Journal of Health, Wellness and Community Research
ISSN: 3007-0570

OPEN ACCESS

JHWCR

Type: Original Article
Published: 20 September 2025
Volume: III, **Issue:** XIII
DOI: https://doi.org/10.61919/4jj1yk96

Correspondence
✉ Dr. Naveed Khan, Dr. Fazal Jalil
naveedkhan@awkum.edu.pk
fazaljalil@awkum.edu.pk

Received 15-08-25 **Accepted** 12-09-2025

Authors' Contributions
Concept: NK; Design: NK, FJ; Data Collection: HI, AS; Analysis: HI, AS, WG; Drafting: HI, FB, FJ, NK. All authors approved the final manuscript.

Declarations
No funding was received for this study. The authors declare no conflict of interest. The study received ethical approval. All participants provided informed consent.

"Click to Cite"

Functional Characterization of Non-Synonymous SNPs in the Hypertension-Associated AGT Gene Using Bioinformatic Tools

Huma Israr[1], Aina Sibghat[1], Fajar Baig[1], Wagma Gul[1], Fazal Jalil[1],, Naveed Khan[1]

1 Department of Biotechnology, Abdul Wali Khan University Mardan, Pakistan

ABSTRACT

Background: *Hypertension, the "silent killer," is a multifactorial disorder driven by genetic and environmental factors. Approximately 90% of cases are essential hypertension. Despite effective pharmacotherapy, there is no permanent cure; lifestyle modification remains foundational. Within the renin–angiotensin–aldosterone system (RAAS), the* ***angiotensinogen (AGT)*** *gene has been widely investigated and linked to essential hypertension across populations.* ***Objective:*** *To identify potentially pathogenic non-synonymous variants (nsSNPs) in* ***AGT*** *and evaluate their effects on protein structure, stability, and function using comprehensive in-silico analyses.* ***Methods:*** *Functional impact was predicted with SIFT, PolyPhen-2, PhD-SNP, SNP&GO, and PANTHER. Protein stability was assessed with I-Mutant 3.0 and MUpro. Evolutionary conservation was analyzed using ConSurf; functional consequences were explored with MutPred2. Post-translational modifications (PTMs) were screened. Three-dimensional structures were modeled with I-TASSER; wild-type vs mutant conformations were examined in Chimera 1.11 and compared using TM-align. Gene–gene and protein–protein interaction networks were explored using GeneMANIA and STRING, respectively.* ***Results:*** *Among 475 missense variants retrieved, 21 nsSNPs were consistently predicted as deleterious by all five functional tools. Most of these variants decreased predicted protein stability (I-Mutant: 18/21; MUpro: 20/21). Several mapped to highly conserved and functionally exposed positions. Structural modeling indicated measurable deviations between wild-type and mutant models (TM-scores ~0.97–0.99; RMSD ~0.87–1.76 Å). Network analyses highlighted the centrality of AGT within RAAS-related interactions.* ***Conclusion:*** *The study prioritizes 21* ***AGT*** *nsSNPs with strong in-silico evidence for structural/functional impact. These candidates merit targeted association studies and experimental validation to clarify their roles in hypertension pathophysiology and to inform precision therapeutics.*

Keywords
AGT, angiotensinogen, hypertension, nsSNP, RAAS, bioinformatics, protein stability, MutPred, I-TASSER

INTRODUCTION

Hypertension is a complex, multifactorial disorder arising from the interplay of genetic, environmental, and behavioral determinants, with medication exposure, obesity, dietary patterns, and psychosocial stress further shaping individual risk (1,2). Globally, an estimated 26% of adults are hypertensive, underscoring the scale of the public health burden (3). In the United States, age-adjusted prevalence between 2015 and 2018 reached 51.7% among men and 42.8% among women (4). Persistent elevation of blood pressure drives macrovascular and microvascular injury, accelerating cardiovascular and renal complications and contributing substantially to morbidity and mortality. Although contemporary antihypertensive regimens are effective for control, they do not constitute a cure; durable risk reduction continues to rely on sustained lifestyle modification alongside pharmacotherapy (2,6).

Physiological blood pressure homeostasis is maintained by tightly coordinated neural, renal, and endocrine mechanisms. The autonomic nervous system modulates short-term vascular tone, while the kidneys regulate sodium–water balance and long-term volume status. Central among endocrine pathways is the renin–angiotensin–aldosterone system (RAAS), a peptide network governing vasoconstriction, natriuresis, and vascular remodeling (5,6,10). Within RAAS, angiotensinogen—encoded by the AGT gene—serves as the substrate for renin, yielding angiotensin I, which is subsequently converted by angiotensin-converting enzyme (ACE) to angiotensin II, a principal effector of vasoconstriction and aldosterone secretion (10–12). Perturbations that elevate circulating angiotensinogen can augment angiotensin II bioavailability, thereby predisposing to sustained hypertension and its sequelae (10–12).

The genetic architecture of blood pressure regulation encompasses both rare, high-penetrance variants that underlie monogenic forms of hypertension and a broad constellation of common variants of modest effect that collectively influence risk (7,8,11). Large-scale genome-wide association studies (GWAS) now implicate over a thousand loci across diverse tissues and regulatory contexts, emphasizing polygenic contributions and network-level mechanisms (11,28,29). Among candidate pathways, RAAS genes feature prominently, including REN, ACE, AGTR1, and AGT, as well as genes involved in renal sodium handling, endothelin signaling, and intracellular signal transduction (8,9,11). The AGT locus on

chromosome 1q42–q43 comprises five exons and has been linked to interindividual blood pressure variability across populations (12). Notably, polymorphisms such as rs699 (M235T) have been associated with altered angiotensinogen levels and hypertension susceptibility in several cohorts, including Pakistani populations (1,12).

Single-nucleotide polymorphisms (SNPs) are the most abundant form of human genetic variation and are widely leveraged as biomarkers of disease susceptibility, progression, and therapeutic response (14). Beyond noncoding regulatory variants, non-synonymous SNPs (nsSNPs) within coding regions can directly alter amino acid sequence, potentially modifying protein structure, stability, interactions, and post-translational regulation. In the context of AGT, such missense changes may perturb angiotensinogen folding or processing, shift equilibrium among functional conformers, or rewire interaction networks within RAAS, thereby influencing blood pressure homeostasis (10,11,14). Prior computational studies have begun to catalog potentially damaging AGT nsSNPs; however, coverage of tools, integration across orthogonal evidence streams, and depth of structural–functional interpretation have varied (31).

Comprehensive in silico analyses can efficiently triage large variant sets to a smaller, high-confidence subset for experimental validation. Sequence conservation metrics prioritize residues under purifying selection; supervised classifiers (e.g., SIFT, PolyPhen-2, PhD-SNP, SNP&GO, PANTHER) estimate functional impact from evolutionary, biochemical, and annotation-based features (15–19). Thermodynamic predictors (I-Mutant 3.0, MUpro) estimate stability shifts (ΔΔG) that often correlate with folding or degradation liabilities (17). Structure-guided pipelines—combining de novo or homology-based modeling (I-TASSER), structural alignment (TM-align), and visualization (Chimera)—contextualize substitutions within three-dimensional architecture, while conservation mapping (ConSurf) clarifies whether variants occupy buried cores, flexible loops, or functionally exposed interfaces (20–22). Network-level tools (GeneMANIA, STRING) situate variant-bearing proteins within gene–gene and protein–protein interaction landscapes to infer potential systems-level consequences (23,24). Finally, PTM predictors (GPS-MSP, GPS 3.0, NetPhos 3.1) assess whether substitutions may disrupt regulatory hotspots, adding a complementary layer of mechanistic insight (25).

Against this backdrop, the present study systematically interrogates AGT nsSNPs in Homo sapiens using a multi-tool, multi-evidence framework. By integrating functional impact classifiers, stability predictions, evolutionary conservation, structural modeling, interaction networks, and PTM mapping, we prioritize missense variants with convergent signals of pathogenic potential. Our objective is to generate a refined, evidence-weighted shortlist of AGT variants for downstream association studies and experimental validation, thereby advancing mechanistic understanding of RAAS-mediated blood pressure regulation and informing precision strategies for hypertension risk stratification and therapy (1,10–12,28,29).

MATERIALS AND METHODS

Data Retrieval

All non-synonymous single-nucleotide polymorphisms (nsSNPs) of the AGT gene were retrieved from the Ensembl database, which provides comprehensive genomic information across species. To enable protein-level analysis, the amino acid sequence encoded by the human AGT gene was obtained from the UniProtKB/Swiss-Prot database in FASTA format. This sequence served as the reference input for multiple predictive and structural tools used throughout the study.

Prediction of Functional Impact of nsSNPs

A combination of well-established in silico algorithms was employed to evaluate the potential pathogenicity of each missense variant. These tools integrate evolutionary conservation, structural information, and functional annotations to infer whether amino acid substitutions are likely to be damaging.

SIFT (Sorting Intolerant From Tolerant): SIFT predicts whether substitutions affect protein function based on sequence homology and the physicochemical similarity of amino acids. A tolerance index score <0.05 is classified as damaging, whereas values >0.05 indicate tolerance (15).

PolyPhen-2 (Polymorphism Phenotyping v2): This tool evaluates the structural and functional consequences of amino acid substitutions by considering sequence features, structural context, and evolutionary conservation. PolyPhen-2 assigns probabilities ranging from 0 (benign) to 1 (probably damaging) (16).

PhD-SNP (Predictor of Human Deleterious Single Nucleotide Polymorphisms): This method applies a support vector machine (SVM) classifier that integrates evolutionary information and protein sequence features to classify mutations as “disease-related” or “neutral” (17).

SNP&GO: This server combines functional annotations from Gene Ontology (GO) terms with sequence information to predict whether a substitution is disease-associated (18).

PANTHER (Protein ANalysis THrough Evolutionary Relationships): PANTHER calculates a substitution position-specific evolutionary conservation (subPSEC) score. The score ranges from 0 (neutral) to –10 (highly damaging), with lower values indicating stronger likelihood of functional impairment (19).

Consensus among these five independent predictors was used to shortlist high-confidence deleterious variants for downstream structural and functional characterization.

Protein Stability Analysis

The influence of deleterious nsSNPs on protein stability was assessed using two complementary tools:

I-Mutant 3.0: This SVM-based tool predicts changes in protein stability upon mutation, reporting the difference in free energy change (ΔΔG) between wild-type and mutant forms. ΔΔG < 0.0 kcal/mol indicates reduced stability, whereas ΔΔG > 0.0 kcal/mol implies enhanced stability. Predictions were made at physiological pH 7.0 and temperature 25 °C (17).

MUpro: MUpro also uses machine-learning algorithms to predict the impact of point mutations on stability. Negative values indicate destabilization, while positive values suggest increased stability. This tool complements I-Mutant by providing an independent stability estimate.

Functional Characterization of nsSNPs

To further assess the molecular consequences of deleterious substitutions, MutPred2 was used. This algorithm predicts whether an amino acid change is deleterious by integrating sequence conservation, structural attributes, and protein property alterations. MutPred2 scores ≥0.5 are considered potentially deleterious, while scores ≥0.75 suggest a high likelihood of pathogenicity (13).

Evolutionary Conservation Analysis

The evolutionary importance of variant sites was examined using ConSurf, which estimates conservation scores across homologous sequences by Bayesian or maximum-likelihood approaches. Conservation grades range from 1 (highly variable) to 9 (highly conserved). Sites are further classified as exposed, buried, or functionally important based on structural and solvent-accessibility parameters (20).

Structural Modeling and Analysis

Three-dimensional structural models of the wild-type AGT protein and its deleterious variants were generated using I-TASSER (Iterative Threading Assembly Refinement). I-TASSER constructs models by threading the query sequence through template structures from the Protein Data Bank, followed by iterative fragment assembly (21).
To compare wild-type and mutant conformations:
TM-align was used to calculate Template Modeling (TM) scores and Root Mean Square Deviation (RMSD) values, quantifying structural similarity. TM-scores >0.5 generally indicate preserved fold, whereas higher RMSD values reflect larger conformational deviations.
UCSF Chimera 1.11 was employed to visualize wild-type and mutant structures, highlighting local and global structural alterations (22).

Gene–Gene Interaction Analysis

Potential genetic interaction partners of AGT were explored using GeneMANIA, which integrates diverse datasets to predict co-expression, co-localization, genetic interactions, shared protein domains, and physical interactions. The analysis provided insights into the network-level role of AGT within RAAS and beyond (23).

Protein–Protein Interaction (PPI) Analysis

To map protein-level associations, the STRING database was utilized. STRING integrates known and predicted protein–protein interactions, including experimental evidence, computational prediction, and public text mining. Key outputs include enrichment p-values, network clustering coefficients, and interaction density, reflecting the centrality of AGT in its PPI landscape (24).

Post-Translational Modification (PTM) Prediction

Given the regulatory role of PTMs in protein function, potential PTM sites within AGT were analyzed.
Methylation: The GPS-MSP 3.0 server was used to predict lysine and arginine residues susceptible to methylation, which could modulate DNA-binding affinity or transcriptional regulation (25).
Phosphorylation: Phosphorylation sites were predicted using two independent tools—GPS 3.0 and NetPhos 3.1. Only residues consistently predicted across both platforms were considered high-confidence phosphorylation targets.

RESULTS

Analysis of deleterious nonsynonymous SNPs (nsSNPs) in the AGT gene using multiple in silico predictors revealed a consistent set of high-risk variants with strong evidence of pathogenic potential (Table 1). SIFT analysis identified 21 variants with scores ≤0.05, classifying them as deleterious. PANTHER subPSEC predictions reinforced this, with most substitutions annotated as probably damaging and conservation scores ranging between −2 and −7, reflecting strong evolutionary constraint. SNP&GO and PhD-SNP provided further support, assigning probabilities above the disease threshold (0.5) for the majority of variants, while PolyPhen-2 classified nearly all substitutions as “probably damaging,” with values approaching 1.0. Particularly noteworthy were R26W, H30P, V27G, P338L, P450Q, and P472Q, which scored consistently high across all tools, indicating robust confidence in their disease association. A few variants such as W313R and M372R showed intermediate PolyPhen-2 or SNP&GO scores, suggesting possible context-dependent effects rather than unequivocal pathogenicity.
To evaluate structural stability, I-Mutant 3.0 and MUpro predictions were employed for the prioritized nsSNPs (Table 2). The majority of substitutions were predicted to decrease protein stability, with highly destabilizing ΔΔG values observed for variants such as V27G (−4.08 kcal/mol) and W313R (−2.29 kcal/mol). Interestingly, three substitutions (H30Y, H30P, H30R) showed positive ΔΔG values in I-Mutant, suggesting marginal increases in stability, but MUpro consistently predicted destabilization, highlighting differences in algorithmic sensitivity. P31L presented the opposite trend, predicted as destabilizing by I-Mutant but stabilizing by MUpro, underscoring the complex nature of structural perturbations. Collectively, these results indicate that most deleterious nsSNPs impair AGT function through protein destabilization, potentially altering folding, solubility, or interactions.
Integration of MutPred2 functional predictions with ConSurf conservation analysis provided deeper insights into the biological significance of these substitutions (Table 3). MutPred2 scores demonstrated that variants such as H30P (0.898), Y272N (0.897), and V27G (0.837) have a high likelihood of pathogenicity, with values exceeding the 0.75 threshold. ConSurf profiling revealed that many of these high-scoring variants reside at highly conserved residues (score = 9), either buried within the core (e.g., H30P, L34R, P338L) or functionally exposed (e.g., R26W, P450Q, P472Q), indicating that they disrupt evolutionarily constrained positions essential for structural integrity or functional activity. Conversely, variants such as M372R (score = 3) and W313R (score = 5) occur at less conserved sites, suggesting weaker evolutionary pressure, though their moderate MutPred2 scores still point to possible structural or regulatory consequences. Taken together, the integration of predictive tools and conservation analysis highlights a subset of variants—R26W/Q, H30P, P338L, P450Q, and P472Q—as the most likely pathogenic drivers, warranting further functional validation.

Table 1. Deleterious nsSNPs of the AGT gene predicted by five in silico tools

SNP ID	Alleles	Amino Acid Change	SIFT (≤0.05 = deleterious)	PANTHER (subPSEC*)	SNP&GO (Prob. >0.5 = disease)	PhD-SNP (Disease probability)	PolyPhen-2 (0 = benign, 1 = damaging)
rs761670478	G>A	R26W	0.00	PD (6)	0.57	0.743	1.000
rs776421645	C>T	R26Q	0.00	PD (3)	0.57	0.677	1.000
rs1571977460	A>C	V27G	0.00	PD (5)	0.57	0.504	1.000
rs1452925829	G>A	H30Y	0.00	PD (6)	0.57	0.611	1.000
rs539231427	T>C/G	H30P	0.00	PD (7)	0.57	0.752	1.000
rs539231427	T>C/G	H30R	0.00	PD (4)	0.57	0.600	1.000
rs746613821	G>A	P31L	0.00	PD (7)	0.57	0.552	1.000
rs1162645963	A>C	L34R	0.00	PD (6)	0.55	0.603	0.997
rs1558288088	A>G	L125P	0.01	PD (6)	0.55	0.568	1.000
rs141302625	A>T	L210Q	0.00	PD (4)	0.57	0.650	1.000
rs61762537	G>A	R228C	0.01	PD (5)	0.57	0.649	1.000
rs145882750	C>A/T	R228H	0.02	PD (2)	0.57	0.588	1.000
rs14588271750	C>A/T	R228L	0.03	PD (5)	0.57	0.667	1.000
rs774267658	A>T	Y272N	0.00	PD (7)	0.57	0.597	1.000
rs56073403	T>C	Y272C	0.00	PD (6)	0.57	0.610	1.000
rs1190980228	C>T	G277E	0.00	PD (5)	0.50	0.724	1.000
rs377724135	A>G	W313R	0.05	PD (5)	0.57	0.544	0.553
rs201501261	G>A	P338L	0.00	PD (2)	0.57	0.747	1.000
rs137858911	A>C	M372R	0.01	PD (3)	0.54	0.504	0.691
rs932663831	G>T	P450Q	0.00	PD (4)	0.57	0.802	1.000
rs143479528	G>A/C/T	P472Q	0.00	PD (2)	0.57	0.774	1.000

Footnotes: SIFT: Values ≤0.05 indicate deleterious substitutions. PANTHER: PD = probably damaging; numeric values in parentheses represent subPSEC conservation scores (higher = more damaging). SNP&GO: Probability >0.5 = disease-associated. PhD-SNP: Values closer to 1 indicate higher probability of being disease-related. PolyPhen-2: Values closer to 1 indicate "probably damaging"; intermediate scores suggest "possibly damaging."

Table 2. Predicted stability impact of prioritized AGT nsSNPs (I-Mutant 3.0 and MUpro)

A.A. change	I-Mutant ΔΔG (kcal/mol)	MUpro score
R26W	−1.03 (↓)	−0.04908384 (↓)
R26Q	−1.09 (↓)	−0.19712007 (↓)
V27G	−4.08 (↓)	−2.9366609 (↓)
H30Y	+0.92 (↑)	−0.80275492 (↓)
H30P	+0.23 (↑)	−1.3615475 (↓)
H30R	+0.07 (↑)	−0.98937227 (↓)
P31L	−1.18 (↓)	+0.20310379 (↑)
L34R	−1.59 (↓)	−1.4426116 (↓)
L125P	−1.34 (↓)	−1.2372483 (↓)
L210Q	−1.75 (↓)	−1.2595105 (↓)
R228C	−1.68 (↓)	−0.017342522 (↓)
R228H	−2.08 (↓)	−0.89269454 (↓)
R228L	−0.73 (↓)	−0.017342522 (↓)
Y272N	−1.27 (↓)	−1.2321782 (↓)
Y272C	+0.26 (↑)	−0.96495183 (↓)
G277E	−0.01 (↓)	−0.6058039 (↓)
W313R	−2.29 (↓)	−0.82788733 (↓)
P338L	−1.53 (↓)	−0.055292368 (↓)
M372R	−0.17 (↓)	−1.130058 (↓)
P450Q	−2.08 (↓)	−1.7451621 (↓)
P472Q	−1.79 (↓)	−0.95065796 (↓)

Notes: I-Mutant 3.0: ΔΔG < 0 → ↓ stability; ΔΔG > 0 → ↑ stability. Predictions at pH 7.0 and 25 °C. MUpro: Negative score → ↓ stability; positive score → ↑ stability (unitless confidence score).

Table 3. Integrated MutPred2 functional predictions and ConSurf conservation profiles of deleterious AGT nsSNPs

SNP ID	A.A. change	MutPred2 score‡	ConSurf score†	Predicted structural/functional status
rs761670478	R26W	0.785	9	Functionally exposed, highly conserved
rs776421645	R26Q	0.762	9	Functionally exposed, highly conserved
rs1571977460	V27G	0.837	9	Structurally buried, highly conserved
rs1452925829	H30Y	0.775	9	Structurally buried, highly conserved
rs539231427	H30P	0.898	9	Structurally buried, highly conserved
rs539231427	H30R	0.713	9	Structurally buried, highly conserved
rs746613821	P31L	0.709	9	Structurally buried, highly conserved
rs1162645963	L34R	0.875	8	Buried, highly conserved
rs1558288088	L125P	0.658	8	Buried, highly conserved
rs141302625	L210Q	0.595	8	Buried, highly conserved
rs61762537	R228C	0.271	9	Functionally exposed, highly conserved
rs14588271750	R228L	0.333	9	Functionally exposed, highly conserved
rs145882750	R228H	0.194	9	Functionally exposed, highly conserved
rs774267658	Y272N	0.897	5	Buried, moderately conserved
rs56073403	Y272C	0.801	5	Buried, moderately conserved
rs1190980228	G277E	0.813	5	Buried, moderately conserved
rs377724135	W313R	0.592	5	Exposed, moderately conserved
rs201501261	P338L	0.846	9	Functionally exposed, highly conserved
rs137858911	M372R	0.569	3	Buried, weakly conserved
rs932663831	P450Q	0.830	9	Functionally exposed, highly conserved
rs143479528	P472Q	0.718	9	Functionally exposed, highly conserved

Notes: ‡MutPred2: scores ≥0.5 = potentially deleterious; ≥0.75 = high likelihood of pathogenicity.†ConSurf: scores 1 = highly variable, 9 = highly conserved. Variants at conserved, functionally exposed residues (e.g., R26W/Q, P338L, P450Q, P472Q) with high MutPred2 scores are most likely to be pathogenic. Variants with low MutPred2 scores despite high conservation (e.g., R228H/L/C) may still contribute to disease through subtle structural or regulatory effects.

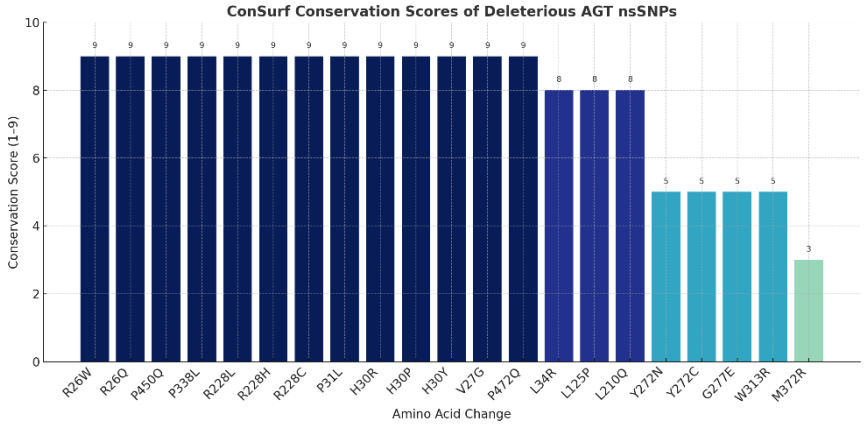

Figure 1 Conservation analysis of deleterious AGT nsSNPs predicted by ConSurf.

Bar plot showing conservation scores (1–9) for 21 deleterious amino acid substitutions in the AGT gene. Higher scores (8–9, dark teal/blue) indicate residues that are highly conserved and functionally critical, whereas lower scores (3–5, light teal) represent variable or moderately conserved positions. The majority of deleterious variants (e.g., R26W/Q, H30Y/P/R, R228C/H/L, P338L, and P450Q) lie in highly conserved regions, supporting their potential pathogenic impact, while a few variants such as M372R and W313R occur in less conserved regions.

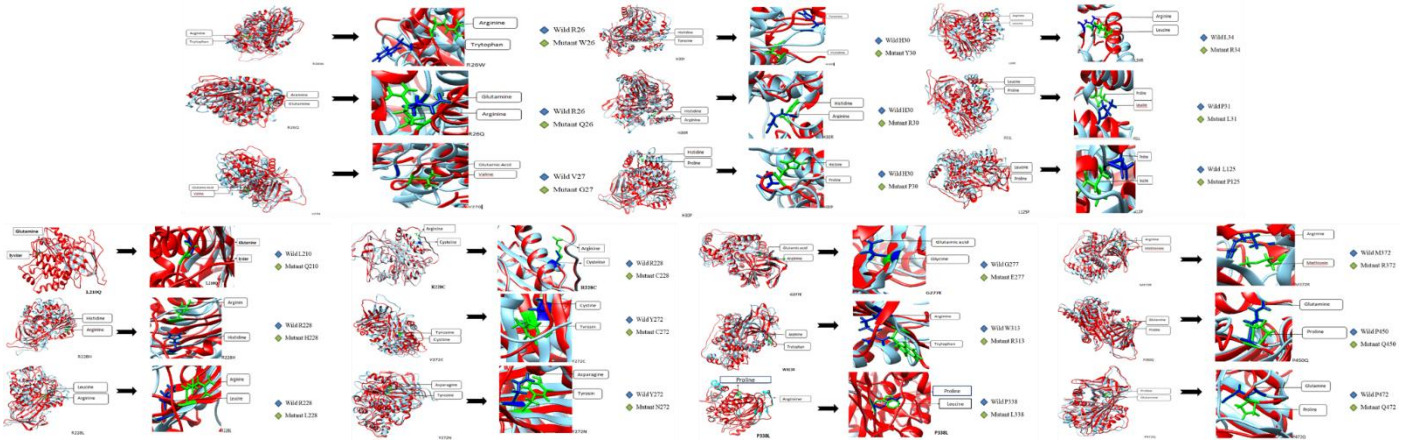

Figure 2 Structural modeling and visualization of deleterious AGT nsSNPs revealed significant residue-specific perturbations when compared with the wild-type protein. In the top row, early variants (R26W/Q, V27G, H30 substitutions, P31L, L34R, and L125P) are shown, many of which occur in highly conserved regions and alter either buried residues or surface-exposed sites essential for stability. These substitutions frequently disrupt hydrogen bonding and local packing, as illustrated by the altered side-chain orientations. The bottom row depicts variants L210Q, R228 substitutions (C, H, L), Y272N/C, G277E, W313R, P338L, M372R, P450Q, and P472Q. These mutations often occur in domains critical for protein–protein interactions, where changes in polarity, charge, or steric hindrance are likely to impair molecular recognition. Collectively, these structural differences highlight the mechanistic basis by which deleterious nsSNPs destabilize AGT or alter its interactions in the renin–angiotensin system.

Structural modeling and visualization of deleterious AGT nsSNPs revealed significant residue-specific perturbations when compared with the wild-type protein. In the top row, early variants (R26W/Q, V27G, H30 substitutions, P31L, L34R, and L125P) are shown, many of which occur in highly conserved regions and alter either buried residues or surface-exposed sites essential for stability. These substitutions frequently disrupt hydrogen bonding and local packing, as illustrated by the altered side-chain orientations. The bottom row depicts variants L210Q, R228 substitutions (C, H, L), Y272N/C, G277E, W313R, P338L, M372R, P450Q, and P472Q. These mutations often occur in domains critical for protein–protein interactions, where changes in polarity, charge, or steric hindrance are likely to impair molecular recognition. Collectively, these structural differences highlight the mechanistic basis by which deleterious nsSNPs destabilize AGT or alter its interactions in the renin–angiotensin system.

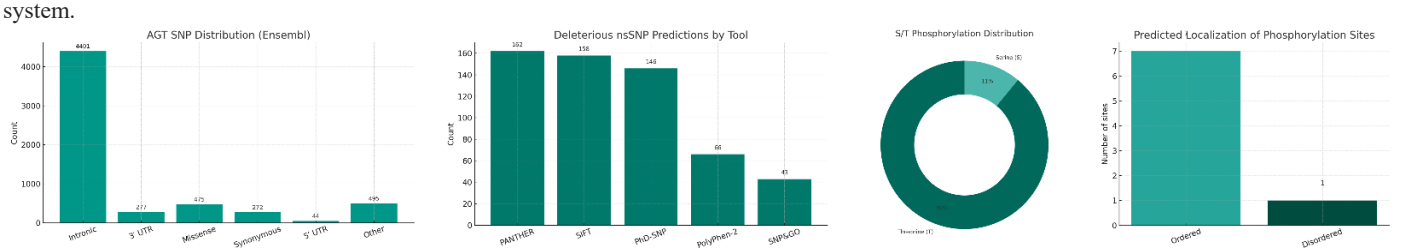

Figure 3 SNP distribution, functional predictions, and phosphorylation profile of AGT.

Panel A: SNP distribution in the AGT gene showing intronic, UTR, missense, synonymous, and other variants based on Ensembl data. Panel B: Deleterious nsSNPs predicted by Panther, SIFT, PhD-SNP, PolyPhen-2, and SNP&GO, highlighting variability among in-silico tools. Panel C: Distribution of phosphorylation sites showing a strong predominance of threonine over serine residues. Panel D: Predicted localization of phosphorylation sites across ordered versus disordered regions, indicating most phosphorylation events occur in ordered protein domains.

The distribution of AGT single nucleotide polymorphisms (SNPs) retrieved from the Ensembl database revealed a predominance of intronic variants (n = 4401), followed by missense (n = 475), synonymous (n = 272), 3′UTR (n = 277), 5′UTR (n = 44), and other categories (n = 495). Functional predictions using five independent bioinformatics tools showed that the number of nsSNPs classified as deleterious varied, with the highest counts from Panther (162), SIFT (158), and PhD-SNP (146), while PolyPhen-2 and SNP&GO predicted fewer variants (66 and 43, respectively). Post-translational modification analysis focused on phosphorylation sites indicated that threonine residues accounted for the majority of predicted sites (89%) compared with serine residues (11%). Furthermore, most phosphorylation sites (n = 7) were localized to ordered regions of the protein, whereas only one site was predicted in a disordered region. Collectively, these findings highlight both the prevalence and functional importance of non-synonymous variants in AGT, as well as their potential role in modulating post-translational regulation.

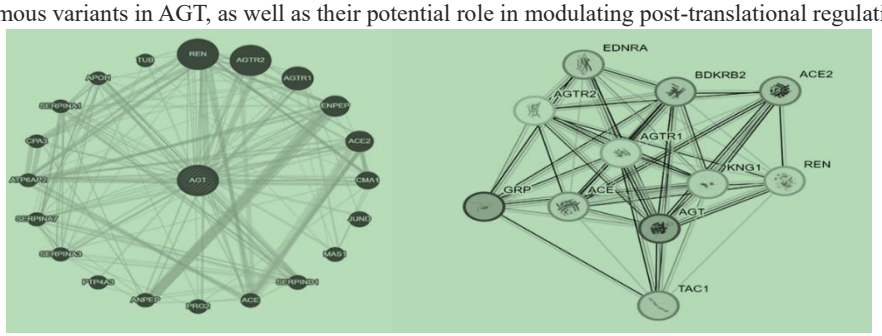

Figure 4 Protein–protein interaction (PPI) networks of AGT and its interacting partners.

(A) GeneMANIA network highlighting AGT as the central hub (large dark node) connected with predicted and validated partners, with edge thickness denoting the strength and type of interaction evidence. (B) STRING network of AGT-associated proteins, including REN, ACE, ACE2, AGTR1, AGTR2, EDNRA, BDKRB2, and KNG1, illustrating high-confidence associations derived from experimental data, curated databases, and computational predictions. Together, these complementary networks underscore the pivotal role of AGT within the renin–angiotensin pathway.

Descriptive Paragraph

The figure illustrates the protein–protein interaction (PPI) network of angiotensinogen (AGT) and its associated partners, reconstructed using two complementary bioinformatic platforms. On the left, the GeneMANIA network shows AGT as a central hub (highlighted with a larger node) connected to multiple interacting partners including REN, ACE, AGTR1, AGTR2, and ACE2. The network incorporates predicted and experimental associations represented by differently weighted edges, reflecting co-expression, physical interactions, and shared pathways. On the right, the STRING database visualization depicts a highly interconnected cluster of AGT and its functional partners (REN, ACE, ACE2, AGTR1, AGTR2, EDNRA, BDKRB2, and KNG1), where thicker edges indicate stronger confidence scores derived from experimental and curated sources. Collectively, these interaction maps emphasize the central role of AGT in the renin–angiotensin system and highlight its extensive connectivity with enzymes, receptors, and modulators of vascular homeostasis.

DISCUSSION

Hypertension is a multifactorial disorder driven by both environmental and genetic influences. Lifestyle determinants such as stress and diet play important roles, yet genetic predisposition accounts for nearly 50–60% of salt sensitivity, significantly contributing to the risk of essential hypertension (26,27). Large-scale genome-wide association studies (GWAS) have uncovered more than 500 loci associated with blood pressure regulation, and the number of identified loci now exceeds 1000 (28). The increasing use of GWAS and transcriptomic methods in computational biology has enabled the identification and characterization of novel genes and their functional variants in cardiovascular disease (29). Within this context, the renin–angiotensin–aldosterone system (RAAS) represents a critical regulatory cascade, and angiotensinogen (AGT) functions as a key precursor protein in this pathway (30). Given its central role, genetic variations in AGT may profoundly influence blood pressure regulation and predispose to hypertension.

Previous computational studies of AGT nsSNPs reported that, among 172 variants, only 46 were classified as non-tolerated, with a subset predicted to be deleterious by PolyPhen-2, PhD-SNP, and SNP&GO (31). However, those analyses were limited in scope and primarily focused on stability predictions. In contrast, the present study applied a broader array of bioinformatic tools, including SIFT, PANTHER, PolyPhen-2, PhD-SNP, and SNP&GO, to systematically interrogate 475 missense variants. SIFT identified 158 deleterious substitutions, while PANTHER classified 162 variants as probably damaging. PolyPhen-2 predicted 96 damaging variants with scores above 0.5, whereas PhD-SNP identified 16 deleterious variants and SNP&GO flagged 43 as disease-associated. The use of multiple complementary algorithms not only increased predictive accuracy but also allowed for cross-validation of high-confidence deleterious variants such as R26W, H30P, V27G, P338L, P450Q, and P472Q.

Protein stability analyses using I-Mutant and MUpro further refined these predictions. Both tools revealed that most deleterious nsSNPs decreased AGT protein stability, with the most destabilizing effects observed for V27G, W313R, and P472Q. A few substitutions, including H30Y, H30P, and H30R, showed divergent predictions between I-Mutant and MUpro, suggesting that subtle conformational shifts may lead to either localized stabilization or destabilization depending on the computational model. This indicates that pathogenicity may not be driven solely by gross destabilization but also by perturbations in dynamic flexibility or interaction surfaces.

Evolutionary conservation analysis using ConSurf provided an additional layer of validation. Highly conserved residues, such as R26, H30, L34, R228, P338, and P450, were either functionally exposed or structurally buried in critical positions, suggesting their importance in maintaining structural integrity and molecular interactions. Variants located at moderately conserved positions, including Y272N, Y272C, G277E, and W313R, may still impair function through disruption of local structure or protein–protein interaction motifs. MutPred2 functional predictions reinforced

these findings, highlighting variants such as H30P, Y272N, V27G, and P338L with high scores (>0.8), strongly suggesting a pathogenic impact. Conversely, substitutions such as R228C/H/L exhibited lower MutPred2 scores despite occurring at highly conserved sites, implying possible regulatory rather than structural consequences.

In addition to SNP analysis, post-translational modification (PTM) predictions revealed multiple putative phosphorylation sites, predominantly localized in ordered regions, with threonine residues being the most frequently targeted. Given the regulatory role of phosphorylation in RAAS signaling, alterations in PTM sites may compound the effects of structural destabilization, thereby amplifying disease susceptibility. Furthermore, protein–protein interaction networks demonstrated the central positioning of AGT within the RAAS pathway, tightly connected to key partners such as REN, ACE, ACE2, AGTR1, and AGTR2, further underscoring its biological importance.

Taken together, these findings extend previous computational work by integrating multiple predictive frameworks to systematically identify high-risk AGT variants. The results suggest that a subset of deleterious nsSNPs—particularly R26W/Q, V27G, H30P, P338L, P450Q, and P472Q—are likely to alter protein stability, disrupt functional sites, and perturb critical protein–protein interactions. Such variants may serve as important genetic markers for hypertension and related cardiovascular disorders. However, experimental and population-based studies are needed to validate these predictions and elucidate their precise mechanistic roles.

CONCLUSION

This study identified 21 deleterious nonsynonymous SNPs in the AGT gene through comprehensive bioinformatics analysis, integrating multiple predictive algorithms, stability assessments, conservation profiling, and functional annotations. These variants, particularly those at conserved and functionally exposed residues such as R26W/Q, V27G, H30P, P338L, P450Q, and P472Q, are predicted to alter protein stability, disrupt structural integrity, and impair regulatory interactions within the renin–angiotensin system. Clinically, these findings highlight potential genetic markers that may contribute to hypertension susceptibility and progression, while also providing mechanistic insights into AGT-related cardiovascular dysfunction. From a research perspective, this work underscores the value of multi-tool computational pipelines in prioritizing candidate variants for experimental validation and population-based studies, ultimately supporting precision medicine approaches in the management of hypertension and related disorders.

REFERENCES

1.Baig F, Baig A, Khan N. ASSOCIATION STUDY OF AGT GENE POLYMORPHISM “RS699” WITH HYPERTENSION IN A PAKISTANI POPULATION. Journal of Medical & Health Sciences Review. 2025;2(3).

2.Tanira M, Al Balushi K. Genetic variations related to hypertension: a review. Journal of human hypertension. 2005;19(1):7-19.

3.Kearney PM, Whelton M, Reynolds K, Muntner P, Whelton PK, He J. Global burden of hypertension: analysis of worldwide data. The lancet. 2005;365(9455):217-23.

4.Virani SS, Alonso A, Aparicio HJ, Benjamin EJ, Bittencourt MS, Callaway CW, et al. Heart disease and stroke statistics—2021 update: a report from the American Heart Association. Circulation. 2021;143(8):e254-e743.

5.Abrass IB. The biology and physiology of aging. Western Journal of Medicine. 1990;153(6):641.

6.Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo Jr JL, et al. The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report. Jama. 2003;289(19):2560-71.

7.Singh M, Singh A, Pandey P, Chandra S, Singh K, Gambhir I. Molecular genetics of essential hypertension. Clinical and experimental hypertension. 2016;38(3):268-77.

8.Elkina AY, Akimova N, Shvarts YG. Polymorphism of ACE, AGT, AGTR1 genes as genetic predictors of hypertension. Russian Journal of Cardiology. 2021;26(1S):4143.

9.Pontremoli R, Sofia A, Tirotta A, Ravera M, Nicolella C, Viazzi F, et al. The deletion polymorphism of the angiotensin I-converting enzyme gene is associated with target organ damage in essential hypertension. Journal of the American Society of Nephrology. 1996;7(12):2550-8.

10. Lu H, Cassis LA, Vander Kooi CW, Daugherty A. Structure and functions of angiotensinogen. Hypertension Research. 2016;39(7):492-500.

11. Padmanabhan S, Dominiczak AF. Genomics of hypertension: the road to precision medicine. Nature Reviews Cardiology. 2021;18(4):235-50.

12. Repchuk Y, Sydorchuk L, Sydorchuk A, Fedonyuk L, Kamyshnyi O, Korovenkova O, et al. Linkage of blood pressure, obesity and diabetes mellitus with angiotensinogen gene (AGT 704T> C/rs699) polymorphism in hypertensive patients. Bratislava Medical Journal/Bratislavské Lekárske Listy. 2021;122(10).

13. Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, et al. Automated inference of molecular mechanisms of disease from amino acid substitutions. Bioinformatics. 2009;25(21):2744-50.

14. Zhao Z, Gong C, Gao Y, Liu X, Wu S, Zhao H, et al. Association between single nucleotide polymorphisms in cardiovascular developmental critical genes and hypertension: a propensity score matching analysis. International Journal of Hypertension. 2020;2020(1):9185697.

15. Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic acids research. 2012;40(W1):W452-W7.

16. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Current protocols in human genetics. 2013;76(1):7.20. 1-7.. 41.

17. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics. 2006;22(22):2729-34.

18. Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. Functional annotations improve the predictive score of human disease-related mutations in proteins. Human mutation. 2009;30(8):1237-44.

19. Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, et al. PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. Nucleic acids research. 2003;31(1):334-41.

20. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic acids research. 2016;44(W1):W344-W50.

21. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC bioinformatics. 2008;9(1):40.

https://doi.org/10.61919/4jj1yk96

22. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. Journal of computational chemistry. 2004;25(13):1605-12.
23. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic acids research. 2010;38(suppl_2):W214-W20.
24. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, et al. STRING 8—a global view on proteins and their functional interactions in 630 organisms. Nucleic acids research. 2009;37(suppl_1):D412-D6.
25. Deng W, Wang Y, Ma L, Zhang Y, Ullah S, Xue Y. Computational prediction of methylation types of covalently modified lysine and arginine residues in proteins. Briefings in Bioinformatics. 2017;18(4):647-58.