Current Research in Structural Biology 11 (2026) 100186

Contents lists available at ScienceDirect

Current Research in Structural Biology

journal homepage: www.journals.elsevier.com/current-research-in-structural-biology

Research Article

Computational hybrid framework integrating clinical-mutation profiling and physics-based simulations to predict structural tolerance of Bruton tyrosine kinase (BTK) and sustained ARQ 531 binding

Abbas Khan [a], Mona A. Sawali [a], Mohanad Shkoor [b], Fakhrul Hasssan [c], Raed M. Al-Zoubi [d,e,f], Fahad M. Alshabrmi [g], Sergio Crovella [e], Abdelali Agouni [a,*]

[a] Department of Pharmaceutical Sciences, College of Pharmacy, QU Health, Qatar University, P.O. Box 2713, Doha, Qatar
[b] Department of Chemistry and Earth Sciences, College of Arts and Science, Qatar University, P.O. Box 2713, Doha, Qatar
[c] Faculty of Rehabilitation and Allied Health Sciences, Riphah International University, Islamabad, Pakistan
[d] Surgical Research Section, Department of Surgery, Hamad Medical Corporation, Doha, Qatar
[e] Department of Biomedical Sciences, College of Health Sciences, QU Health, Qatar University, P.O. Box 2713, Doha, Qatar
[f] Department of Chemistry, Jordan University of Science and Technology, P.O. Box 3030, Irbid, 22110, Jordan
[g] Department of Medical Laboratories, College of Applied Medical Sciences, Qassim University, Buraydah, 51452, Saudi Arabia

ARTICLE INFO

Handling Editor: Dr A Wlodawer

Keywords:
BTK
Leukemia
Mutation screening
All-atom simulation
PCA
Binding free energy

ABSTRACT

Bruton tyrosine kinase (BTK) is a non-receptor tyrosine kinase crucial for relaying signals from the B cell antigen receptor (BCR) in cancerous B lymphocytes. It has been reported that mutations in this protein cause resistance to various covalent and non-covalent drugs. Therefore, we employed a computational genomic mutation screening strategy, combined with molecular simulation, to assess the impact of clinical substitutions on the structure and binding of ARQ 531(Nemtabrutinib), the most potent BTK inhibitor. Using various machine learning algorithms, 62 clinical mutations were identified as deleterious among the 82, while 11 were classified as highly destabilizing using graph signature-based methods. We selected the top mutations that are deleterious and highly destabilizing, which include L408P, Y476D, M477R, C481R, C481Y, and L542P. Molecular docking analysis revealed no significant variations in the bonding network, while molecular simulation results revealed local changes only in the dynamic behavior. The resemblance to the wild type and mutants in certain PCs (principal components) implies that specific structural aspects or dynamics remain preserved despite the mutation, and the single energetic minimum indicates a robust and dominant structural state, reinforcing that these mutations affect the protein locally but not globally. Finally, the total binding free energy (TBE) calculations revealed that ARQ531 exhibits broadly conserved binding energetics across clinically observed BTK mutants, with select variants (notably C481 substitutions) showing moderately enhanced stabilization relative to the wild type. These findings confirm that ARQ 531 remains a significant investigational therapy specifically for overcoming drug resistance in multiple leukemias and B-cell malignancies and can serve as a starting point for designing more robust and effective BTK inhibitors.

1. Introduction

Bruton tyrosine kinase (BTK) is a non-receptor tyrosine kinase crucial for relaying signals from the B cell antigen receptor (BCR) and various other cell surface receptors in both healthy and cancerous B lymphocytes. BCR signaling activation, occurring primarily in secondary lymphoid organs, fuels the uncontrolled growth of malignant B cells, such as those found in chronic lymphocytic leukemia (CLL). In the last decade, BTK inhibitors (BTKi) have gained much popularity and have gradually replaced chemotherapy-based treatment for many individuals with CLL and mantle cell lymphoma (MCL) (Burger, 2019; Mohamed et al., 2009; Quartier et al., 1999). The process of BTK activation begins with a BCR-antigen binding, followed by an activation cascade with downstream signalling events, and thus contributes to B cell survival,

* Corresponding author. Department of Pharmaceutical Sciences, College of Pharmacy, Qatar University, P.O. Box 2713, Doha, Qatar.
E-mail addresses: Abbas.khan@qu.edu.qa (A. Khan), mshkoor@qu.edu.qa (M. Shkoor), fakhr.hassan@riphah.edu.pk (F. Hasssan), ralzoubi@hamad.qa (R.M. Al-Zoubi), Fshbrmy@qu.edu.sa (F.M. Alshabrmi), sgrovella@qu.edu.qa (S. Crovella), aagouni@qu.edu.qa (A. Agouni).

https://doi.org/10.1016/j.crstbi.2026.100186
Received 12 October 2025; Received in revised form 20 February 2026; Accepted 24 February 2026
Available online 26 February 2026

proliferation, and maturation (Smith et al., 2001; Takata and Kurosaki, 1996). After the BCR is engaged and activated, the immune-receptor tyrosine-based activation motifs (ITAMs) located on the cytoplasmic tails of the Igα and Igβ (CD79a/CD79b) proteins cluster into one unit and are phosphorylated (de Gorter et al., 2007). Phosphorylated ITAMs recruit spleen tyrosine kinase (SYK), allowing SYK to form a complex with a B cell linker (BLNK, also known as SLP65 or BASH). After BTK is activated, then CD19, PI3-K is activated, followed by the elevation of cytoplasmic PIP3 levels, which signals further downstream events, activating phospholipase Cγ2 (PLCγ2), which is responsible for calcium and PKC activation. The activation of NF-κB and ERK regulates further transcriptional activation. Without BTK, BCR signalling cannot effectively induce mature peripheral B cell differentiation. This results in an altered development of B cells and impaired functional responses, including impaired cellular proliferation, lower-level expression of activation markers, decreased production of cytokines and antibodies, and impaired responses to infections. Remarkably, the overexpression of BTK in B cells induces spontaneous germinal center formation, antinuclear autoantibody production (Quartier et al., 1999; Smith et al., 2001), and initiates a systemic lupus erythematosus (SLE)-like autoimmune condition. This outcome stems from heightened BCR signaling and increased NF-κB activation. Interestingly, the use of the BTK inhibitor ibrutinib has been shown to have the potential to reverse these effects (Bond and Woyach, 2019).

Variations in the binding of BTKi to BTK may manifest in distinct patterns of BTK mutations, particularly in the context of prolonged treatment required for CLL. Recently published studies in (Blombery et al., 2022; Naeem et al., 2023; Wang et al., 2022) have elucidated the specific mutation patterns associated with the covalent BTK inhibitor zanubrutinib and the noncovalent BTK inhibitor pirtobrutinib in CLL patients. Patients undergoing treatment with the pioneering BTK inhibitor, ibrutinib (Kadri et al., 2017; Gángó et al., 2020; Ahn et al., 2017), or the second-generation BTK inhibitor acalabrutinib (Woyach et al., 2019), tend to experience relapse predominantly characterized by BTK C481S mutations. These mutations lead to decreased binding affinity of the covalent BTK inhibitor to BTK, contributing to therapeutic challenges. The BTK mutational patterns observed in zanubrutinib (Handunnetti et al., 2019) and pirtobrutinib differed significantly. For patients receiving zanubrutinib who were experiencing disease progression, there was a notable increase in BTK L528W mutations. Unlike the C481S mutation, which allows the kinase to retain its activity, the L528W mutation hinders both the binding of covalent inhibitors and adenosine triphosphate, resulting in the generation of a "kinase-dead" protein. Progressions associated with pirtobrutinib have also revealed non-C481 mutations within the kinase domain, including V416L, A428D, M437R, T474I, alongside L528W. Furthermore, in cases involving inactivating C481 mutations (C481F and C481Y), signaling to PLCG2 may persist through a kinase-independent activation of hematopoietic cell kinase (Dhami et al., 2022).

ARQ 531 is a reversible BTK inhibitor that retains activity in ibrutinib-resistant CLL, including models harboring the recurrent C481S BTK mutation and activating PLCγ2 mutations. By inhibiting BTK independently of C481 covalent binding and suppressing downstream BCR signaling through additional kinases such as SYK and LYN, ARQ 531 represents a promising clinical candidate for overcoming resistance mechanisms that limit current BTK-targeted therapies (Liu et al., 2015; Jones et al., 2017). New findings show that using ARQ 531 decreases the ability of BTK to perform its many functions in CLL cells, including its ability to regulate the BCR signaling pathway, support the viability of these cells, enable migration of CLL cells in and out of the bone marrow, and regulate the expression of several necessary genes in CLL cells. When tested in animal models of CLL and Richter's transformation, patients treated with ARQ 531 had better survival than patients treated with ibrutinib. Additionally, ARQ 531 was shown to inhibit the mutations that develop in BTK due to treatment resistance, providing a potential therapeutic option for CLL patients with specific genetic mutations associated with treatment resistance. (Reiff et al., 2018a). ARQ 531 was also shown to inhibit the viability of CLL cells, to inhibit the migration of CLL cells, and to inhibit the expression of genes needed for CLL, including mutations acquired during therapy with ibrutinib or in Richter's transformation; and, in addition, it was shown to be effective against both naive and resistant CLL. Studies using a mouse model demonstrated that mice treated with ARQ 531 survived longer than those treated with ibrutinib. Thus, these results indicate that treatment with ARQ 531 may be able to overcome the mechanisms of resistance seen with selected BTK inhibitors, indicating that further preclinical and early clinical study of ARQ 531 is needed, along with careful monitoring of safety and effectiveness, with the possible use as a broader-spectrum kinase inhibitor (Reiff et al., 2018b). Indirect inhibition of Syk could lead to the investigation of ARQ 531 against acute myeloid leukemia, where Syk plays a key role in leukemia progression. The rationale for further evaluation of the effect of ARQ 531 on other types of cancer exists. ARQ, ARQ 531, or MK-1026 remains a very important investigational clinical candidate for overcoming drug-resistant CLL and B-cell malignancies (Boros et al., 2015; Carnevale et al., 2013). Its primary clinical value lies in its unique mechanism as a reversible, non-covalent BTK inhibitor, which allows it to remain effective even after patients develop resistance to standard covalent inhibitors like ibrutinib (De Nigris et al., 2024; Woyach et al., 2022). Therefore, the current investigation employs an integrated approach combining genomic clinical mutation screening with molecular dynamics simulations to evaluate the impact of selected clinically lethal mutations on the binding of ARQ 531 through binding free energy calculations and essential dynamics analyses. This study aims to elucidate the binding mechanism and to examine whether the selected mutations may underlie potential resistance to ARQ 531, thereby providing insights that could support the development of improved BTK inhibitors and contribute to enhanced clinical outcomes.

2. Materials and methods

2.1. BTK sequence and 3D structure data collection

We performed systematic analyses of the retrieved mutations from the gnomAD database (an online repository of genetic variation information), and detailed information about each SNP reported in the BTK gene in humans (Karczewski et al., 2020). Additionally, the confirmed crystal structure of the BTK protein (**PDB ID:** 6E4F) was accessed and obtained from the Protein Data Bank (http://www.rcsb.org/), a repository for biomolecular structural data (Bender et al., 2017).

2.2. Prediction of deleterious non-synonymous single nucleotide polymorphisms (SNPs) in BTK

Several online servers, including PANTHER (Protein ANalysis THrough Evolutionary Relationships), PredictSNP, SNAP (screening for non-acceptable polymorphisms), SIFT (Sorting Intolerant from Tolerant), PhD-SNP (Predictor of human Deleterious Single Nucleotide Polymorphisms), MAPP (Multivariate Analysis of Protein Polymorphism), and PolyPhen-2 (Polymorphism Phenotyping version 2), were utilized for the prediction of the functional impact of non-synonymous single nucleotide polymorphisms (nsSNPs). PredictSNP, available at https://loschmidt.chemi.muni.cz/predictsnp1/, amalgamates various prediction tools to offer more reliable and accurate forecasts compared to individual tools. It leverages experimental annotations to enhance the predictions made by these integrated tools (Bendl et al., 2014). MAPP (http://mendel.stanford.edu/SidowLab/downloads/MAPP/) assesses the impact of all potential SNPs on protein function by examining the physicochemical variations within aligned protein sequences (Stone and Sidow, 2005). PhD-SNP (http://snps.biofold.org/phd-snp/phd-snp.html) categorizes non-synonymous SNPs (nsSNPs) as either disease-associated or neutral polymorphisms through

their scoring system (ranging from 0 to 1) (Capriotti et al., 2006). PolyPhen-2 (http://www.genetics.bwh.harvard.edu/pp2) determines whether a substitution of one amino acid for another in the protein will affect the protein's structure or functionalities, and score amendments with values of 0 to 1, the non-synonymous SNPs identified in this program represent potentially deleterious mutations, which are characterized by their higher scores through mutations compared to the score of 0 (indicating No loss of function effect from the substituted amino acid) (Adzhubei et al., 2010).

SIFT (http://sift.bii.a-star.edu.sg) is a tool that predicts the impact of a single amino acid substitution on protein functions from the physiochemical properties of the amino acids and the degree of similarity between the sequences containing each pair of amino acids of the two protein sequences receiving scores of 0 to 1 (0-0.05 = deleterious mutations, and 0.05-1 = non-deleterious mutations) (Sim et al., 2012). The SNAP prediction server (https://rostlab.org/services/snap) uses a neural network to provide predictions of the effect of amino acid sequence variants on protein function. With a prediction score range of −100 to 100, where −100 indicates the most likely neutral effect and 100 indicates a very high likelihood that a given single amino acid change will alter the function of the resulting native protein (Hecht et al., 2015). The PANTHER-PSEP tool, available at http://www.pantherdb.org/tools/csnpScoreForm.jsp, also utilizes a unique predictive method based on evolutionary conservation to predict the pathogenicity of non-synonymous mutations associated with human diseases It relies on a unique metric linked to evolutionary conservation. By utilizing related proteins, it reconstructs probable sequences of ancestral proteins at various points in a phylogenetic tree. This enables tracing the historical path of each amino acid, estimating the duration it has remained preserved across ancestral states, and revealing its importance in disease prediction.

2.3. Effect of mutation on structure stability

Protein stability is critical to many biological functions, as proteins make up the majority of cells and organisms and their three-dimensional structures directly impact their ability to function. To assess how the identified mutations affect the stability of BTK 1, we utilized the I-Mutant 2.0 server (https://folding.biofold.org/i-mutant/i-mutant2.0.html), which requires the wild-type and mutant residue positions to determine the effect the substitutions have on the protein's stability. A ΔΔG score that is positive indicates an increase in stability, while a negative ΔΔG score indicates a decrease in stability (Calabrese et al., 2009). DynaMut2 (http://biosig.unimelb.edu.au/dynamut2), was used to assess protein stability based on structural and dynamic changes through the use of Normal Mode Analysis (NMA) (Rodrigues et al., 2021). A negative ΔΔG score indicates destabilization, while a positive ΔΔG score indicates stabilization. Finally, mutations identified as destabilizing by the previous two servers were further analyzed using the mCSM server. This additional analysis, employing a graph-based signature approach (http://biosig.unimelb.edu.au/mcsm/stability), evaluated the impact of mutants on protein stability by scrutinizing Relative Solvent Accessibility (RSA) and ΔΔG (difference in free energy) values for each mutation (Pires et al., 2014). The mCSM server provided valuable insights into how these mutations influenced the stability of the proteins under examination. Additionally, PremPLI has been employed in addition to the evaluation of how missense variants affect the ability of ARQ531 to bind BTK (https://lilab.jysw.suda.edu.cn/research/PremPLI/). A more stringent than usual cutoff (>0.9) as per prediction scores has been applied, by default, to shortlist those mutations that cause more structural and energetic perturbation, for subsequent multiple molecular simulations, instead of their potential clinical frequency or severity. The PredictSNP consensus scores ranging from −1 to +1 were used to determine the order in which the mutations were deemed functionally damaging, with the higher the positive score, the greater the confidence that functional damage will be produced (Sun et al., 2021). Structural stability was evaluated using mCSM, where increasingly negative ΔΔG values correspond to stronger destabilizing effects. Binding relevance was assessed using PremPLI, with higher scores indicating a greater likelihood of protein–ligand interaction disruption. Thresholds were selected to identify mutations ranking highest across these complementary metrics, rather than as absolute biological cutoffs.

2.4. Modeling of wild type and variants of BTK

The structure of the BTK1 protein was obtained from the PDB (ID: 5P9J) and then refined using Chimera software (Webb and Sali, 2016). Following this, mutations were introduced into the wild-type (WT) structure based on the most detrimental mutants identified in earlier analyses. These mutant protein structures were generated by implementing specific point mutations within the initial wild-type (WT) protein structure using Chimera software. Protonation states of titratable residues were assigned at physiological pH using an automated pKa/protonation workflow using PROPKA, followed by visual inspection of residues in the binding site. Catalytic or functional residues were kept in their dominant states under near-neutral conditions: Lys430 was modeled protonated (Lys^{+}), and Asp539 was modeled deprotonated (Asp^{-}) throughout all systems. Histidines were assigned as HIE depending on their local hydrogen-bonding environment. Any missing side chains were rebuilt during protein preparation. No long unresolved loops near the binding site were present, and energy minimization before solvation was carried out. Crystallographic waters were treated consistently across systems: waters directly mediating ligand–protein interactions in the binding pocket were retained, while non-conserved bulk waters were removed before solvation. The interactions were visualized using PyMOL and Discovery Studio Visualizer (for visualization purposes only).

2.5. Molecular dynamics simulation

We utilized the AMBER20 package along with the FF19SB force field (Salomon-Ferrer et al., 2013; Case et al., 2005) to conduct molecular dynamics simulations, examining the structural confirmations of generated complexes within a dynamic setting. Both the wild-type and mutant complexes underwent solvation within a TIP3P water box (10 Å) and were neutralized through the addition of Na+ and Cl-ions. ARQ 531 was parameterized using a General AMBER Force Field small-molecule model (GAFF/GAFF2). Atomic partial charges were derived by a quantum-mechanics–based procedure: the ligand geometry was optimized, and electrostatic potential (ESP) charges were generated and fit using RESP (alternatively, AM1-BCC if used; state one and keep consistent across all mutants). The ligand topology and parameters were generated using standard AMBER tools (Antechamber/Parmchk) and combined with the protein parameters. Additionally, a two-step gentle minimization process was employed to eliminate any unfavorable clashes present within the system. In the initial phase, we used the steepest descent and conjugate gradient minimization techniques in two different steps, performing 12,000 and 6000 minimization cycles, respectively (Watowich et al., 1988; Meza, 2010). Following the minimization, an equilibration of each system after heating from 1 atm to 300K was achieved. Finally, a 300 ns long molecular dynamics simulation (MDS), production phase, was completed, using the particle-mesh Ewald algorithm for treatment of long-range interactions and the SHAKE algorithm for covalent bond treatment (Babar et al., 2022; Salomon-Ferrer et al., 2013).

2.6. Post-simulation analysis

We utilized the CPPTRAJ and PTRAJ tools to assess how the BTK1 mutants impacted the dynamic stability, flexibility, compactness, and hydrogen bonding network within both the wild-type and mutant complexes (Roe and Cheatham, 2013). To gauge structural compactness

over the simulation period, we computed the radius of gyration (Rg). Additionally, we examined structural dynamic stability by determining the Root Mean Square Deviation (RMSD).

The RMSD was calculated by using the following equation:

$$RMSD = \sqrt{\frac{\sum_{i=0}^{N}\left[m_i * \left(X_i - Y_j\right)^2\right]}{M}} \quad \text{(i)}$$

Where N is the number of atoms, m_i is the mass of atom *i*, X_i is the coordinate vector for target atom *i*, Y_i is the coordinate vector for reference atom *i*, and M is the total mass. However, for assessing structural flexibility at the residue level, we computed the Root Mean Square Fluctuation (RMSF). This measurement quantifies the fluctuation of residues over the simulation period rather than focusing on positional variances across the entire complex.

2.7. Binding free energy calculation

The Binding Free Energy (BFE) of the wild type (WT) and mutant complexes was calculated via MM/GBSA in a time-dependent pattern. The MMGBSA.py script utilized here is widely regarded as one of the most effective methods for calculating binding free energies associated with a wide variety of biological interactions, including protein-protein, protein-nucleic acid, and protein-ligand (Hou et al., 2011; Shahraki et al., 2024; Taghizadeh et al., 2024, 2025; Hoseini et al., 2025). The MMGBSA.py script facilitated the binding free energy calculations for the WT and mutant complexes, and provided the components of binding free energy, including electrostatic, van der Waals (vdW), surface area, and Generalized Born solvation energy. The binding free energy was calculated by using the following equation:

$$\text{“}\Delta G(bind) \quad \Delta G\ complex - \Delta G\ receptor + \Delta G\ ligand\text{”}$$

While each component of the total energy was calculated by using the following equation:

$$\text{″}G \quad Gbond + Gele + GvdW + Gpol + Gnpol\text{″}$$

G_{bond}, G_{ele}, and G_{vdW} denote the interactions involving bonding, electrostatic forces, and van der Waals forces, respectively. In contrast, G_{pol} and G_{npol} signify the free energies associated with polar and non-polar components.

2.8. Dimensionality reduction through principal component analysis (PCA)

PCA is a method commonly used to analyze datasets generated by molecular dynamics simulations of biological macromolecules, which reduces the dimensionality of configurational space, both qualitatively and quantitatively, and consequently facilitates the interpretation of large and complex datasets. It is a widely used approach due to its simple use and comparatively low computational cost. PCA analysis is typically performed after removing global translational and rotational motions through structural alignment of the trajectories. The PCA method consists of a series of steps. The first step is to construct and align the coordinates, followed by the calculation of the covariance matrix describing atomic positional fluctuations. The eigenvectors and eigenvalues of the covariance matrix are then determined, and the eigenvectors associated with the highest eigenvalues are regarded as the principal components. These principal components represent the dominant collective motions and major conformational fluctuations of the system and are subsequently used for trajectory analysis and visualization (Kurita, 2019). The following mathematical representation can be used to calculate the covariance matrix (C) from a set of n-dimensional vectors (x_i):

$$C = \frac{1}{N}\, x \sum (i = 1\ to\ N)\left[\ (x_i - \mu)\ x\ (x_i - \mu)^T\right] \quad \text{(x)}$$

Where:

N = number of vectors, μ = mean vector, while T designates the transpose operation.

The following equation can be used to obtain the eigenvectors (V) and eigenvalues (λ) of the covariance matrix C.

$$C\ x\ V = \lambda\ x\ V \quad \text{(xi)}$$

Where *V* represents the eigenvector matrix, while the diagonal matrix of eigenvalues is represented by λ, the eigenvectors with the highest corresponding eigenvalues represent the principal components of the system. To determine the primary motions in our systems, we calculated two principal components, PC1 and PC2, using the unsupervised PCA technique. To do this, we used the CPPTRAJ module of the AMBER21 software package and considered the entire simulation trajectory, which consisted of 10,000 frames for each system. The spatial covariance matrix was computed using the eigenvectors of the eigenvalue matrix and the atomic coordinates of the atoms that made up the molecular dynamics (MD) simulations. An orthogonal coordinate transformation was performed to generate a diagonal eigenvalue matrix from the spatial covariance matrix. The principal components of the systems were determined using the eigenvectors and eigenvalues and were used to represent and illustrate the significant movements that occurred during the simulations and to identify possible conformational transitions.

2.9. Free Energy Landscape (FEL) of the molecular simulation trajectories

The FEL provides important information about the thermodynamics and kinetics of molecular events occurring in solution, and is a useful method for assessing and analyzing biomolecular events, including the folding, aggregation, and binding of molecules (Maisuradze et al., 2010). The FEL is a combination of the free energy (f) of a molecule and the configuration space of that molecule, using collective coordinates (r) to describe the configuration space of the entire molecule being studied. In this study, we used the two principal components identified from PCA to describe and identify the conformational transitions that occurred for each of the systems studied and to provide a characterization of these transitions. The lowest energy minima were identified and represented in the FEL as valleys, while any meta-states that were separated by an energy barrier were shown as peaks.

3. Results and discussion

3.1. Analysis of the genomic mutations in the kinase domain of BTK

We retrieved a total of 74 non-synonymous deleterious substitutions in the kinase domain of BTK using PredictSNP, which combines multiple servers to determine the outcome of each substitution. To evaluate the structural impacts, the mutations were submitted, which reported that among the total 74 mutations, 62 are deleterious, while 12 were classified as non-deleterious (Table 1). PredictSNP, a consensus-based classifier, was used to assess the functional impact of nsSNPs. PredictSNP integrates multiple established prediction tools and applies a normalized consensus scoring scheme to reduce individual method bias. Mutations with PredictSNP scores ranging from 0 to +1 were classified as deleterious, indicating a high likelihood of functional or structural disruption, whereas mutations with scores between −1 and 0 were classified as neutral. This consensus-based classification provides a robust preliminary screening framework, allowing prioritization of high-risk variants for downstream structural and molecular dynamics analyses. Using a standard criterion, the 12 non-deleterious substitutions were discarded, while the remaining 62 were further evaluated for the stability assessment using the mCSM server.

The mCSM server predicted 11 substitutions, including L528W, L512Q, F644S, W581R, A523E, E589G, W563L, P619S, R641H, P619T, and A508D as highly destabilizing. Only two substitutions, i.e., G462V

Table 1
Sequence-based functional analysis of the retrieved non-synonymous substitutions using various machine learning based algorithms. The outcomes, whether deleterious or non-deleterious, were assigned based on the PredictSNP score.

S. No	Variant	PredictSNP	MAPP	PhD-SNP	PolyPhen-1	PolyPhen-2	SIFT	SNAP	Outcomes
1	L408P	0.869	0.858	0.885	0.745	0.811	0.793	0.622	DELETERIOUS
2	G414R	0.869	0.877	0.885	0.745	0.811	0.793	0.848	DELETERIOUS
3	K430E	0.869	0.819	0.817	0.745	0.811	0.793	0.805	DELETERIOUS
4	E445D	0.869	0.657	0.858	0.745	0.647	0.793	0.805	DELETERIOUS
5	G462V	0.869	0.857	0.817	0.745	0.811	0.793	0.805	DELETERIOUS
6	G462D	0.869	0.877	0.817	0.745	0.811	0.793	0.885	DELETERIOUS
7	Y476D	0.869	0.843	0.817	0.745	0.601	0.793	0.805	DELETERIOUS
8	M477R	0.869	0.857	0.858	0.745	0.811	0.793	0.848	DELETERIOUS
9	C481R	0.869	0.589	0.773	0.745	0.650	0.793	0.805	DELETERIOUS
10	C481F	0.869	0.588	0.676	0.745	0.811	0.793	0.720	DELETERIOUS
11	C481Y	0.869	0.571	0.817	0.745	0.811	0.793	0.720	DELETERIOUS
12	C502W	0.869	0.588	0.817	0.745	0.811	0.793	0.720	DELETERIOUS
13	C502F	0.869	0.509	0.875	0.745	0.811	0.793	0.622	DELETERIOUS
14	C506R	0.869	0.877	0.858	0.745	0.675	0.793	0.805	DELETERIOUS
15	C506Y	0.869	0.761	0.875	0.745	0.811	0.793	0.720	DELETERIOUS
16	A508D	0.869	0.877	0.858	0.745	0.650	0.793	0.720	DELETERIOUS
17	M509V	0.869	0.807	0.817	0.745	0.811	0.793	0.720	DELETERIOUS
18	M509I	0.869	0.766	0.875	0.745	0.811	0.793	0.805	DELETERIOUS
19	L512Q	0.869	0.807	0.875	0.745	0.811	0.793	0.622	DELETERIOUS
20	L512P	0.869	0.819	0.817	0.745	0.811	0.793	0.805	DELETERIOUS
21	L518R	0.869	0.920	0.858	0.745	0.675	0.793	0.622	DELETERIOUS
22	R520Q	0.869	0.589	0.858	0.745	0.811	0.793	0.848	DELETERIOUS
23	D521H	0.869	0.783	0.817	0.745	0.811	0.793	0.885	DELETERIOUS
24	D521G	0.869	0.858	0.817	0.745	0.811	0.793	0.848	DELETERIOUS
25	D521N	0.869	0.657	0.817	0.745	0.601	0.528	0.805	DELETERIOUS
26	A523E	0.869	0.621	0.875	0.745	0.542	0.793	0.805	DELETERIOUS
27	R525G	0.869	0.807	0.885	0.745	0.675	0.793	0.720	DELETERIOUS
28	R525Q	0.869	0.409	0.885	0.745	0.811	0.793	0.720	DELETERIOUS
29	R525P	0.869	0.877	0.885	0.745	0.811	0.793	0.805	DELETERIOUS
30	N526K	0.869	0.843	0.875	0.745	0.811	0.793	0.848	DELETERIOUS
31	L528W	0.869	0.775	0.875	0.745	0.811	0.793	0.805	DELETERIOUS
32	V535F	0.869	0.484	0.875	0.745	0.601	0.793	0.556	DELETERIOUS
33	L542P	0.869	0.842	0.858	0.745	0.542	0.793	0.622	DELETERIOUS
34	R544G	0.869	0.633	0.817	0.745	0.562	0.793	0.805	DELETERIOUS
35	R562W	0.869	0.751	0.858	0.745	0.811	0.793	0.720	DELETERIOUS
36	R562P	0.869	0.843	0.817	0.745	0.811	0.793	0.556	DELETERIOUS
37	W563L	0.869	0.877	0.817	0.745	0.811	0.793	0.805	DELETERIOUS
38	E567K	0.869	0.775	0.817	0.745	0.811	0.793	0.848	DELETERIOUS
39	S578Y	0.869	0.657	0.858	0.745	0.811	0.793	0.720	DELETERIOUS
40	W581R	0.869	0.877	0.817	0.745	0.811	0.793	0.848	DELETERIOUS
41	A582V	0.869	0.633	0.773	0.745	0.647	0.793	0.720	DELETERIOUS
42	F583S	0.869	0.807	0.875	0.745	0.811	0.793	0.805	DELETERIOUS
43	E589G	0.869	0.857	0.875	0.745	0.811	0.793	0.805	DELETERIOUS
44	E589K	0.869	0.775	0.817	0.745	0.811	0.528	0.848	DELETERIOUS
45	E589D	0.869	0.751	0.858	0.745	0.675	0.793	0.805	DELETERIOUS
46	S592P	0.869	0.807	0.858	0.745	0.562	0.793	0.848	DELETERIOUS
47	G594E	0.869	0.877	0.858	0.745	0.811	0.793	0.848	DELETERIOUS
48	G594R	0.869	0.877	0.817	0.745	0.811	0.793	0.848	DELETERIOUS
49	Y598C	0.869	0.573	0.858	0.745	0.811	0.793	0.848	DELETERIOUS
50	G613D	0.869	0.766	0.817	0.745	0.647	0.793	0.805	DELETERIOUS
51	P619S	0.869	0.571	0.773	0.745	0.811	0.793	0.622	DELETERIOUS
52	P619A	0.869	0.657	0.858	0.594	0.675	0.793	0.622	DELETERIOUS
53	P619T	0.869	0.560	0.817	0.594	0.675	0.793	0.622	DELETERIOUS
54	A622P	0.869	0.842	0.885	0.745	0.811	0.793	0.556	DELETERIOUS
55	M630K	0.869	0.920	0.885	0.745	0.675	0.793	0.805	DELETERIOUS
56	M630T	0.869	0.807	0.885	0.594	0.675	0.793	0.720	DELETERIOUS
57	C633Y	0.869	0.783	0.817	0.745	0.811	0.793	0.848	DELETERIOUS
58	R641H	0.869	0.718	0.817	0.745	0.811	0.793	0.848	DELETERIOUS
59	R641C	0.869	0.857	0.875	0.745	0.811	0.793	0.848	DELETERIOUS
60	F644S	0.869	0.766	0.817	0.745	0.811	0.528	0.805	DELETERIOUS
61	F644L	0.869	0.633	0.885	0.594	0.601	0.528	0.805	DELETERIOUS
62	L647P	0.869	0.807	0.885	0.745	0.811	0.793	0.805	DELETERIOUS

and S592P, were classified as stabilizing, while the rest were reported to be destabilizing. Usually, the destabilizing mutations cause variations in ligand binding and, therefore, were further considered for subsequent analysis. Using the PremPLI server, these substitutions were analyzed to determine whether they affect ligand binding. Using the co-crystallized complex of BTK with ARQ 531 (ARQ 531), the impact of these substitutions was analyzed. All the substitutions were reported to affect the binding; however, we selected only those whose predicted affinity change was reported to be higher than 0.90. This criterion reported only six substitutions, i.e., C481R with the predicted score of 1.28, Y476D with the predicted score of 1.25, L408P with the predicted score of 1.16, C481Y with the predicted score of 1.06, L542P with the predicted score of 1.01, and M477R with the predicted score of 0.97, which were selected for the functional analysis. The mCSM and PremPLI results for the selected substitutions are shown in Table 2.

Table 2
List of reported clinical mutations screened for structure-based stability analysis and binding variations caused by these substitutions.

S. No	Mutations	PremPLI	Interface	mCSM	Outcome
1	C481R	1.28	Yes	−0.613	Destabilizing
2	Y476D	1.25	Yes	−1.368	Destabilizing
3	L408P	1.16	Yes	−0.731	Destabilizing
4	C481Y	1.06	Yes	−0.93	Destabilizing
5	L542P	1.01	Yes	−1.319	Destabilizing
6	M477R	0.97	Yes	−0.983	Destabilizing
7	L528W	0.90	Yes	−1.798	Destabilizing
8	C481F	0.72	No	−3.438	Highly Destabilizing
9	F583S	0.71	Yes	−1.001	Destabilizing
10	G462V	0.70	No	0.165	Stabilizing
11	F644S	0.67	No	−3.033	Highly Destabilizing
12	K430E	0.67	Yes	−0.645	Destabilizing
13	E589K	0.66	No	−1.702	Destabilizing
14	G594E	0.64	No	−1.493	Destabilizing
15	L512P	0.64	No	−1.178	Destabilizing
16	C506R	0.63	No	−1.23	Destabilizing
17	G594R	0.63	No	−1.046	Destabilizing
18	G414R	0.62	No	−0.851	Destabilizing
19	C502W	0.61	No	−1.523	Destabilizing
20	C506Y	0.61	No	−0.858	Destabilizing
21	W581R	0.59	No	−2.176	Highly Destabilizing
22	A523E	0.58	No	−2.644	Highly Destabilizing
23	G462D	0.57	No	−1.731	Destabilizing
24	C502F	0.54	No	−1.055	Destabilizing
25	G613D	0.54	No	−1.356	Destabilizing
26	Y598C	0.54	No	−1.347	Destabilizing
27	L647P	0.53	No	−1.07	Destabilizing
28	R544G	0.53	No	−1.096	Destabilizing
29	E589G	0.52	No	−2.728	Highly Destabilizing
30	R525G	0.52	No	−1.034	Destabilizing
31	E589D	0.50	No	−1.988	Destabilizing
32	L512Q	0.50	No	−2.166	Highly Destabilizing
33	P619A	0.50	No	−1.992	Destabilizing
34	V535F	0.50	No	−1.356	Destabilizing
35	C633Y	0.49	No	−1.195	Destabilizing
36	R525P	0.49	No	−0.065	Destabilizing
37	S578Y	0.48	No	−1.063	Destabilizing
38	E567K	0.46	No	−1.243	Destabilizing
39	R525Q	0.46	No	−0.226	Destabilizing
40	D521N	0.42	No	−1.578	Destabilizing
41	D521G	0.40	No	−0.918	Destabilizing
42	M630K	0.40	No	−1.259	Destabilizing
43	R641C	0.40	No	−1.737	Destabilizing
44	W563L	0.40	No	−2.82	Highly Destabilizing
45	L518R	0.36	No	−0.784	Destabilizing
46	N526K	0.36	No	−0.117	Destabilizing
47	M630T	0.35	No	−1.808	Destabilizing
48	E445D	0.34	No	−1.195	Destabilizing
49	P619S	0.33	No	−2.53	Highly Destabilizing
50	R641H	0.33	No	−2.172	Highly Destabilizing
51	P619T	0.31	No	−2.128	Highly Destabilizing
52	M509V	0.30	No	−0.999	Destabilizing
53	D521H	0.29	No	−0.717	Destabilizing
54	R562P	0.21	No	−1.169	Destabilizing
55	S592P	0.21	No	0.21	Stabilizing
56	A582V	0.20	No	−0.514	Destabilizing
57	A622P	0.20	No	−0.113	Destabilizing
58	F644L	0.20	No	−1.52	Destabilizing
59	R520Q	0.20	No	−0.674	Destabilizing
60	R562W	0.20	No	−0.488	Destabilizing
61	M509I	0.17	No	−0.518	Destabilizing
62	A508D	0.12	No	−2.454	Highly Destabilizing

3.2. Structure retrieval and binding analysis of BTK-ARQ 531 complex

The co-crystallized complex of BTK with ARQ 531 was retrieved from the Protein Data Bank and analyzed for its interactions. The structure of the BTK-ARQ 531 reported several hydrogen bonding, π-π, π-alkyl, and hydrophobic interactions. As shown, Val458, Glu475, Met477, Cys481, and Asn484 are involved in hydrogen bonding. Among these, Leu408, Val416, Ala428, Leu460, Phe540, and Leu542 are involved in hydrophobic, π-π, π-alkyl, and π-cation interactions. The 3D structure of BTK-ARQ 531 complex, the 2D structure of ARQ 531, and the binding mode of ARQ 531 with BTK are shown in Fig. 1a-c. We identified the top mutants that affect the binding of ARQ 531 using the PremPLI server, which reported six mutations as key substitutions that may impact the binding of ARQ 531. To analyze the impact of the selected substitutions, we generated these mutants using the Chimera rotamers module and minimized them using the conjugate gradient and steepest descent algorithms. The L408P mutation was ranked as the top one affecting the binding of ARQ 531, which reported only two hydrogen bonds, involving Thr410 and Met477. Surprisingly, the other substitutions were observed to have diminished. The interaction pattern of the L408P mutant BTK and ARQ 531 is shown in Fig. 2a.

On the other hand, the interaction pattern for Y476D has been changed, but various contacts were observed, including the hydrogen bonds established by Thr474 and Met477. In the pi-pi and hydrophobic contacts, Val416, Ile429, Lys430, Cys481, Ile472, Leu460, and Phe540 are involved in the interaction with ARQ 531. The interaction pattern of the Y476D mutant BTK and ARQ 531 is depicted in Fig. 2b. The binding mode of C481R was reported to form only a single hydrogen bond with Thr410, along with several hydrophobic and other contacts. The hydrogen bond with Met477 was observed to have diminished. Among the others, Val416, Ala428, Lys430, Leu460, Arg481, Leu528, and Phe540 residues are involved in π-π, π-cation, and π-alkyl interactions. The interaction pattern of the C481R mutant BTK and ARQ 531 is shown in Fig. 2c. In the case of C481Y, Thr474 and Met477 were involved in hydrogen bonding, while Lys430, Met449, Ile472, Phe540, and Leu542 are involved in other interactions. The interaction pattern of the C481Y mutant BTK and ARQ 531 is shown in Fig. 2d. The M477R mutant enhances binding through four hydrogen bonds involving Gly409, Thr474, Arg477, and Asp539. Among the non-hydrogen bonding contacts Leu408, Val416, Lys430, Leu460, Ile472 and Phe540 are involved. The interaction pattern of the M477R mutant BTK and ARQ 531 is shown in Fig. 2e. Interestingly, the L542P variant was found to form only a single hydrogen bond with Gly480, as well as several π-π, π-cation, and hydrophobic interactions. The interaction pattern of the L542P mutant BTK and ARQ 531 is shown in Fig. 2f. This suggests that these mutations alter the binding of ARQ 531 and may not necessarily confer resistance; however, long-term molecular simulation-based results may determine the fate of this drug in the presence of these mutations. Therefore, to gain atomic-level insights, subsequent simulation analysis was performed to understand the impact of these substitutions on the binding of ARQ 531.

3.3. Structural stability assessment through RMSD

Investigation of the dynamic stability using the simulation trajectories reports essential knowledge regarding the functionality of a particular receptor. It has been the most widely used metric for determining the role of a specific mutation in the functional variation and structural stability paradigm. Considering the essential role of RMSD in deciphering critical knowledge regarding the protein structure, we also calculated this parameter as a function of time using the simulation trajectories. As shown in Fig. 3a, the dynamic stability of the wild type revealed a different pattern than that of the mutant. The complex started to increase the RMSD gradually and reached the equilibrium at 1.5 Å. The RMSD stabilized and maintained this level, indicating no significant perturbation during the simulation. An average RMSD of 1.5 Å was reported for the wild type. On the other hand, the RMSD pattern for the L408P reported different behavior than the wild type. The RMSD continues to increase and reaches its maximum level at 100 ns. The complex initially converged with the wild type; however, it then further increased and maintained a similar level of RMSD of 3.0 Å after 125 ns. The higher RMSD level in the L480P is probably associated with the reduced ligand structural constraints that consequently produced a dynamically more flexible conformation due to the induced mutation. The RMSD graph for the L408P is shown in Fig. 3a. Moreover, the Y476D

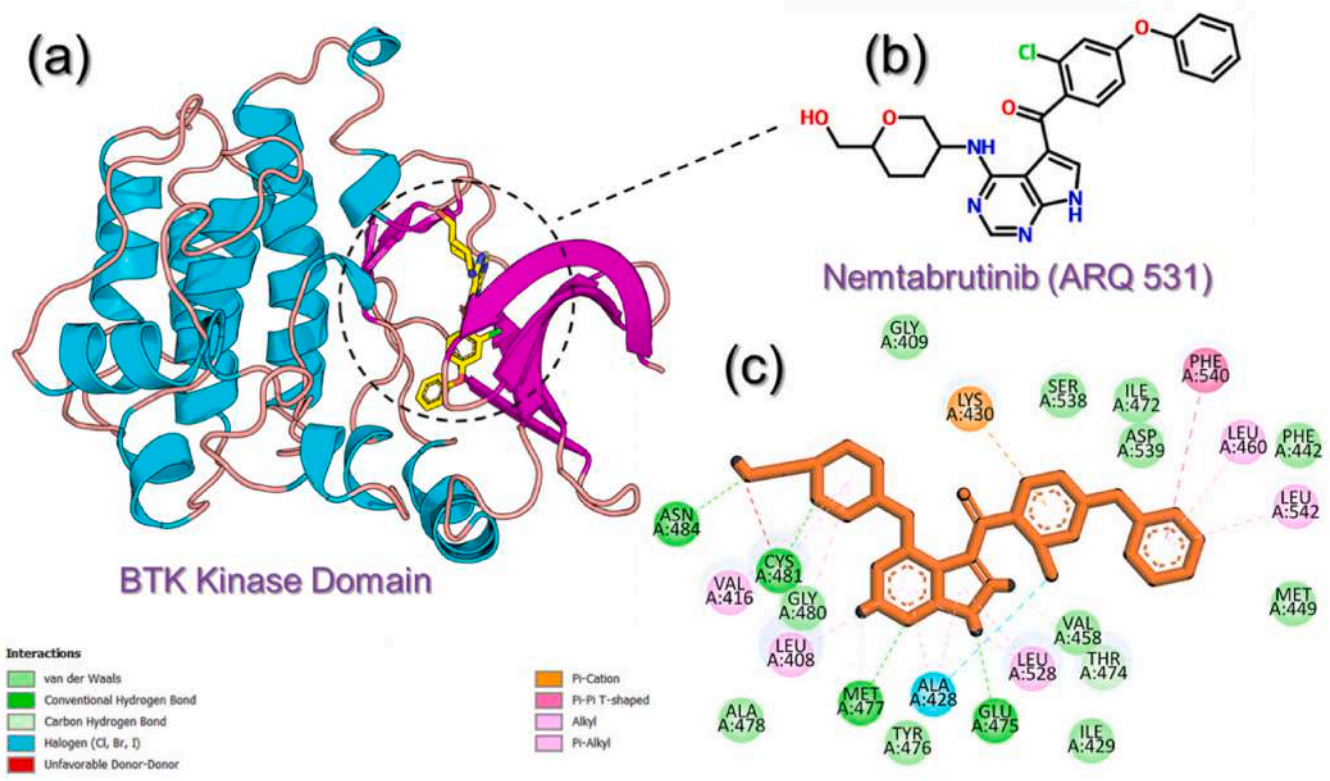

Fig. 1. Structure of BTK and binding analysis. (a) shows the 3D structure of BTK in complex with ARQ 531, **(b)** shows the 2D structure of ARQ 531, while **(c)** shows the 2D interaction pattern of ARQ 531 with BTK. The cyan colour represents the helix distribution in the BTK kinase domain, orange represents loops, and beta-sheets are represented as magenta.

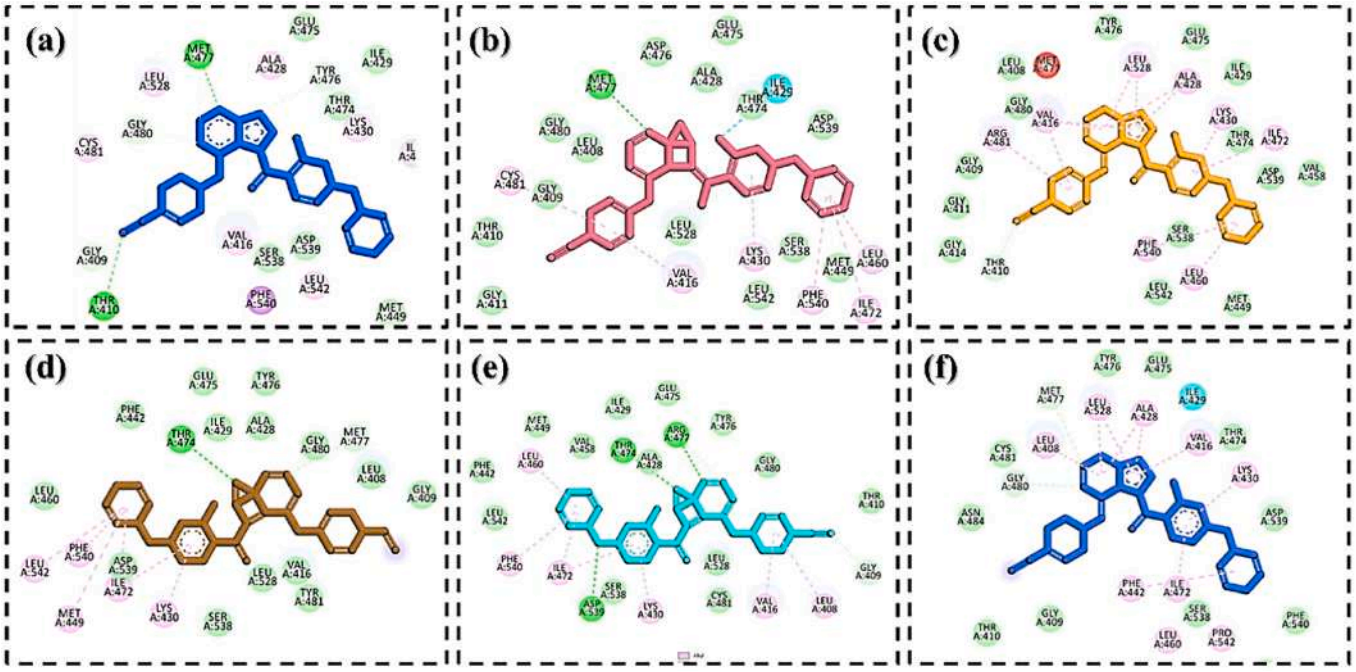

Fig. 2. Interaction analysis of the ARQ 531 with the mutant BTK. (a) shows the interaction of ARQ 531 with the L408P mutant of BTK, **(b)** interaction pattern of ARQ 531 with Y476D, **(c)** shows the interaction pattern of ARQ 531 with C481R, **(d)** shows the interaction pattern of ARQ 531 with C481Y, **(e)** shows the interaction pattern of ARQ 531 with M477R. In contrast, **(f)** shows the interaction pattern of ARQ 531 with the L542P mutant of BTK. Ligand–protein interactions are color-coded as follows: hydrogen bonding (green/light green), van der Waals (pale green), halogen bonds (cyan), π–cation (orange), π–π T-shaped (magenta), alkyl/π–alkyl (pink/light purple), and unfavorable donor–donor contacts (red).

variant exhibited a similar RMSD pattern to the wild type, with no significant perturbation, and achieved stability at the same points. The

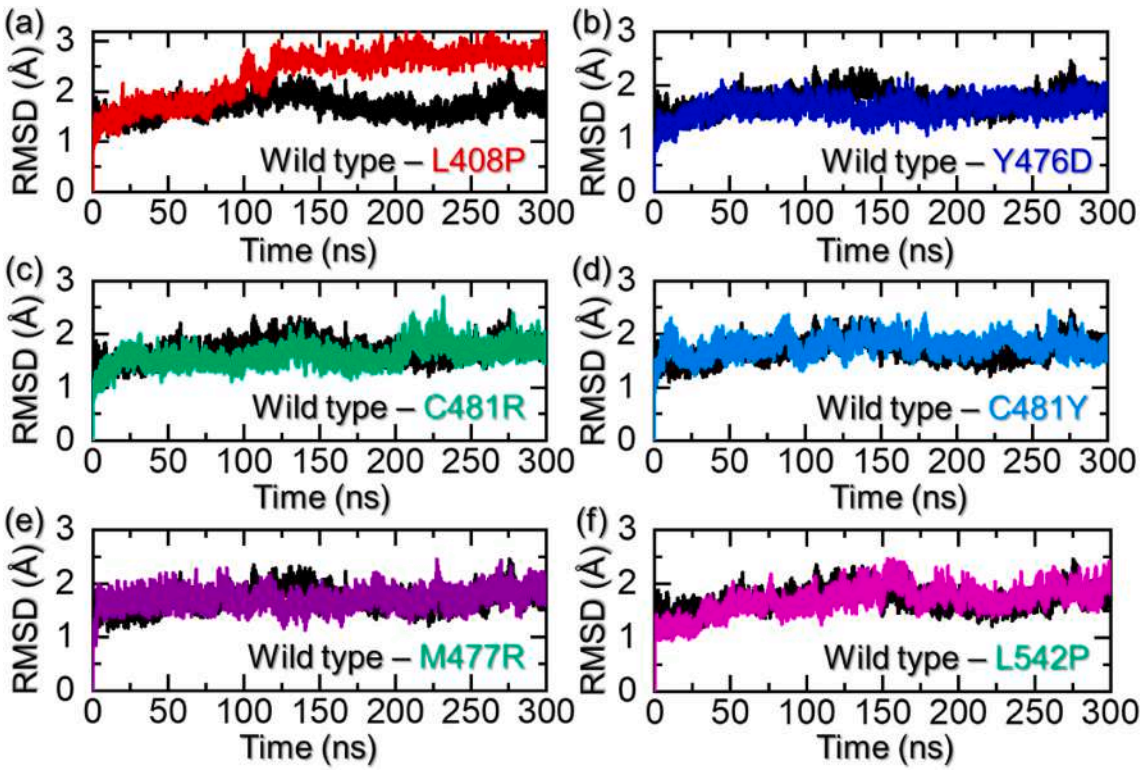

Fig. 3. Dynamic stability assessment through Cα RMSD of the full protein to assess global conformational stability, as a function of time. (a) shows the RMSD results for the wild type and L408P, **(b)** shows the RMSD results for the wild type and Y476D, **(c)** shows the RMSD results for the wild type and C481R, **(d)** shows the RMSD results for the wild type and C481Y, **(e)** shows the RMSD results for the wild type and M477R, while **(f)** shows the RMSD results for the wild type and L542P.

system has already achieved stability, and thus, there is sufficient time for improved sampling. The RMSD results for the Y476D are shown in Fig. 3b. Furthermore, the C481R exhibited a wave-like behavior, characterized by an increase and a decrease in the RMSD pattern. The RMSD reported a similar level to the wild type. Still, an abrupt increment in the RMSD between 200 and 250ns shows that dynamic variation caused by the induced mutation in the structure may consequently cause functional variation. The RMSD results for the C481R are shown in Fig. 3c. Moreover, the C481Y, a gatekeeper residue involved in resistance to the covalent drug ibrutinib, exhibited a continuous increase and decrease in the RMSD during the simulation. Although maintaining a similar level to the wild type, this mutation induced conformational changes that force the complex to attain stability but fail to stabilize at a particular time point. The structure requires a longer simulation run to achieve optimal sampling and stability. The results are shown in Fig. 3d. In the case of M477R, although the complex exhibited dynamically different behavior, it converged with the wild type, showing no significant structural perturbation and therefore demonstrating minimal variations. The behavior of the M477R is more similar to that of the wild type. The RMSD results for the M477R are shown in Fig. 3e. On the other hand, L542P began to increase the RMSD gradually and reached its maximum level at 160 ns. The RMSD then decreased and subsequently attempted to stabilize, but minor deviations were still reported. The RMSD results for the L542P are shown in Fig. 3f. Except for L408P, most mutants exhibit RMSD profiles comparable to the WT, indicating preserved global structural stability, while mutation-specific effects are more evident in local flexibility and interaction analyses. This suggests that the overall structural stability is maintained across the variants. Although 300 ns may not capture rare long-timescale kinase transitions, including full activation-loop or DFG motif switching, the trajectories reached stable plateaus with converged local descriptors, supporting robust comparative analysis of mutation-induced binding-site dynamics within the inhibitor-bound state explored here. However, distinct differences in binding interactions indicate that mutations alter the local environment within the binding site, potentially influencing the binding modes of ARQ531.

3.4. Structural compactness analysis

Analysis of structural compactness is considered one of the primary parameters for showing the binding variations caused by induced mutations. It shows the binding and unbinding events that occurred during the simulation. As shown in Fig. 4a, the wild type reported a uniform pattern of the Rg with no significant increase or decrease in the protein size; however, the L408P determined a similar pattern as the RMSD results. The Rg continues to increase after 55ns and maintains a uniform level throughout the simulation. This indicates that the mutation causes variations in structural properties, which in turn result in functional variance. The Rg results for the L408P are shown in Fig. 4a.

The Rg results for the Y476D showed a gradual decrease in the Rg from a value of 18.80 Å, with the lowest point reached at 80 ns, and maintained minimal deviation values during the simulation (Fig. 4b). The Rg graph for Y476D displayed a more compact characteristic, in which the Rg increased between 95 and 100ns and then decreased back, maintaining the uniform trajectory until the end of the simulation. The highest value for C481R was achieved at 160 ns before the Rg decreased again (Fig. 4c). The difference between the C481R and the C481Y is that the C481Y showed a small increase and decrease in the Rg until the 100 ns time point, where the C481Y began to show an increase in compactness that continued until 120ns and then converged with the wild type during the rest of the simulation and is shown in Fig. 4d.

The M477R appears to be similar to the RMSD results of the two previous mutants in that it was compact and stable in Rg, began at 18.80 Å, and had its lowest value at 120 ns. The Rg for L542P was similar to

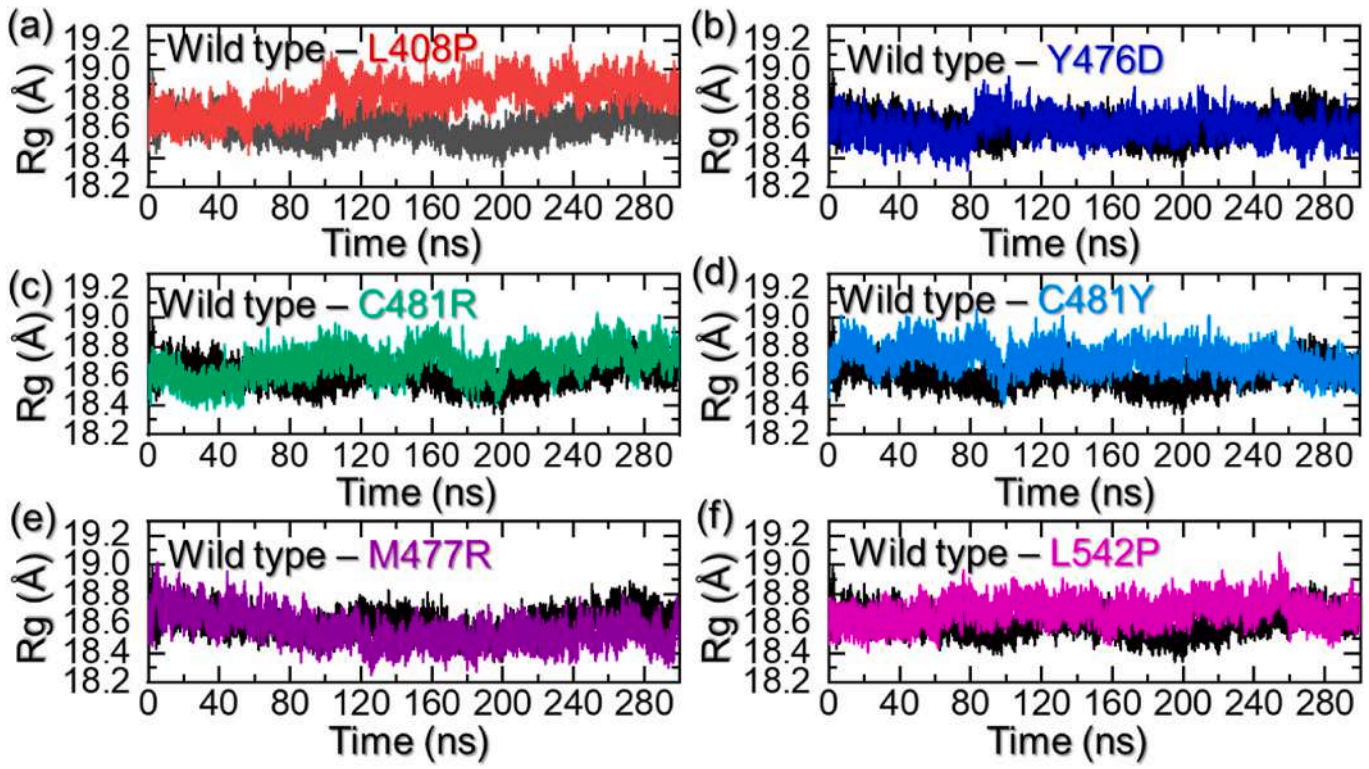

Fig. 4. Structural compactness assessment through Rg calculation as a function of time. **(a)** shows the Rg results for the wild type and L408P, **(b)** shows the Rg results for the wild type and Y476D, **(c)** shows the Rg results for the wild type and C481R, **(d)** shows the Rg results for the wild type and C481Y, **(e)** shows the Rg results for the wild type and M477R, while **(f)** shows the Rg results for the wild type and L542P.

M477R and the wild-type complex, which reported a compact and stable trajectory. Fig. 4e and f demonstrate the Rg results for both M477R and L542P. Although there were small differences observed in Rg values as compared to wild-type complexes and mutant types during the simulation, the degree of compactness for the protein structure remained similar across all tested mutant types. The minor differences between the Rg values of wild-type complexes and the tested mutant complex may indicate additional subtle conformational changes within the binding site. The minor differences in Rg values may also indicate potential influences on the flexibility and access to the binding site for drugs, which impacts the interactions between drugs and the protein across both wild-type and mutant protein variants.

3.5. *Residues flexibility analysis*

The assessment of the flexibility of a residue is an important indicator of how this flexibility contributes to many different roles, such as drug binding, molecular recognition, cascade signaling, protein coupling, enzyme specificity, and design and substrate specificity. The application of the flexibility index can provide insights into key mechanisms, including opening and closing a loop or pushing a substrate, which is a function of the flexibility of the protein. Highly dynamically flexible regions have been examined, and the data indicate that, except for L408P, there is convergence of the flexibility index across all the complexes. Further evaluation of the root-mean-square fluctuations (RMSFs) suggests that the majority of the systems demonstrate similar amplitudes of fluctuation (1–3 Å), whereas the L408P mutant is noted to have an elevation in atomic movement, especially at residues 401–425 and 541–560, consistent with localized destabilization of secondary structural elements proximal to the active-site pocket. Notably, residues 401–425 correspond to β-sheet and loop segments adjacent to the catalytic cleft and hinge region, while residues 541–560 map to the distal C-terminal loop region located near the regulatory lobe of the kinase domain. The enhanced flexibility identified in L408P is likely a result of the helix-breaking properties of proline disrupting the hydrogen bonding to the backbone and thereby inducing higher conformational entropy, which may in turn weaken the coordination of the substrate. The RMSF data are illustrated in Fig. 5a, which indicates which regions of dynamic flexibility exhibit the most fluctuation during the simulations. These flexible regions were projected on the three-dimensional (3D) structure of the BTK protein (Fig. 5b). The multi-representation of the residues, including residues 401–425, 466–471, 490–501, and 541–560, contains the active site of the BTK receptor, and these regions have been characterized as exhibiting a greater degree of dynamic flexibility than the other regions, offering a greater opportunity for ligand recognition and functional diversity. Specifically, residues 466–471 lie within a loop proximal to the ATP-binding pocket, while residues 490–501 encompass a segment of the activation-loop region that modulates access to the catalytic site. The wild-type BTK demonstrates tightly constrained dynamics in the catalytic regions, demonstrating a greater stability of substrate-binding conformations; therefore, numerous mutants demonstrate a unique pattern of fluctuation. Mutants C481R and C481Y, which are located adjacent to the ATP-binding pocket, demonstrate slight increases in RMSF that indicate greater fluctuation at residues 490–501, indicating the increasing complexity of activation-loop dynamics that may impact the binding of inhibitors. Y476D shows slight fluctuation around residues 466–471; the flexibility appears to be secondary to an electrostatic perturbation due to the substitution of aspartate. However, M477R remains relatively stable, exhibiting wild-type-like RMSF; therefore, it appears that M477R is minimally displaced compared to the original base protein topology, whereas L542P exhibits greater flexibility in the final loop (541–560) and demonstrates increased flexibility due to a decreased level of structural anchoring and increased amount of loop breathing. The information presented illustrates various ways that specific substitutions of residues can affect the overall dynamics of the kinase region,

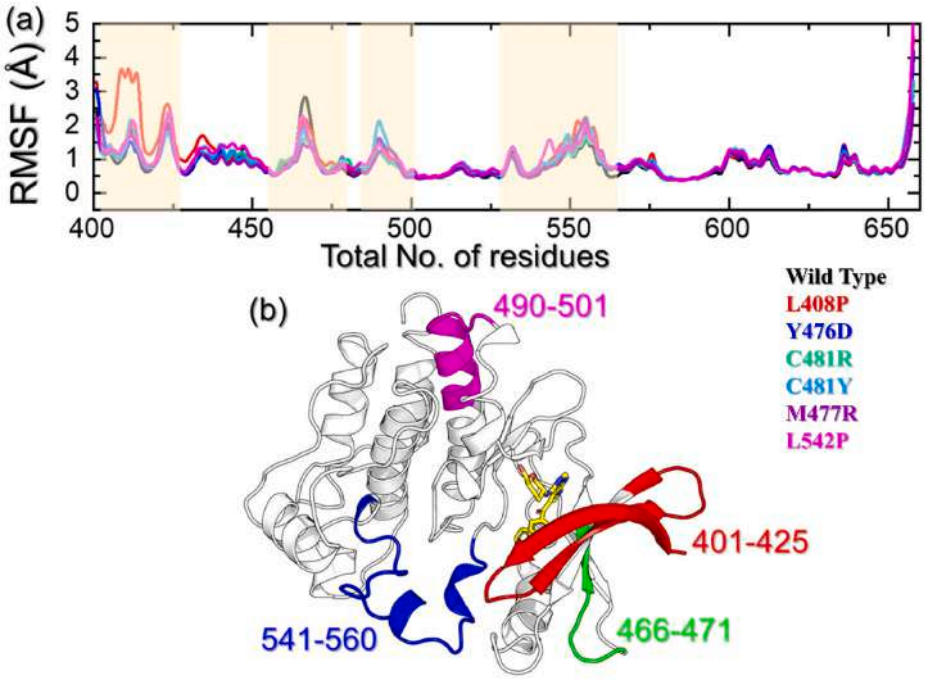

Fig. 5. Dynamic Flexibility analysis of the wild type and mutant complexes. (a) shows the RMSF graphs for the wild type and mutants, while **(b)** shows the regions that determined the higher flexibility during simulation.

contributing to the balance of rigidity that is necessary for successful catalytic activity with the requisite amount of flexibility for the proper accommodation of substrates within the active site. As evident from this study, the comprehensive mapping of flexible residues within the vicinity of the active site demonstrates their direct contribution to molecular recognition and conformational transitions that are integral to BTK functionality. To improve structural context, key regions including the activation loop, hinge region, and ATP-binding pocket are explicitly labeled in Fig. 5b.

3.6. Hydrogen bonding analysis

Macromolecular complexes, primarily composed of biological molecules, are influenced by numerous factors, among which hydrophobic contacts and hydrogen bonding are significant. The water molecules enriched environment of the protein interface facilitates the residues to form hydrogen bonds (Chen et al., 2016). The implications of hydrogen bonding and the mechanism of protein-ligand interactions remain unclear (Chodera and Mobley, 2013). The mechanism by which hydrogen bonds govern molecular docking is fully understood (Taghizadeh et al., 2024; Patil et al., 2010; Olsson et al., 2011). Therefore, understanding the hydrogen bonding landscape in protein–ligand association is crucial. Employing a similar approach, we calculated the total number of hydrogen bonds over the 300 ns simulation time. For the wild type the average number of hydrogen bonds were calculated to be 120, for the L408P the average number of hydrogen bonds were estimated to be 118, for the Y476D the average number of hydrogen bonds were 120, for the C481R the average number of hydrogen bonds were 123, in the C481Y the average number of hydrogen bonds were 120, in M477R the average number of hydrogen bonds were estimated to be 123. In contrast, for the L542P, the average number of hydrogen bonds was estimated to be 121. These results indicate that the global hydrogen-bonding capacity of the BTK–ARQ 531 complexes is largely preserved across the wild type and mutants, consistent with the overall resistance tolerance of ARQ 531. Although the absolute hydrogen bond counts are dominated by intra-protein interactions, this metric remains informative when used comparatively across identically prepared systems and simulation conditions. In a tightly folded kinase domain bound to a small-molecule inhibitor, even modest but consistent differences in total hydrogen bond counts can reflect mutation-induced redistribution of hydrogen-bonding networks that influence local packing, flexibility, and interaction cooperativity. Importantly, the observed preservation of total hydrogen-bond counts does not preclude localized rearrangements of hydrogen-bonding patterns at the binding site; rather, it suggests that mutations primarily induce subtle reorganization of interaction networks rather than wholesale loss of hydrogen bonding. This strongly aligns with the experimental data, which shows that ARQ 531 overcomes the resistance caused by the gatekeeper residues C481S (Reiff et al., 2018b). Hence, this suggests that these mutations alter the hydrogen bond paradigm in these complexes, thereby causing variations in binding at the local interaction-network level, while maintaining overall hydrogen-bonding stability of the complex. The hydrogen bonding graphs for the wild type and mutant complexes are shown in Fig. 6a-d.

3.7. Principal component analysis (PCA)

To cluster the trajectories' motions during the simulation, we performed a PCA analysis. As shown in Fig. 7a–g, most of the complexes retained a constrained motion. Blue and light orange colors represent the two states, while the conformational transition state is represented by purple. These projections correspond to the first two principal components, which together capture the dominant collective motions of the systems and account for the majority of the variance observed during the simulations. The PCA profiles for all mutants are comparable to the wild type. However, compared to the wild type, the L408P mutant has a more non-compact characteristic, which suggests that the mutation affects the protein conformational dynamics in a much less uniform way. The similarity between certain PCs of both mutants and the wild type indicates that the general structure and dynamics have not been dramatically altered despite mutation. The overlap of dominant motion subspaces indicates that the mutations affect the amplitude and distribution of motions rather than introducing entirely new modes of motion in the protein. The wild type has two clearly separated conformational

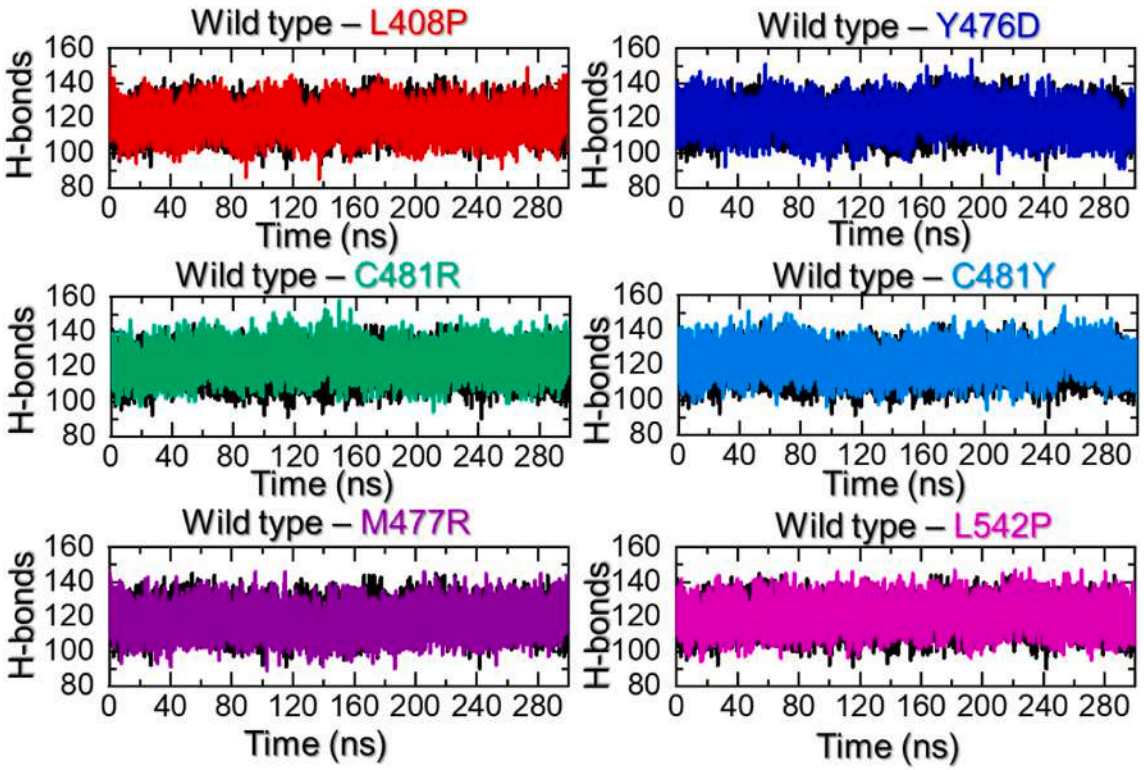

Fig. 6. Hydrogen bonding analysis as a function of time. **(a)** shows the H-bonds results for the wild type and L408P, **(b)** shows the H-bonds results for the wild type and Y476D, **(c)** shows the H-bonds results for the wild type and C481R, **(d)** shows the H-bonds results for the wild type and C481Y, **(e)** shows the H-bonds results for the wild type and M477R, while **(f)** shows the H-bonds results for the wild type and L542P.

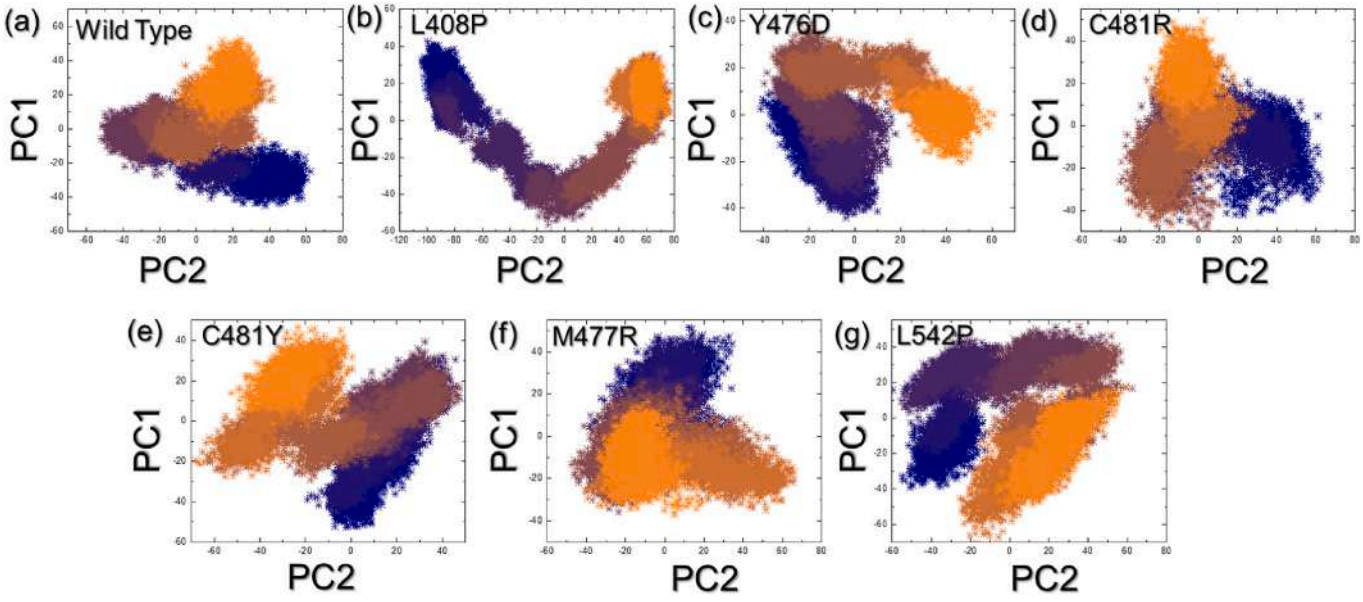

Fig. 7. PCA projections of BTK–ARQ531 complexes showing temporal evolution of conformational sampling from initial (blue) through intermediate transition states (brown) to final equilibrated states (orange) for the wild type and mutant systems. **(a)** shows the PCA graph for the wild type, **(b)** show the PCA graph for the L408P, **(c)** show the PCA graph for the Y476D, **(d)** show the PCA graph for the C481R, **(e)** show the PCA graph for the C481Y, **(f)** show the PCA graph for the M477R while **(g)** show the PCA graph for the L542P mutant system.

basins; therefore, the wild type maintains a balance between open and closed states. The eigenvector clusters for the L408P mutant appear much broader and distorted than the wild type due to the lack of coordinated motion and possible hinge-region destabilization. Therefore, this interpretation is based on a comparison of the relative compactness and dispersion of clusters in the PCA for the mutants, and the comparison should not be interpreted as implying a single or distinct transition. The Y476D and C481R mutants show slight overlap with the wild-type cluster, indicating that while the global fold is similar, the mechanism is altered because of the charge displacement at the active site loop. C481Y appears to follow a similar trajectory to the wild type in terms of cluster dispersion but has a slightly broader conformational

ensemble than the wild type due to localized flexibility rather than global instability. The non-compactness of other PCs shows that localized changes and/or increased flexibility likely occurred in various parts of the mutant systems. These findings highlight relative changes in the density of conformational sampling rather than discrete clusters in PCA. This heterogeneity is an indicator of the disruption that these mutations created in the overall conformational landscape of the protein and illustrates the structural implications of these genetic variations. M477R appears to lie between the two previous categories, as it has a marginally larger conformational envelope, probably due to increased breathing motion of the secondary structural elements. The L542P substitution has caused some degree of unfolding of the peripheral helices, as demonstrated by the separation of clusters in the PCA plot. Together, these findings indicate that, although the overall topology of the mutants is comparable to that of the wild type, many localized deviations and increased fluctuations highlight the unique dynamic effects that each mutation has on the protein's plasticity.

3.8. Free Energy Landscape analysis (FEL)

The analysis of Free Energy Landscape (FEL) shows that both the wild-type and mutant proteins possess one lowest-energy structure. This indicates that the conformations of the proteins during the entire simulation were relatively stable and constrained, meaning that there were no major changes made to the overall structure of the proteins from their wild-type forms due to the genetic mutations made. All three versions of the proteins (the wild type as well as the mutants) appear to have similar energetically favorable structures. The FELs for the proteins were created by projecting all the trajectories (the path/progression of the structure during the simulation) onto the two principal components that made up the conformational densities and using the density estimates to create the relative populations of free energies. The fact that each of the FELs all contained one energetically lowest structural state shows that each protein (including the mutated versions) has a dominant structure that is stable and resulted from the mutations. An additional analysis of the FEL data suggests some small differences in the energies of the different versions of the protein (the mutants) compared to the wild type, and the fact that although they all have the same overall minimum energy structure the wild type version has a narrower/larger energy "bowl" (the "free-energy basin"). This indicates that the wild type is much more stable than the L408P mutant, which has a broader "bowl" and much higher conformational entropy (i.e., less compact) as suggested by its PCA dispersion. It is important to note that the specific size of the free-energy basalt that contains the highest population of the wild-type version indicates that there is only one predominant thermodynamic state, while in the case of the L408P mutant, which has a broader and shallower basin, is more dispersed in the PCA (and therefore likely to have lower energetics density) than the wild type. In addition to the L408P mutant, the Y476D and C481R mutants have a similar profile to that of the wild-type version, both having two energetically comparable minima with uncoupling states that occur around the catalytic core of the protein. The C481Y mutant has maintained its dominant basin, but it has a slightly higher energy than the wild type, indicating that there is no total collapse of the structure, just the (very) slight destabilization; however, the M477R mutant has an energy landscape that is nearly identical to that of the wild-type version with a moderate amount of fluctuation on PC2. Finally, L542P has a considerably flatter energy profile than the wild-type version and indicates the possibility for the presence of many fewer structured intermediates, meaning that it has slightly less stability than the wild type. Overall, these data support the idea that while all three mutants share a predominant energetically stable state, the genetic mutations have introduced slight local energy variances that have changed the conformational preferences for each of the proteins. Importantly, it should be noted that these differences can be described as qualitative changes in complete or nearly complete energy populations, rather than statistically significant energetic separations based on statistical models. The results suggest a structural relationship between the native and mutant versions of the protein because they exhibit similar energetics; however, as evidenced by the dynamic changes in the basin energy topographies, they may produce different effects on the stability of the protein in terms of ligand access and enzyme turnover. The FEL graphs for each of the wild-type and mutant systems are shown in Fig. 8a-g.

3.9. Binding free energy calculation

The validation of docking results can be accomplished through the application of the binding free energy calculation approach, known for its precision, speed, and computational efficiency. This method has been widely used in assessing the binding potential of diverse protein complexes associated with various diseases(Shahraki et al., 2024). Having shown the efficacy of this approach, we conducted binding free energy calculations using the MM/GBSA method. As shown in Table 3, the wild type reported a vdW of −65.60±0.47 kcal/mol, L408P reported a vdW of −66.77±0.35 kcal/mol, Y476D reported a vdW of −66.75±0.34 kcal/mol, M477R reported a vdW of −65.77±0.45 kcal/mol, C481R reported a vdW of −69.29±0.29 kcal/mol, C481Y reported a vdW of −66.54±0.47 kcal/mol, while the L542P mutant reported a vdW of −66.44±0.32 kcal/mol. Moreover, the electrostatic energy for each complex including the wild type reported an electrostatic energy of −2.10±0.24 kcal/mol, L408P reported an electrostatic energy of −1.09 ±0.21 kcal/mol, Y476D reported an electrostatic energy of 0.28±0.16 kcal/mol, M477R reported an electrostatic energy of −0.75±0.21 kcal/mol, C481R reported an electrostatic energy of −0.77±0.20 kcal/mol, C481Y reported and electrostatic energy of −3.89±0.33 kcal/mol while the L542P mutant reported an electrostatic energy of 0.14±0.18 kcal/mol respectively. Finally, the total binding free energy was reported to be −56.87±0.43 kcal/mol for the wild type, −57.42 ±0.34 kcal/mol for the L408P mutant, −56.96±0.31 for the Y476D mutant, −56.41±0.43 kcal/mol for the M477R mutant, −59.45±0.29 kcal/mol for the C481R mutant, −58.26±0.38 kcal/mol for the C481Y mutant while for the L542P mutant the TBE was calculated to be −56.69 ±0.33 kcal/mol. These findings demonstrate that these mutations do not affect the binding of ARQ 531, and they strongly corroborate experimental results indicating that ARQ 531 can overcome the drug resistance caused by these mutations in BTK (Reiff et al., 2018b; Skånland and Mato, 2021). Furthermore, from these data, it can be observed that each system gains free energy in the gas phase but not in the solvent. These findings suggest that the ARQ 531 binding is thermodynamically favored in terms of enthalpy, primarily owing to advantageous interactions in the gas phase. However, it is disfavored in terms of entropy due to the unfavorable effects of solvation. The binding free energy results are shown in Table 3.

4. Conclusions

The current study employs a genomic mutation screening strategy, combined with molecular modeling approaches, to assess the impact of non-synonymous substitutions in BTK on the binding of ARQ 531. Our results revealed that among the 82 mutations, only 62 were classified as deleterious by various algorithms, while the graph-signature-based approach revealed only nine substitutions as highly deleterious. The docking results predicted no significant variation, while the simulation results revealed local changes only in the dynamic behavior. Moreover, the binding free energy results indicate that these mutations do not affect the binding of ARQ 531, which strongly corroborates experimental results showing that ARQ 531 can overcome the drug resistance caused by these mutations in BTK. While only a single replica was simulated for each system, which remains the major limitation of the current study, the extended 300 ns trajectories and the clear convergence of RMSD profiles support the reliability of the observed comparative stability trends among the mutants. While 300 ns may not capture

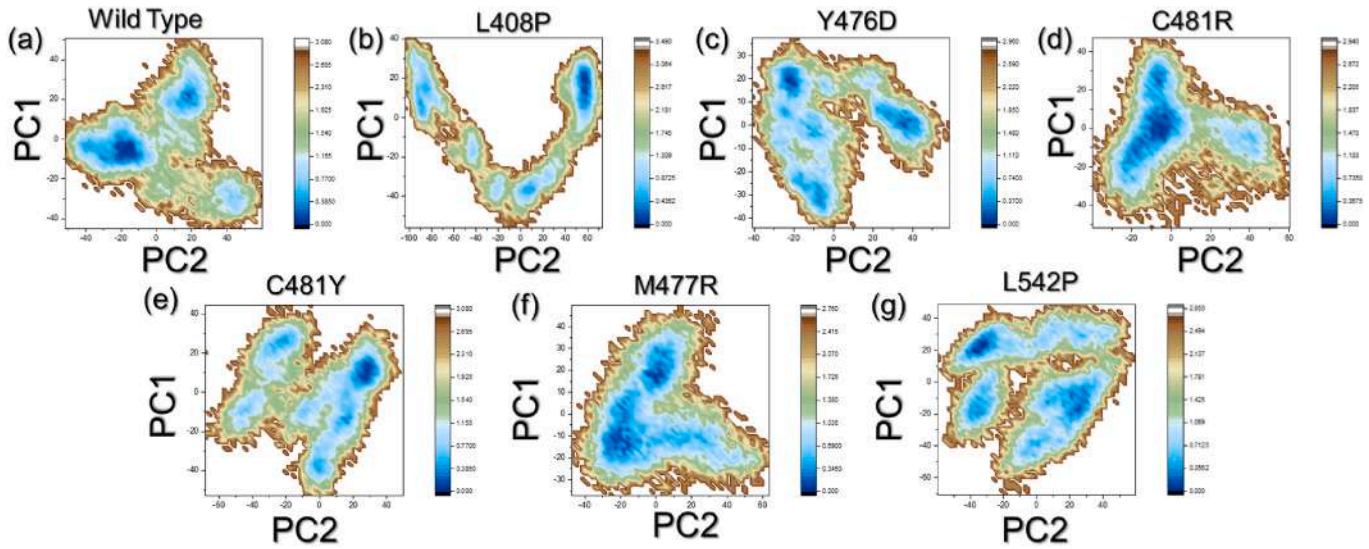

Fig. 8. Free Energy Landscape (FEL) analysis of the wild type and mutant systems. (a) shows the FEL graph for the wild type, **(b)** shows the FEL graph for the L408P, **(c)** shows the FEL graph for the Y476D, **(d)** shows the FEL graph for the C481R, **(e)** shows the FEL graph for the C481Y, **(f)** shows the FEL graph for the M477R, while **(g)** shows the FEL graph for the L542P mutant system.

Table 3
Binding free energy calculation for the wild type and mutant systems using the MM/GBSA approach. All the results are shown in kcal/mol.

Parameters	Wild Type	L408P	Y476D	M477R	C481R	C481Y	L542P
vdW	−65.60±0.47	−66.77±0.35	−66.75±0.34	−65.77±0.45	−69.29±0.29	−66.54±0.47	−66.44±0.32
EEL	−2.10±0.24	−1.09±0.21	0.28±0.16	−0.75±0.21	−0.77±0.20	−3.89±0.33	0.14±0.18
EGB	17.93±0.21	17.45±0.18	16.65±0.16	17.15±0.23	18.11±0.22	19.11±0.23	16.64±0.15
ESURF	−7.10±0.02	−7.00±0.02	−7.13±0.02	−7.04±0.03	−7.49±0.03	−6.93±0.03	−7.03±0.02
ΔG gas	−67.70±0.48	−67.87±0.45	−66.47±0.38	−66.52±0.55	−70.07±0.39	−70.43±0.36	−66.30±0.37
ΔG Solv	10.83±0.21	10.45±0.19	9.5±0.15	10.10±0.22	10.61±0.22	12.17±0.24	9.60±0.15
Total ΔG	**−56.87±0.43**	**−57.42±0.34**	**−56.96±0.31**	**−56.41±0.43**	**−59.45±0.29**	**−58.26±0.38**	**−56.69±0.33**

all rare, long-timescale kinase transitions (e.g., full activation-loop/DFG switching), the trajectories exhibit stable plateaus and converged local descriptors, supporting robust comparative inference of mutation-induced binding-site dynamics within the sampled inhibitor-bound state. These findings confirm that ARQ 531 is the preferred therapy for multiple leukemias and can serve as a starting point for designing more robust and effective BTKi.

Author contributions (CRediT statement)

Conceptualization: **Abbas Khan, Abdelali Agouni**.
Methodology: **Abbas Khan, Mohanad Shkoor, Raed M. Al-Zoubi**.
Software: **Abbas Khan**.
Validation: **Abbas Khan, Mona A. Sawali**.
Formal analysis: **Abbas Khan**.
Investigation: **Abbas Khan, Mohanad Shkoor, Fakhrul Hasssan**.
Resources: **Abdelali Agouni, Sergio Crovella**.
Data curation: **Abbas Khan**.
Writing – original draft: **Abbas Khan**.
Writing – review & editing: **Abbas Khan, Mona A. Sawali, Raed M. Al-Zoubi, Sergio Crovella, Abdelali Agouni**.
Visualization: **Abbas Khan**.
Supervision: **Abdelali Agouni**.
Project administration: **Abdelali Agouni**.
Funding acquisition: **Fahad M. Alshabrmi, Abdelali Agouni**.
All authors have read and approved the final manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was made possible with the support of Qatar National Research Fund (Qatar Research Development and Innovation Council) [grant No. ARG01-0601-230451] and Qatar University grant No. QUPD-CPH-23/24-592. The statements made herein are solely the responsibility of the authors.

Data availability

All relevant data are included within the manuscript. Additional data or information can be provided by the corresponding author upon reasonable request.

References

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. Nat. Methods 7, 248–249.
Ahn, I.E., Underbayev, C., Albitar, A., Herman, S.E., Tian, X., Maric, I., Arthur, D.C., Wake, L., Pittaluga, S., Yuan, C.M., 2017. Clonal evolution leading to ibrutinib resistance in chronic lymphocytic leukemia. Blood, The Journal of the American Society of Hematology 129, 1469–1479.
Babar, Z., Khan, M., Zahra, M., Anwar, M., Noor, K., Hashmi, H.F., Suleman, M., Waseem, M., Shah, A., Ali, S., 2022. Drug similarity and structure-based screening of medicinal compounds to target macrodomain-I from SARS-CoV-2 to rescue the host immune system: a molecular dynamics study. J. Biomol. Struct. Dyn. 40, 523–537.
Bender, A.T., Gardberg, A., Pereira, A., Johnson, T., Wu, Y., Grenningloh, R., Head, J., Morandi, F., Haselmayer, P., Liu-Bujalski, L., 2017. Ability of Bruton's tyrosine kinase inhibitors to sequester Y551 and prevent phosphorylation determines potency

for inhibition of Fc receptor but not B-Cell receptor signaling. Mol. Pharmacol. 91, 208–219.
Bendl, J., Stourac, J., Salanda, O., Pavelka, A., Wieben, E.D., Zendulka, J., Brezovsky, J., Damborsky, J., 2014. PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. PLoS Comput. Biol. 10, e1003440.
Blombery, P., Thompson, E.R., Lew, T.E., Tiong, I.S., Bennett, R., Cheah, C.Y., Lewis, K. L., Handunnetti, S.M., Tang, C.P.S., Roberts, A., 2022. Enrichment of BTK Leu528Trp mutations in patients with CLL on zanubrutinib: potential for pirtobrutinib cross-resistance. Blood Adv. 6, 5589–5592.
Bond, D.A., Woyach, J.A., 2019. Targeting BTK in CLL: beyond ibrutinib. Current hematologic malignancy reports 14, 197–205.
Boros, K., Puissant, A., Back, M., Alexe, G., Bassil, C.F., Sinha, P., Tholouli, E., Stegmaier, K., Byers, R.J., Rodig, S.J., 2015. Increased SYK activity is associated with unfavorable outcome among patients with acute myeloid leukemia. Oncotarget 6, 25575–25587.
Burger, J.A., 2019. BTK inhibitors: present and future. Cancer J. 25, 386.
Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P.L., Casadio, R., 2009. Functional annotations improve the predictive score of human disease-related mutations in proteins. Hum. Mutat. 30, 1237–1244.
Capriotti, E., Calabrese, R., Casadio, R., 2006. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics 22, 2729–2734.
Carnevale, J., Ross, L., Puissant, A., Banerji, V., Stone, R.M., DeAngelo, D.J., Ross, K.N., Stegmaier, K., 2013. SYK regulates mTOR signaling in AML. Leukemia 27, 2118–2128.
Case, D.A., Cheatham III, T.E., Darden, T., Gohlke, H., Luo, R., Merz Jr., K.M., Onufriev, A., Simmerling, C., Wang, B., Woods, R.J., 2005. The Amber biomolecular simulation programs. J. Comput. Chem. 26, 1668–1688.
Chen, D., Oezguen, N., Urvil, P., Ferguson, C., Dann, S., Savidge, T., 2016. Regulation of protein-ligand binding affinity by hydrogen bond pairing. Sci. Adv. 2 (3), e1501240.
Chodera, J.D., Mobley, D.L., 2013. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. Annu. Rev. Biophys. 42, 121–142.
de Gorter, D.J., Beuling, E.A., Kersseboom, R., Middendorp, S., van Gils, J.M., Hendriks, R.W., Pals, S.T., Spaargaren, M., 2007. Bruton's tyrosine kinase and phospholipase Cγ2 mediate chemokine-controlled B cell migration and homing. Immunity 26, 93–104.
De Nigris, E., Yang, X., Zanardo, E., Lejeune, D., Farooqui, M.Z., Gandra, S.R., Laliberté, F., 2024. Healthcare resource utilization and costs of chronic lymphocytic leukemia/small lymphocytic lymphoma patients who relapse or are refractory to ibrutinib. Future Oncol. 20, 2723–2735.
Dhami, K., Chakraborty, A., Gururaja, T.L., Cheung, L.W.-K., Sun, C., DeAnda, F., Huang, X., 2022. Kinase-deficient BTK mutants confer ibrutinib resistance through activation of the kinase HCK. Sci. Signal. 15 eabg5216.
Gángó, A., Alpár, D., Galik, B., Marosvári, D., Kiss, R., Fésüs, V., Aczél, D., Eyüpoglu, E., Nagy, N., Nagy, Á., 2020. Dissection of subclonal evolution by temporal mutation profiling in chronic lymphocytic leukemia patients treated with ibrutinib. Int. J. Cancer 146, 85–93.
Handunnetti, S.M., Tang, C.P.S., Nguyen, T., Zhou, X., Thompson, E., Sun, H., Xing, H., Zhang, B., Guo, Y., Sutton, L.A., 2019. BTK Leu528Trp-a potential secondary resistance mechanism specific for patients with chronic lymphocytic leukemia treated with the next generation BTK inhibitor zanubrutinib. Blood 134, 170.
Hecht, M., Bromberg, Y., Rost, B., 2015. Better prediction of functional effects for sequence variants. BMC Genom. 16, 1–12.
Hoseini, S., Moghadam, A., Palizban, A., Hoseini, F., Taghizadeh, M.S., Choopani, A., Hoseini, F., 2025. Immune boosting by pharmaceutical C-phycocyanin, in vitro and in silico effects on peripheral blood mononuclear cells. Health Biotechnology and Biopharma (HBB) 9, 22–47.
Hou, T., Wang, J., Li, Y., Wang, W., 2011. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J. Chem. Inf. Model. 51, 69–82.
Jones, D., Woyach, J.A., Zhao, W., Caruthers, S., Tu, H., Coleman, J., Byrd, J.C., Johnson, A.J., Lozanski, G., 2017. PLCG2 C2 domain mutations co-occur with BTK and PLCG2 resistance mutations in chronic lymphocytic leukemia undergoing ibrutinib treatment. Leukemia 31, 1645–1647.
Kadri, S., Lee, J., Fitzpatrick, C., Galanina, N., Sukhanova, M., Venkataraman, G., Sharma, S., Long, B., Petras, K., Theissen, M., 2017. Clonal evolution underlying leukemia progression and Richter transformation in patients with ibrutinib-relapsed CLL. Blood Adv. 1, 715–727.
Karczewski, K., Francioli, L., Tiao, G., Cummings, B., Alföldi, J., Wang, Q., Genome Aggregation Database, C., 2020. The Mutational Constraint Spectrum Quantified from Variation in, vol. 141, pp. 434–443.
Kurita, T., 2019. Principal component analysis (PCA). Computer Vision: a Reference Guide, pp. 1–4.
Liu, T.M., Woyach, J.A., Zhong, Y., Lozanski, A., Lozanski, G., Dong, S., Strattan, E., Lehman, A., Zhang, X., Jones, J.A., Flynn, J., Andritsos, L.A., Maddocks, K., Jaglowski, S.M., Blum, K.A., Byrd, J.C., Dubovsky, J.A., Johnson, A.J., 2015. Hypermorphic mutation of phospholipase C, γ2 acquired in ibrutinib-resistant CLL confers BTK independency upon B-cell receptor activation. Blood 126, 61–68.
Maisuradze, G.G., Liwo, A., Scheraga, H.A., 2010. Relation between free energy landscapes of proteins and dynamics. J. Chem. Theor. Comput. 6, 583–595.
Meza, J.C., 2010. Steepest descent. Wiley Interdiscip. Rev.: Comput. Stat. 2, 719–722.
Mohamed, A.J., Yu, L., Bäckesjö, C.M., Vargas, L., Faryal, R., Aints, A., Christensson, B., Berglöf, A., Vihinen, M., Nore, B.F., 2009. Bruton's tyrosine kinase (Btk): function, regulation, and transformation with special emphasis on the PH domain. Immunol. Rev. 228, 58–73.
Naeem, A., Utro, F., Wang, Q., Cha, J., Vihinen, M., Martindale, S., Zhou, Y., Ren, Y., Tyekucheva, S., Kim, A.S., 2023. Pirtobrutinib targets BTK C481S in ibrutinib-resistant CLL but second-site BTK mutations lead to resistance. Blood Adv. 7, 1929–1943.
Olsson, T.S., Ladbury, J.E., Pitt, W.R., Williams, M.A., 2011. Extent of enthalpy–entropy compensation in protein–ligand interactions. Protein Sci. 20, 1607–1618.
Patil, R., Das, S., Stanley, A., Yadav, L., Sudhakar, A., Varma, A.K., 2010. Optimized hydrophobic interactions and hydrogen bonding at the target-ligand interface leads the pathways of drug-designing. PLoS One 5, e12029.
Pires, D.E., Ascher, D.B., Blundell, T.L., 2014. mCSM: predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics 30, 335–342.
Quartier, P., Debré, M., De Blic, J., de Sauverzac, R., Sayegh, N., Jabado, N., Haddad, E., Blanche, S., Casanova, J.-L., Smith, C.E., 1999. Early and prolonged intravenous immunoglobulin replacement therapy in childhood agammaglobulinemia: a retrospective survey of 31 patients. J. Pediatr. 134, 589–596.
Reiff, S.D., Mantel, R., Smith, L.L., Greene, J., Muhowski, E.M., Fabian, C.A., Goettl, V. M., Tran, M., Harrington, B.K., Rogers, K.A., 2018a. The BTK inhibitor ARQ 531 targets ibrutinib-resistant CLL and Richter transformation. Cancer Discov. 8, 1300–1315.
Reiff, S.D., Mantel, R., Smith, L.L., Greene, J.T., Muhowski, E.M., Fabian, C.A., Goettl, V. M., Tran, M., Harrington, B.K., Rogers, K.A., Awan, F.T., Maddocks, K., Andritsos, L., Lehman, A.M., Sampath, D., Lapalombella, R., Eathiraj, S., Abbadessa, G., Schwartz, B., Johnson, A.J., Byrd, J.C., Woyach, J.A., 2018b. The BTK inhibitor ARQ 531 targets ibrutinib-resistant CLL and Richter transformation. Cancer Discov. 8, 1300–1315.
Rodrigues, C.H.M., Pires, D.E.V., Ascher, D.B., 2021. DynaMut2: assessing changes in stability and flexibility upon single and multiple point missense mutations. Protein Sci. 30, 60–69.
Roe, D.R., Cheatham III, T.E., 2013. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J. Chem. Theor. Comput. 9, 3084–3095.
Salomon-Ferrer, R., Gotz, A.W., Poole, D., Le Grand, S., Walker, R.C., 2013. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. J. Chem. Theor. Comput. 9, 3878–3888.
Salomon-Ferrer, R., Case, D.A., Walker, R.C., 2013. An overview of the Amber biomolecular simulation package. Wiley Interdiscip. Rev. Comput. Mol. Sci. 3, 198–210.
Shahraki, Z., Taghizadeh, M.S., Niazi, A., Rowshan, V., Moghadam, A., 2024. Enhancing bioactive compound production in Salvia mirzayanii through elicitor application: insights from in vitro and in silico studies. Food Biosci. 60, 104185.
Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., Ng, P.C., 2012. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res. 40, W452–W457.
Skånland, S.S., Mato, A.R., 2021. Overcoming resistance to targeted therapies in chronic lymphocytic leukemia. Blood Adv. 5, 334–343.
Smith, C.E., Islam, T.C., Mattsson, P.T., Mohamed, A.J., Nore, B.F., Vihinen, M., 2001. The Tec family of cytoplasmic tyrosine kinases: Mammalian Btk, Bmx, Itk, Tec, Txk and homologs in other species. Bioessays 23, 436–446.
Stone, E.A., Sidow, A., 2005. Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. Genome Res. 15, 978–986.
Sun, T., Chen, Y., Wen, Y., Zhu, Z., Li, M., 2021. PremPLI: a machine learning model for predicting the effects of missense mutations on protein-ligand interactions. Commun. Biol. 4, 1311.
Taghizadeh, M.S., Taherishirazi, M., Niazi, A., Afsharifar, A., Moghadam, A., 2024. Structure-guided design and cloning of peptide inhibitors targeting CDK9/cyclin T1 protein-protein interaction. Front. Pharmacol. 15, 2024.
Taghizadeh, M.S., Mirzapour-Kouhdasht, A., Huang, J.-Y., 2025. Unraveling the structure-activity relationship and molecular mechanisms for targeted therapeutic applications of food-derived angiotensin-I-converting enzyme inhibitory peptides. Food Biosci. 74, 107820.
Takata, M., Kurosaki, T., 1996. A role for Bruton's tyrosine kinase in B cell antigen receptor-mediated activation of phospholipase C-gamma 2. J. Exp. Med. 184, 31–40.
Wang, E., Mi, X., Thompson, M.C., Montoya, S., Notti, R.Q., Afaghani, J., Durham, B.H., Penson, A., Witkowski, M.T., Lu, S.X., 2022. Mechanisms of resistance to noncovalent Bruton's tyrosine kinase inhibitors. N. Engl. J. Med. 386, 735–743.
Watowich, S.J., Meyer, E.S., Hagstrom, R., Josephs, R., 1988. A stable, rapidly converging conjugate gradient method for energy minimization. J. Comput. Chem. 9, 650–661.
Webb, B., Sali, A., 2016. Comparative protein structure modeling using MODELLER. Curr. Protoc. Bioinform. 54, 5.6. 1-5.6. 37.
Woyach, J., Huang, Y., Rogers, K., Bhat, S.A., Grever, M.R., Lozanski, A., Doong, T.-J., Blachly, J.S., Lozanski, G., Jones, D., 2019. Resistance to acalabrutinib in CLL is mediated primarily by BTK mutations. Blood 134, 504.
Woyach, J.A., Flinn, I.W., Awan, F.T., Eradat, H., Brander, D., Tees, M., Parikh, S.A., Phillips, T.J., Ghori, R., Reddy, N.M., Farooqui, M.Z.H., Byrd, J.C., Stephens, D.M., 2022. Efficacy and safety of ARQ 531, a wild-type and C481S-Mutated bruton tyrosine kinase inhibitor for B-Cell malignancies: updated analysis of the open-label phase 1/2 dose-expansion Bellwave-001 Study. Blood 140, 7004–7006.