paradox: even if it fails multiple drug-likeness criteria (lowest QED score, poor permeability, high TPSA) it remains therapeutically valuable and effective. The ligand generated by DiffSBDD for the T4 binding site of Y98F-TTR is the best among the studied ligands in terms of theoretical proprieties for drug-likeness. Generated ligand passed the Lipinski's Rule of Five test and reports the higher QED score of 0.91, with excellent lipophilicity (LogP), good membrane permeability (BB. Perm).

Assessing mutation-induced destabilisation and ligand-mediated rescue

The V50M TTR tetramer displayed increased flexibility in regions corresponding to the signal peptides of each monomer and the thyroxine (T4) binding site as represented in Fig. 14. These local fluctuations coincide with a loss of tetrameric symmetry and may underlie an enhanced propensity for aggregation Fig. 15. Notably, all tested ligands restored local rigidity in the V50M variant, reducing root mean square fluctuations (RMSF) to levels comparable to the wild-type protein, consistent with their established stabilising effects on the TTR tetramer⁶⁰. Analysis of root mean square deviation (RMSD) and structural frame comparisons further indicated that ligand-bound V50M variants preserved compact conformations throughout simulations, while unliganded V50M exhibited more pronounced conformational drift, as depicted in Fig. 16).

Discussion

Transthyretin amyloidosis (ATTR) is a genetically heterogeneous disorder characterised by the accumulation of amyloid fibrils derived from destabilised transthyretin (TTR) protein variants. Despite the clinical availability of stabilising agents such as tafamidis and acoramidis, the results of this paper demonstrate that their efficacy is not uniform across all known single-point TTR mutations. This finding carries profound implications for both therapeutic intervention and the fundamental understanding of genotype-phenotype relationships in amyloid diseases.

By combining AlphaFold3-based structure prediction, transformer-based protein embeddings (ESM2), graph-based molecular docking (DiffDock-L), and classical affinity prediction (AutoDock Vina), the structural and functional consequences of TTR mutations have been analysed. The comprehensive docking analysis revealed that while acoramidis generally shows higher average binding affinity than tafamidis, several variants—including W61L and Y98F—exhibit marked resistance to tafamidis. This observation reinforces the notion that the effect of a drug cannot be universally extrapolated across all genotypes, even when the ligand binds to the same pocket in the same protein scaffold.

Table 3 | Molecular properties and drug-likeness metrics of ligand compounds

Property	Acoramidis	Diflunisal	Tafamidis	Tafamidis opt.	Tolcapone	Y98F_T4
Chiral	No	No	No	No	No	No
N chiral atoms	0	0	0	0	0	0
Specified	No	No	No	No	No	No
N rings	2	2	3	2	2	2
Aromatic rings	2	2	3	2	2	2
Aliphatic rings	0	0	0	0	0	0
LogP	2.88	3.04	4.50	4.14	3.07	3.72
TPSA	75.21	57.53	63.33	50.94	103.35	58.20
BB Perm.	-0.95	-0.37	0.22	0.38	-1.63	-0.04
HBA	5	3	3	3	3	4
HBD	2	2	1	2	2	2
MW	292.31	250.20	308.12	267.37	273.24	282.34
Ro3	Yes	Yes	Yes	Yes	Yes	Yes
Ro5	Yes	Yes	Yes	Yes	Yes	Yes
QED Score	0.80	0.86	0.76	0.84	0.39	0.91

Among the reported values are HBA (number of hydrogen bond acceptors), HBD (number of hydrogen bond donors) MW (molecular weight of the compound in daltons (g/mol)).

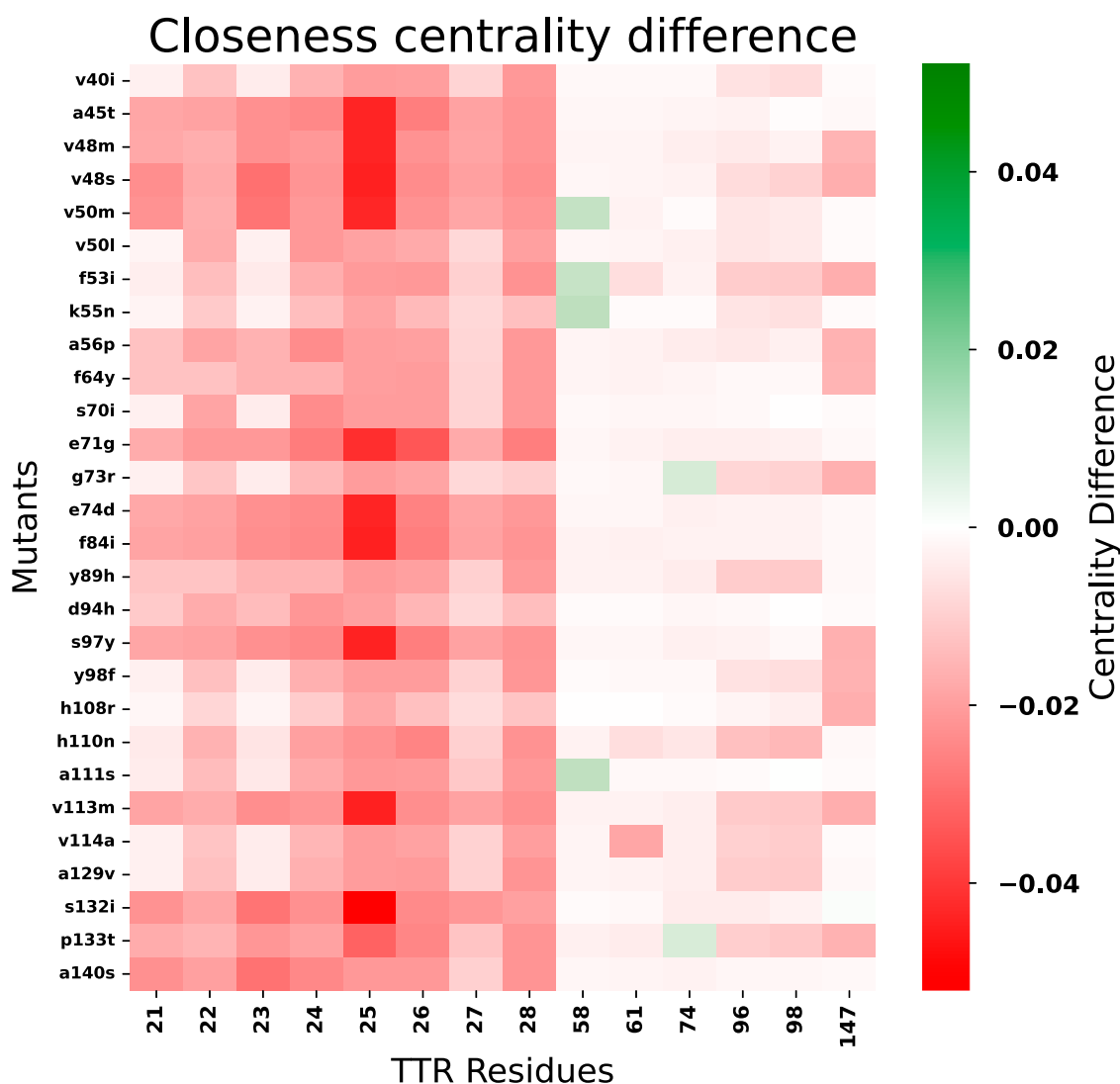


Fig. 7 | Differences in closeness centrality between wild-type and mutant TTR tetramers. Only one chain is displayed for clarity. Mutations shown correspond to those with at least five residues exhibiting altered centrality values.

Fig. 8 | Analysis of Residue Communities in TTR. Residue communities detected in monomeric (left) and tetrameric (right) wild-type TTR structures using the Leiden algorithm. Communities represent densely interconnected regions within the protein contact network.

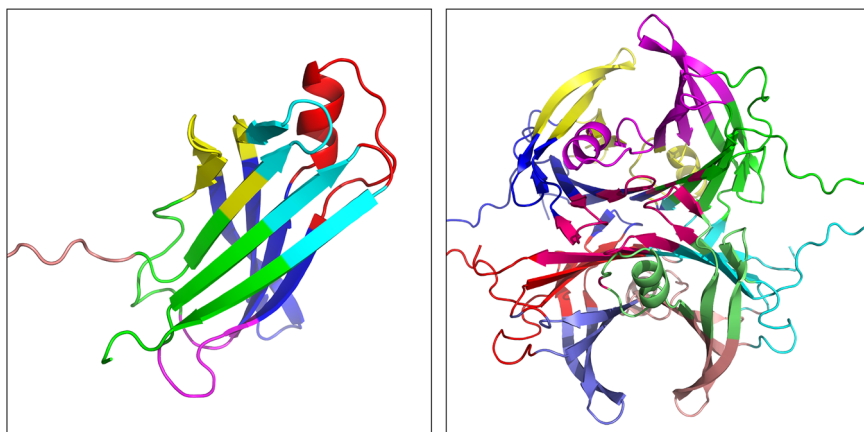


Fig. 9 | Confidence scores of docking between mutants and drugs. Highest-confidence docking poses predicted by DiffDock-L for TTR tetramer variants with acoramidis (top) and tafamidis (bottom). A39D and G73R consistently show low-confidence scores.

This work indicates that mutations exert their influence not only through direct disruption of ligand-contact residues but also via long-range structural perturbations. For example, the Y98F and W61L mutations, although distant from the tafamidis binding site, significantly impair binding affinity, likely by altering the conformational ensemble of the protein. This highlights the necessity of considering the full structural context of a mutation, rather than merely its spatial proximity to the binding site. Docking results, further corroborated by pose alignment and RMSD analysis, demonstrate that traditional assumptions regarding locality in protein-ligand interactions may fall short in

cases where dynamic allostery and subtle conformational shifts are involved.

Between the list of pathogenic mutations that show poor binding affinity with tafamidis, the most clinically prevalent V142I to perform the tafamidis optimisation with DiffSBDD has been selected. The use of generative models like DiffSBDD enabled the design of mutation-specific ligands, with the V142I-optimised analog of tafamidis showing improved binding not only to the target mutation but to a broader subset of destabilising variants. However, this optimisation was not universally beneficial, with some mutations displaying reduced binding affinity to the redesigned

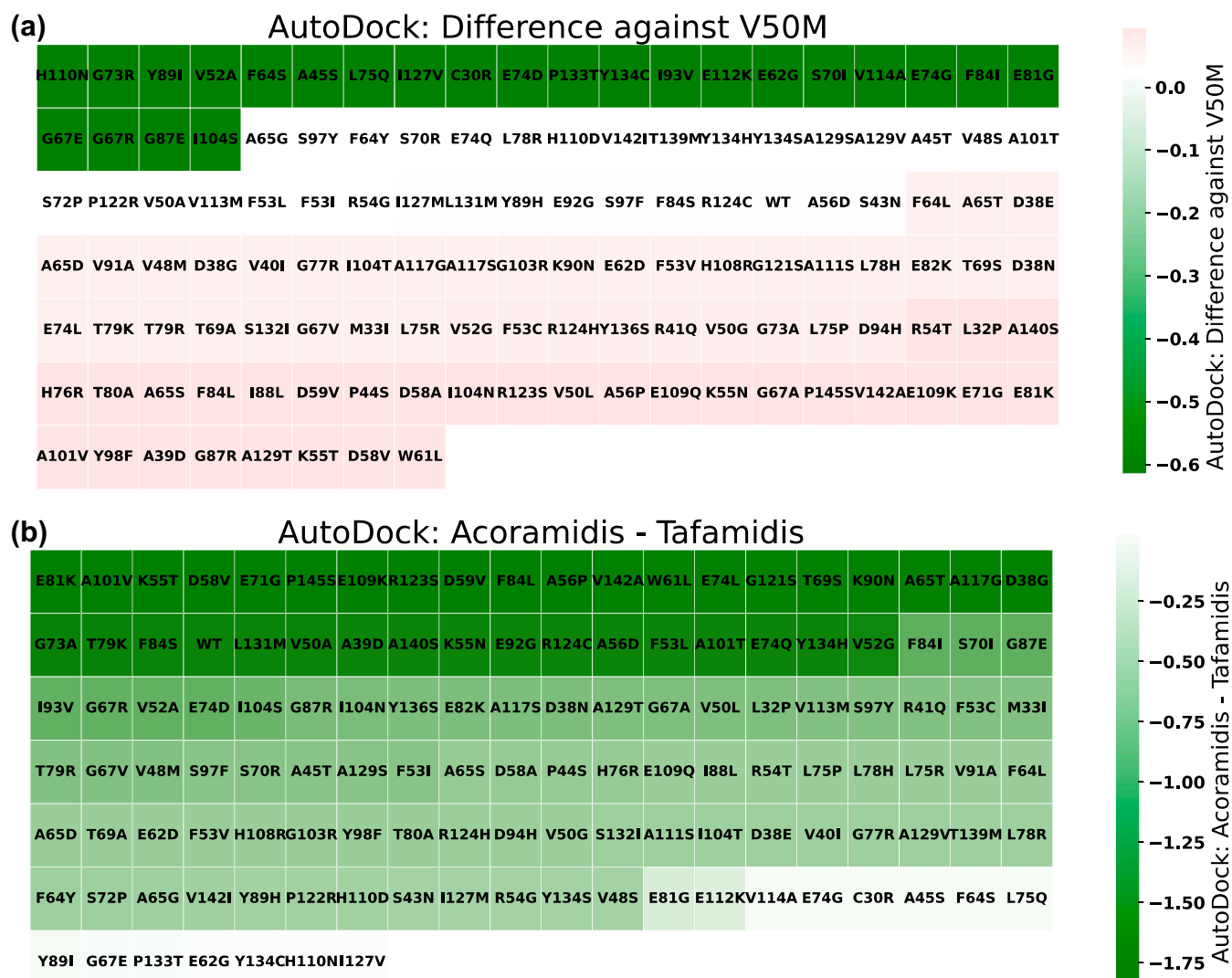


Fig. 10 | Binding Affinity Analysis. Binding affinity differences between TTR mutants and **a** tafamidis and **b** acoramidis. White cells depicts mutations that report approximately no changes in binding affinity, while red and green ones depict mutations with decreased or increased binding affinity.

ligand. These findings suggest that while generative ligand design holds promise for customizing therapies, it may require careful constraint or multi-objective optimisation to avoid adverse trade-offs across the mutational landscape.

Taken together, the results strongly advocate for a synergistic and variant-resolved strategy in the development of TTR-directed therapeutics. Relying on a single stabiliser, even one optimised for a high-prevalence mutation like V50M, may be insufficient in the context of diverse genotypic presentations. A shift toward precision medicine in ATTR is thus both biologically warranted and technologically feasible. Tools such as ESM2 embeddings and UMAP clustering, which has been used to identify benign-like behaviour in uncharacterised mutations, offer a scalable pipeline for preclinical triage and drug-response prediction.

Furthermore, this work opens the possibility for rational repurposing of ligands based on structural similarity in latent embedding space. For example, mutations clustering near known benign variants may benefit from similar therapeutic profiles, while outliers in embedding or docking space could be prioritized for custom ligand design or combinatorial therapy.

Future work should aim to extend these findings beyond computational frameworks. In vitro validation of predicted affinities and conformational shifts will be critical, particularly for variants of unknown significance (VUS). Longitudinal studies tracking clinical outcomes across

genotypes treated with the same stabiliser could provide real-world evidence supporting the in silico predictions. Additionally, the integration of patient-specific omics data may further refine the variant-ligand interaction landscape, ultimately enabling personalised treatment plans.

In conclusion, this study demonstrates the necessity and promise of integrating structural modelling, AI-driven prediction, and generative design in addressing the unmet need for precision therapeutics in TTR amyloidosis. By acknowledging the structural and functional heterogeneity induced by point mutations, this paper suggests a scenario in which each patient’s genotype informs the most effective and targeted treatment strategy.

This study explored the hypothesis that an integrative, structure-based framework—comprising conformational modelling, molecular docking, and generative design—could help prioritize stabilising ligands specifically tailored to various transthyretin (TTR) variants. The results support this hypothesis across multiple areas.

It has been demonstrated that pathogenic TTR mutations lead to variant-specific conformational changes at the binding site. These structural alterations result in measurable differences in binding affinities for known stabilisers, including tafamidis and acoramidis. These observations are consistent with the idea that the varying response to ligands is due to mutation-induced structural heterogeneity, a central premise of this framework.

AutoDock: Optimised - Original Tafamidis

A129T	I104T	G121S	Y136S	E109K	A129S	V52G	I88L	E82K	R54T	D59V	R124C	E62D	G87R	A65S	R123S	D58A	S43N	K55T	A65T
S132I	F84L	G73A	E71G	Y98F	P44S	V91A	F53C	E81K	K90N	Y134S	V50A	WT	W61L	H76R	K55N	V142A	Y89H	P122R	G103R
A111S	R41Q	P145S	A56P	H110D	D58V	D38E	I127M	G67A	A117G	L32P	A129V	A65D	I104N	F53I	E109Q	D38N	L131M	G67V	D94H
S97F	A117S	L75P	Y134H	E92G	F53L	A56D	G77R	V113M	V50L	D38G	R54G	F53V	L75R	S70R	T69S	R124H	M33I	F84S	T79R
S97Y	F64L	A101V	A101T	T80A	L78H	H108R	V48S	T69A	T79K	A45T	L78R	F64Y	S72P	V40I	A65G	V142I	T139M	V50G	A140S
V48M	E74L	F84I	F64S	E81G	P133T	C30R	E62G	L75Q	E112K	Y134C	Y89I	H110N	I104S	G87E	V52A	G67R	I93V	A39D	S70I
A45S	V114A	E74D	G67E	I127V	E74G														

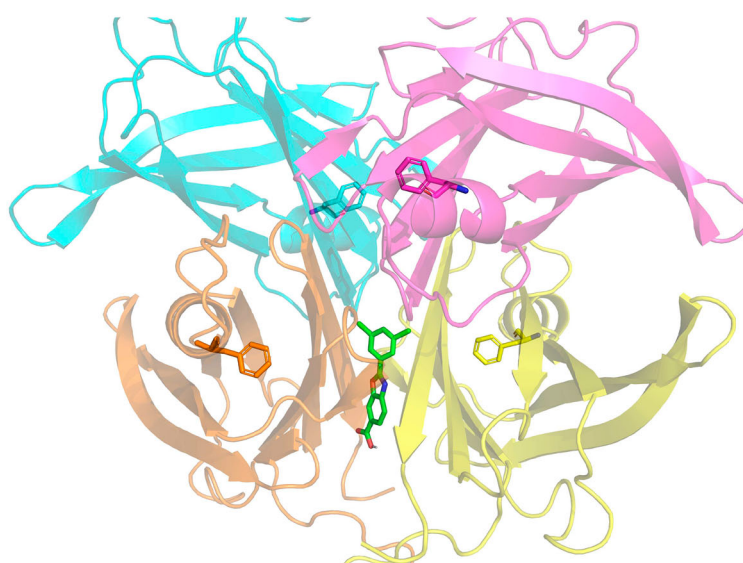
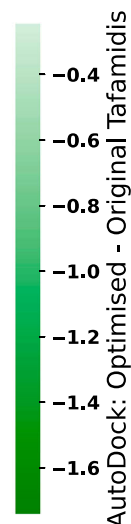


Fig. 11 | Ligand optimisation against the Y98F mutation using DiffSBDD. White cells depicts mutations that report approximately no changes in binding affinity, while red and green ones depict mutations with decreased or increased binding affinity.

AutoDock: Generated - Acoramidis

D38E	G103R	E109Q	Y134S	A111S	Y98F	Y134C	H110N	S132I	V91A	S43N	H110D	F64S	H76R	G77R	P133T	E62D	F53V	V114A	L75Q
Y89I	I88L	V40I	R54T	A65G	A65D	P44S	L78R	P122R	S72P	A65S	I127M	A45S	L75R	T139M	D94H	V50G	F53I	E62G	F64L
A129S	A129V	G67E	Y89H	E81G	I127V	T69A	D58A	V48S	T80A	D38N	Y136S	C30R	R124H	R54G	E82K	I104T	F64Y	L78H	A117S
S97Y	M33I	G87R	T79R	L32P	F53C	A45T	R41Q	S97F	E112K	V48M	V50L	G67V	H108R	V113M	V142I	A129T	G67A	L75P	I104N
S70R	E74G	V52G	T69P	E74Q	V142A	I127F	G121S	E71G	G73A	A140S	WT	A56D	V52A	E109K	F84I	K90N	Y134H	E92G	A117G
D38G	I104S	I93V	F84S	L131M	A101T	S70I	V50A	R124C	G67R	K55N	G87E	F84L	A65T	D59V	T69S	R123S	E81K	A101V	F53L
A56P	K55T	P145S	W61L	T79K	D58V	E74D	A39D	T69I	E74L	E74K	G73E								

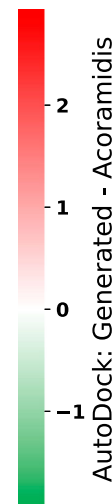


Fig. 12 | Evaluation of de novo generated ligand targeting TTR mutants. White cells depicts mutations that report approximately no changes in binding affinity, while red and green ones depict mutations with decreased or increased binding affinity.

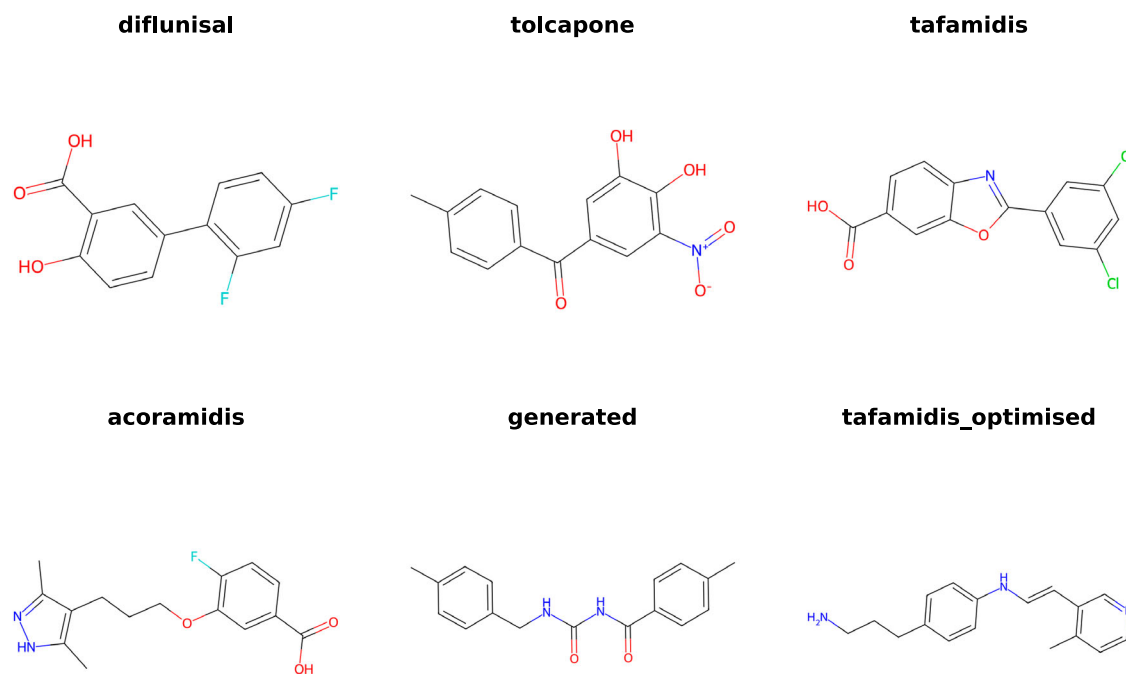


Fig. 13 | Two dimensional structure of ligands. 2D structure of existing ligands (tafamidis, acoramidis, diflunisal, tolcapone), generated ligands (Y98F_T4) and optimised ones (tafamidis_optimised_V142I).

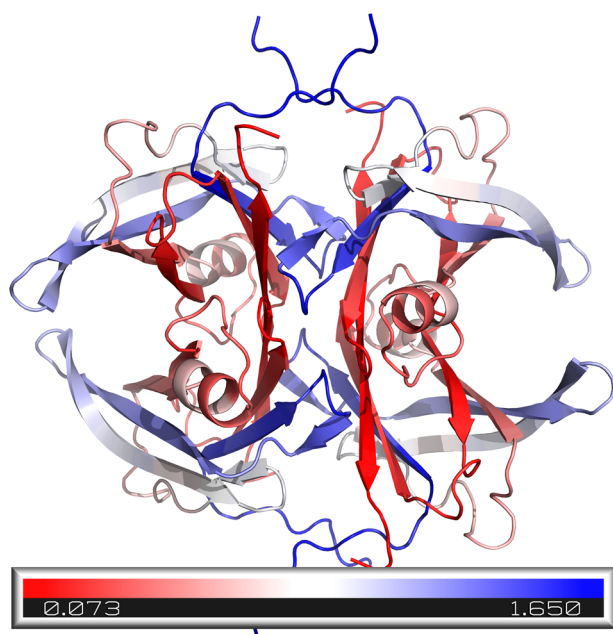


Fig. 14 | Impact of TTR point mutations and ligand binding. RMSF comparison across wild-type and mutant TTR variants. V50M shows increased flexibility in chain C and D.

Building on these insights, a generative design approach to propose novel ligand optimised for binding to representative TTR variants has been applied. The top-scoring candidates, ranked through docking-based affinity predictions, outperformed existing stabilisers in several variant contexts. These findings reinforce the concept that data-driven molecular design can utilise subtle structural variations to develop genotype-specific stabilisers with improved binding profiles.

Importantly, the approach of this paper is fully extensible. Though this initial study was limited to two reference stabilisers and a set of structurally related analogues, the generative pipeline is compatible with various

chemical libraries and alternative TTR conformations. This flexibility allows for broader applications in mutation-guided drug design.

From a clinical standpoint, this article underscores the potential of precision pharmacology in managing TTR amyloidosis. Instead of pursuing a universal stabiliser, the results support the development of personalised ligands tailored to a patient's genetic background. This shift aligns with broader trends in genotype-informed therapy, especially for disorders related to protein misfolding.

However, the study does have limitations. All results are based on in silico predictions; experimental validation of binding and stabilising effects is a vital next step. Additionally, while docking scores provide a useful approximation of affinity, further evaluation of pharmacokinetic and toxicity properties will be necessary for lead optimisation.

In summary, this work provides evidence that mutation-aware ligand generation—grounded in structural modelling and generative chemistry—can inform the design of next-generation TTR stabilisers and serve as a framework for addressing similar amyloid-related diseases.

This approach could be readily applied to other protein misfolding disorders, such as amyotrophic lateral sclerosis (ALS), where mutations in *SOD1* disrupt metal binding and promote aggregation. By modelling *SOD1* variants and optimising ligands to stabilise the native structure, targeted therapeutic strategies could be developed. Similarly, in Alzheimer's disease, the misfolding of tau protein, despite its intrinsic disorder, presents aggregation-prone regions that can be studied using structural embeddings and docking simulations to identify isoform- or modification-specific stabilisers. In systemic light chain (AL) amyloidosis, where immunoglobulin light chains exhibit extensive sequence variability, the framework can be utilised to classify and predict aggregation-prone variants, thereby supporting the design of therapeutic inhibitors. Moreover, in Parkinson's disease and related synucleinopathies, latent space analysis and docking to emergent pockets of α -synuclein variants could inform the development of conformation-specific binders or modulators. Overall, the methodology supports a precision medicine paradigm, where ligand design and therapeutic strategies are tailored to the specific structural and mutational profile of each protein variant. Future work will aim to generalise this platform to additional disease systems, incorporating experimental structures (e.g., cryo-EM), mutagenesis data, and proteostasis models to enhance its clinical and translational utility.

Fig. 15 | Impact of TTR point mutations and ligand binding. RMSD of V50M in apo form and bound to tafamidis or acoramidis, showing rescue of wild-type-like flexibility.

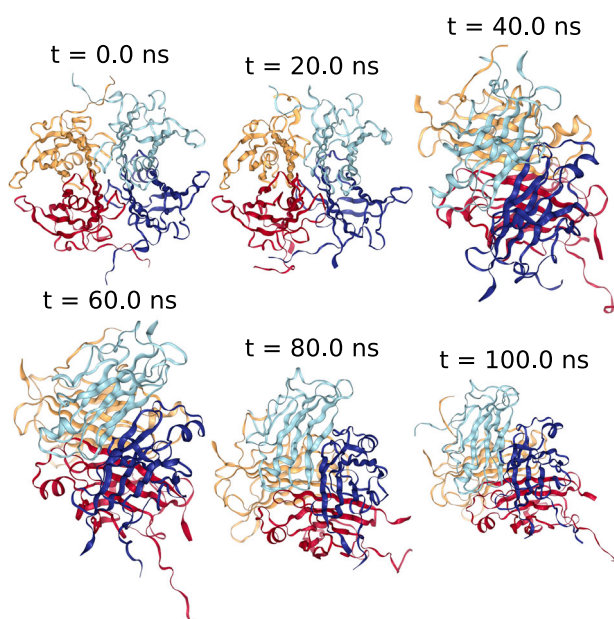
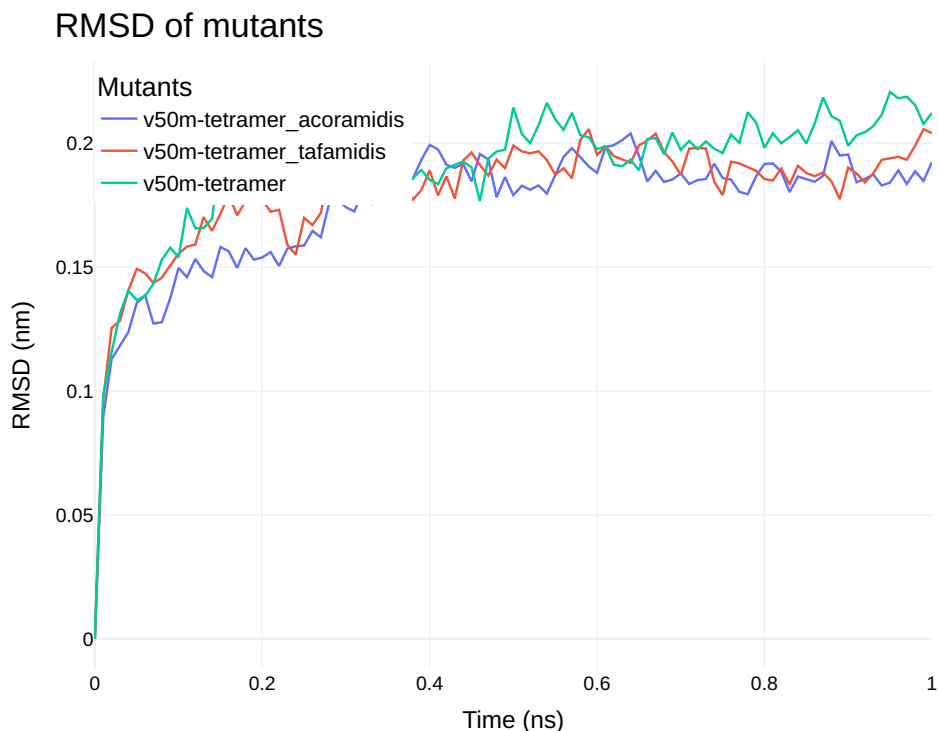


Fig. 16 | Impact of TTR point mutations and ligand binding. Representative frames from the V50M trajectory illustrating progressive structural changes over time.

Methods

Overview of the computational workflow

An overview of the computational pipeline is shown in Fig. 17. Beginning with the wild-type (WT) transthyretin (TTR) sequence and a curated set of single-point mutations from the literature, three-dimensional structural models were generated using AlphaFold3. These models were subsequently analysed using a suite of computational techniques, including molecular dynamics (MD) simulations, network-based metrics, and protein language model embeddings. To assess mutation-specific effects on ligand

interaction, both classical and AI-based molecular docking approaches were applied. Generative models were further used to explore mutation-aware ligand design.

Structure prediction

All single-point variants were modelled using AlphaFold3⁶¹, producing structures in both monomeric and tetrameric forms. The predicted wild-type structure was validated against crystallographic references for consistency. Structural alignments were performed using TM-align⁶², and structural similarity was quantified using the TM-score:

$$\text{TM - score} = \max \left(\frac{1}{L_T} \sum_{i=1}^{L_A} \frac{1}{1 + \left(\frac{d_i}{d_0} \right)^2} \right), \quad (2)$$

where L_A and L_T denote the lengths of the aligned and target proteins, respectively; d_i is the distance between aligned residues; and $d_0 = 1.24\sqrt{[3]L_T - 15 - 1.8}$. TM-scores above 0.5 indicate high structural similarity.

AlphaFold3-predicted models were aligned and validated via root mean square deviation (RMSD) calculations for the available experimentally resolved mutant structures. Validation was carried out on both unbound tetramers and ligand-bound complexes.

Latent space analysis using ESM2 and UMAP

Variant sequences were embedded using the ESM2 transformer model⁴⁰. Each sequence was mapped to a 1280-dimensional vector and projected into two dimensions using Uniform Manifold Approximation and Projection (UMAP)⁴¹. Euclidean distances from the WT embedding were used to classify pathogenicity: variants within a predefined threshold radius were labelled benign, and those beyond as pathogenic. Due to the imbalance between pathogenic annotations ($n = 96$) and benign annotations ($n = 2$), classification performance was assessed using the ROC-AUC metric. The ground truth was obtained through the UniProt Variation API,

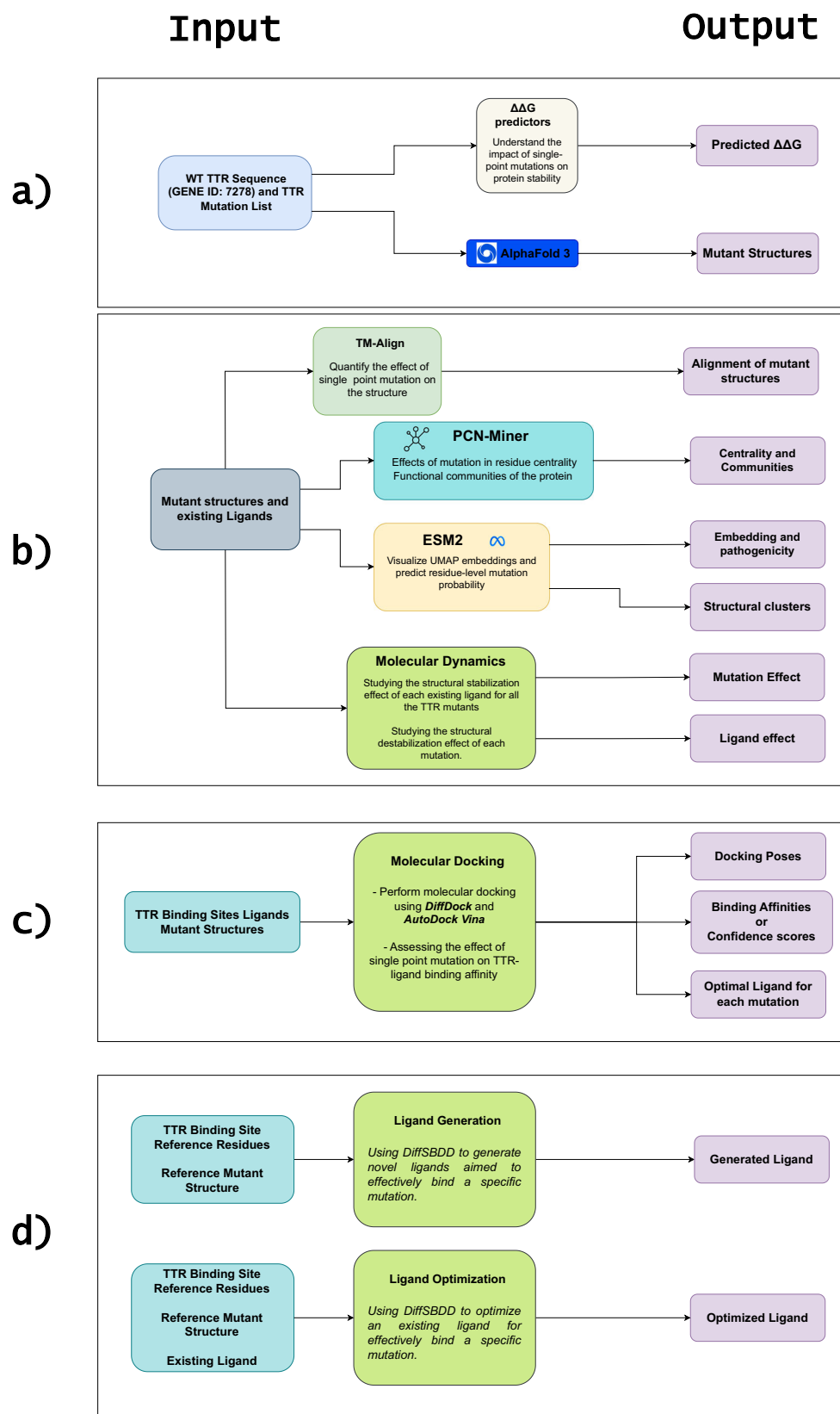


Fig. 17 | Overview of the computational pipeline. **a** Mutant structures are generated using AlphaFold 3 and the impact of point mutations on protein stability is predicted using $\Delta\Delta G$ predictors. **b** The structural effects of these mutations are analysed alongside existing ligands. This involves quantifying structural changes with TM-Align, investigating residue centrality and functional communities with PCN-Miner, visualizing embeddings and predicting pathogenicity with ESM2, and

investigating the structural stabilization/destabilization effects of both mutations and ligands using molecular dynamics simulations. **c** Molecular docking was performed on the TTR binding sites of the mutant structures using DiffDock and AutoDock Vina, evaluating binding affinities and identifying optimal ligands for each mutation. **d** DiffSBDD was used to generate new ligands or optimize existing ones to effectively bind specific mutations.

consolidating 'likely' labels into binary annotations. Variants of uncertain significance were classified using the ROC-optimised threshold.

Benchmarking against state-of-the-art predictors

Performance was benchmarked against three established predictors: AlphaMissense^{46,63}, E-SNPs&GO⁴⁷, and VESM++^{48,64}. AlphaMissense adapts AlphaFold2 to generate pathogenicity scores in the [0, 1] range, mapped to three confidence classes. E-SNPs&GO employs embeddings from ESM-1v⁶⁵ and ProtTrans T5⁶⁶, reduced via PCA and classified using a support vector machine. VESM++ is a co-distilled ensemble of ESM-1b, ESM2-650M, and ESM3, outputting log-likelihood ratios transformed into pathogenicity scores using a sigmoid function. For fair comparison, all models used the same classification threshold.

Prediction of mutation-induced structural changes

Mutation-induced effects on protein stability were predicted using a consensus of sequence- and structure-based tools: mCSM⁶⁷, SDM⁶⁸, DUET⁶⁹, DynaMut2⁷⁰, DDGun⁷¹, and SAAFEC⁷². These tools include energy-based statistical models, machine learning frameworks, and purely sequence-driven approaches. Predicted changes in Gibbs free energy ($\Delta\Delta G$) were combined into a consensus matrix.

Network analysis of protein contact maps

For each TTR structure a Protein Contact Network (PCN)⁷³ was built. Each residue of the protein is represented as a node while edge models residues having distances within the 4–8 Å range^{74,75}. The Euclidean distance between residues i and j was defined as:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}. \quad (3)$$

Centrality measures—degree, closeness, and betweenness—were computed. Closeness was defined as:

$$C_{\text{closeness}}(i) = \frac{n - 1}{\sum_{j \neq i} d(i, j)}, \quad (4)$$

and betweenness as:

$$C_{\text{betweenness}}(i) = \sum_{j \neq k \neq i} \frac{\sigma_{jk}(i)}{\sigma_{jk}}, \quad (5)$$

where σ_{jk} is the number of shortest paths between nodes j and k , and $\sigma_{jk}(i)$ denotes those passing through node i .

Data availability

All data, scripts, and computational workflows used in this study are openly available at: https://github.com/UgoLomoio/ttr_mutations.

Code availability

Code is available at: https://github.com/UgoLomoio/ttr_mutations.

Received: 28 May 2025; Accepted: 24 August 2025;

Published online: 25 September 2025

References

- Benson, M. D. Pathogenesis of transthyretin amyloidosis. *Amyloid* **19**, 14–15 (2012).
- Liz, M. A. et al. A narrative review of the role of transthyretin in health and disease. *Neurol. Ther.* **9**, 395–402 (2020).
- Ferreira, J. A. et al. Correction: structure-based analysis of a19d, a variant of transthyretin involved in familial amyloid cardiomyopathy. *Plos One* **8**, 10–1371 (2013).
- Saponaro, F. et al. Transthyretin stabilization: an emerging strategy for the treatment of alzheimer's disease? *Int. J. Mol. Sci.* **21**, 8672 (2020).
- Costa, J. et al. Transthyretin binding to a-beta peptide - impact on a-beta fibrillogenesis and toxicity. *FEBS Lett.* **582**, 936–942 (2008).
- Maetzler, W. et al. Serum and cerebrospinal fluid levels of transthyretin in lewy body disorders with and without dementia. *Plos One* **7**, e48042 (2012).
- Fleming, T. H. et al. Transthyretin internalization by sensory neurons is megalin mediated and necessary for its neurotogenic activity. *J. Neurosci.* **29**, 3220–3232 (2009).
- Hiram Guzzi, P., Petrizzelli, F. & Mazza, T. Disease spreading modeling and analysis: a survey. *Brief. Bioinforma.* **23**, bbac230 (2022).
- Reixach, N. et al. Tissue damage in the amyloidoses: transthyretin monomers and nonnative oligomers are the major cytotoxic species in tissue culture. *Proc. Natl Acad. Sci.* **101**, 2817–2822 (2004).
- Yee, A. J. et al. A molecular mechanism for transthyretin amyloidogenesis. *Nat. Commun.* **10**, 925 (2019).
- Hammarström, P. The transthyretin protein and amyloidosis—an extraordinary chemical biology platform. *Isr. J. Chem.* **64**, e202300164 (2024).
- Iacocca, M. A. et al. Clinvar database of global familial hypercholesterolemia-associated dna variants. *Hum. Mutat.* **39**, 1631–1640 (2018).
- Stenson, P. D. et al. Human gene mutation database (HGMD®): 2003 update. *Hum. Mutat.* **21**, 577–581 (2003).
- Almeida, Z. L., Vaz, D. C. & Brito, R. M. Transthyretin mutagenesis: impact on amyloidogenesis and disease. *Crit. Rev. Clin. Lab. Sci.* **61**, 1–25 (2024).
- Serpell, L. C., Goldsteins, G., Dacklin, I., Lundgren, E. & Blake, C. C. The "edge strand" hypothesis: prediction and test of a mutational "hot-spot" on the transthyretin molecule associated with fap amyloidogenesis. *Amyloid* **3**, 75–85 (1996).
- Shah, A. Misfolded transthyretin as a novel risk factor for heart failure. *JAMA Cardiol.* **6**, 255–257 (2021).
- Sekijima, Y. Transthyretin (ATTR) amyloidosis: clinical spectrum, molecular pathogenesis and disease-modifying treatments. *J. Neurol. Neurosurg. Psychiatry* **86**, 1036–1043 (2015).
- Manganelli, F. et al. Hereditary transthyretin amyloidosis overview. *Neurological Sci.* **43**, 595–604 (2020).
- Luigetti, M. et al. Gastrointestinal manifestations in hereditary transthyretin amyloidosis: a single-centre experience. *J. Gastrointest. Liver Dis.* **29**, 339–343 (2020).
- Mohan, C. et al. Suspecting and diagnosing transthyretin amyloid cardiomyopathy (ATTR-CM) in india: an indian expert consensus. *Indian Heart J.* **74**, 441–449 (2022).
- Delgado, D. et al. Epidemiology of transthyretin (ATTR) amyloidosis: a systematic literature review. *Orphanet J. Rare Dis.* **20**, 29 (2025).
- Rubin, L. et al. Cardiac amyloidosis: overlooked, underappreciated, and treatable. *Annu. Rev. Med.* **71**, 203–219 (2020).
- Christoffersen, C. et al. Transthyretin tetramer destabilization and increased mortality in the general population. *JAMA Cardiol.* **10**, 155–163 (2024).
- Skrahina, A. et al. Hereditary transthyretin-related amyloidosis is frequent in polyneuropathy and cardiomyopathy of no obvious aetiology. *Ann. Med.* **53**, 1787–1796 (2021).
- He, Y. et al. Association between serum transthyretin and intracranial atherosclerosis in patients with acute ischemic stroke. *Front. Neurol.* **13**, 944413 (2022).
- Tong, Z. et al. Aggregated transthyretin is specifically packaged into placental nano-vesicles in preeclampsia. *Sci. Rep.* **7**, 6694 (2017).
- Kalkunte, S. et al. Transthyretin is dysregulated in preeclampsia, and its native form prevents the onset of disease in a preclinical mouse model. *Am. J. Pathol.* **183**, 1425–1436 (2013).

28. Monu, S. et al. Plasma proteome profiling of coronary artery disease patients: downregulation of transthyretin—an important event. *Mediators Inflamm.* **2020**, 3429541 (2020).
29. Marcoux, J. et al. A novel mechano-enzymatic cleavage mechanism underlies transthyretin amyloidogenesis. *EMBO Mol. Med.* **7**, 1337–1349 (2015).
30. Ruberg, F. L. & Berk, J. L. Transthyretin (TTR) cardiac amyloidosis. *Circulation* **126**, 1286–1300 (2012).
31. Greve, A. M. et al. Association of low plasma transthyretin concentration with risk of heart failure in the general population. *JAMA Cardiol.* **6**, 258–266 (2021).
32. Nichols, K. et al. 1-hour versus 3-hour 99mTc-PYP imaging to evaluate suspected cardiac transthyretin amyloidosis. *Medicine* **102**, e33817 (2023).
33. Koike, H. & Katsuno, M. Expanding the spectrum of transthyretin amyloidosis. *Muscle Nerve* **61**, 3–4 (2019).
34. Zitnik, M. AI-enabled drug discovery reaches clinical milestone: machine learning. *Nat. Med.* **31**, 2490–2491 (2025).
35. Giadone, A. et al. A library of ATTR amyloidosis patient-specific induced pluripotent stem cells for disease modelling and in vitro testing of novel therapeutics. *Amyloid* **25**, 148–155 (2018).
36. Nativi-Nicolau, J. et al. Temporal trends of wild-type transthyretin amyloid cardiomyopathy in the transthyretin amyloidosis outcomes survey. *JACC Cardiooncol.* **3**, 537–546 (2021).
37. Luigetti, M. et al. Serum inflammatory profile in hereditary transthyretin amyloidosis: mechanisms and possible therapeutic implications. *Brain Sci.* **12**, 1708 (2022).
38. Carroll, J. et al. Novel approaches to diagnosis and management of hereditary transthyretin amyloidosis. *J. Neurol. Neurosurg. Psychiatry* **93**, 668–678 (2022).
39. Müllner, D. Modern hierarchical, agglomerative clustering algorithms. Preprint at <https://arxiv.org/abs/1109.2378> (2011).
40. Lin, Z. et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123–1130 (2023).
41. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).
42. Rowczenio, D. M. et al. Online registry for mutations in hereditary amyloidosis including nomenclature recommendations. *Hum. Mutat.* **35**, E2403–E2412 (2014).
43. Almeida, M., Alves, I., Terazaki, H., Ando, Y. & Saraiva, M. Comparative studies of two transthyretin variants with protective effects on familial amyloidotic polyneuropathy: TTR R104H AND TTR T119M. *Biochem. Biophys. Res. Commun.* **270**, 1024–1028 (2000).
44. do Amaral Martins, L. et al. Structural and thermodynamic characterization of a highly amyloidogenic dimer of transthyretin involved in a severe cardiomyopathy. *J. Biol. Chem.* **300**, 107495 (2024).
45. Mazzeo, A. et al. Transthyretin-related familial amyloid polyneuropathy (TTR-FAP): a single-center experience in sicily, an italian endemic area. *J. Neuromuscul. Dis.* **2**, S39–S48 (2015).
46. Cheng, J. et al. Accurate proteome-wide missense variant effect prediction with alphamissense. *Science* **381**, eadg7492 (2023).
47. Manfredi, M., Savojardo, C., Martelli, P. L. & Casadio, R. E-SNPs&GO: embedding of protein sequence and function improves the annotation of human pathogenic variants. *Bioinformatics* **38**, 5168–5174 (2022).
48. Brandes, N., Goldman, G., Wang, C., Ye, C. & Ntranos, V. Genome-wide prediction of disease variant effects with a deep protein language model. *Nat. Genet.* **55**, 1–11 (2023).
49. Zitnik, M. et al. Current and future directions in network biology. *Bioinform Adv.* **4**, vbae099 (2024).
50. Corso, G., Stärk, H., Jing, B., Barzilay, R. & Jaakkola, T. Diffdock: Diffusion steps, twists, and turns for molecular docking. In *International Conference on Learning Representations (ICLR)* (ICLR, 2023).
51. Corso, G., Deng, A., Polizzi, N., Barzilay, R. & Jaakkola, T. Deep confident steps to new pockets: Strategies for docking generalization. In *International Conference on Learning Representations (ICLR)* (ICLR, 2024).
52. Trott, O. & Olson, A. J. Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
53. Coelho, T. et al. Mechanism of action and clinical application of tafamidis in hereditary transthyretin amyloidosis. *Neurol. Ther.* **5**, 1–25 (2016).
54. FDA. FDA approves new treatments for heart disease caused by a serious rare disease, transthyretin mediated amyloidosis, <https://www.fda.gov/news-events/press-announcements/fda-approves-new-treatments-heart-disease-caused-serious-rare-disease-transthyretin-mediated> (2019).
55. FDA. FDA approves drug for heart disorder caused by transthyretin-mediated amyloidosis, <https://www.fda.gov/news-events/press-announcements/fda-approves-drug-heart-disorder-caused-transthyretin-mediated-amyloidosis> (2024).
56. Gillmore, J. D. et al. Efficacy and safety of acoramidis in transthyretin amyloid cardiomyopathy. *N. Engl. J. Med.* **390**, 132–142 (2024).
57. Schneuing, A. et al. Structure-based drug design with equivariant diffusion models. *Nat. Comput. Sci.* **4**, 899–909 (2024).
58. Tolentino-Lopez, L. et al. Outside-binding site mutations modify the active site's shapes in neuraminidase from influenza a h1n1. *Biopolymers* **99**, 10–21 (2013).
59. Palaninathan, S. et al. Novel transthyretin amyloid fibril formation inhibitors: synthesis, biological evaluation, and x-ray structural analysis. *PLoS One* **4**, e6290 (2009).
60. Morris, K. F., Geoghegan, R. M., Palmer, E. E., George, M. & Fang, Y. Molecular dynamics simulation study of ag10 and tafamidis binding to the val122ile transthyretin variant. *Biochem. Biophys. Res. Rep.* **21**, 100721 (2020).
61. Desai, D. et al. Review of AlphaFold 3: transformative advances in drug design and therapeutics. *Cureus* **16**, e63646 (2024).
62. Zhang, Y. & Skolnick, J. Tm-align: a protein structure alignment algorithm based on the tm-score. *Nucleic Acids Res.* **33**, 2302–2309 (2005).
63. Tordai, H. et al. Analysis of alphamissense data in different protein groups and structural context. *Sci. Data* **11**, 495 (2024).
64. ntranoslab. VESM: co-distillation of esm models for variant effect prediction, <https://huggingface.co/spaces/ntranoslab/vesm-variants> (2025).
65. Rives, A. et al. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proc. Natl Acad. Sci.* **118**, e2016239118 (2021).
66. Elnaggar, A. et al. Prottrans: toward understanding the language of life through self-supervised learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 7112–7127 (2022).
67. Pires, D. E., Ascher, D. B. & Blundell, T. L. MCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* **30**, 335–342 (2014).
68. Pandurangan, A. P., Ochoa-Montano, B., Ascher, D. B. & Blundell, T. L. SDM: a server for predicting effects of mutations on protein stability. *Nucleic acids Res.* **45**, W229–W235 (2017).
69. Pires, D. E., Ascher, D. B. & Blundell, T. L. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res.* **42**, W314–W319 (2014).
70. Rodrigues, C. H., Pires, D. E. & Ascher, D. B. DynaMut2: assessing changes in stability and flexibility upon single and multiple point missense mutations. *Protein Sci.* **30**, 60–69 (2021).
71. Montanucci, L., Capriotti, E., Frank, Y., Ben-Tal, N. & Fariselli, P. DDGun: an untrained method for the prediction of protein stability

- changes upon single and multiple point variations. *BMC Bioinforma.* **20**, 1–10 (2019).
72. Li, G., Panday, S. K. & Alexov, E. SAAFEC-SEQ: a sequence-based method for predicting the effect of single point mutations on protein thermodynamic stability. *Int. J. Mol. Sci.* **22**, 606 (2021).
73. Guzzi, P. H. & Roy, S. *Biological network analysis: trends, approaches, graph theory, and algorithms*. Academic Press <https://doi.org/10.1016/C2018-0-01461-9> (2020).
74. Di Paola, L., De Ruvo, M., Paci, P., Santoni, D. & Giuliani, A. Protein contact networks: an emerging paradigm in chemistry. *Chem. Rev.* **113**, 1598–1613 (2013).
75. Guzzi, P. H. et al. Computational analysis of the sequence-structure relation in sars-cov-2 spike protein using protein contact networks. *Sci. Rep.* **13**, 2837 (2023).

Acknowledgements

We acknowledge the support of the PNRR project FAIR - Future AI Research (PE00000013), Spoke 9 - Green-aware AI, under the NRRP MUR program funded by the NextGenerationEU. P.Ve. has been partially supported by Project PNRR-MCNT2- 2023-12377755 “Advancing Lung Cancer Screening: Artificial Intelligence, Multimodal Imaging and Cutting-Edge Technologies for Early Detection and Characterization”.

Author contributions

Ugo Lomoio: Data curation, Software Implementation, Experiments, Writing – original draft preparation. Valentina Carbonari: Data curation, Software Implementation, Experiments, Writing – original draft preparation. Federico Manuel Giorgi: Experiments supervision, Conceived the biological applications, Validated the biology results, Writing – original draft preparation. Pierangelo Veltri: Conceived the main idea of the manuscript, Experiments supervision, Validated the results, Funding, Writing – original draft preparation. Pietro Hiram Guzzi: Conceived the main idea of the manuscript, Experiments supervision, Conceived the biological applications, Validated the results, Funding, Writing – original draft preparation. Pietro Lìo: Conceived the main idea of the manuscript, Experiments supervision, Validated the results, Writing – original draft preparation.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41540-025-00582-2>.

Correspondence and requests for materials should be addressed to Pierangelo Veltri.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025