

Exploratory molecular-dynamics comparison of wild-type, N97S, and I52V calmodulin–CaV1.2 IQ-domain complexes

Sahil Khandekar

sahil04mail@gmail.com

University of Mumbai <https://orcid.org/0009-0003-4158-3820>

Research Article

Keywords: calmodulin, CaV1.2, IQ domain, molecular dynamics, GROMACS, CALM1, structural fluctuation, solvent-accessible surface area, trajectory analysis

Posted Date: October 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11441900/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Background: Calmodulin (CaM) binds the IQ motif in the C-terminal region of the cardiac L-type calcium channel CaV1.2 and contributes to calcium-dependent channel regulation. The structural consequences of CALM1 missense variants can be explored by molecular dynamics (MD), although trajectory descriptors alone cannot establish clinical pathogenicity or binding affinity.

Methods: The CaM–CaV1.2 IQ-domain crystal structure (PDB 2F3Y) was used to prepare wild-type, N97S (current reference nomenclature p.Asn98Ser), and I52V (p.Ile53Val) models. One 100-ns GROMACS trajectory was analysed for each model using the AMBER99SB-ILDN force field, a TIP3P water model, a nominal 0.15 M NaCl concentration, and a stated 310 K simulation temperature. Analyses included backbone root-mean-square deviation (RMSD), Ca root-mean-square fluctuation (RMSF), solvent-accessible surface area (SASA), an SASA-derived buried-area term (Δ SASA), radius of gyration (Rg), and inter-chain polar heavy-atom proximity counts.

Results: Mean overall-complex RMSD values were 0.250, 0.211, and 0.238 nm for wild type, N97S, and I52V, respectively. Mean Rg values were closely grouped (1.593, 1.603, and 1.603 nm). Mean Δ SASA was 29.60, 28.69, and 28.99 nm², while mean Ca RMSF across residues was 0.146, 0.110, and 0.127 nm. The mean distance-based polar-pair count was lower in N97S (13.98) than wild type (17.09), but the method does not validate hydrogen-bond geometry and depends on fixed atom-index selections.

Conclusions: The sampled trajectories show modest differences in global deviation, residue fluctuation, interface-associated SASA, and distance-based contacts, while average complex compactness is similar. Because each condition is represented by a single trajectory, the statistics are descriptive rather than inferential. Metal-ion handling and atom-selection provenance also require verification. The findings are hypothesis-generating and do not establish changes in calcium affinity, channel function, binding free energy, or clinical classification.

1. Introduction

Voltage-gated L-type calcium channels translate membrane depolarization into calcium entry, a signal that contributes to cardiac excitation–contraction coupling and other calcium-dependent cellular processes. CaV1.2 channel activity is regulated by calcium-dependent inactivation (CDI) and calcium-dependent facilitation (CDF), feedback processes that depend on the channel-associated calcium sensor calmodulin (CaM). CaM interacts with intracellular regions of the CaV1.2 pore-forming α 1C subunit, including the C-terminal IQ motif. The architecture and dynamics of this interaction are relevant to how local molecular recognition can influence calcium-current regulation [1–3].

The crystal structure deposited as PDB 2F3Y provides a high-resolution view of calcium-saturated CaM bound to a 21-residue CaV1.2 IQ-domain peptide. The structure was resolved by X-ray diffraction at 1.45 Å and contains a 148-residue CaM chain and the channel-derived peptide chain [1]. Structural work on CaM–CaV1.2 complexes has described hydrophobic interactions and aromatic anchor residues at the

interface, as well as differences between CaM lobes in their contacts with the IQ motif [1,2]. The complex is therefore a reasonable starting template for a focused, explicitly limited comparison of variant-associated changes in sampled conformations and interface descriptors.

Variants in CALM1, CALM2, and CALM3 can cause calmodulinopathies associated with cardiac arrhythmia. The substitution historically referred to as N97S corresponds to p.Asn98Ser in the current NM_006888.6 CALM1 reference sequence. Clinical database records classify p.Asn98Ser as pathogenic/pathogenic-likely pathogenic, and published work has associated CALM1 variants with ventricular tachycardia and sudden cardiac death [4,5]. These clinical observations motivate structural investigation, but a simulation of a small CaM–IQ construct is not a direct model of the full channel or of a patient phenotype.

The comparison substitution historically described as I52V maps to p.Ile53Val in the current CALM1 reference sequence. The current ClinVar record lists p.Ile53Val as a variant of uncertain significance (VUS), supported by a single clinical-laboratory submission [6]. It is therefore not a clinically established benign control. In the present study, I52V is described as a comparison variant, and no benign classification is inferred from the MD outputs.

All-atom MD simulation can provide time-dependent descriptors such as backbone RMSD, residue-level RMSF, solvent-accessible surface area, compactness, and atomic contacts. However, the meaning of any one descriptor is limited: a lower RMSD does not by itself establish greater thermodynamic stability, a lower RMSF does not establish improved function, and a change in buried area does not directly quantify a binding free-energy difference. Sampling limitations can be severe in biomolecular systems, and results from one trajectory may be sensitive to initial conditions and pathway selection [7,8]. Rigorous inference normally requires independent replicas, careful uncertainty estimation, and clearly validated definitions of contacts and energetic observables [7,8].

The aim of this exploratory study was to compare one 100-ns trajectory each for the wild-type 2F3Y-derived complex and two CaM substitutions, N97S/p.Asn98Ser and I52V/p.Ile53Val. The analysis used five categories of descriptors: RMSD, RMSF, SASA/ Δ SASA, Rg, and inter-chain polar-atom proximity. The intention was not to rank clinical risk, estimate calcium affinity, or predict channel physiology, but to document the observed trajectory-level differences and identify methodological checks necessary for a reproducible follow-up study.

2. Materials and Methods

2.1. Study design and reporting scope

This was a retrospective analysis of three completed MD trajectories initiated from a common crystallographic template. The systems were wild type, N97S/p.Asn98Ser, and I52V/p.Ile53Val, with one 100-ns production trajectory for each condition (300 ns of cumulative simulated time). The three

trajectories represent different model systems, not replicate runs of a single system. No hypothesis-testing statistics or p-values were calculated.

Quantitative summaries were recalculated from the trajectory CSV exports supplied for this project. The consolidated comparative figure was used as a cross-check for the intended system labels and trajectory profiles. Supplementary Table S1 records the CSV filenames used for each reported descriptor; this file-to-system mapping must be checked against the original trajectory folders before public posting.

2.2. Structural template and variant nomenclature

The starting coordinate template was PDB 2F3Y, an X-ray structure of a Ca²⁺-calmodulin/CaV1.2 IQ-domain complex. RCSB PDB annotates CaM as chain A (148 residues) and the channel-derived IQ peptide as chain B (21 residues); the entry reports a resolution of 1.45 Å [1,9]. Variant coordinates were generated by substituting amino acids in the template in PyMOL. The wild-type input was identified by the author as 2F3Y_WT.pdb. The N97S and I52V labels refer to the structural-model numbering used in the project; for clinical and sequence reporting, they map to CALM1 p.Asn98Ser and p.Ile53Val, respectively [5,6].

The variant-residue mapping should be documented in a sequence-alignment supplement before final posting. This is particularly important because structural residue identifiers, historical variant names, and reference-transcript HGVS numbering are not necessarily identical. Residue-specific mechanistic claims are not made here.

2.3. Force field, solvation, and ion concentration

The reported system-preparation commands used GROMACS and the AMBER99SB-ILDN protein force field selected by `—ffamber99sb —ildn`, with the TIP3P water model selected by `—watertip3p` [10,11]. The complex was centered in a dodecahedral periodic box with a minimum solute-to-box distance of 1.0 nm using `gmxditconf`, followed by solvation with the coordinate file `spc216.gro` using `gmxsolvate`. Sodium and chloride ions were added with `gmx ≥ nion` for neutralization and a nominal salt concentration of 0.15 M. The preparation commands used Particle Mesh Ewald (PME) electrostatics and 1.0-nm Coulomb and van der Waals cutoffs in the supplied MDP definitions [12].

PDB 2F3Y contains crystallographically modeled calcium and magnesium ions [1,9]. The command transcript available for this manuscript does not document a complete metal-ion retention and parameterization procedure, and the supplied input does not establish whether the crystallographic ions remained in the final production topology. We therefore describe the simulations as 2F3Y-derived CaM–IQ models and do not interpret them as a verified calcium-bound ion-coordination simulation. Before a definitive submission, the original coordinate files, `Topol.top`, `ions.gro`, `production.tpr`, and GROMACS log files should be inspected to establish the exact ion composition and parameters.

2.4. Energy minimization and equilibration

Energy minimization used steepest descent for up to 50,000 steps with $emstep = 0.01$ and $em \rightarrow l = 1000.0$ (GROMACS energy-gradient units). The supplied command then performed 100-ps NVT equilibration with a 4-fs integration step ($dt = 0.004$, $nsteps = 25000$), positional restraints ($def \in e = -DPOSRES$), the velocity-rescaling thermostat (V-rescale), and a target temperature of 310 K [13]. The thermostat time constant was set to 0.1 ps for both $Prote \in$ and $Non - Prote \in$ groups.

The subsequent 100-ps NPT equilibration used the same 4-fs time step, positional restraints, target temperature of 310 K, and isotropic pressure coupling. The provided MDP selected the C-rescale barostat with a 2-ps pressure-coupling time constant, target pressure of 1 bar, and compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$. Production employed the Parrinello–Rahman barostat with isotropic pressure coupling, a 2-ps pressure-coupling time constant, and the same target pressure and compressibility [14]. Hydrogen virtual-site generation was requested through $gmxdbs2gmx - vsitehydro \geq ns$; all bonds were constrained using LINCS. The 4-fs time step is reported as configured, not as an independently validated stability optimum. The underlying energy, temperature, pressure, and constraint logs should be checked before resubmission.

2.5. Production molecular dynamics

Each system was simulated for 25,000,000 production steps at a 4-fs time step, corresponding to 100 ns per system. The production MDP used $\int egra \rightarrow r = md, nstxout - compressed = 2500$, $nste \neq rgy = 2500$, $nstlog = 5000$, and $nstcalce \neq rgy = 100$. A 2,500-step coordinate-output interval corresponds to 10 ps at the stated time step. The run command requested GPU acceleration for short-range nonbonded interactions and PME, with six OpenMP threads and CPU pinning ($-nbgpu - \pm egpu - n \rightarrow mp6 - \pi non$).

The exact GROMACS version, GPU model, precision mode, operating environment, random seeds/initial velocities, and full log outputs were not provided in the material available for this draft. These details should be added from the original logs or run metadata if recoverable; they are not inferred here.

2.6. Trajectory processing and analysis software

Trajectories were processed with $gmxconvert - tpr$ and $gmxtxtrjconv$ to retain protein atoms, remove periodic-boundary jumps ($- pbcnojump$), center the protein ($- center - pbcmol$), and fit the trajectory by rotation and translation ($- fitrot + trans$) using the backbone as the fitting group. The output topology and trajectory used by the Python scripts were named $md_{prote \in o} nly. tpr$ and $md_c \leq an. xtc$. Python analyses used MDAnalysis along with NumPy, pandas, and Matplotlib [15]. The

package versions were not recorded in the supplied scripts and should be recovered for a versioned reproducibility record.

The RMSD, Rg, and polar-atom proximity scripts sampled every tenth saved trajectory frame. Given the 10-ps production coordinate interval, this corresponds to an approximately 100-ps sampling interval for these time-series metrics. The SASA output contained 10,001 time points over 100 ns, corresponding to the 10-ps coordinate interval. RMSF was calculated over the available fitted trajectory for the selected Ca atoms. Moving averages were used only for visualization and not as substitutes for the raw data in the descriptive statistics.

2.7. Structural descriptors

2.7.1. Backbone RMSD

Backbone RMSD was calculated relative to the first trajectory frame by superposing the selected backbone atoms and then evaluating the root mean square displacement. Overall-complex RMSD and chain-specific RMSD values were recorded in nanometers. RMSD is a measure of deviation from the chosen reference under the selected fitting protocol; it is not a direct free-energy or stability measurement. The computation used the RMSD implementation in MDAnalysis [15,16].

2.7.2. Ca RMSF

Per-residue RMSF was calculated for protein Ca atoms on the processed trajectory and converted from ångströms to nanometers by division by 10. The scripts use the topology residue numbering for the plotted x-axis and show a boundary after residue 148 to distinguish the CaM region from the IQ-peptide region. The current PDB annotation states 164 modeled residues across the two chains despite 169 deposited sequence positions [9]; unresolved residues and topology numbering must be considered when mapping individual RMSF peaks to biological residue numbers. RMSF describes positional fluctuation about a time-average position after alignment and should not be interpreted as a direct measure of functional impairment [15].

2.7.3. Radius of gyration

Radius of gyration was calculated with the MDAnalysis $A \rightarrow mGroup.radius_{ofgyration}$ method for the overall protein complex and individual chains, then converted from ångströms to nanometers. Rg summarizes the mass-distribution compactness of the selected atoms and does not by itself reveal whether an interface is functionally intact or energetically favorable.

2.7.4. Solvent-accessible surface area and buried area

Solvent-accessible surface area (SASA) was calculated with *gmxsasa* for the protein complex and individual chains. The analysis script calculated a combined buried surface term as $\Delta\text{SASA} = \text{SASA}(\text{chain A}) + \text{SASA}(\text{chain B}) - \text{SASA}(\text{complex})$. We retain that exact formula and call the resulting value “combined buried SASA” or “ ΔSASA ” to distinguish it from interface-area conventions that divide this difference by two. This distinction is important when comparing values across studies [17]. SASA and ΔSASA are geometric descriptors; neither alone gives a binding free energy.

2.7.5. Inter-chain polar heavy-atom proximity

The supplied Python script selected atoms whose names began with N, O, or S in two hard-coded atom-index ranges and counted every cross-chain pair with a distance of $\leq 3.5 \text{ \AA}$. Because the script does not identify donor and acceptor roles, hydrogen atoms, angular geometry, or salt-bridge chemistry, this metric is reported as an inter-chain polar heavy-atom proximity count. It is not reported as a validated hydrogen-bond count or salt-bridge count. The fixed atom-index ranges should be checked against each variant topology because amino-acid substitutions can change atom counts and therefore shift atom indices. Geometrically validated hydrogen-bond analysis would require explicit donor–acceptor selections and distance/angle criteria [18].

2.8. Descriptive statistics and uncertainty

For time-series descriptors, reported summaries are the arithmetic mean and sample standard deviation across sampled trajectory frames. For RMSF, the mean and standard deviation are calculated across the residue positions represented in the CSV. These standard deviations summarize within-trajectory temporal variation or between-residue variation; they are not uncertainty estimates across independent MD replicas. Frame-to-frame observations are temporally correlated, so no standard errors, confidence intervals, or hypothesis-test p-values are reported. This descriptive approach follows general guidance that simulation uncertainty and sampling limitations should be explicitly communicated [7,8].

3. Results

3.1. Overview of the compared trajectories

All three production trajectories extended to approximately 100 ns, and the comparative time series showed distinct temporal profiles despite broadly similar mean complex radii of gyration. Summary values below were calculated from the CSV exports supplied by the author. Because each model is represented by one run, differences should be read as properties of the sampled trajectories, not as population-level estimates or statistical evidence of mutation effects.

The descriptive summary of the five principal trajectory descriptors is shown in Table 1. Time-averaged means should be interpreted alongside the corresponding time series: wild-type RMSD rose noticeably

late in the trajectory, whereas the N97S RMSD profile was less elevated in the final part of the run. Such profile differences may reflect different conformational paths, but without replicas and structural inspection they do not establish the cause or biological significance of those paths.

Table 1. Descriptive summaries of the principal trajectory metrics.

Descriptor	Wild type	N97S (p.Asn98Ser)	I52V (p.Ile53Val)
Complex backbone RMSD (nm)	0.250 ± 0.083	0.211 ± 0.032	0.238 ± 0.043
CaM-chain backbone RMSD (nm)	0.233 ± 0.081	0.196 ± 0.028	0.229 ± 0.042
IQ-chain backbone RMSD (nm)	0.131 ± 0.031	0.124 ± 0.031	0.141 ± 0.043
Mean Ca RMSF (nm; mean ± residue SD)	0.146 ± 0.064	0.110 ± 0.059	0.127 ± 0.054
Overall complex Rg (nm)	1.593 ± 0.012	1.603 ± 0.013	1.603 ± 0.013
Complex SASA (nm ²)	95.158 ± 2.250	97.068 ± 2.266	95.026 ± 2.177
Combined buried SASA, ΔSASA (nm ²)	29.601 ± 1.283	28.688 ± 1.172	28.986 ± 1.007
Polar heavy-atom proximity pairs ≤3.5 Å	17.09 ± 3.21	13.98 ± 3.21	16.29 ± 4.10

Note: Values are mean ± standard deviation across frames except Ca RMSF, where the standard deviation is across residue positions. These are descriptive within-run summaries, not independent-replicate uncertainties. The polar-pair metric is not a geometrically validated hydrogen-bond count. ΔSASA is the sum of chain SASAs minus complex SASA and is not divided by two.

3.2. Backbone deviation and global compactness

The mean overall-complex RMSD was 0.250 ± 0.083 nm for wild type, 0.211 ± 0.032 nm for N97S, and 0.238 ± 0.043 nm for I52V (Table 1, Figure 2A). In the supplied composite plot, the wild-type RMSD trace rises after roughly 65 ns and reaches its highest values during the final 15 ns. N97S remains comparatively level across the later trajectory, whereas I52V exhibits a smaller late increase. These differences describe deviations from the selected first-frame reference and may include conformational rearrangement; they should not be equated with thermodynamic stability or peptide dissociation.

The average overall Rg values were 1.593 ± 0.012 nm for wild type, 1.603 ± 0.013 nm for N97S, and 1.603 ± 0.013 nm for I52V (Table 1, Figure 2B). The two variant means were approximately 0.010 nm above the wild-type mean, a small difference relative to the observed temporal variation. Accordingly, the global compactness summary does not indicate a large difference among the sampled complexes. It

does not exclude local changes at the CaM–IQ interface, and it cannot determine whether the peptide remains in the same interaction register throughout the trajectory.

3.3. Solvent exposure and interface-associated area

The mean complex SASA was $95.158 \pm 2.250 \text{ nm}^2$ for wild type, $97.068 \pm 2.266 \text{ nm}^2$ for N97S, and $95.026 \pm 2.177 \text{ nm}^2$ for I52V. N97S therefore had a mean complex SASA approximately 1.91 nm^2 above wild type, while the I52V mean was close to wild type. These area differences reflect the sampled solvent-exposed surfaces, and no single residue-level cause was established by the analysis.

The combined buried SASA term (ΔSASA) was $29.601 \pm 1.283 \text{ nm}^2$ in wild type, $28.688 \pm 1.172 \text{ nm}^2$ in N97S, and $28.986 \pm 1.007 \text{ nm}^2$ in I52V. Relative to wild type, the mean ΔSASA values were approximately 3.1% lower for N97S and 2.1% lower for I52V. These modest differences are geometric comparisons from one trajectory per model and should not be translated into percentage changes in affinity. Since the reported term is not divided by two, it should be compared only with work using the same convention or converted consistently before cross-study comparisons.

3.4. Residue-level Ca fluctuations

The mean RMSF across the 164 Ca positions represented in the supplied per-residue tables was $0.146 \pm 0.064 \text{ nm}$ in wild type, $0.110 \pm 0.059 \text{ nm}$ in N97S, and $0.127 \pm 0.054 \text{ nm}$ in I52V (Table 1, Figure 3). The RMSF curves are heterogeneous, with peaks at several residue positions rather than a uniform shift across the sequence. The N97S dataset has a lower mean RMSF across the analyzed positions; this is a description of the fitted trajectory and does not itself demonstrate greater stability, reduced function, or clinical pathogenicity.

The plots label chain A as CaM through topology residue 148 and chain B as the IQ peptide after that boundary. RCSB reports 164 modeled residues but 169 deposited residues for the structure [9]. The correspondence between topology residue numbers and biological/reference sequence positions should therefore be checked before interpreting peaks as mutation-proximal effects. The current analysis does not identify statistically enriched flexible regions or perform a pre-specified mutation-site comparison.

3.5. Inter-chain polar heavy-atom proximity

The mean number of selected inter-chain N/O/S heavy-atom pairs separated by no more than $3.5 \text{ \AA}$ was 17.09 ± 3.21 for wild type, 13.98 ± 3.21 for N97S, and 16.29 ± 4.10 for I52V (Table 1, Figure 2D). The N97S count was approximately 18.2% lower than wild type, and the I52V count approximately 4.7% lower. The respective time series also show changing contact counts over the trajectory, rather than a constant interface-contact level.

These values must be interpreted conservatively for two reasons. First, the calculation counts any selected cross-chain polar heavy-atom pair inside a distance cutoff; it does not evaluate donor–hydrogen–acceptor alignment, so it cannot establish hydrogen bonding or salt bridging. Second, atom selections are specified by fixed atom-index intervals. Because residue substitutions can change side-chain atom counts, an identical numeric index interval is not guaranteed to select exactly the same chain atoms across all topologies. The observed count differences are therefore exploratory and require reanalysis with chain-ID and explicit chemical selections before they are used as an interface mechanism.

3.6. Chain-specific compactness and SASA

Chain-level means provide context for the complex summaries (Table 2). CaM-chain Rg means were tightly grouped near 1.655–1.657 nm across the three conditions. The IQ-peptide Rg means were lower in the two variant datasets than in wild type (0.917 nm versus 0.813 and 0.825 nm). This difference could reflect peptide conformational differences, residue mapping, selection composition, or trajectory-specific sampling; without inspection of aligned structures, it should not be ascribed to a specific physical cause. Chain SASA values indicate that the N97S model has the highest average complex SASA and peptide SASA, but differences remain descriptive.

Table 2. Chain-specific descriptive summaries.

Descriptor	Wild type	N97S (p.Asn98Ser)	I52V (p.Ile53Val)
CaM-chain Rg (nm)	1.656 ± 0.013	1.657 ± 0.012	1.655 ± 0.014
IQ-chain Rg (nm)	0.917 ± 0.016	0.813 ± 0.024	0.825 ± 0.014
CaM-chain SASA (nm ²)	100.203 ± 2.009	100.711 ± 1.803	99.827 ± 1.814
IQ-chain SASA (nm ²)	24.556 ± 0.712	25.045 ± 0.875	24.185 ± 0.812

Note: Chain-specific quantities depend on the topology selections and residue ranges used during analysis. Means ± frame-wise standard deviation are reported for the trajectory-derived time-series variables.

4. Discussion

This study provides a descriptive comparison of three 100-ns trajectories derived from the CaM–CaV1.2 IQ-domain structural template 2F3Y. The most consistent cross-system observation is that the average global radius of gyration remains similar among wild type, N97S, and I52V. At the same time, the supplied RMSD profiles follow different time-dependent paths, the two variant datasets have modestly lower mean ΔSASA than wild type, and the distance-based polar-pair metric is lower for N97S. These

observables are complementary but do not converge on a single, independently validated measure of binding strength or biological stability.

The structural framework is biologically relevant because the CaM–IQ interaction participates in CaV1.2 regulation. The 2F3Y and related structures establish that CaM lobes recognize hydrophobic features of the IQ peptide, and functional studies of the channel demonstrate that residues within the IQ region can influence CDI and CDF [1,2]. However, the present simulation includes only the CaM–IQ peptide complex, not the full CaV1.2 channel, membrane environment, intracellular channel domains, or a controlled calcium-binding titration. Its outputs cannot be used to infer electrophysiological effects or to reproduce the full disease mechanism of a CALM1 variant.

The N97S/p.Asn98Ser model shows a lower mean RMSD and lower mean RMSF than wild type, accompanied by lower average Δ SASA and fewer selected polar heavy-atom proximity pairs. One possible interpretation is that the sampled N97S trajectory occupies a distinct conformational path with a different pattern of inter-chain proximity. Yet this interpretation is not mechanistically resolved: lower RMSD can arise from remaining close to the chosen initial reference, lower RMSF can reflect sampling of a more restricted local ensemble, and a lower distance-based pair count does not necessarily mean fewer hydrogen bonds. These measurements therefore motivate reanalysis of local contacts and conformational states rather than supporting a conclusion that the N97S interface is definitively weakened or stabilized.

The I52V/p.Ile53Val data are intermediate for several summaries: the mean RMSD and RMSF fall between the wild-type and N97S values; mean Δ SASA is slightly below wild type; and the polar-pair count is also below wild type but above N97S. This profile should not be interpreted as benignity. The current ClinVar record classifies p.Ile53Val as a VUS [6], and a single simulation cannot reclassify a clinical variant. In particular, the similarity or difference of MD descriptors does not substitute for clinical evidence, functional experiments, segregation information, or an accepted variant-classification framework.

An additional issue is the treatment of calcium and magnesium ions. The experimental 2F3Y complex is annotated as calcium–calmodulin bound to the IQ peptide and contains crystallographic metal ions [1,9]. However, the supplied *pdb2gmx* command transcript does not show how non-protein ions were retained and parameterized in the simulation topology. If those ions were absent or inadequately parameterized, the simulated CaM conformational ensemble may not correspond to the experimentally observed calcium-saturated state. This is a potentially consequential reproducibility issue, not a cosmetic reporting detail. The original topology and run files should be inspected and the manuscript updated accordingly before a definitive scientific claim is made.

The contact analysis also requires a methodological upgrade. Inter-chain N/O/S heavy-atom pairs within 3.5 Å are a useful coarse proximity measure, but atom-level distance alone is not a conventional validation of hydrogen bonds. A rigorous hydrogen-bond analysis should use explicit donor and acceptor groups and a stated donor–acceptor distance and angular criterion; salt bridges should be defined

separately from charged side-chain groups. Moreover, the current fixed atom-index intervals may not define homologous selections in all three variants. Repeating the analysis with chain IDs, atom names, residue identities, and verified residue mapping would make the comparison chemically more interpretable.

Finally, statistical uncertainty cannot be resolved by calculating a standard deviation across hundreds or thousands of saved frames from one run. Consecutive frames are correlated and share the same initial conditions and trajectory history. Best-practice discussions of molecular simulations emphasize independent replicas and explicit sampling/uncertainty analysis when drawing comparative conclusions [7,8]. The current means and standard deviations describe the sampled runs, not uncertainty across possible runs. A reasonable follow-up would repeat each system from independent initial velocities, confirm stability of energy and pressure observables, compare replicate-level statistics, and then assess whether the differences in RMSD, R_g , $\Delta SASA$, residue fluctuations, and validated interface contacts are reproducible.

5. Limitations

Several limitations define the scope of the present work. First, each condition was represented by one 100-ns trajectory. The comparisons therefore cannot quantify between-replica variability and do not support inferential statistical testing. Any differences may reflect the sampled conformational paths rather than a reproducible consequence of the amino-acid substitution.

Second, the supplied commands do not establish whether the calcium and magnesium ions present in the crystallographic template were retained in the final production system, or how any retained metal ions were parameterized. Because CaM is calcium-responsive, the production topology and coordinate files must be inspected before the simulations are described as metal-bound or used to discuss metal-dependent mechanisms.

Third, the inter-chain polar-atom count is a distance-cutoff proximity metric rather than a validated hydrogen-bond or salt-bridge calculation. Fixed atom-index selections should be replaced with topology-aware chain and chemical selections and checked for identical chain definitions across variants. Fourth, residue-number mapping requires verification because the PDB contains fewer modeled residues than deposited sequence positions and the structural labels differ from current CALM1 reference nomenclature.

Fifth, exact software versions, complete log files, ion-topology records, and full input files were not part of the materials that could be verified for this draft. Finally, the current study does not include replicate-level convergence analysis, a validated binding free-energy protocol, calcium-affinity calculations, full-channel modelling, or electrophysiological data. Accordingly, mechanistic and clinical conclusions remain outside the scope of this analysis.

6. Conclusion

Three single 100-ns trajectories of a 2F3Y-derived CaM–CaV1.2 IQ-domain complex were compared using backbone RMSD, Ca RMSF, Rg, SASA-derived Δ SASA, and inter-chain polar heavy-atom proximity counts. Mean overall compactness was similar across systems, while the sampled variants showed modestly lower Δ SASA and different fluctuation/contact profiles relative to wild type. These observations are descriptive and hypothesis-generating. They do not establish changes in binding affinity, calcium coordination, channel function, or clinical pathogenicity. Before the findings can support stronger conclusions, the metal-ion treatment, variant-to-trajectory mapping, residue/atom selections, and independent-replica reproducibility need to be verified.

Declarations

Data and code availability

The project repository identified by the author is available at <https://github.com/sahilkhandekar04/calmodulin-cav1.2-md-simulation>. The repository should be versioned before submission and include the exact topology (*Topol.*) and included parameter files), processed coordinate and trajectory files or an appropriate archive, GROMACS logs, MDP files, the analysis scripts, per-system CSV outputs, a README with software versions, and a clear mapping from each trajectory to its variant. The numerical results reported here were calculated from CSV exports provided for manuscript preparation. The repository contents and revision identifier should be rechecked by the author before posting.

Author contributions

Sahil Ravindra Khandekar: project conception, variant model preparation, molecular-dynamics setup and execution, analysis scripting, data curation, visualization, and manuscript drafting. The author should confirm this statement reflects the actual work performed.

Funding

This work received no specific funding from public, commercial, or not-for-profit funding agencies.

Competing interests

The author declares no competing interests.

Ethics approval and consent to participate

Not applicable. The work consists of in silico analysis of a publicly available protein structure and public variant-annotation records; no new human-participant or animal data were collected.

Consent for publication

Not applicable.

Acknowledgements

None.

Supplementary information

Supplementary Information accompanies this manuscript and provides the reported GROMACS command sequence, MDP settings, analysis definitions, file-to-system mapping used for the current tables, and a submission reproducibility checklist. The supplementary file is a companion document and should be uploaded separately if the Research Square portal offers a supplementary-material field.

References

1. Fallon JL, Halling DB, Hamilton SL, Quirocho FA. Structure of calmodulin bound to the hydrophobic IQ domain of the cardiac Ca(v)1.2 calcium channel. *Structure*. 2005;13(12):1881–1886. <https://doi.org/10.1016/j.str.2005.09.021>. PDB entry 2F3Y: <https://doi.org/10.2210/pdb2F3Y/pdb>.
2. Van Petegem F, Chatelain FC, Minor DL Jr. Insights into voltage-gated calcium channel regulation from the structure of the CaV1.2 IQ domain–Ca²⁺/calmodulin complex. *Nat Struct Mol Biol*. 2005;12:1108–1115. <https://doi.org/10.1038/nsmb1027>.
3. Ben-Johny M, Yang PS, Niu J, Yang W, Joshi-Mukherjee R, Yue DT. Conservation of Ca²⁺/calmodulin regulation across Na and Ca²⁺ channels. *Cell*. 2014;157(7):1657–1670. <https://doi.org/10.1016/j.cell.2014.04.035>.
4. Nyegaard M, Overgaard MT, Søndergaard MT, Vranas M, Behr ER, Hildebrandt LL, et al. Mutations in calmodulin cause ventricular tachycardia and sudden cardiac death. *Am J Hum Genet*. 2012;91(4):703–712. <https://doi.org/10.1016/j.ajhg.2012.08.015>.
5. National Center for Biotechnology Information. ClinVar: NM_006888.6(CALM1):c.293A>G (p.Asn98Ser), pathogenic/likely pathogenic aggregate records. Variation ID 39758. <https://www.ncbi.nlm.nih.gov/clinvar/variation/39758/>. Accessed 20 October 2026.

6. National Center for Biotechnology Information. ClinVar: NM_006888.6(CALM1):c.157A>G (p.Ile53Val), uncertain significance. Variation ID 475380.
<https://www.ncbi.nlm.nih.gov/clinvar/variation/475380/>. Accessed 20 October 2026.
7. Grossfield A, Patrone PN, Roe DR, Schultz AJ, Siderius DW, Zuckerman DM. Best practices for quantification of uncertainty and sampling quality in molecular simulations. *Living J Comput Mol Sci*. 2018;1(1):5067. <https://doi.org/10.33011/livecoms.1.1.5067>.
8. Knapp B, Ospina L, Deane CM. Avoiding false positive conclusions in molecular simulation: the importance of replicas. *J Chem Theory Comput*. 2018;14(12):6127–6138.
<https://doi.org/10.1021/acs.jctc.8b00391>.
9. RCSB Protein Data Bank. 2F3Y: Calmodulin/IQ domain complex.
<https://www.rcsb.org/structure/2F3Y>. Accessed 20 October 2026.
10. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*. 2010;78(8):1950–1958.
<https://doi.org/10.1002/prot.22711>.
11. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983;79(2):926–935.
<https://doi.org/10.1063/1.445869>.
12. Darden T, York D, Pedersen L. Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. *J Chem Phys*. 1993;98:10089–10092. <https://doi.org/10.1063/1.464397>.
13. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys*. 2007;126:014101. <https://doi.org/10.1063/1.2408420>.
14. Parrinello M, Rahman A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys*. 1981;52(12):7182–7190. <https://doi.org/10.1063/1.328693>.
15. Michaud-Agrawal N, Denning EJ, Woolf TB, Beckstein O. MDAAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J Comput Chem*. 2011;32(10):2319–2327.
<https://doi.org/10.1002/jcc.21787>.
16. Theobald DL. Rapid calculation of RMSDs using a quaternion-based characteristic polynomial. *Acta Crystallogr A*. 2005;61(4):478–480. <https://doi.org/10.1107/S0108767305015266>.
17. Shrake A, Rupley JA. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J Mol Biol*. 1973;79(2):351–371. [https://doi.org/10.1016/0022-2836\(73\)90011-9](https://doi.org/10.1016/0022-2836(73)90011-9).
18. Baker EN, Hubbard RE. Hydrogen bonding in globular proteins. *Prog Biophys Mol Biol*. 1984;44(2):97–179. [https://doi.org/10.1016/0079-6107\(84\)90007-5](https://doi.org/10.1016/0079-6107(84)90007-5).
19. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1–2:19–25. <https://doi.org/10.1016/j.softx.2015.06.001>.

Figures

Study design and conceptual representation of the CaM-IQ model

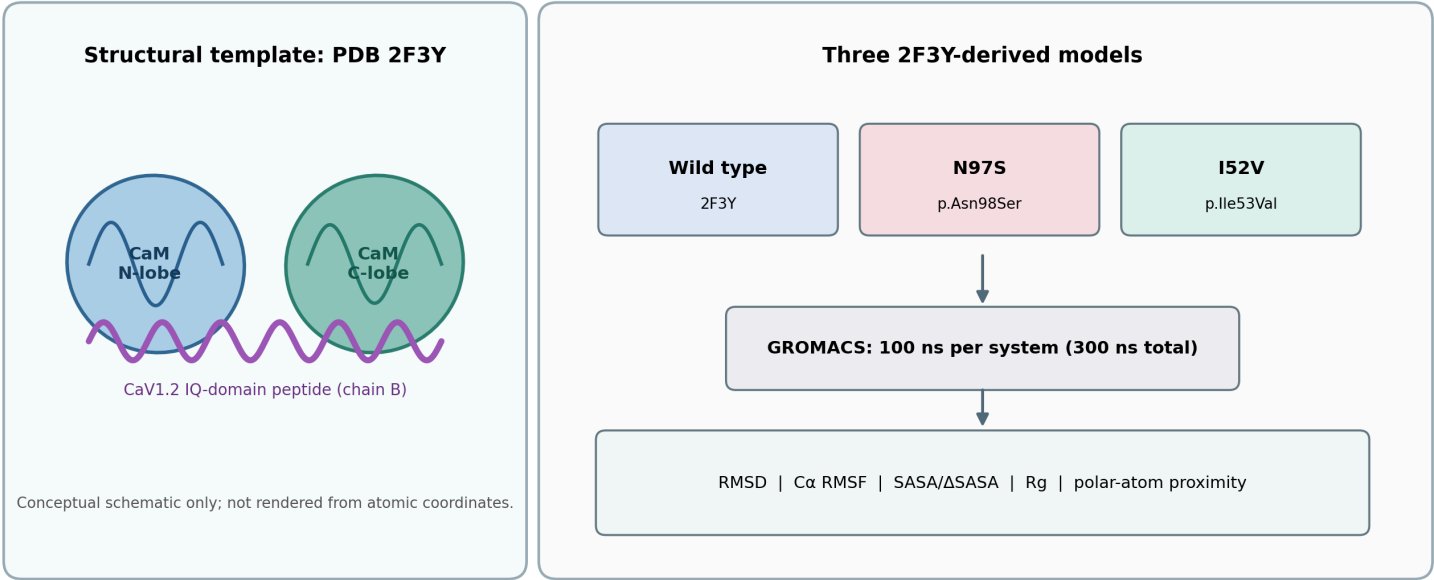


Figure 1

Study overview and conceptual model representation. The left panel is a simplified, non-atomic schematic of the CaM–IQ-domain assembly and is not an image rendered from PDB coordinates. PDB 2F3Y is the structural template; the right panel summarizes the three model systems, 100-ns production simulations, and trajectory descriptors analyzed. The historically used variant labels are accompanied by current CALM1 reference-sequence nomenclature.

Comparative trajectory descriptors across the three systems

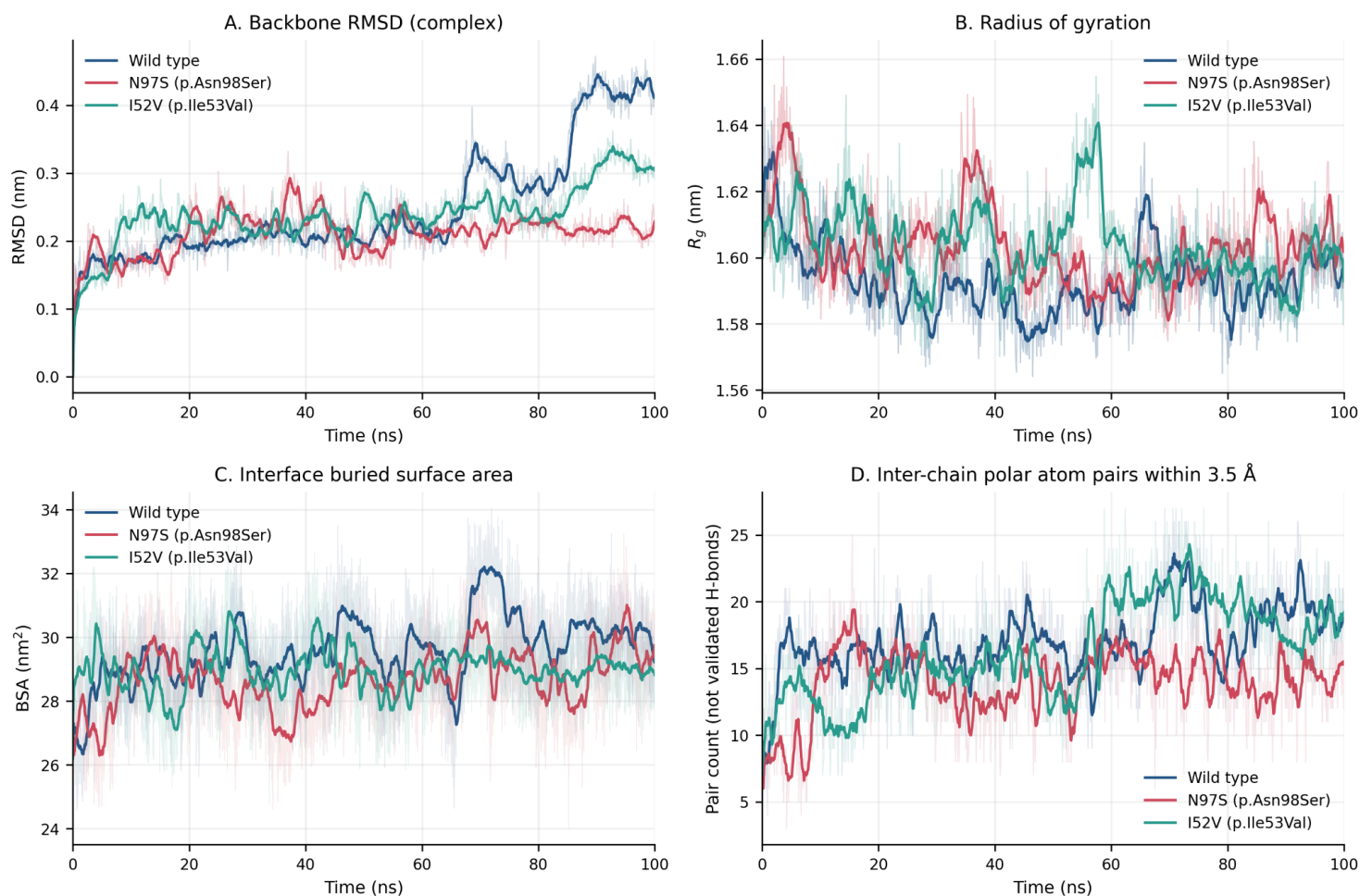


Figure 2

Comparative time-series descriptors across the three systems. (A) Backbone RMSD of the complete CaM-IQ model; the pale traces show sampled values and darker traces are smoothed for visualization. (B) Overall-complex radius of gyration. (C) Combined buried SASA, calculated as SASA(chain A) + SASA(chain B) - SASA(complex). (D) Count of selected inter-chain polar heavy-atom pairs within 3.5 Å. Panel D is a proximity metric, not a geometrically validated hydrogen-bond count. Smoothing does not enter the numerical summaries reported in Table 1.

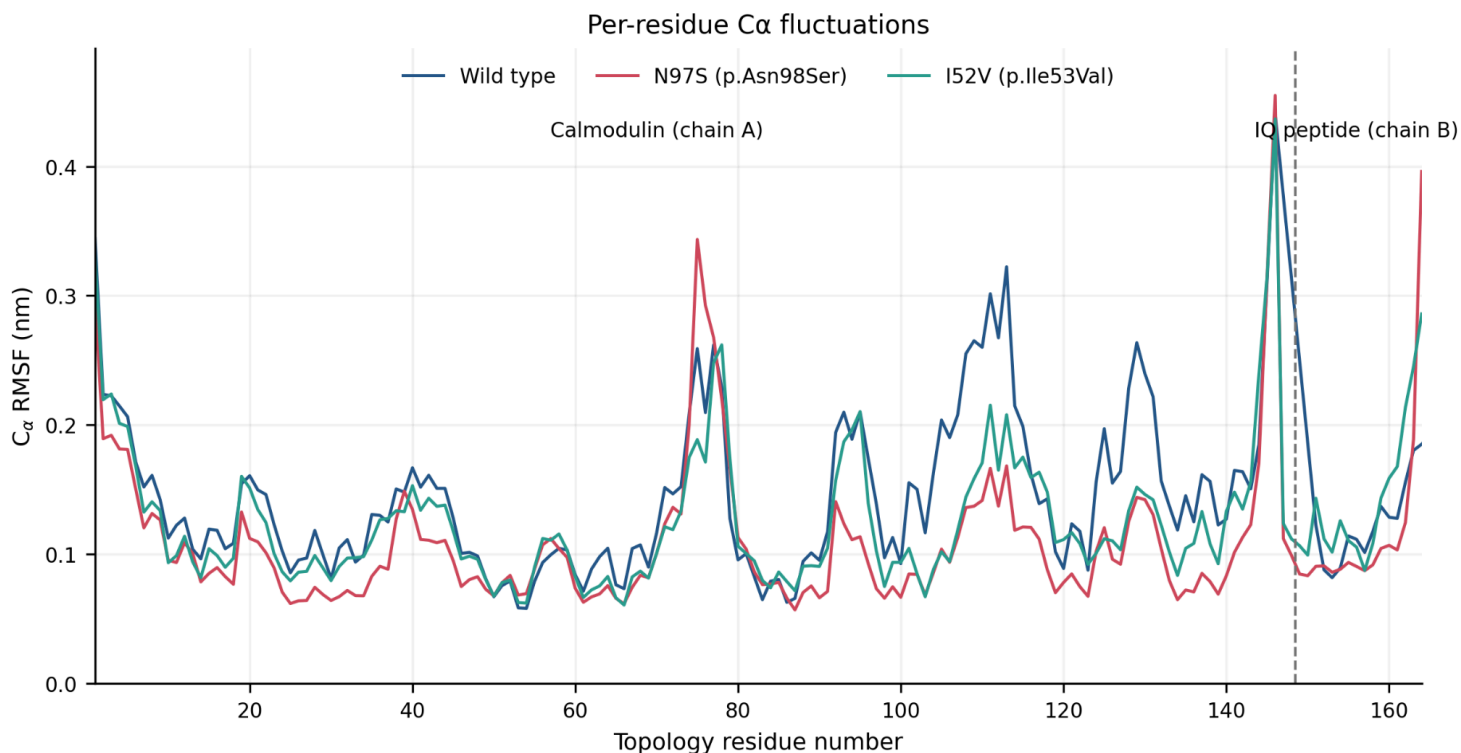


Figure 3

Per-residue C α RMSF across the three analyzed trajectories. RMSF values are expressed in nanometers. The dashed boundary after residue 148 reflects the script's chain-separation convention (CaM chain A versus IQ peptide chain B). The largest peaks are retained as descriptive observations; residue-level mapping should be verified against the exact topology and PDB/reference alignment before mutation-site-specific interpretation.

Supplementary Files

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

- [CalmodulinCaV1.2IQSupplementaryInformation.docx](#)