Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Interpretable ML of Receptor Dynamics Reveals AT1R Allostery and a Negative Allosteric Modulator

Nagarajan Vaidehi

NVaidehi@coh.org

Beckman Research Institute of City of Hope https://orcid.org/0000-0001-8100-8132

Hanyu Chen

Irell and Manella Graduate School of Biological Sciences, Beckman Research Institute of the City of Hope https://orcid.org/0000-0002-5648-9939

Yoon Namkung

Research Institute of the McGill University Health Center

Zahra Asadi Jafari

Research Institute of the McGill University Health Center

Wijnand van der Velden

Beckman Research Institute of the City of Hope https://orcid.org/0000-0003-0785-3762

Elizaveta Mukhaleva

City of Hope

Grady Kestler

Beckman Research Institute of the City of Hope

Andrei Rodin

City of Hope https://orcid.org/0000-0002-2570-0332

Sergio Branciamore

City of Hope https://orcid.org/0000-0002-2556-8765

Stéphane Laporte

Research Institute of the McGill University Health Centre https://orcid.org/0000-0002-0633-543X

Article

Keywords:

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11606425/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Interpretable Machine Learning Model of Receptor Dynamics Reveals AT1R Allostery and a Negative Allosteric Modulator

Hanyu Chen[1,2], Yoon Namkung[3], Zahra Asadi Jafari[3], Wijnand J. C. van der Velden[1], Elizaveta Mukhaleva[1], Grady Kestler[1], Andrei S. Rodin[1,2]*, Sergio Branciamore[1,2]*, Stephane A. Laporte[3,4]*, Nagarajan Vaidehi[1,2]*

[1]*Department of Computational and Quantitative Medicine, Beckman Research Institute of the City of Hope, Duarte, CA, USA.*

[2]*Irell and Manella Graduate School of Biological Sciences, Beckman Research Institute of the City of Hope, Duarte, CA, USA.*

[3]*Department of Medicine, Research Institute of the McGill University Health Center, McGill University, Montréal, Québec, H4A 3J1, Canada.*

[4]*Department of Pharmacology and Therapeutics, McGill University, Montréal, Québec, H3G 1Y6, Canada.*

*Corresponding author. Email: nvaidehi@coh.org, stephane.laporte@mcgill.ca, sbranciamore@coh.org, arodin@coh.org

Abstract

Allosteric modulation of G protein-coupled receptors (GPCRs) offers opportunities to control receptor signaling, but connecting receptor dynamics to functionally relevant binding sites remains challenging. We developed an interpretable machine learning Bayesian Network (BN) modeling combined with molecular dynamics simulations to prioritize residues and binding sites for allosteric modulation. By representing each residue through its interaction energy time series, the BN model identifies residues with local and long-range conditional dependencies within receptor ensembles. Applied to the angiotensin II type 1 receptor (AT1R), comparison of AngII/G_q-bound active and antagonist-bound inactive ensembles revealed state-dependent reorganization of allosteric communication residue networks spanning the orthosteric pocket and intracellular G-protein interface. Using differential network properties, we identified residue communities enriched in mutationally sensitive residue positions. Retrospective benchmarking against published alanine-scanning data and prospective experimental testing of selected non-alanine substitutions supported the functional relevance of the network-prioritized positions. Ensemble-based binding site mapping revealed increased accessibility within a pre-existing intracellular region surrounded by residues with high positive differential weighted degrees, indicating greater network participation in the active state ensemble. Integrating this conformational accessibility with network-derived prioritization guided virtual screening we identified a small-molecule fragment Q2 whose pharmacological properties support a proposed negative allosteric modulation mechanism. Together, these results demonstrate how interpretable machine learning modeling of receptor dynamics can connect residue-level dependencies to functional pocket prioritization and guide small-molecule allosteric modulator discovery.

Introduction

G protein–coupled receptors (GPCRs) mediate diverse physiological processes and remain major therapeutic targets[1–3]. Their ability to interconvert among multiple conformational states enables bidirectional allosteric communication network between the orthosteric ligand-binding pocket and the intracellular G-protein coupling interface, involving residues distributed throughout the receptor[4–10]. Small molecules that bind outside the orthosteric site can modulate this communication network and offer opportunities for subtype-selective pharmacology[11,12]. However, identifying pockets that are both accessible to ligands and functionally coupled to receptor signaling, also known as allosterically coupled sites, remains challenging. Some pockets are transient or cryptic, becoming accessible only in particular conformations, whereas others are apparent in available structures but require further assessment of their functional relevance[13,14]. The functional relevance of such potential binding pockets may thus emerge from the dynamics of the receptor, G protein and ligand bound to the receptor.

Molecular dynamics (MD) simulations capture the conformational dynamics underlying GPCR allostery, including receptor motions, correlated movement of amino acid residues, and the transient formation of cryptic small molecule binding pockets. A major outstanding challenge, however, is not simply identifying such pockets, but determining which pockets are functionally coupled to the orthosteric site and therefore represent bona fide sites for allosteric modulation. Addressing this challenge requires methods that can identify and quantify allosteric communication in a data driven fashion between potential binding pockets and functionally important receptor sites and domains across receptor conformational ensembles.

Protein network approaches provide a framework for describing residue interactions and investigating communication between distant functional receptor's regions across conformational ensembles[15,16]. Analyses of fluctuations in residue contacts have identified collective dynamics and functionally relevant coupling[17,18]. In the commonly used dynamic network analysis[15,16,19], the importance of individual edges is solely based on correlated motions of pairs of connected nodes. However, this can often lead to spurious correlations, complicating their interpretation as network connections[20]. Complementary approaches have used neural relational inference to learn long-range interactions from MD trajectories, evolutionary coupling to prioritize allosteric residues, and structure-based pocket mapping to identify candidate GPCR modulator sites[21–23]. PocketMiner predicts cryptic-pocket locations from individual protein structures, while the recent AE-

PocketMiner approach additionally predicts their allosteric coupling[24,25]. These advances motivate a complementary question: can changes in residue-level statistical dependencies between distinct receptor signaling ensembles help prioritize functionally sensitive sites and candidate modulator drug binding pockets?

To address this question, we apply Bayesian Network Modeling (BNM), an interpretable unsupervised machine learning method[26]. Using the thermodynamic properties of amino acid residues calculated from the MD simulation trajectories as input, we generated the BN model, which is a sparse directed acyclic graph that encodes the joint probability distribution through local conditional dependencies among residues. In this network, amino acid residues form the nodes, and the edges connecting the nodes are the probability of direct co-dependencies between the dynamic input properties of the residues connected by the edge[27–29]. By accounting for dependencies among multiple variables, BNM can reduce indirect associations present in pairwise correlation analyses[30]. Building on our previously developed BNM, we used differential Bayesian network modeling (DBNM) to rigorously compare residue-level conditional dependency strengths between active and inactive receptor state ensembles (Fig. 1a). DBNM is enabled by our novel BN model scoring criterion[28] that allows direct intra- and inter-network comparisons of weighted degrees of nodes. Changes in each residue's weighted degree quantify changes in network participation and generate testable hypotheses about functionally sensitive positions and candidate modulator pockets.

We apply our framework to the angiotensin II type 1 receptor (AT1R), a prototypical class A GPCR, activated by angiotensin II (AngII) and central to cardiovascular regulation[36,37]. Although high-resolution structures have captured AT1R in active and inactive states[34,35], the dynamic communication that links the ligand-binding pocket, intracellular interface, and cryptic allosteric sites remains incompletely understood. We used MD-derived interaction energies to infer state-specific networks and retrospectively benchmarked the resulting residue rankings against published whole-receptor alanine-scanning data and mutants' $G\alpha_q$ signaling. We then prospectively tested selected non-alanine substitutions at high-ranking positions without statistically detected effects in the alanine scan, extending the experimental assessment of their functional sensitivity. Guided by this analysis, we prioritized an intracellular pocket that is limited in the active structure but expands across MD-sampled conformations and is surrounded by residues exhibiting pronounced state-dependent changes in network participation. Virtual ligand screening and experimental testing identified Q2, a small molecule hit that putatively binds this

pocket and exhibits pharmacological properties consistent with negative allosteric modulation of AngII signaling. Computational and functional analyses further support Q2 interaction with the prioritized pocket. Together, these findings demonstrate how state-dependent residue networks can guide both experimental identification of functionally important positions and prospective discovery of small-molecule modulators. Beyond the identification of Q2, the principal advance is the connection between receptor dynamics, experimentally testable residue predictions, and functionally informed pocket prioritization, providing a proof of principle for dynamics-guided GPCR allosteric drug discovery.

Results

Alanine-scanning mutagenesis effects on AT1R $G\alpha_q$ coupling

Previously, we generated a library of 359 AT1R alanine mutants or glycine in the case of native alanine, to identify residues critical for G protein coupling, using the downstream PKC-c1b BRET biosensors (Fig. 1) that report endogenous $G\alpha_{q/11}$ activation[36]. We determined the pharmacological EC_{50} and Emax values of AngII across mutants using dose–response curves, and receptor expression levels were titrated to define the minimum threshold required for maximal $G\alpha_{q/11}$ response. Mutants expressing less than half the wild-type (WT) receptor level (30 in total) were excluded from analysis. The resulting dataset provided a comprehensive map of residues affecting Gαq coupling and was subsequently used to retrospectively evaluate the residue rankings generated by DBNM. Notably, measurements on the mutants were not used as input variables in Bayesian network structure learning, which was only performed using MD-derived residue interaction energies.

Bayesian network models of active and inactive AT1R ensembles

Starting from the respective structures, we performed all-atom MD simulations on the fully active state of AT1R bound to AngII and $G\alpha_qG_{\beta\gamma}$ (pdb ID: 7F6G) and the antagonist (ZD7) bound inactive state of AT1R (pdb ID: 4YAY) in a POPC lipid bilayer (see Methods for details). For each conformational state, we performed five MD runs each 1 μs long followed by BN model analysis using our software BaNDyT[27]. We calculated the nonbonded interaction energy of each residue with its surrounding residues within 12Å radius in AT1R for every MD snapshot and used them as input variables to generate the BN model (Fig. 1a) for both the inactive and active states of AT1R. The two networks therefore represent the dynamics of AT1R in the presence of ligand-

and transducer conformational ensembles: an AngII/Gαq-engaged active complex and a ZD7-bound inactive receptor. In each resulting BN, AT1R residues are represented as nodes, the edges connecting two nodes encode conditional probability co-dependencies among their interaction energies during the dynamics, and edge weights quantifying the strength of these co-dependencies. The distribution of residue interaction energies used as inputs to the BNM is shown in Supplementary Fig. 1A, B, revealing that residues in the intracellular loops (ICLs) display greater energy spread than residues in the transmembrane (TM) helices in both the active and inactive states. As inputs to BaNDyT, we discretized the continuous residue interaction energy using a maximum-entropy binning procedure[37], which yields an approximately equal-frequency partition resulting in eight bins (see Methods for details). We fitted separate BN models for the active and inactive states of AT1R (Fig. 1). For each residue (or node), the weighted degree (Wd) which is the sum of the edge strengths connected to the node, quantifies its aggregated co-dependencies with neighboring nodes in the probabilistic space, irrespective of their spatial proximity. To evaluate the statistical robustness of the Wd, we performed a bootstrap procedure in which the original MD-derived interaction energy matrix was resampled 1,000 times and mutual information was recomputed for each resampled dataset. These mutual information values were used to reweight the fixed BN topology for every replicate, yielding a distribution of Wd values per residue. Wd estimates were highly consistent across bootstrap samples in both active and inactive states, and the resulting differential weighted degree (ΔWd) showed reproducible state-dependent shifts (Supplementary Fig. 2). These results support the stability of Wd and ΔWd estimates to resampling of the MD-derived energy data, conditional on the inferred network topology. It is important to note that the BN models do not represent energy decomposition but capture conditional dependencies among residue interaction energies across MD conformational ensembles. These dependencies describe statistical relationships among residues rather than their mean interaction energies.

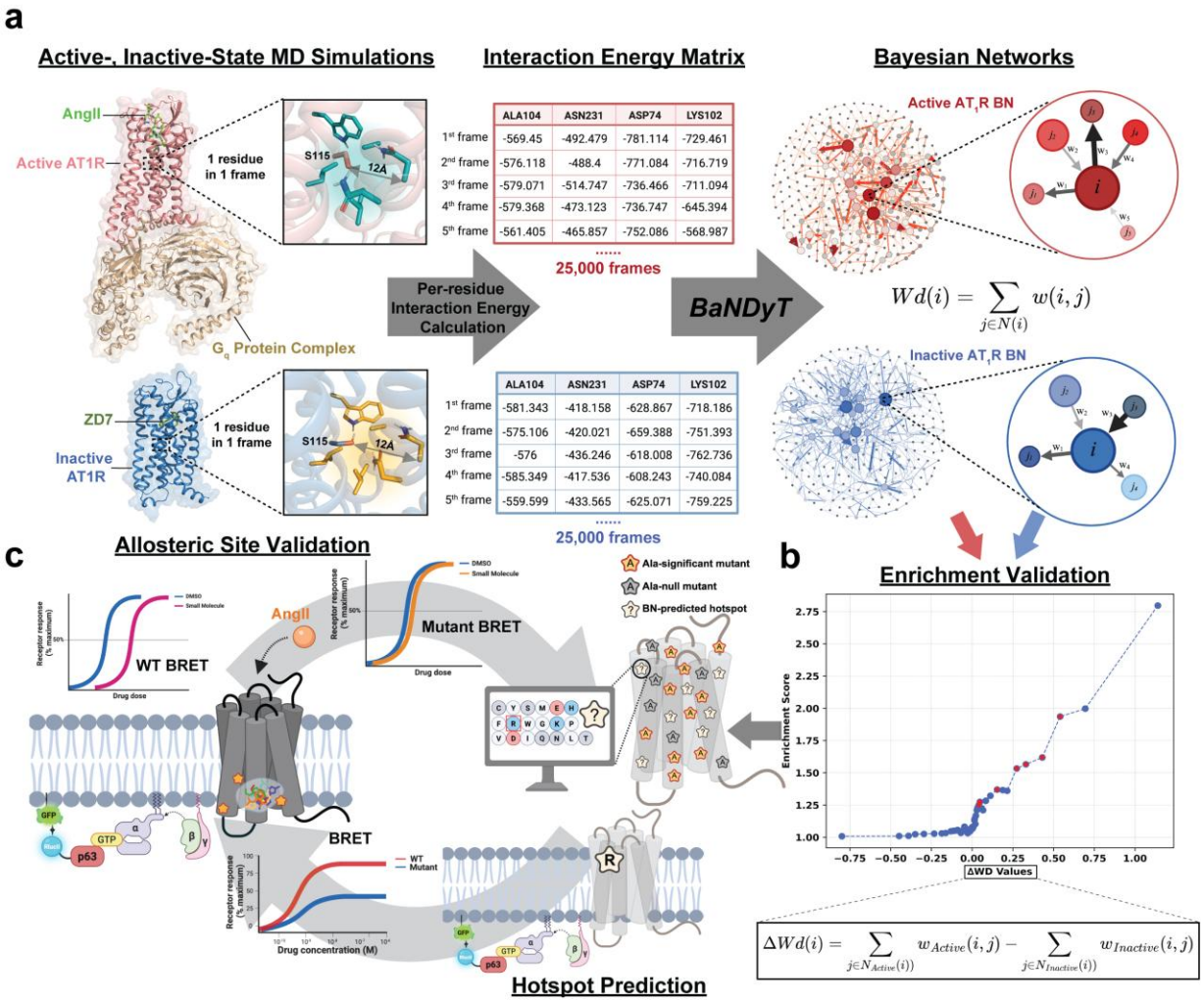


Figure 1. ***Workflow of BN analysis and the BRET experimental assays.*** *(**a**) MD simulations were done on fully active (AngII:AT1R:Gα_q) and antagonist ZD7 bound inactive (AT1R) states. The residue interaction energies were calculated from five trajectories (25,000 MD snapshots) and used as input variables to generate state-specific Bayesian networks for the active (red) and inactive (blue) states. The difference in the weighted degree for each node (residue) was calculated (ΔWd_i), and residues were ranked accordingly. (**b**) Retrospective benchmarking of the calculated ΔWd for ranked residues from Bayesian networks using enrichment analysis with the experimental positives from alanine scanning mutagenesis. Enrichment curves were generated by calculating, at each rank threshold, the score for experimental positives relative to background, with statistical significance assessed using Fisher's exact test (red points indicate $p < 0.05$). (**c**) Iterative-predict-test cycle: Predicted non-alanine mutations for high ranking ΔWd residue positions that showed no statistically detected effect upon alanine mutation, were tested for effects on AngII-mediated Gα_q activation. Gα_q BRET biosensor (Gα_q, P63-RlucII, and rGFP-CAAX) assay used for generating dose–response curves to assess receptor activation in wild-type and mutant AT1R. Also shown is the procedure to predict and test small molecules to modulate AngII mediated AT1R activation. Together, these assays show the iterative workflow used to predict residues that are hotspots in allosteric communication pathways from MD and Bayesian network analysis and to test them using BRET assays for experimental validation of allosteric sites.*

Bayesian network modeling reveals state-dependent residue dependencies in AT1R

To compare information flow within each AT1R conformational state, we calculated the Wd of each residue from the BN models of the active and inactive AT1R ensembles and performed a quartile analysis of Wd. Residues in the top quartile of Wd in each state were grouped into three functional regions: ligand-binding site, G-protein interface, and Mediators (residues that are not in ligand-binding site or G-protein interface) to quantify how information is transferred across the receptor. BN edges (representing co-dependencies in the dynamics) connecting residues at a distance greater than 10 Å separation between the Cα atoms were classified as allosteric edges. In the inactive state, residues in the top quartile of Wd, exhibit co-dependencies primarily within the G-protein interface region (33%) and between G-protein interface and mediator residues (29%), with limited coordinated motion observed within the ligand-binding site (3%) and mediator domains (18%) (Fig. 2a). The higher co-dependencies within the G-protein interface region indicate the more coordinated motion among the intracellular part of the receptor to keep the $G\alpha_q$ protein binding cavity from opening in the inactive state (Fig. 2a, b). In the active-state AngII-$G\alpha_q$ signaling complex, co-dependencies within the G-protein interface decrease, indicating reduced coordinated motion at the intracellular interface compared to the inactive state, while the mediator residues emerge as a central hub. Notably, strong interdomain links between Mediators and ligand-binding site (28%) and between Mediators and G-protein interface (16%) indicate extensive allosteric network (Fig. 2d).

Consistent with the redistribution of the connectivity of residues in the top quartile of Wd, visualization of allosteric (>10 Å) dependencies reveals a difference in allosteric communication network in the active and inactive states. The active state network has multiple direct edges from the extracellular region to the intracellular region while such long-range network edges are sparse in the inactive state (Supplementary Figs. 3A and 3B). Although receptor activation and G protein engagement cannot be fully disentangled here, this reorganization encodes functionally relevant information flow and provides a structural framework for interpreting residue-level network metrics.

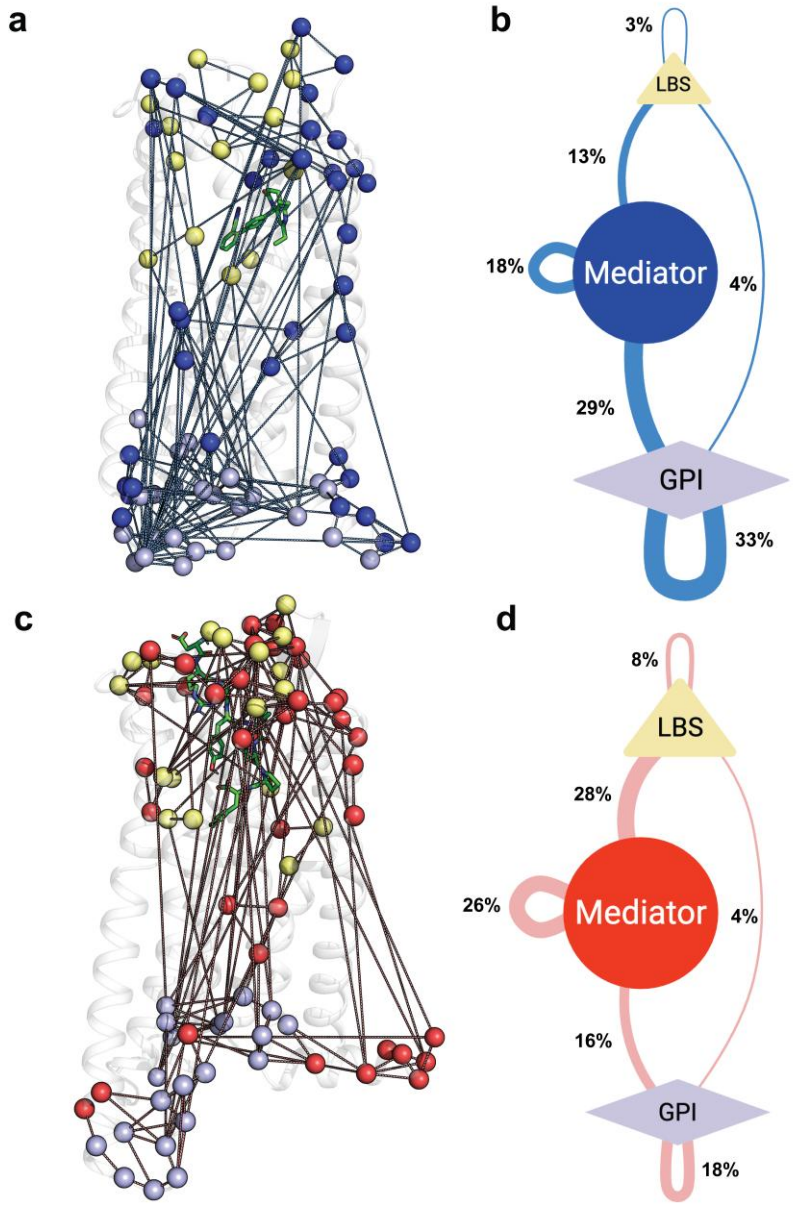


Figure 2. ***State-dependent patterns of residue dependencies in inactive and active AT1R ensembles.*** ***(a)*** *The Bayesian network edges among residues that rank in the top quartile by Wd in the inactive AT1R state shown on the structure of the inactive state of AT1R (PDB ID: 4YAY). Colored spheres represent residues by functional regions: ligand-binding site (yellow), mediator region (blue) and G-protein interface residues (purple) as identified by residue contact analysis in the AT1R: Gα_q interface from MD simulations of the active state.* ***(b)*** *Summary of the pattern of allosteric communication between different structural regions of AT1R in the inactive state. The thickness of the blue lines connecting the different structural regions is proportionate to the sum of all the edge strengths between residue pairs located in each of the three structural regions of AT1R. The shape sizes are proportionate to the number of residues in each group in the top quartile of Wd.* ***(c)*** *The Bayesian network edges among residues that rank in the top quartile by Wd in the active AT1R state shown on the structure of the active state of AT1R (PDB ID: 7F6G). Colored spheres represent residues by functional regions: ligand-binding site (yellow), mediator region (red) and G-protein interface residues (purple) as identified by residue contact analysis in the AT1R: Gα_q interface*

*from active state MD simulations. (**d**) Summary of the pattern of allosteric communication between different structural regions of AT1R in the active state. The thickness of the pink lines connecting the different structural regions is proportionate to the sum of all the edge strengths between residue pairs located in each of the three structural regions of AT1R. The shape sizes are proportionate to the number of residues in each group in the top quartile of Wd in the active state of AT1R.*

Differential weighted degree from BN models recapitulates the residues that regulate $G\alpha_q$ coupling to AT1R

We next asked whether differences in Wd between the active- and inactive-state BN models could identify positions important for $G\alpha_q$ signaling. We calculated ΔWd which is the difference in Wd of a residue in the BN derived from the AngII/$G\alpha_q$-bound active-state ensemble and in the ZD7-bound inactive-state ensemble (Fig. 1a). Positive and negative values therefore indicate greater co-dependencies among residues in the active and inactive ensembles, respectively. To clarify, this comparison quantifies differences in the probabilistic dependencies among fluctuating residue–residue interaction energies, rather than differences in mean interaction energies or the magnitude of residue displacements. We hypothesize that residues exhibiting substantial differences in network participation between the two ensembles are candidate sites for mutations that alter receptor signaling. We refer to this comparative analysis as differential Bayesian network modeling (DBNM).

To evaluate the functional relevance of ΔWd, we compared retrospectively the resulting residue rankings with previously published whole-receptor alanine-scanning data for AT1R:$G\alpha_q$ signaling[36]. Hence, the BN models and ΔWd rankings were derived exclusively from MD simulation data, without using the alanine-scanning results for model construction or optimization. The functional impact upon alanine mutation was quantified as the change in either $LogEC_{50}$, Emax or/and relative activity (RA; $\Delta\log(Emax/EC_{50})$) for AT1R:$G\alpha_q$ signaling compared to WT[36]. We considered EC_{50}, Emax, and logRA as separate but complementary pharmacological readouts because they capture distinct features of the concentration–response relationship and provide different insights into receptor-mediated signalling responses. We found that residues with higher positive ΔWd, indicating greater conditional dependencies in the active state, were significantly enriched among functionally impactful alanine mutants (Fig. 1b). The enrichment scores were calculated as the fraction of experimental positives within the top-*n* ranked residues, normalized to the overall fraction of positives in the dataset (random expectation, ES = 1). This enrichment driven by the ΔWd outperformed rankings based on Wd in either state alone (Supplementary Fig.

5C-F), demonstrating that ΔWd captures state-dependent predictive information relevant to $G\alpha_q$ coupling. Residues ranked by either their positive ΔWd (greater network participation in the AngII/$G\alpha_q$-bound ensemble than in the ZD7-bound ensemble) or negative ΔWd (greater network participation in the in the ZD7-bound ensemble than AngII/$G\alpha_q$-bound ensemble) both correlated with experimental outcomes (Supplementary Fig. 5A, B). Thus, ΔWd provides a quantitative criterion for ranking residues in the receptor by their likelihood of functionally regulating $G\alpha_q$ coupling. Of note, the sign of the ΔWd indicates whether a residue exhibits greater or lower network connectivity in one receptor state compared to the other but does not predict whether mutation of that residue will enhance or reduce receptor activity.

Residues sorted by their ΔWd are shown in a waterfall plot (Fig. 3a). Residue $G22^{1.26}$, for instance, showed significantly higher co-dependency with both spatially neighbor and allosteric residues in the active state compared to the inactive state (Fig. 3b). $R126^{3.50}$ showed low ΔWd values which stems from its high weighted degree in both conformational states (Fig. 3c). In contrast, residue $F250^{6.45}$ next to the PIF motif ($Ile^{3.40}$, $Pro^{5.50}$ and $Phe^{6.44}$) and part of the TM5/TM6 aromatic residue cluster ($F208^{5.51}$/$F249^{6.44}$/$F250^{6.45}$) that was found to distinguish between inactive and active AT1R conformation states[38,39], exhibited a negative ΔWd, reflecting reduced network connectivity in the active state relative to the inactive state (Fig. 3d).

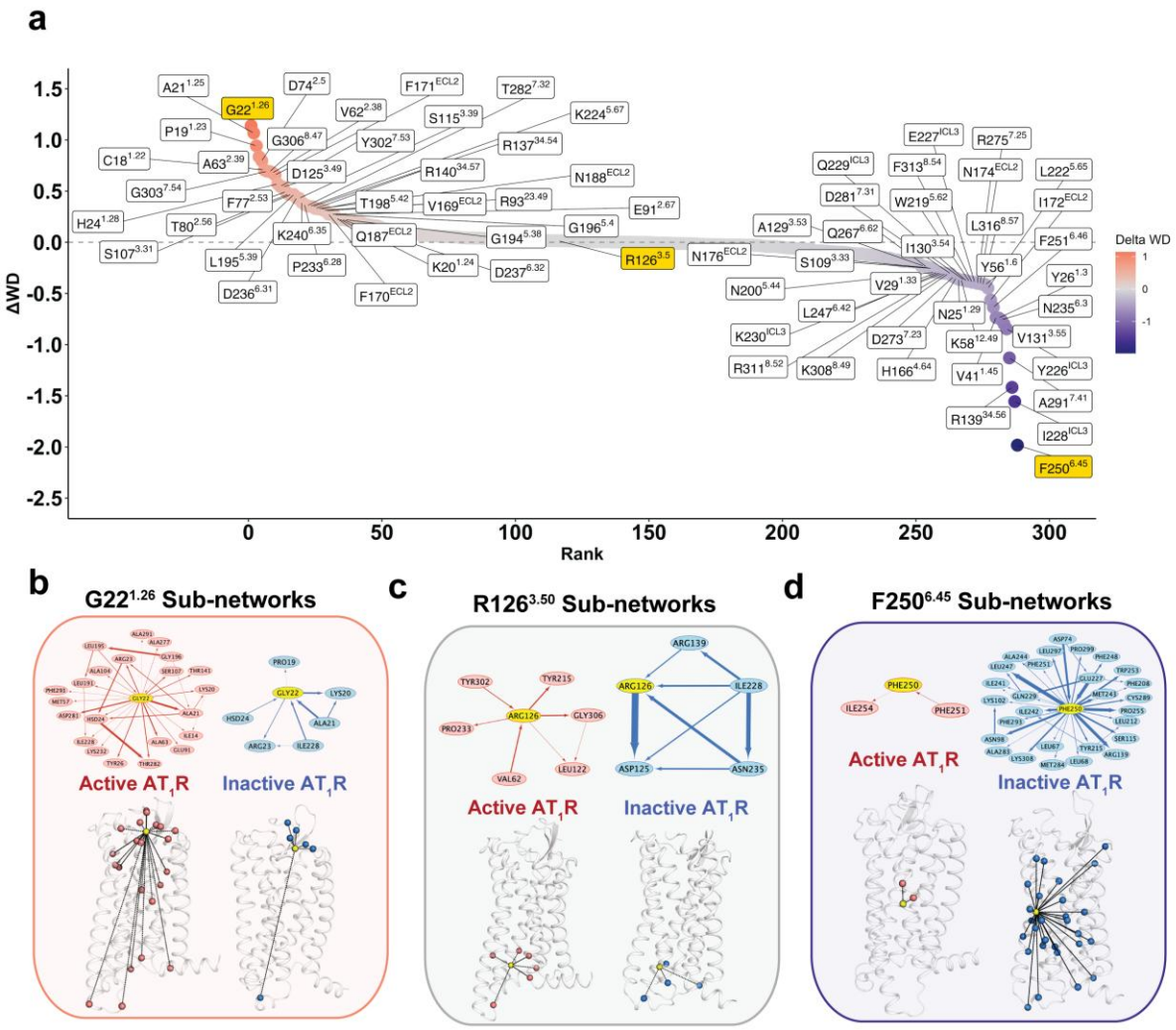


Figure 3. ***Weighted degree difference between active and inactive states follows the power law*** *(**a**) Waterfall plot showing the ranked ΔWd of all 289 residues common to active (PDB ID: 7F6G) and inactive (PDB ID: 4YAY) state AT1R structures. (**b**) Sub-network of Gly22$^{1.26}$ with a high positive ΔWd shown in active and inactive states. (**c**) Sub-network of Arg126$^{3.50}$ with ΔWd close to zero in active and inactive states. (**d**) Sub-network of Phe250$^{6.45}$ with high negative ΔWd shown in active and inactive states.*

Hierarchical clustering reveals residue communities involved in allosteric communication and targetable for allosteric modulator identification

Having shown that ΔWd captures state-dependent rewiring of residue-level networks (Fig. 3), we next asked whether residues with high ΔWd values organize into communities across the receptor. Because high ΔWd values highlight residues whose network participation changes between active and inactive states, we performed quartile-based analyses on residues with positive and negative ΔWd values and applied hierarchical clustering based on pairwise Cα–Cα distances in the corresponding structures. Residues in each ΔWd quartile were mapped

onto the receptor structures, and the resulting distance matrices were hierarchically clustered to assess whether high-ranking residues are spatially adjacent or distributed across the receptor (Supplementary Fig. 6).

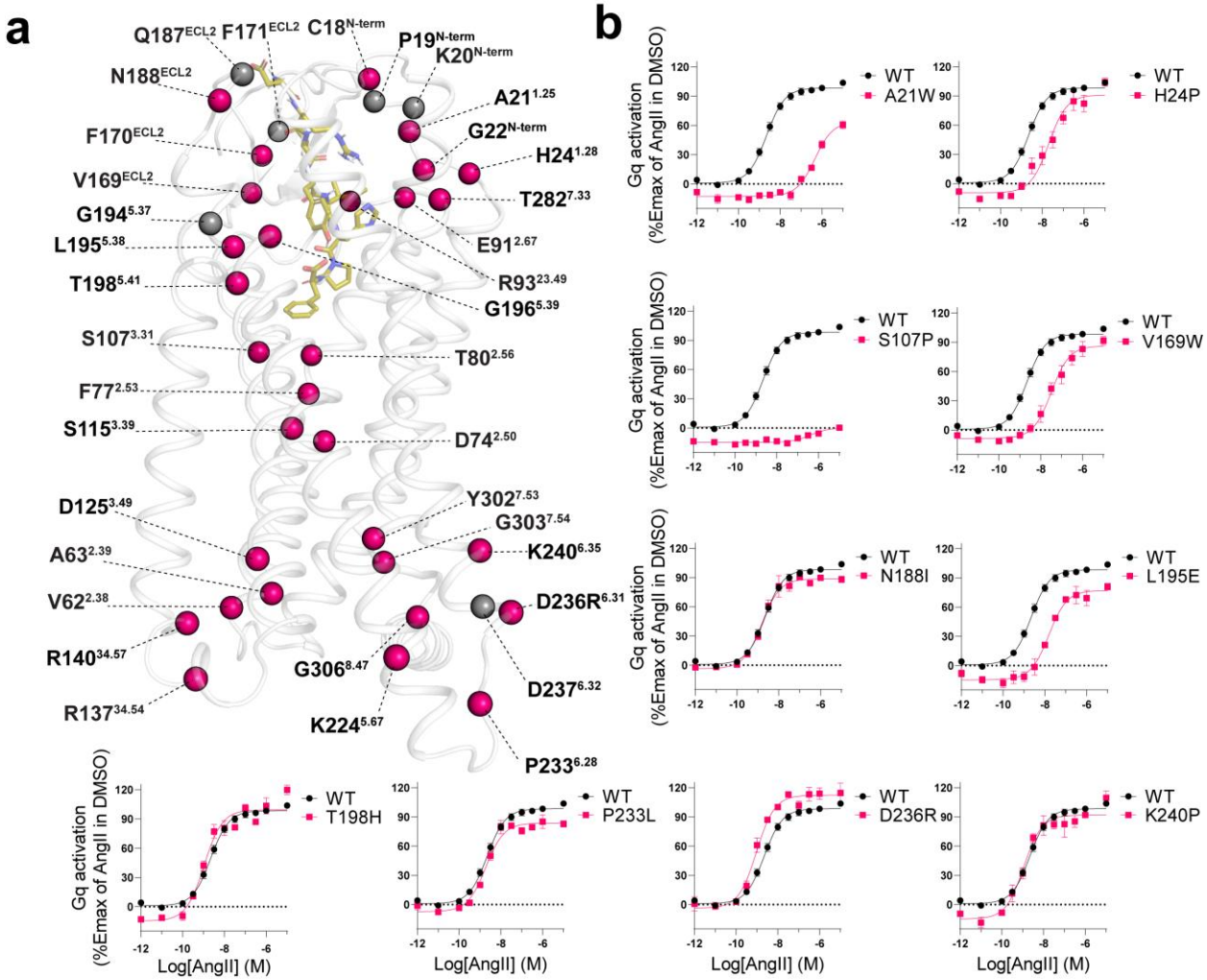
In the active state, residues in the top quartile of positive ΔWd, reflecting increased connectivity relative to the inactive state, form continuous clusters of residues spanning the ligand-binding site, mediator region, and G-protein interface, indicating strong allosteric coupling across these functional domains (Supplementary Fig. 6A). In contrast, residues in the top quartile of negative ΔWd form more localized clusters, enriched near the inner core and intracellular interface of AT1R with fewer continuous links between the ligand-binding site, mediator, and G-protein interface (Supplementary Fig. 6I), indicating a rewired communication pattern that favors stabilization of the inactive ensemble. These contrasting connectivity patterns emphasize the mediator region's central role in dynamically coupling functional sites in AT1R and highlight its potential as a strategic target for therapeutic intervention. Lower-ranked ΔWd quartiles show progressively diffuse spatial distributions, lacking the coherent communities observed among top-ranking residues (Supplementary Figs. 6B-D and 6F-H for positive ΔWd; Supplementary Figs. 6J-L and 6N-P for negative ΔWd).

Targeted non-alanine substitutions reveal functional sensitivity at network-prioritized positions

Quartile and clustering analyses identified residue communities with high positive differential weighted degree (ΔWd) enriched in functionally sensitive positions (Fig. 1). Among the 36 residues in the top ΔWd quartile, 16 showed no statistically significant effect on AT1R-$G\alpha_q$ coupling in the published alanine/glycine scan[36]. These positions may represent false negatives in mutants, and they may require alternative substitutions to reveal their functional sensitivity in AT1R. To distinguish between these possibilities, we performed an *in-silico* deep mutational residue scan evaluating all amino acid substitutions. For each mutant, we computed changes between mutant and the WT in interaction energies between the AT1R and $G\alpha_q$ protein (ΔΔG_affinity) and in total energy of the AT1R:$G\alpha_q$ system (ΔΔG_stability) (Supplementary Data 2). Candidate variants were selected for experimental testing using calculated changes in receptor-$G\alpha_q$ interaction and system stability, together with inspection of local contacts (Methods). We predicted more than one non-alanine mutation in some of the residue positions.

All tested mutants retained sufficient cell-surface expression for functional assessment. At 10 of the 16 positions, at least one tested substitution significantly altered AngII potency (EC_{50}), maximal response (Emax), or relative activity (RA) compared with WT (Fig. 4 and Supplementary Table 1). The phenotypes were heterogeneous. $V169W^{ECL2}$ and $H24P^{1.28}$ reduced AngII potency while retaining near-WT maximal responses, whereas $A21W^{1.25}$, $S107P^{3.31}$, and $L195E^{5.39}$ affected both potency and Emax. N188I/H reduced Emax and $P233L^{6.28}$ reduced Emax and RA. Conversely, $T198H^{5.42}$, $D236R^{6.31}$, and $K240P^{6.35}$ increased Emax. No statistically significant functional effects were detected for the tested substitutions at the remaining six positions: $P19H^{1.23}$, $K20D^{1.24}$, $F171D^{ECL2}$, $Q187W^{ECL2}$, $G194H^{5.38}$, $D237G^{6.32}$. Four of these positions ($K20^{1.24}$, $Q187^{ECL2}$, $G194^{5.38}$, and $D237^{6.32}$) ranked near the lower end of the first positive ΔWd quartile (Fig. 3a; Supplementary Data 1), suggesting more modest contributions relative to higher-ranked positions, although alternative substitutions may reveal additional functional effects. We acknowledge that some predicted non-alanine mutations involve substantial charge reversals, bulky substitutions, or proline substitutions that could potentially disrupt receptor structure. However, these mutants exhibited robust expression and a range of functional effects, from decreased to enhanced receptor activity, suggesting that many of these substitutions did not cause severe structural perturbations.

Collectively, these findings establish DBNM-derived ΔWd rankings as a rational framework for (i) prioritizing residues and mutations for experimental validation, (ii) potentially using this method across other GPCRs where there is no resource-intensive mutational scanning data with a targeted strategy for identifying determinants of receptor allosteric regulation. (iii) This capability could accelerate the discovery of functionally important allosteric sites and inform the development of allosteric modulators as we demonstrate below.



Best Mutant on Each Position	Bayesian Network ΔWD Rank	GPCRdb Numbering	ΔΔG interaction energy b/w AT1R-Gα$_q$	AT1R Mutant Stability	Experimental Outcome		RA ΔLog(Emax/EC50)
					pEC$_{50}$ (CL95%)	Emax ± SEM	
WT	**N/A**	**N/A**	**N/A**	**N/A**	**8.65 (8.57 - 8.74)**	**100**	**0**
A21W	2	1.25	-1.612	100.2	6.34* (6.19 - 6.49)	76.6 ± 3.6*	-2.39 ± 0.14*
H24P	11	1.28	1.611	104.305	7.69* (7.45 - 7.95)	100.5 ± 4.7	-0.89 ± 0.24*
S107P	16	3.31	1.45	31.086	6.38* (5.56 - 7.21)	15 ± 3.2*	-2.36 ± 0.57*
V169W	29	ECL2	0.212	77.433	7.49* (7.31 - 7.68)	94.7 ± 3.9	-0.88 ± 0.21*
N188I	26	ECL2	-5.318	-6.549	8.81 (8.68 - 8.95)	92.4 ± 2.9*	-0.15 ± 0.12
L195E	17	5.39	4.353	20.318	7.81* (7.65 - 7.98)	91.9 ± 3.5*	-1.15 ± 0.1*
T198H	27	5.42	0.562	41.73	8.97 (8.79 - 9.15)	113.9 ± 4.8*	0.29 ± 0.16
P233L	21	6.28	-2.45	-6.946	8.72 (8.59 - 8.86)	91 ± 3.1*	-0.23 ± 0.07*
D236R	18	6.31	-6.577	2.236	9.02 (8.88 - 9.16)	116.2 ± 3.9*	0.18 ± 0.1
K240P	20	6.35	0.853	3.653	8.95 (8.74 - 9.15)	106.8 ± 5.1*	0.24 ± 0.15

Figure 4. ***Experimental validation of BN-predicted high scoring residues with non-alanine mutations.*** *(**a**) Structural visualization of residues ranked in the top quartile by ΔWd by BN model analysis (Data S1). Colored spheres indicate prediction and testing outcomes: pink = true positives (predicted and tested by alanine or non-alanine mutation(s), gray = tested with non-alanine mutation and have not shown a significant effect. (**b**) Dose–response curves using the $G\alpha_q$ BRET biosensor ($G\alpha_q$, P63-RlucII, and rGFP-caax) for predicted non-alanine mutations at high-ranking ΔWd residue positions that showed no significant effect with alanine substitution. Data are normalized to WT, and error bars indicate SEM from three biological replicates (also see Supplementary Fig.7 in Supporting Information for data on all mutants). (**c**) Table of the non-alanine mutations predicted by a computational deep mutational scan. 3rd column contains the residue-based numbering system used for GPCRs. 4th column is the difference in the interaction energy of the $G\alpha_q$ protein with AT1R in the mutant versus the wild type (WT). The difference in total energy of the mutants versus the WT, the mutants' stability is shown in 5th column. More details in the Methods section. The 6th and 7th columns of the table are the changes in EC_{50} and Emax for each mutation in response to AngII-mediated AT1R activation, as measured by BRET. *Indicating statistically significant difference compared to the WT receptor ($p < 0.05$), obtained with unpaired Student's t-test.*

Effect of allosteric residues on AngII binding

The mutant dataset shows that in addition to residues within the ligand-binding pocket, alanine substitutions at distant positions from the orthosteric site, like M134$^{34.51}$, T141$^{4.38}$, F301$^{7.52}$, and F309$^{8.50}$ which are in the intracellular region of AT1R (Fig. 5), decreased AT1R-mediated $G\alpha_q$ response. In our previous study, we showed that these residues affect the binding affinity of AngII to AT1R[36] using radiolabeled AngII binding experiments. T141A$^{4.38}$ resulted in 75% loss in AngII binding compared to the WT. Mutants M134A$^{34.51}$, F301A$^{7.52}$, and F309A$^{8.50}$ showed AngII binding loss by 50%, 100%, and 100% respectively (Supplementary Fig. 10 in [36]). In our current BN analysis, each of these residues showed strong allosteric edges and co-dependencies with residues directly in the AngII binding site or in the second shell of AngII binding site (Fig. 5). Specifically, T141$^{4.38}$ showed strong allosteric co-dependencies (high edge weight in BNM connecting T141$^{4.38}$ and G22$^{N\text{-term}}$) with G22$^{N\text{-term}}$, which impaired AngII binding when mutated to Ala; M134$^{34.51}$ showed allosteric co-dependencies with ligand binding site residue N188^{ECL2} (as identified from MD simulations of the active state of AT1R); F301$^{7.52}$ and F309$^{8.50}$ both strongly co-dependent with nearby G306$^{8.47}$, which showed allosteric co-dependency with A21$^{N\text{-term}}$. Despite not contacting AngII, A21$^{N\text{-term}}$ is in proximity of the ligand, and has a high positive ΔWd with mutation to Trp impaired both $LogEC_{50}$ and Emax (Fig. 4b). Through these allosteric co-dependencies observed in the BNM, mutations at M134$^{34.51}$, T141$^{4.38}$, F301$^{7.52}$, and F309$^{8.50}$ could destabilize AngII-AT1R interactions indirectly by perturbing the allosteric network. Thus, BNM unifies prior mutational observations within a coherent, network-based framework of AT1R allostery.

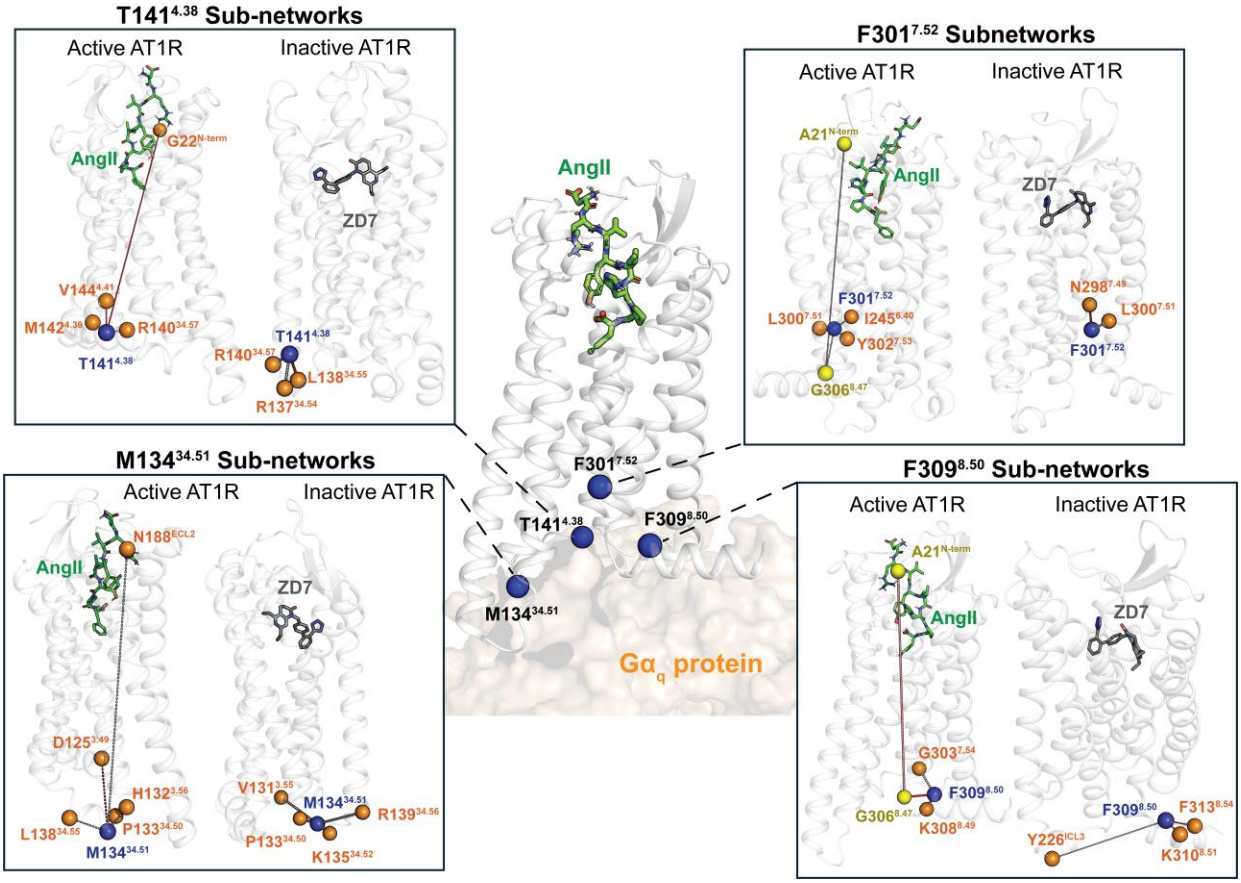


Figure 5. ***Bayesian network model provides insight into the allosteric effects of intracellular residues that modulate AngII binding:*** *Four intracellular residues T141$^{4.38}$, M134$^{34.51}$, F301$^{7.52}$, and F309$^{8.50}$ (blue spheres) affect AngII binding at a distance[36]. The edges derived from the Bayesian network highlights sub-networks (orange spheres: first-shell subnetwork; yellow spheres: second-shell sub-network) centered on these four residues that connect allosterically to the AngII binding pocket in the active state of AT1R. Mutagenesis studies show that perturbations at these sites alter AngII binding (Supplementary Fig. 10 in [36]) and EC$_{50}$ values, consistent with their predicted allosteric roles (see Fig. 3).*

Identification and characterization of a small molecule modulator, Q2 of AngII:AT1R

We next tested whether residue communities implicated in Gα_q coupling could be leveraged to find small molecules, allosteric modulators of AngII:AT1R:Gα_q signaling. We clustered the active-state MD simulation trajectories of AT1R by root mean squared deviation (RMSD) in Cα atom coordinates and this resulted in 10 conformational clusters (details in Methods). We extracted a representative conformation for each of the ten conformational clusters and used the in-house software, *FindBindSite*[40] on each representative structure to identify putative small molecule binding sites (see Methods for details). The spatial locations of the diverse set of small molecule binding pockets found across the ensemble are shown in Fig. 6a. We prioritized pockets that were consistently present across highly occupied (>75% of the population)

conformational clusters from the MD ensemble (Supplementary Fig. 8) and overlapped with the community of the top quartile of high positive-ΔWd residues identified by the DBNM (Fig. 6c). Positive ΔWd identifies hotspot residue regions in the structure to target for allosteric modulators but it does not predict the effect of a ligand acting at that region. We used this metric to prioritize functionally coupled pockets, whereas the direction of modulation was determined experimentally. This overlap ensured that the selected pocket was not only persistent across the dynamic ensemble, but also functionally enriched key residues involved in allosteric co-dependencies with other residues. The prioritized intracellular binding pocket was present in the cryo-EM structure of the active state of AT1R-Gα_q complex, but its volume increased during the MD simulations to accommodate small molecules. Using FBS we identified MD conformation that displayed a larger accessible region than the cryo-EM structure (Supplementary Fig. 9). This supports conformation-dependent expansion of a pre-existing intracellular pocket. Together with the differential network analysis, this comparison motivated selection of the region for virtual screening. Structure-based virtual screening of a 756,340-compound Enamine library against this pocket yielded 14 high-scoring candidates, selected based on docking rank, physicochemical filters, and overlap with DBNM-derived interaction hotspots (Fig. 6b). The rotamer conformation of the residues in the binding site of the selected hit molecules was optimized (see Methods for details).

Nine commercially available compounds (Q1 to Q9; Supplementary Table 2) were evaluated for modulation of AT1R-mediated Gα_q activation using a Gα_q/P63 translocation BRET biosensor[41]. Preincubation of AT1R-expressing cells with 200 μM Q2 produced an approximately two-fold right-shift in the EC_{50} of AngII-induced Gα_q signaling (Fig. 6d, e). Although Q8 and Q9 increased maximal Gα_q responses by ~25%, they were excluded from further study due to compromised luciferase activity and poor solubility (precipitation/aggregation) (Supplementary Table 2). Q2 was therefore selected as an initial small molecule (fragment: 296.32 Da)[42] hit for further pharmacological characterization (Fig. 6f-l).

Table 1. ***Q2 concentration dependent EC_{50} changes in AngII-mediated $G\alpha_q$ and β-arrestin activation.***

	LogEC$_{50}$ (EC$_{50}$ in nM)		EC$_{50}$ Fold change	
Q2 concentration	$G\alpha_q$	β-arrestin	$G\alpha_q$	β-arrestin
0 μM (DMSO)	-8.92 ± 0.04 (1.2)	-8.32 ± 0.03 (4.8)	1	1
1 μM	-8.93 ± 0.06 (1.2)	-8.34 ± 0.06 (4.6)	1	1
10 μM	-8.85 ± 0.07 (1.4)	-8.36 ± 0.07 (4.4)	1.2	0.9
50 μM	-8.76 ± 0.05 (1.7)	-8.24 ± 0.06 (5.7)	1.4	1.2
200 μM	-8.61 ± 0.05** (2.5)	-7.99 ± 0.05** (10.3)	2.0**	2.2**
500 μM	-8.33 ± 0.05*** (4.7)	-7.68 ± 0.04**** (21.1)	3.9***	4.4****

LogEC50 values were obtained from the data presented in Fig.6f. Each value represents the mean ± SEM of three independent experiments. Fold changes were obtained from dividing EC50 by the value in the absence of Q2 (DMSO) in the $G\alpha_q$ and β-arr2 pathway, respectively. $^{**}p < 0.01$, $^{****}p < 0.0001$, two-way ANOVA, compared to the value obtained in the absence of Q2 (DMSO).

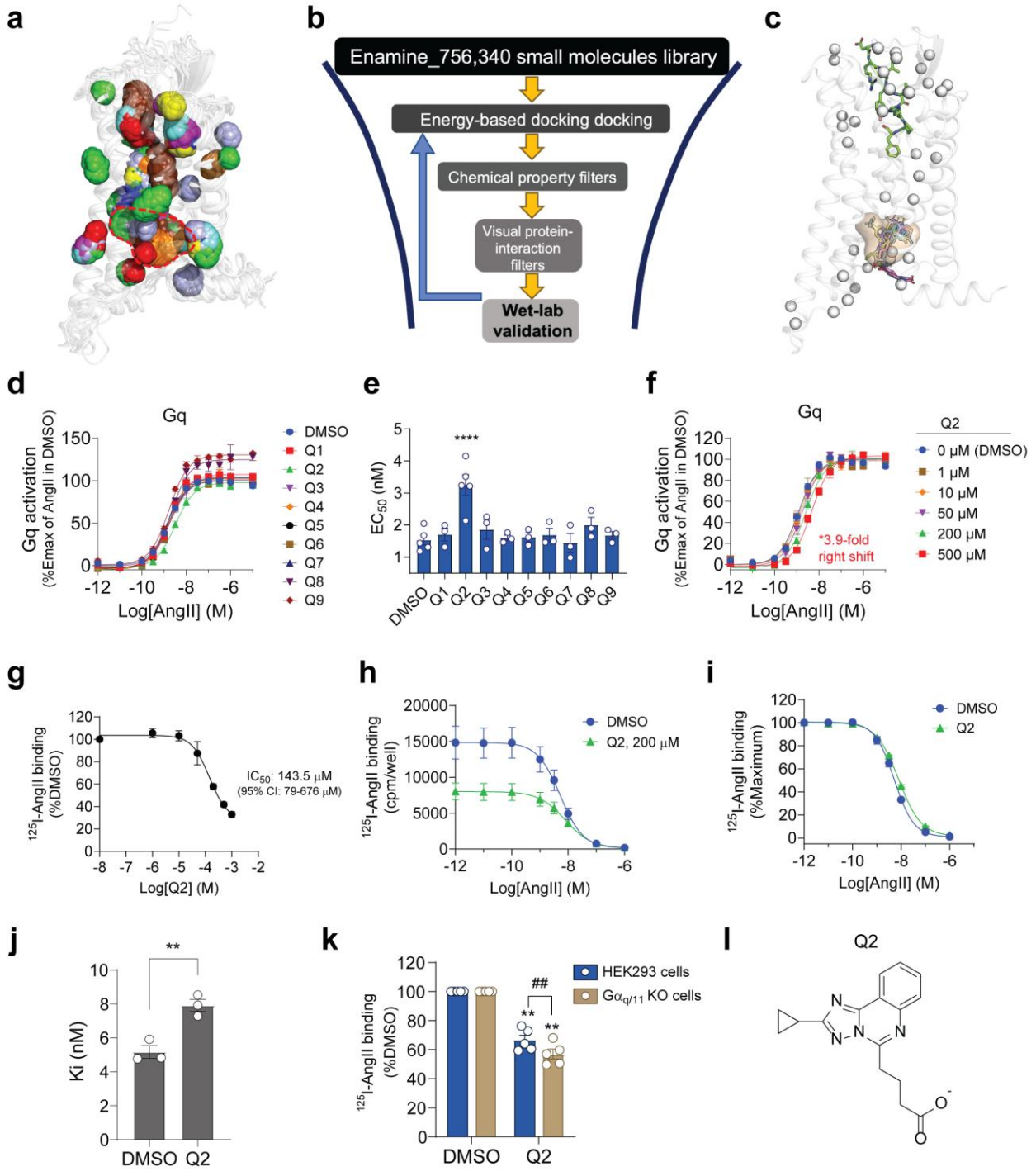


Figure 6. ***Identification of the allosteric pocket and allosteric modulator Q2 in AT1R.*** *(**a**) Ten representative structures overlaid from the ten conformational clusters of the active state MD simulation ensemble of AngII and Gα$_q$ bound AT1R used for identifying putative allosteric small molecule binding sites. The predicted small molecule binding pockets are shown as colored spheres. Each binding pocket is color-coded per conformation, with the merged overlay highlighting the consensus binding pockets. The prioritized allosteric pocket (shown in red dashed line) was selected for Virtual Ligand Screening. (**b**) Workflow for Virtual Ligand Screening (VLS) starting from ~756,000 small molecules (details in the Methods section). Ligands were filtered through Glide docking, physicochemical property thresholds,*

*pocket-specific filters, and validated against BN predictions and mutational data. (**c**) The selected allosteric pocket 'Q' shown in (A), together with the poses of top nine tested compounds. The top quartile ΔWd residues are shown in white spheres. (**d** and **e**) Functional screening of nine top VLS hits using a BRET-based $G\alpha_q$ biosensor. (**d**) AT1R-mediated $G\alpha_q$ activation as measured by $G\alpha_q$ BRET assays. HEK293 cells expressing AT1R and the $G\alpha_q$ BRET biosensor ($G\alpha_q$, P63-RlucII, and rGFP-caax) were pretreated with 200 μM of Q compounds before AngII stimulation. Normalized BRET dose–response curves are shown relative to vehicle control. Data are from more than three independent experiments and expressed as mean ± SEM. (**e**) Plot of EC_{50} in nM from concentration response curves of AngII-mediated $G\alpha_q$ activation. ****, **p** < 0.0001; one-way ANOVA, compared to the value of DMSO. (**f**) Concentration dependent effect of Q2. AT1R along with $G\alpha_q$ sensor transfected cells were preincubated with different concentrations of Q2 then stimulated with various concentrations of AngII. Data are mean ± SEM of three independent experiments. (**g-j**) Effects of Q2 on radioligand binding and AngII competition. (**g**) Competition binding of ^{125}I-AngII to AT1R was performed in intact HEK293 cells in the presence of increasing concentrations of Q2. Data are expressed as percentage of maximal specific binding relative to vehicle (DMSO). Data represent mean ± SEM of three independent experiments. (**h**) Competition binding of ^{125}I-AngII to AT1R was performed in intact HEK293 cells using increasing concentrations of AngII in the absence or presence of 200 μM Q2. Data represent total binding (cpm/well) and are shown as mean ± SEM from three independent experiments. (**i**) Data in (**h**) were normalized as a percentage of its own maximum binding. (**j**) Estimated Ki values obtained from the curves in (**i**). Data are expressed as mean ± SEM with individual values. **, p < 0.01, unpaired Student's t-test, compared to the value of DMSO. (**k**) Q2 effects on ^{125}I-AngII binding in $G\alpha_{q/11}$ Knock-out (KO) cells. ^{125}I-AngII binding to AT1R was performed in AT1R transfected HEK293 cells and $G\alpha_{q/11}$ KO cells in the absence (DMSO) or presence of 200 μM Q2. Data were normalized as a percentage of its own maximal binding (DMSO). **, p < 0.01, unpaired Student's t-test, compared to the value of DMSO in each cell line. ##, p < 0.01, paired Student's t-test, compared between Q2-treated HEK293 and $G\alpha_{q/11}$ KO cells. (**l**) Structure of Q2.*

Q2 displays pharmacological properties consistent with negative allosteric modulation of AngII:AT1R:$G\alpha_q$ complex

We characterized the effects of Q2 on AngII signaling and binding using functional and radioligand binding assays to assess a proposed allosteric mechanism. Q2 produced a concentration-dependent reduction in apparent AngII potency, with an approximately fourfold rightward shift in the functional concentration-response curve at 500 μM (Fig. 6f and Table 1). Consistent with this effect, Q2 partially inhibited ^{125}I-AngII binding over the tested concentration range, with an apparent IC_{50} of approximately 143 μM (Fig. 6g). At 200 μM, Q2 reduced AngII radioligand binding by approximately 50% (Fig. 6h). Normalizing the AngII competition curves to their respective maxima revealed a rightward shift, with the estimated AngII K_i increasing from 5 to 8 nM in the presence of Q2 (Fig. 6i; $p < 0.01$, Fig. 6j), which is consistent with an allosteric effect reported by other small allosteric molecules[43,44]. Moreover, Q2 reduced radioligand binding in both parental HEK293 and $G\alpha_{q/11}$-knockout cells (Fig. 6k). Its effect on AngII binding therefore

does not require $G\alpha_{q/11}$ binding to AT1R. This finding argues against a mechanism involving solely the displacement of $G\alpha_{q/11}$ from AT1R to regulate AngII binding.

We next examined whether Q2 altered the functional effects of different orthosteric antagonists. At 200 μM, Q2 right shifted the ZD7155 and candesartan inhibition curves, whereas no effects on losartan were observed (Supplementary Fig. 10). This pattern is consistent with possible probe-dependent modulation, which is a characteristic of allosteric modulators[45–47]. To assess receptor selectivity, we tested the negative allosteric modulation effect of Q2 across a panel of six $G\alpha_q$-coupling class A GPCRs. Q2 produced the largest rightward agonist-potency shift at AT1R, a smaller shift at the prostaglandin F2α receptor, and weaker or minimal effects at the M1 muscarinic, thromboxane TPα, oxytocin, and bradykinin B2 receptors under the conditions tested (Supplementary Fig. 11). These findings indicate that Q2 does not uniformly inhibit signaling across the tested $G\alpha_q$-coupled receptors and preferentially inhibits AT1R signaling under the conditions tested. Such receptor-dependent activity has also been reported for the micromolar-affinity β_2-adrenoceptor negative allosteric modulator compound 15[43,48]. Q2 also reduced AngII-induced β-arrestin2 coupling to AT1R (Table 1), indicating that its functional effects were not restricted to the $G\alpha_q$ signaling. Mechanistic modeling focused on $G\alpha_q$ coupling in this study, as our previous AT1R:β-arrestin2 homology model did not adequately reproduce the corresponding alanine-scanning BRET data[36].

Collectively, these pharmacological findings support the proposed negative allosteric mechanism of Q2. The persistence of its effect on AngII binding in $G\alpha_{q/11}$-knockout cells, considered alongside its predicted intracellular binding pose (Fig. 6c and 7a), is consistent with communication between the intracellular receptor region and the extracellular orthosteric site. We next investigated the functional involvement of the proposed pocket through mutational analysis.

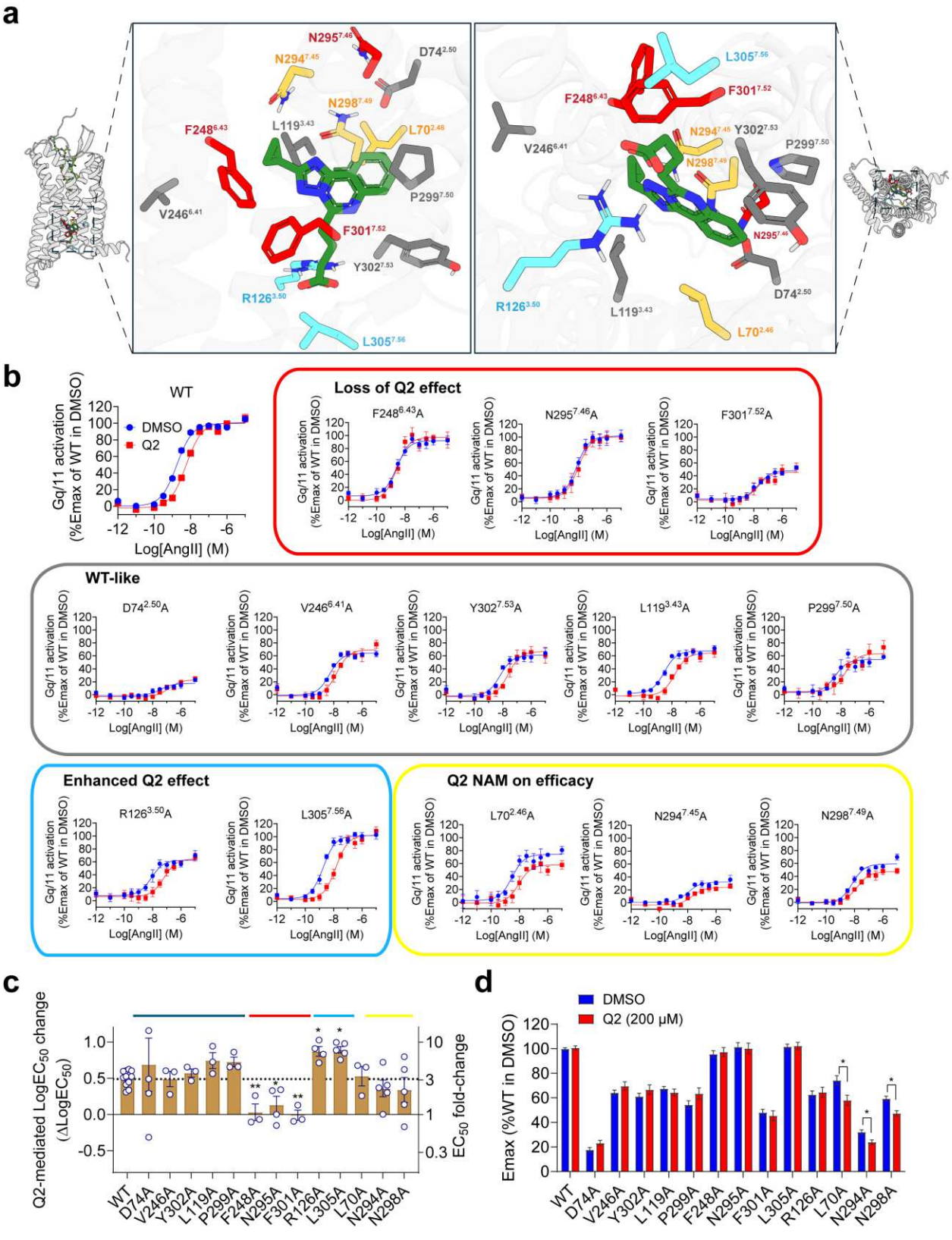


Figure 7. Model of the proposed intracellular Q2 pocket and effects of surrounding residue mutations on Q2 modulation. *(**a**) Front and bottom view of residues surrounding the predicted Q2 (green sticks)*

*binding site in AT1R and tested experimentally. Residues labeled in red represent substitutions that markedly reduced the Q2-induced AngII potency shift; residues in blue enhanced the Q2 negative allosteric modulator (NAM) effect upon mutation; residues in yellow show changes in efficacy of Q2 mediated NAM effect and residues in gray indicate mutations without significant functional effect. (**b**) AngII-concentration response curves of the AT1R mutants. Data are normalized to the Emax of WT in vehicle control (DMSO) and expressed as mean ± SEM of more than three independent experiments. (**c**) Q2-mediated logEC_{50} changes. LogEC_{50} of each mutant's AngII responses in the absence (DMSO) or presence of Q2 (200 μM) were obtained and calculated Q2-mediated logEC_{50} changes by subtracting logEC_{50} value of DMSO from that of Q2 treated. *, $p < 0.05$; **, $p < 0.01$, one-way ANOVA, compared to the value of WT. (**d**) Emax of AngII-mediated $G\alpha_{q/11}$ activation of the mutants in the absence or presence of Q2. The maximum asymptote of each curve in (**b**) is plotted. *, $p < 0.05$, unpaired Student's t-test, compared to value of control (DMSO).*

Mutational analyses support the functional relevance of the proposed intracellular Q2 binding pocket

To further evaluate the intracellular Q2 binding pocket, we tested different alanine substitutions of residues surrounding the proposed intracellular docked Q2 pose (Fig. 7a). These substitutions produced heterogeneous effects on Q2-mediated modulation. Substitutions at F248A$^{6.43}$, N295A$^{7.46}$, and F301A$^{7.52}$ markedly reduced the Q2-induced rightward shift in AngII potency (Fig. 7a, red sticks; Fig. 7b), indicating that these residues are critical for Q2 binding and/or affecting the negative allosteric modulation of Q2. In contrast, mutations at R126A$^{3.50}$ and L305A$^{7.56}$ enhanced the Q2 negative allosteric modulator (NAM) effects, while substitutions at L70A$^{2.46}$, N294A$^{7.45}$, and N298A$^{7.49}$ predominantly reduced the maximal AngII response (Fig. 7b-d and Supplementary Table 3). Other substitutions at predicted sites (D74A$^{2.50}$, V246A$^{6.41}$, Y302A$^{7.53}$, L119A$^{3.43}$ and P299A$^{7.50}$) did not affect Q2 NAM activity. These results indicate that residues within the proposed pocket influence Q2 pharmacology, either through direct interactions with Q2 or through effects on the receptor conformational changes underlying its modulation. The alanine substitutions that modestly affect Q2 affinity without altering NAM activity, suggest that the dynamic Q2-binding pocket accommodates compensatory interactions that preserve Q2 binding and NAM function at AT1R.

To investigate the relationship between Q2-mediated negative allosteric modulation and orthosteric antagonist binding, we tested the effects of Q2 and sartans on selected AT1R mutants (Fig. 8). We substituted three residues within the orthosteric ligand-binding site F77A$^{2.53}$, S109A$^{3.33}$, and Y113A$^{3.37}$, previously reported to affect losartan binding without substantially affecting AngII binding (Fig. 8a–d, g–i)[34]. We compared the inhibitory effects of Q2 and losartan

on AngII-induced $G\alpha_q$ signaling using concentrations that produced comparable right-shift inhibition responses at the WT receptor. F77A$^{2.53}$ substitution markedly reduced the inhibitory effect of losartan while enhancing Q2-mediated negative allosteric modulation compared with WT (Fig. 8b, g, h). In contrast, S109A$^{3.33}$ preserved Q2 NAM activity at levels comparable to WT while enhancing the inhibitory effect of losartan (Fig. 8c, g, h). The Y113A$^{3.37}$ substitution abolished the inhibitory effects of both ligands (Fig. 8d, g, h). We also examined whether mutation of F248A$^{6.43}$, a residue in the predicted binding site of Q2, could influence the inhibitory effects of the orthosteric antagonists losartan and candesartan. Interestingly, F248A$^{6.43}$ selectively impaired losartan-mediated inhibition but preserved the inhibitory effect of candesartan (Fig. 8e, i). These findings support bidirectional allosteric communication between the orthosteric ligand-binding pocket and the intracellular allosteric region of AT1R and suggest that mutations within the orthosteric pocket differentially influence Q2-mediated negative allosteric modulation, whereas a mutation within the predicted intracellular Q2-binding region selectively affects the inhibitory action of an orthosteric antagonist.

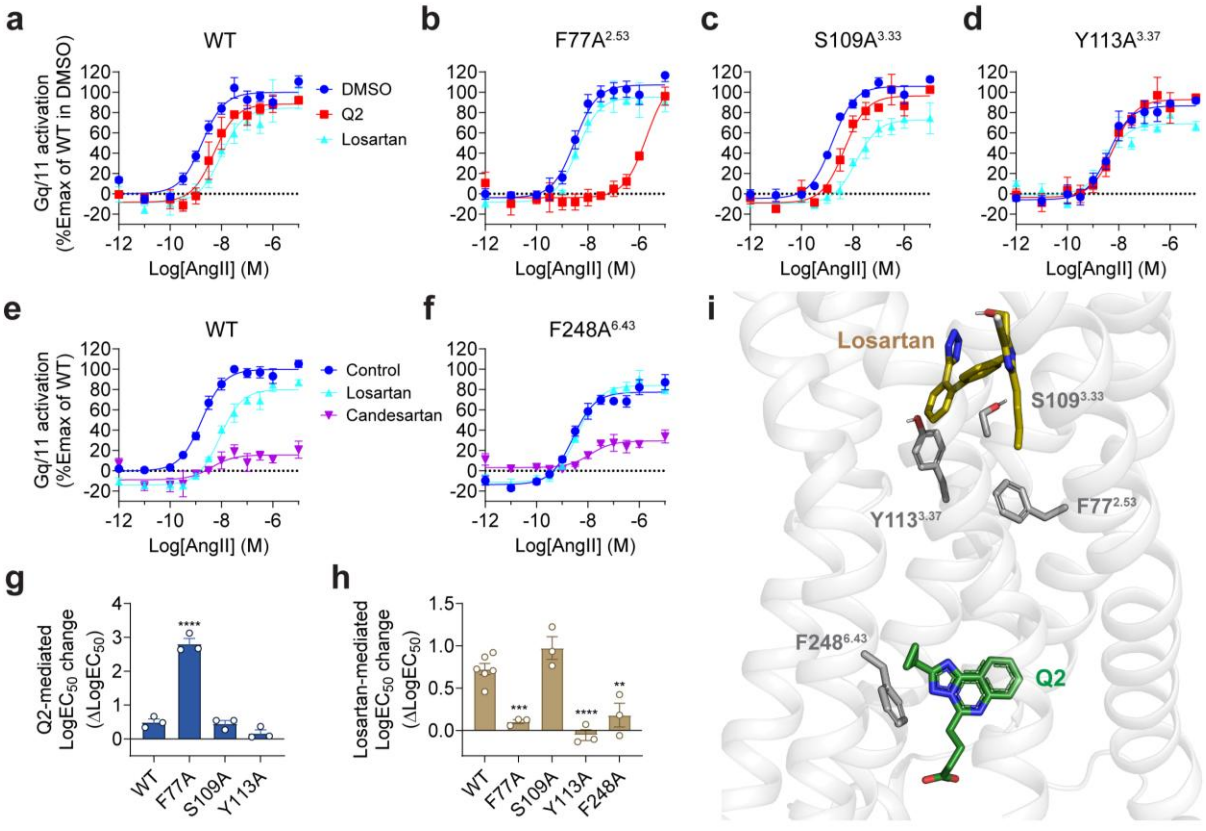


Figure 8. Effects of orthosteric- and Q2-pocket mutations on AT1R modulation by Q2 and sartans. *(**a-d**) AngII concentration-response curves for WT, F77A$^{2.53}$, S109A$^{3.33}$, and Y113A$^{3.37}$ AT1R with vehicle, Q2 (200 μM), or losartan (10 nM). (**e, f**) WT and F248A$^{6.43}$ responses with vehicle, losartan, or candesartan*

*(3 nM). **(g,h)** Q2- and losartan-induced shifts in $logEC_{50}$, respectively, relative to vehicle in AT1R WT and different mutants. **(i)** Locations of the tested residues relative to the orthosteric site and the modeled Q2 pose. Data are normalized to the Emax of WT in vehicle control (DMSO) and expressed as mean ± SEM of more than three independent experiments. **, $p < 0.01$; ***, $p < 0.001$;****, $p < 0.0001$, one-way ANOVA, compared to the value of WT.*

Distinct molecular mechanisms underlying enhanced Q2-mediated allosteric inhibition in R126A$^{3.50}$ and L305A$^{7.56}$ AT1R mutants

To investigate the enhanced Q2-mediated inhibition observed for R126A$^{3.50}$ and L305A$^{7.56}$, we performed all-atom MD simulations of WT and mutant AT1R bound to AngII, $G\alpha_qG_{\beta\gamma}$, and Q2 (Fig. 9; Supplementary Figs. 12–13). Clustering of Q2 poses revealed increased ligand flexibility from WT to L305A$^{7.56}$ and R126A$^{3.50}$, with the dominant clusters representing 98.2%, 72.1%, and 21.3% of their respective ensembles. In WT, Q2 maintained a relatively stable intracellular pose (Fig. 9a, Supplementary Fig. 12A), interacting with R126A$^{3.50}$ and disrupting the Y215$^{5.58}$–Y302$^{7.53}$ packing (Supplementary Fig. 13B) that stabilizes the active receptor conformation.

Principal Component Analysis showed substantial overlap between the $G\alpha_q$ α5 helix conformations from Q2-bound WT and active WT AT1R simulations (Fig. 9c), suggesting that Q2 primarily modulates intracellular allosteric networks rather than directly obstructing G-protein engagement. In R126A$^{3.50}$, Q2 sampled deeper and more diverse poses, forming alternative contacts with Y302$^{7.53}$, K308$^{8.49}$, and occasionally residue N357$^{G.H5.24}$ in $G\alpha_q$ (Fig. 9b, Supplementary Fig. 12B). The R126A$^{3.50}$ simulations revealed a distinct α5 helix conformational basin associated with straightening of the distal C-terminus of α5 helix ("wavy hook"[49]), displacement from the receptor cavity, and disruption of the D125$^{3.49}$–Y356$^{G.H5.23}$ interaction (Fig. 9c–e). Although Q2–N357$^{G.H5.24}$ hydrogen bonding occurred in some frames, it did not uniquely define this conformational state.

In L305A$^{7.56}$, removal of the bulky leucine side chain reorganized the intracellular TM3/TM5/TM7 region, allowing Q2 to interact with Y215$^{5.58}$ while retaining its contact with R126A$^{3.50}$. The resulting R126$^{3.50}$–Q2–Y215$^{5.58}$ polar triad may sequester Y215$^{5.58}$ from Y302$^{7.53}$, further destabilizing the active-state intracellular network (Supplementary Figs. 13 D, E). Together, these observations support a model in which Q2 acts as a negative allosteric modulator by rewiring intracellular polar interaction networks and affecting orthosteric ligand binding rather than directly and sterically blocking G protein engagement.

Together, these simulations suggest distinct mechanisms for enhanced Q2 inhibition: altered $G\alpha_q$ α5 helix conformational dynamics in R126A$^{3.50}$ and stabilization of an inhibitory

intracellular polar network in L305A$^{7.56}$. These findings support the functional relevance of the proposed intracellular allosteric pocket without establishing changes in Q2 binding affinity. The caveat is that because simulations were initiated from the docked Q2 pose, they characterize its stability and associated conformational effects rather than independently validating the binding location. More broadly, the integrated MD–DBNM approach provides a framework for identifying and experimentally investigating dynamically accessible allosteric pockets.

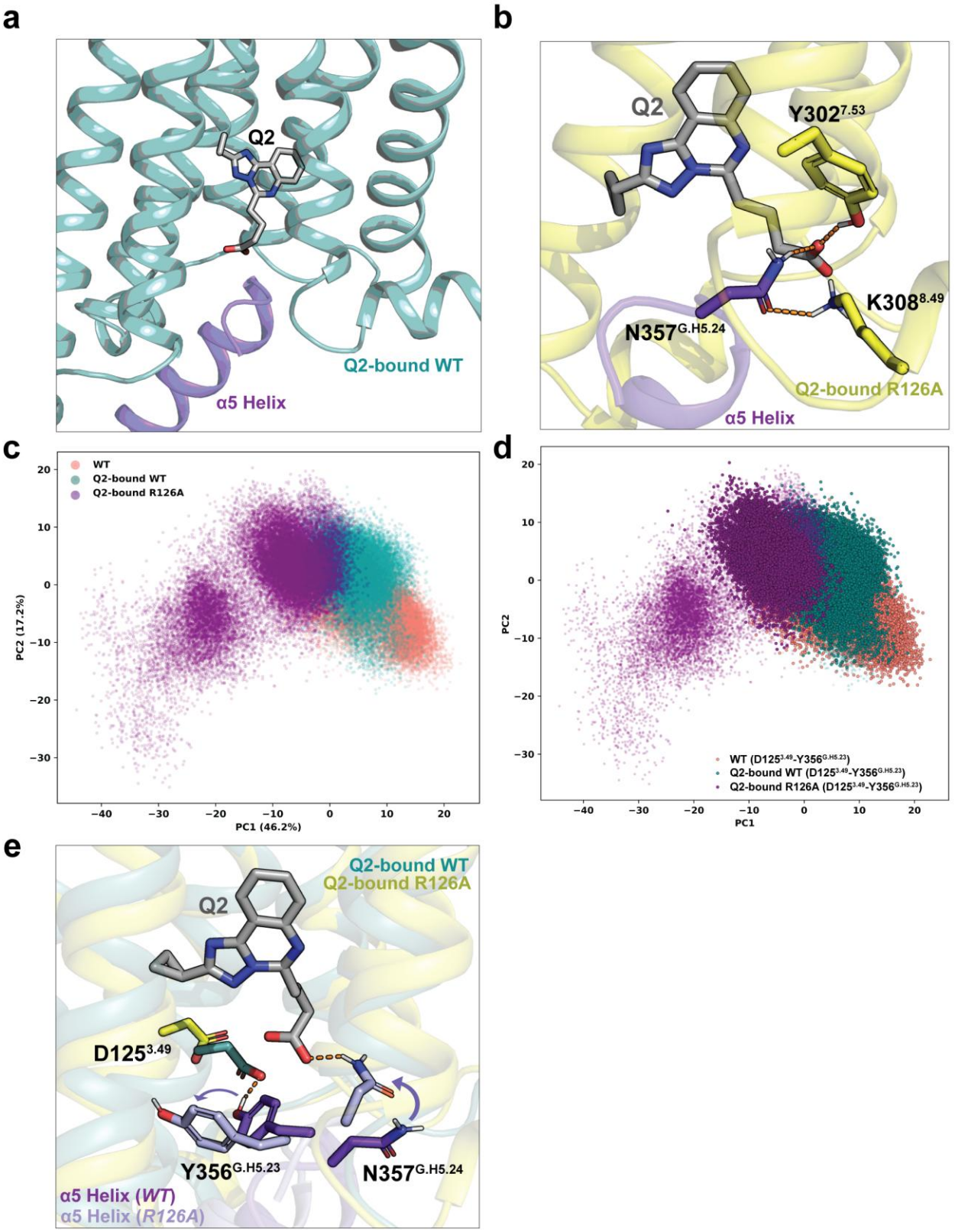


Figure 9. ***Q2 destabilizes productive G protein coupling by disrupting a conserved α5 polar lock revealed by the R126A mutant.*** *(**a**) Structural view of Q2 bound to the intracellular allosteric pocket of AT1R in the WT receptor–G protein complex, shown relative to the Gα_q α5 helix. (**b**) In the R126A$^{3.50}$ mutant, Q2 forms*

*additional interactions with residues lining the intracellular pocket and engages the α5 helix through contacts involving N357$^{G.H5.24}$, illustrating ligand proximity to the coupling interface. (**c**) Principal component analysis (PCA) of α5 helix conformations for WT (orange), Q2-bound WT (teal), and Q2-bound R126A$^{3.50}$ (purple) systems. (**d**) Conditioning the α5 PCA on the presence of the conserved receptor–G protein D125$^{3.49}$–Y356$^{G.H5.23}$ polar contact. (**e**) Structural comparison of representative WT-like and R126A α5 conformations highlighting the conserved D125$^{3.49}$–Y356$^{G.H5.23}$ polar interaction that stabilizes productive coupling in the WT complex and is destabilized in the R126A mutant.*

Discussion

AngII mediated AT1R signaling arises from ensembles of interconverting conformations [39,50–52]. A central challenge is connecting these conformational dynamics to experimentally testable predictions about which residues influence receptor signaling and which allosteric pockets within receptors can be targeted for modulation[23,53,54]. In this study, we demonstrate the power of a systematic computational framework that integrates molecular dynamics (MD) simulations with an interpretable machine learning approach, Bayesian Network Modeling (BNM), to identify residues exhibiting strong conditional dependencies with other receptor residues, quantified by their weighted degree (Wd). We show that differences in residue weighted degree (ΔWd) between active and inactive AT1R states successfully identify functionally important residues whose mutations significantly alter Gαq coupling. Several residues initially classified as false positives based on alanine scanning exhibited significant functional effects upon the introduction of our predicted non-alanine substitutions. In a companion study, we demonstrate the broader applicability of this approach to AT1R and the prostaglandin FP receptor, examining their coupling selectivity to both $G\alpha_q$ and $G\alpha_{12/13}$[55]. Taken together, these findings establish Bayesian network-derived residue rankings as an effective strategy for prioritizing functionally important residue positions involved in receptor-wide allosteric communication and G-protein coupling.

Importantly, in this present study, we have used this information on the functionally important residues in AT1R to demonstrate the utility of our computational framework for dynamics-guided discovery of GPCR allosteric modulators. As proof of principle, we integrated AT1R residue communities enriched in highly positive weighted-degree residues with putative small-molecule binding pockets identified from MD conformational ensembles. Using virtual screening of a large chemical library against these prioritized pockets, we identified Q2, a low-affinity small-molecule fragment exhibiting negative allosteric modulation of AT1R.

Multiple complementary experimental assays support the proposed receptor-dependent allosteric mechanism of Q2. First, Q2 reduced AngII radioligand binding in both parental and

$G\alpha_{q/11}$-knockout cells, demonstrating that this effect does not require $G\alpha_{q/11}$ and cannot be explained solely by displacement of $G\alpha_{q/11}$ from AT1R. Second, mutations within the predicted intracellular binding pocket differentially enhanced or diminished Q2 modulation, supporting the functional relevance of this receptor region. Third, screening Q2 against six $G\alpha_q$-coupled GPCRs revealed receptor-dependent effects, with the largest agonist-potency shift observed for AT1R. Although some promiscuity toward other receptors is expected for a non-optimized, low-affinity small molecule as shown for other GPCRs[48], our findings rule out nonspecific effects arising from membrane partitioning or colloidal aggregation of Q2. Finally, Q2 exhibited probe-dependent effects on the orthosteric antagonists losartan and candesartan, providing further evidence of allosteric modulation of the orthosteric binding site.

Additionally, mutations within the predicted intracellular pocket also differentially affected inhibition by losartan and candesartan, despite both being orthosteric antagonists. Conversely, mutations within the orthosteric binding pocket altered Q2 modulation, often without corresponding changes in losartan activity. These reciprocal, ligand-dependent mutational effects provide functional evidence for allosteric communication between the orthosteric binding site and the predicted intracellular Q2-binding region. Such differential effects have precedence in the M_4 muscarinic receptor, where an orthosteric-pocket mutation reduced acetylcholine affinity while increasing allosteric modulator affinity, and additional mutagenesis identified residues governing cooperativity between the two sites[56,57].

Although Q2 remains an early-stage, low-affinity hit, the convergence of mutagenesis, pharmacological profiling, and computational analyses supports the proposed allosteric mechanism. However, these findings do not independently establish the precise Q2 binding pose or location. Rather, the principal advance of this study is the demonstration that integrating state-specific MD simulations, interpretable Bayesian network modeling, dynamic pocket identification, and virtual screening can translate computationally predicted allosteric communication networks into experimentally testable small-molecule modulators**.** By prioritizing functionally coupled receptor regions and reducing the chemical and structural search space, this approach complements empirical screening strategies and offers a potentially transferable framework for allosteric modulator discovery across GPCRs.

Materials and Methods

Computational methods and data analysis

MD simulations

MD simulations were performed using the GROMACS package (version 2016/2019) with the Chemistry HARvard Molecular Mechanics (CHARMM) 36 force field for proteins, palmitoyl-oleoyl-phosphatidylcholine (POPC) lipids, ions, and using CHARMM Transferable Intermolecular Potential with 3 Points (TIP3P) water as solvent [58,59]. We performed separate MD simulations starting from the active-state cryo-EM structure of the AT1R-AngII-$G\alpha_q$-$G\beta_1$-$G\gamma_2$ complex (PDB ID: 7F6G), the inactive-state crystal structure of AT1R-ZD7 complex (PDB ID: 4YAY), the AT1R-AngII-Q2-$G\alpha_q$-$G\beta_1$-$G\gamma_2$ complex, the R126A_AT1R-AngII-Q2-$G\alpha_q$-$G\beta_1$-$G\gamma_2$ complex, and the L305A_AT1R-AngII-Q2-$G\alpha_q$-$G\beta_1$-$G\gamma_2$ complex. Sar^1-AngII was reverted to AngII in 7F6G, and molecules other than listed proteins and ligands were removed in structure preparation. The soluble cytochrome b562 was removed in 4YAY, and ICL3 was reconstructed during structure preparation. The AT1R-Q2 complex was taken from the glide-dock results, with $G\alpha_q$-$G\beta_1$-$G\gamma_2$ complex remains the same as active-state structural setup. For AT1R mutants, the mutated residue is computationally swapped to Ala, and the side chain rotamer conformation of residues within 5Å of the swapped residue is energetically optimized. The hydrogen atoms were added, protein chain termini were capped with neutral acetyl and methyl amide groups, and histidine-protonated states were assigned using Maestro (Schrödinger). The simulation box was created using CHARMM-GUI. We used the OPM (orientation of proteins in membranes) structure of PDB: 7F6G for alignment of the TM helices of protein structure and inserted a pre-equilibrated POPC bilayer. Final system dimensions were as follows: AT1R-AngII- $G\alpha_q$-$G\beta_1$-$G\gamma_2$ 120 × 120 × 179 Å, including 371 lipids, 58170 waters and 150 mM NaCl; AT1R-ZD7 75 × 75 × 122 Å, including 130 lipids, 13801waters and 150mM NaCl; AT1R-AngII-Q2- $G\alpha_q$-$G\beta_1$-$G\gamma_2$ 135 × 135 × 185 Å, including 489 lipids, 78019 waters and 150 mM NaCl; L305A_AT1R-AngII-Q2- $G\alpha_q$-$G\beta_1$-$G\gamma_2$ 140 × 140 × 186 Å, including 531 lipids, 84956 waters and 150 mM NaCl; R126A_AT1R-AngII-Q2- $G\alpha_q$-$G\beta_1$-$G\gamma_2$ 140 × 140 × 186 Å, including 533 lipids, 85312 waters and 150 mM NaCl. After minimizing the AT1R-ligand complex, we equilibrated it using an NVT (constant temperature, constant volume) ensemble (1 ns long) and consequently with an NPT (constant temperature, constant pressure) ensemble where we gradually reduced the position restraints from 5 to 0 kcal/mol per square angstrom (each step was 5 ns long). In the last step of equilibration, we performed 5 ns of unrestrained NPT simulations before running a total of five

production MD simulations, each 1000 ns long. The snapshots were stored every 20 ps, and every 10th snapshot of the entire 1000 ns × 5 runs amounting to 5 μs of simulation time was used for analysis (25,000 snapshots total) (Supplementary Fig. 14).

Residue-wise interaction energy calculation

To quantify residue-level energetic contributions to allosteric communication, nonbonded interaction energies were computed for each AT1R residue across the MD trajectories. The nonbonded interaction energy (Eij) between any two residues (i and j) is composed of two separate energy terms: The van der Waals (vdW) interaction energy (VLJ) given by the Lennard-Jones (LJ) potential (Eq. 1) and the electrostatic interaction energy (VC) given by Coulombic potential (Eq. 2). Specifically, the short-range (within 12 Å) Coulombic and van der Waals forces were calculated using gmx energy module in GROMACS [60] and extracted from the energy log file. The computed interaction energy is defined as the sum of the LJ and Coulombic interaction energies (Eq. 3) for every MD snapshot (25,000). All favorable interaction energies are <0 kJ/mol and hence have a negative sign.

$$V_{LJ}(r_{ij}) = 4\varepsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right), \text{where } \sigma_{ij} = \frac{1}{2}(\sigma_i + \sigma_j) \text{ and } \varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \quad (1)$$

$$V_c(r_{ij}) = 138.935485\ \frac{q_i q_j}{\varepsilon_r r_{ij}} \quad (2)$$

$$E_{ij} = V_{LJ}(r_{ij}) + V_c(r_{ij}) \quad (3)$$

This procedure yielded a time series of interaction energy values for each residue, forming the input features for Bayesian network modeling. Distributions of these interaction energies for intracellular loop (ICL) residues and transmembrane (TM) residues are shown in Supplementary Fig. 1A, B.

Construction of differential Bayesian network models

DBNM starts by constructing two distinct $\mathcal{G}_A, \mathcal{G}_B$ BN models for the two systems (datasets) to be contrasted. In this study, we have applied it to delineating the differences between the active state and inactive state of AT1R, aiming to identify the structural differences in the interaction networks. For each $\mathcal{G}$ we obtain the weighted degree for all nodes as $D_i = \sum_{j \in \Gamma_i} w_{ij}$, where Γ_i is the set of

neighbors of the node i and w_{ij} is the commensurate, universal-scale, edge strength [28]. We have previously demonstrated that the weighted degree D_i is a strong correlate of a node's functional significance, with high weighted degree predictive of nodes (both residues and residue contacts) playing key roles in protein function and stability[30]. Here, we are interested in identifying salient nodes that also differentiate between the two systems being contrasted. To that effect, we define $\Delta D_i(\mathcal{G}_\mathrm{A}, \mathcal{G}_\mathrm{B}) = D_i(A) - D_i(B)$ as the weighted degree difference for the node i between the two BN $\mathcal{G}_\mathrm{A}, \mathcal{G}_\mathrm{B}$. This measure quantifies the extent of changes in the dependency structure as identified by BNs in the two systems, directly pointing to the most important differential modulator(s) for these two systems. Then, we identify the set vertex V^* predominantly important for the system A (or B) as $\mathrm{V}_\mathrm{A}^* = \{\mathrm{v} \in V \mid \Delta_i(\mathcal{G}_\mathrm{A}, \mathcal{G}_\mathrm{B}) \geq Q_3\big(\Delta(\mathcal{G}_\mathrm{A}, \mathcal{G}_\mathrm{B})\big)\}$ (and $V_B^* = \{\mathrm{v} \in V \mid \Delta_i(\mathcal{G}_\mathrm{A}, \mathcal{G}_\mathrm{B}) \leq Q_1\big(\Delta(\mathcal{G}_\mathrm{A}, \mathcal{G}_\mathrm{B})\big)\}$, where Q_1, Q_3 are the first and third quartiles, respectively. The overlay of the V^* nodes onto the 3D structures for the two systems allows visual identification of the protein substructures corresponding to the key differences.

Testing robustness of BNM: bootstrap resampling of weighted degree and differential weighted degree

To evaluate the robustness of BN-derived residue network participation, we performed a nonparametric bootstrap procedure in which the MD-derived residue interaction energy matrix was resampled 1,000 times. For each bootstrap replicate, 5,000 MD frames were randomly drawn with replacement from the 25,000-frame original trajectory. Pairwise mutual information (MI) between residues was recomputed for each resampled dataset. This procedure retains the dependency graph structure inferred from the full dataset while allowing edge weights to fluctuate according to sampling variability in the underlying MD ensemble. For each residue r, the 1,000 bootstrap replicate i values formed the distribution: $\mathrm{Wd}_r = \{Wd_r^{(1)}, Wd_r^{(2)}, \ldots, Wd_r^{(1000)}\}$ Mean Wd values were computed for active and inactive networks separately: $\mu_r^X = \frac{1}{1000} Wd_{r,X}^{(i)}, X \in \{\text{active,inactive}\}$. State-dependent network participation shifts were quantified as: $\Delta Wd_r = \mu_r^{\text{active}} - \mu_r^{\text{inactive}}$. Bootstrap variances were propagated assuming independence of active and inactive estimates: $\sigma_{\Delta r} = \sqrt{(\sigma_r^{\text{active}})^2 + (\sigma_r^{\text{inactive}})^2}$. These propagated uncertainties were used to generate confidence intervals and assess the stability of state-dependent effects.

Structural partitioning and definition of allosteric dependencies

To obtain the list of amino acid residues that make contact between ligand-receptor and receptor-G protein we utilized the Get_Contacts software (https://getcontacts.github.io/). To interpret information flow across AT1R, residues were categorized into three structural regions: 1. Ligand-binding site: residues contacting AngII (active state) or ZD7155 (5,7-diethyl-1-[[4-[2-(2H-tetrazol-5-yl)phenyl]phenyl]methyl]-3,4-dihydro-1,6-naphthyridin-2-one) (inactive state) with greater than 20% frequency across MD snapshots. 2. G protein interface: residues contacting the Gαq subunit with >20% frequency in the active-state simulations. 3. Mediator residues: all remaining AT1R residues. Edges connecting residue pairs whose Cα atoms were separated by more than 10 Å were classified as long-range or "allosteric" dependencies. Edge strengths within and between these regions were summed to quantify patterns of information flow across receptor states.

Mutual Information (MI) analysis of residue pairs

To evaluate residue dependencies identified by the BN, we performed MI analysis on residue interaction energies extracted from MD simulations. MI was calculated for all residue pairs to capture statistical dependencies irrespective of their physical proximity. BN edge mapping was then combined with MI values to assess whether BN-highlighted interactions were supported by statistically significant dynamical coupling (Supplementary Fig. 4). Residue pairs with Cα–Cα distances greater than 30 Å were specifically examined to evaluate long-range communication (Supplementary Fig. 3). Structural mapping of the top 25% MI-supported, BN-derived residue pairs was conducted to visualize connections across functional regions, including ligand-binding sites, mediator regions, and G-protein interaction interfaces, in both active and inactive states (Fig. 2).

RMSD-based clustering of the transmembrane backbone in active-state AT1R

RMSD-based clustering was performed using GROMACS rms module by aligning using backbone atoms of transmembrane residues (GPCRdb) and then clustering using a cutoff of 0.061 nm. The representative structure from each of the ten clusters was outputted as a PDB and rendered in Schrodinger PYMOL.

Method of identifying druggable binding sites using the code FindBindSite

FindBindSite[40] was used to identify putative druggable small-molecule binding pockets where potential allosteric modulators may bind. FindBindSite is a workflow that involves docking a small (60,000) and diverse library of small molecules to the entire protein structure. We used Glide to dock the 60,000-molecule library to AT1R. After docking this diverse library of 60,000 molecules, we clustered the regions with the highest docked ligand density and favorable dock score and ranked the sites. FindBindSite then outputs the rank of each site and the centers of the clusters that show ligand occupancy above the cut-off ligand density.

In silico deep mutational scanning

Computational mutational scanning was performed using the Schrodinger Maestro *Residue Scanning* module. Active-state structure of the AT1R-AngII (PDB ID: 7F6G) was loaded into the panel, and residues P19, K20, A21, H24, S107, V169, F171, Q187, N188, G194, L195, T198, P233, D236, D237, K240, were selected to be mutated to all other 19 amino acids. For each residue, all 19 possible non-native amino acid substitutions were modeled. For each mutant, two energetic metrics were computed: (i) the change in interaction energy of AT1R with $G\alpha_q$ ($\Delta\Delta G_{affinity}$), and (ii) the change in total receptor stability of the mutant with respect to WT ($\Delta\Delta G_{stability}$). These values were summarized for the subset of mutants tested experimentally in Fig.4 and Supplementary Table 1. Postmutational rotamer conformation refinement was limited to within 5 Å of mutated residue. Stability was optimized with exhaustive enumeration. Candidate mutations were prioritized for experimental testing based on three criteria: (i) a significant change in AT1R-Gαq interaction energy (either stabilizing or destabilizing) relative to other substitutions at the same position; (ii) predicted receptor stability within 200 kcal/mol of the wild-type ensemble; and (iii) maintenance or strengthening of interhelical hydrogen bonds and van der Waals packing interactions relative to the wild-type structure.

Enrichment analysis

Residues were ranked by their differential weighted degree (ΔWd) values, calculated as the difference between the weighted degree in the active-state BN model and the inactive-state BN model. Experimental "positives" were defined as residues showing a statistically significant change ($\mathbf{p} < 0.05$, Student's *t*-test) in at least one pharmacological parameter (EC_{50}, Emax, or relative activity) compared to wild-type in the alanine scanning mutagenesis assays. For each

integer *n* from 1 to the total number of ranked residues, the enrichment score (ES) was computed as $\mathbf{ES} = \boldsymbol{p}/(\frac{\boldsymbol{P}}{\boldsymbol{N}} * \boldsymbol{n})$,where $\boldsymbol{n}$ is sample size of the subset; $\boldsymbol{N}$ is total sample size; $\boldsymbol{p}$ is true positive rate of the subset; $\boldsymbol{P}$ is total true positive rate. An ES > 1 indicates that the top-n residues are enriched for experimental positives relative to the background distribution. Enrichment curves were plotted by sequentially calculating ES for increasing n values, with every sixth point shown for visualization in Fig. 1.

Fisher's exact test

To assess the statistical significance of enrichment at specific ranked subsets, Fisher's exact test was performed comparing the number of experimental positives and negatives in the subset versus the remainder of the ranked list. The 2 × 2 contingency table consisted of: (i) Experimental positives in subset (ii) Experimental negatives in subset (iii) Experimental positives in remainder (iv) Experimental negatives in remainder. A one-tailed Fisher's exact test ($p < 0.05$) was used to determine whether the proportion of positives in the subset was greater than expected by chance.

MI and BNM analyses uncover allosteric communication pathways

Given the clear hierarchical clustering and differential connectivity patterns identified among residues in active versus inactive states, we next sought to quantitatively validate these interactions at the residue-pair level using MI analysis combined with BN edge mapping. MI captures statistical dependencies between residue pairs irrespective of their structural proximity, providing an additional validation layer for the BN-derived residue dependencies and their roles in state-specific receptor dynamics. Analysis of residue pairs highlighted by BN edges at long distances (>30 Å) revealed critical insights into genuine allosteric communication within AT1R. These residue pairs, despite their significant physical separation, are identified by our model to have dynamical dependencies, indicating that their interaction energies were statistically coupled in a state-dependent manner. This long-range coupling suggests these interactions facilitate communication between spatially distant functional regions of the receptor, aligning well with classical definitions of allosteric interactions. Structural mapping of these long-range interactions further supported their functional relevance, linking residues in ligand-binding and G-protein coupling sites through intermediary mediators in both active and inactive states (Supplementary Fig. 4).

Experimental methods

Reagents

Human AngII [Asp[1]-Arg[2]-Val[3]-Tyr[4]-Ile[5]-His[6]-Pro[7]-Phe[8]], oxytocin, bradykinin, carbachol, ZD7155, poly-L-ornithine, ANTI-FLAG® M2-Peroxidase (HRP) antibody, and SIGMAFAST OPD were purchased from Sigma-Aldrich. Prostaglandin F2α (PGF2α), U-46619, and candesartan were purchased from Cayman chemical. Losartan was synthesized at University of Sherbrooke (QC, Canada). Dulbecco's Modified Eagle Media (DMEM), fetal bovine serum (FBS), and other cell culture additives were purchased from Gibco, Life Technologies. Linear polyethyleneimine MW 25000 (PEI) was purchased from Polysciences. Sixteen percent paraformaldehyde (PFA) was purchased from ThermoFisher Scientific. BSA (Bovine Serum Albumin) was purchased from Wisent. Coelenterazine 400a was purchased from Nanolight Technology. Q5 high-fidelity DNA polymerase, Gibson Assembly Master mix, DpnI, and other PCR reagents were purchased from New England BioLabs.

DNA constructs and mutagenesis

Alanine mutants of Signal peptide-FLAG-tagged human AT1R (sp-flag-AT1R) construct were described previously. $G\alpha_q$ DNA was obtained from cDNA.org (www.cDNA.org). P63-RlucII, βarr2-RlucII, and rGFP-CAAX were described previously[41,61]. Mutations were introduced into the sp-flag-AT1R using a two-fragment PCR approach, as previously described [62]. In short, forward and reverse site-directed mutagenesis primers (mutant-F and mutant-R) were designed with 21 bp of Gibson homology for Gibson assembly recombination. Stepdown PCR was used to make mutations, where two separate PCR reactions (primer pairs of mutant-F with Ori-R primers and mutant-R with Ori-F primers, respectively) were carried out to divide the vector in half. Each half of the PCR samples thereafter were purified and Gibson assembled. Reassembled vectors were transformed into *E. coli* and colonies were selected and amplified. Mutations were confirmed by Sanger sequencing (Genome Quebec CES). All primer sequences are listed in the Data S3.

Cell culture and transfection

HEK293 and $G\alpha_{q/11}$ KO cells were cultured in DMEM supplemented with 10% FBS and 20 µg/ml gentamicin. Cells were authenticated (GeneCopoeia™) and tested for mycoplasma contamination

periodically (ABM™, Mycoplasma PCR Detection kit). Cells were grown in 5% CO_2 and 90% humidity. Cells were seeded at a density of 9,000 cells per well in a white (for BRET) or clear (for ELISA) 96-well plate, and the next day, transfected with receptor and BRET sensor constructs using PEI transfection reagent. Briefly, 1 µg of total DNA in 100 µl of PBS was mixed with 100 µl of PBS containing 2.5-3 µl of 1 mg/ml PEI. For $G\alpha_q$ signaling ($G\alpha_q$/P63 sensor), 100 ng of AT1R DNA with 20 ng of $G\alpha_q$, 30 ng of P63-RlucII, and 80 ng of rGFP-CAAX were used. For the β-arr2 sensor, 150 ng of AT1R, 20 ng of βarr2-RlucII, and 80 ng of rGFP-CAAX DNA were used. For $G\alpha_{q/11}$ signaling by endogenous $G\alpha_q$ and $G\alpha_{11}$, 150 ng of AT1R DNA with 30 ng of P63-RlucII and 80 ng of rGFP-CAAX were used. For ELISA, 300 ng of AT1R DNA was used. Empty pcDNA was used to make up 1 µg of the total DNA amount. After 20 min incubation, the DNA/PEI complexes were dispensed into cells in 96-well plates (15 µl/well). All assays were performed 48 h post-transfection.

BRET measurements

HEK293 cells transfected in a white 96-well plate were washed once with Tyrode's buffer (140 mM NaCl, 2.7 mM KCl, 1 mM $CaCl_2$, 12 mM $NaHCO_3$, 5.6 mM D-glucose, 0.5 mM $MgCl_2$, 0.37 mM NaH_2PO_4, 25 mM HEPES, pH 7.4) and left in Tyrode's buffer. Coelenterazine 400a (final concentrations of 2.5 µM) was added 1 min before ligand stimulation. Cells were then stimulated with various concentrations of AngII for 3 min for the $G\alpha_q$/P63 sensor or 5 min for the β-arrestin translocation sensor, prior to BRET measurements. For compound treatment, compounds were preincubated at 37 °C for 20 min, then 10 more min at RT before ligand stimulation. BRET signals were measured using a Synergy2 (BioTek) microplate reader with emission filters, 410/80 nm and 515/30 nm for detecting the RlucII (*Renilla* luciferase) (donor) and GFP10/rGFP (acceptor) light emissions, respectively. The BRET ratio was determined by calculating the ratio of the light emitted by GFP10/rGFP over the light emitted by the RlucII. The effects of the compounds on luciferase activity were quantified based on the 410-nm readout.

Whole-cell ELISA

HEK293 cells were transfected with pcDNA, WT, or mutant receptors in poly-L-ornithine-coated, transparent 96-well plates to quantify the expression of receptors on the cell surface. 48 hours after transfection, cells were washed with PBS containing 1 mM $MgCl_2$ and 1 mM $CaCl_2$ (PBS-CM) and fixed with 4% PFA. Following blocking with 1% BSA in PBS-CM, cells were treated with

ANTI-FLAG® M2-HRP antibody (1:4000) for one hour. After two PBS- CM and two PBS washes, 100 µl of SIGMAFAST OPD solution was added to each well. In order to stop the reaction, after 10 min, 25 µl of 3 M HCl was added. Then a Synergy 2 microplate reader (Bio-Tek) was used to read the plate at an absorbance of 492 nm. The non-specific signal from pcDNA-transfected cells was deducted to acquire a specific signal.

Intact cell radioligand binding

[^{125}I]-AngII was prepared with the iodogen method as previously described[63]. AngII binding and dose displacements were done using [^{125}I]-AngII as a tracer. HEK293 and $G\alpha_{q/11}$ KO[64] cells were seeded at a density of 1 × 10^6 cells per 100 mm dish, and the next day, cells were transfected with 3 µg of AT1R along with 3µg of empty pcDNA using PEI methods (DNA: PEI, 1:2-2.5 ratio). The following day, cells were detached using TrypLE and re-seeded onto poly-L-ornithine-coated 24-well plate at a density of 1.5-2x10^5 cells per well. The next day, cells in a 24-well plate were washed once with PBS, then incubated in binding buffer (50 mM Tris, 100 mM NaCl, 5 mM $MgCl_2$, 0.2% BSA, pH 7.2). Dose displacement experiments were done by incubating the cells for 1 h at room temperature with ~ 0.1 nM ^{125}I-AngII as tracer and increasing concentrations of Q2 or AngII in the absence (DMSO) or presence of 200 µM of Q2. Q2 was preincubated at 37 °C for 5–10 minutes, followed by cooling to room temperature for 10 min before the addition of AngII or ^{125}I-AngII. After a 1-hour incubation period, the cells were washed three times with ice-cold PBS to remove any unbound radioactivity. Cells were solubilized in 0.5 M NaOH/0.05% SDS, and radioactivity was counted using a Wizard 1470 automatic γ-counter.

Data analysis

Data were normalized to the maximal response (Emax) of AngII for the WT receptor or WT receptor in control (vehicle only) treatment. $LogEC_{50}$ and Emax were derived from BRET concentration-response curves using GraphPad Prism 10 software (GraphPad Software Inc.). Curves represent the best fits. The data from a typical BRET experiment followed a normal probability distribution, and three experiments were considered sufficient to accurately estimate the SD. Student's *t* tests were therefore used to compare differences in the $logEC_{50}$, Emax, and log ($Emax/EC_{50}$) of each mutant to WT in the same pathway, and RAs. One-way ANOVA tests were used for comparisons of effects of different compounds on WT and Q2 effects on different mutant receptors.

References and Notes

1. Hilger, D. The role of structural dynamics in GPCR-mediated signaling. *FEBS J.* **288**, 2461–2489 (2021).
2. Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B. & Gloriam, D. E. Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discov.* **16**, 829–842 (2017).
3. Sriram, K. & Insel, P. A. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol. Pharmacol.* **93**, 251–258 (2018).
4. Vaidehi, N. & Kenakin, T. The role of conformational ensembles of seven transmembrane receptors in functional selectivity. *Curr. Opin. Pharmacol.* **10**, 775–781 (2010).
5. Ma, N., Nivedha, A. K. & Vaidehi, N. Allosteric communication regulates ligand-specific GPCR activity. *FEBS J.* **288**, 2502–2512 (2021).
6. Latorraca, N. R., Venkatakrishnan, A. J. & Dror, R. O. GPCR Dynamics: Structures in Motion. *Chem. Rev.* **117**, 139–155 (2017).
7. Thaker, T. M., Kaya, A. I., Preininger, A. M., Hamm, H. E. & Iverson, T. M. Allosteric mechanisms of G protein-Coupled Receptor signaling: a structural perspective. *Methods Mol. Biol. Clifton NJ* **796**, 133–174 (2012).
8. Bock, A. & Bermudez, M. Allosteric coupling and biased agonism in G protein-coupled receptors. *FEBS J.* **288**, 2513–2528 (2021).
9. Kenakin, T. Theoretical Aspects of GPCR–Ligand Complex Pharmacology. *Chem. Rev.* **117**, 4–20 (2017).
10. DeVree, B. T. *et al.* Allosteric coupling from G protein to the agonist-binding pocket in GPCRs. *Nature* **535**, 182–186 (2016).
11. Dror, R. O. *et al.* Structural basis for modulation of a G-protein-coupled receptor by allosteric drugs. *Nature* **503**, 295–299 (2013).
12. Moore, M. N. *et al.* Design of allosteric modulators that change GPCR G protein subtype selectivity. 2024.11.20.624209 Preprint at https://doi.org/10.1101/2024.11.20.624209 (2024).
13. Hollingsworth, S. A. *et al.* Cryptic pocket formation underlies allosteric modulator selectivity at muscarinic GPCRs. *Nat. Commun.* **10**, 3289 (2019).
14. Aleksandrov, A. & Simonson, T. Molecular Dynamics Simulations Show That Conformational Selection Governs the Binding Preferences of Imatinib for Several Tyrosine Kinases. *J. Biol. Chem.* **285**, 13807–13815 (2010).

15. Wu, N., Barahona, M. & Yaliraki, S. N. Allosteric communication and signal transduction in proteins. *Curr. Opin. Struct. Biol.* **84**, 102737 (2024).
16. Petrizzelli, F. *et al.* Connecting the dots: A practical evaluation of web-tools for describing protein dynamics as networks. *Front. Bioinforma.* **2**, (2022).
17. Johnson, Q. R., Lindsay, R. J. & Shen, T. CAMERRA: An analysis tool for the computation of conformational dynamics by evaluating residue–residue associations. *J. Comput. Chem.* **39**, 1568–1578 (2018).
18. Doshi, U., Holliday, M. J., Eisenmesser, E. Z. & Hamelberg, D. Dynamical network of residue–residue contacts reveals coupled allosteric effects in recognition, catalysis, and mutation. *Proc. Natl. Acad. Sci.* **113**, 4735–4740 (2016).
19. Boonserm, P. *et al.* Allosteric Signal within the Receptor-Binding Domain of the SARS-CoV-2 Spike Protein Mediated by a Class 3 Monoclonal Antibody Revealed through Molecular Dynamics Simulations and Protein Residue Networks. *ACS Omega* **9**, 4684–4694 (2024).
20. Feizi, S., Marbach, D., Médard, M. & Kellis, M. Network deconvolution as a general method to distinguish direct dependencies in networks. *Nat. Biotechnol.* **31**, 726–733 (2013).
21. Zhu, J., Wang, J., Han, W. & Xu, D. Neural relational inference to learn long-range allosteric interactions in proteins from molecular dynamics simulations. *Nat. Commun.* **13**, 1661 (2022).
22. Xie, J., Zhang, W., Zhu, X., Deng, M. & Lai, L. Coevolution-based prediction of key allosteric residues for protein function regulation. *eLife* **12**, e81850 (2023).
23. Hedderich, J. B. *et al.* The pocketome of G-protein-coupled receptors reveals previously untargeted allosteric sites. *Nat. Commun.* **13**, 2567 (2022).
24. Meller, A. *et al.* Predicting locations of cryptic pockets from single protein structures using the PocketMiner graph neural network. *Nat. Commun.* **14**, 1177 (2023).
25. Zhang, S., Mishra, P., Kelly, D., Kumar, R. & Bowman, G. R. AE-PocketMiner Uses Attention to Simultaneously Predict Cryptic Pockets and Their Allosteric Coupling. *bioRxiv* 2026.05.21.726899 (2026) doi:10.64898/2026.05.21.726899.
26. Gogoshin, G., Boerwinkle, E. & Rodin, A. S. New Algorithm and Software (BNOmics) for Inferring and Visualizing Bayesian Networks from Heterogeneous Big Biological and Genetic Data. *J. Comput. Biol. J. Comput. Mol. Cell Biol.* **24**, 340–356 (2017).

27. Mukhaleva, E. *et al.* BaNDyT: Bayesian Network Modeling of Molecular Dynamics Trajectories. *J. Chem. Inf. Model.* https://doi.org/10.1021/acs.jcim.4c01981 (2025) doi:10.1021/acs.jcim.4c01981.
28. Gogoshin, G. & Rodin, A. S. Minimum uncertainty as Bayesian network model selection principle. *BMC Bioinformatics* **26**, 100 (2025).
29. Gogoshin, G., Branciamore, S. & Rodin, A. S. Synthetic data generation with probabilistic Bayesian Networks. *Math. Biosci. Eng. MBE* **18**, 8603–8621 (2021).
30. Mukhaleva, E. *et al.* Bayesian network models identify cooperative GPCR:G protein interactions that contribute to G protein coupling. *J. Biol. Chem.* **300**, 107362 (2024).
31. Manglik, A., Wingler, L. M., Rockman, H. A. & Lefkowitz, R. J. β-Arrestin–Biased Angiotensin II Receptor Agonists for COVID-19. *Circulation* **142**, 318–320 (2020).
32. Lymperopoulos, A., Borges, J. I., Carbone, A. M., Cora, N. & Sizova, A. Cardiovascular angiotensin II type 1 receptor biased signaling: Focus on non-Gq-, non-βarrestin-dependent signaling. *Pharmacol. Res.* **174**, 105943 (2021).
33. Dasgupta, C. & Zhang, L. Angiotensin II receptors and drug discovery in cardiovascular disease. *Drug Discov. Today* **16**, 22–34 (2011).
34. Zhang, D. *et al.* Structural insights into angiotensin receptor signaling modulation by balanced and biased agonists. *EMBO J.* **42**, e112940 (2023).
35. Zhang, H. *et al.* Structure of the Angiotensin Receptor Revealed by Serial Femtosecond Crystallography. *Cell* **161**, 833–844 (2015).
36. Cao, Y. *et al.* Unraveling allostery within the angiotensin II type 1 receptor for Gαq and β-arrestin coupling. *Sci. Signal.* **16**, eadf2173 (2023).
37. Rosenblueth, E., Karmeshu, & Han Ping Hong. Maximum entropy and discretization of probability distributions. *Probabilistic Eng. Mech.* **2**, 58–63 (1987).
38. Nivedha, A. K. *et al.* Decoding Angiotensin II Type 1 Receptor Allosteric Communication to Gq and β-arrestin. 2021.05.17.444454 Preprint at https://doi.org/10.1101/2021.05.17.444454 (2021).
39. Wingler, L. M., McMahon, C., Staus, D. P., Lefkowitz, R. J. & Kruse, A. C. Distinctive Activation Mechanism for Angiotensin Receptor Revealed by a Synthetic Nanobody. *Cell* **176**, 479-490.e12 (2019).
40. Li, H., Kasam, V., Tautermann, C. S., Seeliger, D. & Vaidehi, N. Computational Method To Identify Druggable Binding Sites That Target Protein–Protein Interactions. *J. Chem. Inf. Model.* **54**, 1391–1400 (2014).

41. Avet, C. *et al.* Effector membrane translocation biosensors reveal G protein and βarrestin coupling profiles of 100 therapeutically relevant GPCRs. *eLife* **11**, e74101 (2022).
42. de Souza Neto, L. R. *et al.* In silico Strategies to Support Fragment-to-Lead Optimization in Drug Discovery. *Front. Chem.* **8**, (2020).
43. Liu, X. *et al.* Mechanism of intracellular allosteric β2AR antagonist revealed by X-ray crystal structure. *Nature* **548**, 480–484 (2017).
44. Liu, X. *et al.* Mechanism of β2AR regulation by an intracellular positive allosteric modulator. *Science* **364**, 1283–1287 (2019).
45. Valant, C., Felder, C. C., Sexton, P. M. & Christopoulos, A. Probe Dependence in the Allosteric Modulation of a G Protein-Coupled Receptor: Implications for Detection and Validation of Allosteric Ligand Effects. *Mol. Pharmacol.* **81**, 41–52 (2012).
46. Cheng, L. *et al.* Orthosteric ligand selectivity and allosteric probe dependence at Hydroxycarboxylic acid receptor HCAR2. *Signal Transduct. Target. Ther.* **8**, 364 (2023).
47. Cooper, S. L. *et al.* Probe dependence of allosteric enhancers on the binding affinity of adenosine A1 -receptor agonists at rat and human A1 -receptors measured using NanoBRET. *Br. J. Pharmacol.* **176**, 864–878 (2019).
48. Ahn, S. *et al.* Allosteric "beta-blocker" isolated from a DNA-encoded small molecule library. *Proc. Natl. Acad. Sci.* **114**, 1708–1713 (2017).
49. Glukhova, A. *et al.* Rules of Engagement: GPCRs and G Proteins. *ACS Pharmacol. Transl. Sci.* **1**, 73–83 (2018).
50. Wingler, L. M. *et al.* Angiotensin Analogs with Divergent Bias Stabilize Distinct Receptor Conformations. *Cell* **176**, 468-478.e11 (2019).
51. Wingler, L. M. *et al.* Angiotensin and biased analogs induce structurally distinct active conformations within a GPCR. *Science* **367**, 888–892 (2020).
52. Nivedha, A. K., Lee, S. & Vaidehi, N. Biased agonists differentially modulate the receptor conformation ensembles in Angiotensin II type 1 receptor. *J. Mol. Graph. Model.* **118**, 108365 (2023).
53. Vaidehi, N. & Bhattacharya, S. Allosteric Communication Pipelines in G-protein-coupled receptors. *Curr. Opin. Pharmacol.* **30**, 76–83 (2016).
54. Hollingsworth, S. A. *et al.* Cryptic pocket formation underlies allosteric modulator selectivity at muscarinic GPCRs. *Nat. Commun.* **10**, 3289 (2019).

55. Boora, T. S. *et al.* Allosteric pathways govern Gα protein coupling selectivity at promiscuous GPCRs. 2026.08.28.747882 Preprint at https://doi.org/10.64898/2026.08.28.747882 (2026).
56. Leach, K., Davey, A. E., Felder, C. C., Sexton, P. M. & Christopoulos, A. The Role of Transmembrane Domain 3 in the Actions of Orthosteric, Allosteric, and Atypical Agonists of the M4 Muscarinic Acetylcholine Receptor. *Mol. Pharmacol.* **79**, 855–865 (2011).
57. Nawaratne, V., Leach, K., Felder, C. C., Sexton, P. M. & Christopoulos, A. Structural Determinants of Allosteric Agonism and Modulation at the M4 Muscarinic Acetylcholine Receptor. *J. Biol. Chem.* **285**, 19012–19021 (2010).
58. Huang, J. & MacKerell, A. D. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).
59. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).
60. Abraham, M. J. *et al.* GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **1–2**, 19 (2015).
61. Namkung, Y. *et al.* Monitoring G protein-coupled receptor and β-arrestin trafficking in live cells using enhanced bystander BRET. *Nat. Commun.* **7**, 12178 (2016).
62. Heydenreich, F. M., Miljuš, T., Milić, D. & Veprintsev, D. B. High-throughput Site-directed Scanning Mutagenesis Using a Two-fragment PCR Approach. *Bio-Protoc.* **10**, e3484 (2020).
63. Laporte, S. A. *et al.* The tyrosine within the NPXnY motif of the human angiotensin II type 1 receptor is involved in mediating signal transduction but is not essential for internalization. *Mol. Pharmacol.* **49**, 89–95 (1996).
64. Schrage, R. *et al.* The experimental power of FR900359 to study Gq-regulated biological processes. *Nat. Commun.* **6**, 10156 (2015).

Acknowledgments

We thank Dr. Supriyo Bhattacharya for his valuable input on this work. This work was funded by grants from the National Institutes of Health grants R01-GM117923, R35 GM156498 to N.V., NLM R01LM013138 grant to A.S.R., and NIH NLM R01-LM013876 to N.V., A.S.R and S.B., Dr. Susumu Ohno Chair in Theoretical Biology (held by A.S.R.), Canadian Institutes of Health Research (PJT-162368 and PJT-173504 to S.A.L.)

Author contributions:

HC developed the DBNM framework, and performed computational analyses including ΔWd hotspot identification, virtual ligand screening for candidate allosteric modulators, and MD simulations of mutant AT1R systems to elucidate the mechanism of Q2 modulation. YN performed experimental evaluation of small-molecule effects on WT and mutant AT1R signaling. ZAJ experimentally tested receptor mutations predicted from the DBNM hotspot analysis. WJCvdV performed the WT AT1R MD simulations. EM contributed to early methodological development. GK developed the robustness test for the networks. HC, YN and ZAJ prepared the figures and wrote the manuscript with input from all authors. HC, YN, EM, ASR, SB, SAL and NV conceptualized the study. ASR, SB, SAL, and NV supervised the project and secured funding. All authors discussed the results and approved the final manuscript.

Competing interests:

The authors declare no competing interests.

Data and materials availability:

The molecular MD simulation snapshots for all the simulations in this work are deposited in the GPCRMD database. The code for running the BN models is available in GitHub: https://github.com/bandyt-group/bandyt. Networks for active- and inactive-state AT1R provided as graphml files as supplementary material. Experimental data have been provided to the journal as attachments.

Supplementary Materials

In a separate file

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- NetworkS2ActiveAT1Rinteractionenergygraph.xml
- NetworkS4InactiveAT1Rinteractionenergygraph.xml
- NetworkS1ActiveAT1Rinteractionenergy100.unk
- NetworkS3InactiveAT1Rinteractionenergy100.unk
- SupplementaryData3PrimersequencesforAT1R.ods
- SupplementaryData1BayesianNetworkMetricsWeightedDegree.xlsx
- RevisedSI.docx
- SupplementaryData2DeepMutationalScanningdatatopquartile.xlsx