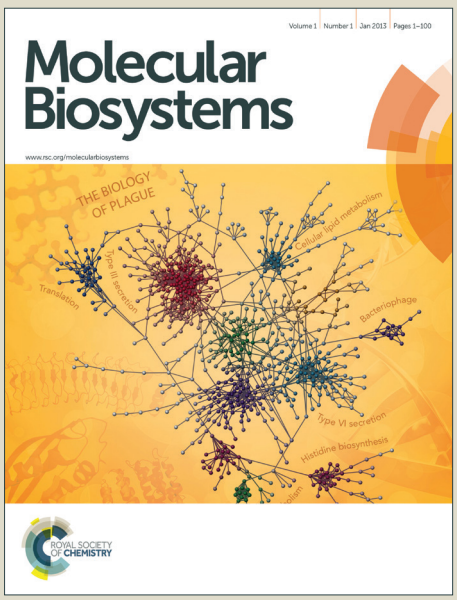
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

Conformational analysis of mouse Nalp3 domain structures by molecular dynamics simulation, and binding sites analysis

Bikash R Sahoo[1, 2*], Jitendra Maharana[1, 3], Gopal K Bhoi[1], Santosh K Lenka[1], Mahesh C Patra[1], Manas R Dikhit[4], Praveen K Dubey[5], Sukanta K Pradhan[1], Bijay K Behera[3]

[1]Department of Bioinformatics, Centre for Post-Graduate Studies, Orissa University of Agriculture and Technology, Bhubaneswar-751003, Odisha, India

[2]Laboratory of Molecular Biophysics, Institute of Protein Research, Osaka University, Osaka Prefecture 5650871, Japan

[3]Biotechnology Laboratory, Central Inland Fisheries Research Institute, Barrckpore-700120, Kolkata, West Bengal

[4]Rajendra Memorial Research Institute of Medical Sciences, ICMR, Agam Kuan, Patna-800007, Bihar, India

[5]Immunology Frontier Research Centre, Osaka University, Osaka Prefecture 5650871, Japan

*Corresponding author:

Bikash Ranjan Sahoo,

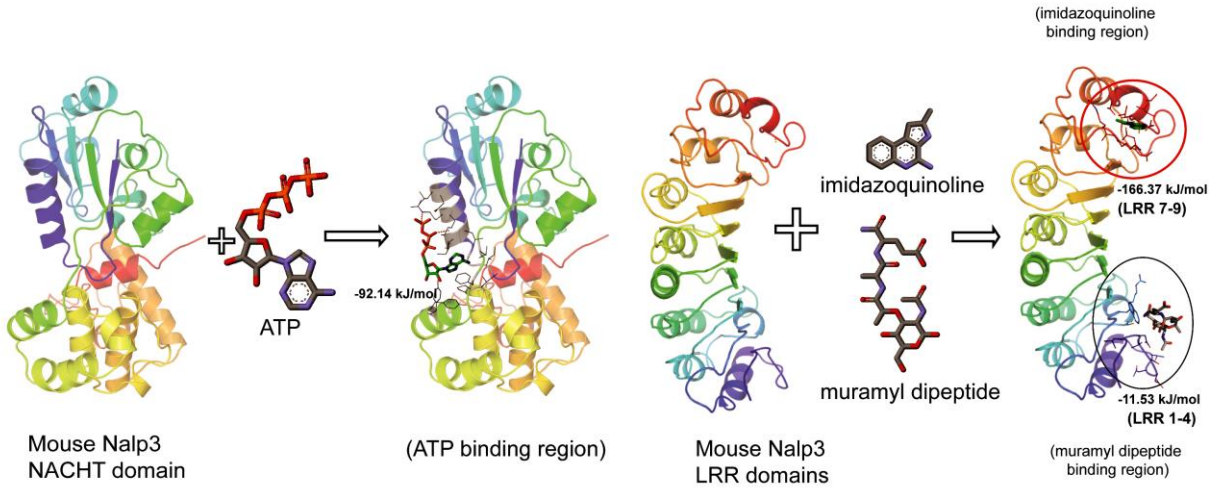
Department of Bioinformatics, Centre for Post-Graduate Studies,

Orissa University of Agriculture and Technology, Bhubaneswar-751003,

Odisha, India.

Email: bikash.bioinformatics@gmail.com
bikash.bioinformatic@protein.osaka-u.ac.jp
Phone: +91-674-2565760, +81-80-4761-7742

Graphical abstract



Highlight of the work:

- Binding site analysis of adenosine triphosphate, muramyl dipeptide and imidazoquinoline with mouse Nalp3 domains and free energy calculation.

Abstract:

Scrutinizing various nucleotide-binding oligomerization domain (NOD)-like receptors (NLR) genes in higher eukaryotes is very important for understanding the intriguing mechanism of host defense against pathogens. The nucleotide-binding domain (NACHT), leucine-rich repeat (LRR) and pyrin domain (PYD) containing protein 3 (Nalp3) is an intracellular innate immune receptor, and is associated with several immune system related disorders. Despite of Nalp3's protective role during pathogenic invasion, the molecular feature, and structural organization of this crucial protein is poorly understood. Using comparative modeling and molecular dynamics simulations, we have studied the structural architecture of Nalp3 domains, and characterized the dynamic and energetic parameters of adenosine triphosphate (ATP) binding in NACHT and pathogen derived ligands muramyl dipeptide (MDP) and imidazoquinoline with LRR domains. The results anticipated walker A, B and extended walker B motifs as the key ATP binding regions in NACHT that mediates self-oligomerization. Analysis of binding sites of MDP and imidazoquinoline revealed LRR7-9 being the most energetically favored site of imidazoquinoline interaction. However, binding free energy calculations using Molecular Mechanics/Possion-Boltzman Surface Area (MM/PBSA) method advocated that MDP is incompatible for activating Nalp3 molecule in monomeric form and suggest its complex nature with NOD2 or other NLRs for MDP recognition. The high affinity binding of ATP with NACHT is correlated to the experimental data for human NLRs. Our binding site prediction for imidazoquinoline in LRR warrants further investigation via in vivo models. This is the first study that provides ligand(s) recognition in mouse Nalp3 and its spatial structural arrangements.

Keywords: nucleotide-binding and oligomerization domain; Nalp3; muramyl dipeptide; imidazoquinoline; Molecular Mechanics/Possion-Boltzman Surface Area

1. Introduction:

The recognition of pathogen/damage-associated molecular patterns (PAMPs/DAMPs) through pattern recognition receptors (PRRs) has been studied in a number of higher and lower eukaryotes.[1] Upon interaction with PAMPs, PRRs trigger activation of innate immune genes, and protect the host organisms from various kinds of infections and inflammations.[2] However, the interaction between host innate immune system and microorganisms is very intricate to analyze. Among the different types of PRRs, the newly discovered nucleotide-binding oligomerization domain-containing protein (NOD) like receptors (NLRs) with leucine rich repeats (LRR) plays an important role in sensing intracellular PAMPs or DAMPs.[3, 4] NLRs are mainly comprised of 3 domains i.e. a variable N-terminal domain, a central NACHT (nucleotide-binding domain) domain and a C-terminal LRR domain. These are further divided into different subfamilies,[5] of which the NACHT, LRR and PYD (pyrin domain) domains-containing protein 3 (Nalp3) or cryopyrin protein belongs to Nalp subfamily that is characterized by N-terminal PYD, central NACHT, and C-terminal LRR domain.[6, 7] Nalp3 is normally present in the cytoplasm primarily in an inactive form and becomes active when the LRR domain is engaged with an agonist. The phenomenon is thought to be attributed to the conformational rearrangement of Nalp3 molecule, which exposes the oligomerization domain and subsequently the effector domain (PYD).[8, 9] Nalp3 plays an important role in inflammation, and is considered as a proximal sensor of cellular stress and danger signals[10] that forms a caspase-1 activating molecular complex and allows the activation of interleukin (IL)-1β. The involvement of Nalp3 with inflammatory diseases, specifically chronic infantile neurological cutaneous and articular syndrome has been studied recently.[11] However, its role in these diseases is not clearly understood yet. Nalp3 is able to respond to a variety of signals including adenosine triphosphate (ATP), nigericin, maitotoxin,

Staphylococcus aureus and *Listeria monocytogenes*,[12] ribonucleic acid (RNA)[13] and uric acid crystals (monosodium urate and calcium pyrophosphate dehydrate) released from dying cells.[14] Muramyl dipeptide (MDP), a key activating ligand of NOD2, is also reported as an activator of the Nalp3 inflammasome.[15] However, the consideration of MDP as Nalp3 ligand has been challenged in recent years, as IL-1β is produced when macrophages are stimulated with bacterial ligands followed by ATP.[16]

Previous studies showed no significant difference in IL-1β production between Nalp3 and wild-type mice peritoneal macrophages stimulated with MDP or lipids like lipopolysaccharide (LPS).[9,11] So, the exact role of MDP in IL-1β secretion is still elusive. In this study, we investigated the possible role of MDP in Nalp3 activation, and structural and functional characteristics of Nalp3 using computer algorithms. To date there is no experimental report available on the structure and PAMPs/DAMPs interacting mechanism of Nalp3. In order to understand the structural architecture and molecular interaction of Nalp3 domains leading to activation of the Nalp3 signal transduction, we modeled three dimensional (3D) structures of NACHT and LRR domains, and elucidated their interaction with ligands using molecular docking and long-range molecular dynamics (MD) simulations. We studied the intermolecular interactions of ATP with NACHT, MDP and imidazoquinoline with LRR domain. The protein-ligand complexes were examined for their bonding patterns, conformational variability, and binding free energies. For the first time, we elucidated the structural arrangements of NACHT and LRR domains in mouse Nalp3 and studied their biological functions which could be useful for therapeutic applications on Nalp3 related disorders.

2. Computational methods

2.1 Initial model preparation of NACHT and LRR domains

The amino acid sequence of mouse Nalp3 was retrieved from UniProtKB (http://www.uniprot.org/) database (UniProt ID: Q8R4B8). The NACHT and LRR domains of mouse Nalp3 were aligned with those of other homologous organisms using ClustalW program,[17] and the alignments were exported to ESPript 3.0 server[18] for graphical presentation. The domain alignments identified the biologically important regions with reference to the reported evidences. The primary sequence of NACHT and LRR domains were scanned against various web servers such as Genesilico,[19] 3D-Jury,[20] I-Tasser[21] and LOMETS[22] to identify the best templates for model building. Multi-template modeling followed by loop refinement was executed using Modeller 9.12[23] program based on the highly homologous templates. Predicted secondary structures were utilized for determining accuracy of the modeled proteins. For each individual domains of mouse Nalp3 protein, 100 models were constructed. The models having lowest discrete optimized potential energy (DOPE) score were considered for further studies, and were checked for stereo chemical accuracy using SAVES (http://nihserver.mbi.ucla.edu/SAVES/), MolProbity,[24] and WHAT IF[25] web servers. The models were further optimized by conducting energy minimization in Swiss PDB Viewer[26] followed by structural refinement at 3Drefine.[27] The refined models were cross checked in the above servers and were prepared for molecular dynamics (MD) simulation.

2.2 Molecular dynamics (MD) simulation

The simulation systems were prepared by embedding the refined protein models in a cubic water box with 23279 (for NACHT model) and 9148 (for LRR model) water molecules and 0.15 M NaCl concentration. A minimum distance of 12 Å was kept between surface of the protein and the simulation box. The simulations were carried out in GROMACS 4.5.5 [28, 29] with OPLS-AA/L all-atom force field [30, 31] for proteins and TIP4P for water. A steepest descent

algorithm using a tolerance of 1000 kJ mol^{-1} nm^{-1} and step size of 0.01 nm was used to minimize the systems. The minimized systems were equilibrated for 1 nanosecond (ns) under the NPT ensemble (temperature: 300 K and pressure: 1 atm) conditions, where the backbone atoms of the proteins were harmonically retrained throughout the equilibration. A production run of 50 ns was carried out for both NACHT and LRR systems. Built-in modules of Gromacs and VMD 1.9.1[32] programs were utilized to analyze the MD trajectories and the quality of the simulations. All graphs were generated using Grace-5.1.23 (http://plasma-gate.weizmann.ac.il/Grace/). The final snapshot obtained at the end of simulation was considered to represent the structures of NACHT and LRR proteins. The stereo chemical quality/parameters of final optimized models were verified using SAVES, ProSA,[33] ProQ,[34] and MolProbity[24] web resources. Vadar[35] and GeNMR[36] web servers were used to investigate the standard deviations, packing effects, bumps in the proposed models. VMD was used for time-dependent secondary structure analyses of the models. Structure visualizations were done using PyMOL (academic license) (http://www.pymol.org/), Discovery Studio Visualizer v3.5 (Accelrys) and VMD. All computations were done in a corei5 processor of 3.10 GHz (ASUS- ET2701) well equipped with CentOS 6.3 (http://www.centos.org/).

2.3 Principal component analysis (PCA)

To identify the most prominent structural motions during the MD simulation in NACHT and LRR models, PCA was performed using *g_covar* and *g_anaeig* programs. The backbone atoms (N-Cα-Cβ) were considered for this analysis. Among the generated eigenvectors (ev), ev1 showed the dominant model of motion and was considered as the principal component in our simulations. The ev with 100 frames obtained from the *g_anaeig* program was visualized in PyMOL by the help of porcupine plots depicting a graphical view of the motion along the

trajectories. Each Cα atom in a porcupine plot has a cone pointing in the direction of the smoothened motion of individual subdomains along the trajectories. Each Cα atom in a porcupine plot has an arrow pointing in the direction of the motion of the atom, the length of the stem of the arrow reflects the maximum displacement of the motion. The obtained porcupine plots of NACHT and LRR domains were graphically presented in PyMOL.

2.4 Molecular docking analysis

The 2D structures of ATP (CID: 5957), MDP (CID: 451714) and imidazoquinoline (CID: 5351568) were obtained from the PubChem database (http://pubchem.ncbi.nlm.nih.gov/), and their 3D structures were built using PRODRG2 server [37] subjecting to full chirality and charges followed by energy minimization. Molecular docking was carried out using AutoDock 4.2[38] following previously described methods [39-41]. To investigate the possible binding sites of ATP in NACHT we adopted two strategies. First we generated a complete grid embedding the whole NACHT model and set both ATP and protein as flexible to identify the best binding pose and lowest interaction energy. Second, we generated grid around the predicted ATP binding sites as suggested by UniProt (222-229 aa), and previously reported evidences [42, 43]. In LRR model, multiple grids were generated for both MDP and imidazoquinoline spanning different number of LRR repeats including a grid with all 9 LRRs. The best docking poses with high binding affinity (lowest binding energy) and more numbers of H-bonds were considered for MD simulation studies to observe molecular interaction at flexible and dynamic conditions.

2.5 MD simulation of complex

A set of 7 ligand bound complexes of NACHT and LRR domains of mouse Nalp3 protein were selected from the molecular docking results based on binding energy score and number of H-bond interactions. These complex structures were subjected to MD simulations using Gromacs.

The topologies of the ligand structures were prepared and protonated with PRODRG2 server. The procedure and parameters of MD simulations were same as described for free proteins in section 2.2. Position restraints were applied to protein and ligands in the complexes during equilibration phases of 0.2 and 0.5 ns under constant volume (NVT) and constant pressure (NPT) condition, respectively. The temperature was maintained at 300 K via Berendsen weak coupling method in both NVT and NPT condition. Additionally, the pressure was maintained at 1 bar by Parrinello-Rahman barostat method in NPT condition. Upon completion of temperature and pressure equilibration phases, final production MD simulations were conducted for 10 ns for each of 7 complexes.

2.6 Binding free energy calculations from the MD complexes

In recent years the binding free energy following the MM/PBSA approach has been widely used that combines internal energy, solvation energy based on electrostatic and nonpolar contributions, and the entropy.[44] In this study we calculated the binding free energy for the 7 different complexes considering their snapshots collected from MD simulations. The GMXABPS tool[45] was used to calculate the binding free energy employing MM/PBSA approach using Gromacs and APBS. For each complex, 1000 snapshots were extracted from MD trajectory using trajectory, topology and index files generated from each of the 7 MD simulations of NACHT and LRR domains of mouse Nalp3. Binding free energy calculation was carried out as described below.

$$\Delta G_{bind} = < G_{Protein\text{-}ligand\ complex} - G_{protein} - G_{ligand} > \quad (1)$$

Where the $G_{complex}$, $G_{protein}$ and G_{ligand} are the free energies of the complex, protein and ligand respectively. The brackets indicate that the binding free energy is calculated according to the

single trajectory method (STM).[46] The free energy terms used in equation-1 was described in details by Spiliotopoulos et al. [45] in the GMXAPBS tool.

2.7 In-silico site directed mutagenesis

To fortify the accuracy of our binding site predictions for ATP in NACHT, MDP and imidazoquinoline in LRR domain, we mutated the important interacting amino acid residues to alanine, proline and cysteine followed by re-docking using AutoDock 4.2.[38] Same grid and docking parameters were used for the docking analysis, and the effect of mutagenesis on binding affinity was analyzed.

3. Results and Discussion

3.1 Sequence analysis and structure modeling of NACHT and LRR domains in mouse Nalp3

The mouse Nalp3 protein is comprised of 3 domains among which the NACHT (216-532) and LRR (739-988) domains plays an important role by interacting with ATP and PAMPs/DAMPs, respectively. Sequence alignment of NACHT and LRR domains with Nalp3, NOD2 and NOD1 sequences of mouse and human was presented in Fig. 1. The potential ATP binding residues were well conserved in NACHT domain with conserved walker-A (common ATP binding p-loop in ATP/GTP binding proteins) "**G** *A/E/D* **AG** *I/S/V* **GK** *T/S*" and walker-B motifs (Fig. 1a).[47-49] The distance matrix analysis showed that mouse Nalp3-NACHT shared 92 and ~ 43 % similarities with human Nalp3-NACHT and NOD-NACHT respectively. The sequence alignment of LRR domain in mouse Nalp3 indicated remarkable homology of LRR regions across the Nalp3 and NOD groups in mouse and human (Fig. 1b). Mouse Nalp3-LRR shared 93 % and ~ 45 % similarity with human Nalp3-LRR and NOD-LRR respectively. However, the sequence identities was very low (~28%) between Nalp3 and NOD-NACHT/LRR domains. The leucine residues were observed to be well conserved in Nalp3 and NOD sequences.

The 3D models of NACHT domain were constructed using the crystal structures of NLRC4 (PDB ID: 4KXF); apoptotic protease-activating factor 1 bound to ADP (PDB ID: 1Z6T), and CED-4/CED-9 complex (PDB ID: 2A5Y), as these templates possessed top scores in different threading web servers. The models of LRR domain were built based on the crystal structure of mouse ribonuclease inhibitor (PDB ID: 3TSR) and porcine ribonuclease inhibitor (PDB ID: 2BNH). To ensure the accuracy of the template selection procedure, we compared the secondary structures of templates with the predicted secondary structures of NACHT and LRR domains using PSI-PRED[50] (Fig. S1, ESI). The results revealed good secondary structure conservation across the sequence length. The structural artifacts generated during modeling procedures were corrected using energy minimization and subsequent refinement with 3Drefine program. The validation of the refined models for accuracy of stereo chemical parameters using various structure validation web servers revealed good scores.

3.2 Stability of simulation systems of NACHT and LRR domains

MD simulations were performed for NACHT and LRR models to ensure stability of each model over a simulation time period of 50 ns. Soon after first 10 ns, both the NACHT and LRR models achieved a stable conformation throughout the simulation with an average backbone root mean square deviation (RMSD) of ~6.9 and ~3.35 Å, respectively (Fig. 2a). The radius of gyration analysis showed the models maintain a compact shape and size with gyration radii (Rg) of ~20 nm (Å) and ~21 nm (Å) for NACHT and LRR models, respectively (Fig. 2b). To investigate the fluctuations of individual residues in both the domains, root mean square fluctuations (RMSF) of Cα atoms were calculated during the 50 ns MD simulations. In NACHT system, the N-terminal region (~250-290 aa) and C-terminal regions (~455-500 aa) showed maximum fluctuations (up to 4 Å). In LRR, RMSF analysis showed that the N-terminal LRR1-2,

LRR6 and C-terminal LRR8-9 regions exhibited higher order of flexibility (up to 4 Å) (Fig. 2c and d). The flexible region between walker A/B motifs in NACHT indicated its involvement in nucleotide binding.[48] The most flexible loops at N and C-terminal of LRR domain suggested its contribution in facilitating the rearrangement of the domains during the Nalp3 activation or PAMPs/DAMPs interaction [39, 41, 51]. The secondary structure analysis from the MD trajectory showed that the first two α-helices of NACHT attained a Pi (symbol)-helix conformation close to N-termini. The β-sheets and other α-helical regions were largely intact throughout the simulations (Fig. S2a, ESI). All the 9 β-sheets of the LRR model retained their secondary structure during the MD simulations. The 5^{th} and 6^{th} α-helices showed structural changes from α-helix to 3_{10}helix conformation that remained in a well equilibrated state over the simulation time (Fig. S2b, ESI). Altogether, no significant changes were noticed in the secondary structural elements of the build models during the 50 ns MD simulations.

3.3 3D model evaluation

The final snapshots of the simulated models were validated with several structure validation programs. Primarily, analysis of backbone dihedral angles using Ramachandran plot showed ~ > 99.5 % of residues were in allowed regions for both NACHT and LRR models (Table 1) (Fig. S3, ESI). Both models exhibited good agreement to their primary sequences as revealed by the Verify-3D scores. ERRAT program indicated that all the non-bonded atoms in the constructed models were accurately predicted (Table 1). Analysis of NACHT and LRR models in ProSA and ProQ revealed that the Z-scores of the models were within the accepted range (Table 1) (Fig. S3, ESI). Analysis of MolProbity server suggested that the backbone bond angles and bond lengths of our proposed models were highly accurate. Further validation reports of NACHT and LRR models by Vadar and GeNMR are presented in Table 1. All these

validation scores ascertained that the constructed models using standard homology protocol is reasonably good to carry out further interaction study though MD simulations. The overall structure of NACHT domain can be broken down into two halves, N-terminal and C-terminal, connected by a 13 residues loop (Fig. 3a). The N-terminal half starts with a β-sheet structure comprised of 6 aa residues followed by six consecutive α-helices (α1- α6), several loops and 3 more β-sheets. However, the C-terminal half comprised of two β-sheets and comparatively nine longer α-helices connected by intermediate loops (Fig. 3a). The LRR domain consists of 9 LRR repeats with each repeat having one α-helix (in convex surface) and one β-sheet (in concave surface) connected by a loop (Fig. 3b). The semicircular horseshoe shape of Nalp3-LRR highly resembles the previously reported LRR structures of NOD2 and NOD1.[39, 41] The comparison of secondary structures derived from primary sequences (Fig. S1, ESI) and proposed 3D models of NACHT and LRR (Fig. S4, ESI) showed good conservations and fortified the reliability of the proposed 3D models.

3.4 Principal component analysis

The global motions of the predicted mouse Nalp3-NACHT and LRR models during the MD simulations were observed by analytical methods of PCA.[52] The porcupine plot analysis showed that a region comprising of residues ~456-486 aa of NACHT domain that is close to C-terminus with two β-sheets, one α-helix, and two loops (τ) (β–τ–α–τ–β) showed a firm upward and outward motion yielding a more flattened NACHT structure (Fig. 4a). The regions (~244-276 aa) close to N-terminal comprised of two α-helices, one β-sheet, and two loops (τ) in a fashion α–τ–α–τ–β showed an outward motion along with a small upward motion (Fig. 4a). The overall motions of these two regions close to the terminals generated a little more gap between

the N-terminal and C-terminal halves. The outward and opposite motions of the walker A and B motif regions increased the gap between them. The expanded conformations at the top end of walker A and B indicated a wide surface area is required for nucleotide binding in NACHT domain. The movements of walker A and B was shown in the Electronic supplementary video (Video, ESI). The walker A and B domains are shown in dotted secondary structures drawing by VMD in the animation (Video.mpg). In LRR domain, specifically the region LRR7-9 showed a maximum upward and inward motion in comparison to other LRR domains (Fig. 4b). These motions generated more compact structure with curved concave surfaces and β-sheets (Fig. 4b). The upward motion and bent concave surface of LRR7-9 suggest its critical biological function in PAMPs/DAMPs interaction. The importance of C-terminal region has also been previously reported in recognition of MDP and iE-DAP in NOD2 and NOD1.[41, 53]

3.5 Docking analysis

Docking analysis can be employed to investigate the best orientation and binding affinity of a ligand to its receptor. To address the possible ATP, MDP and imidazoquinoline binding sites in mouse Nalp3, docking experiments was performed by comparing the ligand conformations, position, and orientation in the complex. The interacting residues in NACHT domain with ATP predicted by docking analysis were presented in Table 2a (Fig. 5a). AutoDock results signified that both the grids (as described in materials and methodology) presented nearly equal binding affinities for ATP (Table 2a). The AutoDock results showed MDP depicted good binding energy (BE) at LRR 1-4, 3-6 and 6-9 with better ligand efficiency that measures the BE per atom of MDP to LRR model (Table 2b, Fig. 5b, c and d). Among these the C-terminal LRR6-9 regions showed highest BE, ligand efficiency and existence of hydrogen bonds (H-bonds), and shared well agreement with the previous studies that showed the C-terminal LRRs of NOD2 are critical

for bacterial recognition.[51, 53] Interaction of imidazoquinoline with LRR model in AutoDock predicted LRR 1-4 and 6-9 as the critical binding regions. The BE, ligand efficiency and H-bonds varied at both regions, and were comparatively low at the central region (Table 2c). The involvement of different amino acid residues in imidazoquinoline interaction were presented in Fig 5e, f and g. To further investigate the docking predictions and to understand the best binding regions in LRR domains for MDP and imidazoquinoline, we carried out MD simulation for each six ensembles followed by H-bond analysis, PCA, binding free energy calculations and site-directed mutagenesis analysis.

3.6 MD simulation analysis of complexes

The interaction of ATP at N-terminal sites of NACHT showed a good conformational stability and was in agreement with the previous reports.[42, 48] In human Nalp3, the walker A and walker B mutants significantly affect the ATP-binding activity. These mutants are associated with diseases by involving in IL-1β productions.[48] The binding mode of ATP with NACHT in mouse also presented the active sites at walker A and extended walker B regions that have been investigated in NOD1 and NOD2 by mutation analysis.[42] The binding of ATP close to walker A and extended walker B in mouse also suggested its involvement in disease associated cryopyrin. In LRR domain, the docking analysis showed the MDP and imidazoquinoline shared a strong binding activity close to the N and C-terminal regions. The RMSD analysis of backbone atoms in complexes during 10 ns showed a stable plateau during the 10 ns MD simulations in all complexes (Fig. 6a). In addition, the stability of MDP and imidazoquinoline at C-terminal were little unfavorable with a rising RMSD value of around 3.20 Å (Fig. 6a). RMSD of NACHT-ATP complex was < 1.5 Å, and that of MDP and imidazoquinoline at other regions (LRR1-4 and LRR3-6) in the complexes were < 2.5 Å (Fig. 6a). In LRR complex the MD trajectories showed

low RMSD values at the N-terminal regions. The radius of gyration analysis of all MD complexes also presented well compacted conformations with respect to their center of mass (Fig. 6b). The hydrogen, hydrophobic, electrostatic and van der Walls interaction residues (Table 3) in mouse Nalp3-NACHT with ATP molecule after MD simulation showed that, the ATP retained its position at the active site with a little conformational changes (Fig. 7a). The H-bond analysis of ATP-NACHT complex (Fig. 7b) also exhibited a good conservation with respect to simulation time period. The MDP-mNalp3LRR complex at N-terminal regions showed at the beginning the Asp (744,747 and 801) and Arg (771 and 776) residues contributed all H-bonds (Table 3), however, after 10 ns MD simulation no H-bond was retained between protein and MDP (Fig. 5b and 8a). At central region also after MD simulation most of the non-bonded contributions disappeared (Fig. 5c and 8b). This suggests the inappropriate binding mode or unselective MDP catalytic activity in Nalp3. The C-terminal (LRR6-9) analysis of MDP showed comparative little good bonded and non-bonded interaction (Fig. 5d and 8c) stability after 10 ns MD simulation (Table 3). However, significant conformational changes in the complex were noticed. The H-bond analysis also exhibited fluctuations in the H-bond numbers during the MD simulation for MDP at all defined sites (Fig. 8d, e and f). The low affinity of MDP may be due to less conserved residues at either N- and C-terminal regions or different LRR motif spatial arrangements. The MDP binding crucial residues in NOD2 as reported[53] are very poorly conserved in Nalp3-LRR, and this may influence the less binding stability of MDP in Nalp3-LRR domain. The steady fluctuations of MDP binding in Nalp3 also suggested a different mechanism of MDP interaction and signaling. The interaction of MDP may be influenced in heterogenic environment. Nalp3 possibly form complex with NOD2 or other NLRs to recognize MDP with a different conformation and better binding stability like NOD2-Nalp1 complex.[16, 54]

Structural analysis of LRR-imidazoquinoline complex at three different LRR regions revealed a good conservation of interacting residues at LRR7-9 (C-terminus). However, the interaction at N-terminal and central regions presented only non-bonded interactions (Fig 9a, b and c) at the end of simulation (Table 3). The H-bond analysis presented a steady H-bond number conservation at C-terminal sites in comparison to N-terminal and central region (Fig. 9d, e and f). This suggests imidazoquinoline may interact at the C-terminal region of LRR-domain. To ensure the predictions binding free energy analysis was conducted.

3.7 Binding free energy analysis and site-directed mutagenesis

The binding affinities of ATP with NACHT domain, and MDP/imidazoquinoline with the LRR model at different regions were calculated using the MM/PBSA method (Table 4). The binding free energy (BFE) calculations were performed for each complex by extracting 1000 snapshots from 2-10 ns. The BFE for each snapshot was computed as described in the material and methods. It should be noted here that experimental binding energy for ligand receptor interactions are not available for comparison. BFE of complex between ATP and NACHT was calculated to be -92.1415 kJ mol^{-1} indicating ATP's favorable conformation at the predicted active site (Table 4). In absence of experimental BE calculations for ATP interaction in previous studies of human Nalp3-, NOD1- and NOD2-NACHT domain[42, 48] this analysis showed a substantial BE is generated to stabilize the complex. The large non-polar and van der Walls interaction plays a vital role in stabilizing the Nalp3-NACHT and ATP complex. BFE analysis for MDP at three different binding sites showed that MDP yielded poor binding energy value (-11.5325kJ/mol) with a standard error of ±35.17 at LRR1-4 regions. Other binding sites of MDP resulted in positive binding energy (ΔG_{bind}), and thus suggesting energetic instability of the complex. But, the large standard error value of ΔG_{bind} at LRR1-4 than the mean ΔG_{bind} value for

MDP also directed the inappropriate/nonentity monomeric interaction of Nalp3-LRR domain (Table 4). The binding conformations and free energy may be elevated in presence of some other molecules that induces some conformational changes in Nalp3. Binding energy analysis of imidazoquinoline at different sites presented higher catalytic activity at N and C-terminal regions in terms of ΔG_{bind} value (Table 4), and suggested energetically favorable conformations of imidazoquinoline in mouse LRR domain. The calculated binding free energies of imidazoquinoline by MM-PBSA showed a good correlation with the molecular docking and the MD simulation results in terms of ligand orientations, bonded and non-bonded interactions. Comparing the H-bonds, non-bonded interactions, ΔG_{bind} value and standard error ΔG_{bind} values suggested LRR7-9 region as more appropriate binding sites for imidazoquinoline. In the other hand the very low binding free energy of MDP suggested that, it may not be a putative ligand for Nalp3 activation and signal transduction, or it may recognize MDP in a different fashion by forming complex with NOD2 or other NLR proteins.

The *in silico* mutagenesis approach was adopted in order to characterize important residues of Nalp3 that bind the ligands. The results showed that the residues Gly225, Gly227, Leu231, Lys234, Lys373, Ile366 and Leu409 were highly essential for ATP binding, since mutating these residues significantly altered the conformation and orientation of the ATP in Nalp3-NACHT domain. We have replaced the important ATP binding residues in Nalp3-NACHT with either Alanine, Cysteine or Proline and performed docking calculations. These mutagenesis yielded low binding energies of -6.1, -6.23 and -6.5 kcal mol^{-1}, respectively. The alanine scanning of walker A and B motifs significantly affect the ATP binding by lowering the BE to -3.69 and -4.89 kcal mol^{-1} respectively. However, no mutations resulted a complete loss of ATP recognitions. The low binding energies estimated after substitution mutations indicated a

reduced Nalp3-NACHT association with ATP. Alanine, Cysteine and Proline scanning of imidazoquinoline interacting residues in modeled Nalp3-LRR at LRR1-4 and 7-9 that yielded good BFE showed little binding energy variations. At LRR1-4 the Alanine, Cysteine and Proline scanning has binding energy of -5.4, -5.95 and -6.1 kcal mol^{-1}, respectively, and at LRR7-9 the BE calculated are -5.42, -4.25 and -6.87 kcal mol^{-1}, respectively. The comparison of docking scores between mutant Nalp3-LRR-imidazoquinoline and wild type Nalp3-LRR-imidazoquinoline complex together with the observed conformational changes suggest that the mutations have minimal effect on the interaction between Nalp3-LRR and imidazoquinoline. However, the cysteine scanning showed a comparatively lower docking score (-4.25 kcal mol^{-1}) by compensating the contribution of second and third Nitrogen-atoms in imidazoquinoline. The nucleophilic thiol side chain in the mutant Nalp3-LRR model also minimizes the electrostatic contributions, which indicated the probability of LRR7-9 as a site for imidazoquinoline recognition instead of LRR1-4 regions. Virtual alanine scanning followed by BE calculation using GMXAPBS tool was carried out for ATP and imidazoquinoline binding site residues. Mutation of key ATP binding residues presented at walker A and B regions (G227, K234, and P408) shows remarkable binding energy variation. Walker A residues shows comparatively higher effects on the binding affinity of ATP with NACHT domain (Table 4). The mutation of residue F369 shows no significant alteration in the BE. Mutational analysis of imidazoquinoline at LRR7-9 on residues L971, C987 and E988 presents very little BE changes. Among these, the L971 shows a comparatively higher effect on the imidazoquinoline binding affinity in mouse Nalp3-LRR domain (Table 4). The residues affecting the BE of ATP and imidazoquinoline are thought to be critical for Nalp3 signaling and transduction.

Conclusion

The innate immunity plays a key role to ward off invading pathogen and protect the host from many diseases. However, the information on newly identified classes of Nalp is sparse in public domain. This work revealed the 3D structures of mouse Nalp3 domains (NACHT and LRR) along with their dynamic features. In addition a consistency among the proposed models, molecular dynamic features and docking analysis show the reliability and robustness of the study. Overall, the change in the structure conformations, dynamics and interactions were unveiled in both NACHT and LRR domains of mouse Nalp3 receptor. This analysis explained the favorable conformations of NACHT with ATP, and LRR with ligand (MDP and imidazoquinoline) bound states based on docking, molecular dynamics and binding free energy analysis. This analysis also investigated the favorable and unfavorable mouse Nalp3-ligand complexes, and investigated the catalytic amino acid residues in NACHT and LRR domains. This work presented the complex mechanism in mouse Nalp3 and warrants an *in vivo* investigation for ATP and imidazoquinoline binding site analysis and Nalp3 signaling. Low-binding affinity of MDP also suggested a complex form of Nalp3 with NOD2 or other NLRs is required to stabilize the Nalp3-MDP complex in mouse.

Acknowledgements

The authors are thankful to Department of Biotechnology, Ministry of Science and Technology, Government of India for providing infrastructure facility to carry out this research. We would like to thank Sushma Martha, Asst. Prof., Department of Bioinformatics, Centre for Post-Graduate Studies, Orissa University of Agriculture and Technology, Bhubaneswar, Orissa; Mr. Bikram Parida, technical assistant, Institute of Minerals and Materials Technology, Bhubaneswar, Odisha, India, and Kahnu Charan Maharana, Scientist I, Sanjay Gandhi Post

Graduate Institute of Medical Science, Uttar Pradesh, India for helpful suggestions and critical reading of this manuscript.

References

1. C.A. Jr. Janeway, *Cold Spring Harb. Symp. Quant. Biol.*, 1989, **54**, 1-13.
2. S. Akira, S. Uematsu and O. Takeuchi, *Cell*, 2006, **124**, 783-801.
3. J.P. Ting, R.C. Lovering, E.S. Alnemri, J. Bertin, J.M. Boss, B.K. Davis, R.A. Flavell, S.E. Girardin, A. Godzik, J.A. Harton, H.M. Hoffman, J.P. Hugot, N. Inohara, A. Mackenzie, L.J. Maltais, G. Nunez, Y. Ogura, L.A. Otten, D. Philpott, J.C. Reed, W. Reith, S. Schreiber, V. Steimle and P.A. Ward, *Immunity*, 2008, **28**, 285-287.
4. J.P.Y. Ting, D.L. Kastner and H.M. Hoffman, *Nat. Rev. Immunol.*, 2006, **6**, 183-195.
5. J.A. Harton, M.W. Linhoff, J.H. Zhang and J.P.Y. Ting, *J. Immunol.*, 2002, **169**, 4088-4093.
6. H.M. Hoffman, F.A. Wright, D.H. Broide, A.A. Wanderer and R.D. Kolodner, *Am. J. Hum. Genet.*, 2000, **66**, 1693-1698.
7. M. Proell, S.J. Riedl, J.H. Fritz, A.M. Rojas and R. Schwarzenbacher, *PloS One*, 2008, **3**, e2119.
8. F. Martinon and J. Tschopp, *Trends. Immunol.*, 2005, **26**, 447-454.
9. F.S. Sutterwala, Y. Ogura, M. Szczepanik, M. Lara-Tejero, G.S. Lichtenberger, E.P. Grant, J. Bertin, A.J. Coyle, J.E. Galán, P.W. Askenase and R.A. Flavell, *Immunity*, 2006, **24**, 317-327.
10. K. Vilaysane, J. Chun, M.E. Seamone, W. Wang, R. Chin, S. Hirota, Y. Li, S.A. Clark, J. Tschopp, K. Trpkov, B.R. Hemmelgarn, P.L. Beck and D.A. Muruve, *J. Am. Soc. Nephrol.*, 2010, **21**, 1732-1744.

11. L. Ferrero-Miliani, O.H. Nielsen, P.S. Andersen and S.E. Girardin, *Clin. Exp. Immunol.*, 2007, **147**, 227-235.
12. S. Mariathasan, D.S. Weiss, K. Newton, J. McBride, K. O'Rourke, M. Roose-Girma, W.P. Lee, Y. Weinrauch, D.M. Monack and V.M. Dixit, *Nature*, 2006, **440**, 228-232.
13. T.D. Kanneganti, N. Ozören, M. Body-Malapel, A. Amer, J.H. Park, L. Franchi, J. Whitfield, W. Barchet, M. Colonna, P. Vandenabeele, J. Bertin, A. Coyle, E.P. Grant, S. Akira and G. Núñez, *Nature*, 2006, **440**, 233-266.
14. F. Martinon, V. Petrilli, A. Mayor, A. Tardivel and J. Tschopp, *Nature*, 2006, **440**, 237-241.
15. F. Martinon, L. Agostini, E. Meylan and J. Tschopp, *Curr. Biol.*, 2004, **14**, 1929-1934.
16. Q. Pan, J. Mathison, C. Fearns, V.V. Kravchenko, J. Da Silva Correia, H.M. Hoffman, K.S. Kobayashi, J. Bertin, E.P. Grant, A.J. Coyle, F.S. Sutterwala, Y. Ogura, R.A. Flavell, R.J. Ulevitch and *J. Leukoc. Biol.*, 2007, **82**, 177-183.
17. M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson and D.G. Higgins, *Bioinformatics*, 2007, **23**, 2947-2948.
18. P. Gouet, X. Robert and E. Courcelle, *Nucleic Acids Res.*, 2003, **31**, 3320-3323.
19. M.A. Kurowski and J.M. Bujnicki, *Nucleic Acids Res.*, 2003, **31**, 3305-3307.
20. K. Ginalski, A. Elofsson, D. Fischer and L. Rychlewski, *Bioinformatics*, 2003, **19**, 1015-1018.
21. Y. Zhang, *BMC Bioinformatics*, 2008, **9**, 40.
22. S. Wu and Y. Zhang, *Nucleic Acids Res.*, 2007, **35**, 3375-3382.
23. A. Sali and T.L. Blundell, *J. Mol. Biol.*, 1993, **234**, 779-815.

24. V.B. Chen, W.B. 3rd Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson and D.C. Richardson, *Acta. Crystallogr. D. Biol. Crystallogr.,* 2010, **66**, 12-21.
25. G. Vriend, *J Mol Graph.,* 1990, **8**, 52-56.
26. N. Guex and M.C. Peitsch, *Electrophoresis,* 1997, **18**, 2714-2723.
27. D. Bhattacharya, J. Cheng, *Proteins,* 2013, **81**, 119-131.
28. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark and H.J. Berendsen, *J. Comput. Chem.,* 2005, **26**, 1701-1718.
29. B. Hess, C. Kutzner, D. Van Der Spoel and E. Lindahl, *J. Chem. Theory. Comp.,* 2008, **4**, 435-447.
30. W.L. Jorgensen, D.S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.,* 1996, **118**, 11225-11236.
31. G.A. Kaminski, R.A. Friesner, J. Tirado-Rives and W.L. Jorgensen, *J. Phys. Chem. B.,* 2001, **105**, 6474-6487.
32. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.,* 1996, **14**, 33-38.
33. M. Wiederstein and M.J. Sippl, *Nucleic Acids Res.,* 2007, **35**, 407-410.
34. B. Wallner and A. Elofsson, *Protein Sci.,* 2003, **12**, 1073-1086.
35. L. Willard, A. Ranjan, H. Zhang, H. Monzavi, R.F. Boyko, B.D. Sykes and D.S. Wishart, *Nucleic Acids Res.,* 2003, **31**, 3316-3319.
36. M. Berjanskii, P. Tang, J. Liang, J.A. Cruz, J. Zhou, Y. Zhou, E. Bassett, C. MacDonell, P. Lu, G. Lin and D.S. Wishart, *Nucleic Acids Res.,* 2009, **37**, 670-677.
37. A.W. Schuttelkopf and D.M. van Aalten, *Acta. Crystallogr. D. Biol. Crystallogr.,* 2004, **60**, 1355-1363.

519 38. G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell and A.J.
520 Olson, *J Comput Chem.,* 2009, **30**, 2785-2791.
521 39. J. Maharana, B. Swain, B.R. Sahoo, M.R. Dikhit, M. Basu, A.S. Mahapatra, P. Jayasankar
522 and M. Samanta, *Fish. Physiol. Biochem.,* 2013, **39**, 1007-1023.
523 40. B.R. Sahoo, M. Basu, B. Swain, M.R. Dikhit, P. Jayasankar and M. Samanta, *Biomed Res*
524 *Int.,* 2013, 185282, DOI: 10.1155/2013/185282 .
525 41. B.R. Sahoo, B. Swain, M.R. Dikhit, M. Basu, A. Bej, P. Jayasankar and M. Samanta, *Appl.*
526 *Biochem. Biotechnol.,* 2013, **170**, 1282-1309.
527 42. B. Zurek, M. Proell, R.N. Wagner, R. Schwarzenbacher and T.A. Kufer, *Innate. Immun.,*
528 2012, **18**, 100-111.
529 43. J. Mo, J.P. Boyle, C.B. Howard, T.P. Monie, B.K. Davis and J.A. Duncan, *J. Biol. Chem.,*
530 2012, **287**, 23057-23067.
531 44. P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan,
532 W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case and T.E. 3rd Cheatham, *Acc,*
533 *Chem, Res.,* 2000, **33**, 889-897.
534 45. D. Spiliotopoulos, A. Spitaleri and G. Musco, *PLoS One,* 2012, **7**, e46902.
535 46. S. Huo, I. Massova and P.A. Kollman, *J. Comput. Chem.,* 2002, **23**, 15-27.
536 47. M. Saraste, P.R. Sibbald and A. Wittinghofer, *Trends. Biochem. Sci.,* 1990, **15**, 430-434.
537 48. J.A. Duncan, D.T. Bergstralh, Y. Wang, S.B. Willingham, Z. Ye, A.G. Zimmermann and
538 J.P. Ting, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 8041-8046.
539 49. J.A. MacDonald, C.P. Wijekoon, K.C. Liao and D.A. Muruve, *IUBMB Life,* 2013, **65**, 851-
540 862.
541 50. L.J. McGuffin, K. Bryson and D.T. Jones, *Bioinformatics,* 2000, **16**, 404-405.

51. J.A. Mótyán, P. Bagossi, S. Benko and J. Tozser. *BMC Bioinformatics*, 2013, **14**, 275.

52. A. Amadei, A.B. Linssen and H.J. Berendsen., *Proteins*, 1993, **17**, 412-425.

53. T. Tanabe, M. Chamaillard, Y. Ogura, L. Zhu, S. Qiu, J. Masumoto, P. Ghosh, A. Moran, M.M. Predergast, G. Tromp, C.J. Williams, N. Inohara and G. Núñez, *EMBO J.*, 2004, **23**, 1587-1597.

54. L.C. Hsu, S.R. Ali, S. McGillivray, P.H. Tseng, S. Mariathasan, E.W. Humke, L. Eckmann, J.J. Powell, V. Nizet, V.M. Dixit and M. Karin, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 7803-7808.

Figure legends

Fig. 1 Multiple sequence alignment of mouse Nalp3-NACHT and LRR domains with respective domain sequences of NOD1, NOD2 and Nalp3 of human and mouse. **A)** NACHT domain alignment. The potential ATP binding sites in NACHT domains of Nalp3, NOD1 and NOD2 are shown inside the rectangular box, **B)** LRR domain alignment. The 9 different LRR motifs are marked in the consensus sequence. The alignments were generated by ClustalW program and are presented in ESPript 2.2 server. Consensus residues are shown in box and are highlighted, and the leucine-rich repeat (LRR) regions are indicated in using "LxxLxLxxNxL" motif

Fig. 2 Conformational analysis of mouse Nalp3-NACHT and LRR domains. **A)** Cα RMSD, **B)** radius of gyration (Rg), **C)** Cα RMSF of Nalp3-NACHT domain, **D)** Cα RMSF of Nalp3-LRR domain

Fig. 3 Models for the Nalp3-NACHT and LRR domain. **A)** NACHT, **B)** LRR, The cartoon presentation of both models in PyMOL presents the α-helices (red), β-sheets (yellow) and loops (green)

564 **Fig. 4** Prevalent motions in Nalp3-NACHT and LRR domain using principal component analysis.
565 **A)** Porcupine plot of the first eigenvector in NACHT model, **B)** LRR model. The flexible N and
566 C-terminals are presented in different colors. In NACHT, the light blue and purple blue
567 represents the walker A and B, respectively with an outward motions. The 9 LRR regions in
568 LRR model are presented in 9 different colors
569 **Fig. 5** Molecular interaction of ligands with mouse Nalp3 domains in AutoDock 4.2. **A)** ATP
570 and mNalp3-NACHT domain**, B)** MDP and mNalp3-LRR domain at LRR1-4, **C)** MDP and
571 mNalp3-LRR domain at LRR3-6, **D)** MDP and mNalp3-LRR domain at LRR6-9, **E)**
572 imidazoquinoline and mNalp3-LRR domain at LRR1-4, **F)** imidazoquinoline and mNalp3-LRR
573 domain at LRR6-8, **G)** imidazoquinoline and mNalp3-LRR domain at LRR7-9. The ligands
574 (ATP, MDP and imidazoquinoline) are shown as sticks, protein (mNalp3-NACHT/LRR) models
575 as cartoons, and hydrogen bonds as red dotted lines
576 **Fig. 6** Stability parameters for mNalp3-NACHT/LRR complexes over MD simulation time
577 period. **A)** Cα RMSD, **B)** radius of gyration (Rg). Different colors are used to represent the
578 seven complexes in the graph as shown in the legends
579 **Fig. 7** Structure analysis of mNalp3-NACHT and ATP complex after MD simulation. **A)**
580 Binding site residues of ATP in mNalp3-NACHT domain. The ATP is shown as stick and
581 mNalp3-NACHT model as cartoon, **B)** hydrogen bond (H-bond) variations between ATP and
582 mNalp3-NACHT model
583 **Fig. 8** Binding site analysis of mNalp3-LRR and MDP complexes after 10 ns MD simulation. **A)**
584 orientation of MDP at LRR1-4, **B)** orientation of MDP at LRR3-6, **C)** orientation of MDP at
585 LRR6-9. The MDP is shown as stick and mNalp3-LRR model as cartoon, **D)** H-bond variations

in MDP- mNalp3-NACHT complex at LRR1-4, **E)** H-bond variations in MDP- mNalp3-NACHT complex at LRR3-6, **F)** H-bond variations in MDP- mNalp3-NACHT complex at LRR6-9

Fig. 9 Conformational analysis of mNalp3-LRR and imidazoquinoline complexes after 10 ns MD simulation. **A)** Orientation of imidazoquinoline at LRR1-4, **B)** orientation of imidazoquinoline at LRR6-8, **C)** orientation of imidazoquinoline at LRR7-9. The imidazoquinoline is shown as stick and mNalp3-LRR model as cartoon, **D)** H-bond variations in imidazoquinoline-mNalp3-NACHT complex at LRR1-4, **E)** H-bond variations in imidazoquinoline-mNalp3-NACHT complex at LRR6-8, **F)** H-bond variations in imidazoquinoline-mNalp3-NACHT complex at LRR7-9

Supplementary materials

Fig. S1 Secondary structure prediction by PSI-PRED **A)** mNalp3-NACHT domain, **B)** mNalp3-LRR domain. Representation of helix, b-sheets and loops are shown in the legend. Amino acid numbers depicted in the figure represents numbers 216-532 (mNalp3-NACHT) and 739-988 (mNalp3-LRR) in mouse Nalp3 protein

Fig. S2 Secondary structure assignment as a function of MD simulation time. **A)** mNalp3-NACHT domain, **B)** mNalp3-LRR domain. Legends of different structural elements are presented. Amino acid numbers depicted in the figure represents numbers 216-532 (mNalp3-NACHT) and 739-988 (mNalp3-LRR) in mouse Nalp3 protein

Fig. S3 Validation reports of mNalp3-NACHT and LRR models by Ramachandran plot and ProSA analysis

Fig. S4 Secondary structure assignments of **A)** mNalp3-NACHT and **B)** LRR domains from their 3D-models by ProFunc program. Representation of secondary structural elements are given at the right hand bottom corner

609 **Video**

610 The separation of walker A and B forming a wider space during the MD simulation was presented
611 in the video. The walker A and B regions were highlighted with dotted representation in VMD
612 program.

Table 1 Model validation reports of modeled mouse Nalp3 domains

Servers		**Mouse Nalp3-NACHT**	**Mouse Nalp3-LRR**
Procheck	Most favored regions (%)	81.70	82.70
	Additionally allowed regions (%)	16.60	15.90
	Generously allowed regions (%)	1.40	0.90
	Disallowed regions (%)	0.30	0.40
	Overall G-factor	-0.07	-0.04
Verify3D	Averaged 3D-1D score > 0.2	89.62	100.00
ERRAT	Overall quality	90.29	92.98
ProSA	Z-score	-6.83	-4.89
ProQ	LG score	6.28	6.34
	MaxSub	0.56	0.53
MolProbity	Residues with bad bonds (%)	0.00	0.00
	Residues with bad angles (%)	0.06	0.00
Vadar	Standard deviation of χ1 pooled	2.27	1.69
	Mean H-bond energy	0.65	0.65
	Generously allowed Ω angles (%)	-1.60	-1.60
	Packing defects (%)	-0.36	0.36
GeNMR	Ramachandran outside of most favored	1.34	1.48
	Bump score	0.34	0.09
	Radius gyration score	1.34	1.96

Table 2 Molecular docking analysis of ATP, MDP and imidazoquinoline with mouse Nalp3 domain models

Grid Area	Binding Energy kcal mol^{-1}	Ligand Efficiency	Hydrogen bonds	
A) NACHT-ATP				
Full grid	-8.00	-0.19	4	
Walker A and B	-8.59	-0.28	3	
B) LRR-MDP				**LRR contributor**
Full Grid	-4.11	-0.12	7	LRR4 - LRR8
LRR 1-4	-3.71	-0.11	6	LRR1 - LRR4
LRR 2-5	-3.78	-0.11	5	LRR3 - LRR7
LRR 3-6	-4.19	-0.12	8	LRR2 - LRR7
LRR 4-7	-3.09	-0.09	9	LRR4 – LRR9
LRR 5-8	-3.94	-0.12	9	LRR4 – LRR9
LRR 6-9	-4.26	-0.13	9	LRR4 – LRR9
C) LRR- imidazoquinoline				
Full Grid	-5.88	-0.39	3	LRR 1 – 2
LRR 1-2	-5.69	-0.38	2	LRR 1 – 2
LRR 1-4	-5.68	-0.38	3	LRR 1 – 2
LRR 4-8	-5.06	-0.34	2	LRR 6 – 9
LRR 6-8	-5.15	-0.34	2	LRR 6 – 9
LRR 5-9	-5.42	-0.36	3	LRR 8 - 9
LRR 7-9	-5.45	-0.36	3	LRR 8 - 9

Table 3 Interaction analysis of mouse Nalp3-ligand complexes before and after MD simulation

Complex	**Interactions**	**After docking**	**After 10 ns MD simulation**
NACHT-ATP			
	H-bond	Lys234, Lys373, Ile366, Leu409	Arg233, Lys234
	Hydrophobic	Leu231, Phe369,Leu307	Tyr277, Phe369, Leu409
	Electrostatic	Gly225, Gly227, Leu367, Tyr381, Pro408	Gly227, Ile230, Tyr381,
	van der walls	Ile226, Ile230, Ile407	Ile226, Leu231, Leu237, Leu367, Ile407, Pro408
LRR-MDP(1-4)			
	H-bond	Asp744, Asp747, Arg771, Arg776, Asp801	-
	Hydrophobic	Leu743, Leu745, Trp773	Leu745, Trp773
	Electrostatic	Ser746, Gly775, Ser803	Thr749, Arg776, Cys777
	van der walls	Glu799, Asp804	Asp744, Asn748,
LRR-MDP (3-6)			
	H-bond	Arg771, Glu799, Asp801, Lys828, Arg856, Lys885	Arg856
	Hydrophobic	Trp830	-
	Electrostatic	Tyr858	-
	van der walls	-	-
LRR-MDP (6-9)			
	H-bond	Arg856, Lys885, Glu944, Asp946, Asn972	Lys885, Tyr915
	Hydrophobic	Tyr858, Tyr915	-
	Electrostatic	Arg917, Leu973, Gly974	Arg856
	van der walls	-	Trp830, Arg917
LRR-Imidazoquinoline (1-4)			
	H-bond	Leu743, Leu745	-
	Hydrophobic	Trp773	Trp773
	Electrostatic	Gly775, Arg776	-
	van der walls	Cys777, Asp801, Ser803, Asp804	Leu745, Ser746, Asp747, Met755, Leu774, Gly755, Arg776, Cys777, Ser803
LRR-Imidazoquinoline (6-8)			
	H-bond	Leu916, Glu944, Asp946	-
	Hydrophobic	Tyr915	Trp830, Tyr858, Tyr915
	Electrostatic	Leu945	Glu944, Asn972
	van der walls	Val889, Asn972, Leu973, Gly974	Leu888, Val889, Leu916, Arg917, Leu945, Asp946, Gly974
LRR-Imidazoquinoline (7-9)			
	H-bond	Leu962, Ser959	Cys987, Glu988
	Hydrophobic	-	Leu971
	Electrostatic	Thr960, Arg969	-
	van der walls	Thr963, Asn965, Leu968, Val984, Glu988	Cys955, Leu958, Ser959, Leu962, Leu968, Arg969, Leu973, Val984, Thr985, Leu986

Table 4 Binding free energy (kJ mol^{-1}) calculation by MM/PBSA method in mouse Nalp3-ligand complexes

Conformations	**[1]ΔG_{bind}**	**Polar Contribution**		**[4]ΔG_{polar}**	**Non-polar Contribution**		**[7]$\Delta G_{nonpolar}$**
		[2]ΔG_{coul}	**[3]ΔG_{ps}**		**[5]ΔG_{vdw}**	**[6]ΔG_{nps}**	
NACHT-ATP	-92.14 (70.64)	-17.31 (20.33)	101.65 (28.67)	84.33	-159.55 (70.46)	-16.92 (0.93)	-176.48
LRR-MDP(1-4)	-11.53 (35.17)	247.39 (69.03)	-188.09 (90.48)	59.29	-60.82 (20.55)	-10.00 (1.87)	-70.82
LRR-MDP(3-6)	32.03 (53.68)	-122.17 (238.00)	199.71 (247.50)	77.54	-36.97 (48.22)	-8.53 (3.96)	-45.51
LRR-MDP(6-9)	19.00 (57.15)	-106.81 (170.50)	192.82 (52.84)	86.01	-54.89 (52.84)	-12.11 (1.85)	-67.00
LRR-imidazoquinoline(1-4)	-141.38 (94.49)	-45.00 (32.33)	87.54 (34.06)	42.54	-173.45 (92.93)	-10.46 (1.18)	-183.92
LRR-imidazoquinoline(6-8)	27.14 (114.46)	-81.40 (71.77)	128.78 (51.99)	47.38	-9.90 (102.31)	-10.33 (0.53)	-20.23
LRR-imidazoquinoline(7-9)	-166.37 (89.52)	-36.35 (27.48)	82.35 (28.76)	46.00	-200.99 (88.61)	-11.37 (1.30)	-212.37
NACHT-ATP alanine scanning							
GLY227-ALA227	-85.06 (75.13)	-13.43(16.02)	96.91 (24.78)	83.48	-147.32 (74.88)	-21.21 (1.15)	-168.54
LYS234-ALA234	-86.41 (81.04)	-19.57 (25.24)	108.62 (32.11)	89.05	-162.31 (80.93)	-13.15 (0.89)	-175.46
PHE369-ALA369	-94.35(71.70)	-18.22 (20.94)	103.06 (27.85)	84.84	-161.38 (71.12)	-17.81 (1.06)	-179.19
PRO408-ALA408	-89.98 (73.07)	-16.38 (18.96)	95.43 (27.11)	79.05	-149.21 (72.70)	-19.82 (1.45)	-169.03
LRR-imidazoquinoline (7-9) alanine scanning							
LEU971-ALA971	-152.77 (93.23)	-31.44 (29.51)	85.39 (27. 06)	53.94	-202.32 (91.72)	-14.39 (1.13)	-206. 71
CYS987-ALA987	-166.65 (82.49)	-27.76 (31.82)	83.39 (32.09)	55.63	-212.42 (80.88)	-9.85 (1.61)	-222.28
GLU988-ALA988	-158.30 (83.26)	-29.71 (33.75)	76.38 (31.19)	46.67	-192.42 (81.01)	-12.54 (0.98)	-204.97

[1]Binding free energy.
[2]Coulombic term.
[3]Polar solvation terms.
[4]Polar solvation energy.
[5]van der Waals energy.
[6]Nonpolar solvation energy.
[7]Nonpolar solvation terms.
Standard errors are presented in parenthesis.

Fig. 1

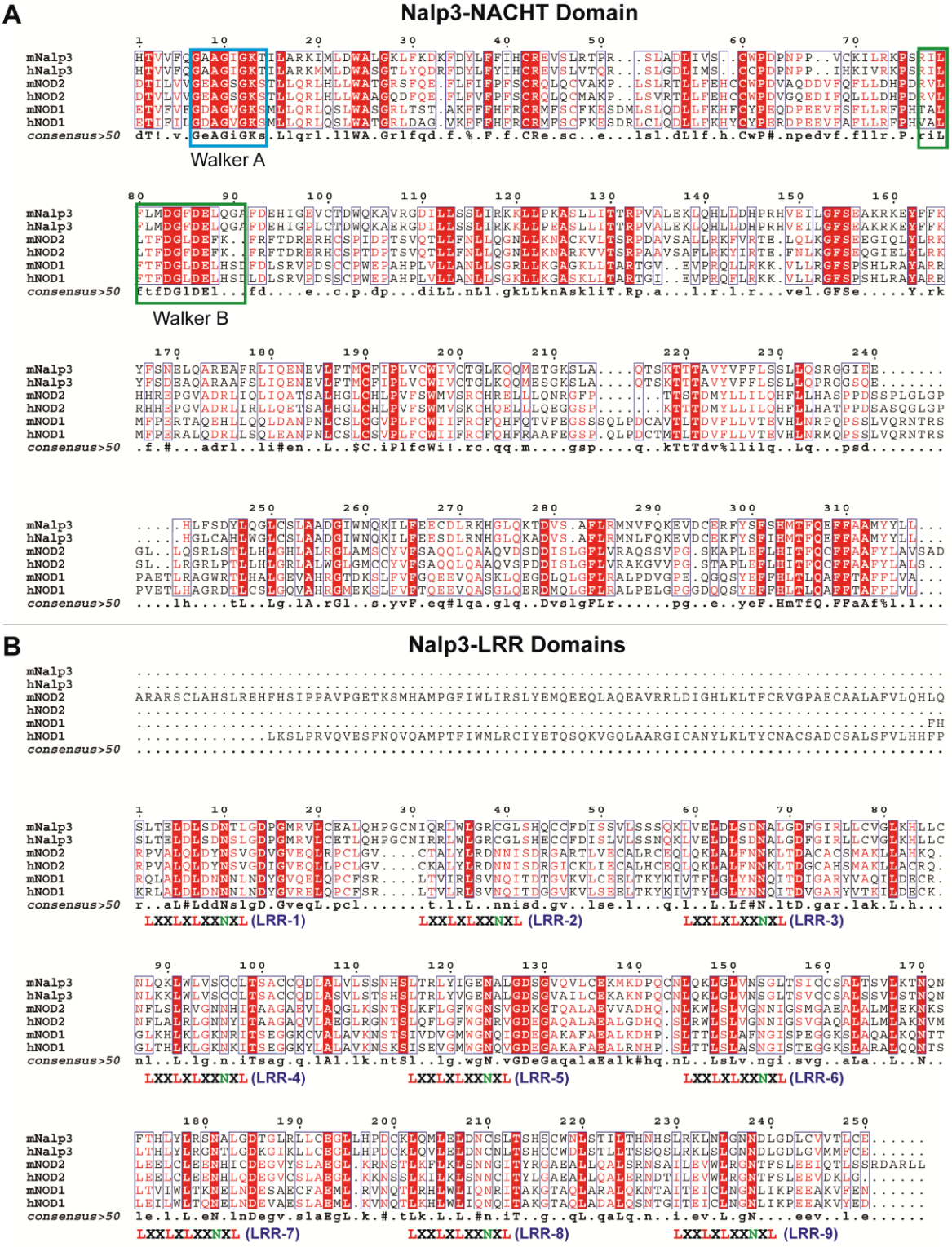


Fig. 2

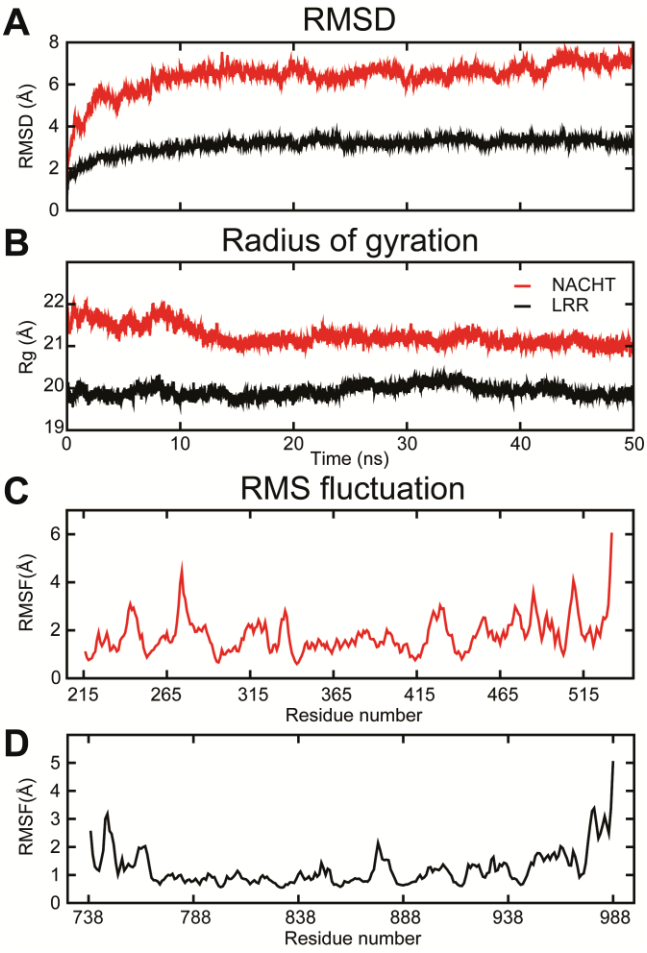


Fig. 3

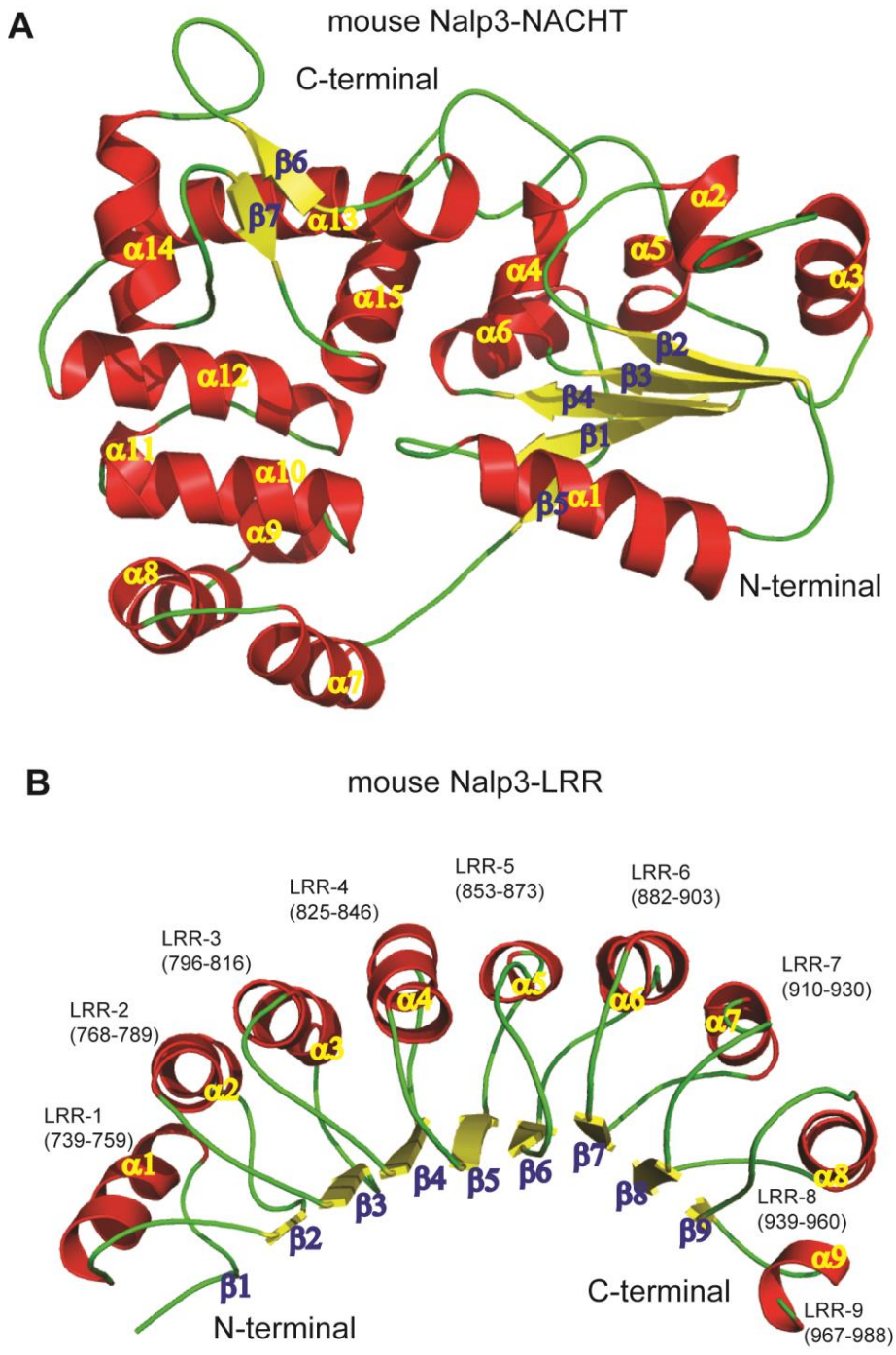


Fig. 4

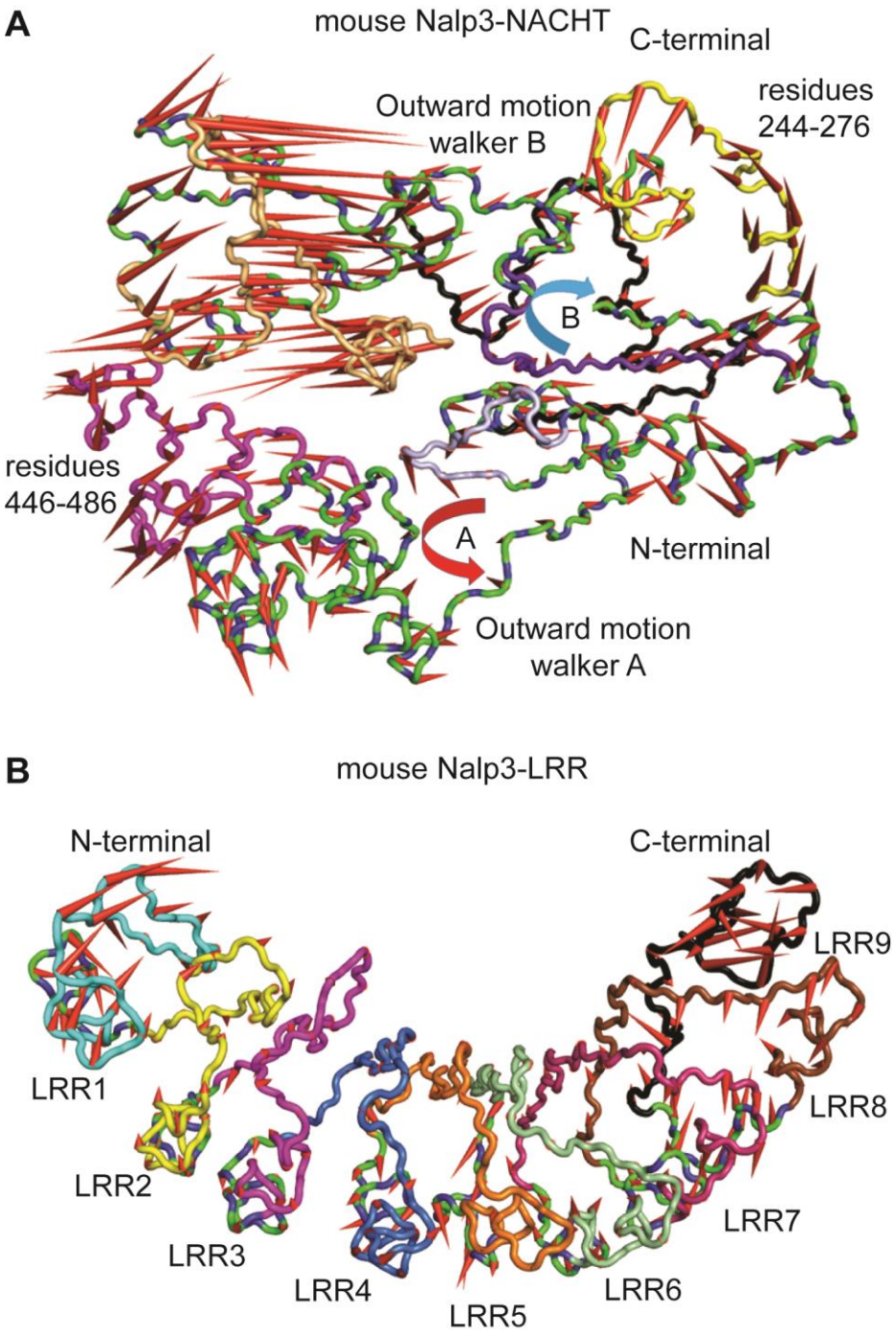


Fig. 5

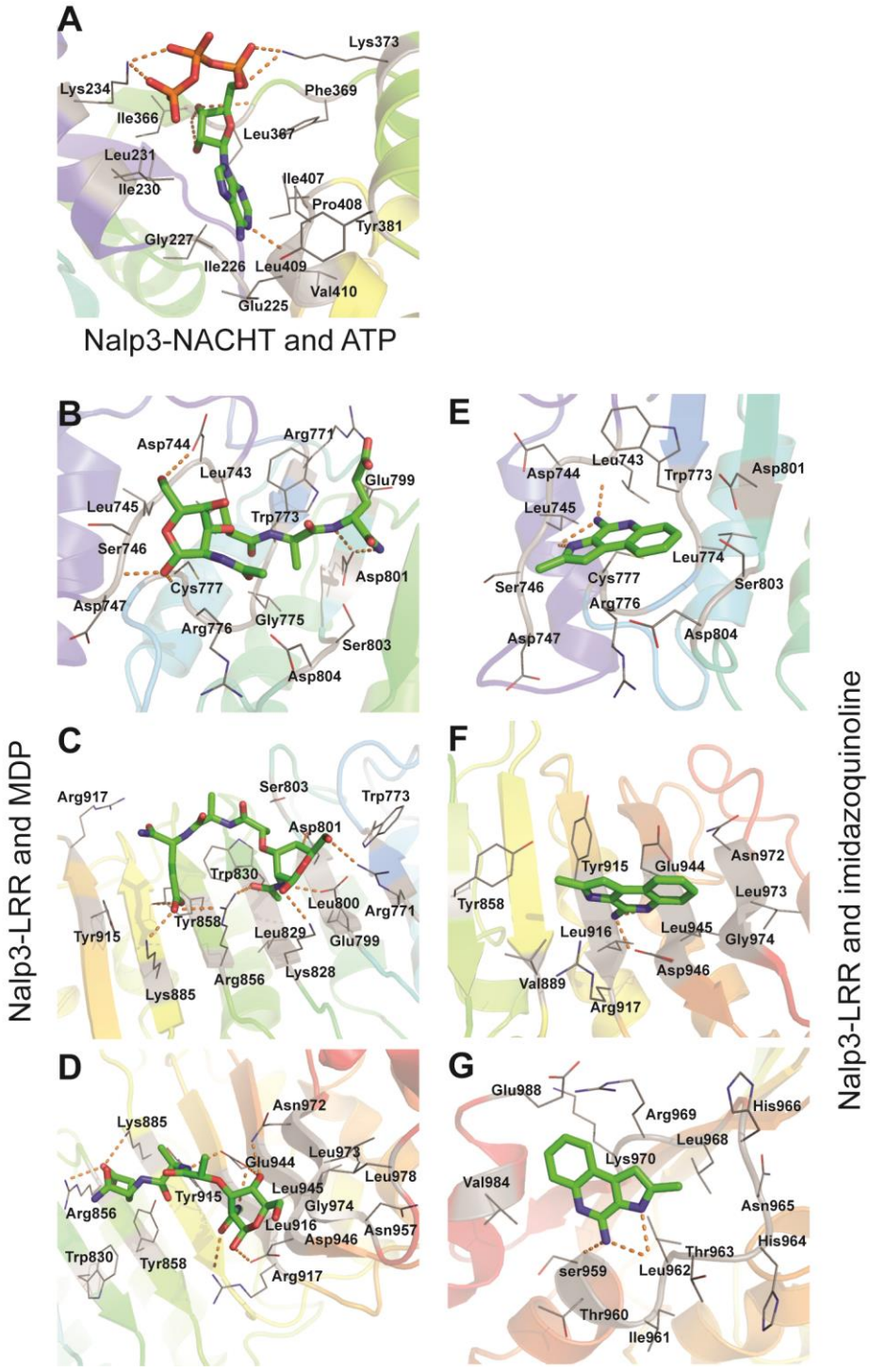


Fig. 6

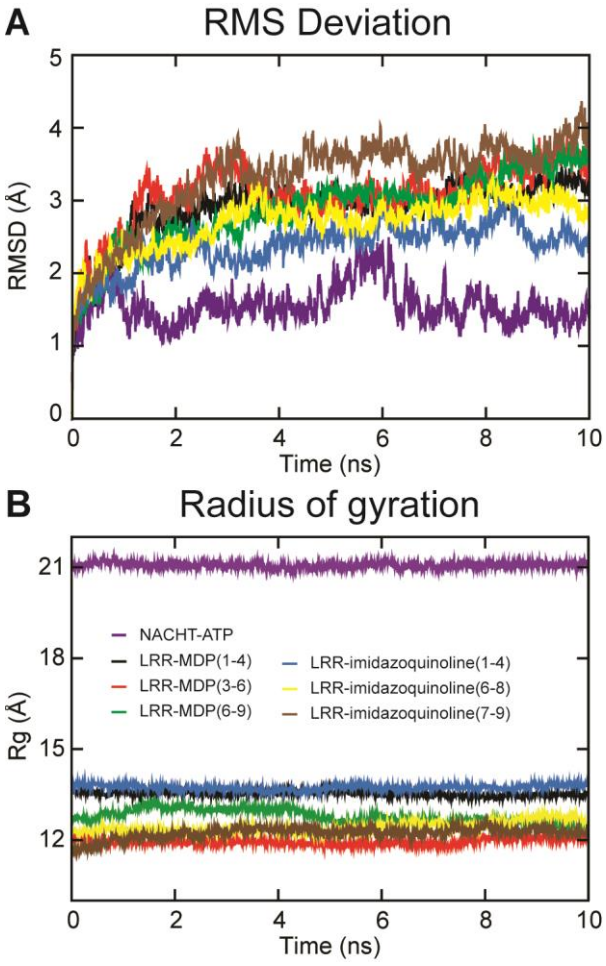


Fig. 7

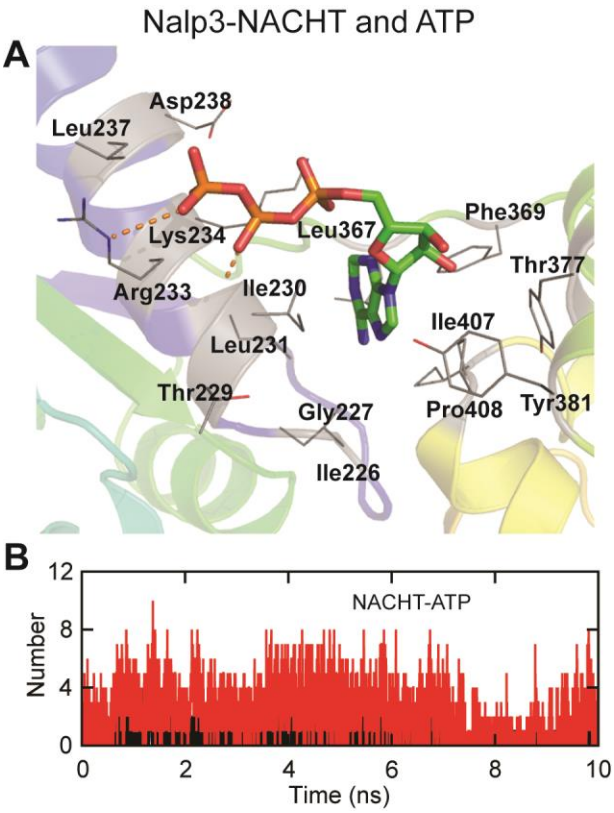


Fig. 8

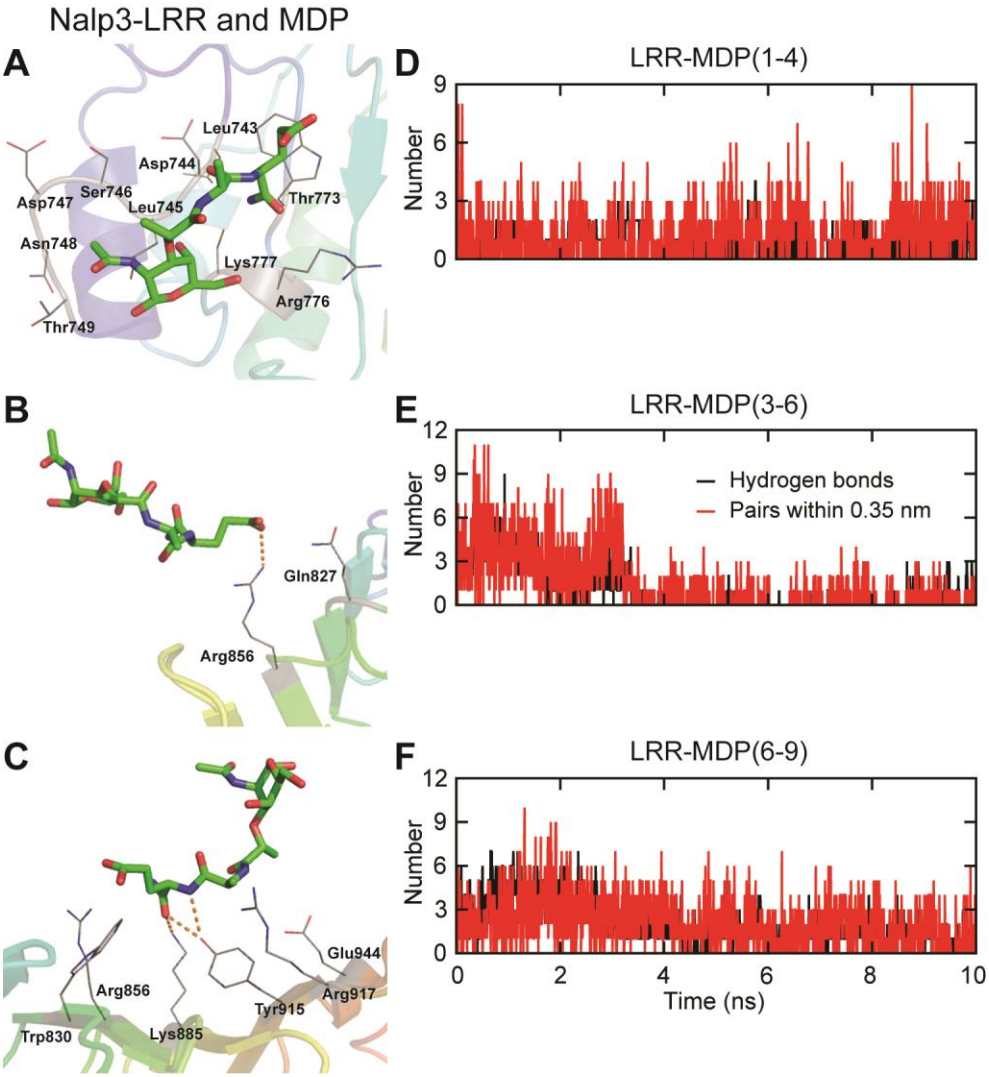


Fig. 9