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 Zn$^{2+}$-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 downstream signaling, which leads to pleiotrophic 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), 3Dgissaw, CPH models (Lund et al., 2002), ESyPred3D (Lambert et al., 2002) and Geno3d (Comb 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 databank, therefore, this study considered some of them (Table 4) as multiple templates using Psi-Blast (Altshul et al., 1997), with matrix BLOSUM62 (Altshul et al., 2005). Aligned sequences of all the templates with Btk sequence were allowed to model using the modeler sv3 (Sai 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 Zn\(^{2+}\)ion is mediated by a this zinc finger motif (Smith et al., 1994; Murayama et al., 2008). Mutations affecting PH and TH domains affect Zn\(^{2+}\) binding, hence, leading to the generation of extremely unstable protein (Vihinen et al., 1997; Hyvonen and Saraste, 1997). Btk is a metalloprotein enzyme, requiring Zn\(^{2+}\) 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).

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

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

<table>
<thead>
<tr>
<th>Model</th>
<th>Z-score</th>
<th>Model length</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 DJIGSAW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-2.576</td>
<td>2-170</td>
</tr>
<tr>
<td>GENO3D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-24.18</td>
<td>1-169</td>
</tr>
<tr>
<td>2</td>
<td>-24.18</td>
<td>1-169</td>
</tr>
<tr>
<td>3</td>
<td>-24.18</td>
<td>1-169</td>
</tr>
<tr>
<td>SWISS-MODEL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-2.072</td>
<td>2-170</td>
</tr>
<tr>
<td>2</td>
<td>-4.824</td>
<td>270-350</td>
</tr>
<tr>
<td>3</td>
<td>-4.228</td>
<td>216-273</td>
</tr>
<tr>
<td>4</td>
<td>-3.130</td>
<td>181-271</td>
</tr>
<tr>
<td>5</td>
<td>-6.019</td>
<td>130-196</td>
</tr>
<tr>
<td>6</td>
<td>-1.241</td>
<td>346-478</td>
</tr>
<tr>
<td>CPHMODELS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-3.966</td>
<td>397-654</td>
</tr>
<tr>
<td>ESYPred3D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.653</td>
<td>252-657</td>
</tr>
</tbody>
</table>
Table 3. Ramachandran plot values of all the models generated by all the web servers obtained through RAMPAGE.

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

Table 4. Multiple templates used for MODELLER.

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

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
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
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
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 Modeler and SAM_T08 server via threading approach along with their Z-score, packing quality control and model length.

<table>
<thead>
<tr>
<th>Model</th>
<th>Z-score</th>
<th>Packing quality control</th>
<th>Model length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeller</td>
<td>1</td>
<td>-1.895</td>
<td>1-659</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.609</td>
<td></td>
</tr>
<tr>
<td>2 SAM_T08</td>
<td>1</td>
<td>-0.132</td>
<td>1-659</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.739</td>
<td></td>
</tr>
</tbody>
</table>
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.
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 5f).

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 helixes 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
(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 (Holinkoki-Feder, 1998). Due to this conversion, it is considered as the disruption of peptide substrate binding site (Figure 5i).

Holinkoki-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 5i). 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.
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.

<table>
<thead>
<tr>
<th>S/N</th>
<th>Position</th>
<th>Nucleotide change</th>
<th>Protein</th>
<th>Mutant protein stability change</th>
<th>ΔΔG</th>
<th>Model figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>535</td>
<td>tGTT-TTT</td>
<td>Val-Phe</td>
<td>0.78 kcal/mol (destabilizing)</td>
<td>5a</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>537</td>
<td>GTA-GAA</td>
<td>Val-Glu</td>
<td>3.36 kcal/mol (destabilizing)</td>
<td>5b</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>538</td>
<td>aTCT-CCT</td>
<td>Ser-Pro</td>
<td>1.86 kcal/mol (destabilizing)</td>
<td>5c</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>541</td>
<td>GGC-GAC</td>
<td>Gly-Asp</td>
<td>1.03 kcal/mol (destabilizing)</td>
<td>5d</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>542</td>
<td>CTG-CCG</td>
<td>Leu-Pro</td>
<td>1.89 kcal/mol (destabilizing)</td>
<td>5e</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>544</td>
<td>AGG-AAG</td>
<td>Arg-Lys</td>
<td>1.78 kcal/mol (destabilizing)</td>
<td>5f</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>544</td>
<td>AGGt-AGT</td>
<td>Arg-Ser</td>
<td>1.85 kcal/mol (destabilizing)</td>
<td>5g</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>544</td>
<td>cAGG-GGG</td>
<td>Arg-Gly</td>
<td>2.01 kcal/mol (destabilizing)</td>
<td>5h</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>559</td>
<td>TTT-TCT</td>
<td>Phe-Ser</td>
<td>1.08 kcal/mol (destabilizing)</td>
<td>5i</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>562</td>
<td>cCGG-TGG</td>
<td>Arg-Trp</td>
<td>-0.04 kcal/mol (stabilizing)</td>
<td>5j</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>562</td>
<td>CGG-CCG</td>
<td>Arg-Pro</td>
<td>0.69 kcal/mol (destabilizing)</td>
<td>5j</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>563</td>
<td>TGG-TTG</td>
<td>Trp-Leu</td>
<td>1.89 kcal/mol (destabilizing)</td>
<td>5l</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>565</td>
<td>CCA-CTA</td>
<td>Pro-Leu</td>
<td>1.28 kcal/mol (destabilizing)</td>
<td>5m</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>565</td>
<td>cCCA-ACA</td>
<td>Pro-Thr</td>
<td>1.24 kcal/mol (destabilizing)</td>
<td>5n</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>566</td>
<td>aCCG-TCG</td>
<td>Pro-Ser</td>
<td>1.19 kcal/mol (destabilizing)</td>
<td>5o</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>567</td>
<td>GAAg-GAC</td>
<td>Glu-Asp</td>
<td>1.00 kcal/mol (destabilizing)</td>
<td>5p</td>
<td></td>
</tr>
</tbody>
</table>
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., 2008). 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


BiGRID [http://www.thebiogrid.org/about.php]

Bioinformatics Toolkit [http://toolkit.tuebingen.mpg.de/]


Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC (2002). Structure validation by

Iqbal et al.       3287

PDB [http://www.pdb.org/pdb/home/home.do]
PoPMuSiC v2.0 [http://babylone.ulb.ac.be/PoPV2/index.php?pg= single]
String [http://string.embl.de/newstring_cgi/show_network_section.pl]
WhatIf [http://swift.cmbi.ru.nl/servers/html/]

