Received: 25 August 2023 | Revised: 19 October 2023 | Accepted: 24 October 2023

DOI: 10.1002/jcb.30495

RESEARCH ARTICLE

Journal of Cellular Biochemistry WILEY

Bioinformatics insights into CENP-T and CENP-W protein-protein interaction disruptive amino acid substitution in the CENP-T-W complex

Suryakanta Mohanty[1] | **Rajendra Bhadane**[2] | **Shashank Kumar**[1]

[1]Molecular Signaling & Drug Discovery Laboratory, Department of Biochemistry, Central University of Punjab, Guddha, Bathinda, India

[2]Institute of Biomedicine, Research Unit for Infection and Immunity, University of Turku, Turku, Finland

Correspondence
Shashank Kumar, Molecular Signaling & Drug Discovery Laboratory, Department of Biochemistry, Central University of Punjab, Guddha, Bathinda-151401, Punjab, India.
Email: shashankbiochemau@gmail.com

Rajendra Bhadane, Institute of Biomedicine, Research Unit for Infection and Immunity, University of Turku, Kiinamyllynkatu 10, FI-20520, Turku, Finland.
Email: rajendra.bhadane@utu.fi

Funding information
Department of Science and Technology, Ministry of Science and Technology, India, Grant/Award Numbers: DST-FIST grant to Department of Biochemistry, CUPB, SR/PURSE/2023/220

Abstract

Kinetochores are multi-protein assemblies present at the centromere of the human chromosome and play a crucial role in cellular mitosis. The CENP-T and CENP-W chains form a heterodimer, which is an integral part of the inner kinetochore, interacting with the linker DNA on one side and the outer kinetochore on the other. Additionally, the CENP-T-W dimer interacts with other regulatory proteins involved in forming inner kinetochores. The specific roles of different amino acids in the CENP-W at the protein-protein interaction (PPI) interface during the CENP-T-W dimer formation remain incompletely understood. Since cell division goes awry in diseases like cancer, this CENP-T-W partnership is a potential target for new drugs that could restore healthy cell division. We employed molecular docking, binding free energy calculations, and molecular dynamics (MD) simulations to investigate the disruptive effects of amino acids substitutions in the CENP-W chain on CENP-T-W dimer formation. By conducting a molecular docking study and analysing hydrogen bonding interactions, we identified key residues in CENP-W (ASN-46, ARG-53, LEU-83, SER-86, ARG-87, and GLY-88) for further investigation. Through site-directed mutagenesis and subsequent binding free energy calculations, we refined the selection of mutant. We chose four mutants (N46K, R53K, L83K, and R87E) of CENP-W to assess their comparative potential in forming CENP-T-W dimer. Our analysis from 250 ns long revealed that the substitution of LEU83 and ARG53 residues in CENP-W with the LYS significantly disrupts the formation of CENP-T-W dimer. In conclusion, LEU83 and ARG53 play a critical role in CENP-T and CENP-W dimerization which is ultimately required for cellular mitosis. Our findings not only deepen our understanding of cell division but also hint at exciting drug-target possibilities.

KEYWORDS
amino acid substitution, binding free energy, CENP-T, CENP-W, dimer formation, molecular dynamics (MD) simulation, protein-protein interaction

1 | INTRODUCTION

Kinetochores, massive proteinaceous scaffolds, assemble on centromeric chromatin and accurately attached to spindle microtubules, facilitating the separation of sister chromatids during mitosis. Comprising around 30 core subunits, kinetochores possess the capability to recruit additional regulatory proteins. The inner kinetochore primarily encompasses the constitutive centromere-associated network (CCAN), a complex formed by 16 subunits, which orchestrates the recruitment of outer kinetochore components to centromeric DNA. This creates robust interface capable of withstanding spindle forces.[1] All members of the CCAN are collectively known as centromere proteins (CENPs). The inner kinetochore CCAN complex is comprises proteins like CENP-T/W/X/S, CENP-H/K/I/M, CENP-C, CENP-N/L, CENP-O/P/Q/U, and CENP-R. The CENP-T/W/S/X complex, a heterotetramer, directly engages with both the centromeric chromatin nucleosome and the outer kinetochore protein complex. CENP-T consists of 561aa and contains five α helical structures with a molecular weight of 60,423 Dalton. CENP-W consists of 88aa and contains three α helical structures with a molecular weight of 10,061 Dalton.[2,3] The C-terminal of the CENP-T interacts with the histone fold domain of the DNA, while the elongated N-terminal interacts with the external kinetochore components.[4,5] CENP-T forms a direct binding with CENP-W, another member of the CENP-T/W/X/S complex.[6] The CENP-T-W heterodimer is pivotal for kinetochore localization in human cells.[3] CENP-S and CENP-X, being smaller proteins, combine with CENP-T/W to form a complex. CENP-S/X is not vital for CENP-T recruitment to kinetochores; however, it appears that the CENP-T/W/S/X complex forms a nucleosome-like structure at centromeres.[3] Specifically, only CENP–T or CENP–W possesses DNA-binding regions, which are absent in CENP–S or CENP–X.[7] Recently, Yatskevich et al.[8] reported the cryo-EM structure of the human CCAN complex bound to CENP-A reconstituted onto α-satellite DNA. The CENP-T/W/S/X histone-fold enhances DNA binding and partially wraps the linker DNA segment of the α-satellite (Figure 1). The study was focused on the interaction of the CENP-T/W/S/X complex with the DNA. The CENP-W/T complex assembles at centromeres during the late S of the cell cycle and is indispensable for cellular mitosis.[9] The CENP-S/X hetero-dimer is not mandatory for mitosis but contributes to kinetochore stabilization.[10] Depletion of CENP-W results in substantial mitotic delay, disorganized pro-metaphases, and the generation of multipolar spindles in-vitro.[11] Notably, CENP-T/-W is not inherited at centromeres; instead, new deposition (during the S phase) is needed in each cell cycle for kinetochore function.[9] In earlier research, it was demonstrated that the C-terminal histone fold domains (HFDs) of the CENP-T and CENP-W mediate the formation of CENP-T-W heterodimer. HFDs are capable of establishing stable protein–protein interactions.[6] In previous studies, the impact of amino acid substitutions in the DNA-binding domain of CENP-T and CENP-W on the localization of the CENP-T/W complex to the kinetochore has been demonstrated.[3] Information about the essential amino acid residues responsible for the formation of CENP-T and CENP-W interaction-mediated heterodimer formation is still lacking.

Protein–protein interactions (PPIs) are crucial for executing essential biological process. Employing structure-based site-directed mutagenesis offers a sustainable approach for comprehending and modifying specific aspects of protein function.[12–14] The knowledge-based substitution of amino acid residue(s) in a target protein of the PPI complex may alter the binding affinity and stability. Investigating correlation between amino acid substitution and its impact on protein-protein interaction and stability provides furnishes vital insight into the crucial amino acid residues implicated in complex formation. Nonetheless, experimental techniques to study these critical amino acid residues within PPI are often resource-intensive and time-consuming. Advances in computer-based tools and software present an effective solution to this challenge.[15] In this study, our objective is to identify the fundamental amino acid residues involved in the interactions between CENP-T and CENP-W by introducing mutations in the CENP-W chain (Figure 2). The knowledge garnered about key amino acid residues within CENP-T/W heterodimer has potential to illuminate the intricate functional role of the dimer in the human cell cycle. Moreover, insights into protein-protein interaction could guide the design and identification of small molecule or peptide inhibitors targeting the formation of the CENP-T/W heterodimer, opening up avenues for therapeutic intervention.

2 | METHODOLOGY

2.1 | Protein retrieval and preparation

The 3D structure of receptor molecule was selected from Protein Data Bank (PDB) to retrieve structure of the chosen receptor proteins, CENP-T and CENP-W, bound with the human CCAN complex (PDB ID: 7R5S).[8] Subsequently, the structure was manually curated in Pymol by adding hydrogen and removing the water molecules and other chains.[16]

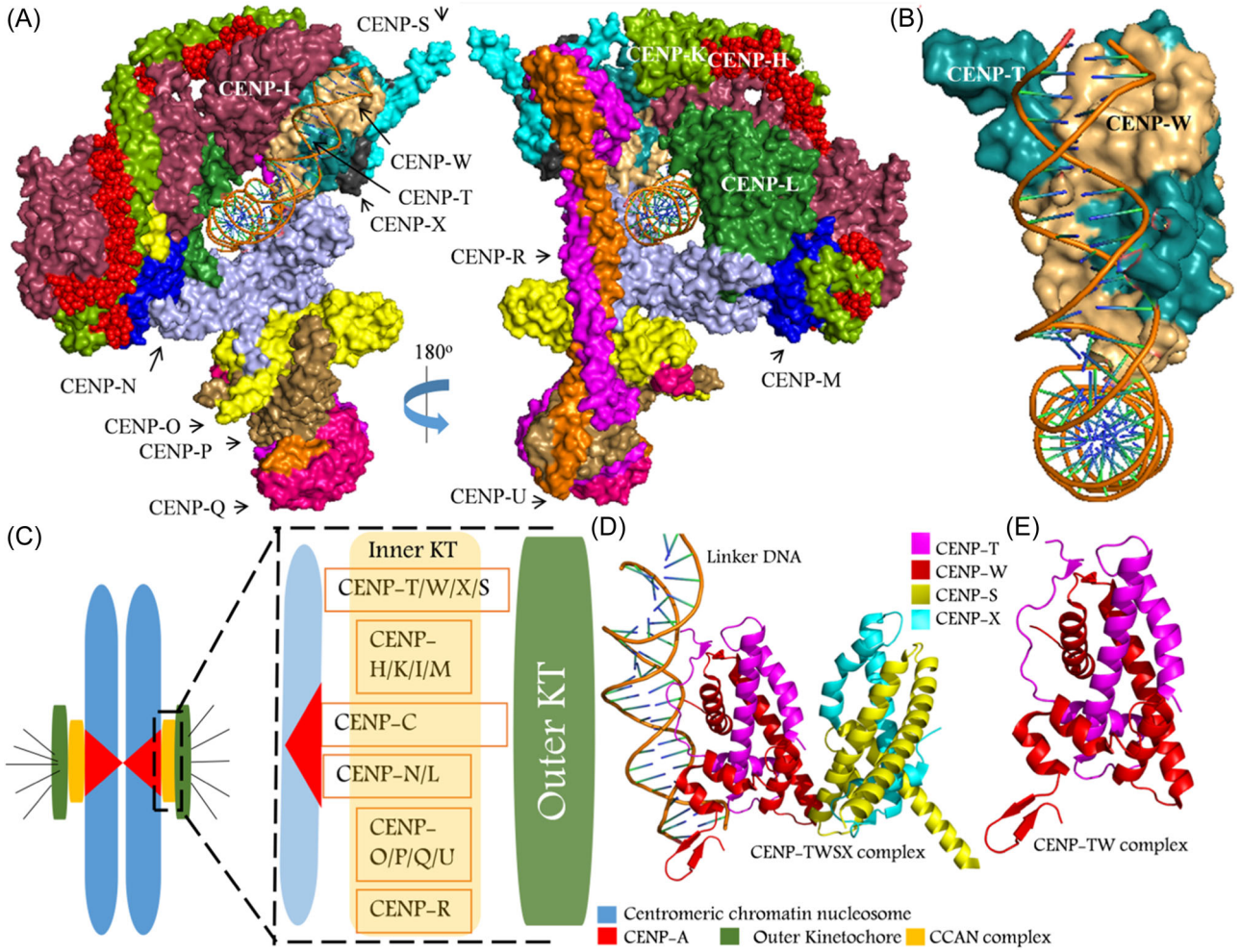

FIGURE 1 Representation of the human Chromosome, Centromere, Kinetochore Assembly, and CENP complex. (A) CCAN complex architecture (PDB ID:7R5S) (B) Separate visualization of the CENP-TW region of the CCAN complex (C) Cartoon structure representation of chromosome, centromere, and kinetochore assembly. (D) The 3D structure of human CENP-T-W-S-X complex with linker DNA segment (PDB ID: 7R5S) and (E) The 3D structure of CENP-T-W dimer (PDB ID:7R5S).[8]

2.2 | Protein-protein docking

Cluspro, was employed to analyze protein complexes 3D structures. Utilizing a combination of rigid body docking and conformational sampling algorithms, it explores potential binding orientations between two protein structures.[17] For protein-protein interactions analysis, The CENP-T chain served as receptor, while CENP-W acted as the ligand, enabling the identification of pivotal amino acids involved in the interaction.

2.3 | Selection of protein-protein interaction site

Initially, PyMol was utilized to visualize the interaction sites in CENP-T-W complex. The selection of the probable protein-protein interaction site for further investigation was guided by following criteria:

1. The site comprising higher number of amino acid residues involved in CENP-T-W complex formation.
2. The region involved in the CENP-T-W-DNA interactions was excluded from the study.

2.4 | Site directed mutagenesis in CENP-W

The approach to mutagenesis was based on the protocol for site-directed mutagenesis.[18] The name of the amino acids considered for the mutagenesis and the respective substituent amino acid with their properties are summarized in Table 1. Mutagenesis in

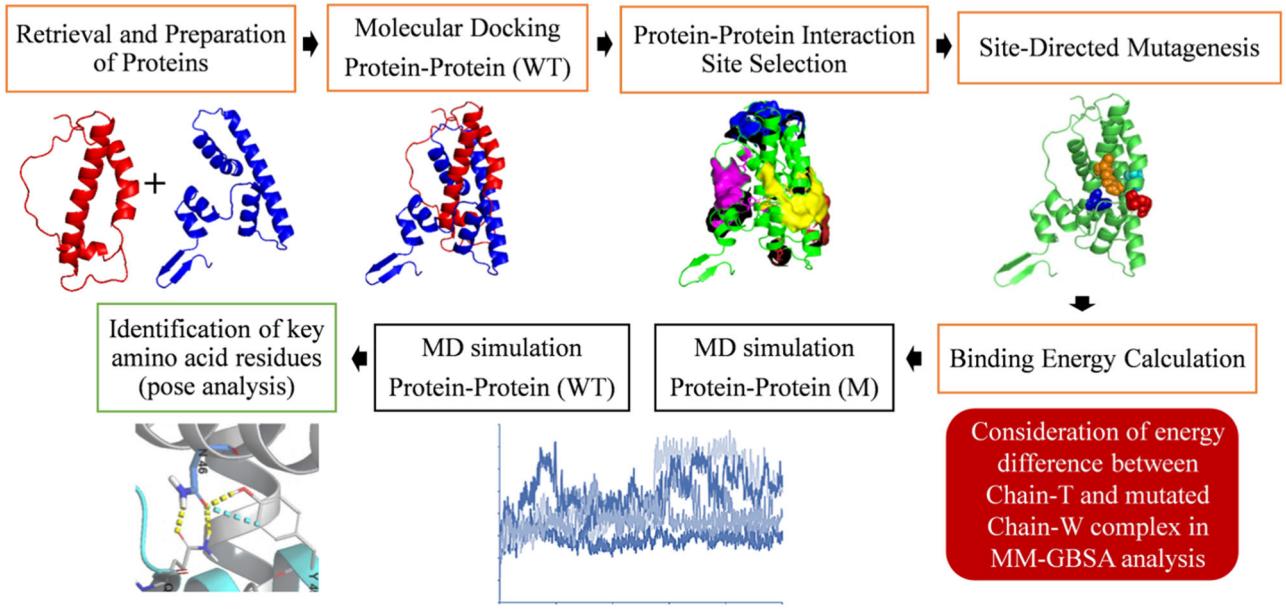

FIGURE 2 Flowchart illustrating the study design employed for the current investigation. WT-Wild type; M-Mutant; MD-Molecular dynamics.

TABLE 1 CENP-W amino acid residues involved in H-Bond interaction with CENP-T in WT complex and the properties of residues used for substitution to create mutant CENP-W

CENP-W residues interacting with CENP-T at Site 3	Properties of CENP-W residues interacting with CENP-T	CENP-W residue used for substitution	Properties of mutant CENP-W amino acid residues
ASN-46	Polar uncharged	LYS	Polar, positively charged, H-donor/acceptor
ARG-53	Positively charged, H-donor group	LYS	Polar, positively charged, H-donor/acceptor
LEU-83	Nonpolar, Hydrophobic	LYS	Polar, positively charged, H-donor/acceptor
SER-86	Polar uncharged	GLU	Negatively charged
ARG-87	Positively charged	GLU	Negatively charged
GLY-88	Nonpolar aliphatic	TRP	Non-polar aromatic

CENP-W protein was incorporated by utilizing PyMOL Mutagenesis Wizard.[16] Using PyMol, residues within 4 Å radius were identified, and specific amino acids within site 3 of CENP-W were chosen, as depicted in Figure 3A. The hydrogen bond interaction distance between chain-T and chain-W was measured employing Ligplot+ v.2.2.5.

2.5 | Molecular dynamics (MD) simulations

MD simulations were performed on both the wild-type (WT) and mutated CENP-T-W protein-protein complexes with the OPLS4 force field.[19] Each complex underwent a 250-ns MD simulation using Desmond (Schrödinger Release 2023-2: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, USA, 2021).[20] These simulations played a crucial role in refining the CENP-T-W protein-protein complex. The simulation systems were prepared using the System Builder tool, with explicit solvation modeled using single point charge (SPC) water.[21] To maintain the system's electroneutrality and simulate physiological conditions, Na^+ or Cl^- ions were added, resulting in a 0.15 M salt concentration. Orthorhombic simulation boxes with Periodic Boundary Conditions (PBC) and a 10-Å buffer space between the solute and the

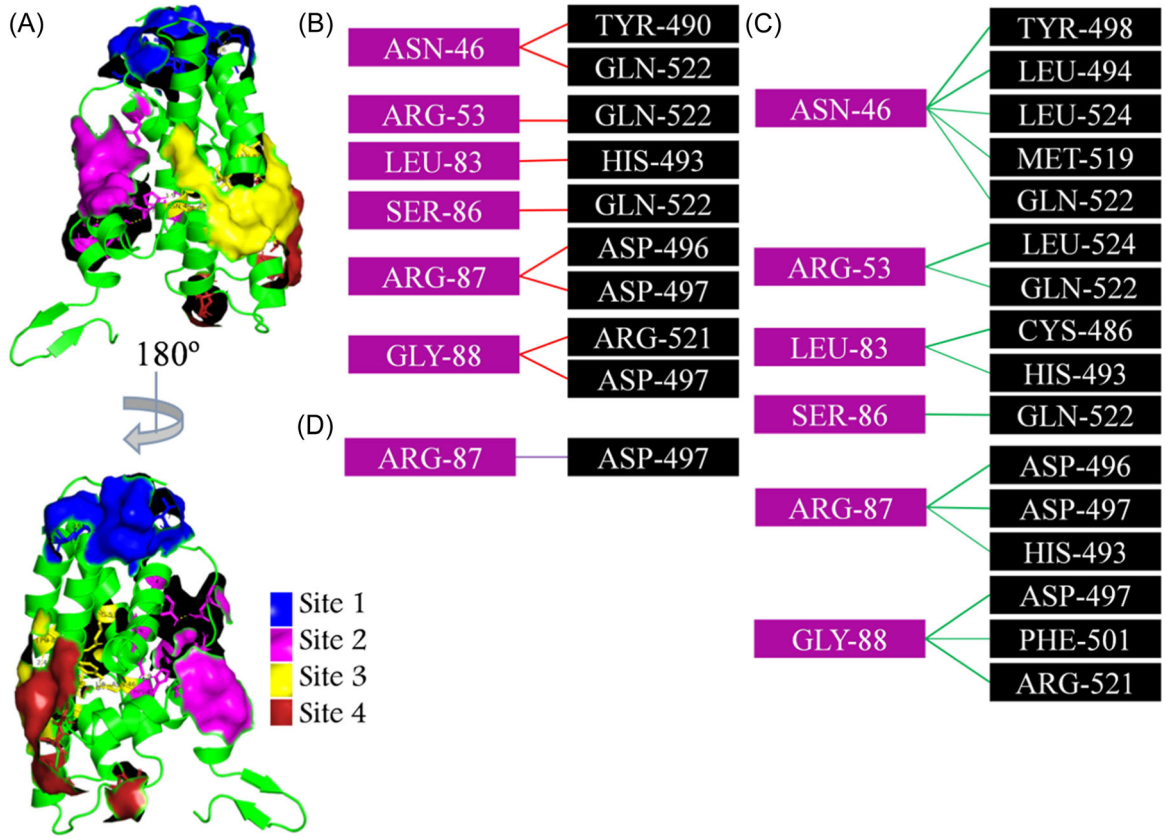

FIGURE 3 Visualization of CENP-T-W complex and amino acid interaction involved in the interaction at site-3. (A) Two views of surface stick representation of protein-protein interaction sites between CENP-T-W complexes (PDB- 7R5S) representing different sites. Representation of amino acids involved in interactions with in 4 Å distance (B) Hydrogen bond interaction (C) Hydrophobic interaction and (D) Salt bridge interaction. Amino acids represented in purple and black color boxes represent the residues of CENP-W and CENP-T proteins respectively.

box edge were defined. The simulation systems underwent minimization and brief equilibration before running the production simulations in a stepwise manner. The production simulations were carried out in the NPT ensemble for 250 ns, utilizing a 2-fs time step. The temperature was maintained at 300 K using the Nosé-Hoover chain thermostat, while the pressure was maintained at 1.01325 bar with the Martyna-Tobias-Klein barostat.[22–24] Isotropic coupling was applied, and the pressure and temperature relaxation times were set to 1 ps and 2 ps, respectively. To handle short-range Coulombic interactions, a cutoff radius of 9.0 Å was employed. The long-range electrostatic interactions were calculated using the u-series decomposition of the Coulomb potential.[25] The resulting simulation trajectories were analyzed using the Simulation Interactions Diagram tool in Maestro. The data generated from the interaction analysis were further processed using Microsoft Excel365 to create graphical representations, facilitating detailed analysis and interpretation.

2.6 | Binding energy calculation

The binding energy of the WT and mutated complexes were calculated by Cluspro online tool and MMGBSA analysis. The binding free energy of the WT and mutant protein-protein complex was determined from the trajectory file obtained after 250 ns long MD simulation using prime MM-GBSA approach employing the VSGB2.1 implicit solvation model and the OPLS4 force field to calculate the binding free energy.[26–28] The binding free energy (ΔG_{bind}) was computed using the previously established equation described in our protocol[29]:

$$\Delta G_{bind} = G_{\mathrm{CENP-T-W}} - (G_{\mathrm{CENP-T}} + G_{\mathrm{CENP-W}})$$

where $G_{\mathrm{CENP-T-W}}$ is the free energy of the CENP-T-W complex, $G_{\mathrm{CENP-T}}$ is the free energy of the CENP-T and $G_{\mathrm{CENP-W}}$ is the free energy of the CENP-W protein. ΔG_{bind} is calculated for each mutated CENP-W in this way.

3 | RESULTS AND DISCUSSION

3.1 | CENP-T-W interaction analysis

We initiated protein-protein docking of the CENP-T and CENP-W chains (PDB ID: 7R5S)[7] to generate 10 docked poses. The selection of the best docked poses was based on the least docking energy. Figure 3 illustrates the interactions between CENP-T and CENP-W chains within a 4 Å distance. Table S1 provides details about the key amino acids involved in hydrogen bonding, hydrophobic interactions, and salt bridges formed between the proteins. In the subsequent step of our analysis, we aimed to identify potential binding sites in the dimer for further investigation. After shortlisting four possible binding sites (Figure 3A), we excluded site 1 due to its main role in protein-DNA interactions. Poor interactions at sites 2 and 4 also led to their elimination from further analysis, leaving only site 3 for further investigation.

Figure 3B–D illustrate the amino acid interactions at site 3, which was selected for further study due to the involvement of a higher number of residues in the protein-protein interaction (PPI). The results revealed that hydrogen bonding, hydrophobic, and salt-bridge interactions were primarily formed at site 3. Specifically, ASN-46, ARG-53, LEU-83, SER-86, ARG-87, and GLY-88 residues of the CENP-W chain formed hydrogen bonds with TYR-490, GLN-552, GLN-522, HIS-493, GLN-552, ASP-497, and ARG-521 residues of CENP-T, respectively. Among these, ARG-87 and GLY-88 formed two hydrogen bonds each, while ARG-53, LEU-83, SER-86, and ARG-87 formed one hydrogen bond each with the CENP-T chain. At site 3, a total of five CENP-T residues (TYR-490, GLN-552, HIS-493, ASP-497, and ARG-521) were involved in hydrogen bond formation during the protein-protein interaction.

Additionally, the ASN-46, ARG-53, LEU-83, SER-86, ARG-87, and GLY-88 residues of the CENP-W chain engaged in hydrophobic interactions with TYR-498, LEU-494, LEU-524, MET-519, GLN-522, LEU-524, GLN-552, CYS-486, HIS-493, GLN-552, ASP-497, PHE-501, and ARG-521 residues of CENP-T, respectively. In this context, ASN-46, ARG-53, LEU-83, SER-86, ARG-87, and GLY-88 contributed to five, two, two, one, three, and three hydrophobic interactions, respectively, with the CENP-T chain. At site 3, a total of 10 CENP-T residues (TYR-498, LEU-494, LEU-524, MET-519, GLN-522, CYS-486, HIS-493, ASP-497, PHE-501, and ARG-521) were involved in hydrophobic interactions during the protein-protein interaction. Furthermore, the ARG-87 residue of the CENP-W chain formed a salt bridge interaction with the ASP-497 residue of the CENP-T chain.

These results provide valuable insights into the critical interactions at site 3, which play a pivotal role in the protein-protein interaction between CENP-T and CENP-W chains, thereby contributing to potential scientific research publication.

3.2 | Site directed mutagenesis and H-bond formation analysis

To select suitable amino acid pairs within the designated site of the CENP-T-W complex, we focused on six amino acid residues from the CENP-W chain, namely ASN-46, ARG-53, LEU-83, SER-86, ARG-87, and GLY-88. These residues exhibited potential interactions with CENP-T within a 4 Å region of the protein-protein interface at site 3. These amino acid residues were subsequently chosen for site-directed mutagenesis with the objective of reducing the binding interactions.

The mutation of CENP-W amino acid residues was based on the broad categories' classification of amino acids, aiming to diminish the binding interactions. Table 1 summarizes the properties of the residues used for the substitution to create the mutant CENP-W. Within site 3, ASN-46 of CENP-W formed hydrogen bond interactions with CENP-T chain TYR-490 and GLN-522 residues at distances of 2.74 Å and 2.97–2.80 Å, respectively. Additionally, ARG-53, LEU-83, and SER-86 of CENP-W exhibited independent hydrogen bond interactions with GLN-522, HIS-493, and GLN-522 of CENP-T at distances of 2.72 Å, 3.29 Å, and 2.79 Å, respectively. Furthermore, CENP-W residues ARG-87 and GLY-88 interacted with CENP-T ASP-496 and ASP-497 at distances of 2.97–2.83 Å and 2.88 Å, respectively. Additionally, ARG-87 interacted with CENP-T ARG-521 and ASP-497 at distances of 2.85–2.65 Å and 2.78 Å, respectively, through hydrogen bond interactions.

The structural analysis of the CENP-W-T complex with CENP-W mutations revealed significant deviations in the complex's conformation. The mutations primarily affected loop and α-helix regions, indicating their sensitivity to structural changes (Figure 4A). The observed alterations included both relaxation and compactness-related shifts, suggesting potential changes in protein dynamics and packing. These findings are depicted in Figure 4B, which shows the superimposed structures of the wild-type complex and the complex containing CENP-W mutants. The structural deviations are clearly visible, emphasizing the impact of the mutations on the protein complex. The results underscore the functional relevance

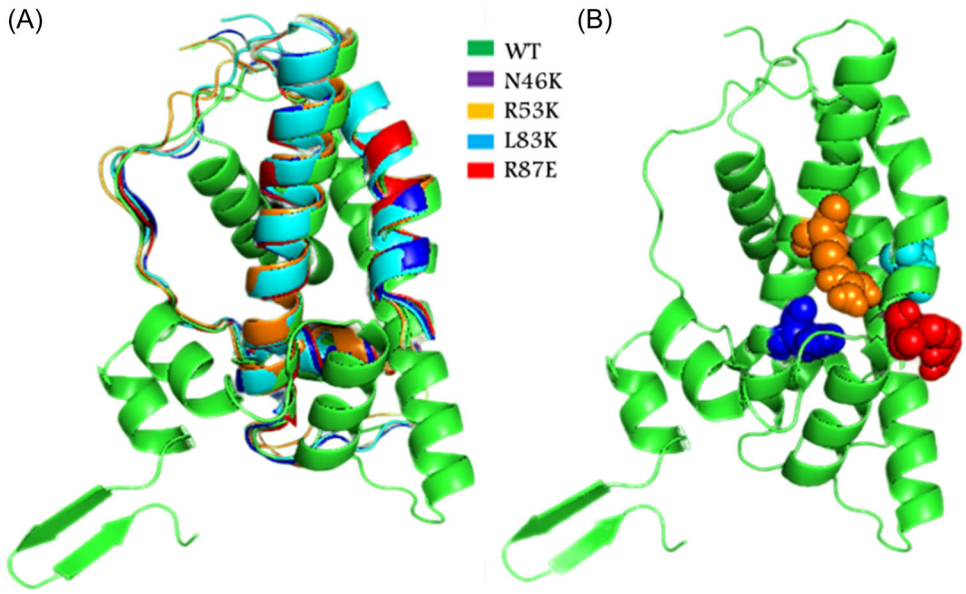

FIGURE 4 Superimposed structures of the CENP-T and CENP-W complex and point mutations. (A) Superimposed structures of the CENP-T and CENP-W containing WT CENP-T complexed with CENP-W mutants. (B) Amino acid residues used to form CENP-W mutants.

of the considered mutations in CENP-W, potentially impacting its interactions and cellular processes. However, further experimental assays are needed to fully understand their functional implications.

In summary, the superimposed structural analysis, depicted in Figure 4, highlights the potential impact of CENP-W mutations on the complex's conformation and offers insights into their functional significance, setting the stage for further investigations.

3.3 | Binding energy calculation for WT and mutant complexes

After performing mutagenesis, we assessed the binding energy using Cluspro for both the wild-type CENP-T-W complex and the complex containing the mutated amino acids in the CENP-W protein chains. Table 2 displays the binding energy values for the wild-type complex and the complex with mutated amino acids, along with the energy difference between the two. The results revealed that substituting LEU-83, ARG-87, ASN-46, and ARG-53 residues in the CENP-W chain with LYS, GLU, LYS, and LYS, respectively, led to a significant increase in the binding energies of the respective complexes. Based on these findings, we selected this particular mutant complex for further investigation of its binding free energy using the MM-GBSA approach.

TABLE 2 Binding energy calculation for WT and mutated CENP-T-W complexes

CENP-W residue	Substitution (Mutant) with the residue	Binding energy (kcal/mol)	Energy difference
WT	-	−2542.3	-
LEU-83	**LEU-LYS**	**−2456.8**	**85.5**
SER-86	SER-GLU	−2538.5	3.8
ARG-87	**ARG-GLU**	**−2466.3**	**76**
GLY-88	GLY-TRP	−2551.0	8.7
ASN-46	**ASN-LYS**	**−2495.2**	**47.1**
ARG-53	**ARG-LYS**	**−2509.0**	**33.3**

Note: Based on the data we selected the mutant represented in bold letter for MD simulation and subsequent binding free energy calculations.

3.4 | Effect of amino acid mutation on protein backbone and amino acid fluctuations

MD simulation is a valuable tool for assessing the stability and flexibility of biomolecular systems.[30,31] In this study, we analyzed the simulation trajectories obtained from the MD simulations of the WT and mutant CENP-W-T complexes (N46K, R53K, L83K, and R87E). To assess the convergence of each system, we calculated the backbone Root Mean Square

Deviation (RMSD) of the protein chains. RMSD measures the variance between a protein's backbone atom locations from its initial energy-minimized structure. We separately analyzed the RMSD results for both the CENP-W and CENP-T chains (Figure 5A, B). The RMSD analysis showed that the CENP-T chain remained relatively stable in all cases, with no significant difference between WT and mutant CENP-W complexes. When analysing the stability of CENP-W, we observed significant differences in RMSD, except for the N46K mutation, which converged well below the WT CENP-W. The RMSD of R53K and L83K mutants was significantly higher, indicating that lysine mutations negatively impact the stability of CENP-W. The RMSF values for backbone atoms at each time point of the trajectories of the WT and mutant CENP-T-W complex were determined to investigate how the amino acid substitutions affect the dynamics of the backbone atoms. A higher RMSF value denotes an increased flexibility in the complex during the MD simulation period. It has been reported that alteration in RMSF is largely associated with the protein function.[32] We analyzed RMSF of the WT/mutant CENP-W of the respective complex and CENP-T-W complex residues separately to get the comparative effect of the amino acid substitutions on CENP-W chain and CENP-T-W complex (Figure 5C, D). However, while observing the Root Mean Square Fluctuations (RMSF), we noticed increased fluctuations in CENP-T upon N46K, R53K, and L83K mutations. This suggests that these mutations may have a deleterious effect on protein-protein interactions. In case of N46K mutant fluctuations increased around amino acid 520–530 and towards the C-terminal end, implying the destabilizing effect of N46K mutant. The R53K mutation other hand mostly increased the fluctuation around 505–525 amino acids. The L83K mutation on the other hand mostly increased fluctuation in amino acid 520–530 and C-terminal end amino acids from 540 to 550 which could be due to drifting of CENP-W helix due to Lys83 mutant. Conversely, R87E mutation showed no significant difference in RMSF fluctuations.

Overall, the MD simulation and RMSD/RMSF analysis provided valuable insights into the structural dynamics and stability changes induced by these mutations in the CENP-W-T complex. These findings contribute to our understanding of the molecular consequences of the mutations and their potential impact on protein function and interactions.

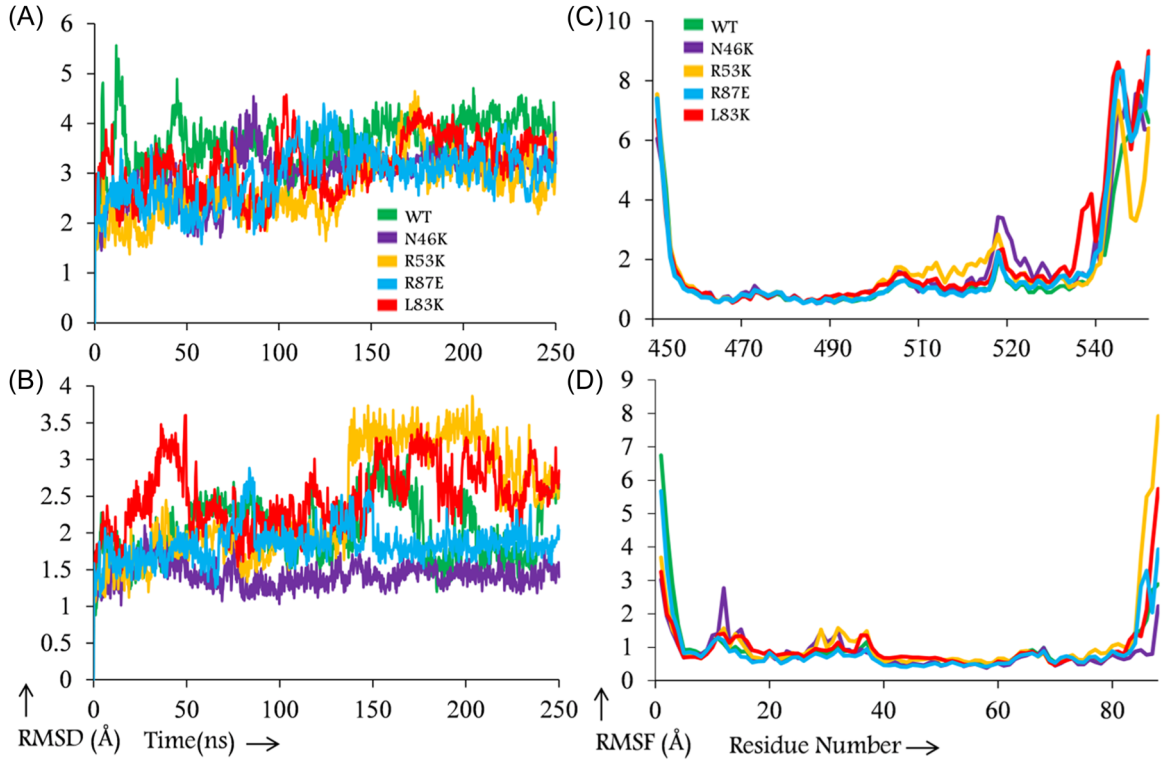

FIGURE 5 RMSD and RMSF analysis of CENP-T-W complexs during the 250 ns MD simulation period (A) RMSD of WT CENP-T in complex with WT and mutant CENP-W (B) RMSD of WT and mutant CENP-W in complex with WT CENP-T. (C) RMSF of WT CENP-T in complex with WT and mutant CENP-W (D) RMSF of WT and mutant CENP-W in complex with WTCENP-T.

3.5 | Effect of amino acid substitution on structural integrity of the CENP-W-T dimer

The radius of gyration (Rg) analysis is a powerful tool used to assess the compactness and conformational stability of biomolecular structures during molecular dynamics (MD) simulations.[33] In our study, we applied Rg analysis to investigate the impact of specific amino acid substitutions in the CENP-W chain on the structural dynamics of the CENP-T-W complex. The Rg values obtained during the 250 ns MD simulations provided valuable insights into the conformational changes occurring in the wild-type and mutant complexes over time. For the wild-type CENP-T-W complex, the observed decrease in Rg during the initial 50 ns of simulation indicates a shift towards a more compact conformation (Figure 6A). This stabilization at 17.7 Å up to 220 ns suggests that the complex maintains a relatively stable and well-packed structure during this period. The slight decrease in Rg during the last 30 ns suggests some additional structural adjustments or fluctuations taking place towards the end of the simulation.

The Rg analysis of the CENP-W mutants, specifically the R87E substitution, revealed rhythmic fluctuations in the Rg values throughout the entire simulation (Figure 6A). Interestingly, during the initial 50–100 ns period, the Rg value of the R87E mutant complex showed a distinct increase compared to the WT, indicating a potential alteration in its compactness. These fluctuations may indicate dynamic structural changes induced by the R87E substitution, possibly affecting the overall stability of the complex. The L83K substitution in the CENP-W chain resulted in an increase in Rg values compared to the wild-type complex, peaking at 100 ns. However, beyond 100 ns, the Rg values of the L83K mutant followed a pattern similar to that of the wild-type, suggesting a possible restoration of stability or accommodation of the mutation-induced changes in the complex. The Rg values of the R53K mutant were comparable to the WT during the initial 50 ns period. However, from 50 ns onwards, the Rg values of the R53K mutant were significantly higher than those of the WT, indicating a considerable expansion or structural perturbation in this mutant. The N46K substitution also initially showed higher Rg values, indicating a potential destabilization or conformational change induced by this mutation. During the 80–230 ns period, the N46K mutant exhibited substantial fluctuations in Rg values, suggesting a dynamic interplay of structural changes and flexibility at this site.

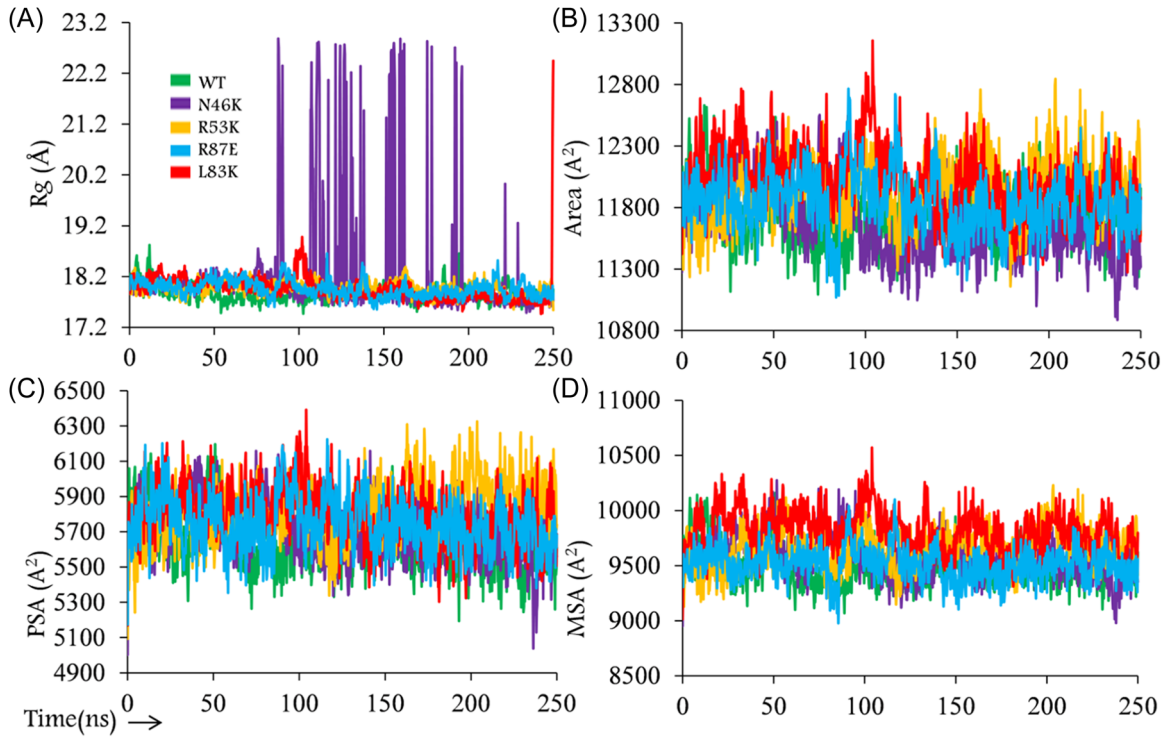

FIGURE 6 Trajectory analysis of the CENP-T-W complex during 250 ns MD simulation period. (A) Rg (B) SASA (C) PSA and (D) MSA pattern of the CENP-T-W WT dimer compared to WT CENP-T and point mutation carrying CENPT-W chain dimers. N46K, R53K, L83K, and R87E represent the mutations at reside number 46, 53, 83, and 87 respectively in CENPT-W chain of the CENP-T-W dimer.

Overall, the Rg analysis indicated that the amino acid substitutions in the CENP-W chain had distinct effects on the compactness and conformational stability of the CENP-T-W complex. The decrease in compactness observed in the mutants suggests that these specific substitutions may influence the complex's structural organization, potentially affecting its function and interactions. These findings add to our understanding of the structural consequences of the CENP-W mutations and provide important insights into their potential functional implications in the context of the CENP-T-W complex. However, it is important to note that further experimental validation and functional assays are required to fully elucidate the biological significance of these observed structural changes and their impact on the cellular processes involving the CENP-T-W complex. Our study serves as a valuable starting point for future investigations to unravel the molecular mechanisms underlying CENP-T-W function and regulation.

Solvation accessible surface area (SASA) analysis was performed to investigate the impact of amino acid substitutions on the unfolding and exposure of the hydrophobic core residues in the CENP-T-W complex (Figure 6B). The SASA values showed notable differences between the wild-type and mutant dimers. The SASA value for the wild-type dimer remained relatively constant at about 11 600 $Å^2$ throughout the MD simulation. The R87E substitution exhibited fluctuations in SASA values similar to the WT, particularly during the last 50 ns of the simulation. In contrast, both the R53K and L83K substitutions led to increased SASA values (11 300–13 200 $Å^2$) compared to the wild-type complex. Interestingly, during the initial 125 ns, the L83K mutant showed elevated SASA values, while during the last 125 ns, the R53K mutant exhibited increased SASA values compared to the WT. On the other hand, the N46K substitution resulted in decreased SASA values compared to the wild-type complex during the 100–250 ns simulation period. These findings suggest that the amino acid substitutions in the CENP-W chain induced alterations in the protein folding, leading to changes in the exposure of hydrophobic core residues of the complex. Furthermore, the increase in SASA values in the R53K and L83K mutants indicates a greater exposure of hydrophobic core residues to the surrounding water, potentially reducing the stability of the protein folding.

We further analyzed the polar surface area (PSA) in the wild-type and mutant CENP-W chain-containing CENP-T-W complex (Figure 6C). The PSA values for the wild-type complex fluctuated between 5300 and 6100 $Å^2$, with a predominant concentration around 5600 $Å^2$ during the simulation period. The amino acid substitutions increased the fluctuation in the PSA values in the mutant CENP-T-W complex compared to the wild-type. The increase in PSA values followed the order of WT < N46K < R87E < L83K < R53K, indicating that the mutations in the CENP-W chain affected the surface contribution of polar areas in the complex chains.

We also assessed the molecular surface area (MSA) as another parameter of MD simulation to evaluate the stability of the protein complex.[34] Lower MSA values are indicative of comparatively stable protein complex (Figure 6D). The MSA value of the wild-type complex was largely around 9500 $Å^2$ during the MD simulation period. However, the R53K and L83K mutant complex exhibited increased MSA values compared to the WT. The R53K mutant initially showed decreased MSA values during the first 40 ns, followed by a subsequent increase compared to the WT. The L83K mutant consistently showed higher MSA values throughout the simulation compared to the wild-type and other mutant complex. In contrast, the N46K mutant showed increased MSA during the initial 100 ns but subsequently maintained values within the range of the wild-type complex. Remarkably, the R87E mutant complex displayed minimal deviation in MSA values and maintained a pattern similar to the wild-type complex throughout the 250 ns MD simulation period. Overall, the amino acid substitutions in the CENP-W chain increased the MSA values compared to the WT CENP-T-W complex, with minimal alteration observed in the R87E mutant. These results highlight the potential effects of specific substitutions on the stability and dynamics of the CENP-T-W complex, which may have implications for its biological function. However, further experimental investigations and functional assays are warranted to fully comprehend the functional consequences of these structural changes in the context of the CENP-T-W complex.

3.6 | Effect of amino acid substitution on key amino acid residue interaction in CENP-W-T complex

The binding free energy calculations and molecular dynamics (MD) simulations provided valuable insights into the impact of amino acid mutations on the stability and conformational dynamics of the CENP-T-W complex. In an effort to gain a more comprehensive understanding of the effects of these mutations on the interactions within the complex during the MD simulations, we analyzed the trajectories of the N46K, R53K, L83K, and R87E mutants.

In the N46K mutant, the substitution of ASP with LYS led to an extended side chain that initially interacted with GLY-522 and PHE490. However, after 100 ns, clashes with neighboring amino acids caused the LYS-46-containing helix to drift away from the CENP-T chain, resulting in increased gaps and loss of hydrophobic contacts during the simulations (Figure 7). These findings suggest that the N46K mutation induces conformational changes that affect the stability of the CENP-T-W complex.

Similarly, in the R53K mutant, the ARG-53 residue in the wild-type complex formed strong hydrogen-bonded interactions with GLY-522. While the mutation to LYS-53 maintained these interactions at the beginning of the simulation, the Arg side chain eventually flipped back, forming intramolecular interactions with GLU-56 and GLU-57. This disruption of critical bridges between CENP-W and CENP-T further compromised the stability of the complex (Figure 8).

On the other hand, the Leu83 residue is deeply buried in the hydrophobic pocket, surrounded by PHE-50 from the CENP-W chain, TYR-490, and HIS-493 from CENP-T. Its mutation to LYS-83 resulted in significant alterations to the hydrophobic pocket, making it more solvent-exposed than the WT complex. As a consequence, LYS-83 caused the entire helix to drift away from the CENP-T chains during the MD simulations, leading to a loss of critical interactions (Figure 9). These observations suggest that the L83K mutation destabilizes the CENP-T-W complex by disrupting key hydrophobic interactions.

In contrast, the wild-type complex demonstrated stable hydrogen-bond and salt bridge interactions between the ARG-87 residue of CENP-W and the CENP-T ASP-497 residue during the initial MD simulation. As the simulation progressed, R87 also formed an intramolecular hydrogen bond with the LYS-84 residue and a pi-cation interaction with the CENP-T HIS-493 residue (Figure 10). The R87E mutation disrupted these interactions, leading to partial compensation through interactions with GLU-87 and HIS-493. This indicates that the R87E substitution induces subtle changes in the interactions of the CENP-T-W complex.

3.7 | Binding free energy assessment

After completing the MD simulations, we sought to investigate the impact of the mutations on the binding

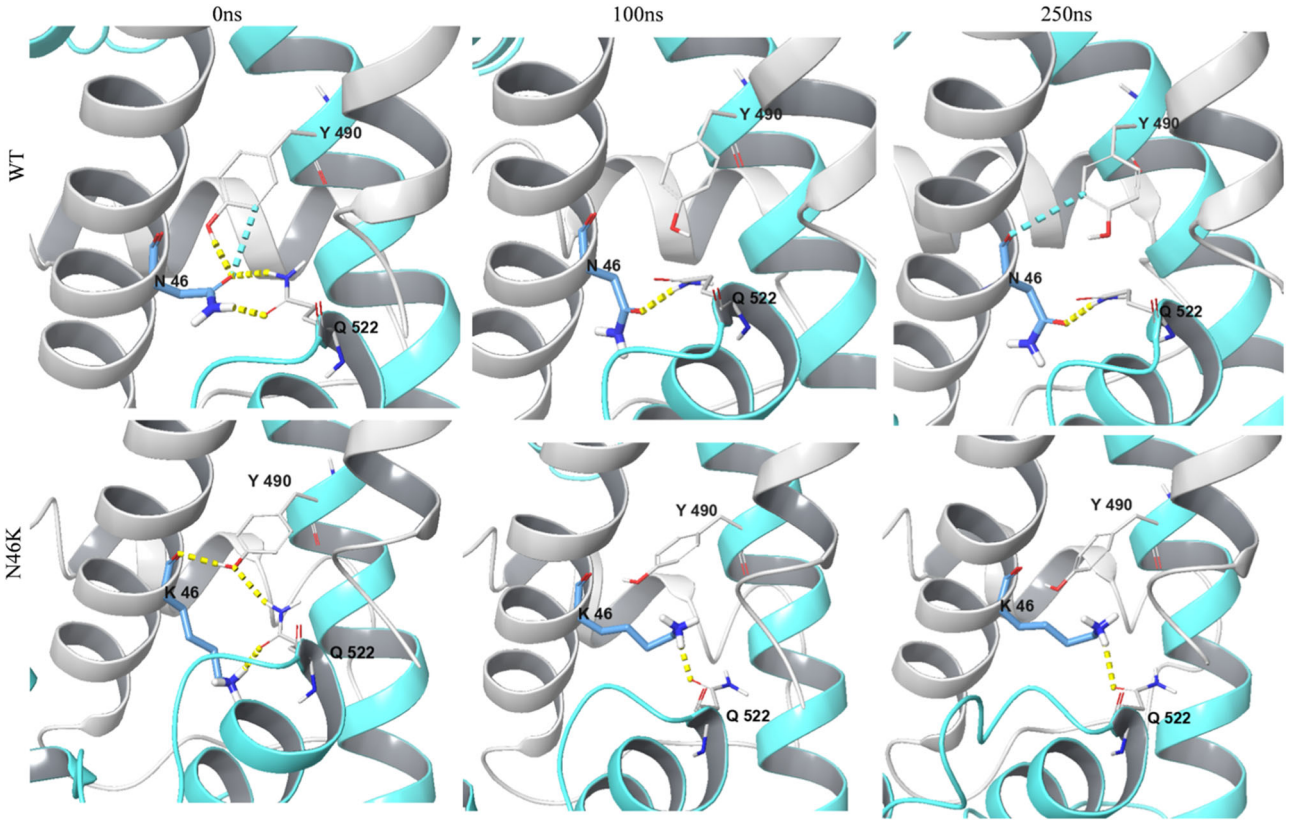

FIGURE 7 Effect of N46K substitution on interaction between CENP-T and WT/mutant CENP-W chains at the site-3 during the simulation period. The poses were extracted at 0, 100, and 250 ns MD simulation period.

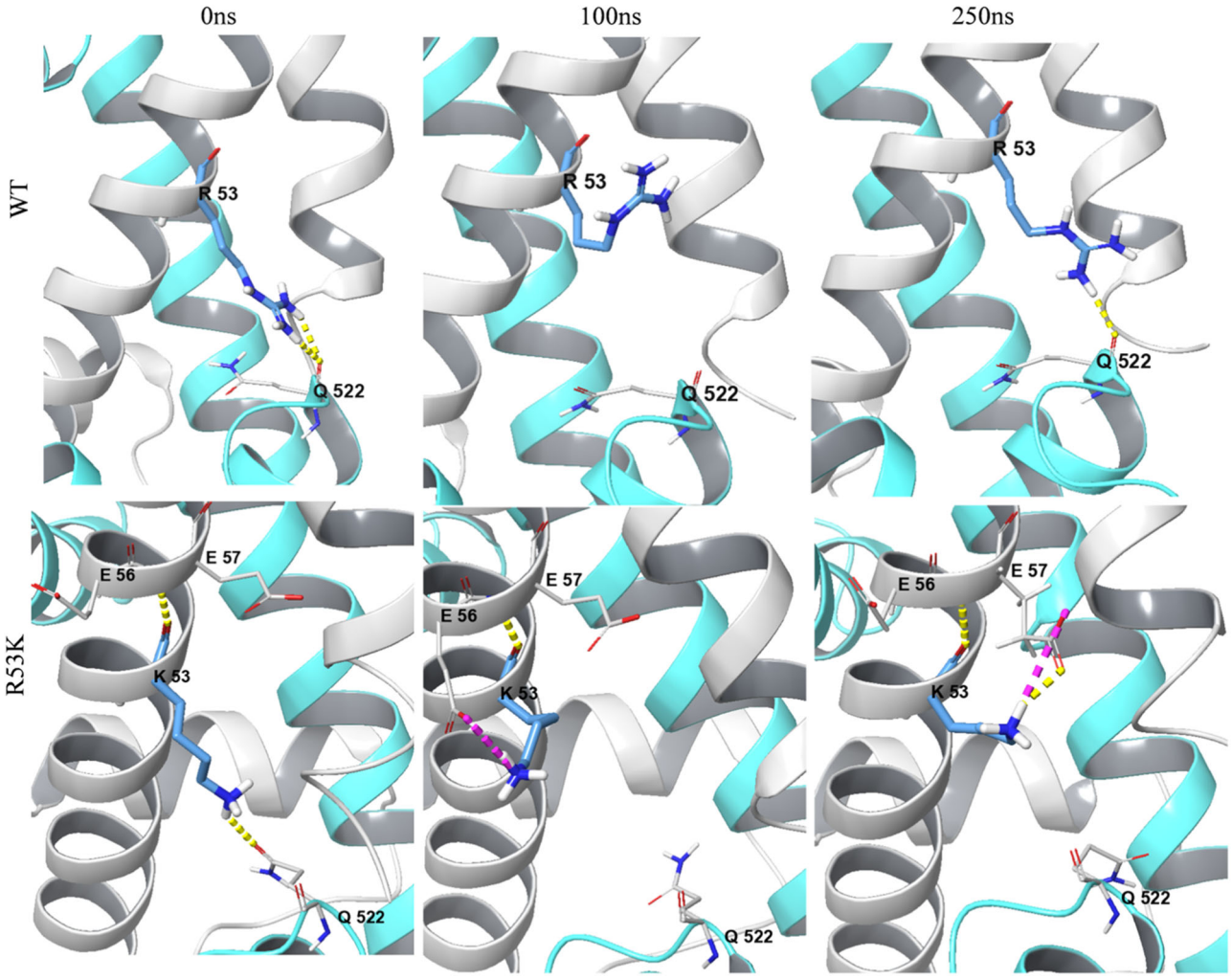

FIGURE 8 Effect of R53K substitution on interaction between CENP-T and WT/mutant CENP-W chains at the site-3 during the simulation period. The poses were extracted at 0, 100, and 250 ns MD simulation period.

free energy (BFE) of the CENP-T-W complex. To accomplish this, we employed the widely used and computationally efficient MM-GBSA (Molecular Mechanics-Generalized Born Surface Area) method.[35] MM-GBSA is well-suited for calculating the binding free energy of protein-protein complexes and evaluating their structural stability.[36] Specifically, it has been successfully utilized to identify critical amino acids involved in protein-protein interactions (PPI) and estimate how mutations at the interface alter the binding free energy.[36]

Previous studies have demonstrated that only a small number of key residues at the PPI interface can significantly influence the binding free energy of the protein-protein complex.[37] Therefore, MM-GBSA has the potential to reveal the important amino acid residues in PPI and elucidate the effects of substitutions on the protein-protein energetics.[36] For our analysis, we calculated the binding free energy of the test complexes, including the wild-type CENP-T-W complex (WT-CENP-T-W) and the CENP-T-W complex with mutations (N46K/R53K/L83K/R87E in the CENP-W chain), over the 250 ns simulation period.

The changes in binding free energy for the amino acid substituted complexes were compared to the wild-type complex, and the results are presented in Figure 11A. Our findings revealed that the amino acid substitutions at positions 46, 53, 83, and 87 in the CENP-W chain led to a decrease in the binding free energy of the mutant complexes compared to the WT CENP-T-W complex (Figure 11B). The order of change in binding free energy was found to be L83K > N46K > R53K > R87E during the 250 ns simulation period (Figure 11C).

These results indicate that LEU-83, ASN-46, and ARG-53 residues of CENP-W play crucial roles in the formation of the CENP-T-W complex. The decrease in binding free energy in these mutated complexes suggests potential alterations in the stability and interaction

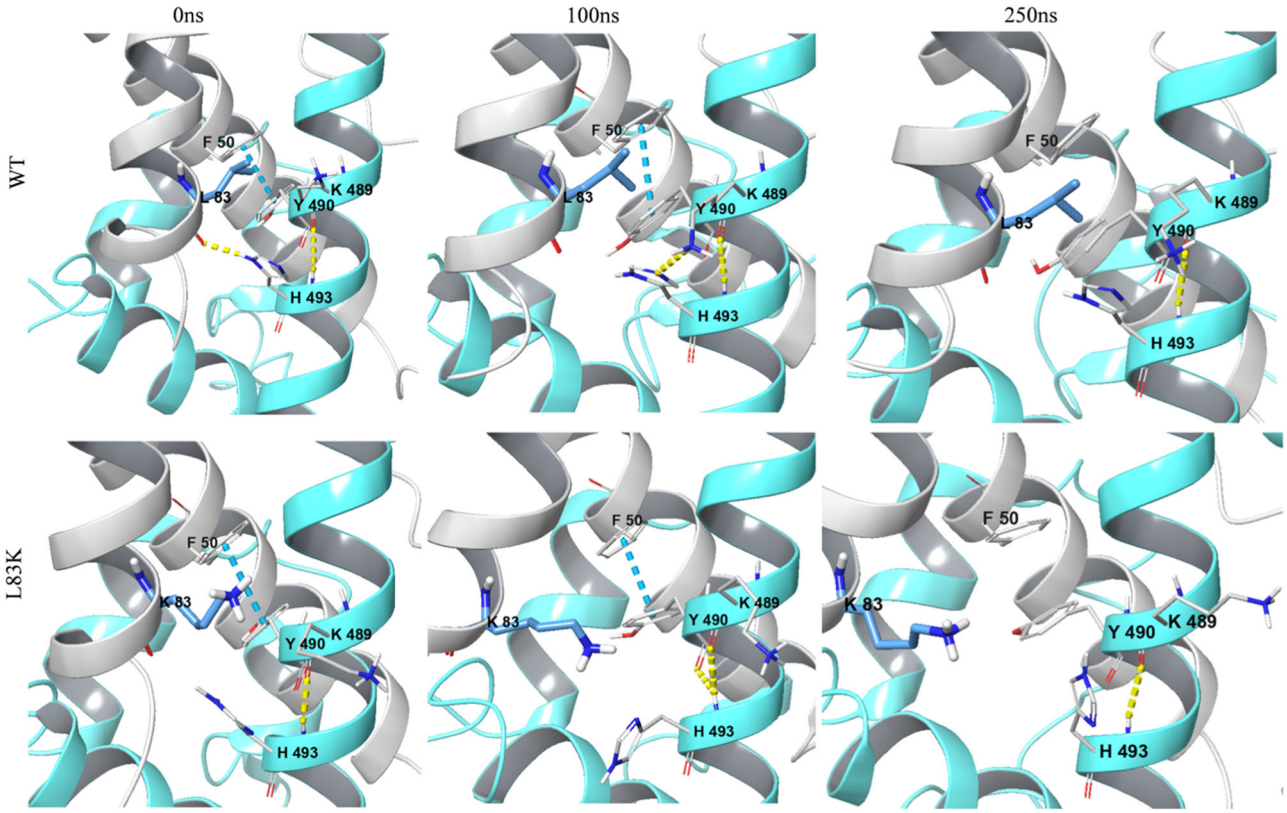

FIGURE 9 Effect of L83K substitution on interaction between CENP-T and WT/mutant CENP-W chains at the site-3 during the simulation period. The poses were extracted at 0, 100, and 250 ns MD simulation period.

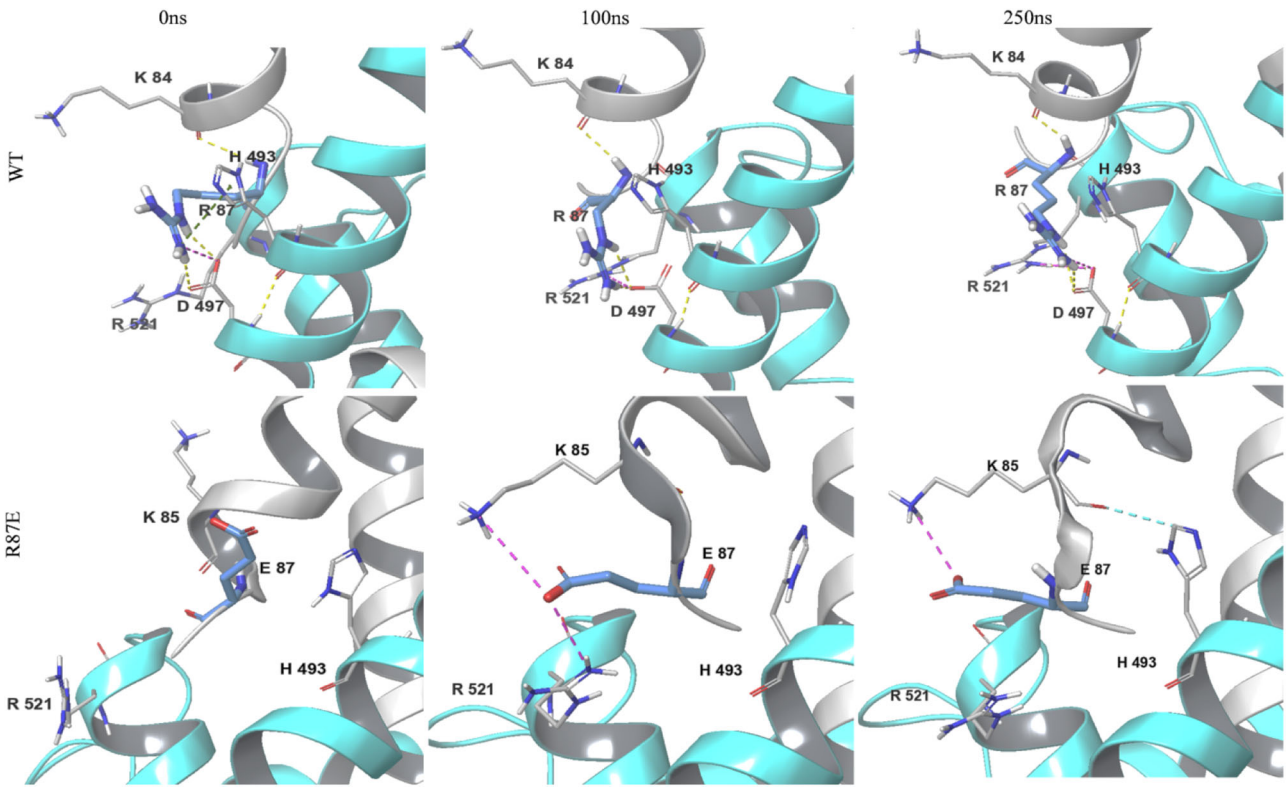

FIGURE 10 Effect of R87E substitution on interaction between CENP-T and WT/mutant CENP-W chains at the site-3 during the simulation period. The poses were extracted at 0, 100, and 250 ns MD simulation period.

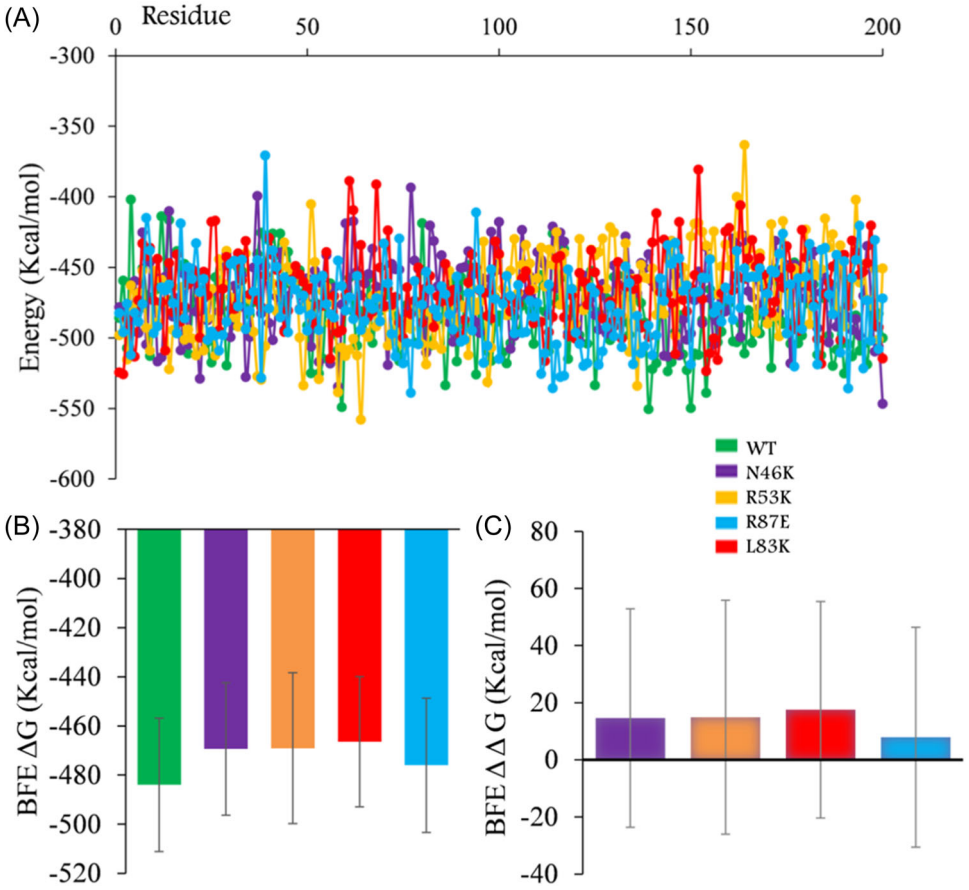

FIGURE 11 Binding free energy calculation for WT and mutant CENP-T-W complex. (A) Binding free energy change of the CENP-T-W complexes during 250-ns long MD simulation. (B) Binding free energy change (ΔG) in Kcal/mol for the mutant complex compared with the WT complex. (C) The predicted binding affinity for selected mutant. The ΔΔG was calculated by subtracting the binding free energy of WT protein from that of the mutant protein.

patterns within the protein-protein interface. These findings shed light on the significance of specific amino acid residues in governing the stability and energetics of the CENP-T-W complex and provide valuable insights into the functional implications of these mutations in the context of the protein complex formation. However, further experimental investigations and functional assays are warranted to validate and extend these observations and fully comprehend the biological consequences of these structural changes.

4 | CONCLUSION

The CENP-T-W heterodimer of the CENP-T/W/X/S hetero-tetramer holds substantial significance within the inner kinetochore assembly, playing a pivotal role in cellular mitosis. The CENP-W chain forms critical interactions with the CENP-T chain, creating a dimer that subsequently engages with the linker DNA.[6,8] The amino acids implicated in the protein-protein interaction (CENP-T-W) are not only crucial for the DNA-dimer interaction but also integral to the overall cellular mitosis process.[6,8,9] In this current study, we have employed *in silico* site-directed mutagenesis (amino acid substitution), molecular docking, MD simulation, binding free energy calculation to meticulously investigate the roles played by key amino acids in the formation of a stable CENP-T-W dimer. Notably, amino acid substitutions (N46K, R53K, L83K, and R87E) within the CENP-W chain led to pronounced increases in binding energy along with discernible structural alterations in their respective mutant complexes, when juxtaposed with the wild-type CENP-T-W dimer. A meticulous trajectory analysis, encompassing factors such as RMSD, RMSF, SASA, Rg, and more, undertaken during the 250 ns MD simulation period, unveiled that the introduced amino acid substitutions (N46K, R53K, L83K, and R87E in the CENP-W chain) had a substantial impact on both stability and conformational dynamics within the dimers.

Furthermore, calculations involving MM-GBSA corroborated the notion that amino acid substitutions had a

discernible impact, leading to an elevated binding free energy within the mutant dimers compared to the wild-type CENP-T-W dimer. A closer examination of the interface interaction pose between the wild-type and mutant CENP-T-W dimers spotlighted R53 and L83 as potential key residues located at site-3, indicating a plausible role in upholding interfacial stability. The structural elucidations gleaned from our investigation are poised to substantially enrich prospective experimental inquiries aimed at unraveling the intricate role of CENP-T-W in cellular mitosis.

AUTHOR CONTRIBUTIONS

Shashank Kumar, conceived and designed the study, and wrote the manuscript; Suryakanta Mohanty, and Rajendra Bhadane, performed the experiments; Shashank Kumar, Suryakanta Mohanty, and Rajendra Bhadane, analyzed and interpreted the data; Shashank Kumar, wrote the original draft of the manuscript. Shashank Kumar supervised the work, Rajendra Bhadane and Shashank Kumar critically reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS

S.K. acknowledges DST-India for providing DST-PURSE grant [SR/PURSE/2023/220]. S.K. also thanks DST-India for granting a Departmental grant in the form of a DST-FIST grant to the Department of Biochemistry at the Central University of Punjab in Punjab, India. R.B. wish to acknowledge CSC–IT Center for Science, Finland, for generous computational resources.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Rajendra Bhadane at https://doi.org/10.5281/zenodo.8241336. The data generated during the current study are available in the Zenodo repository with the DOI identifier https://doi.org/10.5281/zenodo.8241336.

ORCID

Rajendra Bhadane http://orcid.org/0000-0002-7330-4801
Shashank Kumar http://orcid.org/0000-0002-9622-0512

REFERENCES

1. Perpelescu M, Fukagawa T. The ABCs of CENPs. *Chromosoma*. 2011;120:425-446.
2. Bateman A, Martin MJ, Orchard S, et al. UniProt Consortium UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res*. 2023;51(D1):D523-D531. doi:10.1093/nar/gkac1052
3. Nishino T, Takeuchi K, Gascoigne KE, et al. CENP-TWSX forms a unique centromeric chromatin structure with a histone-like fold. *Cell*. 2012;148:487-501.
4. Takahashi K, Chen ES, Yanagida M. Requirement of mis6 centromere connector for localizing a cenp-a-like protein in fission yeast. *Science*. 2000;288:2215-2219.
5. Gascoigne KE, Cheeseman IM. Kinetochore assembly: if you build it, they will come. *Curr Opin Cell Biol*. 2011;23:102-108.
6. Hori T, Amano M, Suzuki A, et al. CCAN makes multiple contacts with centromeric dna to provide distinct pathways to the outer kinetochore. *Cell*. 2008;135:1039-1052. doi:10.1016/j.cell.2008.10.019
7. Takeuchi K, Nishino T, Mayanagi K, et al. The centromeric nucleosome-like CENP–T–W–S–X complex induces positive supercoils into DNA. *Nucleic Acids Res*. 2014;42:1644-1655.
8. Yatskevich S, Muir KW, Bellini D, et al. Structure of the human inner kinetochore bound to a centromeric CENP-A nucleosome. *Science*. 2022;376:844-852.
9. Prendergast L, Van Vuuren C, Kaczmarczyk A, et al. Premitotic assembly of human CENPs -T and -W switches centromeric chromatin to a mitotic state. *PLoS Biol*. 2011;9:e1001082.
10. Amano M, Suzuki A, Hori T, et al. The CENP-S complex is essential for the stable assembly of outer kinetochore structure. *J Cell Biol*. 2009;186:173-182.
11. Kaczmarczyk A, Sullivan KF. CENP-W plays a role in maintaining bipolar spindle structure. *PLoS One*. 2014;9:e106464.
12. Chiappori F, D'Ursi P, Merelli I, Milanesi L, Rovida E. In silico saturation mutagenesis and docking screening for the analysis of protein-ligand interaction: the endothelial protein C receptor case study. *BMC Bioinformatics*. 2009;10:S3.
13. Dehury B, Somavarapu AK, Kepp KP. A computer-simulated mechanism of familial Alzheimer's disease: mutations enhance thermal dynamics and favor looser substrate-binding to γ-secretase. *J Struct Biol*. 2020;212:107648. doi:10.1016/j.jsb.2020.107648
14. Dehury B, Raina V, Misra N, Suar M. Effect of mutation on structure, function and dynamics of receptor binding domain of human SARS-CoV-2 with host cell receptor ACE2: a molecular dynamics simulations study. *J Biomol Struct Dyn*. 2021;39:7231-7245. doi:10.1080/07391102.2020.1802348
15. Rodrigues CHM, Myung Y, Pires DEV, Ascher DB. mCSM-PPI2: predicting the effects of mutations on protein–protein interactions. *Nucleic Acids Res*. 2019;47:W338-W344.
16. DeLano WL. Pymol: an open-source molecular graphics tool. *CCP4 Newsletter Protein Crystallography*. 2002;40:82-92.
17. Kozakov D, Hall DR, Xia B, et al. The ClusPro web server for protein–protein docking. *Nat Protoc*. 2017;12:255-278.
18. Etemadi M, Rose J, Kool E. *Department of chemistry 184-biological chemistry laboratory manual*. Howard Hughes Medical Institute: Stanford University; 2009.
19. Lu C, Wu C, Ghoreishi D, et al. OPLS4: improving force field accuracy on challenging regimes of chemical space. *J Chem Theory Comput*. 2021;17:4291-4300. doi:10.1021/acs.jctc.1c00302

20. Bowers KJ, Chow E, Xu H. Scalable algorithms for molecular dynamics simulations on commodity clusters. *Proceedings of the ACM/IEEE Conference on Supercomputing, SC'06, Tampa, FL, USA, 11 November 2006*. ACM Press; 2006.
21. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: Pullman B, ed. *In Intermolecular Forces*. Springer; 1981: 331-342.
22. Nosé S. A unified formulation of the constant temperature. *Molecular Dynamics Methods J Chem Phys*. 1984a;81: 511-519.
23. Nosé S. A molecular dynamics method for simulations in the canonical ensemble. *Mol Phys*. 1984b;52:255-268.
24. Martyna GJ, Tobias DJ, Klein ML. Constant pressure. *J Chem Phys*. 1994;101:4177-4189.
25. Predescu C, Lerer AK, Lippert RA, et al. The u-series: a separable decomposition for electrostatics computation with improved accuracy. *J Chem Phys*. 2020;152:084113.
26. Jacobson MP, Pincus DL, Rapp CS, et al. A hierarchical approach to all-atom protein loop prediction. *Proteins: Struct, Funct, Bioinf*. 2004;55:351-367.
27. Jacobson MP, Friesner RA, Xiang Z, Honig B. On the role of the crystal environment in determining protein side-chain conformations. *J Mol Biol*. 2002;320:597-608.
28. Li J, Abel R, Zhu K, Cao Y, Zhao S, Friesner RA. The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins: Struct, Funct, Bioinf*. 2011;79:2794-2812.
29. Bhadane R, Salo-Ahen OMH. High-throughput molecular dynamics-based alchemical free energy calculations for predicting the binding free energy change associated with the selected omicron mutations in the spike receptor-binding domain of SARS-CoV-2. *Biomedicines*. 2022;10:2779. doi:10.3390/biomedicines10112779
30. Rana N, Singh AK, Shuaib M, et al. Drug resistance mechanism of M46I-mutation-induced saquinavir resistance in HIV-1 protease using molecular dynamics simulation and binding energy calculation. *Viruses*. 2022;14:697. doi:10.3390/v14040697
31. Singh AK, Kushwaha PP, Prajapati KS, Shuaib M, Gupta S, Kumar S. Identification of FDA approved drugs and nucleoside analogues as potential SARS-CoV-2 A1pp domain inhibitor: an in silico study. *Comput Biol Med*. 2021;130:104185. doi:10.1016/j.compbiomed.2020.104185
32. Berhanu WM, Masunov AE. Molecular dynamic simulation of wild type and mutants of the polymorphic amyloid NNQNTF segments of elk prion: structural stability and thermodynamic of association. *Biopolymers*. 2011;95:573-590.
33. Singh AK, Prajapati KS, Kumar S. Hesperidin potentially interacts with the catalytic site of gamma-secretase and modifies notch sensitive genes and cancer stemness marker expression in colon cancer cells and colonosphere. *J Biomol Struct Dyn*. 2022;41:8432-8444. doi:10.1080/07391102.2022.2134213
34. Ferdausi N, Islam S, Rimti F, et al. Point-specific interactions of isovitexin with the neighboring amino acid residues of the hACE2 receptor as a targeted therapeutic agent in suppressing the SARS-CoV-2 influx mechanism. *J Adv Veterinary Animal Res*. 2022;9:230-240. doi:10.5455/javar.2022.i588
35. Chen F, Liu H, Sun H, et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein–protein binding free energies and re-rank binding poses generated by protein–protein docking. *Phys Chem Chem Phys*. 2016;18:22129-22139.
36. Simões ICM, Costa IPD, Coimbra JTS, Ramos MJ, Fernandes PA. New parameters for higher accuracy in the computation of binding free energy differences upon alanine scanning mutagenesis on protein–protein interfaces. *J Chem Inf Model*. 2017;57:60-72.
37. Potapov V, Cohen M, Schreiber G. Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details. *Protein Eng Des Sel*. 2009;22: 553-560.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mohanty S, Bhadane R, Kumar S. Bioinformatics insights into CENP-T and CENP-W protein-protein interaction disruptive amino acid substitution in the CENP-T-W Complex. *J Cell Biochem*. 2023;1-16. doi:10.1002/jcb.30495