

Full Length Research Paper

Molecular modeling of mutant kinase domain of Btk, a Tec family member, for structure prediction

Hina Iqbal, Asif Mir and Rani Faryal*

Department of Biosciences, COMSATS Institute of Information Technology, 44000, Islamabad, Pakistan.

Accepted 15 March, 2011

Bruton's tyrosine kinase (Btk), a member of non-receptor tyrosine kinases, is one of the crucial kinases for the B-cell maturation and mast cell activation. This work was planned to develop the full length Btk protein structure via different approaches. The threading approach provided suitable full length Btk protein structure. Furthermore, this structure was utilized to predict the consequences of 16 selected missense mutations in the kinase domain (402 to 651) around Y551, a first tyrosine to be autophosphorylated and important for downstream signaling. Valuable information gathered from this work is an insight for analyzing protein structural changes and their effects on protein stability. The amino acid residues at positions 554, 559 and 562 were considered critical as per their involvement and presence near the catalytic site, peptide substrate binding pocket and their linkage with other amino acid residues resulting in the disturbance of structural integrity.

Key words: Bruton's tyrosine kinase (Btk), *in silico*, kinase domain, comparative homology modeling, autophosphorylation.

INTRODUCTION

Bruton's tyrosine kinase (Btk) belongs to the family of non-receptor tyrosine kinases namely Tec family. This is the second largest family of cytoplasmic protein tyrosine kinases and consists of five members; Btk, Itk, Tec, Rlk/Txk and Bmx. Btk is expressed in B cells and plays a crucial role in B-cell maturation and mast cell activation. Structurally, Btk protein has C-terminal kinase domain followed by two Src homology domains (SH2 and SH3). At the N-terminal of the SH3 domain, Btk possesses two proline rich regions (PRR) with Zn²⁺-binding region known as Btk motif (BH). BH motif and the PRR together form the Tec homology domains. At the amino terminus, Btk possesses a pleckstrin homology domain (Lopez-Herrera et al., 2008). There is credible evidence that Btk is involved in the signaling of B cell receptor (BCR) following its stimulation. Activation of Src kinase Lyn after

stimulation of BCR, leads to the recruitment of Btk to the plasma membrane via its PH domain. Btk is phosphorylated on tyrosine (Y) at the position 551 via Src kinase involving SLP-65. Phosphorylation of the Y551 leads to conformational change in the Btk protein and is important for the autophosphorylation of Y223 in the SH2 domain. This phenomenon is important for the full activation of Btk and is integral for the down stream signaling, which leads to pleiotropic effects; such as PLC-gamma activation, generation of second messengers like inositol (1, 4, 5) - triphosphate (IP3), diacylglycerol and calcium, cyto-skeletal remodeling, cell proliferation through the regulation of gene transcription and expression (Futatani et al., 1998; Manna et al., 2007). The block in differentiation from pro-B to pre-B cells results in a selective defect in the humoral immune response characteristic of human X-linked Agammaglobulinemia (XLA). Mutations of Bruton's tyrosine kinase (Btk) gene have been identified as the cause of XLA (Danielian et al., 2003).

Although, the structures of the phosphorylated and unphosphorylated Btk kinase domain are known but understanding the kinase selectivity profiles of Btk inhibitors has been hampered by the lack of availability of a high resolution, ligand-bound Btk structure. The protein-folding

*Corresponding author. E-mail: ranifaryal@comsats.edu.pk.

Abbreviations: Btk, Bruton's tyrosine kinase; SH, Src homology; PH, pleckstrin homology; NCBI, National Center for Biotechnology Information; PDB, protein data bank; aa, amino acid; PTK, protein tyrosine kinase; Y, tyrosine.

problem (predicting the three-dimensional configuration of a protein given the sequence of amino acids) is one of the important and crucial parts of computational biology. Experimental techniques to determine protein structures such as x-ray crystallography, are not able to keep up with recent huge improvements in the ability to sequence DNA (and thereby, proteins), making a theoretical solution to the problem that is becoming even more urgent. Therefore, a solution to these problem is needed that would have extraordinary scientific implications and open the way to rational drug design without the need for laborious wet laboratory work that is the need of time.

MATERIALS AND METHODS

Sequence retrieval

Btk gene (MIM no. 300300) is present on X chromosome at q arm at position 21.3 to 22. It contains 19 exons transcribing 16 transcripts out of which 10 are producing protein products of different lengths. The amino acid sequences of Btk and its variants were obtained from NCBI and Ensemble databases. The full length coding protein of Btk is of 659 amino acid long, whereas, other isoforms are of various lengths that is., 483, 280, 252, 216, 183, 118, 58, 55. Domains of Btk are PH, SH3, SH2, SH1 (kinase domain) whereas there exists a Btk motif (BH) in Btk protein (Table 1).

It was confirmed that, there is no full length Btk crystalline structure existing in protein databank. However, all of the domains have been structurally identified. Therefore, the major module of the study was to model the full length 3-D structure of Btk.

Full length 3-D structure prediction

All the crystalline structure and PDB files of the domains of Btk have been preserved in the protein databank. Therefore, it was easy to select the templates for Btk. In this study, two strategies were used to model the full length Btk.

Comparative homology modeling

SWISS-MODEL (Arnold et al., 2006), 3Djigsaw, CPH models (Lund et al., 2002), ESyPred3D (Lambert et al., 2002) and Geno3d (Combet et al., 2002) automatically build models by using their own set of modeling algorithms. These web servers used different templates to model the Btk structure but the major drawback was that, they built Btk model of different lengths instead of full length (Table 2).

Threading approach

According to different studies at different time, different domains were modeled and also analyzed mutationally (Vihinen et al., 1994). PDB files of all the domains and Btk motif exists in protein data bank, therefore, this study considered some of them (Table 4) as multiple templates using Psi-Blast (Altschul et al., 1997), with matrix BLOSUM62 (Altschul et al., 2005). Aligned sequences of all the templates with Btk sequence were allowed to model using the modeler 9v3 (Sali et al., 1995) managed by Bioinformatics toolkit server (Biegert et al., 2006). A full length model was built (Figure 1). For this study, another web server SAM_T08 was also use

(Karplus, 2009; Katzman et al., 2008; Shackelford and Karplus, 2007; Karchin et al., 2003; Karplus et al., 2001) with the same threading approach to build the full length model (Figure 2).

Superimposing models

In this study, desktop application was used that is., Viewerlite, to analyze the existing domains with the models built by different servers (Modeller and Sam_T08).

Models evaluation

To evaluate the models built, this study used the RAMPAGE server (Lovell et al., 2002). This server allowed analyzing the residues in the favored, allowed and disallowed regions (Figures 3 and 4, for models built by Modeller and Sam_T08). Packing quality control was checked by what if server (Table 5).

Mutation analysis

Mutations were retrieved by the databases that is, BTKbase (Väliaho et al., 2006) and HGMD (Cooper et al., 2005) (Table 6). For the detection of the effects of selected reported missense Btk mutations on the structure of Btk, what if web server was used to mutate the Btk protein (Figure 5a to p). The mutated structures were analyzed with wild structure by imposing it use on Viewerlite. However, mutant protein stability changes were analyzed by a web tool PoPMuSiC v2 (Dehouck et al., 2009) (Table 6).

Pockets generation

Pockets of Btk were generated by CASTp server (Binkowski et al., 2003) and it generated 108 pockets of Btk protein.

RESULTS AND DISCUSSION

Btk, a member of non-receptor tyrosine kinases, is one of the prime kinase that is involved in the maturation of pre-B cell to pro-B cell. Absence or abnormal Btk protein is responsible for the development of X-Linked Agammaglobulinemia (XLA) in humans. Btk, a member of cytoplasmic protein tyrosine kinase (PTK), belonging to the Tec family (Bolen, 1995), is composed of Src homology 2 and 3 (SH2 and SH3) domains along with catalytic kinase domain. Being a Tec family kinase, it contains an N-terminal pleckstrin homology (PH) domain which includes a Btk motif and a proline-rich region with conserved zinc finger motif. The Btk motif comprises of only 26 residues which characteristically include three fully conserved cysteines and a histidine (Vihinen et al., 1994). Binding and coordination of Btk to the Zn²⁺ ion is mediated by a this zinc finger motif (Smith et al., 1994; Murayama et al., 2008). Mutations affecting PH and TH domains affect Zn²⁺ binding, hence, leading to the generation of extremely unstable protein (Vihinen et al., 1997; Hyvonen and Saraste, 1997). Btk is a metalloprotein enzyme, requiring Zn²⁺ for optimal activity and stability, but the deletion of the PH-TH domain doublet or a trunca-

Table 1. Domains of Btk and its isoforms according to SMART Jorg et al., (1998).

Number	Size	PH	BH	SH3	SH2	Kinase Domain	Other
1	659	4-135	135-171	217-273	279-368	402-651	-
2	483	4-135	135-171	217-273	279-347	-	-
3	252	-	-	-	1-83	21-244	-
4	183	-	-	-	1-83	21-175 (Unclassified specificity)	-
5	280	-	-	-	1-83	117-272 (Unclassified specificity)	-
6	483	4-135	135-171	217-273	279-353	255-479	-
7	216	-	-	-	1-83	117-216 (Unclassified specificity)	-
8	55	-	-	-	1-39	-	11-41 (Sel 1 like repeats)
9	58	-	-	-	-	-	-
10	118	-	-	-	1-83	-	17-75 (Helix_turn_helix, arsenical resistance operon repressor)

Table 2. Z-score with models length generated by different servers via comparative homology modeling.

Model	Z-score	Model length
3 DJIGSAW		
1	-2.576	2-170
GENO3D		
1	-24.18	1-169
2	-24.18	1-169
3	-24.18	1-169
SWISS-MODEL		
1	-2.072	2-170
2	-4.824	270-350
3	-4.228	216-273
4	-3.130	181-271
5	-6.019	130-196
6	-1.241	346-478
CPHMODELS		
1	-3.966	397-654
ESYPRED3D		
1	-0.653	252-657

Table 3. Ramachandran plot values of all the models generated by all the web servers obtained through RAMPAGE.

Model	Ramachandran plot value		
	Number of residue in favored region (%)	Number of residue in allowed region (%)	Number of residue in outlier region (%)
3DJIGSAW			
1	93.4	4.2	2.4
GENO3D			
1	82.0	15.6	2.4
2	83.8	12.6	3.6
3	86.2	12.0	1.8
SWISS-MODEL			
1	95.2	2.4	2.4
2	77.2	19.0	3.8
3	66.1	32.1	1.8
4	80.9	14.6	4.5
5	75.4	20.0	4.6
6	87.8	10.7	1.5
CPHMODELS			
1	93.0	6.6	0.4
MODELLER			
1	88.1	8.8	3.0
SAM_T08			
1	97.2	2.2	0.6

Table 4. Multiple templates used for MODELLER.

Number	Template	Amino acid	Domain	Region covering
1	1K2P_A	283	Bruton's tyrosine kinase	381-659
2	1BTK_A	169	Ph domain and Btk motif	2-170
3	2GE9_A	125	The SH2 domain	269-387
4	1AWW_A	67	SH3 domain	211-275

tion extending into the N-terminal part of the SH3 domain renders Btk slightly more nuclear, but not unstable (Mohamed et al., 2000). In addition, mutation of Y223, but not Y551, increases the fraction of Btk in the nucleus (Mohamed et al., 2009). These mutations gave insight that mutation affecting the PH-TH is not directly affecting the kinase activity. The present study is targeted at the modeling of kinase domain; therefore, only mutations affecting the kinase domain are discussed. This work determined the effect of missense mutations and development of mutant proteins and their effect on downstream signaling due to conformational changes that resulted from mutations. Full length Btk structure was determined *in silico* using two approaches that is,

comparative homology modeling that approximates the 3D structure of a target protein for which only the sequence is available and threading approach, that is based on the observation that many protein structures in the PDB are very similar.

Model building and selection

Different web servers were used for comparative homology modeling like SWISS-MODEL, 3Djigsaw, CPHmodels, ESyPred3D and Geno3d. All the models generated by these servers were evaluated by RAMPAGE server along with their Z-scores (Table 2 and Table 3). Selection criteria

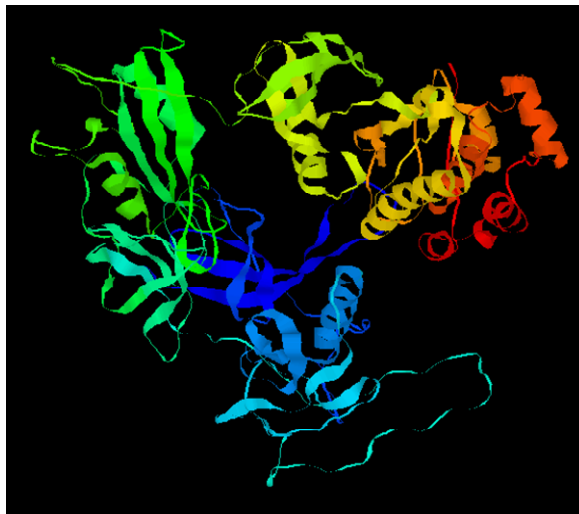


Figure 1. Full length BTK model from MODELLER 9v3 in ViewerLite version 5.0.



Figure 2. Full length BTK model from SAM_T08 in ViewerLite version 5.0.

criteria was based upon the full length Btk structure determination rather than short stretches of varying lengths generated by these web servers. The threading approach was applied by using two methods. Firstly, manually templates were selected (Table 4) then, modeling of protein was done by MODELER 9v3. Other method was by using a SAM_T08 server. SAM_T08 server automatically selected multiple templates to build the best model and in this case it selected around 45 templates out of which 20 were top reported for alignment to generate the Btk model. These templates belonged to the members of the same and different families of kinases. The models generated by these servers were also evaluated by RAMPAGE server and packing quality was checked by what if web server (Table 5 and Table 3). On the basis of evaluation by RAMPAGE server model generated by

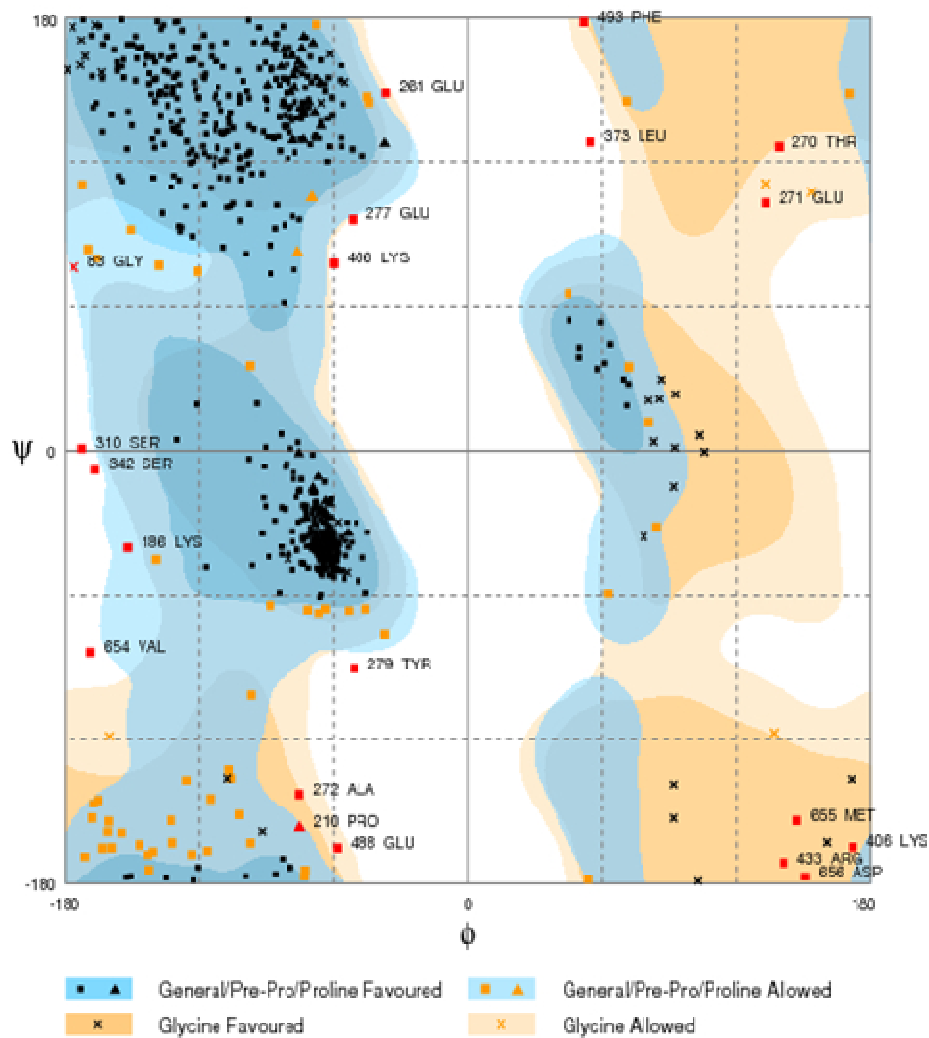
SAM_T08, server was found to be favorable.

Model validation

In order to validate the predicted model generated by SAM_T08 server, this study selected approach of superimposing existing crystalline structures of all the domains of Btk present at protein data bank (PDB). This also suggested the positive sign of model selection.

Pocket generation and selection

One of the important features of a functional unit in a protein is a function of domains and motifs. In order to



Ⓔ

Figure 3 A

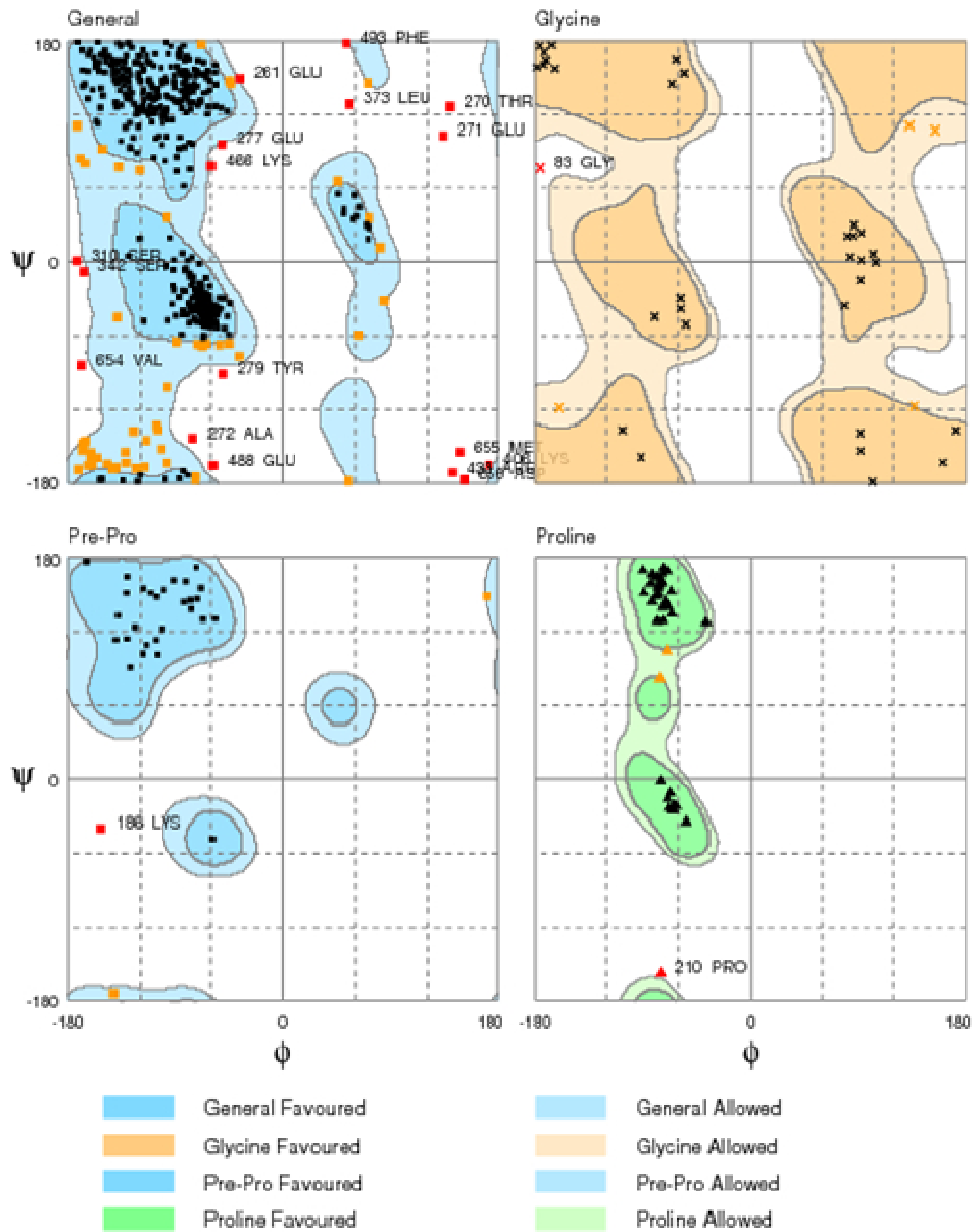
determine the role of kinase domain and effect of its mutations, 108 pockets were generated and analyzed by CASTp server in a full length Btk structure. On the basis of presentation of both tyrosine residues, one (Y551) involved in phosphorylation with Src kinase and other (Y223) residue involved in autophosphorylation, pocket was selected (Figure 6).

Mutation analysis

A limited wet lab data is available on mutations in Btk and its consequences on mature protein. Kinase domain of Btk is the first domain to be phosphorylated, while Btk is stimulated by other Src kinase. It is considered as a prime domain and is integral for down stream signaling. Therefore, kinase domain was selected to predict the

effect of mutations on change in catalytic site and its implications on full Btk structure and processes involved directly due to this domain and its consequences. Mutation analysis was carried out for mutations in area thought to be main site for kinase domain's catalytic activity. There were sixteen reported missense mutations around Y551 (Table 6). Missense mutations in the kinase domain of Btk have been found to abolish Btk expression completely (Futatani et al., 1998; Gaspar et al., 1998), presumably as a result of protein instability (Gaspar et al., 2000).

Mutation at position 535, valine is substitute with phenylalanine (Conley et al., 1998). Amino acid valine, a non polar aliphatic upon missense mutation changed to aromatic amino acid (phenylalanine). It can be inferred from Figure 5a that, the change from aliphatic side chain to aromatic hydrophobic R-group led to exposed aliphatic

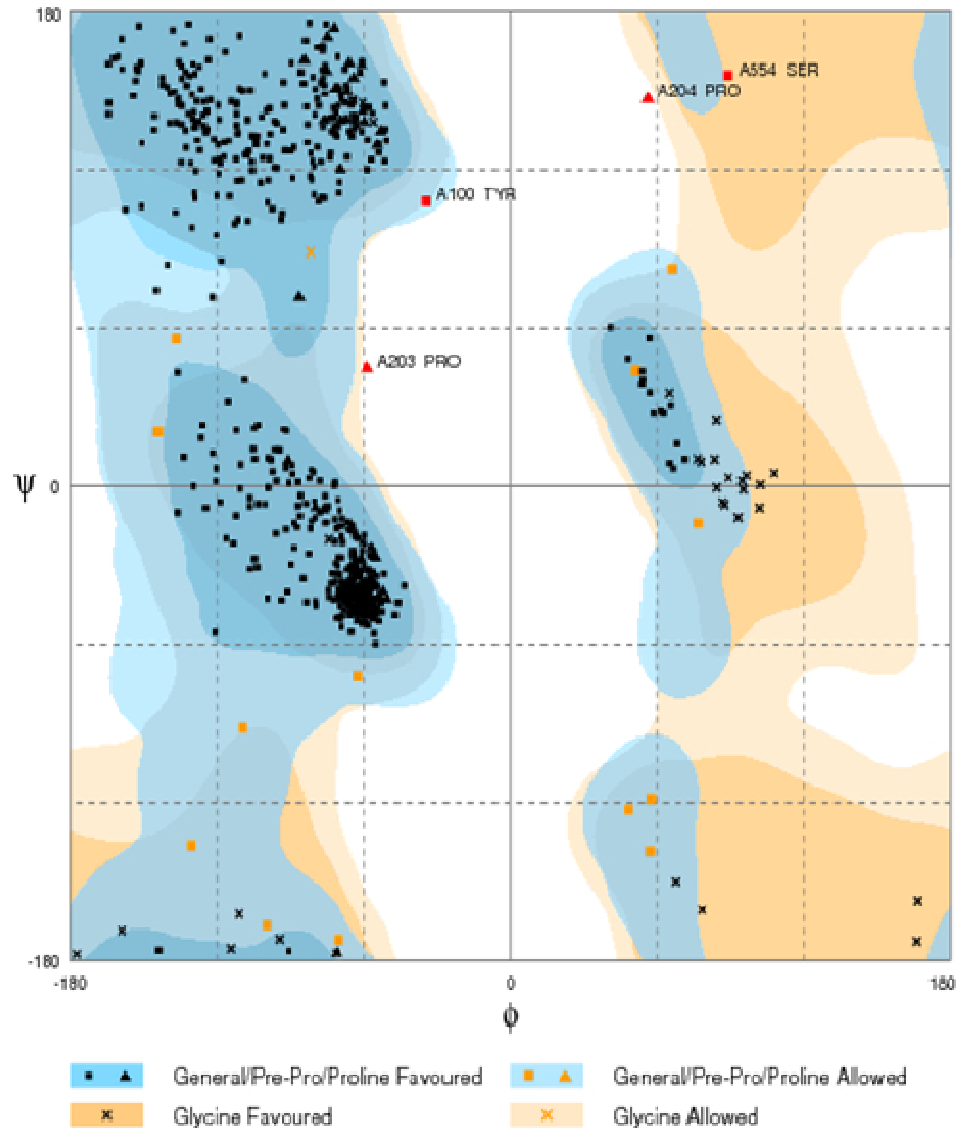


b

Figures 3. Ramachandran plot values (a, b) showing number of residues in favored, allowed and outlier region of model by

Table 5. Models built by Modeller and SAM_T08 server via threading approach along with their Z-score, packing quality control and model length.

Model	Z-score	Packing quality control	Model length
Modeller			
1	-1.895	-1.609	1-659
2 SAM_T08			
1	-0.132	-1.739	1-659



a

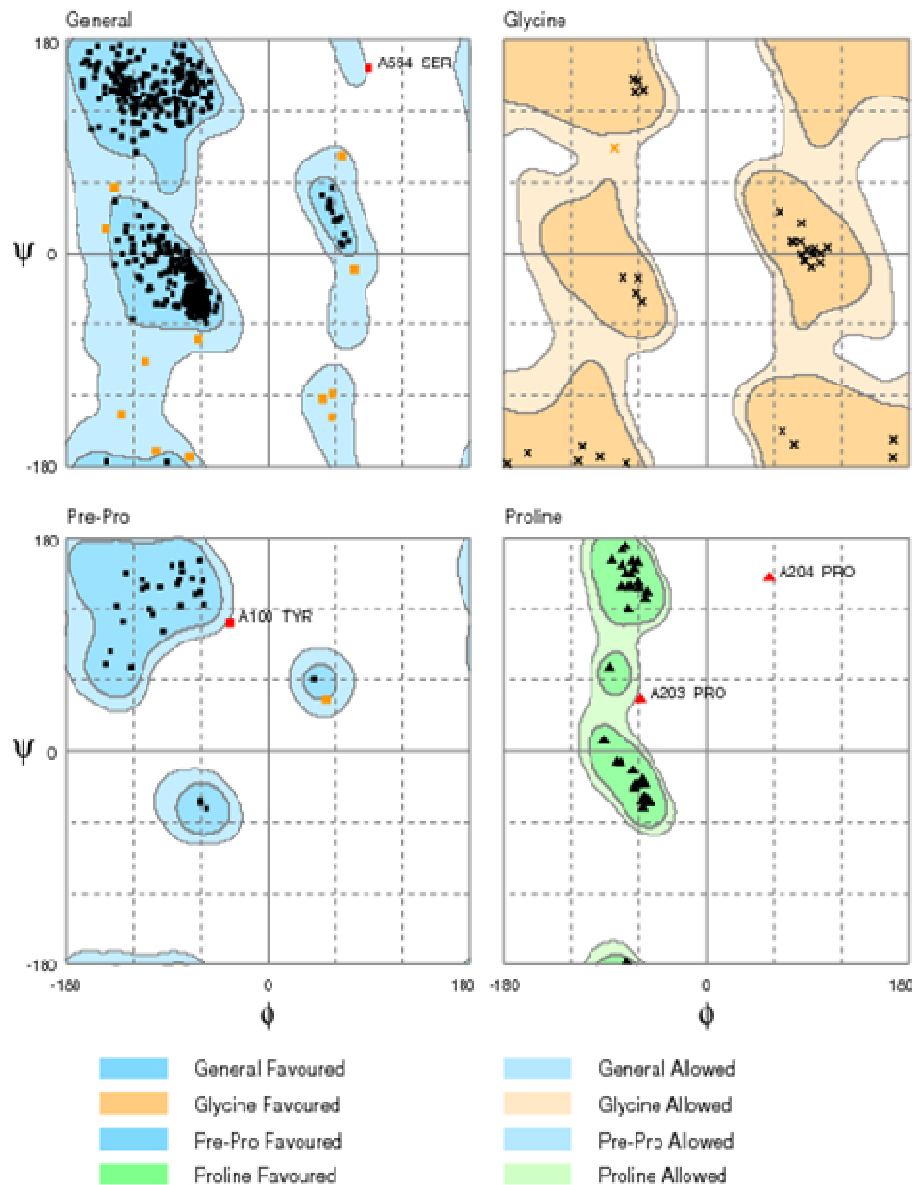
chain readily accessible to non-accessible side group. According to PoPMuSiC v2 web server, mutant protein stability changes analyzer, this mutation destabilizes the mutant protein that might shows the role of this amino acid (Valine) in the proper topology of active site.

In case of mutation at position 537, valine is mutated with glutamic acid. Valine is a non polar amino acid which is substituted with negatively charged hydrophilic amino acid (glutamic acid). In structural analysis (Figure 5b), there is an increased availability of side chain. This accounts for patients reported with low expression of Btk (Kanegane et al., 2000).

Danielian et al. (2003) reported sixteen missense mutations out of which four were selected due to their presence in kinase domain near Y551 residue. They reported that mutations occurred at positions 538 (Ser538Pro), 541 (Gly541Asp), 542 (Leu542Pro) and 562

(Arg562Trp). In case of serine to proline at position 538, the resulting phenotypes are XLA compatible. As depicted in Figure 5c, this mutation led to change in orientation of side chain more towards the interior of the protein due to hydrophobic nature of protein as proline gives structural rigidity.

In case of second mutation (Gly541Asp), glycine (a non polar aliphatic amino acid) replaced by aspartic acid (negatively charged aliphatic amino acid) is readily exposed and can interact with positive charges (Figure 5d). Third mutation (Leu542Pro) in kinase domain was reported in two patients of the same family. In this mutation, leucine at position 542 is substituted with proline (Figure 5e). Leucine is an aliphatic non polar hydrophobic amino acid which on substitution with proline led to the change in the orientation of side chain more towards the interior of the protein due to structural rigidity produced by proline.



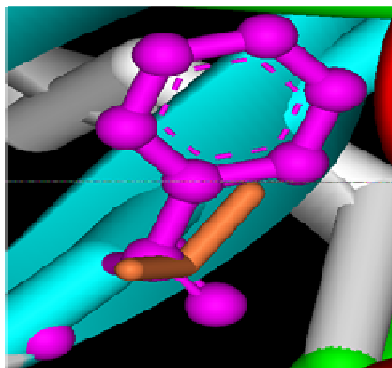
1

Figure 4. Ramachandran plot values (a, b) showing number of residues in favored, allowed and outlier region of model by SAM_T08 (Ramachandran plot).

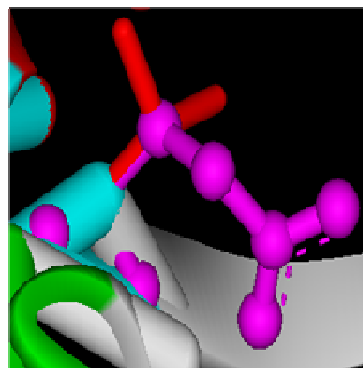
Danielian et al. (2003) reported mutation at position 562 (Arg562Trp). The position 562 is found to be critical as amino acid (arginine), is located near the catalytic site of protein and interact with the side chain of amino acid tryptophan at position 563, whereas, tryptophan at 563 keeps the integrity of PTK specific regions and also the side chain of 563 interacts with alanine at position 582 (act as a sandwich between 562 and 582 amino acids), however, this mutation disturbs the integrity (Vihinen et al., 1994). Arginine is a polar aliphatic amino acid; it is considered to bind the phosphate anion and is often found in the active centres of proteins that bind phosphorylated substrates. However, tryptophan is an

aromatic non polar amino acid that ultimately inactivates the protein to bind with phosphorylated substrates (Figure 5j).

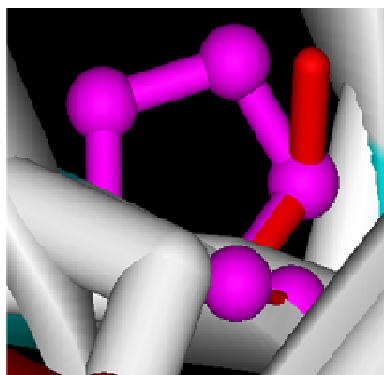
The amino acid arginine at position 544 is considered important due to its direct linkage to Y-551 via the hydroxyl group rather than phosphate. Transphosphorylation of Y-551 appears to trigger the exchange of hydrogen bonded pairs from E445/R544 to E445/K430 resulting in relocation of helices inducing Btk activation. Although, it is a less conserved residue but results in a sever XLA on mutation. Three reported missense mutations at this position are (1) Arginine to lysine (Kobayashi et al., 1996) (Figure 5f); (2) arginine to serine



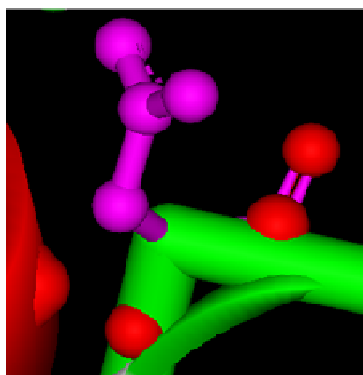
A



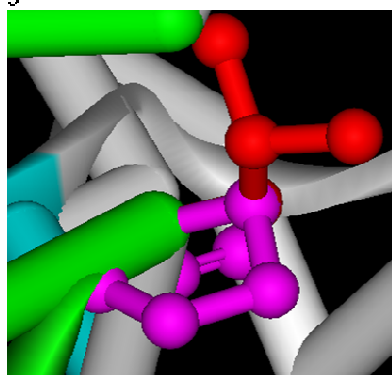
B



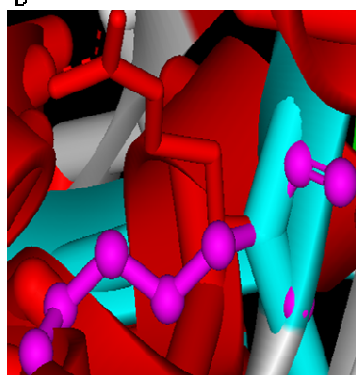
C



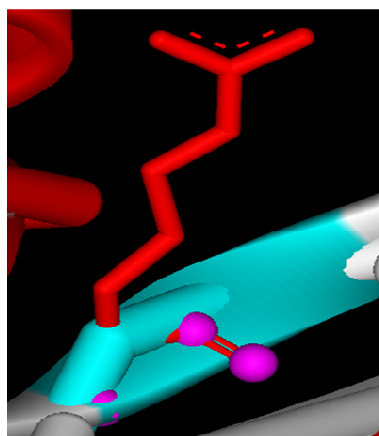
D



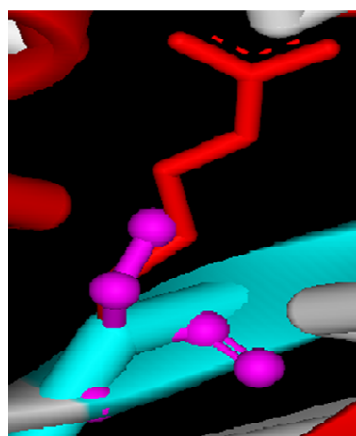
E



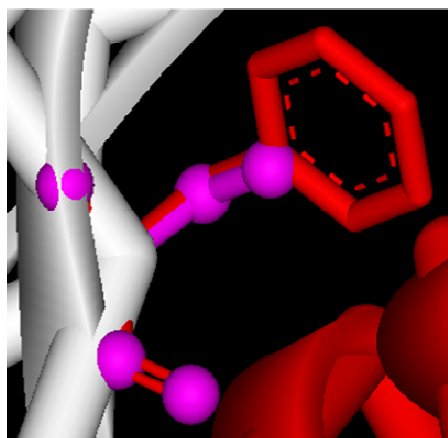
F



G



H



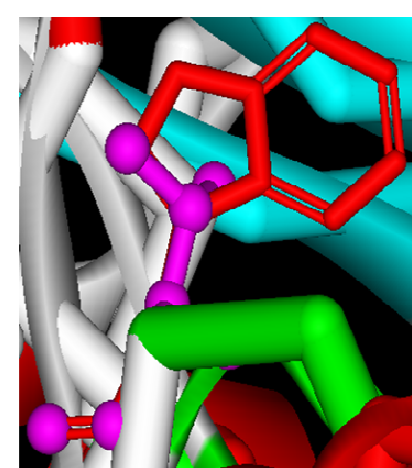
I



J



K



L

(Rodríguez et al., 2001) (Figure 5h); (3) arginine to glycine (Orlandi et al., 1999) (Figure 5g). Arginine is different from lysine residue in side chain length. It has more hydrogen bonding capability and interaction sites (Mao et al., 2001). Mutants of this residue may result in destabilized phosphorylation of Y-551. Lysine, serine and glycine are all structurally different residues resulting in the disruption of interactions. According to studies, it is suggested that mutations at position 544 are unlikely to abolish the kinase activity entirely. PopMusic v2 also suggested the destabilized structures of mutant of 544 position.

Position 559 is considered to be directly or indirectly involved in peptide substrate binding pocket (Mao et al., 2001). Phenylalanine at this position is an aromatic hydrophobic amino acid that on mutation converted into serine an aliphatic hydrophilic amino acid resulting into the exposure of the side chain (Holinoki-Feder, 1998). Due to this conversion, it is considered as the disruption of peptide substrate binding site (Figure 5i).

Holinoki-Feder (1998) also reported one more mutation at position 563 (tryptophan to leucine). Tryptophan-563

along with phenylalanine-559 is also involved either directly or indirectly in the peptide substrate binding and also, has hydrophobic interaction with alanine at position 523. This mutation may alter the substrate binding site as tryptophan is an aromatic hydrophobic amino acid, whereas leucine is an aliphatic amino acid with exposed side chains that may alter the conformation of the peptide substrate binding site (Figure 5l). Two groups reported two different mutations at same position 565. Stewart et al. (2001) reported the conversion of proline into leucine (Figure 5m) and Jo et al. (2003) reported the proline into threonine (Figure 5n). Due to structural rigidity produced by proline, the side chain is more interior; however, the leucine is an aliphatic amino acid with more bonding sites. So mutations at this position may alter the conformation of the protein structure leading to the unstable protein.

Proline, next to position 565 produces more structural rigidity when combined. Mutation at this position (Pro566Ser; Rodríguez et al., 2001) showed the altered conformation of the protein leading to the destabilized protein kinase activity (Table 6) (Figure 5o).

The proposed pocket of this study also contains amino

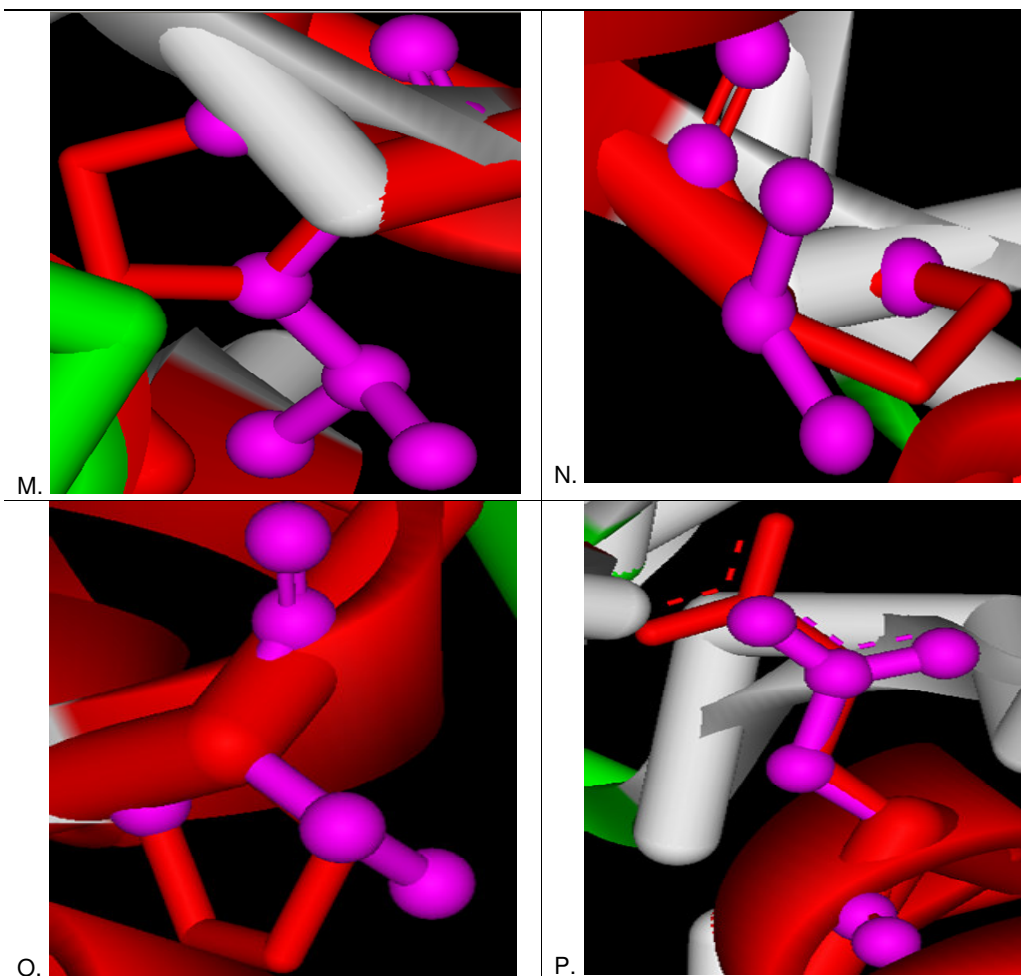


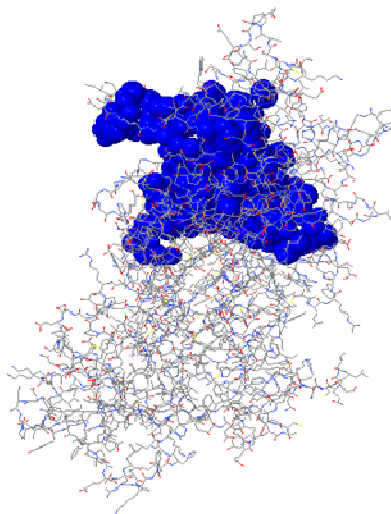
Figure 5. Superimposed mutated structures of kinase domain on predicted normal BTK structure in ViewerLite version 5.0: A, Valine (V) to phenylalanine (F) conversion at position 535. Valine is shown in STICK (Red) display style and phenylalanine is in BALL and STICK (Pink) display style; B, valine (V) to glutamic acid (E) conversion at position 537. Valine is shown in STICK (Red) display style and glutamic acid is in BALL and STICK (Pink) display style; C, serine (S) to proline (P) conversion at position 538. Serine is shown in STICK (Red) display style and proline is in BALL and STICK (Pink) display style; D, glycine (G) to aspartic acid (D) conversion at position 541. Glycine is shown in BALL and STICK (Red) display style and aspartic acid is in BALL and STICK (Pink) display style; E, leucine (L) to proline (P) conversion at position 542. Glycine is shown in BALL and STICK (Red) display style and proline is in BALL and STICK (Pink) display style; F, arginine (R) to lysine (K) conversion at position 544. Arginine is shown in STICK (Red) display style and lysine is in BALL and STICK (Pink) display style; G, arginine (R) to glycine (G) conversion at position 544. Arginine is shown in STICK (Red) display style and glycine is in BALL and STICK (Pink) display style; H, arginine (R) to serine (S) conversion at position 544. Arginine is shown in STICK (Red) display style and serine is in BALL and STICK (Pink) display style. I, phenylalanine (F) to serine (S) conversion at position 559. Phenylalanine is shown in STICK (Red) display style and serine is in BALL and STICK (Pink) display style; J, arginine (R) to proline (P) conversion at position 562. Arginine is shown in STICK (Red) display style and proline is in BALL and STICK (Pink) display style; K, arginine (R) to tryptophan (W) conversion at position 562. Arginine is shown in STICK (Red) display style and tryptophan is in BALL and STICK (Pink) display style; L, tryptophan (W) to leucine (L) conversion at position 563. Tryptophan is shown in STICK (Red) display style and leucine is in BALL and STICK (Pink) display style; M, proline (P) to leucine (L) conversion at position 565. Proline is shown in STICK (Red) display style and leucine is in BALL and STICK (Pink) display style; N, proline (P) to threonine (T) conversion at position 565. Proline is shown in STICK (Red) display style and threonine is in BALL and STICK (Pink) display style; O, proline (P) to serine (S) conversion at position 566. Proline is shown in STICK (Red) display style and serine is in BALL and STICK (Pink) display style; P, glutamic acid (E) to aspartic acid (D) conversion at position 567. Glutamic acid is shown in STICK (Red) display style and aspartic acid is in BALL and STICK (Pink) display style.

```

1-  MAAVILESIFLKR SQQKKKT SPLNFKKRLFLLT VHKLSYYEYDFERGRRG
51-  SKKGSIDVEKITCVETVYPEKNPPPERQIPRRGEESSSEMEQISIIERFPY
101- PFGVVYDEGPLYVFSPT EELRKRWIHQLKNVIRYNSDLVQKYHPCFWIDG
151- QYLCCSQTAKNAMGCOILENRNGSLKPGSSHRKTKKRLPPTPEEDQILK
201- PLPPEPAAAPVSTSELKKVVALYDYMMPNANDLQLRKGDEYFILLESNLP
251- WWRARDKNGQECYTPSNYVTEAEDSIEMYEWYSKHMTRSQAQQLLKQEGK
301- EGGFIVRDSSKAGKYTVSVFAKSTGDPQGVIRHYVVCSTPQSQYYLAEKH
351- LFTSTIPELINYHQHNSAGLIISRLKYPVSDQNKNA PSTAGLGYGSWEIDPK
401- DLTFFLKELGTFQFVVYKYGKWRGQYDVAIKMIKEGSMSEDEFIEEAKVMM
451- NLSHEKLVQLYGVGTKQRPIFIITEYMANGGLLNYLREMRHRFQTQQLLE
501- MCKDVCCEAMEYLESKGF LHRDLAARNCLVNDQGVVKVSD FGLSR YVLDDE
551- YTS SVGSKFPVRWSPPEVLMYSKFSSKSDI WAFGVL MWEIYSLGKMPYER
601- FTNSETA EHIAGQLRLYRPHLASEKVYTIMYSCWH EKADE RPTFKIILLSN
651- LDVMDDEES

```

A



B

Figure 6. Pocket of BTK structure generated by CASTp server. A, All the amino acids (highlighted with blue) involved in this pocket; B, Pocket with BTK structure having tyrosine 551.

Table 6. Selected reported missense mutations of kinase domain with their effect on the stability of protein by PopMusic v2.

S/N	Position	Nucleotide change	Protein	Mutant protein stability change $\Delta\Delta G$	Model figure
1	535	tGTT-TTT	Val-Phe	0.78 kcal/mol (destabilizing)	5a
2	537	GTA-GAA	Val-Glu	3.36 kcal/mol (destabilizing)	5b
3	538	aTCT-CCT	Ser-Pro	1.86 kcal/mol (destabilizing)	5c
4	541	GGC-GAC	Gly-Asp	1.03 kcal/mol (destabilizing)	5d
5	542	CTG-CCG	Leu-Pro	1.89 kcal/mol (destabilizing)	5e
6	544	AGG-AAG	Arg-Lys	1.78 kcal/mol (destabilizing)	5f
7	544	AGGt-AGT	Arg-Ser	1.85 kcal/mol (destabilizing)	5g
8	544	cAGG-GGG	Arg-Gly	2.01 kcal/mol (destabilizing)	5h
9	559	TTT-TCT	Phe-Ser	1.08 kcal/mol (destabilizing)	5i
10	562	cCGG-TGG	Arg-Trp	-0.04 kcal/mol (stabilizing)	5j
11	562	CGG-CCG	Arg-Pro	0.69 kcal/mol (destabilizing)	5j
12	563	TGG-TTG	Trp-Leu	1.89 kcal/mol (destabilizing)	5l
13	565	CCA-CTA	Pro-Leu	1.28 kcal/mol (destabilizing)	5m
14	565	cCCA-ACA	Pro-Thr	1.24 kcal/mol (destabilizing)	5n
15	566	aCCG-TCG	Pro-Ser	1.19 kcal/mol (destabilizing)	5o
16	567	GAAg-GAC	Glu-Asp	1.00 kcal/mol (destabilizing)	5p

acids of position 566 and 567. Glutamic acid at position 567 has more bonding capability when compared with aspartic acid (Figure 5p). Glutamic acid pairs with arginine 641 and also contributes in the activation of Btk (Mao et al., 2001). On mutation, protein has changed conformation resulting in a destabilized mutant protein.

Conclusions

This study clearly demonstrates the significance of *in silico* protein modelling and it can be useful for the prediction of change in protein structure and its functional consequences. In all the selected mutation cases, a useful protein was predicted by using different bioinformatics web servers and tools. It gave the study insight into changes in resultant Btk protein. This study showed that, selected mutations are involved in destabilizing the protein by changing their interactions with other residues of the protein. At the end, it can be easily summarized that such model generation save the time for lengthy experimental testing for preliminary studies. However, studied mutations validation can be done by using wet laboratory experiments in future.

ACKNOWLEDGEMENT

We acknowledge the effort and contribution of Human Molecular Genetics laboratory for physical contribution and providing us with technical assistance.

REFERENCES

- Altschul S, Madden T, Schaffer A, Zhang J, Zhang Z, Miller W, Lipman D (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25: 3389-3402.
- Altschul SF, John C, Wootton E, Michael G, Richa A, Aleksandr M, Alejandro AS, Yi KY (2005). Protein database searches using compositionally adjusted substitution matrices. *FEBS J.* 272: 5101-5109.
- Arnold K, Bordoli L, Kopp J, Schwede T (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*, 22: 195-201.
- Biegert A, Mayer C, Remmert M, Söding J, Lupas A (2006). The MPI Toolkit for protein sequence analysis. *Nucl. Acids Res.* 34: W335-339
- Binkowski TA, Naghibzadeh S, Liang J (2003). CASTp: Computed Atlas of Surface Topography of proteins. *Nucl. Acids Res.* 31: 3352-3355. BioGRID [<http://www.thebiogrid.org/about.php>]
- Bioinformatics Toolkit [<http://toolkit.tuebingen.mpg.de/>]
- Bolen JB (1995) Protein tyrosine kinases in the initiation of antigen receptor signaling. *Curr. Opin. Immunol.* 7: 306-311.
- Combet C, Jambon M, Deléage G, Geourjon C (2002). Geno3D: Automatic comparative molecular modelling of protein. *Bioinformatics*, 18: 213-214.
- Conley ME, Mathias D, Treadaway J, Minegishi Y, Rohrer J (1998). Mutations in Btk in patients with presumed X-linked agammaglobulinemia. *Am. J. Hum. Genet.* 62: 1034-1043.
- Cooper DN, Stenson PD, Chuzhanova NA (2005). The Human Gene Mutation Database (HGMD) and Its Exploitation in the Study of Mutational Mechanisms. *Current Protocols in Bioinformatics*. DOI: 10.1002/0471250953.bi0113s12.
- Danielian S, El-Hakeh J, Basílico G, Oleastro M, Rosenzweig S, Feldman G, Berozdnic L, Galicchio M, Gallardo A, Giraudi V, Liberatore D, Rivas EM, Zelazko M (2003). Bruton tyrosine kinase gene mutations in Argentina. *Hum. Mutat.* 21: 451.
- Dehouck Y, Grosfils A, Folch B, Gillis D, Bogaerts P, Rooman M (2009). Prediction of protein stability changes upon mutations using statistical potentials and neural networks: PoPMuSiC 2.0. *Bioinformatics*, 25: 2537-2543.
- Futatani T, Miyawaki T, Tsukada S, Hashimoto S, Kunikata T, Arai S, Kurimoto M, Niida Y, Matsuoka H, Sakiyama Y, Iwata T, Tsuchiya S, Tatsuzawa O, Yoshizaki K, Kishimoto T (1998). Deficient expression of Bruton's tyrosine kinase in monocytes from X-linked agammaglobulinemia as evaluated by a flow cytometric analysis and its clinical application to carrier detection. *Blood.* 91: 595-602.
- Gaspar HB, Ferrando M, Caragol I, Hernandez M, Bertran JM, De Gracia X, Lester T, Kinnon C, Ashton E, Espanol T (2000). Kinase mutant Btk results in atypical X-linked agammaglobulinemia phenotype. *Clin. Exp. Immunol.* 120: 346-350.
- Gaspar HB, Lester T, Levinsky RJ, Kinnon C (1998). Bruton's tyrosine kinase expression and activity in X-linked agammaglobulinemia (XLA): the use of protein analysis as a diagnostic indicator of XLA. *Clin. Exp. Immunol.* 111: 334-338.
- Holinoki-Feder E, Weiss M, Brandau O, Jedeke KB, Nore B, Bäckesjö CM, Vihinen M, Hubbard SR, Belohradsky BH, Smith CI, Meindl A (1998). Mutation screening of the BTK gene in 56 families with X-linked agammaglobulinemia (XLA): 47 unique mutations without correlation to clinical course. *Pediatrics*, 101: 276-284.
- HPRD Human Proteins Reference Database [<http://www.hprd.org>]
- Hyvonen M, Saraste M (1997). Structure of the PH domain and Btk motif from Bruton's tyrosine kinase: molecular explanations for X-linked agammaglobulinemia. *EMBO J.* 16:3396-3404.
- Jo EK, Wang Y, Kanegane H, Futatani T, Song CH, Park JK, Kim JS, Kim DS, Ahn KM, Lee SI, Park HJ, Hahn YS, Lee JH, Miyawaki T (2003). Identification of mutations in the Bruton's tyrosine kinase gene, including a novel genomic rearrangements resulting in large deletion, in Korean X-linked agammaglobulinemia patients. *J. Hum. Genet.* 48: 322-326.
- Jorg S, Frank M, Peer B, Chris PP (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl. Acad. Sci.* 95: 5857-5864.
- Kanegane H, Tsukada S, Iwata T, Futatani T, Nomura K, Yamamoto J, Yoshida T, Agematsu K, Komiyama A, Miyawaki T (2000). Detection of Bruton's tyrosine kinase mutations in hypogammaglobulinemic males registered as common variable immunodeficiency (CVID) in the Japanese Immunodeficiency Registry. *Clin. Exp. Immunol.* 120: 512-517.
- Karchin R, Cline M, Mandel-Gutfreund Y, Karplus K (2003). Hidden Markov models that use predicted local structure for fold recognition: alphabets of backbone geometry. *Proteins: Structure, Function, and Genetics.* 51: 504-514.
- Karplus K (2009). SAM-T08: HMM-based Protein Structure Prediction. *Nucleic Acids Research. Web Server Issue. Nucleic Acids Research Advance Access published online on May 29, 2009.* doi:10.1093/nar/gkp403
- Karplus K, Karchin R, Barrett C, Tu S, Cline M, Diekhans M, Grate L, Casper J, Hughey R (2001). What is the value added by human intervention in protein structure prediction? *Proteins: Structure Function and Genetics.* 45: 86-91.
- Katzman S, Barrett C, Thiltgen G, Karchin R, Karplus K (2008). Predict-2nd: a tool for generalized protein local structure prediction. *Bioinformatics.* 24: 2453-2459.
- Kobayashi S, Iwata T, Saito M, Iwasaki R, Matsumoto H, Naritaka S, Kono Y, Hayashi Y (1996). Mutations of the Btk gene in 12 unrelated families with X-linked agammaglobulinemia in Japan. *Hum. Genet.* 97: 424-430.
- Lambert C, Leonard N, De Bolle X, Depiereux E (2002). ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics*, 18: 1250-1256.
- Lopez-Herrera G, Berron-Ruiz L, Mogica-Martinez D, Espinosa-Rosales F, Santos-Argumedo L (2008). Characterization of Bruton's tyrosine kinase mutations in Mexican patients with X-linked agammaglobulinemia. *Mol. Immunol.* 45: 1094-1098.
- Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC (2002). Structure validation by

- Alpha geometry: phi,psi and Cbeta deviation. *Proteins: Structure, Function and Genetics*. 50: 437-450.
- Lund O, Nielsen M, Lundegaard C, Worning P (2002). CPHmodels 2.0: X3M a Computer Program to Extract 3D Models. Abstract at the CASP5 conferenceA102.
- Manna D, Albanese A, Park WS, Cho W (2007). Mechanistic basis of differential cellular responses of phosphatidylinositol 3,4-bisphosphate- and phosphatidylinositol 3,4,5-trisphosphate-binding pleckstrin homology domains. *J. Biol. Chem.* 282: 32093-32105.
- Mao C, Zhou M, Uckun FM (2001). Crystal structure of Bruton's tyrosine kinase domain suggests a novel pathway for activation and provides insights into the molecular basis of X-linked agammaglobulinemia. *J. Biol. Chem.* 276: 41435-41443.
- Mohamed AJ, Vargas L, Nore BF, Bäckesjö CM, Christensson B, Smith CI (2000). Nucleocytoplasmic shuttling of Bruton's tyrosine kinase. *J. Biol. Chem.* 275:40614-40619.
- Mohamed AJ, Yu L, Bäckesjö CM, Vargas L, Faryal R, Aints A, Christensson B, Berglöf A, Vihinen M, Nore BF, Smith CI (2009). Bruton's tyrosine kinase (Btk): function, regulation, and transformation with special emphasis on the PH domain. *Immunol. Rev.* 228(1): 58-73.
- Murayama K K-MM, Mishima C, Akasaka R, Shirouzu M, Fukui Y, Yokoyama S (2008). Crystal structure of the Bruton's tyrosine kinase PH domain with phosphatidylinositol. *Biochem. Biophys. Res. Commun.* 377:23-28.
- OMIM [<http://www.ncbi.nlm.nih.gov/omim/>]
- Orlandi P, Ritis K, Moschese V, Angelini F, Arvanitidis K, Speletas M, Sideras P, Plebani A, Rossi P (1999). Identification of nine novel mutations in the Bruton's tyrosine kinase gene in X-linked agammaglobulinaemia patients. *Hum. Mutat.* 15: 117.
- PDB [<http://www.pdb.org/pdb/home/home.do>]
- PoPMuSiC v2.0 [<http://babylone.ulb.ac.be/PoPV2/index.php?pg=single>]
- Rodríguez MC, Granados EL, Cerdán AF, Casariego GF(2001). Molecular analysis of Bruton's tyrosine kinase gene in Spain. *Hum. Mutat.* 18: 84.
- Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M (1995). Evaluation of comparative protein modelling by MODELLER. *Proteins*. 23: 318-326.
- Shackelford G, Karplus K (2007). Contact Prediction using Mutual Information and Neural Nets. *Proteins: Structure, Function, and Bioinformatics*. 69: 159-164.
- Smith CI, Islam KB, Vorechovský I, Olerup O, Wallin E, Rabbani H, Baskin B, Hammarström L (1994). X-linked agammaglobulinemia and other immunoglobulin deficiencies. *Immunol Rev.* 138: 159-83.
- Stewart DM, Lian L, Nelson DL (2001). The clinical spectrum of Bruton's agammaglobulinemia. *Curr. Allergy Asthma Rep.* 1: 558-565.
- String [http://string.embl.de/newstring.cgi/show_network_section.pl]
- Väliaho J, Smith CI, Vihinen M (2006). BTKbase: the mutation database for X-linked agammaglobulinemia. *Hum. Mutat.* 27: 1209-1217.
- Vihinen M, Nilsson L, Smith CI (1994) Tec homology (TH) adjacent to the PH domain. *FEBS Lett.* 350: 263-265.
- Vihinen M, Nore BF, Mattsson PT, Bäckesjö CM, Nars M, Koutaniemi S, Watanabe C, Lester T, Jones A, Ochs HD, Smith CI (1997). Missense mutations affecting a conserved cysteine pair in the TH domain of Btk. *FEBS Lett.* 413(2): 205-210.
- Whatif [<http://swift.cmbi.ru.nl/servers/html/>]