Jurnal Teknologi

Full Paper

M630L MUTATION DISRUPTS THE STRUCTURE CONFORMATION OF BRUTON TYROSINE KINASE (BTK) DOMAIN IN PATIENT WITH X-LINKED AGAMMAGLOBULINEMIA: INSIGHTS FROM IN SILICO

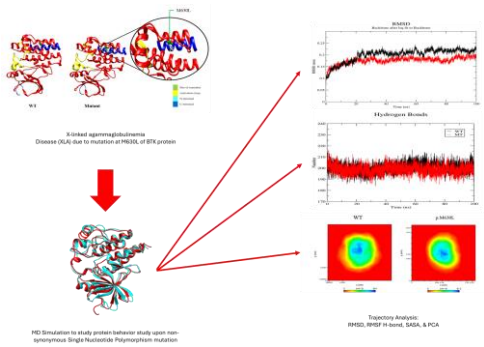
Muhammad Syahrul Nizam Mohd Ruzani[a], Siti Fatimah Mohd Tahab[b], Adiratna Mat Ripen[c], Saharuddin Mohamad[a]*, Azzmer Azzar Abdul Hamid[d]*, Che Muhammad Khairul Hisyam Ismail[d], Amir Feisal Merican[a]

[a]Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia
[b]Allergy and Immunology Research Centre, Institute for Medical Research, Jalan Pahang, 50588 Kuala Lumpur, Malaysia
[c]Centre of Research in Systems Biology, Structural Bioinformatics and Human Digital Imaging (CRYSTAL), University of Malaya, 50603 Kuala Lumpur, Malaysia
[d]Research Unit for Bioinformatics and Computational Biology (RUBIC), Kulliyyah of Science, International Islamic University Malaysia, Bandar Indera Mahkota, 25200 Kuantan, Pahang, Malaysia

Article history
Received
25 March 2024
Received in revised form
18 July 2024
Accepted
7 August 2024
Published Online
22 December 2024

*Corresponding author
saharuddin@um.edu.my

Graphical abstract



Abstract

X-Linked Agammaglobulinemia (XLA) is a rare inherited disease, attributed to mutations found in the Bruton Tyrosine Kinase (BTK) gene. This research outlines the application of Molecular Modelling and Simulation to predict the effects of a novel non-synonymous Single Nucleotide Polymorphism (nsSNP) reported in the BTK kinase domain protein of a male XLA patient. The mutation at position c.1888A>T causes a substitution of p.M630L within the kinase domain of the BTK protein. Functional assessment using in silico prediction tools (SIFT, Polyphen-2, PROVEAN, MutationAssessor, PANTHER, FATHMM, MutPred and MutationTaster) predicted the mutation to be deleterious, potentially disrupting both the structure and function of the protein. Nevertheless, the mutation is not located at the active site of the kinase domain. Consequently, the molecular dynamics (MD) simulations were performed to investigate the impact of amino acid substitution to the three-dimensional (3D) conformational structure of the BTK protein, contributing to the disease phenotype observed in the reported patient. MD analysis, clustering analysis and Principle Component Analysis (PCA) revealed that the 3D conformational structure of the kinase domain of mutant protein to be more compact and rigid compared to the wildtype. Given that the alterations in protein structure can influence their functional characteristics, the p.M630L mutation within the BTK protein kinase domain might disrupt the protein function, potentially impeding the maturation of B cells and contributing to the onset of XLA disease.

Keywords: X-linked Agammaglobulinemia (XLA), Bruton Tyrosine Kinase (BTK), Non-synonymous Single Nucleotide Polymorphism (nsSNP), Molecular Dynamics (MD) simulation, Cluster analysis, Principle Component Analysis (PCA)

Abstrak

Agamaglobulinemia Berpaut-X (XLA), sebuah penyakit warisan yang jarang ditemui disebabkan oleh mutasi dalam gen Kinase Tirosina Bruton (BTK). Kajian ini menggariskan aplikasi Pemodelan dan Simulasi Dinamik Molekul (MD) untuk meramalkan kesan Polimorfisma Nukleotida Tunggal Tidak Bersinonim (nsSNP) yang baru dilaporkan dalam protein domain kinase BTK, seorang pesakit lelaki XLA. Mutasi pada kedudukan c.1888A>T menyebabkan penggantian p.M630L berlaku dalam protein domain kinase BTK. Penilaian fungsi protein menggunakan alat jangkaan *in silico* (SIFT, Polyphen-2, PROVEAN, MutationAssessor, PANTHER, FATHMM, MutPred dan MutationTaster) meramalkan mutasi tersebut berpotensi merosakkan dan mungkin mengganggu struktur dan fungsi protein tersebut. Walaubagaimanapun, mutasi tersebut tidak terletak di tapak aktif domain kinase. Oleh itu, simulasi dinamik molekul dilakukan untuk menyiasat impak penggantian asid amino ke atas struktur konformasi tiga dimensi (3D) protein BTK, akhirnya menyumbang kepada fenotip penyakit yang diperhatikan pada pesakit yang dilaporkan. Analisis MD, analisis pengklusteran dan Analisis Komponen Utama (PCA) mendedahkan struktur konformasi 3D protein domain kinase mutan menjadi lebih padat dan kaku berbanding dengan jenis asli. Mengambil kira bahawa perubahan dalam struktur protein boleh mempengaruhi ciri-ciri fungsinya, ini mencadangkan bahawa mutasi p.M630L dalam protein domain kinase BTK mungkin mengganggu fungsi protein BTK, menghalang pematangan sel B dan menyumbang kepada permulaan penyakit XLA.

Kata kunci: Agamaglobulinemia Berpaut-X (XLA), Kinase tirosina Bruton (BTK), Polimorfisma Nukleotida Tunggal Tidak Bersinonim (nsSNP), Simulasi Dinamik Molekul (MD), Analisis Kluster; Analisis Komponen Utama (PCA)

1.0 INTRODUCTION

X-linked agammaglobulinemia (XLA) results from a mutation in the Bruton tyrosine kinase (BTK) gene, leading to defective development and maturation of B cell in the bone marrow. This condition is exclusively observed in male infants, at the ratio of 1:200,000 [1, 2]. Mutation in the BTK gene suppresses the production of BTK protein and affects the development of B lymphocytes (also known as B cells) causing failure to produce mature B cells [3, 4]. Previous reports indicate that individuals with XLA typically exhibit reduced antibody levels, despite having normal quantities of pre-B cells. This occurrence is attributed to the dysfunction of the BTK gene [3, 4]. Such a condition contributes to the recurrence of pyogenic bacterial infections, including pneumonia, bronchitis, otitis, conjunctivitis, and sinusitis, which can potentially lead to organ damage in the long run [5, 6]. The BTK gene is located at the long arm of the X chromosome from Xq21.3 to Xq22, at approximately 101349447 to 101390796 base pairs in the human genome (GRCh38.p7). The gene consists of 19 exons extending over 37.5kb of the genomic sequences, and encodes a protein of 659 amino acid residues [6-9]. The BTK gene product (BTK protein) is a cytoplasmic non-receptor tyrosine kinase of the TEC family, consisting of five structural domains namely pleckstrin homology (PH), tec homology (TH), src homology 3 (SH3), src homology 2 (SH2) and catalytic kinase (TK also known as SH1) [10-12]. According to findings by Zheng *et al*. mutations occurring in any of the five domains can impact the activity of the BTK protein, consequently influencing the development and proliferation of B cells [13].

Non-synonymous single nucleotide polymorphism (nsSNP) mutations have often been reported to cause deleterious effects leading to various diseases and organ dysfunction. For instance, a mutation involving a substitution of a hydrophobic residue (alanine) with a hydrophilic residue (serine) in the TUBB8 protein affects the folding of β-tubulin structure, causing a dominant negative effect by impairing the microtubule behavior and spindle assembly leading to infertility due to oocyte maturation arrest [14]. Another study demonstrated that the PITX2 mutation (Lys88Glu) occurring at the DNA binding site led to the formation of stable heterodimers, causing reduced promoter-binding activity and subsequently contributing to Axenfeld-Rieger syndrome [15]. While Di Maas *et al*. and Steffen *et al*. reported the association of Arg215Trp substitution in Nijmegen Breakage Syndrome 1 (NBS1) with colorectal cancer that is mainly found in Czech infants [3, 16]. In the case of XLA disease, Mohammad *et al*. performed a thorough analysis of deleterious mutations in the BTK gene. From the extensive sequence and structural examination, two specific pathogenic mutations, namely E589G and M630K, demonstrate substantially induced changes in the structural conformation and stability of the BTK

protein, leading to the development of XLA disease [17].

Interestingly, a transcriptomic profiling of monocytes from XLA patients was conducted by our groups, Mirsafian *et al*. which revealed a mutation at c.1888A>T, causing a p.M630L substitution within the kinase domain of the BTK protein [7]. The active site region of the BTK protein, crucial for protein-substrate binding and catalytic efficiency, was identified between residues 539 to 567 [18]. Given that the substitution does not occur at the active site of the BTK kinase domain, it is intriguing to explore how the p.M630L mutation leads to a biologically dysfunctional protein in XLA patients, besides the previously reported E589G and M630K mutations [17]. Amino acid substitutions resulting from nsSNP mutations have been hypothesized to alter the 3D conformational structure of a protein, thereby compromising its biological function and contributing to the disease phenotype in patients. Molecular dynamics analysis was performed to elucidate the impact of the p.M630L mutation on the 3D conformational structure of BTK. To date, this is a novel mutational study associated with XLA which corroborate the significant changes to the BTK as similarly shown in previous study. Mutational effect at M630L in BTK has been declared as main destabilizer for functional domain, thus affecting B-cell maturation [7].

2.0 METHODOLOGY

2.1 Transcriptome Dataset

The RNASeq dataset of BTK gene in XLA Patient 1 was retrieved from Mirsafian *et al*. [7]. The sequence was analyzed using FastQC [19] and mapped to the reference genome (GRCh38) using HISAT2 [20].

2.2 Non-synonymous Single Nucleotide Polymorphism (nsSNP) Verification and Pathogenic Predictions

The potential nsSNP from the transcriptome dataset was screened and identified using Genome Analysis Toolkits (GATK) [21]. The nsSNP was further refined and analyzed using web-based ANNOVAR (wANNOVAR) software [22]. Prediction of the pathogenicity of nsSNP (p.M630L) was conducted using nine in silico functional assessment tools namely, SIFT [23], Polyphen-2 [24], PROVEAN [25], MutationAssessor [26], PANTHER [27], FATHMM [28], A-GVGD [29], MutPred [30], and MutationTaster [31].

2.3 Generation of Mutant Protein and 3D Structural Mapping of the nsSNP Location

Initial coordinates of wildtype Bruton Tyrosine Kinase (BTK) kinase domain protein structure, (PDB ID: 5U9D, chain A, resolution 1.33 Å), was retrieved from Protein Data Bank (PDB) [32]. The BTK kinase domain is comprised of 271 amino acid residues (starting from Gly389 to Ser659) and is associated with compound 3 (ligand ATP-competitive BTK inhibitor) [33]. All water molecules and the ligand were removed from the crystal structure. In silico mutagenesis was performed to obtain mutant BTK kinase domain (p.M630L) by reconstituting the wildtype BTK kinase domain (WT) protein residues with its polymorphic residues using UCSF Chimera [34].

2.4 Molecular Dynamics Simulation (MDS)

In order to reveal the structural consequences of the single mutation in the BTK kinase domain, we performed molecular dynamics (MD) simulation using GROMACS 5.1.4 [35]. MD was conducted at physiological temperature and neutral pH. The system was solvated using explicit flexible extended simple point charge (SPC/E) water model [35], embedded in a cubic box (with 8.565 nm edges) at ≥10 Å marginal radius, and neutralized by adding 7 sodium (Na+) counter ions. Steepest descent energy minimization was performed for 50000 steps to give each solvated structure a maximum force below 1000 kJ mol-1nm-1. After energy minimization, the system was equilibrated at constant temperature (300 K) and pressure (1 bar) with a time step of 2 fs. The temperature was kept constant using a Berendsen thermostat and the electrostatic interaction was calculated using the particle mesh Ewald summation method [36]. Finally, 100 ns of MD simulation were performed in triplicates for both WT and p.M630L structures. Comparative analysis of structural deviations in both WT and p.M630L such as root mean square deviation (RMSD), root mean square fluctuation (RMSF), secondary structure count, solvent accessible surface area (SASA), intermolecular protein-solvent hydrogen bond (cut-off radius: 0.35 nm), and atomic distribution were averagely computed using built-in functions in GROMACS package. The graphs were generated using GRACE software (http://plasma-gate.weizmann.ac.il/Grace/).

2.5 Cluster Analysis and Principal Component Analysis (PCA)

Results from MD simulation was further analysed using cluster analysis and principal component analysis (PCA) to justify the structural deviations caused by the nsSNP mutation. Cluster analysis was performed using the Gromos method in the GROMACS package (gmx cluster) and analysed at 0.17 nm cut-off. In order to analyse the overall motions in the simulations, PCA was conducted using built in functions of GROMACS software (gmx covar and gmx anaeig). Covariance matrices were constructed using Ca backbone trajectories. The first two eigenvectors with the largest eigenvalues were used to build the 2D projection for each independent trajectory. The results were then visualized using UCSF

Chimera [34] and GRACE software (http://plasma-gate.weizmann.ac.il/Grace/).

3.0 RESULTS AND DISCUSSION

The 3D conformational protein structure BTK kinase domain with PDB ID 5U9D and the corresponding p.M630L substitution was shown in Figure 1.

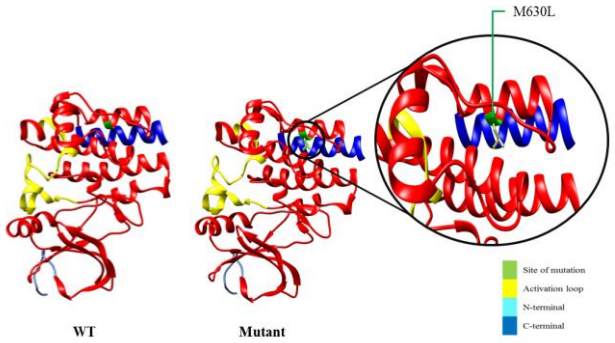


Figure 1 The illustration represents the nsSNP (p.M630L) in the BTK kinase domain of an XLA patient (P1). In the WT structure, the mutation site is highlighted in green, with the mutation labeled as M630L. The activation loops of the protein are highlighted in yellow, while cyan corresponds to the N-terminal, and blue corresponds to the C-terminal of the protein structure

We performed functional prediction analysis to predict the consequences of p.M630L substitution on BTK protein biological function using different tools as described in Table 1.

Table 1 The potential functionality of the novel nsSNP within the BTK gene of a XLA patient

Mutation Prediction Tool	Dataset	Score	Pathogenicity Prediction
SIFT	N/A	0	Damaging
Polyphen-2	HDiv	0.878	Probably damaging
	HVar	0.925	Damaging
PROVEAN	N/A	-2.54	Damaging
MutationAssessor	N/A	2.055	Medium altered function
PANTHER	N/A	1036	Probably damaging
FATHMM	N/A	-1.87	Damaging
A-GVGD	GD	14.3	Interfere with function
MutPred	N/A	0.66	Loss of catalytic residue at V626 (p = 0.0231)
MutationTaster	N/A	1	Damaging

Hdiv: HumDiv, HVar: HumVar, GD: Grantham deviation, N/A: Not available

All the nine functional prediction tools (SIFT, Polyphen-2, PROVEAN, MutationAssessor, PANTHER, FATHMM, A-GVGD, MutPred and MutationTaster) predicted the mutation to be deleterious and may interfere with structure and function. This prediction agreed with the innate immune function dysregulation reported before in the XLA patient [7]. In their study, the M630L variant identified in the patient's transcriptome has been anticipated to be deleterious and detrimental to the function of BTK [7]. A similar study evaluated and screened all mutations in BTK from the public domain database through sequence and structure analyses. In their findings, the mutational position of 630 made the BTK alter its function [17]. However, the study only highlights the mutation of methionine to the lysine residue but not to leucine as identified from the patient's transcriptome in our earlier work. In the following section, we have further examined this M630L using MD simulation, uncovering how this mutant also impairs the functional properties of the BTK enzyme.

3.1 MD Simulation of Wildtype and Mutant BTK Kinase Domain

Molecular dynamics (MD) simulations were performed to evaluate the structural behaviour of the mutant protein (p.M630L) in comparison to the wildtype BTK protein (WT). Both the WT and p.M630L proteins had similar total energy levels, demonstrating their stability across the whole 100 ns simulation period (Figure 2). The structural deviations for both WT and p.M630L proteins was assessed by calculating the root mean square deviation (RMSD) of the Ca backbone during the 100 ns molecular dynamics (MD) simulation, as shown in Figure 3. Significantly, a minor variation was observed in the p.M630L from around 20 ns to the end of the simulations. This observation indicates a difference in the structural conformation between the two BTK kinase domain configurations. The WT protein had an average RMSD value of 0.21 nm, which is higher than the value of 0.17 nm for the p.M630L mutant. This indicates that the mutant protein underwent less structural motions and closely resembled as its starting conformation compared to the wildtype protein. This finding contradicts to the earlier BTK simulation study 17, as the average values and deviation patterns are not corroborated. Nevertheless, our p.M630L demonstrated a quicker attainment of equilibrium, commencing at 20 nanoseconds. Reaching equilibrium or a plateau in RMSD indicates that the protein has undergone structural relaxation and has settled into a stable state within the simulation time frame. Thus, atomic fluctuations by residues are crucial to infer BTK's local conformation and dynamics pattern.

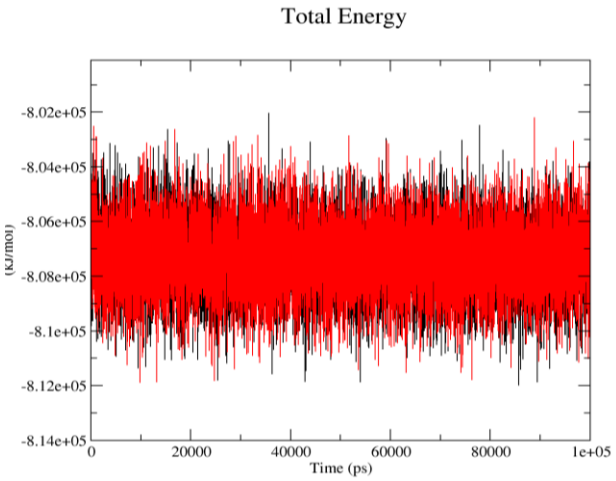


Figure 2 The total energy of WT (black line) and p.M630L mutant (red line) of the BTK proteins throughout the molecular dynamic simulation. The total energy remained constant, indicating stability in the total energy of both systems

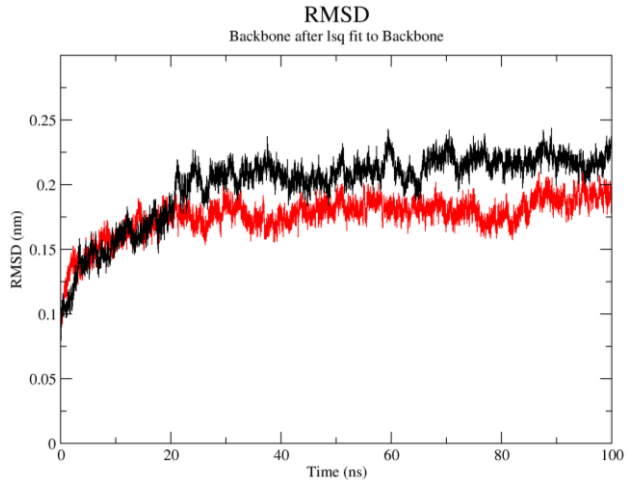


Figure 3 The RMSD with respect to Ca of WT (black line) and p.M630L mutant (red line) kinase domain throughout molecular dynamic simulation. The p.M630L mutant of kinase domain exhibits less structural deviation as compared to WT

The root mean square fluctuation (RMSF) value of the Ca atoms in both proteins was computed during the simulation to analyse the dynamic behaviour of the amino acid residues that contribute to conformational changes (Figure 4). Both the wild-type (WT) and p.M630L mutant showed Ca with RMSF values greater than 0.1 nm in the upstream (N-terminal) region and around the active site. The catalytic domain of BTK kinase is regulated by the phosphorylation of tyrosine residues including Tyr551 in the activation loop of the kinase domain, which participate in kinase activation. The p.M630L mutation in the BTK kinase domain results in more peaks at the activation loop compared to the wild type (WT), demonstrating the occurrence of conformational alterations in the secondary structure. However, the mutated amino acid maintains a stable fluctuation with a Ca RMSF value < 0.1 nm.

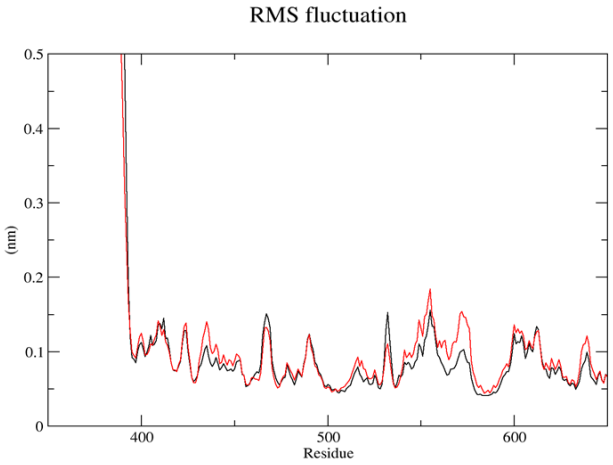


Figure 4 The amino acid conformational changes for WT (black line) and p.M630L BTK kinase domain (red line) with respect to Ca

The changes in the secondary structure of both MD simulations were subsequently evaluated using the Database of Secondary Structure in Proteins (DSSP). The analysis using DSSP showed a little change in the secondary structure of p.M630L towards the end of the 100 ns period (Figure 5). The high quantity of coil, bend, and turn in p.M630L, compared to the WT indicates the possibility of structural deformation at the active site or activation region in the mutant protein, resulting from individual amino acid substitution. This finding was corroborated by the RMSF analysis, demonstrating structural conformation changes in the BTK kinase domain.

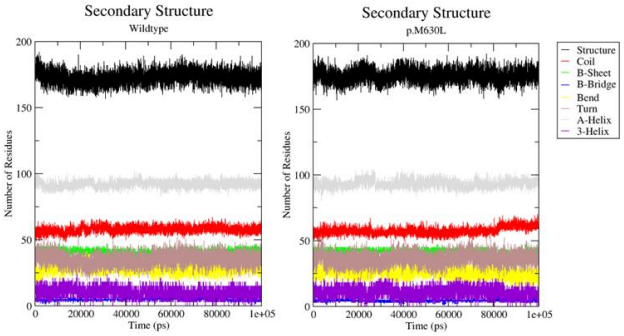


Figure 5 The predicted secondary structure counts for WT and p.M630L BTK kinase domain throughout 100 ns by the Database of Secondary Structure in Proteins (DSSP)

Solvent accessible surface area (SASA) has been used as a decisive factor in determining protein folding and stability. It provides valuable insights into the conformational dynamics, hydrophobicity, and interactions of the protein with the solvent, all of which contribute to its functionality. SASA can also predict

protein hydration, which may play a role in protein function, such as the accessibility of critical residues involved in the biological mechanisms during B-cell maturation processes [37, 38]. Fluctuations in SASA levels signify alterations in exposed amino acid residues, which in turn can potentially impact the protein's tertiary structure. The p.M630L structure exhibited a slightly higher SASA value, reaching up to 143 nm^2 from the initial time to 20 ns in contrast to the WT. However, the average surface area values appeared nearly comparable for both p.M630L and WT, both measuring around 141 nm^2 (Figure 6). This is due to the p.M630L having a lower SASA value after half of the simulation time. Given that the residue at position 630 within the 3D structure of the BTK kinase domain is deeply embedded in the protein's core and the amino acid substitution from methionine to leucine is hydrophobic, this substitution might not induce substantial alterations in SASA owing to the rigidity of the structure. Further analysis of the 3D structural conformational flexibility of both WT and p.M630L involved assessing protein-solvent intermolecular hydrogen bonds. These hydrogen bonds play a crucial role in upholding the folding and stability of protein structures [39].

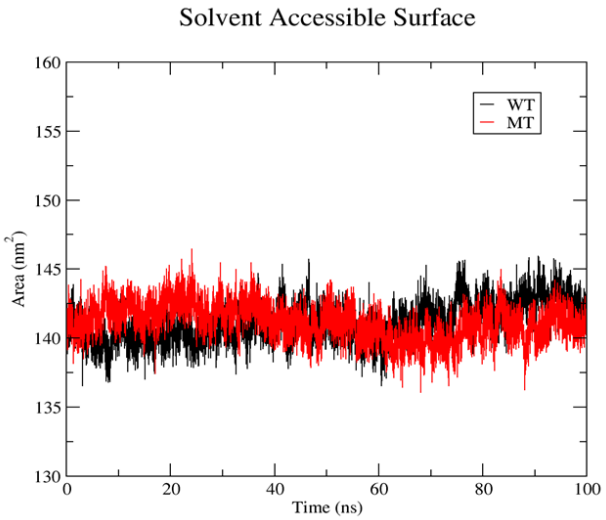


Figure 6 The assessment of solvent accessible surface area (SASA) of WT and p.M630L mutant BTK kinase domain throughout 100 ns of molecular dynamic simulation. The average surface area values appeared nearly comparable for both p.M630L and WT, both measuring around 141 nm^2

The investigation continuously observed that the p.M630L variant had a slightly reduced number of hydrogen bonds compared to the wild type (WT) throughout the 100 ns length (Figure 7). Therefore, the results convey that the nsSNP mutation can produce a p.M630L protein that exhibits a more condensed and inflexible three-dimensional structure than the wild-type (WT) protein

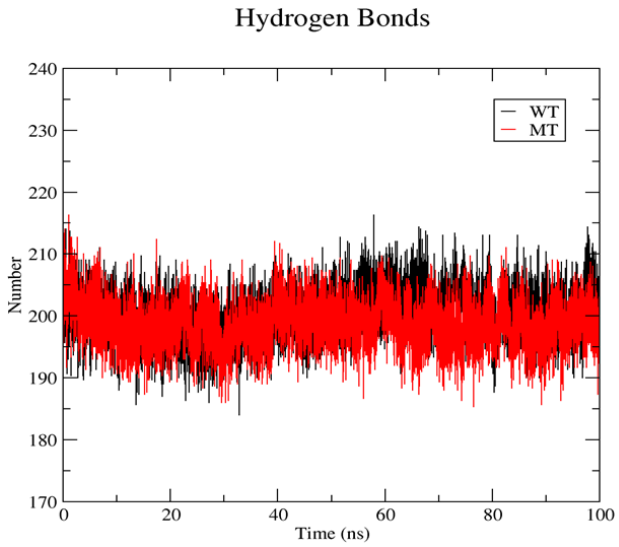


Figure 7 The number of intermolecular hydrogen bonds between the BTK protein and the solvent formed over a 100 ns molecular dynamic simulation. The WT variant is depicted by the black line, while the p.M630L variant is represented by the red line

Figure 8 shows a notable disparity in the arrangement of atoms between the mutant and wildtype proteins. Specifically, the p.M630L variant displays a greater maximal density (61.2 nm^{-3}) compared to the wildtype (49.1 nm^{-3}). This observation indicates that the p.M630L mutation causes significant changes in the orientation and distribution of atoms. The results obtained from the molecular dynamics (MD) simulation provide strong evidence that the non-synonymous single nucleotide polymorphism (nsSNP) p.M630L mutation has the potential to cause significant alterations in both the structure and function of the BTK kinase domain. Hence, this mutation is likely to increase the compactness and rigidity of the protein conformation with secondary structure deformation at the activation region.

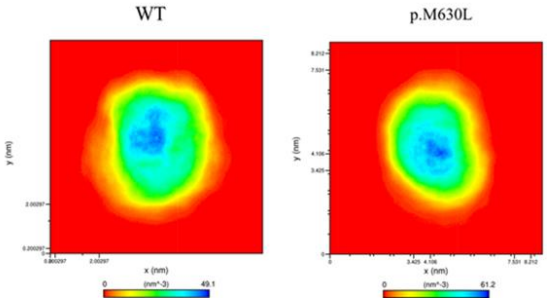


Figure 8 The atomic distribution of WT and P.M630L mutant of BTK kinase domain. The p.M630L variant demonstrates a higher maximum density at 61.2 nm-3 in contrast to the WT's 49.1 nm-3, indicating the significant changes in orientation and atomic distribution attributed to the p.M630L mutation

3.2 Cluster Analysis and Principal Component Analysis

To gain better insights into the structural conformational differences between the WT and p.M630L proteins, cluster analysis was employed on the MD trajectories of both protein structures. As detailed in Table 2, the MD trajectories of the WT protein formed 19 distinct clusters, with RMSD values spanning from 0.0655 nm to 0.345 nm and an average RMSD of 0.0925 nm. In contrast, the MD trajectories for p.M630L were clustered into 12 groups, displaying RMSD values ranging from 0.0644 nm to 0.301 nm and an average RMSD of 0.0841 nm. We superimposed the trajectory structures to identify differences in the 3D conformational structures between the WT and p.M630L proteins. The RMSD value for the superimposed average structure was 1.094 nm, with higher RMSD values observed for clusters 2 and 3 (Table 3). The superimposed structures were visualized in Figure 9. The difference in the superimposition of protein structures can lead to substantial alterations in the protein's conformation, stability, and interactions, potentially impacting its biological function.

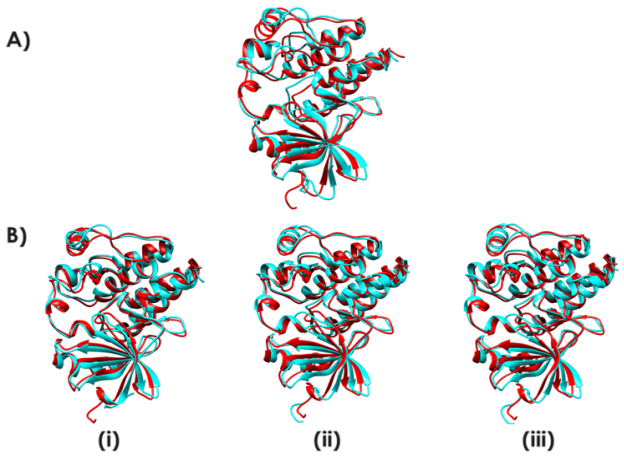


Figure 9 Cluster analysis between WT and p.M630L. (A) Superimposed average structure of WT and p.M630L (B) Superimposed structures of clusters at RMSD cut-off 1.7 Å. (i) First cluster (ii) Second cluster (iii) Third cluster. Structures in cyan represent the WT, red represent p.M630L

Table 2 Cluster analysis of WT and p.M630L at 0.17 nm cut-off

Structure	WT	p.M630L
RMSD range (nm)	0.0655 - 0.345	0.0644 - 0.301
Average RMSD (nm)	0.0925	0.0841
No. of clusters	19	12

Table 3 RMSD of superimposed average structures and top three clusters with the highest number of trajectories

Superimposed WT and p.M630L	RMSD (nm)
Average Structure	1.094
Cluster 1	1.073
Cluster 2	1.186
Cluster 3	1.186

The PCA analysis was also performed on the MD trajectories of both WT and p.M630L 3D conformational structures to examine correlated motion differences, to provide additional support for the MD analysis. The first two eigenvectors reveal significant fluctuations in the system, with most internal motions of the protein structure restricted to a small dimension within the essential subspace. Figure 10 represents the trajectory projections for the first two eigenvectors of both WT and p.M630L structures over the 100 ns simulation period.

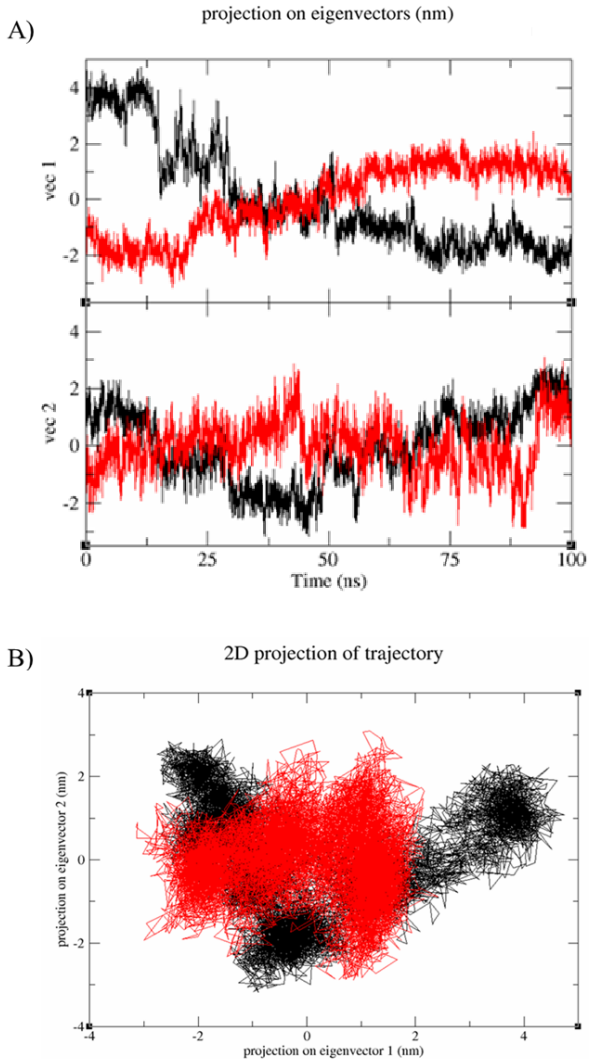


Figure 10 PCA for WT and p.M630L mutant at 2 eigenvectors (A) linear (1D) projection on eigenvectors (B) 2-Dimensional (2D) projection of trajectory. Black line: WT; red line: p.M630L mutant

The structural divergence between the proteins shows a 9.0% overlap, with an average RMSD value of 0.091 nm. Trajectories of p.M630L demonstrate reduced dynamic movements, covering a lesser region of phase space along the eigenvector 1 and eigenvector 2 planes compared to the WT. Another study conducted by Pandey *et al*. to observe the

deleterious single nucleotide polymorphism towards neurodegenerative disorder. The researcher performed PCA analysis to demonstrate the mutational effects of JD-1 and VPS35 protein which causes protein structure to become more compressed, affecting the flexibility of the protein compared to its native form [40]. Thus, our analyses suggest potential evidence of the deleterious effect of the p.M630L mutation towards conformation of the BTK protein, affecting its biological function due to alteration of protein flexibility.

Our study addressed the deleterious effect of the p.M630L mutation on the BTK kinase domain, potentially hindering the proper functioning of the BTK protein. This is potentially due to alteration of protein accessible regions that are critical for biological function during the process of B-cells maturation in the XLA patient resulting due to overall protein conformation changes resulting from single subunit mutation. In 1998, Lipscomb *et al*. reported that substituting methionine with leucine caused an increased protein stability in T4 lysozymes [41]. Furthermore, a study conducted by Udhaya *et al*. showcased how the substitution of the hydrophobic amino acid valine with leucine at position 53 within the PAX6 gene contributes to the pathogenicity of the mutation. This effect is attributed to the accessibility of the protein, as leucine is comprised of a shorter side chain compared to valine [42]. Furthermore, several studies also reported that nsSNP mutations altered the flexibility and stability of protein, leading to detrimental effects and diseases [43, 44]. We identified that the p.M630L substitution within the BTK kinase domain induces significant structural alterations in its 3D protein structure, resulting in a more compact and rigid conformation.

In another case of XLA diseases, the mutation was found in the promoter region, causing unexpressed B-lymphocytes or monocytes in the XLA patient [45]. In our cases, the single subunit mutation was found in the kinase region of the BTK protein. Although the mutated residue is not within the active site, it leads to an alteration in overall protein flexibility. This results in changes to the accessible regions of protein residues that are critical for biological function. Our findings suggest that, in addition to subunit mutations at residues E589G and M630K, the mutation at c.1888A>T causing substitution of residue M630L also needs to be considered for XLA disease treatment through gene editing technologies such as the CRISPR-associated protein 9 system.

Despite the advantages of MD simulation in exploring detailed protein behavior and conformational changes due to single subunit mutations, there is a limitation that can lead to variations in results when compared to experimental data. This is due to various factors in MD simulation, such as the selection of accurate force field parameters, adequate computational resources, and appropriate modeling assumptions [46]. Therefore, further validation through experimental studies is necessary to support the hypothesis and demonstrate the structural changes resulting from the M630L mutation, using techniques such as X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy.

4.0 CONCLUSION

A novel nsSNP with a p.M630L substitution in the BTK protein of an XLA patient was previously reported. Computational analyses of the mutation predicted a damaging effect on the structure and function, despite the substitution occurring outside the active region of the BTK kinase domain protein. MD analysis was performed to observe alterations in the 3D conformational structure of the mutant p.M630L BTK kinase domain, resulting in dysfunctionality of the BTK protein. The MD simulation indicated that the mutant p.M630L exhibited a more compact and rigid 3D conformation with structural deformation of activation region compared to the wildtype BTK kinase domain protein. These findings conclude that the novel nsSNP p.M630L is likely to disrupt the native conformation of protein structure, consequently leading to dysfunctionality in the BTK protein.

Acknowledgement

The authors express their gratitude to the Director General of Health Malaysia for granting permission to publish this paper. This research received support from the Ministry of Higher Education Malaysia Fundamental Research Grant Scheme (FRGS): FP050-2016 and the University of Malaya High Impact Research (HIR) Grant: UM. S/P/HIR/MOHE/30.

Conflict of Interest

The author declares no affiliations or involvement with any organization or entity, and has no financial interests related to the subject matter or materials discussed in this manuscript.

References

[1] Valiaho, J., C. I. Smith, and M. Vihinen, 2006. BTKbase: The Mutation Database for X-linked Agammaglobulinemia. *Human Mutation*. 27(12): 1209–17.
[2] Plebani, A., *et al*. 2002. Clinical, Immunological, and Molecular Analysis in a Large Cohort of Patients with X-linked Agammaglobulinemia: An Italian Multicenter Study. *Clinical Immunology*. 104(3): 221–30.
[3] Maas, A. and R. W. Hendriks. 2001. Role of Bruton's Tyrosine Kinase in B Cell Development. *Journal of Immunology Research*. 8(3-4): 171–81.
[4] Genevier, H. C., *et al*. 1994. Expression of Bruton's Tyrosine Kinase Protein within the B Cell Lineage. *European Journal of Immunology*. 24(12): 3100–3105.
[5] Cunningham-Rundles, C. and P. P. Ponda. 2005. Molecular Defects in T- and B-cell Primary

Immunodeficiency Diseases. *Nature Reviews Immunology*. 5(11): 880–92.
[6] Vihinen, M., *et al*. 1999. Mutations of the Human BTK Gene Coding for Bruton Tyrosine Kinase in X-linked Agammaglobulinemia. *Human Mutation*. 13(4): 280–5.
[7] Mirsafian, H., *et al*. 2017. Transcriptome Profiling of Monocytes from XLA Patients Revealed the Innate Immune Function Dysregulation due to the BTK Gene Expression Deficiency. *Scientific Reports.* 7(1): 1–13.
[8] Sayers, E. W., *et al*. 2010. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Research.* 39(suppl_1): D38-D51.
[9] Vetrie, D., *et al*. 1993. The Gene Involved in X-linked Agammaglobulinaemia is a Member of the src Family of Protein-tyrosine Kinases. *Nature.* 361(6409): 226–233.
[10] Sharma, S., *et al*. 2016. Identification of a Structurally Novel BTK Mutation that Drives Ibrutinib Resistance in CLL. *Oncotarget.* 7(42): 68833.
[11] Mohamed, A. J., *et al*. 2009. Bruton's Tyrosine Kinase (Btk): Function, Regulation, and Transformation with Special Emphasis on the PH Domain. *Immunological Reviews.* 228(1): 58–73.
[12] Smith, C.I., *et al*. 2001. The Tec Family of Cytoplasmic Tyrosine Kinases: Mammalian Btk, Bmx, Itk, Tec, Txk and Homologs in other Species. BioEssays. 23(5): 436–46.
[13] Zheng, B., *et al*. 2014. A Novel Bruton's Tyrosine Kinase Gene (BTK) Missense Mutation in a Chinese Family with X-linked Agammaglobulinemia. *BMC Pediatrics*. 14(1): 1-5.
[14] Xiang, J., *et al*. 2018. Human Oocyte Maturation Arrest Caused by a Novel Missense Mutation in TUBB8. *Journal of International Medical Research.* 46(9): 3759–3764.
[15] Saadi, I., *et al*. 2003. Dominant Negative Dimerization of a Mutant Homeodomain Protein in Axenfeld-Rieger Syndrome. *Molecular and Cellular Biology*. 23(6): 1968–82.
[16] Steffen, J., *et al*. 2004. Increased Cancer Risk of Heterozygotes with NBS1 Germline Mutations in Poland. *International Journal of Cancer*. 111(1): 67–71.
[17] Mohammad, T., *et al*. 2020. Impact of Amino Acid Substitution in the Kinase Domain of Bruton Tyrosine Kinase and its Association with X-linked Agammaglobulinemia. *International Journal of Biological Macromolecules*. 164: 2399–2408.
[18] Roskoski, R., Jr. 2016. Ibrutinib Inhibition of Bruton Protein-tyrosine Kinase (BTK) in the Treatment of B Cell Neoplasms. *Pharmacological Research*. 113(Pt A): 395–408.
[19] Andrews, S. 2017. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010.
[20] Kim, D., B. Langmead, and S. L. Salzberg. 2015. HISAT: A Fast Spliced Aligner with Low Memory Requirements. *Nature Methods*. 12(4): 357–60.
[21] Van der Auwera, G. A., *et al*. 2013. From FastQ Data to High Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics.* 43(1110): 11 10 1–11 10 33.
[22] Chang, X. and K. Wang. 2012. wANNOVAR: Annotating Genetic Variants for Personal Genomes Via the Web. *Journal of Medical Genetics.* 49(7): 433–6.
[23] Kumar, P., S. Henikoff, and P. C. Ng. 2009. Predicting the Effects of Coding Non-synonymous Variants on Protein Function using the SIFT Algorithm. *Nature Protocols*. 4(7): 1073–81.
[24] Adzhubei, I. A., *et al*. 2010. A Method and Server for Predicting Damaging Missense Mutations. *Nature Methods.* 7(4): 248–9.
[25] Choi, Y. and A.P. Chan. 2015. PROVEAN Web Server: A Tool to Predict the Functional Effect of Amino Acid Substitutions and Indels. *Bioinformatics*. 31(16): 2745–7.
[26] Reva, B., Y. Antipin, and C. Sander. 2011. Predicting the Functional Impact of Protein Mutations: Application to Cancer Genomics. *Nucleic Acids Research.* 39(17): e118.
[27] Mi, H., *et al*. 2016. PANTHER Version 10: Expanded Protein Families and Functions, and Analysis Tools. *Nucleic Acids Research.* 44(D1): D336–42.
[28] Shihab, H. A., *et al*. 2013. Predicting the Functional, Molecular, and Phenotypic Consequences of Amino Acid Substitutions using Hidden Markov Models. *Human Mutation.* 34(1): 57–65.
[29] Tavtigian, S.V., *et al*. 2006. Comprehensive Statistical Study of 452 BRCA1 Missense Substitutions with Classification of Eight Recurrent Substitutions as Neutral. *International Journal of Cancer.* 43(4): 295–305.
[30] Li, B., *et al*. 2009. Automated Inference of Molecular Mechanisms of Disease from Amino Acid Substitutions. *Bioinformatics*. 25(21): 2744–50.
[31] Schwarz, J. M., *et al*. 2010. MutationTaster Evaluates Disease-causing Potential of Sequence Alterations. *Nature Methods.* 7(8): 575–6.
[32] Berman, H.M., *et al*. 2000. The Protein Data Bank. *Nucleic Acids Research.* 28(1): 235–242.
[33] Cuozzo, J. W., *et al*. 2017. Discovery of a Potent BTK Inhibitor with a Novel Binding Mode by Using Parallel Selections with a DNA-Encoded Chemical Library. *ChemBioChem.* 18(9): 864–871.
[34] Pettersen, E.F., *et al*. 2004. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *Journal of Computational Chemistry*. 25(13): 1605–1612.
[35] van der Spoel, D., P. J. Van Maaren, and H. J. Berendsen. 1998. A Systematic Study of Water Models for Molecular Simulation: Derivation of Water Models Optimized for Use with a Reaction Field. *The Journal of Chemical Physics.* 108(24): 10220–10230.
[36] Darden, T., D. York, and L. Pedersen. 1993. Particle Mesh Ewald: An N· log (N) method for Ewald Sums in Large Systems. *The Journal of Chemical Physics*. 98(12): 10089–10092.
[37] Ehrlich, L., *et al*. 1998. Prediction of Protein Hydration Sites from Sequence by Modular Neural Networks. *Protein Engineering, Design and Selection.* 11(1): 11–9.
[38] Savojardo, C., *et al*. 2020. Solvent Accessibility of Residues Undergoing Pathogenic Variations in Humans: From Protein Structures to Protein Sequences. *Frontiers in Molecular Biosciences.* 7: 626363.
[39] Pace, C. N., *et al*. 2014. Contribution of Hydrogen Bonds to Protein Stability. *Protein Science*. 23(5): 652–61.
[40] Pandey, S., *et al*. 2020. An In Silico Analysis of Deleterious Single Nucleotide Polymorphisms and Molecular Dynamics Simulation of Disease Linked Mutations in Genes Responsible for Neurodegenerative Disorder. *Journal of Biomolecular Structure and Dynamics.* 38(14): 4259–4272.
[41] Lipscomb, L. A., *et al*. 1998. Context-dependent Protein Stabilization by Methionine-to-leucine Substitution Shown in T4 Lysozyme. *Protein Science.* 7(3): 765–773.
[42] Udhaya Kumar, S., *et al*. 2015. Functional and Structural Characterization of Missense Mutations in PAX6 Gene. *Frontiers in Biology*. 10: 377–385.
[43] Zhang, Z., *et al*. 2012. Analyzing Effects of Naturally Occurring Missense Mutations. *Computational and Mathematical Methods in Medicine*. 2012: 805827.
[44] Wang, Z. and J. Moult. 2001. SNPs, Protein Structure, and Disease. *Human Mutation*. 17(4): 263–70.
[45] García-Morato, M. B., *et al*. 2020. A Mutation in the Promoter Region of BTK Causes Atypical XLA. Heliyon. 6(9).
[46] Barbhuiya, S. and B. B. Das. 2033. Molecular Dynamics Simulation in Concrete Research: A Systematic Review of Techniques, Models and Future Directions. *Journal of Building Engineering*. 2023: 107267.