Mechanistic insights into the deleterious role of nasu-hakola disease associated TREM2 variants

Raju Dash, Ho Jin Choi, Il Soo Moon*

Department of Anatomy, Dongguk University Graduate School of Medicine, Gyeongju 38066,
Republic of Korea

* Correspondence:

Il Soo Moon

Email: moonis@dongguk.ac.kr

Abstract

Recently, critical roles of genetic variants in Triggering Receptor Expressed on Myeloid cells 2 (TREM2) for myeloid cells to Alzheimer's disease have been aggressively highlighted. However, little studies focused to the deleterious role of Nasu-Hakola disease (NHD) associated TREM2 variants. In order to get insights into the contributions of these variants in neurodegeneration, we investigated the influences of three well-known NHD associated TREM2 mutations (Y38C, T66M and V126G) on the loss-of-function by using conventional molecular dynamics simulation. Compared to the wild type, the mutants produced substantial differences in the collective motions in the loop regions, which not only promotes structural remodelling in complementarity-determining region 2 (CDR2) loop but also in CDR1 loop, through changing the inter and intra-loop hydrogen bonding network. In addition, the structural studies from free energy landscape showed that Y38, T66 and V126 are crucial for maintaining structural features of CDR1 and CDR2 loops, while their mutation at this position produced steric clash and thus contributes to the structural impact and loss of ligand binding. These results revealed that the presence of the mutations in TREM2 ectodomain induced flexibility and promotes structural alterations. Dynamical scenarios, which are provided by the present study, may be critical to our understanding of the role of the three TREM2 mutations in neurodegenerative diseases.

Keywords: TREM2, Nasu-Hakola disease, variants, Molecular Dynamics Simulation

Introduction

Accumulation of amyloid- β (A β) in brain parenchyma is the main hallmark of Alzheimer's disease, which is implicated to the slowly progressive decline of cognitive function, as a result of intracellular neurofibrillary tangles formation, synapse loss and cell death ^{1,2}. Although the precise mechanisms and molecular determinants of this neurodegenerative disease remain incomplete, recent studies covering with whole genome sequencing identified that the altered genetic loci including Triggering Receptor Expressed on Myeloid cells 2 (TREM2) is correlated with a drastically increased risk of progression to Alzheimer's disease ^{3,4}.

TREM2 is expressed especially in osteoclasts, microglia, alveolar macrophages, and other mononuclear phagocytes, which is a V-type immunoglobulin (Ig) domain-containing transmembrane protein ⁴. The activation of TREM2 is initiated by anionic lipid such as bacterial lipopolysaccharide (LPS) and phospholipids, and also several putative ligands like apolipoprotein J (apoJ) and apolipoprotein E (apoE) are reported to bind with TREM2 ⁵⁻⁷. Upon the binding of ligands, TREM2 recruits protein tyrosine kinase SYK through an adapter protein, which is associated with the transmembrane region of TREM2, is known as DNAX-activating protein of 12 kDa (DAP12) ⁸⁻¹⁰. Subsequently, a cascade of signaling events, including MAPK and PI3K activations is followed and thus regulates the phagocytosis of cellular debris and inflammatory responses of microglia ¹⁰⁻¹². In both early and mid-term AD, TREM2 plays a protective role ¹³, where its overexpression is involved to clearing the soluble and insoluble A β 42 aggregates from the brain ^{14,15}. On the other hand, TREM2 is reported to avert the accumulation and diffusion of A β by modulating microglial activation around amyloid plaques ¹⁶⁻¹⁹.

At molecular level, the genetic variations in TREM2 are linked with the nasu-hakola disease (NHD) and frontotemporal dementia, where NHD is characterized by demyelination, early-onset dementia, and bone cyst lipoma and linked with the ectodomain Y38C, T66M, and V126G mutations ²⁰⁻²². According to findings of *Kober et al* ²³ it has been appeared that the NHD mutations may cause the changes of buried residues in TREM2, and promote misfolding with inconsistent impact on surface expression and aggregation. Previous studies explained the influence of TREM2 structure and function by AD associated variants including R47H. However, the deleterious role of NHD-associated mutations is still need to be studied.

The present study considered Y38C, T66M, and V126G mutants of TREM2 in order to characterize the dynamic behavior of TREM2 protein in wild and mutated state by classical molecular dynamics simulation.

Results and Discussion

Wild and Mutant type of TREM2

The natural human TREM2 structure is composed of 230 amino acids, a polypeptide chain that consists of three distinct regions including N-terminal mature ectodomain (ECD, residues 19-174), membrane spanning region (residues 175-195) and C-terminal cytosolic tail (residues 196-230). The rest of amino acids, especially residues 1-18 act as the signaling peptide in TREM2 signaling cascade. As shown in **Figure 1**, the tertiary structure of TREM2 ECD domain is mainly composed of nine β -strands (β A - β F), including three major complementarity-determining regions, dubbed as CDR loops, mainly CDR1 (residues from Pro³⁷ to Arg⁴⁷), CDR2 (residues Thr⁶⁶ to Arg⁷⁶) and CDR3 (His¹⁴⁴ to Glu¹¹⁷), respectively. Previous studies showed that CDR2 maintain a stable conformation in normal condition by maintaining hydrogen bonding with the CDR1 loop, which is necessary for ligand interactions²⁴. However, the hydrogen bonding network is lost by R47H variants, as it is involved in the network maintenance, which facilitates conformational remodeling of CDR2 loop. Interestingly, the mutations of our interest, including Y38C and T66M are located in both CDR1 and CDR2 regions, respectively. Therefore, how these mutations contribute to pathological behavior of TREM2 is systematically analyzed in the following subsections.

Alteration of conformational stability by TREM2 variants

In last decades, the molecular simulation has brought a new platform to detail characterize the structural configuration of macromolecules in various environments at microscopic scale, which is related to their functions and interactions with other molecular species^{25,26}. In this context, molecular dynamics simulations of wild type and other three substitutions (Y38C, T66M and V126G) were conducted for 100 ns of each system to describe their structural dynamics and stability. The conformational stability of the wild and mutant types in the simulation have been analyzed by calculating RMSD for the backbones of all proteins and rendered in **Figure 2**. The RMSD analysis revealed that the wild type and Y38C achieved

equilibrium after 5 ns and maintained till to the end of simulation, while rest of the mutants, including T66M and V126G showed system stability following 20 ns. The native and Y38C mutant structures showed similar way of deviation till 100 ns from their starting structure, resulting in backbone RMSD of ~ 1.4 to $1.7\text{\AA}$ during the simulations. However, T66M and V126G mutant structures showed significantly different deviations from the wild and Y38C structures, ranging from ~ 1.9 to $3\text{\AA}$. Since the magnitude of fluctuations of four different systems after relaxation period were maintained stable over the time, the simulation is considered to attain convergence, i.e. a stable conformation²⁷, that may lead to produce stable trajectories which is appropriate basis for further analysis. Therefore, the conformational changes in the wild and mutant structure were evaluated by the radius of gyration (Rg) (**Figure 3**), which describes the overall dimension of protein structure in terms of compactness. Higher values in Rg represent the lower compactness^{28,29}. As shown in **Figure 3 a & b**, the overall dimension of Y38C and T66M structure maintained similar behaviour as like wild type structure, while V126G (**Figure 3c**) structure showed increased compactness after 50 ns, till to the end of simulation. The SASA, which account for solvent accessibility, showed similar profile like Rg for mutants and wild structure. Whereas native and mutant (Y38C and T66M) structures showed similar fashion of deviation from the initial structure (**Figure 4 a & b**), other than V126G structure increased the SASA after 50 ns (**Figure 4c**). Increase in SASA of mutant denotes its relatively bloated nature as compared to the native structure³⁰. Although RMSD analysis depicted the alterations in TREM2 conformation by T66M and V126G mutations, V126G alone showed structure alteration by means of Rg and SASA, which concluded that this mutation decreased the overall dimension of the protein that leads to the increase in solvent accessibility in order to destabilized the protein folding and affecting protein-protein interactions³¹.

Figure 5 rendered RMSF of the each protein backbone, which shows the local changes in protein structure due to the mutations. The local residue fluctuations in the protein serve as an important factor to describe the biological functions, as the many functional sites in the protein have the structural fluctuations and uniquely coupled³²⁻³⁴. As represented in **Figure 5a**, TREM2 structure contains ten loops, including three CDR loops; all were mostly fluctuated throughout the simulations. The degree of fluctuation is high in mutant structures than the wild type. Interestingly, Y38C and T66M mutant structures showed high fluctuations in the CDR1 and CDR2 regions, respectively. On the other hand, V126G mutant increased overall local flexibility

of the protein, especially in the regions including residues 50 to 60 and 85 to 95, CDR1 loop and N-terminal tail, respectively. This analysis suggested that mutations have significant impact on the local structural rearrangement rather than overall structure. However, to view more insight, per residue SASA results of all wild and mutant structure have been calculated and represented in **Figure 5b**. The mutant structures increased local solvent accessibility of the protein, especially by T66M in CDR1 and CDR2 regions (**Figure 5b**). Comparatively, the deviations in SASA were also observed by the Y38C and V126G mutants, and it was noticed in the CDR1 and CDR2 loop regions. The SASA of per residue elucidates the structural conformational changes to the protein surface, where lower value means thermodynamically stability as well as buried inside the core of the protein³⁵⁻³⁷. This analysis together with RMSF study suggested that nasu-hakola disease-associated TREM2 variants provide larger surface area than the native, which is thermodynamically unstable and have major impact on ligand binding.

Effect of mutations in protein dynamics

In different functional states, protein undergoes various conformational transitions by the global domain motions, which is facilitated by the collective motion of backbone atoms³⁸. Principally, various regions including the regions having secondary structures move in a correlated manner³⁹. For additional insights into concerted dynamics of TREM2 in wild and mutant types, DCCM and principle component analyses were performed on the MD trajectories from different system. The DCCM covers total correlated motions among the protein residues, where the strong correlated motions between specific residues are highlighted by red regions (Figure 6). On the other hand, highly anti-correlation is expressed by blue regions. The analysis revealed that mutations induced completely different motions than the wild type structure. Comparatively, Y38C (**Figure 6b**) produced significant anti-correlative motion between residues 17-27 (CDR1 loop, residues 37 to 47) and residues 45 to 56 (CDR2 loop, residues 65 to 76) as well as 65 to 74 (residues, 85 to 94). On the other hand, T66M (**Figure 6c**) slightly reduced the degree of both correlated and anti-correlated motions that exist in the wild type structure, although no significant correlation has been seen in the DCCM analysis. In case of V126G mutant structure (**Figure 6d**), the result showed that the mutation increased the correlative intra-residue movement, specifically between the loop 34 to 40 (residues, 54 to 60) and c-terminal

portion 80 to 105 (residues, 100 to 125). The negative correlation was also seen in the CDR1 and CDR2 loop regions, indicating that the deleterious effect in TREM2 functions.

To further support this explanation, principal component analysis (PCA) was conducted based on C_{α} atoms (**Figure S1**). In order to get insights about the dynamic mechanical properties of the simulated system, PCA analysis is usually performed²⁶, by reducing the complexity of classifying collective motions. The dynamics of two different proteins is best achieved via characterization of its phase space behavior, which is directly associated with the protein stability and protein function^{40,41}. At this point, the overall combined motion of the C_{α} atoms of protein structures were described by the eigenvectors of the covariance matrix, which is correlated with the coincident eigenvalues. As resulted in the PCA analysis, the simulated systems including, wild type, Y38C, T66M and V126G showed 19.81%, 33.43%, 20.94% and 29.67% of variance in the first three PCAs, respectively.

The motions of different proteins can be visualized through the projection of trajectories in phase space of the first two principal components (PC1, PC2), where mutants show different patterns of conformational spaces, denoting the significant changes in protein conformation. On these projections, the Y38C showed uniform conformational distribution in a larger phase space, describing the higher flexibility in contrast to the native (**Figure S1b**). Furthermore, T66M structure also showed conformational transitions in the simulation, representing the less flexibility over the time (**Figure S1c**). On the contrary, V126G represented higher periodic jumps, indicating greater fluctuations compared to the native, as it occupied a larger phase space (**Figure S1d**).

The variations in atomic movement were represented by all first PCAs (**Figure 7**), as PC1 reveal the most dynamic motion of all systems. Here the displacement of atoms is denoted by wide tubes, whereas the narrow tubes marked the regions that stayed rigid during the simulation. The structural deviations among the wild type and mutations were clearly depicted in the first PCAs (**Figure 7**), where the contribution of each residue to the first principal component is represented in the plot which is rendered in the bottom panel (**Figure 7e**) describing the local mobility during the simulation. As can be seen, Y38C induced flexibility in several loops of TREM2 including CDR1, CDR2 and CDR3 regions (**Figure 7b**), while CDR2 was greatly fluctuated in T66M mutant structure (**Figure 7c**). The conformational changes in the wild type structure were reflected in the loop region, residues 47 to 60; this region also deviated in V126G

mutant structure (**Figure 7d**). The PCA analysis supports the results described by the RMSF and DCCM analysis, which ultimately concluded that the mutants have higher structural flexibility than the wild type demonstrating local deleterious effects. Since Y38C and T66M TREM2 produced large flexibility in CDR regions, these mutations may bring different structural orientations in these regions. The following section further addresses on this issue.

Mutation-induced conformational changes

Although MD simulations well explained the conformational changes in the protein structure, the changes in the secondary structures caused by the mutations are still needed to investigate. Therefore, the variations in the secondary structural elements for both wild and mutant proteins during the simulation were characterized by the simulation interaction diagram of Schrödinger 2017-1 (LLC, New York, NY, USA)⁴². During the simulation, the appearance of secondary structural elements determines the structural flexibility of a protein molecule, where conformations such as α -helix and β -sheet are naturally tend to be more rigid than the coils and turn conformations²⁶. The total percentage of secondary structural elements in total simulation time is shown in **Figure S2**. Interestingly, the mutants increased the total secondary structures of TREM2 protein, when compared to wild type. Moreover, the CDR1 (residue index 17-27) and CDR2 (residue index 46-57) regions were gained short α -helix conformation in the mutant structure, where the percentage of helix formation is higher in the Y38C structure (**Table S1**). On the contrary, V126G increased total secondary structure elements in the TREM2, where the increase of α -helix conformation in CDR2 region was higher in comparison to other mutants. Previous studies with Alzheimer's disease associated mutant showed that substitution of ARG⁴⁷ by His facilitates CDR2 loop remodelling to short α -helix which contributes to loss of ligand binding²⁴. Together with these data, our study also suggested that mutants in this study also facilitate not only conformation remodelling of CDR2 loop but also in CDR1 regions.

Mechanistic insights into conformational remodelling

It is inferred that the intra-residue hydrogen bonding influences the secondary structure in biological macromolecule⁴³. Thus hydrogen bond occupancy analysis within CDR1 and CDR2 loops in the wild as well as in mutant types was further considered, as both CDR1 and CDR2 loops undergoes conformational remodelling (**Figure S2**). As shown in the **Figure 8**, mutations

induced diverse intra and inter loop hydrogen bond interactions. Furthermore, these loops were identified as the flexible regions from RMSF and PCA studies, indicating that increasing flexibility by mutations render residues to make different interaction network. According to the RMSF study, the substitution of tyrosine with cystine in 38th position was seen to induce high flexibility next to D39, which formed significantly elevated hydrogen bonding interactions within CDR1 loop, especially with M41 and H43 (**Figure 8a**). The hydrogen bonding occupancy between D39 and H43 in Y38C structure was maintained to 12.34% during simulation, while it was 3.91% in wild type. Furthermore, D39 maintained 6.6% H-bond occupancy with M41 residue; however, it was missing in wild type structure. This observation indicates that maintaining intra-loop hydrogen bond with M41 and H43 by D39 may facilitate short helix formation of CDR1 loop in Y38C mutant. In case of CDR2 loop, Y38C mutant reduced hydrogen bonding between S73 and W70, L72 and S75, while it was increased in the wild type. Besides, Y38C maintained two new hydrogen bonds, including T66 and S73, S73 and L69, which were lacking in wild type structure (**Figure 8b**).

Previous studies showed that the conformational change of CDR2 loop is facilitated by the π - π stacking interactions between the imidazole moieties of H47 and H67, which caused to swing H67 approximately by 180°²⁴. Therefore, inter-loop interaction between CDR1 and CDR2 was investigated and found that Y38C increased the hydrogen bonding between H47 and H67 (**Figure 8c**). This observation indicates the conformational changes in CDR1 loop. Similarly, T66M mutant showed the significant local fluctuations in S65 and H67 residues, compared to wild type; H67 maintained H-bonding interaction with L72 more than 20%. Interestingly, mutation in T66 increased the hydrogen bonding capacity with R47 more than the wild structure. However, an unusual hydrogen bond interaction between R47 and D65 by replacing M66 was seen, which may contribute to CDR1 conformational change. However, the interaction between M66 and R47 was more dominant, as a result the degree of helix formation in CDR1 loop was seen to decrease subsequently in the simulation (**Figure S2c**). On the other hand, V126G mutation reduced the hydrogen bonding between R47 and T66, while induced interaction between R47 and N68, which might allows the conformational changes in CDR1 loop. In addition, conformational changes in CDR2 loop could be induced by numerous hydrogen interactions of H67 with M41, H43 and T66.

In order to validate the findings from H-bond analysis, additional free energy landscape analysis was introduced to select most stable conformer from the simulation ensemble. The FEL is rendered in **Figure 8d**, where different colors represent the corresponding structures with different energies. The conformation having the lower energy states can be found in the blue area, are usually more stable than the other conformations (red area) generated during the simulation

44.

The thermodynamic stability of a protein is represented by the depth of an energy minimum, where the kinetic stability of a protein is indicated by the height of barriers separated by different energy minima. This depth of energy also suggests the change of conformation from one to another, where the width of an energy minimum associated with the width of the conformational ensemble within the energy well ⁴⁵. As represented in the **Figure 8d**, the region of the conformational space corresponding to the basin was seen to change in both wild and mutant types, indicating that the overall conformational stability of the protein was affected by the mutations. In order to understand the disordered state, the conformation in the favorable energy minima has been visualized and analyzed the intra-residue hydrogen bonding network through residue interaction network analysis. The lowest minima in wild type trajectory was found with corresponding Rg and RMSD to average value of 1.33 and 1.32 nm, respectively, containing no conformational changes in CDR1 and CDR2 loops. In detailed structural view, T66 was found to interact with R47 by forming double hydrogen bonds. Comparatively, a short helix was found in the CDR2 loop in the most stable conformer of mutant trajectories. In Y38C mutant structure, M41 was seen to be maintained H-bond with D39. Furthermore, R47 formed hydrogen bonding interactions with both H67 and L72 in Y38C, while in case of T66M; this residue maintained Pi-Pi stacking interaction with H67. On the other hand, V126G structure showed an unusual hydrogen bonding between 66 and 67 residues as well as between R47 and S65. Although no conformational change in the CDR1 loop was visualized in the structure, the interaction between D39 and M43 was also seen within CDR1 loop. Most importantly, all mutants, especially Y38C and T66M, produced steric clashes with neighbouring side chains, while mutation with glycine at 126 position induced side chain steric clash within Y38 residue, which induced flexibility in both CDR1 and CDR2 loops thus promoting conformational remodelling.

Conclusion

The effect of genetic variants in the loss of TREM2 function is crucial to understand its involvement with late-onset Alzheimer's disease (AD). Using molecular dynamics simulation, this study represented novel findings on the deleterious role of NHD variants, which promotes structural alteration in TREM2 ectodomain. All three mutations (Y38C, T66M and V126G) induced structural remodelling in both CDR1 and CDR2 loop by inducing steric clash, increased flexibility and altered hydrogen bonding patterns. This study supports previous findings, and provides additional insight on the mechanism behind the loss of ligand binding, which is critical to our understanding of the role of TREM2 in neurodegenerative diseases.

Methods

Preparation of simulation system

The three dimensional crystal structure of TREM2 ectodomain was retrieved from the protein databank (<http://www.rcsb.org/pdb>)⁴⁶, with PDB id of 5ELI⁴⁷. The structure was initially prepared by adding bond orders, hydrogen and charges; and also refined by removing water molecules and optimizing it at neutral pH. The structure was further fixed by correcting amide groups of asparagines, some thiol and hydroxyl groups, protonation states of glutamic acids, aspartic acids and histidines. In order to adjust the heavy atom Root Mean Square Deviation (RMSD), minimization was applied by using Optimized Potentials for Liquid Simulation (OLPS3) force field until it reached to 0.30 Å. The mutant structures, consisting of Y38C, T66M, and V126G, were constructed by computational mutagenesis methodology, using the Mutate Residues script from Schrödinger suite 2017-1 (LLC, New York, NY, USA)⁴⁸.

Additional short molecular dynamics refinement simulation was performed for better resolute quality of the structure. Thus applying YAMBER3 force field⁴⁹, molecular dynamics simulation was conducted for 500 ps at pH 7.4 and 298K temperature within the solvent density of 0.997. The simulation was run by the YASARA software with a default md_refine macro, and the final structure was selected based on the lowest energy.

Molecular dynamics simulation

Molecular dynamics simulation approach was used to study the changes in dynamics behavior of protein due to mutations, by the implementing Desmond module of Schrödinger suite 2017-1 (LLC, New York, NY, USA) ^{50,51}. Here, OLPS3 all-atom force field was utilized to illustrate the molecular behavior ⁵²⁻⁵⁴. The prepared structures including wild type, Y38C, T66M, and V126G were solvated in the presence of explicit solvent in a triclinic periodic boundary box. Each systems was submerged to a Monte-Carlo equilibrated TIP3P solvation model extending to approximately 10Å in each direction, which was known to give the best experimental output ⁵⁵. For the neutralization of the solvation system, additional counter ions were added to the water model. The physiological condition as well as the ionic strength of the solvent system was also maintained by placing the default salt (NaCl) concentration of 0.15 M to the simulation box. The default relaxation protocol ^{56,57} was used to relax the system, which contains eight stages. Followed by Brownian dynamics, the second stage simulation was started at the temperature of 10K in NVT ensemble for 12 ps and restraints on solute heavy atoms. Afterward, 12 ps simulation was done in same ensemble at temperature of 10K and restraints on solute heavy atoms in third stage. In the same temperature, the fourth stage was begun by running 12 ps simulation under the NPT ensemble at 1 bar pressure by keeping restraints on the solute heavy atoms. The protein cavity in the stage five solvated by using the solvate pocket script. Followed by 12 ps simulation in NPT ensemble in stage six with restraints on the solute heavy atoms, 24ps simulation was run in the stage seven with no restraints on the solute heavy atoms. Both stages were performed at the 300 K in 1 bar pressure. Finally 100 ns MD simulation was done for each system with no constraints applied and resulted trajectories were subjected for further analysis. Throughout the simulation, the integration of motion, RESPA ⁵⁸ integrator was employed with 2 fs as inner time step with and M-SHAKE ⁵⁹ algorithm to constrain all covalent bonds connecting hydrogen atoms. Particle mesh Ewald method was applied for calculating the long-range electrostatic interaction [22], while 9.0 Å was used for short-range electrostatic contacts; uniform density approximation was select for the cutoff of the long-range van der Waals (VDW) interaction cutoff. The conditions in the simulation were maintained by Nose–Hoover thermostats ⁶⁰ at 300 K constant temperature and 1 atmosphere constant pressure using the Martyna–Tobias–Klein method ⁶¹.

The resultant trajectories were used to evaluate the conformational change and protein stability by means of Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation

(RMSF) and SSE (Secondary Structure Elements) by using Simulation Interactions Diagram panel of Schrödinger 2017-1 (LLC, New York, NY, USA). Whereas, other analysis including, radius of gyration (Rg). Solvent Accessible Surface Area (SASA) and hydrogen bond occupancy were calculated by VMD software⁶². The dynamic cross-correlation maps were generated from trajectories to explain the time-correlated motions in the protein, which were assembled by Bio3D [41] software through R programming. The time-correlated information between protein atoms *i* and *j* (*c_{ij}*) is represented as a matrix in DCCM, which was obtained with the following expression:

$$DCCM_{i,j} = \frac{\langle \vec{d}_i \cdot \vec{d}_j \rangle}{\sqrt{\langle d_i^2 | d_j^2 \rangle}}$$

The displacement between the current and average position of *i*th and *j*th atoms is represented by *d*, and the mean time over the all trajectories is represented by the angle brackets. The calculated values in DCCM ranged between the −1 to +1, denoting negative and positive correlation, respectively. Principle Component Analysis (PCA) was further included in the study to describe the collective motions of TREM2 variants³⁹. The eigenvectors in PCA were calculated by the superimposing atomic coordinates to the reference structure without translational and rotational movements. Each eigenvector displays the mean squared displacements (MSD) of atoms along the corresponding eigenvector, which was associated with an eigenvalue. The mathematical details have been described previously^{63,64}.

Free energy landscape (FEL)

Free energy landscape is used to map the all possible conformational changes in macromolecules through highlighting their corresponding energy levels, which required to probe the spatial position of interacting molecules in a system^{65,66}. In free energy landscape analysis, Gibb's free energy is calculated as a function of protein enthalpy and entropy, which describes the protein stability. It also highlights the various conformational states related to protein's structure-function correlation. In this study, the energy bases of conformational diversity in various TREM2 structures were investigated by the free energy landscapes, using following equation:

$$G_i = -k_B T \ln \left(\frac{N_i}{N_{\max}} \right)$$

Where, Boltzmann's constant is represented by k_B , T defined the temperature, which is set to 300 K. N_i is the population of bin i and $N_{\max}$ is the population of the most populated bin. An artificial barrier scale is set to the bin with no population, as the lowest provability. Color-code modes are used to display different energy levels.

Residue Interaction Network

The most stable three dimensional coordinates of wild and mutant types were transferred to RING server ⁶⁷, in order to analyze the Residue Interaction Network (RIN), which represents intraresidue interactions in a exhaustive network view. In the network model, a protein residue is represented as the node, while the mode of interactions is described as the edge. The result from RING server is then subjected to Cytoscape for constructing the interactive RIN 3.2.1 ⁶⁸, using the plug-in RINalyzer. The types of interactions are described by the dashed or dotted edges, including the salt bridge, hydrogen bond van der Waal interactions.

References

- 1 Selkoe, D. J. & Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO molecular medicine* **8**, 595-608 (2016).
- 2 Jack Jr, C. R. *et al.* Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *The Lancet Neurology* **12**, 207-216 (2013).
- 3 Guerreiro, R. *et al.* TREM2 variants in Alzheimer's disease. *New England Journal of Medicine* **368**, 117-127 (2013).
- 4 Jonsson, T. *et al.* Variant of TREM2 associated with the risk of Alzheimer's disease. *New England Journal of Medicine* **368**, 107-116 (2013).
- 5 Atagi, Y. *et al.* Apolipoprotein E is a ligand for triggering receptor expressed on myeloid cells 2 (TREM2). *Journal of Biological Chemistry* **290**, 26043-26050 (2015).
- 6 Bailey, C. C., DeVaux, L. B. & Farzan, M. The triggering receptor expressed on myeloid cells 2 binds apolipoprotein E. *Journal of Biological Chemistry*, jbc. M115. 677286 (2015).

- 7 Yeh, F. L., Wang, Y., Tom, I., Gonzalez, L. C. & Sheng, M. TREM2 binds to apolipoproteins, including APOE and CLU/APOJ, and thereby facilitates uptake of amyloid-beta by microglia. *Neuron* **91**, 328-340 (2016).
- 8 Daws, M. R. *et al.* Pattern recognition by TREM-2: binding of anionic ligands. *The Journal of Immunology* **171**, 594-599 (2003).
- 9 Cannon, J. P., O'Driscoll, M. & Litman, G. W. Specific lipid recognition is a general feature of CD300 and TREM molecules. *Immunogenetics* **64**, 39-47 (2012).
- 10 Poliani, P. L. *et al.* TREM2 sustains microglial expansion during aging and response to demyelination. *The Journal of clinical investigation* **125**, 2161-2170 (2015).
- 11 Peng, Q. *et al.* TREM2-and DAP12-dependent activation of PI3K requires DAP10 and is inhibited by SHIP1. *Science signaling* **3**, ra38-ra38 (2010).
- 12 Zhao, Y. *et al.* TREM2 Is a Receptor for beta-Amyloid that Mediates Microglial Function. *Neuron* **97**, 1023-1031.e1027, doi:10.1016/j.neuron.2018.01.031 (2018).
- 13 Cheng, J. *et al.* TREMs in Alzheimer's disease: Genetic and clinical investigations. *Clinica Chimica Acta* **463**, 88-95 (2016).
- 14 Jiang, T. *et al.* Upregulation of TREM2 ameliorates neuropathology and rescues spatial cognitive impairment in a transgenic mouse model of Alzheimer's disease. *Neuropsychopharmacology* **39**, 2949 (2014).
- 15 Jiang, T. *et al.* TREM2 overexpression has no improvement on neuropathology and cognitive impairment in aging APPswe/PS1dE9 mice. *Molecular neurobiology* **54**, 855-865 (2017).
- 16 Wang, Y. *et al.* TREM2 lipid sensing sustains the microglial response in an Alzheimer's disease model. *Cell* **160**, 1061-1071 (2015).
- 17 Wang, Y. *et al.* TREM2-mediated early microglial response limits diffusion and toxicity of amyloid plaques. *Journal of Experimental Medicine*, jem. 20151948 (2016).
- 18 Yuan, P. *et al.* TREM2 haploinsufficiency in mice and humans impairs the microglia barrier function leading to decreased amyloid compaction and severe axonal dystrophy. *Neuron* **90**, 724-739 (2016).
- 19 Keren-Shaul, H. *et al.* A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* **169**, 1276-1290. e1217 (2017).

- 20 Guerreiro, R. *et al.* A novel compound heterozygous mutation in TREM2 found in a Turkish frontotemporal dementia-like family. *Neurobiology of aging* **34**, 2890. e2891-2890. e2895 (2013).
- 21 Guerreiro, R. J. *et al.* Using exome sequencing to reveal mutations in TREM2 presenting as a frontotemporal dementia-like syndrome without bone involvement. *JAMA neurology* **70**, 78-84 (2013).
- 22 Le Ber, I. *et al.* Homozygous TREM2 mutation in a family with atypical frontotemporal dementia. *Neurobiology of aging* **35**, 2419. e2423-2419. e2425 (2014).
- 23 Kober, D. L. *et al.* Neurodegenerative disease mutations in TREM2 reveal a functional surface and distinct loss-of-function mechanisms. *eLife* **5**, e20391, doi:10.7554/eLife.20391 (2016).
- 24 Sudom, A. *et al.* Molecular basis for the loss-of-function effects of the Alzheimer's disease-associated R47H variant of the immune receptor TREM2. *Journal of Biological Chemistry* **293**, 12634-12646 (2018).
- 25 Elber, R. Perspective: Computer simulations of long time dynamics. *The Journal of chemical physics* **144**, 060901 (2016).
- 26 Hosen, S. Z., Dash, R., Junaid, M., Mitra, S. & Absar, N. Identification and structural characterization of deleterious non-synonymous single nucleotide polymorphisms in the human SKP2 gene. *Computational biology and chemistry* **79**, 127-136 (2019).
- 27 Schreiner, W., Karch, R., Knapp, B. & Ilieva, N. Relaxation estimation of RMSD in molecular dynamics immunosimulations. *Computational and Mathematical Methods in Medicine* **2012** (2012).
- 28 Priya Doss, C. G., Chakraborty, C., Chen, L. & Zhu, H. Integrating in silico prediction methods, molecular docking, and molecular dynamics simulation to predict the impact of ALK missense mutations in structural perspective. *BioMed research international* **2014** (2014).
- 29 Tanwar, H. *et al.* Chapter Five-A Computational Approach to Identify the Biophysical and Structural Aspects of Methylenetetrahydrofolate Reductase (MTHFR) Mutations (A222V, E429A, and R594Q) Leading to Schizophrenia. *Advances in Protein Chemistry and Structural Biology* **108**, 105-125 (2017).

- 30 Eisenberg, D. & McLachlan, A. D. Solvation energy in protein folding and binding. *Nature* **319**, 199 (1986).
- 31 Anwar, M. A. & Choi, S. Structure-activity relationship in TLR4 mutations: atomistic molecular dynamics simulations and residue interaction network analysis. *Scientific reports* **7**, 43807 (2017).
- 32 Eaton, W. A., Henry, E. R. & Hofrichter, J. Application of linear free energy relations to protein conformational changes: the quaternary structural change of hemoglobin. *Proceedings of the National Academy of Sciences* **88**, 4472-4475 (1991).
- 33 Ikeguchi, M., Ueno, J., Sato, M. & Kidera, A. Protein structural change upon ligand binding: linear response theory. *Physical review letters* **94**, 078102 (2005).
- 34 Levy, Y. & Onuchic, J. N. Mechanisms of protein assembly: lessons from minimalist models. *Accounts of chemical research* **39**, 135-142 (2006).
- 35 Lee, Y. *et al.* Dissecting the critical factors for thermodynamic stability of modular proteins using molecular modeling approach. *PloS one* **9**, e98243 (2014).
- 36 Benson, N. C. & Daggett, V. A comparison of multiscale methods for the analysis of molecular dynamics simulations. *The Journal of Physical Chemistry B* **116**, 8722-8731 (2012).
- 37 Agrahari, A. & Doss, C. G. P. Impact of I30T and I30M substitution in MPZ gene associated with Dejerine–Sottas syndrome type B (DSSB): A molecular modeling and dynamics. *Journal of theoretical biology* **382**, 23-33 (2015).
- 38 Gaieb, Z. & Morikis, D. Detection of Side Chain Rearrangements Mediating the Motions of Transmembrane Helices in Molecular Dynamics Simulations of G Protein-Coupled Receptors. *Computational and structural biotechnology journal* **15**, 131-137 (2017).
- 39 Ichiye, T. & Karplus, M. Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins: Structure, Function, and Bioinformatics* **11**, 205-217 (1991).
- 40 Amadei, A., Linssen, A. B. & Berendsen, H. J. Essential dynamics of proteins. *Proteins: Structure, Function, and Bioinformatics* **17**, 412-425 (1993).
- 41 Agrahari, A. K., George, P. D. C., Siva, R., Magesh, R. & Zayed, H. Molecular insights of the G2019S substitution in LRRK2 kinase domain associated with Parkinson's disease: A molecular dynamics simulation approach. *Journal of theoretical biology* (2019).

- 42 Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637 (1983).
- 43 Bai, Y. & Englander, S. W. Hydrogen bond strength and beta-sheet propensities: the role of a side chain blocking effect. *Proteins* **18**, 262-266, doi:10.1002/prot.340180307 (1994).
- 44 Guan, S. *et al.* Exploration of binding and inhibition mechanism of a small molecule inhibitor of influenza virus H1N1 hemagglutinin by molecular dynamics simulation. *Scientific reports* **7**, 3786 (2017).
- 45 Gershenson, A., Gierasch, L. M., Pastore, A. & Radford, S. E. Energy landscapes of functional proteins are inherently risky. *Nature chemical biology* **10**, 884 (2014).
- 46 Zardecki, C., Dutta, S., Goodsell, D. S., Voigt, M. & Burley, S. K. (ACS Publications, 2016).
- 47 Kober, D. L. *et al.* Neurodegenerative disease mutations in TREM2 reveal a functional surface and distinct loss-of-function mechanisms. *eLife* **5**, doi:10.7554/eLife.20391 (2016).
- 48 Raju, D. *et al.* Molecular Insight and Binding Pattern Analysis of Shikonin as a Potential VEGFR-2 Inhibitor. *Current Enzyme Inhibition* **13**, 1-10, doi:<http://dx.doi.org/10.2174/1573408013666161227162452> (2017).
- 49 Krieger, E., Darden, T., Nabuurs, S. B., Finkelstein, A. & Vriend, G. Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins: Structure, Function, and Bioinformatics* **57**, 678-683 (2004).
- 50 Tools, M. (Version, 2011).
- 51 Fazil, M. H. U. T. *et al.* Comparative structural analysis of two proteins belonging to quorum sensing system in *Vibrio cholerae*. *Journal of Biomolecular Structure and Dynamics* **30**, 574-584 (2012).
- 52 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics* **79**, 926-935 (1983).
- 53 Bowers, K. J. *et al.* in *Proceedings of the 2006 ACM/IEEE conference on Supercomputing*. 84 (ACM).
- 54 Shaw, D. E. A fast, scalable method for the parallel evaluation of distance-limited pairwise particle interactions. *Journal of computational chemistry* **26**, 1318-1328 (2005).

- 55 Shaik, M. S., Liem, S. Y. & Popelier, P. L. Properties of liquid water from a systematic refinement of a high-rank multipolar electrostatic potential. *The Journal of chemical physics* **132**, 174504 (2010).
- 56 Mitra, S. & Dash, R. Structural dynamics and quantum mechanical aspects of shikonin derivatives as CREBBP bromodomain inhibitors. *Journal of Molecular Graphics and Modelling* **83**, 42-52 (2018).
- 57 Arifuzzaman, M. *et al.* A Computational workflow for the identification of the potent inhibitor of type II secretion system traffic ATPase of *Pseudomonas aeruginosa*. *Computational biology and chemistry* **76**, 191-201 (2018).
- 58 Humphreys, D. D., Friesner, R. A. & Berne, B. J. A multiple-time-step molecular dynamics algorithm for macromolecules. *The Journal of Physical Chemistry* **98**, 6885-6892 (1994).
- 59 Kräutler, V., Van Gunsteren, W. F. & Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *Journal of computational chemistry* **22**, 501-508 (2001).
- 60 Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Physical review A* **31**, 1695 (1985).
- 61 Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **101**, 4177-4189 (1994).
- 62 Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* **14**, 33-38 (1996).
- 63 Shlens, J. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100* (2014).
- 64 Salmas, R. E., Yurtsever, M. & Durdagi, S. Investigation of inhibition mechanism of chemokine receptor CCR5 by micro-second molecular dynamics simulations. *Scientific reports* **5** (2015).
- 65 Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. The energy landscapes and motions of proteins. *Science* **254**, 1598-1603 (1991).
- 66 Lange, O. F. & Grubmüller, H. Generalized correlation for biomolecular dynamics. *Proteins: Structure, Function, and Bioinformatics* **62**, 1053-1061 (2006).

- 67 Martin, A. J. *et al.* RING: networking interacting residues, evolutionary information and energetics in protein structures. *Bioinformatics* **27**, 2003-2005 (2011).
- 68 Baryshnikova, A. Exploratory analysis of biological networks through visualization, clustering, and functional annotation in cytoscape. *Cold Spring Harbor Protocols* **2016**, pdb. prot077644 (2016).

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

ISM conceived the idea and designed the experimental work. RD and HJC carried out the experiments. ISM, RD and HJC analyze the data and wrote the manuscript.

Funding

The study was supported by the Basic Science Research Program through the national research foundation of Korea (NRF-2018R1A2B6002232 to ISM)

Figure Legends

Figure 1. Wild and mutant structures of TREM2 ectodomain. Cartoon depiction of TREM2 wild type ectodomain along with the indication of domain boundaries (a). Three dimensional view of NHD-associated mutated positions, Y38C (b), T66M (c) and V126G (d). Original residue is colored as green, while mutated as yellow.

Figure 2. Root-mean-square deviation (RMSD) of the C_{α} atoms for the wild-type and mutant TREM2 structure at 100 ns. Here, dark green line represents the wild-type, while the blue, orange and violate lines stand for the Y38C, T66M and V126G mutants of RMSDs. In addition, bottom panel of the figure is represented with the density plots for both wild and mutant structures, illustrating the distribution of sampled conformations during simulation.

Figure 3. The radius of gyration for the Y38C (a), T66M (b), and V126G (c) TREM2 mutant in comparison to wild type structure. Here, dark green line represents the wild-type, while the blue, orange and violate lines stand for the Y38C, T66M and V126G mutant structure.

Figure 4. The change of total solvent accessible surface area (SASA) for the Y38C (a), T66M (b), and V126G (c) TREM2 mutant in comparison to wild type structures. Here, dark green line represents the wild-type, while the blue, orange and violate lines stand for the Y38C, T66M and V126G mutant structure.

Figure 5. Investigation of local structural effect of mutations. a) Root mean square fluctuation of the C_{α} atoms for the wild type and mutant types of TREM2. a) The SASA values, that computed on a per residue basis for the wild type and mutant types of TREM2. In all plots, residues index is level with colored bar, representing the structural propensity. In each graph, dark green line represents the wild-type, while the blue, orange and violate lines stand for the Y38C, T66M and V126G mutant structure.

Figure 6. Calculated dynamic cross-correlation matrix of the C_{α} atoms around their mean positions for 100 ns molecular dynamics simulations. The extent of correlated motion and anti-correlated motions are color-coded, ranging from red to blue. Here red color denoted positive correlation, while negative described by blue color. Wild type (a), Y38C (b), T66M (c) and V126G (d).

Figure 7. Representation of the atomic displacements in PCA1, calculated from the simulation of wild type (a), Y38C (b), T66M (c) and V126G (d), which indicates significant alterations in the movements. The green marked circles of each panels describes the regions of greater flexibilities. Residue-wise loadings for PC1, where dark green line represents the wild-type, while the blue, orange and violate lines stand for the Y38C, T66M and V126G mutant structure (e).

Figure 8. Hydrogen bond occupancy analysis for within the CDR1 (a), CDR2 (b), between the CDR1 and CDR2 loop (c). Free energy contour map depicted based on the Radius of gyration and RMSD to average (d), where deeper dark blue color area in the maps indicates lower energy. In the figure each graph is represented with the three dimensional structure having the lowest energy, along with residue interaction network. Each node in network describes the residue, while edge represented as the interaction. The color of edge signifying the mode of interaction, including green for hydrogen bond, orange of van der walls contact, cyan color for salt bridge, pi-pi stacking as blue and pink for disulfide bridge.

Supplementary data

Table S1. The summery of secondary structure element throughout the protein structure in different simulation systems.

Figure S1. Principle Component Analysis (PCA) regarding the TREM2 protein in four different systems, wild type (a), Y38C (b), T66M (c) and V126G (d). Each panel represents the two-dimensional plots between eigenvectors (EV) 1, 2, and 3 in three cross profiles, PC1-PC2, PC1-PC3 and PC2-PC3 for all systems. Throughout the x and y axes, each dot denotes the one conformation of the complex. The spread of blue and red color dots described the degree of conformational changes in the simulation, where color scale from blue to white to red is equivalent to simulation time. The blue indicates initial timestep, white is intermediate and final timestep is represented by red color.

Figure S2. Residue-specific average secondary structural preferences for wild type (a) Y38C (b), T66M (c) and V126G (d) structure, during 100 ns simulation. Here, dark blue color stand for Beta sheets and red color for alpha helix conformation.

Figure 1. Wild and mutant structures of TREM2 ectodomain.

Figure 2. Root-mean-square deviation (RMSD) of the C α atoms for the wild-type and mutant TREM2 structure at 100 ns.

Figure 3. The radius of gyration for the Y38C (a), T66M (b), and V126G (c) TREM2 mutant in comparison to wild type structure.

Figure 4. The change of total solvent accessible surface area (SASA) for the Y38C (a), T66M (b), and V126G (c) TREM2 mutant in comparison to wild type structures.

Wild Y38C T66M V126G

Figure 5. Investigation of local structural effect of mutations.

Figure 6. Calculated Dynamic Cross-correlation Matrix (DCCM) of the C α atoms around their mean positions for 100 ns molecular dynamics simulations.

Figure 7. Representation of the atomic displacements in PCA1, calculated from the simulation of wild type (a), Y38C (b), T66M (c), and V126G (d), which indicates significant alterations in the movements.

Figure 8. Hydrogen bond occupancy and free energy landscape analysis