Dynamic play between human N-α-acetyltransferase D and H4-mutant histones: Molecular dynamics study

Shravan B. Rathod[a,*] and Kinshuk Raj Srivastava[b,*]

[a]*Department of Chemistry, Smt. S. M. Panchal Science College, Talod, Gujarat, India*

[b]*Division of Medicinal and Process Chemistry, CSIR-Central Drug Research Institute (CDRI), Lucknow, Uttar Pradesh, India*

*Corresponding author:

Shravan B. Rathod (shravanathorizon93@gmail.com)

Kinshuk Raj Srivastava (kinshuk.srivastava@cdri.res.in)

Abstract

N-terminal acetyltransferases (NATs) are overexpressed in various cancers. Specifically in lung cancer, human N-α-acetyltransferase D (hNatD) is upregulated and prevents the histone H4 N-terminal serine phosphorylation, leading to the Epithelial-to-mesenchymal transition (EMT) of cancer cells. hNatD facilitates histone H4 N-α-terminal serine acetylation and halts the CK2α-mediated serine phosphorylation. In the present study, we report the effects of four N-terminal mutant (S1C, R3C, G4D and G4S) histone H4 peptides on their bindings with hNatD by employing a molecular dynamics simulation. We also used graph theory-based analyses to understand residue correlation and communication in hNatD under the influence of WT and MT H4 peptides. Results show that S1C, R3C and G4S mutant peptides have significant stability at the catalytic site of hNatD. However, S1C, G4D and G4S peptides disrupt hNatD structure. Additionally, intramolecular hydrogen bond analysis reveals greater stability of hNatD in complex with R3C peptide. Further, intermolecular hydrogen bond analysis of acetyl-CoA with hNatD and its RMSD analysis in five complexes indicate that cofactor has greater stability in WT and R3C complexes. Our findings support previously reported experimental study on impacts of H4 mutations on its hNatD-mediated acetylation catalytic efficiency. The Betweenness centrality (BC) analysis further gives insight into the hNatD residue communication dynamics that can be exploited to target hNatD using existed or novel drug candidates therapeutically.

Keywords: Epigenetic modification, Histone acetylation, Oncohistone mutations, N-α-acetyltransferase D, Molecular docking, Molecular dynamics simulations, Residue correlation

Introduction

Oncogenesis is complex and multistep process involves mutations in epigenetic enzymes. Current experimental evidences indicate that histones are found to be mutated in numerous cancers. For instance, histone H3 was reported mutated in rare paediatric gliomas and sarcomas (Wan et al., 2018). Additionally, in large B-cell lymphomas, histone H1 was found mutated (Yusufova et al., 2021). Nacev et al. reported cancer associated histone mutations recently (Nacev et al., 2019).

Epigenetic histone tail modifications portray vital role in gene expression at the transcription site for the protein synthesis. While modification at the nucleosome region facilitates the chromatin condensation and regulates various nuclear events (Biancotto et al., 2010). Among the phosphorylation, methylation, acetylation, deacetylation, ubiquitylation, and sumoylation Post-translational modifications (PTMs), acetylation and methylation have been widely studied (Cavalieri, 2020). N-α-terminal acetylation is most frequently occurring covalent modification in approximately 90% soluble proteins in humans and around 70% in yeast (Aksnes et al., 2016; Arnesen et al., 2009; J. L. Brown & Roberts, 1976; Starheim et al., 2012). This covalent modification is involved in cellular metabolism, apoptosis, protein translocation, membrane attachment, protein-protein interactions, protein regulation and degradation (Forte et al., 2011; Hwang et al., 2010; Starheim et al., 2012; Yi et al., 2011).

N-α-acetyltransferases (NATs) catalyse N-terminal acetylation of alpha-amino group of protein by using acetyl-coenzyme A (Ac-CoA). There have been total six NATs reported in humans according to their specific subunits and substrates (Starheim et al., 2012). hNatD is also known as Patt1 (Protein acetyltransferase) or Nat4 exclusively acetylates the histone H2A and H4 compared to other NATs those target numerous other substrates (Magin et al., 2015; Mullen et al., 1989; Tran et al., 2011). Chen A. et al. have recently reported that hNatD is overexpressed in lung cancer and it elevates the cancer progression by blocking the Protein kinase 2-alpha (CK2α)-mediated histone H4 serine phosphorylation. hNatD acetylates the histone H4 serine at alpha-amino position competitively and increases the rate of Epithelial-to-mesenchymal transition (EMT) of cells in lung cancer (Ju et al., 2017). The hNatD acetylates the N-terminus residues of H4 and H2A by using acetyl-CoA cofactor located inside the catalytic tunnel in hNatD. Figure 1 illustrates the acetylation of H4/H2A by hNatD. Recently, how the activity and catalytic efficiency of hNatD are altered under the influence of WT and MTs pentapetides binding to hNatD has been reported (Ho & Huang, 2021). However, the molecular level understanding of H4 (WT & MTs) binding dynamics still lacks.

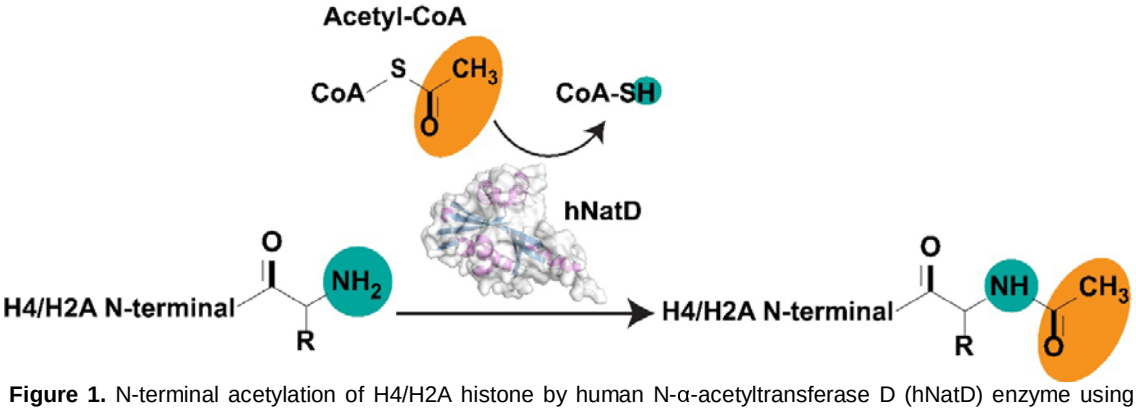

Figure 1. N-terminal acetylation of H4/H2A histone by human N-α-acetyltransferase D (hNatD) enzyme using acetyl-CoA cofactor.

In the present study, our aim was to investigate the impacts of four mutations at the histone H4 tail on its binding at the hNatD acetylation site and, to probe the structural and dynamic characteristics of hNatD to target it therapeutically in the future. We used Molecular dynamics simulation (MDS) approach to probe the binding events of hNatD and histone H4 (WT and MTs) tails. Additionally, Residue interaction network (RIN) based analyses were carried out to further understand hNatD structural dynamics.

Materials and methods

Protein preparations

The crystal structure (PDB: 4U9W) (Magin et al., 2015) of hNatD (a.a.: 1-207) with histone H4 tail and Acetyl-CoA was retrieved from the Protein data bank (PDB). SWISS-MODEL(Schwede et al., 2003) available at http://swissmodel.expasy.org was employed to add missing residues in the β6-β7 loop and N-terminal α0-helix in hNatD. Additionally, the missing acetyl group in Acetyl-CoA was added using Builder tool in PyMOL 2.4.1 (Schrödinger LLC., 2010).

Protein-peptide docking

In the crystal structure, histone H4 tail had five residues (SGRGK). To fully investigate the impacts of four mutations (S1C, R3C, G4D, and G4S) of histone H4 tail on its binding with hNatD, we wanted to add another five residues to the H4 tail. Thus, the HPEPDOCK protein-peptide docking web server (Zhou et al., 2018) available at http://huanglab.phys.hust.edu.cn/hpepdock/ was employed in which modelled hNatD was uploaded as a receptor and for ligand (H4 tail), 10 aa short sequence (SGRGKGGKGL) of H4 tail was provided in the ligand box. In the case of sequence query, HPEPDOCK utilizes MODPEP program (Yan et al., 2017) to generate ensembles of peptides and further these ensembles of peptides are docked with receptor using MDock program suite (Huang & Zou, 2011), in which rigidity is applied to the receptor and flexibility is given to peptide during the docking. To determine the 3D interactions of H4 peptide with hNatD, Discovery Studio v.20.1 (BIOVIA, Dassault Systemes, Discovery Studio Visualizer, 20.1.0.19295, San Diego: Dassault Systemes, 2020) was used.

Molecular dynamics simulation

Molecular dynamics simulation was employed to scrutinize binding between the histone H4 (WT & MTs) and hNatD. Four mutations in histone H4 peptide were inserted using Mutagenesis tool in PyMOL. These four mutants and wildtype complexes were further subjected to the 100 ns Molecular dynamics (MD) simulation. Molecular dynamics simulation was carried out using GROMACS 2020.1 version (Abraham et al., 2015). The updated version of CHARMM36 force filed (charmm36-feb2021) (Best et al., 2012; MacKerell et al., 1998) and TIP3P water model (Jorgensen et al., 1983) were employed. To build the Acetyl-CoA topology, CHARMM general force field (CGenFF) (Allouche, 2012; Yu et al., 2012) available at (https://cgenff.umaryland.edu/) was utilized. For the simulations, these four complexes were placed at the center inside the cubic box with a 10 Å from the box edges. Further, positive (Na^+) and negative (Cl^-) ions were added to neutralize the solvation system. Initially, the system was energy minimized using the steepest descent algorithm (Chen et al., 2013) to reduce the steric clashes between the atoms. For the optimization of long-range electrostatic forces, Verlet cut-off (Verlet, 1967) and Particle mesh Ewald (PME) schemes (Darden et al., 1993) were employed. During the energy minimization protocol, nsteps was set to 50,000 and energy cut off was kept 10 kJ/mol. Energy minimization was followed by two subsequent equilibrations (NVT and NPT) at 300 K temperature for 100 ps (50,000 steps) with a 2 fs time step gap. Then, MD simulations was carried out for the 100 ns using leap-frog integrator (Lippert et al., 2007), Verlet cut-off scheme, and PME at 300 K (modified Berendsen thermostat (Berendsen et al., 1984)) and 1 bar pressure (Parrinello-Rahman method (Parrinello & Rahman, 1981)). LINCS (Linear constraint solver) algorithm (Hess et al., 1997) was utilized to constrain bonds and angles. CHARMM36 force filed and python script were downloaded from the MacKerell lab website (http://mackerell.umaryland.edu/charmm_ff.shtml#gromacs). Finally, the MD trajectories were analyzed for the Root mean square deviation (RMSD), Solvent accessible surface area (SASA), Radius of gyration (Rg), distances, Root mean square fluctuation (RMSF) and, inter and intramolecular hydrogen bonds in the complexes.

Protein-peptide interactions analysis

To investigate interactions between histone H4 and hNatD, ten frames from the 100 ns each simulation were extracted at every 10 ns from 10 ns to 100 ns. The web-based tool, PDBePISA (https://www.ebi.ac.uk/pdbe/pisa/) was employed to calculate numbers and types of interactions in 10 frames of each four complexes; hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C, hNatD-H4_G4D and hNatD-H4_G4S. We considered dominant interactions in 10 frames of each complex and they are illustrated. PDBePISA calculates interface area and number of interacting residues between the proteins and solvation energy along with interactions. To calculate van der Waals, hydrogen bond, electrostatic and stabilization energies of WT and MT complexes, PIMA web server (Mathew et al., 2016) available at http://caps.ncbs.res.in/pima/ was also employed.

Linear mutual information

To probe the dynamical behaviour and fluctuations of proteins, Dynamical cross-correlation (DCC) and Linear mutual information (LMI) are used widely (Negre et al., 2018; Penkler et al., 2018; Sethi et al., 2009; Van Wart et al., 2014). Considering the limitation of DCC that it does not calculate correlation of simultaneously moving atoms in perpendicular directions (Tekpinar et al., 2021), we utilized normalized LMI (nLMI). To perform the nLMI on complexes, we used recently developed python 3-based tool, *correlationplus* (Tekpinar et al., 2021). To calculate nLMI, we provided 100 ns MD trajectory files (. XTC format) as an input to the program. From the 100 ns trajectories, it used 10,000 frames from each trajectory to calculate nLMI. Additionally, we calculated residue betweenness to find significant correlated residues. Commands to run *correlationplus* is given in Table S1.

Electrostatic potential (ESP)

The eF-surf web server available at https://pdbj.org/eF-surf/top.do was used to calculate electrostatic potential surface of hNatD. This server requires only protein structure file in pdb format as an input file. Avogadro (Hanwell et al. 2012) was used to generate ESP surface of the acetyl-CoA while PyMOL APBS Electrostatics plugin was used to generate ESP of histone H4 WT and MTs tails.

Results and discussion

Structural features of hNatD

The structural features of hNatD are resemble to the previously reported other NATs (N-acetyltransferases) and it contains α/β mixed GCN5-like fold (Figure 2) (Liszczak et al., 2011, 2013; Liszczak & Marmorstein, 2013). From the comparison of hNatD with NatA and NatE, it was observed that hNatD has some specific additional structural features which are loops formed by α1-α2 and β6-β7 and, unique N-terminal helix (α0). The hNatD sequence alignments among various organisms revealed that a specific region, 55-216 a.a. (~150 residues) is highly conserved in hNatDs. For the substrate recognition by NatA and NatE, they have β6-β7 hairpin loop region. However, there is significant structural difference observed at the catalytic site of hNatD. In hNatD, β6-β7 loop is located slightly away from the substrate binding site (Magin et al., 2015).

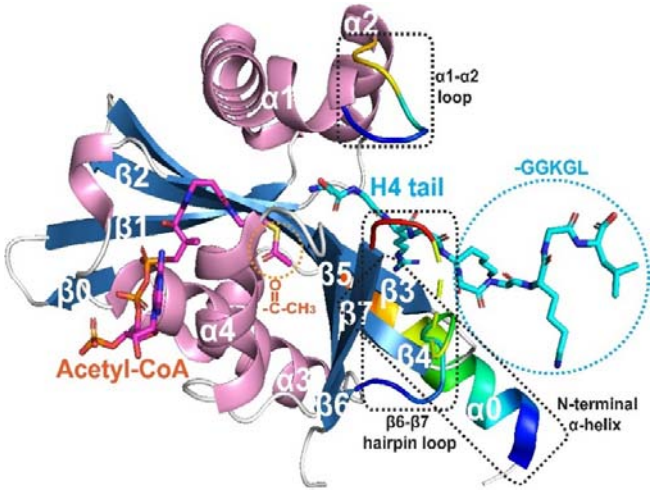

Figure 2. Structure of hNatD with acetyl-CoA and histone H4-tail (PDB: 4U9W). In this crystal structure, acetyl group was missing that was added and another five amino acids were added to the H4 peptide. Circular dashed rings indicate the modification sites. Three important structural features of hNatD, α1-α2 loop, β6-β7 loop and N-terminal α-helix are represented in rainbow color and shown in rectangular dashed boxes.

In addition to this, α1-α2 loop further protects catalytic region. Another interesting feature is helix-loop-strand forming N-termini region (24-51 a.a.). This region stabilizes the whole hNatD by wrapping around the hydrophobic patch of substrate recognition site. Five main residues, P38, L39, F42, F45 and Y48 of N-terminal interact with the L53, I57, C59, W107, L109, F124, F126, L151, F154, I158, L161 and M162 residues of catalytic site. These observations indicate that hNatD has specific structural features to recognize histone H4 or H2A.

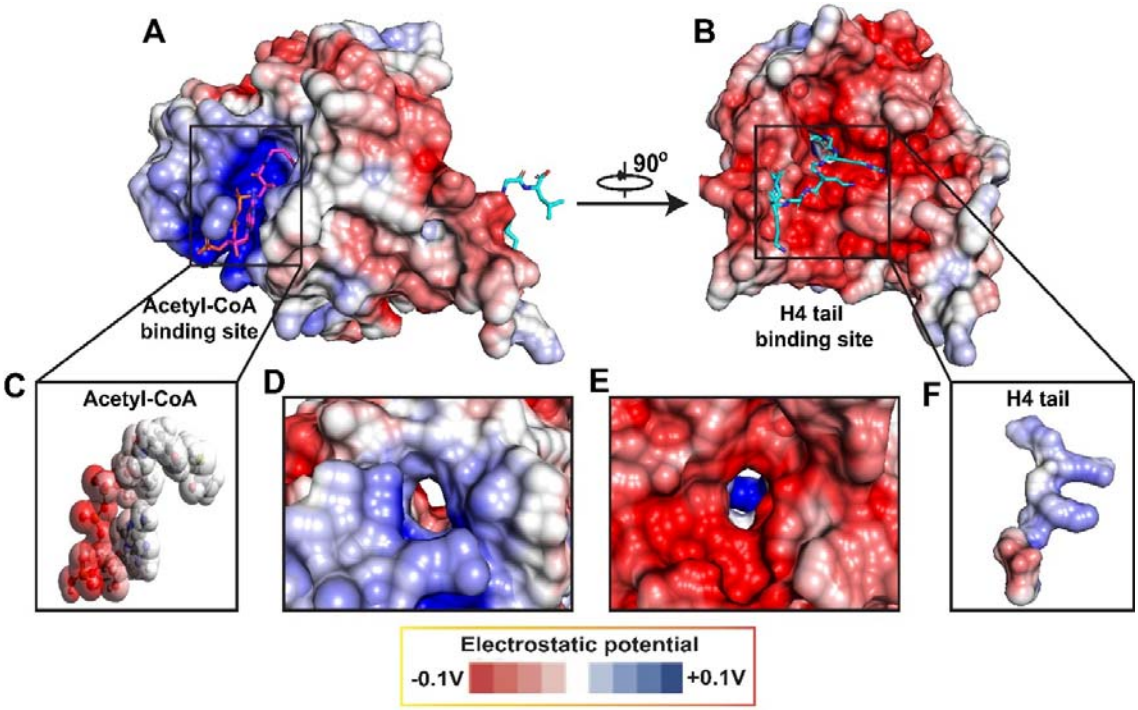

Figure 3. Electrostatic potential (ESP) surface representations. (A) Positively charged binding site of acetyl-CoA. (B) Negatively charged binding site of histone H4 tail. (C) ESP of acetyl-CoA (D) Catalytic tunnel from acetyl-CoA binding site to H4 tail binding site. (E) Catalytic tunnel from H4 tail binding site.to acetyl-CoA binding site. (F) ESP of histone H4 tail.

The Electrostatic potential (ESP) surface analysis is crucial to understand molecular mechanisms such as intermolecular association of protein and small ligands, actions of drug candidate, enzyme

catalysis and protein-DNA interactions complementarity (Weiner et al., 1982). In this study, we also calculated electrostatic potential surface of the hNatD. It can be seen from Figure 2, left-hand side on hNatD has positively charged residues (R134, R135, K136 and K140) containing cleft which harbours acetyl-CoA (Figure 3(A)). As acetyl-CoA has negatively charged phosphate backbone (Figure 3(C)), it perfectly positioned itself into this cleft. In contrast to that on the right-hand side on hNatD, histone H4 binding site has negatively charged residues (E76, E83, E86, E87, D90, D114, E116, E126 and E194) which recognises positively charged histone H4 tail (Figure 3(B)). Histone H4 tail (SGRGKGGKGL) has arginine and lysine basic residues which make this surface positively charged (Figure 3(F)). hNatD exhibits long tunnel to harbour acetyl-CoA and histone H4 tail (Figure 3(D-E)).

Protein-peptide docking

Molecular docking is used widely to investigate the bindings of protein with protein, peptide, small ligand, DNA and RNA in Computer-aided drug design (CADD) and other computational analyses to complement experimental findings (Aamir et al., 2018; Azam & Abbasi, 2013; Blanco et al., 2019; de Ruyck et al., 2016; Panda et al., 2020; Parisien et al., 2012; Rathod et al., 2020). In this study, to probe the impacts of four mutations (S1C, R3C, G4D, and G4S) in histone H4 tail (10 a.a.), we first docked histone H4 tail with hNatD using HPEPDOCK protein-peptide web server (Zhou et al., 2018). From the top 100 predictions, best model was selected based on docking score (-172.891 kcal/mol). This docked model was structurally aligned with crystal structure (PDB: 4U9W) in PyMOL and RMSD was found 0.220 Å (Figure S1). Further, we obtained protein-peptide 3D interaction profile using Discovery Studio v.20.1 tool. Figure 4 illustrates best docking pose (Figure 4(A)) and 3D interactions between the hNatD and histone H4 tail (Figure 4(B)).

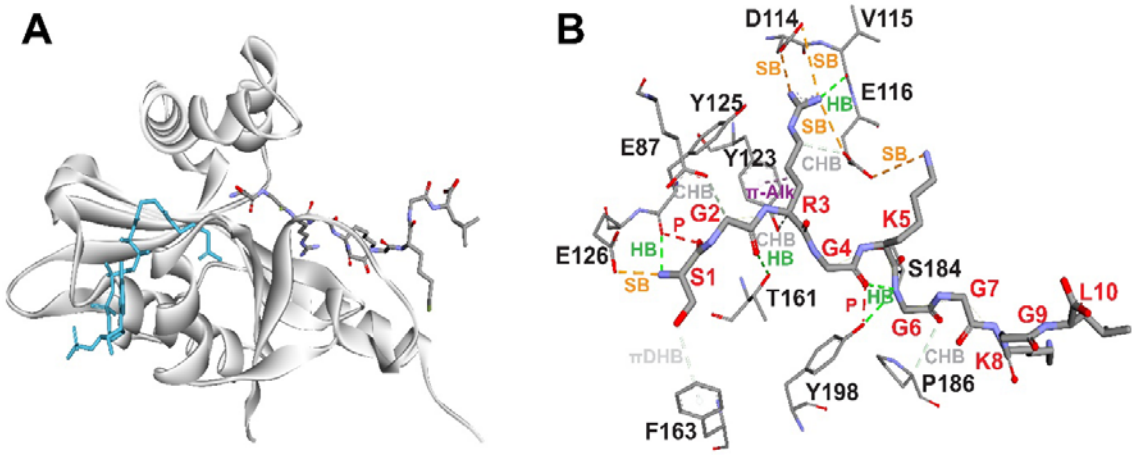

Figure 4. (A) H4 peptide docked structure. Acetyl-CoA and H4 peptide are shown in stick representation and have cyan and grey colors respectively. (B) 3D-interactions between hNatD and H4 peptide. H4 peptide amino acids' labels are colored in red and hNatD in black. HB: hydrogen bond, SB: salt bridge, CHB: carbon hydrogen bond, π-Alk: pi-alkyl bond, πDHB: pi-donor hydrogen bond and P: polar.

It can be seen from the Figure 3(B) that histone H4 tail is surrounded with mainly acidic and aromatic residues. Histone tail forms salt bridges with the acidic residues and hydrogen bonds with other residues of the hNatD. And, two aromatic residues (Y123 and F163) from the hNatD interact through donating pi electrons to the peptide residues. Additionally, polar interactions (carbon-hydrogen bond) can be observed between hNatD and histone H4 tail.

Impacts of histone H4 tail binding on the dynamics of hNatD

To target hNatD therapeutically, histone H4 interacting residues and their communication are significant. Thus, we analysed the impacts of WT and MT histone H4 decapeptides bindings on dynamic behaviour of hNatD through employing MD simulation and nLMI analyses. Ligand binding to protein induces structural alternations in proteins (Celej et al., 2003; Kumar & Jernigan, 2021; Settanni et al., 2013). To understand which regions of the hNatD fluctuate largely, we aligned MD trajectory frames from 0 ns to 100 ns at every 10 ns for each case (Figure 5). The average RMSDs of

hNatD after aligning frames in each complex were observed 0.91 Å, 0.91 Å, 0.87 Å, 1.18 Å and 0.85 Å for hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C, hNatD-H4_G4D and hNatD-H4_G4S respectively. Among these five complexes, hNatD-H4_R3C (Figure 5(C)) and hNatD-H4_G4S (Figure 5(E)) showed considerably less deviation in comparison with other complexes (Figure 5(A-B, D)). Radar plots illustrate the H4 head and tail dynamics in all five complexes (Figure 5(F-G)). In radar plots, distances of Cα-atom of first residue in every frame (10 ns-100 ns at every 10 ns) with respect to Cα-atom of first residue in 0 ns frame were calculated and plotted. It can be observed from the Figure 5(F) that heads of WT and G4D histones showed greater deviation (20 Å-25 Å) from their original positions at 0 ns frame in comparison with S1C, R3C and G4S MTs histones which showed at around 5 Å-10 Å. However, in case of H4 histones, distances of Cα-atom of last residue (L10) in every frame (10 ns-100 ns at every 10 ns) with respect to Cα-atom of last residue in 0 ns frame were calculated and plotted. However, in this case, except R3C, every histone showed large fluctuations and deviations were around 15 Å-25 Å (Figure 5(G)).

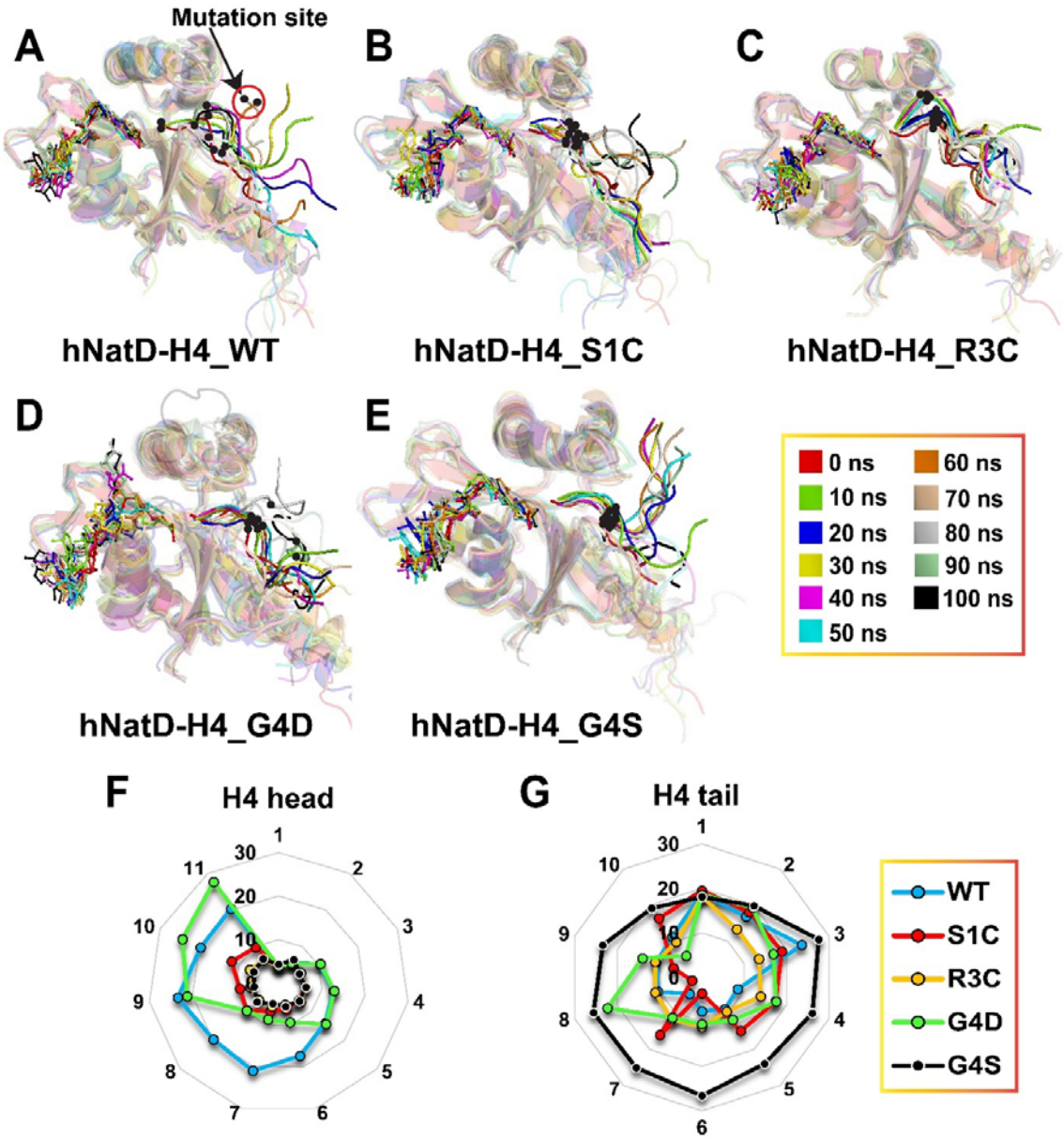

Figure 5. Structural alignments of 100 ns MD simulation frames at every 10 ns from 0 ns to 100 ns. (A) hNatD-H4_WT complex. (B) hNatD-H4_S1C complex. (C) hNatD-H4_R3C complex. (D) hNatD-H4_G4D complex. (E) hNatD-H4_G4S complex. (F) Radar plot of histone H4 (WT & MTs) heads' movements. (G) Radar plot of histone H4 (WT & MTs) tails' movements. In radar plots, labels on circular path indicate the number of frame (1-11) whereas vertical values are distance ranges (0-30 Å) between H4 heads or tails frames. Black dots designate the mutation position in H4 tail.

To understand the histone H4 tail behaviour inside the catalytic site of hNatD, we also calculated the RMSDs of WT and MTs histone H4 tails (Figure 6(A)). Mutant histone H4_G4S showed larger deviation (RMSD: 1.21 nm) from the initial position which is responsible to the higher flexibility of its tail. The average RMSDs of H4_WT, H4_S1C, H4_R3C and H4_G4D histone H4 tails were observed 0.96 nm, 0.97 nm, 0.88 nm and 0.94 nm respectively. Thus, except H4_R3C, all the WT and MTs H4 tails have greater dynamicity inside the hNatD cavity. We also determined dynamics of acetyl-CoA inside the hNatD cavity under the influence of histone H4 WT and MTs tails bindings to the hNatD. Figure 6(B) shows the RMSD of acetyl-CoA in hNatD of all five complexes. Acetyl-CoA showed greater fluctuations (0.54 nm) inside the binding pocket in hNatD of hNatD-H4_G4D as compared to other complexes. The average RMSDs of acetyl-CoA in hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C and hNatD-H4_G4S were found

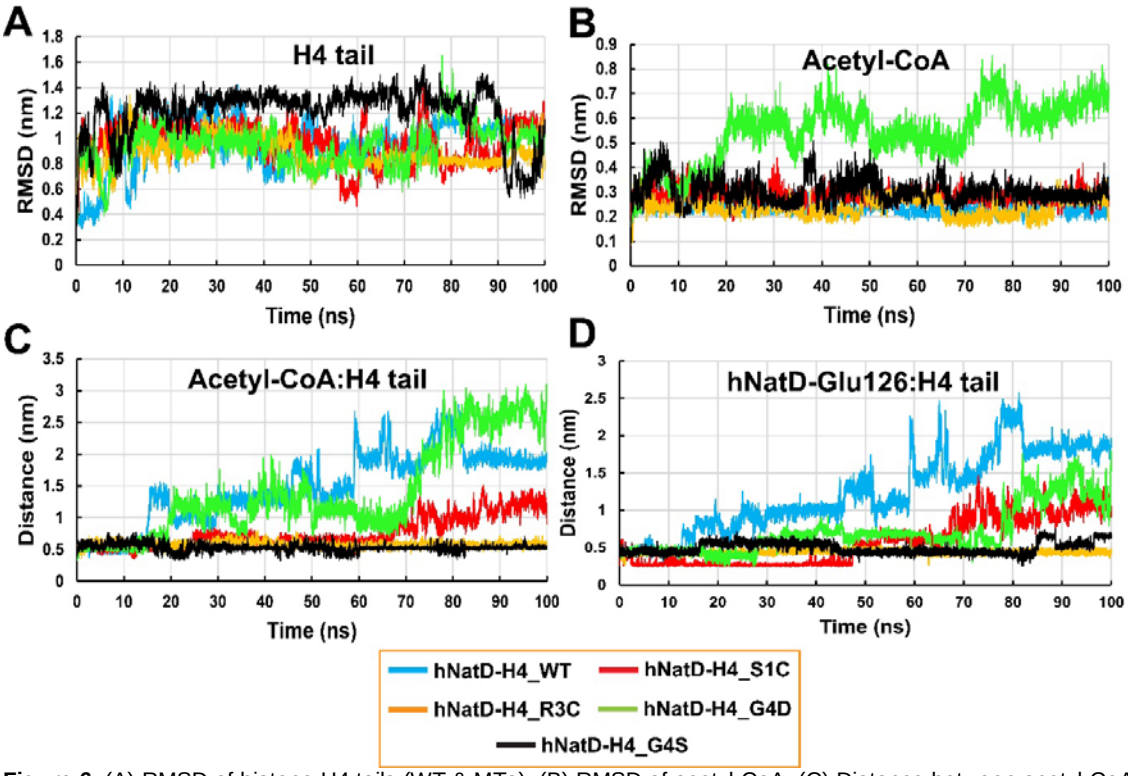

Figure 6. (A) RMSD of histone H4 tails (WT & MTs). (B) RMSD of acetyl-CoA. (C) Distance between acetyl-CoA sulfur (S) atom and nitrogen (N) atom of amino group in residue 1 (Serine/Cysteine) of histone H4 tails (WT & MTs). (D) Distance between hNatD E126 carboxylic oxygen (O) atom and nitrogen (N) atom of amino group in residue 1 (Serine/Cysteine) of histone H4 tails (WT & MTs).

0.24 nm, 0.28 nm, 0.24 nm and 0.30 nm respectively. To find the cause of this trend, we looked at the electrostatic surface of the WT and MTs H4 histones (Figure S2) and hNatD at the binding pocket (Figure 3(B)). The histone H4 binding cavity has the dominant glutamic acidic residues which make pocket negatively charged surface. Except the H4_G4D, all histones have positively charged surface (Figure S2). In case of H4_G4D, glycine smaller amino acid is substituted with acidic aspartic acid residue which carries negatively charged carboxylic group (Figure S2(D)). Thus, due to the repulsion between the H4_G4D and hNatD catalytic cleft, this mutation significantly disrupts the residue network in hNatD_H4_G4D and also the acetyl-CoA fluctuates largely (Figure 5(D)).

The hNatD acetylates the histone H4 and H2A N-terminals using acetyl-CoA cofactor (Magin et al., 2015). For the acetylation, the distance between the amino group of residues in histones and sulfur atom of acetyl-CoA must be in close proximity. Thus, we calculated distances between amino group nitrogen of S1/C1 and sulfur atom of cofactor in WT and MTs complexes (Figure 6(C)). In WT and G4D, the position of the first residue deviated greatly from its original position thus, distances were observed significantly large (~5-25 Å) during the 100 ns simulation. However, first residue in S1C and G4S was observed highly stable throughout the simulation and average distance was noted 5 Å. In R3C, distance was observed approximately 5 Å until 65 ns and then, it was increased and reaches to the around 15 Å at the end of the simulation. Additionally, E126 residue in hNatD was confirmed to be important to catalyse the acetyltransferase reaction not by directly acting as deprotonating agent but by increasing the nucleophilicity of S1/C1 amino group through inducing dipole moment from the carboxylic group. Also, it is believed that this residue orients the S1/C1 amino group towards the acetyl-CoA to facilitate reaction and it is considered essential for the catalysis (Magin et al., 2015). Thus, we calculated the distance between the carboxylic oxygen (-CO(O̲H)) of E126 and amino group nitrogen (N) atom of S1/C1 histone tail (Figure 6(D). It can be observed that except WT, in all other complexes, until around 65 ns, distance was remained approximately 5 Å. And, more importantly, it was remained stable during the entire simulations of hNatD-H4_R3C and hNatD-H4_G4S complexes. Hence, this data suggests that E126 is vital residue for the catalysis and also validate the previous finding.

After performing 100 ns simulations, the dynamic behaviour of hNatD was determined using various post-MD trajectory analyses. The Root mean square deviation (RMSD) of hNatD in five complexes (hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C, hNatD-H4_G4D and hNatD-H4_G4S) was calculated using backbone atoms to understand the conformational changes in hNatD under the influence of histone H4 tail (WT & MTs) binding. The RMSD can be calculated using following equation;

$$\boldsymbol{RMSD} = \left(\frac{\mathbf{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{N}}\left(\boldsymbol{r'}_{\boldsymbol{i}}(\boldsymbol{t})\right) - \left(\boldsymbol{r}_{\boldsymbol{i}}\left(\boldsymbol{t}_{\boldsymbol{ref}}\right)\right)^{\mathbf{2}}\right)^{\frac{\mathbf{1}}{\mathbf{2}}} \quad (1)$$

Where, N indicates the total number of particles (atoms), r_i is the position of atom i at reference time t_{ref}, r' is the position of the selected atom at time t after aligning with reference frame. Figure 7(A) shows the RMSD plot of hNatD in studied five complexes. The average RMSD of hNatD in hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C, hNatD-H4_G4D and hNatD-H4_G4S was observed 0.25 Å, 0.25 nm, 0.24 nm, 0.28 nm and 0.27 nm respectively. Thus, data indicates that hNatD has more structural deviations and fluctuations in hNatD-H4_G4D and hNatD-H4_G4S as compared to others. It can be seen from the RMSD plots also in Figure 7(A). The hNatD in hNatD-H4_R3C has the highest stability (RMSD: 0.24 nm) and throughout the 100 ns simulation, RMSD oscillates between 0.20 nm and 0.25 nm. However, in case of hNatD-H4_G4D and hNatD-H4_G4S, there were significant changes noted during the simulation and RMSDs were fluctuated between 0.15 nm and 0.43 nm. Figure S3-S7 shows hNatD regions which are significantly dynamic in nature. In all the complexes, mainly four regions, N-terminal α0-helix (~1-20 a.a.), α1-α2 loop (~68-81 a.a.), α4-helix (~166-176 a.a.) and β6-β7 loop (~180-198 a.a.) exhibit strong deviation except in case of hNatD-H4_R3C, in which only β6-β7 loop showed deviation. Additionally, α2-helix (~81-90 a.a.) in hNatD-H4_G4S showed deviation addition to remaining four regions compared to others.

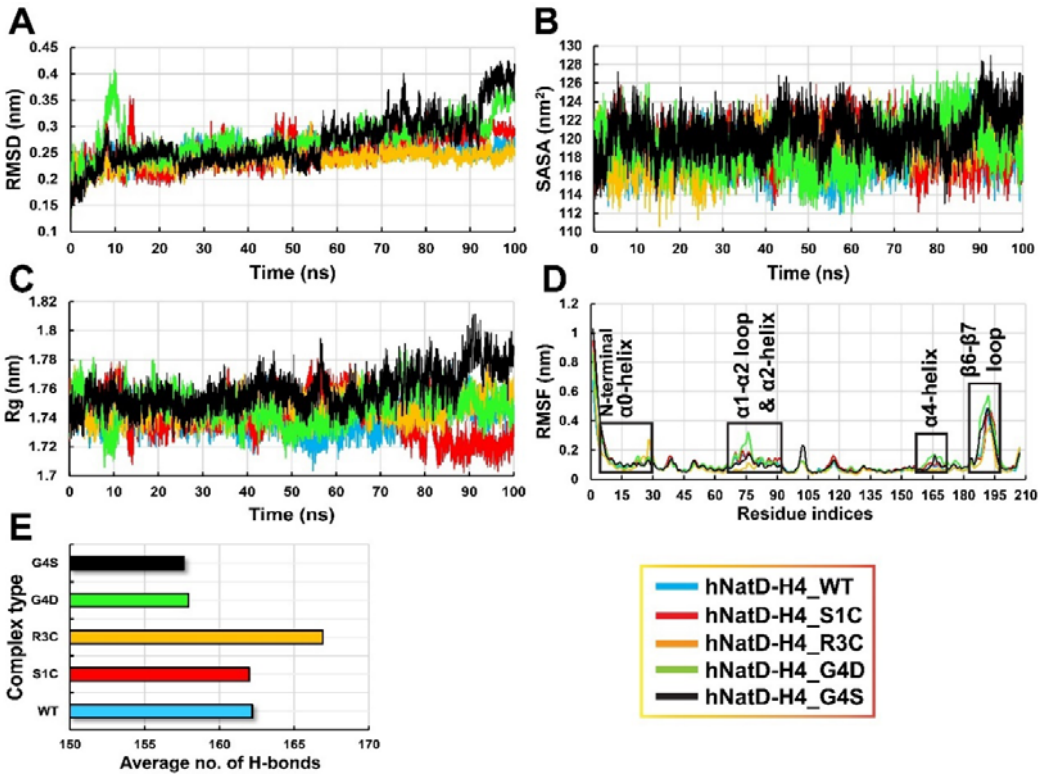

Figure 7. MD trajectory post analyses of five complexes (WT & MTs). (A) RMSDs (nm) plots of hNatD. (B) SASA (nm^2) plots of hNatD. (C) Rg (nm) plots of hNatD. (D) RMSF (nm) plots of hNatD. (E) Average intramolecular hydrogen bonds in hNatD.

From the RMSD plots in Figure 7(A), it can be noted that there is a peak (0.40 nm) at 10 ns and then RMSD was decreased to normal (0.25 nm) in hNatD-H4_G4D. Afterwards it showed gradual increment and reached to the ~0.35 nm at the end of simulation. However, hNatD in hNatD-H4_G4S showed four significant changes in RMSD during 0-25 ns, 26-55 ns, 56-90 ns and 91-100 ns times.

Moreover, Solvent accessible surface area (SASA) and Radius of gyration (Rg) were calculated to investigate the hNatD folding and compactness. The SASA and Rg give similar trend in the plots. Both indicate the opening or closing of hNatD during the simulation. The SASA is the total surface area which is in contact with solvent (water) whereas, Rg is the average distance of each atom from the centre of gravity in protein. The Rg is calculated using below equation;

$$\boldsymbol{R_g} = \left(\frac{|\boldsymbol{r_i}|^2\, \boldsymbol{m_i}}{\sum_i \boldsymbol{m_i}}\right)^{\frac{1}{2}} \tag{2}$$

Where, r_i and m_i are the position and mass of the atom i respectively in protein. The average SASA values of hNatD in hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C, hNatD-H4_G4D and hNatD-H4_G4S were found 118.3 nm^2, 119.2 nm^2, 118.9 nm^2, 119.2 nm^2 and 120.6 nm^2 respectively. The higher value of SASA or Rg indicates the reduction in protein compactness and lower value suggests gain in the compactness. It can be observed from the SASA (Figure 7(B)) and Rg (Figure 7(C)) plots that the folding and compactness of hNatD in hNatD-H4_WT and hNatD-H4_G4D are following similar trend. Initially, the SASA values gradually decline from around 121 nm^2 to 116 nm^2 until ~55 ns. After that both the values begin to rise and reach to approximately 124 nm^2 at ~80 ns and then again decline to 118 nm^2 at the end of simulation. In case of hNatD-H4_S1C, the value fluctuates between around 115 nm^2 and 124 nm^2 for the initial 70 ns time and, then it decreases to 118 nm^2 at the end of simulation. For the initial 70 ns, the overall SASA value of hNatD in hNatD-H4_R3C was remained below to the SASA value of hNatD in hNatD-H4_S1C. However, the SASA value of hNatD in hNatD-H4_R3C was observed slightly higher (~2 nm^2) as compared to SASA value of hNatD in

hNatD-H4_S1C for the last 30 ns. Finally, the SASA value of hNatD in hNatD-H4_G4S was significantly altered (~115 nm^2- 126 nm^2) over the course of 100 ns simulation which reveals large structural deviation of hNatD as compared to in other complexes. The average Rg values of hNatD were observed 1.74 nm, 1.74 nm, 1.75 nm, 1.75 and 1.76 in hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C, hNatD-H4_G4D and hNatD-H4_G4S respectively.

To determine the fluctuation of each residue in hNatD, Root mean square fluctuation (RMSF) of Cα atoms in residues was calculated for the all complexes. The RMSF value gives details about the interacting residues of hNatD with acetyl-CoA or histone H4 tail. RMSF is calculated by employing following equation;

$$\boldsymbol{RMSF} = \left(\frac{1}{T}\sum_{t=1}^{T}\left\langle\left(r'_i(t)\right) - \left(r_i\left(t_{ref}\right)\right)^2\right\rangle\right)^{\frac{1}{2}} \quad (3)$$

Where, T is the simulation time (time steps), r_i is the position of atom i at reference time, t_{ref} and, r' indicates the position of the selected atom at time t after aligning with the reference time frame. Figure 7(D) shows the RMSF plots of hNatD. Highly dynamic residue regions of hNatD are illustrated in Figure S3-S7 in hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C, hNatD-H4_G4D and hNatD-H4_G4S complexes respectively. It can be seen from the RMSF plots that two regions, β6-β7 loop (~180-198 a.a.) at the C-terminal and α1-α2 loop (~68-81 a.a.) showed the highest fluctuation compared to other dynamic regions. hNatD residues in hNatD-H4_G4D have the greater fluctuations whereas residues in hNatD-H4_R3C have the lowest degree of fluctuations. Five dynamic regions, N-terminal α0-helix (~1-20 a.a.), α1-α2 loop (~68-81 a.a.), α2-helix (~81-90 a.a.), α4-helix (~166-176 a.a.) and β6-β7 loop (~180-198 a.a.) in hNatD are part of the catalytic site in which histone H4 tai and acetyl-CoA are located.

Protein stability is attributed to the intramolecular hydrogen bonds network and this network provides stability to the protein folds (Hubbard & Kamran Haider, 2010). To investigate how WT and MT histone H4 tail bindings to the hNatD disrupt intramolecular hydrogen bond network in hNatD, we calculated total number of hydrogen bond formation in each complex during the 100 ns time. Figure 7(E) shows average number of hydrogen bonds in five complexes. Figure S8 represents the intramolecular hydrogen bonds variation during 100 ns simulations of studied five complexes. The average number of intramolecular hydrogen bonds were found 162, 162, 167, 158 and 158 in hNatD of hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C, hNatD-H4_G4D and hNatD-H4_G4S complexes respectively. Thus, hNatD in hNatD-H4_R3C has the greater stability and hNatDs in hNatD-H4_G4D and hNatD-H4_G4S have the least stability. However, hNatDs in hNatD-H4_WT and hNatD-H4_S1C have the stability between prior complexes. Hence, it can be concluded that R3C mutation in histone H4 tail enhances the protein stability through increasing the number of hydrogen bonds in hNatD whereas, G4D and G4S mutations significantly disturb the hNatD hydrogen bond network and destabilize hNatDs. Also, we obtained intermolecular hydrogen bonds between H4/acetyl-CoA and hNatD during entire simulation in five complexes (Figure S9). The average number of hydrogen bonds between H4 and hNatD was found 6, 6, 8, 6 and 6 in hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C, hNatD-H4_G4D and hNatD-H4_G4S complexes respectively (Figure S9(A-E)). Thus, R3C mutant H4 tail has greater stability inside the catalytic site in comparison with other WT and MT histones. And, the average number of hydrogen bonds between acetyl-CoA and hNatD was observed 9, 7, 8, 7 and 6 in hNatD-H4_WT, hNatD-H4_S1C, hNatD-H4_R3C, hNatD-H4_G4D and hNatD-H4_G4S complexes respectively (Figure S9(F-J)). Hence, the stability of acetyl-CoA in WT and S1C complexes is greater as compared to other complexes. Thus, this can be considered governing cause for the reduction in catalytic efficiency of hNatD in mutant complexes.

Furthermore, we generated Normalized linear mutual information (nLMI) correlation metrics to probe the inter-residues communication network in hNatD under the influence of histone H4 tail binding. The following equation is used to calculate the LMI between the residue i and j;

$$\boldsymbol{LMI_{ij}} = \frac{1}{2}\left[\ln\left(\boldsymbol{detC_i}\right) + \ln\left(\boldsymbol{detC_j}\right) - \ln\left(\boldsymbol{detC_{ij}}\right)\right] \quad (4)$$

$C_i = \langle x_i^T x_i \rangle$, $C_j = \langle x_j^T x_j \rangle$, and $C_{ij} = \langle (x_i, x_j)^T (x_i, x_j) \rangle$ where $x_i = R_i - \langle R_i \rangle$ and $x_j = R_j - \langle R_j \rangle$ in which R_i and R_j are the position vectors of atom i and j. For the calculation, residues which have LMI ≥ 0.3 and distance threshold ≤ 7Å were considered to be interacting with each other in the protein. In normalized LMI correlation, it gives 0 value for LMI in case of no correlation and 1 when atoms are completely correlated.

To understand the impacts of WT and MTs H4 tails on hNatD residue network during the simulation, we performed linear mutual correlation analysis on the MD trajectories. The nLMI plots are given in Figure S10. It is apparently clear from Figure S10 that among five complexes of histone H4 tail (WT & MTs) with hNatD, there is considerably lower correlation between residues in hNatD of hNatD-H4_R3C complex as compared to other complexes. As depicted in Figure S10 that five significantly fluctuating parts (α4-helix, β6-β7 loop, α1-α2 loop, α2-helix and N-terminal α0-helix) of hNatD in five complexes have remarkable differences in correlation.

Additionally, we calculated the correlation difference in hNatD of WT and MT complexes (Figure 8). It can be clearly seen from the Figure 8 that MTs have significant correlation difference. The large correlation difference was observed in hNatD-H4_R3C. In this complex, a large number of residues show slightly positive correlation and some residues are appeared to be having higher correlation (Figure 8(B)). Thus, residues communicate cooperatively in this complex. However, in other MTs (S1C, G4D and G4S) complexes, anticorrelation motions were observed (Figure 8(A,C-D)). This indicates that in these MTs, the residue network of the hNatD is disrupted significantly. Recently, Ho et al. have investigated the catalytic efficiency of hNatD for pentapeptide of H4 WT and MT tails (Ho & Huang, 2021). Their finding shows that the catalytic efficiency of N-α-acetylation of MT histone tails was remarkably reduced. The K_m values for WT, S1C, R3C, G4D and G4S were found 5.9 μM, ~500 μM, 17 μM, 272 μM and 14 μM respectively. The higher values of K_m for S1C and G4D may be due to the significant disruption of hNatD structure as it has anticorrelation residue networks in these two complexes. And, the lower values of K_m for another two mutants, R3C and G4S may be considered

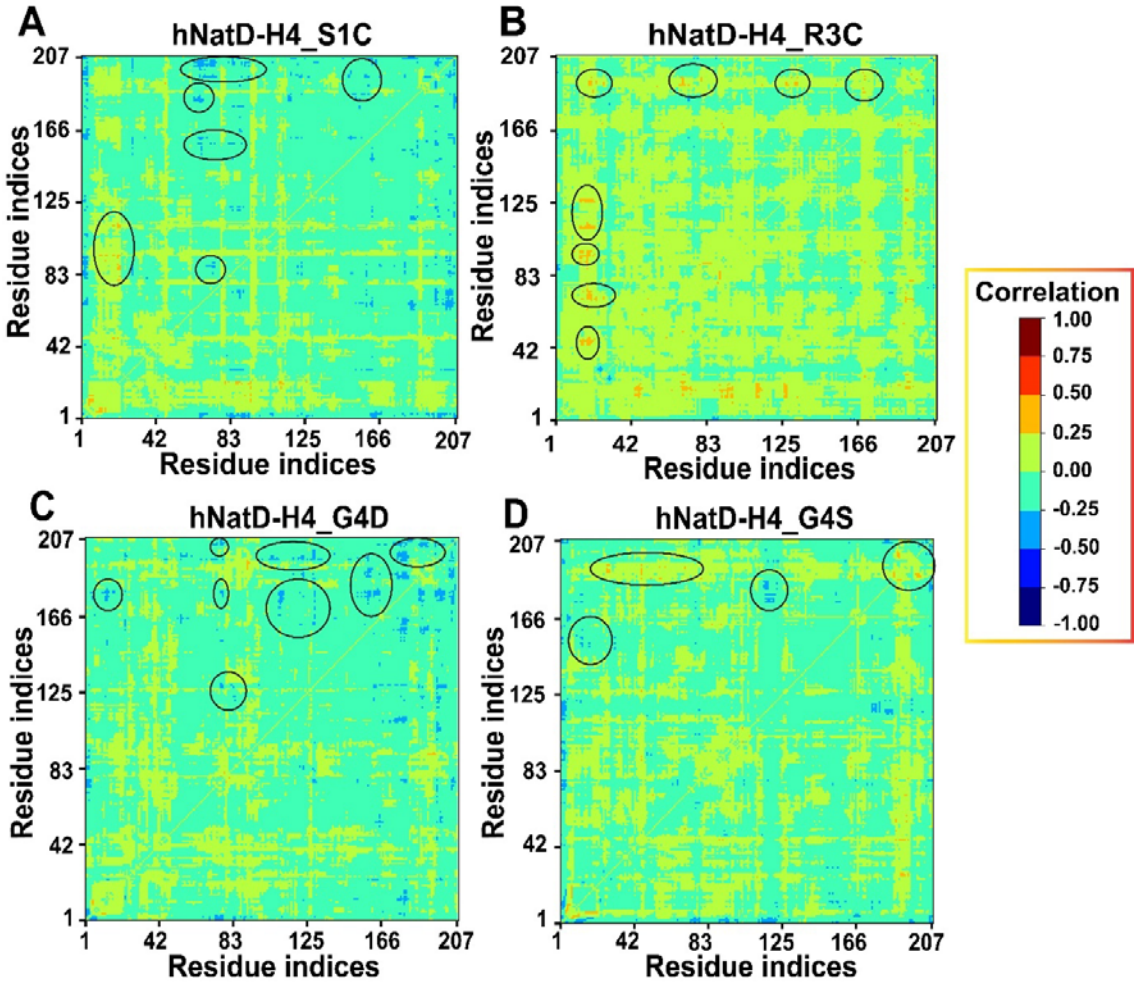

Figure 8. The nLMI correlation difference between WT and MTs complexes of hNatD and histone H4. (A) hNatD-H4_S1C (B) hNatD-H4_R3C (C) hNatD-H4_G4D (D) hNatD-H4_G4S. The degree of correlation corresponds to the color bar.

due to absence of anticorrelation motions. The WT, R3C and G4S histone tails do not alter the dynamic residue network of catalytic site markedly thus, their catalytic efficiencies were preserved largely as compared to S1C and G4D tails. Hence, our molecular level investigation supports the previously reported finding.

Furthermore, we calculated Betweenness centrality (BC) in complexes. The centrality analysis indicates the degree of communication between residues. The higher BC values reveal protein region where significant inter-domain communication would be present (D. K. Brown et al., 2017). From Table 1, considering all residues which are in close proximity inside the catalytic site, we determined position of every residue in various regions of hNatD in five complexes (Table S2). It was noted that regions which are part of catalytic site have substantial communication (Figure 9). Most notable regions such as β3, β4, β5 and β7 have been found to have significant connection in every complex whereas only some number of residues from the α0, α1, α2 and α3 are communicate with former regions. Apart from these, residues from the loops (α1-α2, α2-β2, α3-β5, α4-β5 and β6-β7) have notable communication with former regions. Also, there are only six residues (R112, F113, Q128, L160, I200 and S202) common among five complexes and other are different in each complex thus, mutations lead remarkable changes in the hNatD structure.

Table 1. The large Betweenness centrality (BC) containing residues of hNatD in studied five complexes with histone H4 tail (WT & MTs).

Complex	Correlated residues
hNatD-H4_WT	S43, D90, R92, A93, W94, Y95, H110, R112, F113, V115, L122, Y125, Q128, L129, L160, T161, V162, S185, S197, I200, S202
hNatD-H4_S1C	V16, D62, M71, T89, W94, F111, R112, F113, E120, Y123, Q128, A150, K156, V158, L160, F163, H165, F179, P186, I200, S202
hNatD-H4_R3C	V12, L26, S43, C46, S75, R112, F113, E120, V121, L122, Q128, A150, M155, K156, V158, L160, T161, F179, S197, I200, S202
hNatD-H4_G4D	D10, S43, S75, R92, W94, S109, F111, R112, F113, E120, L122, Y123, E126, Q128, V158, M159, L160, T161, S197, I200, S202
hNatD-H4_G4S	D10, S43, T64, S75, T89, W94, A98, R112, F113, E120, V121, L122, Y123, Q128, A150, M159, L160, T161, S197, I200, S202

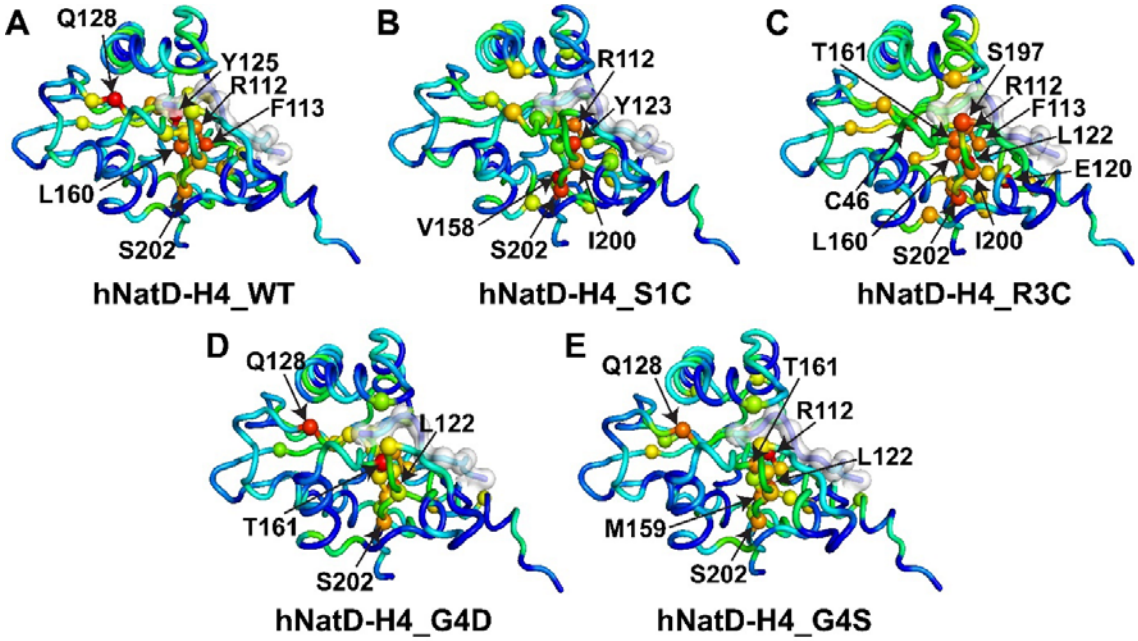

Figure 9. Betweenness centrality (BC) of hNatD in WT and MT complexes. (A) hNatD-H4_S1C (B) hNatD-H4_R3C (C) hNatD-H4_G4D (D) hNatD-H4_G4S. Spheres are the residues which have greater communication in hNatD. Red color has the highest degree of betweenness and blue has the least. Histone H4 tails are shown in cartoon-surface representation.

Furthermore, in addition to the beta strands and helices, loop regions are also communicating in hNatD. It was noted that there are a greater number of residues show communication from the loop regions in MTs complexes as compared to WT complex of hNatD. The β6-β7 loop communication was observed in all the complexes. In WT, another loop (α2-β2) is part of the communication. However, another type of loops such as α0-β0, α1-α2, α3-β5 and α4-β5 were observed to participate to the communication in MTs complexes. Among these five loops in WT and MTs, except α3-β5 loop every loop is part of acetyl-CoA and substrate recognition sites. Thus, it is clear that during the substrate recognition and acetylation, these beta strands, helices and loops dynamically communicate and work cooperatively. It can be seen from the Figure X that in hNatD-H4_WT and hNatD-H4_R3C, the residue contact network is maintained at the catalytic sites as compared to other MT complexes. Hence, it implies that catalytical core of the hNatD in these two complexes are not influenced greatly.

Interaction network between H4 tail and hNatD

To understand the interaction dynamics between the histone H4 tails (WT & MTs), we extracted the 10 frames from 10 ns to 100 ns at every 10 ns from the each complex and calculated the interactions and plotted dominant interactions. From the Figure 2(B), we know that the catalytic site of hNatD is highly acidic in nature compared to previous hNatA (Magin et al., 2015). Thus, dominant interactions come from acidic residues in all the complexes. Results indicate that as compared to hNatD-H4_WT, the total number of stable interactions in MTs were observed large in number. And, this is the cause of significant deflection of WT histone tail inside the catalytical core in hNatD (Figure 5(A)). In MTs, greater number of residues interact with the histone H4 tail and stabilize it inside the hNatD (Figure 5(B-E)).

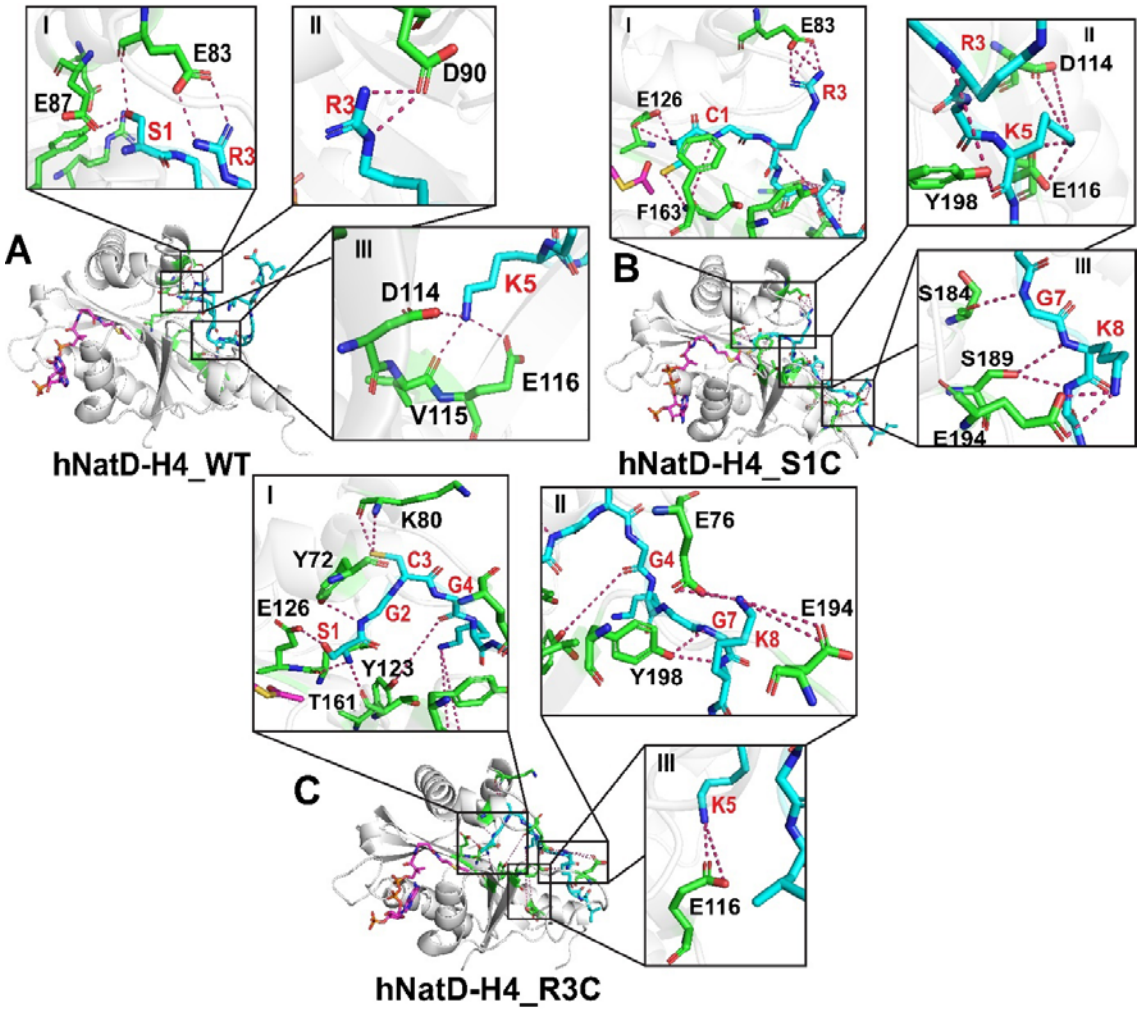

Figure 10. Interactions between histone H4 tails and hNatD. (A) hNatD-H4_WT (B) hNatD-H4_S1C (C) hNatD-H4_R3C. Histone H4 residues are colored in red while hNatD residues are in black.

In hNatD-H4_WT, histone H4 tail interacts with E83 and E87 from α2-helix, D90 from α2-β2 loop and, D114, V115 and E116 from the β3 strand through S1, R3 and K5 residues (Figure 10(A)). The hydrogen bonds and salt bridges are the dominant types of interactions between H4 tail and hNatD. E83 forms hydrogen bonds and salt bridges with S1 hydroxyl and R3 amino groups. Also, E87 interacts to S1 hydroxyl and R3 amino groups through forming hydrogen bonds and salt bridges with amino groups of S1 and R3 residues. The K5 sidechain amino group interacts with the D115, V115 and E116 hNatD residues by forming hydrogen bonds and salt bridges.

In MTs, new residues from the hNatD appear and increase the binding affinity of H4 tail with hNatD. Thus, the interaction network is largely altered in MTs in comparison with WT. The most important

residue, E126 was observed to be participating in interactions with all MT H4 tails and this might be cause of the stability of MT H4 tails at the binding site. In case of hNatD-H4_S1C, E116 became dominant and E194 from the β6-β7 loop was emerged to interact with H4 tail. In addition to that, two aromatic residues (F163 of β5 strand and Y198 of β7 strand) were also found to have interactions with H4 tail (Figure 10(B)). In this case, E83 only interacts to R3 by forming hydrogen bonds and salt bridges. And, E126 of β4 strand forms hydrogen bonds and salt bridges with the amino groups of C1 residue. The thiol group in C1 also forms hydrogen bonds with F163 residue. The V115 residue was disappeared and new residue, Y198 was observed to interact with the R3 and K5 through forming hydrogen bonds. Like E116, D114 also interacts with the K5 sidechain through hydrogen bonds and salt bridge. Furthermore, S184 and S189 from the β6-β7 loop bind to the G7 and K8 of the H4 tail by forming hydrogen bonds.

In hNatD-H4_R3C, E83 was vanished and new acidic residue E76 from the α1-α2 loop was found to involved in interactions (Figure 10(C)). E76 interacts to K8 residue of the H4 tail by forming hydrogen bond and salt bridge. In this mutation, basic residue, arginine is substituted with cysteine polar residue. Hence, K80 from α1-α2 loop was found to interact with the cysteine thiol group by forming hydrogen bonds. Along with E126, another residue (T161 from β5 strand) was observed to interact with S1 residue of the H4 tail. Y72 residue from the α1-helix was found to have interaction with G2 of histone tail. Y123 from the β4 strand another new residue that was observed to bind with G4 residue. E116 and E194 interact with K5 and K8 residues of the histone H4 respectively like in hNatD-H4_S1C complex. However, Y198 interacts with G7 and K8 only unlike in hNatD-H4_S1C complex where it interacts with R3 and K5.

At the fourth position in histone H4, there were two mutations studied (G4D and G4S). Figure 11 illustrates the interactions of G4S and G4D H4 with hNatD. It can be seen from the Figure 11 (A-B) that there are large number of hNatD residues interact with the G4S H4 as compared to G4D H4 tail. The G4D mutation causes significant alternation in the hNatD structure due to the acidic sidechain of the aspartic acid. The H4 binding site is highly acidic in nature (Figure 3(B)) thus, this carboxylic group repels surrounding residue through electrostatic interactions. Along with hNatD structural alternation, this residue also significantly impacts the position of acetyl-CoA inside the hNatD in comparison with other complexes (Figure 5(D)).

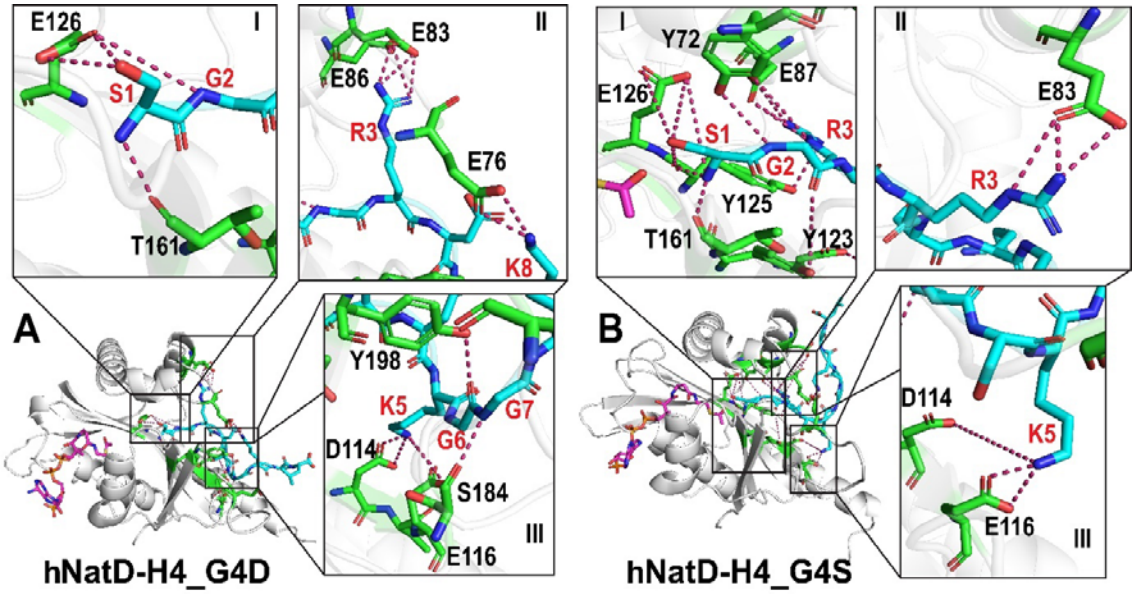

Figure 11. Interactions between histone H4 tails and hNatD. (A) hNatD-H4_G4D (B) hNatD-H4_G4S. Histone H4 residues are colored in red while hNatD residues are in black.

The commonly interacting residues are E83, D114, E116, E126 and T161. However, E83, D114 and E116 are the dominant residues in hNatD-H4_G4D and, E116 and E126 are in hNatD-H4_G4S. These residues form hydrogen bonds and salt bridges with the MT H4 tails. In case of hNatD-

H4_G4D, E126 interacts with the S1 and G2 through hydrogen bonds only. T161 forms hydrogen bond with the S1 residue. Two acidic residues (E83 and E86) interact to the sidechain of the R3 whereas another acidic residue, E76 interacts with the K8 through hydrogen bond and salt bridge. The K5 residue surrounds with two acidic residues, D114 and E116 and interacts by forming hydrogen bonds and salt bridges. S184 and Y198 interact with the G7 and G6 through only hydrogen bonds. However, there were two aromatic residues, Y72 and Y123 observed to have interactions with G2 and, Y125 with S1 and R3 by forming hydrogen bonds only in hNatD-H4_G4S complex. Like in hNatD-H4_G4D, D114 and E116 interact with the K5 residue of the H4 in hNatD-H4_G4S. The H-bond and salt bridge interactions between H4 tails and hNatD in five complexes are given in Table S3.

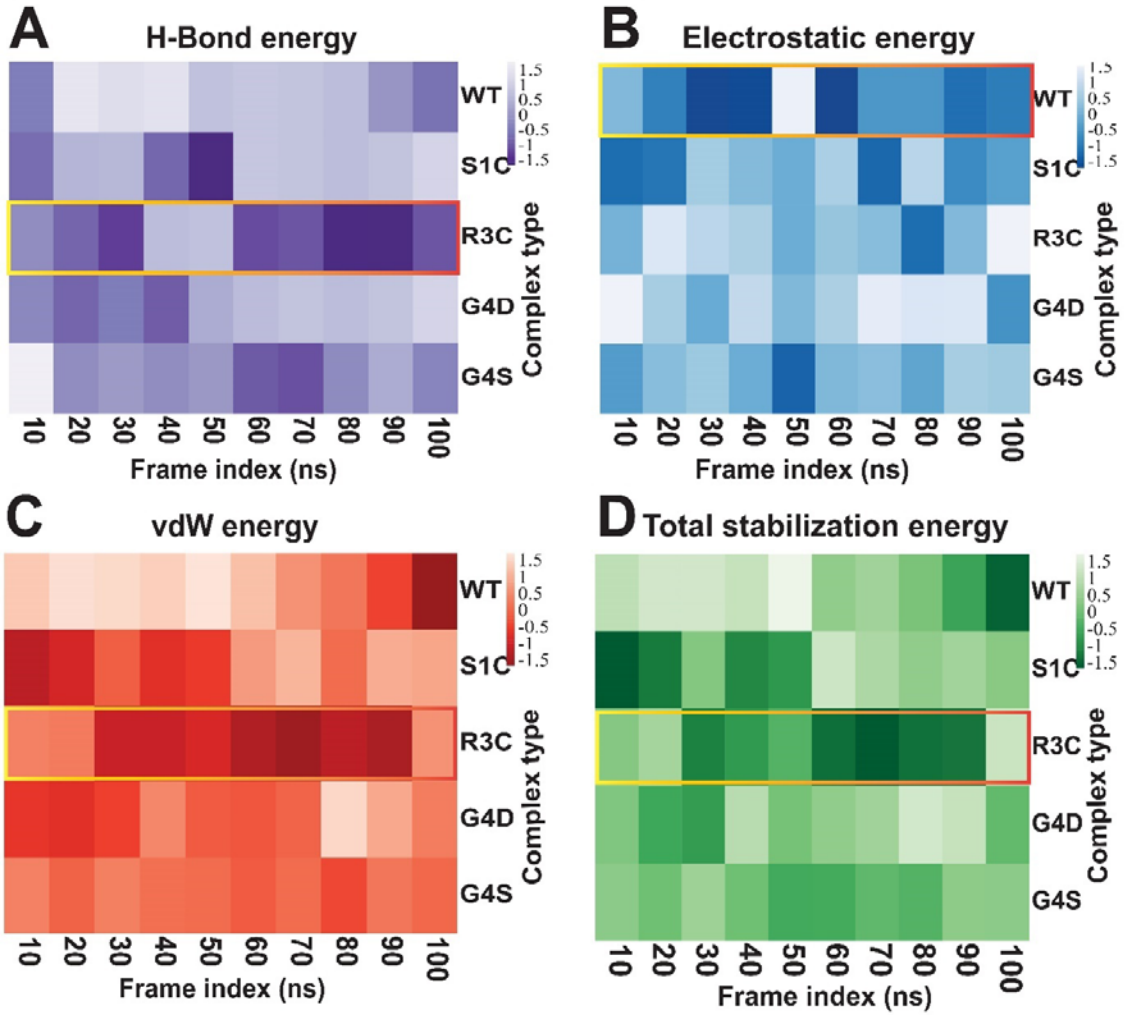

Figure 12. Heatmaps of energy contributions (kJ/mol) to the interactions calculated using PIMA web server. (A) Hydrogen bond energy (B) Electrostatic energy (C) van der Waals energy (D) Total stabilization energy. The yellow-red rectangular box shows the overall greater energy contribution to the complex.

Also, we calculated hydrogen bond, electrostatic and van der Waals energies of the WT and MTs complexes to investigate the binding energetics. These three energies were calculated for every 10 ns frame from the 10 ns to 100 ns in each complex (Figure 12). The total stabilizing energy of the complex was calculated using adding all three energies. These energies are plotted in the form of heatmaps and represented in Figure 12. The values of these energies are given in Table S4. It can be seen from the Figure 10 that hNatD-H4_R3C has the highest amount of stabilization energy. The maximum contribution to total stabilization energy comes from the van der Waals and less from the

hydrogen bond energy. The hydrogen bond, and van der Waals energies in MT complexes are large as compared to WT complex. However, electrostatic energy was observed higher in WT complex.

Conclusion

In this study, we examined the impacts of WT and MTs (S1C, R3C, G4D and G4S) histone H4 decapeptides on their bindings with hNatD by using 100 ns all-atom MD simulation. Our results support previous finding that the mutant H4 histones reduce the catalytic efficiency of hNatD. The MD post-trajectory analyses revealed that S1C, G4S and G4D mutants remarkably alter the residue network in hNatD. The intramolecular hydrogen bond analysis suggested that there is considerable number of losses of hydrogen bonds in hNatD of hNatD-H4_G4D and hNatD-H4_G4S complexes whereas large number of hydrogen bonds were increased in hNatD of hNatD-H4_R3C complex during the entire simulations. This implies that R3C mutant binding to hNatD brings stability in hNatD in comparison with WT and other MTs complexes. The Linear mutual information (LMI) and Betweenness centrality (BC) suggest that S1C, G4D and G4S significantly disrupt the catalytic site residue network as compared to R3C mutation in H4 histone. Thus, this might be the cause of notable reduction in catalytic efficiency of hNatD in these three mutant complexes. Further, interaction analysis supports that E126 is the important residue for the acetyltransferase mechanisms as it is dominantly found to have interactions with numerous residues of MTs histones in MD frames. Additionally, intermolecular hydrogen bond and RMSD analyses of acetyl-CoA predict the higher stability of acetyl-CoA inside the WT complex of hNatD and R3C complex. Also, we report here structural and dynamic aspects and Residue interactions network (RIN) of hNatD to target it to control the cell proliferation in the lung cancer conditions.

Acknowledgements

Shravan B. Rathod is thankful to his Chemistry Department for providing computational and infrastructure facilities. He is also thankful to Dr. Mustafa Tekpinar from the Pasteur Institute, France for explaining correlation plots and valuable suggestions on draft.

Conflict of interest

The authors declare no potential conflict of interest.

Author contributions

Shravan B. Rathod: Conceptualization, Investigation, Methodology, Data curation, Formal analysis, Writing-original draft, Writing-review & editing. **Kinshuk Raj Srivastava:** Conceptualization, Investigation, Supervision, Methodology, Data curation, Formal analysis, Writing-original draft, Writing-review & editing.

References

Aamir, M., Singh, V. K., Dubey, M. K., Meena, M., Kashyap, S. P., Katari, S. K., Upadhyay, R. S., Umamaheswari, A., & Singh, S. (2018). In silico prediction, characterization, molecular docking, and dynamic studies on fungal SDRs as novel targets for searching potential fungicides against fusarium wilt in tomato. *Frontiers in Pharmacology*, *9*, 1–28. https://doi.org/10.3389/fphar.2018.01038

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindah, E. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001

Aksnes, H., Drazic, A., Marie, M., & Arnesen, T. (2016). First Things First: Vital Protein Marks by N-Terminal Acetyltransferases. *Trends in Biochemical Sciences*, *41*(9), 746–760. https://doi.org/10.1016/j.tibs.2016.07.005

Allouche, A. (2012). Software News and Updates Gabedit — A Graphical User Interface for

Computational Chemistry Softwares. *Journal of Computational Chemistry*, *32*, 174–182. https://doi.org/10.1002/jcc

Arnesen, T., Van Damme, P., Polevoda, B., Helsens, K., Evjenth, R., Colaert, N., Varhaug, J. E., Vandekerckhove, J., Lillehaug, J. R., Sherman, F., & Gevaert, K. (2009). Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(20), 8157–8162. https://doi.org/10.1073/pnas.0901931106

Azam, S. S., & Abbasi, S. W. (2013). Molecular docking studies for the identification of novel melatoninergic inhibitors for acetylserotonin-O-methyltransferase using different docking routines. *Theoretical Biology and Medical Modelling*, *10*(1), 1–16. https://doi.org/10.1186/1742-4682-10-63

Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, *81*(8), 3684–3690. https://doi.org/10.1063/1.448118

Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., & MacKerell, A. D. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ1 and χ2 Dihedral Angles. *Journal of Chemical Theory and Computation*, *8*(9), 3257–3273. https://doi.org/10.1021/ct300400x

Biancotto, C., Frigè, G., & Minucci, S. (2010). Histone modification therapy of cancer. *Advances in Genetics*, v.*70*, 341–386. https://doi.org/10.1016/B978-0-12-380866-0.60013-7

Blanco, J. D., Radusky, L. G., Cianferoni, D., & Serrano, L. (2019). Protein-assisted RNA fragment docking (RnaX) for modeling RNA–protein interactions using ModelX. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(49), 24568–24573. https://doi.org/10.1073/pnas.1910999116

Brown, D. K., Penkler, D. L., Amamuddy, O. S., Ross, C., Atilgan, A. R., Atilgan, C., & Bishop, Ö. T. (2017). MD-TASK: A software suite for analyzing molecular dynamics trajectories. *Bioinformatics*, *33*(17), 2768–2771. https://doi.org/10.1093/bioinformatics/btx349

Brown, J. L., & Roberts, W. K. (1976). Evidence that approximately eighty per cent of the soluble proteins from Ehrlich ascites cells are Nα acetylated. *Journal of Biological Chemistry*, *251*(4), 1009–1014. https://doi.org/10.1016/s0021-9258(17)33793-6

Cavalieri, V. (2020). Histones, their variants and post-translational modifications in zebrafish development. *Frontiers in Cell and Developmental Biology*, *8*, 1–8. https://doi.org/10.3389/fcell.2020.00456

Celej, M. S., Montich, G. G., & Fidelio, G. D. (2003). Protein stability induced by ligand binding correlates with changes in protein flexibility. *Protein Science*, *12*(7), 1496–1506. https://doi.org/10.1110/ps.0240003

Chen, C., Huang, Y., Ji, X., & Xiao, Y. (2013). Efficiently finding the minimum free energy path from steepest descent path. *Journal of Chemical Physics*, *138*(16), 1–9. https://doi.org/10.1063/1.4799236

Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, *98*(12), 10089–10092. https://doi.org/10.1063/1.464397

de Ruyck, J., Brysbaert, G., Blossey, R., & Lensink, M. F. (2016). Molecular docking as a popular tool in drug design, an in silico travel. *Advances and Applications in Bioinformatics and Chemistry*, *9*(1), 1–11. https://doi.org/10.2147/AABC.S105289

Forte, G. M. A., Pool, M. R., & Stirling, C. J. (2011). N-terminal acetylation inhibits protein targeting to the endoplasmic reticulum. *PLoS Biology*, *9*(5). e1001073. https://doi.org/10.1371/journal.pbio.1001073

Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A Linear Constraint Solver for molecular simulations. *Journal of Computational Chemistry*, *18*(12), 1463–1472.

https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H

Ho, Y. H., & Huang, R. (2021). Effects of Oncohistone Mutations and PTM Crosstalk on the N-Terminal Acetylation Activities of NatD. *ACS Chemical Biology*, 1–21. https://doi.org/10.1021/acschembio.1c00840

Huang, S. Y., & Zou, X. (2011). Construction and test of ligand decoy sets using MDock: Community structure-activity resource benchmarks for binding mode prediction. *Journal of Chemical Information and Modeling*, *51*(9), 2107–2114. https://doi.org/10.1021/ci200080g

Hubbard, R. E., & Kamran Haider, M. (2010). Hydrogen Bonds in Proteins: Role and Strength. *eLS*. https://doi.org/10.1002/9780470015902.a0003011.pub2

Hwang, C. S., Shemorry, A., & Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. *Science*, *327*(5968), 973–977. https://doi.org/10.1126/science.1183147

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, *79*(2), 926–935. https://doi.org/10.1063/1.445869

Ju, J., Chen, A., Deng, Y., Liu, M., Wang, Y., Wang, Y., Nie, M., Wang, C., Ding, H., Yao, B., Gui, T., Li, X., Xu, Z., Ma, C., Song, Y., Kvansakul, M., Zen, K., Zhang, C. Y., Luo, C., … & Zhao, Q. (2017). NatD promotes lung cancer progression by preventing histone H4 serine phosphorylation to activate Slug expression. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-00988-5

Kumar, A., & Jernigan, R. L. (2021). Ligand Binding Introduces Significant Allosteric Shifts in the Locations of Protein Fluctuations. *Frontiers in Molecular Biosciences*, *8*, 1–8. https://doi.org/10.3389/fmolb.2021.733148

Lippert, R. A., Bowers, K. J., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., & Shaw, D. E. (2007). A common, avoidable source of error in molecular dynamics integrators. *Journal of Chemical Physics*, *126*(4), 2006–2008. https://doi.org/10.1063/1.2431176

Liszczak, G., Arnesen, T., & Marmorsteins, R. (2011). Structure of a ternary Naa50p (NAT5/SAN) N-terminal acetyltransferase complex reveals the molecular basis for substrate-specific acetylation. *Journal of Biological Chemistry*, *286*(42), 37002–37010. https://doi.org/10.1074/jbc.M111.282863

Liszczak, G., Goldberg, J. M., Foyn, H., Petersson, E. J., Arnesen, T., & Marmorstein, R. (2013). Molecular basis for N-terminal acetylation by the heterodimeric NatA complex. *Nature Structural and Molecular Biology*, *20*(9), 1098–1105. https://doi.org/10.1038/nsmb.2636

Liszczak, G., & Marmorstein, R. (2013). Implications for the evolution of eukaryotic amino-terminal acetyltransferase (NAT) enzymes from the structure of an archaeal ortholog. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(36), 14652–14657. https://doi.org/10.1073/pnas.1310365110

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., … & Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B*, *102*(18), 3586–3616. https://doi.org/10.1021/jp973084f

Magin, R. S., Liszczak, G. P., & Marmorstein, R. (2015). The molecular basis for Histone H4- and H2A-specific amino-terminal acetylation by NatD. *Structure*, *23*(2), 332–341. https://doi.org/10.1016/j.str.2014.10.025

Mathew, O., & Sowdhamini, R. (2016). PIMA: Protein-Protein Interactions in Macromolecular Assembly - a web server for its Analysis and Visualization. *Bioinformation*, *12*(1), 9–11. https://doi.org/10.6026/97320630012009

Mullen, J. R., Kayne, P. S., Moerschell, R. P., Tsunasawa, S., Gribskov, M., Colavito-Shepanski, M.,

Grunstein, M., Sherman, F., & Sternglanz, R. (1989). Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast. *EMBO Journal*, *8*(7), 2067–2075. https://doi.org/10.1002/j.1460-2075.1989.tb03615.x

Nacev, B. A., Feng, L., Bagert, J. D., Lemiesz, A. E., Gao, J. J., Soshnev, A. A., Kundra, R., Schultz, N., Muir, T. W., & Allis, C. D. (2019). The expanding landscape of 'oncohistone' mutations in human cancers. *Nature*, *567*(7749), 473–478. https://doi.org/10.1038/s41586-019-1038-1

Negre, C. F. A., Morzan, U. N., Hendrickson, H. P., Pal, R., Lisi, G. P., Patrick Loria, J., Rivalta, I., Ho, J., & Batista, V. S. (2018). Eigenvector centrality for characterization of protein allosteric pathways. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(52), E12201–E12208. https://doi.org/10.1073/pnas.1810452115

Panda, P. K., Arul, M. N., Patel, P., Verma, S. K., Luo, W., Rubahn, H. G., Mishra, Y. K., Suar, M., & Ahuja, R. (2020). Structure-based drug designing and immunoinformatics approach for SARS-CoV-2. *Science Advances*, *6*(28), 1–15. https://doi.org/10.1126/sciadv.abb8097

Parisien, M., Freed, K. F., & Sosnick, T. R. (2012). On docking, scoring and assessing protein-DNA complexes in a rigid-body framework. *PLoS ONE*, *7*(2), e32647. https://doi.org/10.1371/journal.pone.0032647

Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, *52*(12), 7182–7190. https://doi.org/10.1063/1.328693

Penkler, D. L., Atilgan, C., & Tastan Bishop, Ö. (2018). Allosteric Modulation of Human Hsp90α Conformational Dynamics. *Journal of Chemical Information and Modeling*, *58*(2), 383–404. https://doi.org/10.1021/acs.jcim.7b00630

Rathod, S. B., Prajapati, P. B., Punjabi, L. B., Prajapati, K. N., Chauhan, N., & Mansuri, M. F. (2020). Peptide modelling and screening against human ACE2 and spike glycoprotein RBD of SARS-CoV-2. *In Silico Pharmacology*, *8*(1), 1–9. https://doi.org/10.1007/s40203-020-00055-w

Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Research*, *31*(13), 3381–3385. https://doi.org/10.1093/nar/gkg520

Sethi, A., Eargle, J., Black, A. A., & Luthey-Schulten, Z. (2009). Dynamical networks in tRNA: Protein complexes. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(16), 6620–6625. https://doi.org/10.1073/pnas.0810961106

Settanni, G., Serquera, D., Marszalek, P. E., Paci, E., & Itzhaki, L. S. (2013). Effects of Ligand Binding on the Mechanical Properties of Ankyrin Repeat Protein Gankyrin. *PLoS Computational Biology*, *9*(1), e1002864. https://doi.org/10.1371/journal.pcbi.1002864

Starheim, K. K., Gevaert, K., & Arnesen, T. (2012). Protein N-terminal acetyltransferases: When the start matters. *Trends in Biochemical Sciences*, *37*(4), 152–161. https://doi.org/10.1016/j.tibs.2012.02.003

Tekpinar, M., Neron, B., & Delarue, M. (2021). Extracting Dynamical Correlations and Identifying Key Residues for Allosteric Communication in Proteins by correlationplus. *Journal of Chemical Information and Modeling*, *61*(10), 4832–4838. https://doi.org/10.1021/acs.jcim.1c00742

Tran, J. C., Zamdborg, L., Ahlf, D. R., Lee, J. E., Catherman, A. D., Durbin, K. R., Tipton, J. D., Vellaichamy, A., Kellie, J. F., Li, M., Wu, C., Sweet, S. M. M., Early, B. P., Siuti, N., Leduc, R. D., Compton, P. D., Thomas, P. M., & Kelleher, N. L. (2011). Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature*, *480*(7376), 254–258. https://doi.org/10.1038/nature10575

Van Wart, A. T., Durrant, J., Votapka, L., & Amaro, R. E. (2014). Weighted implementation of suboptimal paths (WISP): An optimized algorithm and tool for dynamical network analysis. *Journal of Chemical Theory and Computation*, *10*(2), 511–517. https://doi.org/10.1021/ct4008603

Wan, Y. C. E., Liu, J., & Chan, K. M. (2018). Histone H3 Mutations in Cancer. *Current Pharmacology*

Reports, *4*(4), 292–300. https://doi.org/10.1007/s40495-018-0141-6

Weiner, P. K., Langridge, R., Blaney, J. M., Schaefer, R., & Kollman, P. A. (1982). Electrostatic potential molecular surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(12), 3754–3758. https://doi.org/10.1073/pnas.79.12.3754

Yan, Y., Zhang, D., & Huang, S. Y. (2017). Efficient conformational ensemble generation of protein-bound peptides. *Journal of Cheminformatics*, *9*(1), 1–13. https://doi.org/10.1186/s13321-017-0246-7

Yi, C. H., Pan, H., Seebacher, J., Jang, I. H., Hyberts, S. G., Heffron, G. J., Vander Heiden, M. G., Yang, R., Li, F., Locasale, J. W., Sharfi, H., Zhai, B., Rodriguez-Mias, R., Luithardt, H., Cantley, L. C., Daley, G. Q., Asara, J. M., Gygi, S. P., Wagner, G., … & Yuan, J. (2011). Metabolic regulation of protein N-alpha-acetylation by Bcl-xL promotes cell survival. *Cell*, *146*(4), 607–620. https://doi.org/10.1016/j.cell.2011.06.050

Yu, W., He, X., Vanommeslaeghe, K., & MacKerell, A. D. (2012). Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *Journal of Computational Chemistry*, *33*(31), 2451–2468. https://doi.org/10.1002/jcc.23067

Yusufova, N., Kloetgen, A., Teater, M., Osunsade, A., Camarillo, J. M., Chin, C. R., Doane, A. S., Venters, B. J., Portillo-Ledesma, S., Conway, J., Phillip, J. M., Elemento, O., Scott, D. W., Béguelin, W., Licht, J. D., Kelleher, N. L., Staudt, L. M., Skoultchi, A. I., Keogh, M. C., … & Melnick, A. M. (2021). Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture. *Nature*, *589*(7841), 299–305. https://doi.org/10.1038/s41586-020-3017-y

Zhou, P., Jin, B., Li, H., & Huang, S. Y. (2018). HPEPDOCK: A web server for blind peptide-protein docking based on a hierarchical algorithm. *Nucleic Acids Research*, *46*(W1), W443–W450. https://doi.org/10.1093/nar/gky357