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

Chen Mao‡¶, Min Zhou‡, and Fatih M. Uckun¶¶

From the Departments of ¶¶Structural Biology and ¶¶Immunology, Parker Hughes Cancer Center, St. Paul, Minnesota 55113

Bruton’s tyrosine kinase (BTK), a member of the BTK/TEC family of cytoplasmic protein-tyrosine kinases (PTKs), is intimately involved in signal transduction pathways regulating survival, activation, proliferation, and differentiation of B lineage lymphoid cells. Mutations in the human btk gene are the cause of X-linked agammaglobulinemia, a male immune deficiency disorder characterized by a lack of mature, immunoglobulin-producing B lymphocytes. We have determined the x-ray crystal structure of the Bruton’s tyrosine kinase domain in its unphosphorylated state to a 2.1 Å resolution. A comparison with the structures of other tyrosine kinases and a possible mechanism of activation unique to Bruton’s tyrosine kinase are provided.

Bruton’s tyrosine kinase (BTK), a member of the BTK/TEC family of cytoplasmic protein-tyrosine kinases (PTKs), is intimately involved in signal transduction pathways regulating survival, activation, proliferation, and differentiation of B lineage lymphoid cells (1–5). BTK participates in signal transduction pathways initiated by the binding of a variety of extracellular ligands to their cell surface receptors (3). Following ligation of B cell antigen receptors, BTK activation by the concerted actions of the PTKs LYN and SYK is required for ligation of B cell antigen receptors, BTK activation by the ligand binding site (2, 12–17). Phosphorylation of Tyr-223 may partially activate BTK and cause XLA (11).

BTK has two regulatory tyrosine residues, Tyr-223 and Tyr-551, which participate in kinase activation (12). BTK is initially activated by trans-phosphorylation of Tyr-551 on the presumed “activation loop” (A loop) of the kinase domain by a SRC family PTK (e.g. LYN), followed by stimulating autophosphorylation of the Tyr-223 residue within the SH3 domain ligand binding site (2, 12–17). Phosphorylation of Tyr-223 may function to disrupt an intramolecular TEC homology domain-SH3 domain interaction, allowing the BTK TEC homology domain to bind SH3 domains of SRC family PTK and PH domain to bind a proline-rich region of CBL (3, 18, 19). The molecular mechanism by which the phosphorylation of Tyr-551 within the kinase domain leads to the initial activation of BTK has not been deciphered. Similarly, there is no structural understanding of how XLA is caused by various mutations within the BTK kinase domain.

In this paper, we report the x-ray crystal structure of the kinase domain of murine BTK to 2.1 Å resolution. The structure was determined by multiple isomorphous replacement and shows the architecture of a typical kinase and reveals a unique mechanism of activation. The N-terminal lobe (residues 397–475) contains five strands of antiparallel β sheets and one α helix. The C-terminal lobe (residues 479–659) contains a 4-helix bundle flanked by a short antiparallel β sheet and four additional helices. The N-terminal and C-terminal lobes are connected by a linker region (residues 475–479) and form a cleft at the ATP binding site. In contrast to previous predictions, the catalytic cleft of the BTK-KD is not occluded by the A loop or by any other portion of the KD. The A loop in the unphosphorylated BTK-KD structure adopts a unique non-ninhibitory conformation very similar to the active state conformation of phosphorylated LCK-KD, and hence does not limit substrate access to the active site. Because of the inactive conformation of helix C, however, the enzyme is not in the active state. Based on the BTK structure, we propose that the trans-phosphorylation of Tyr-551 can lead to BTK activation by

SH2, and a C-terminal SH1 kinase domain. The SH2 domain mediates binding to tyrosine-phosphorylated peptide motifs on other molecules, and the SH3 domain mediates binding to proline-rich motifs. Mutations in the SH1 domain, SH2 domain, as well as the PH domain of human BTK have been found to cause maturational blocks at early stages of B cell ontogeny leading to XLA (9). BTK-deficient mice generated by introducing PH domain or SH1 domain mutations in embryonic stem cells exhibit defective B cell development and function (10). Thus, different domains of BTK are important for its physiologic functions. The crystal structure of the PH domain has been determined and contributed to a better structural understanding of how point mutations of the PH domain can inactivate BTK and cause XLA (11).

This paper is available on line at http://www.jbc.org

Received for publication, May 28, 2001, and in revised form, August 28, 2001
Published, JBC Papers in Press, August 29, 2001, DOI 10.1074/jbc.M104828200

© 2001 by The American Society for Biochemistry and Molecular Biology, Inc.

The abbreviations used are: BTK, Bruton’s tyrosine kinase; PTK, protein-tyrosine kinase; XLA, X-linked agammaglobulinemia; PH domain, pleckstrin homology domain; SH domain, Src homology domain; A loop, activation loop; KD, kinase domain; PCR, polymerase chain reaction; DTT, dithiothreitol; IBK, insulin receptor kinase; AMP-PNP, adenosine 5’-(β,γ-imino)triphosphate; LCK, lymphocyte kinase; SRC, Rous sarcoma virus; cAPK, cyclic-AMP-dependent kinase; rTEV, recombinant tobacco etch virus.
Data for the outermost resolution shell are given in parentheses.

\[ R_{\text{sym}} = \frac{1}{N} \sum_{h} \left( \sum_{i} (I(h) - \langle I \rangle) \right)^2 \sum_{i} I(h) + 100, \]

where \( I(h) \) is the \( i \)th measurement of reflection \( h \) and \( \langle I \rangle \) is the mean value of \( N \) equivalent reflections. \( R_{\text{free}} = 2\langle F_{\text{obs}} \rangle - F_{\text{calc}} \sqrt{2 \langle F_{\text{obs}} \rangle^2 + \langle F_{\text{calc}} \rangle^2} \) for all centric reflections. Phasing power = \( \text{rms}(F_{\text{obs}})/E \), where \( F_{\text{obs}} \) is the heavy atom structure factor amplitude, and \( E \) is the residual lack of closure. \( R_{\text{crys}} = \sum |F_{\text{obs}}| - |F_{\text{calc}}| \sum |F_{\text{obs}}| \), where summation is over data used in the refinement. \( R_{\text{crys}} \) is the same as \( R_{\text{crys}} \) but only includes 5% of data excluded from refinement.

| Crystal | Nat 1 | Nat 2 | EMP | Au |
|---------|-------|-------|-----|----|
| X-ray source, wavelength (Å) | CHESS, \( \lambda = 1.009 \) | CHESS, \( \lambda = 1.009 \) | CHESS, \( \lambda = 1.009 \) | CHESS, \( \lambda = 1.009 \) |
| Resolution limits (Å) | 2.1 | 3.0 | 3.0 | 3.0 |
| Unique reflections | 25,284 | 10,020 | 11,391 | 10,028 |
| Completeness (%) | 86.7 (79)* | 96.5 (79) | 88.3 (68) |
| \( R_{\text{sym}} \) (%) | 5.4 (20) | 8.7 (21) | 8.0 (19) | 8.7 (23) |
| MIRES analysis | | | | |
| Resolution range (Å) | 15–4 | 15–4 | 66 | 22 |
| Sites | 4 | 6 | 1.21 | 0.73 |
| Cullis R factor (%) | | | 0.6 |
| Phasing power | | | | |
| Overall POM | | | | |
| Refinement | | | | |
| Resolution range (Å) | 5–2.1 | 20,370 | 30% | 80% |
| Reflections used (\( F > 2\sigma \)) | 4,180 | | | |
| Total nonhydrogen atoms (protein) | 180 | | | |
| Water molecules | 22.1 | | | |
| \( R_{\text{sym}} \) (%) | 28.7 | | | |
| \( R_{\text{free}} \) (%) | 0.06 | | | |
| R\( \text{free}\) from ideal bond length | 1.5 | | | |
| Average B (Å\(^2\)) for 95% of protein atoms with a B factor < 40.0 Å\(^2\) | 19.8 | | | |

*Reactions were not used in the resolution bins near or on ice rings.

The fractions that contain BTK-KD were pooled together, dialyzed against a solution that contained 20 mM Tris/HCl, 100 mM NaCl, 2 mM DTT, and 1 mM EDTA. Then the protein was concentrated to 5 mg/ml, digested with 6 μg/mg rTEV protease at 4 °C overnight, and then concentrated to 3 ml, which was then loaded on a Superdex 200 column (20/60), equilibrated with a solution of 20 mM Tris/HCl, pH 8.5, 50 mM NaCl, and 2 mM DTT. The pure BTK-KD fraction was concentrated to 12 mg/ml for crystallization.

**Dynamic Light Scattering**—The pure protein at a concentration of 2 mg/ml was used for dynamic light scattering studies. The data indicated that the BTK-KD sample was monodisperse as measured immediately after purification and became polydisperse in 30 min at room temperature. Therefore, all of the crystallization trays were set up at 4 °C.

**Cryocrystallization**—The purified BTK-KD was crystallized using the hanging drop method, with 2 μl of protein and 1 μl of reservoir solution in the drop and equilibrated with 600 μl of reservoir solution, which contains 20% polyethylene glycol 1000, 100 mM Tris, pH 8.0, and 1 mM DTT using VDP trays (Hampton Research, Laguna Hills, CA). The crystals began to appear in a week. It took 2–3 weeks to reach the full size of 200 x 200 x 200 μm.

**Data Collection**—Two native data sets and two heavy atom derivative data sets were collected using a CCD detector and high intensity synchrotron radiation source at the Cornell High Energy Synchrotron Source (CHESS). Data statistics are summarized in Table I. One crystal was soaked for 24 h in 5 mM KAu(CN)\(_5\), and another crystal was soaked in 0.1 mM ethylene mercury phosphate. All crystals were flash cooled and kept in liquid nitrogen with 30% polyethylene glycol 1000 solution as a natural cryoprotectant. All data sets were processed using DENZO and SCALEPACK programs (20) and then analyzed and scaled using the CCP4 program suite (21).

**Molecular Replacement**—The native data set was first analyzed for molecular replacement solution using AmoRe (22). The self-rotation function calculation failed to indicate any noncrystallographic symmetry, suggesting that there is either no noncrystallographic symmetry or that the noncrystallographic symmetry axis is too close to the crystallographic axis (in fact, the 2-fold axis was later found to be close to the x axis). The search models of polyglycine and polyalanine from crystal structures of kinase domains including LCK, insulin receptor kinase (IRR), fibroblast growth factor receptor kinase, haematopoietic cell...
kinase, and SRC kinase, as well as the homology BTK-KD model revealed a clear solution in rotation function searches with a top peak 10–25% higher than the second solution, depending on the search model used in the calculation. Interestingly, the search model of LCK with the A loop yielded better results, but the apo-IRK model with the A loop produced noisy results. This agrees with the final refined BTK structure in which the A loop in BTK is similar to phospho-LCK but not to IRK. The following translational searches using both the polyglycine/polyalanine models and the models that maintain conserved residues failed to produce a convincing solution with confidence, probably because of the existence of the two molecules in the asymmetric unit and the significant difference between the search models and the target structure.

Multiple Isomorphous Replacement—Meanwhile, a heavy atom derivatized search using a Patterson function calculation (using the RSPS program) followed by difference Fourier calculation (FFT) in the CCP4 suite (21) based on the data sets from the ethylene mercury phosphate-soaked crystal and the K2Au(CN)2-soaked crystal led to the identification of four mercury sites and six gold sites which were then used to phase all reflections at 4.0 Å and above by MLPHARE. The correct hand of the protein was determined by a sign of anomalous occupancies during the heavy atom site refinement. The map calculated based on the combination of phase information from the mercury and gold derivatives including anomalous diffraction data of mercury atoms in the ethylene mercury phosphate derivative was subjected to several rounds of phase improvement using the DM program with a solvent flattening and histogram matching method. The improved maps demonstrated clear boundaries between the target protein and solvent region. The map indicated two BTK molecules in one asymmetric unit. Therefore, two polyglycine/polyalanine kinase search models were fit manually into the electron density, which revealed how the two molecules are related by noncrystallographic symmetry. The matrix that relates the orientation of the two molecules was generated and then refined to two different matrices that corresponded to the N-terminal lobe domain and the C-terminal lobe domain, respectively, using MAMA and 6D_IMP (23–26). The refined matrices were then used in multidomain averaging of the electron density map using 6D_AVE. The final map was improved considerably from the previous map and the α helix and β strand structures and many large amino acids were clearly visible. The 500 residues and ~4000 atoms for two BTK molecules based on the BTK amino acid sequence were fit readily into the electron density. The entire amino acid sequences of BTK fragments from residue 397 to residue 659 were mostly traceable in the electron density except for a few disordered residues on the surface.

Crystallographic Refinement—The kinase domain structures of the two BTK molecules were refined using simulated annealing in X-PLOR (27), and numerous structural adjustments were performed with the help of CHAIN (28) and O programs (29). The refinement statistics are summarized in Table I.

Results

Purification and Crystallization of BTK-KD—His-tagged BTK-KD was purified using a nickel column, digested with rTEV protease to cleave off the His tag, and concentrated, as described under “Experimental Procedures.” The purity of the protein was confirmed by Coomassie Blue staining and Western blot analysis using an anti-BTK antibody. Anti-phosphotyrosine Western blot analysis and in vitro kinase assays of the purified BTK-KD showed that it is not phosphorylated but is capable of undergoing autophosphorylation as well as phosphorylating a glutathione S-transferase fusion protein of Igα in a 5-min kinase reaction. Purified BTK-KD was concentrated to 1.2 mg/ml and crystallized using the hanging drop method. The obtained crystals were analyzed using a CCD detector and high intensity synchrotron radiation source.

Structure Determination—The x-ray crystal structure of BTK-KD was determined by multiple isomorphous replacement. Data statistics are summarized in Table I. The final structure has been refined at 2.1 Å resolution to an R value of 22% with all amino acid residues falling into favored or generously allowed regions in the Ramachandran plot (except for glycine residues), as indicated by PROCHECK (30). The average B factor for all non-hydrogen atoms is 21 Å² and is below 20 Å² for more than half of the protein atoms. Two short regions that are disordered in the electron density map include part of the β1 strand (residues 409–412) and part of the A loop (residues 550 and 555). Other regions that display a visible electron density in the original map and have high B factors include the loop from residues 546 to 549, the loops around residue 467, and the N-terminal end of helix C which has B factors mostly in the 20–30 Å² range.

Discussion

Overall Architecture of the BTK-KD—The crystal structure of the BTK-KD in its unphosphorylated state is shown in Fig. 1. BTK-KD is packed in a dimeric form in the crystal lattice but is present mainly in monomeric form in solution (data not shown). BTK-KD has a two-lobe fold reminiscent of the topology of other PTK kinase domain structures (31–33). The secondary structure of the BTK-KD is labeled in Fig. 1 using the established nomenclature (34, 35). The N-terminal lobe (residues 397–475) contains five strands of antiparallel β sheets (β1–β5) and one α helix (C helix) (shown in yellow, Fig. 1). The C-terminal lobe (residues 479–659) contains a four-helix bundle (αD, αE, αF, and αH) flanked by a short antiparallel β sheet (β6, β8, and β9) and four additional helices (αI, αD, αE, and αH). The N- and C-terminal lobes are connected by a linker region (residues 475–479) and form a cleft at the ATP binding site.

The rotation of the N-terminal lobe relative to the C-terminal lobe varies among the different KD structures, and the ATP binding cleft between the two lobes is closed when substrates (or analogs) are bound (22, 36). Therefore, both lobes of a KD need to adopt a mandatory closed conformation for the KD to achieve a catalytically active state. Notably, the crystal structure of the unphosphorylated BTK-KD revealed that its N- and C-terminal lobes adopt a closed conformation very similar to the reported conformation of the lobes of the phosphorylated LCK-KD (rms deviation between the backbones of the central portion of β3 and β5 < 1 Å) (34), and the central regions of the β3 and β5 strands are almost superimposable when the C-terminal lobes are overlaid. However, the β1 and β2 strands, very much like the corresponding β strands of the cAPK-KD, adopt a more “closed” conformation than their counterparts in LCK-KD. The rotation needed to superimpose the N-terminal lobe of the BTK-KD onto the N-terminal lobe of the cAPK-KD is only 5.2°, whereas an 11° rotation is required to overlay the N-terminal lobe of the LCK-KD with the N-terminal lobe of the cAPK-KD (22). By comparison, the magnitudes of the rotation needed to open the BTK-KD and LCK-KD N-terminal lobes to match the conformation observed in the apo-IRK structure are 22° and 17° (22), respectively.

In contrast to the similarity of the closed conformation of their N- and C-terminal lobes, the C-terminal ends of the helix αC in unphosphorylated BTK-KD versus phosphorylated LCK-KD are markedly different. The distance between the Cα position of residue 440 in BTK-KD and its counterpart in LCK is 6.7 Å. From the pivot point at residue 452, helix αC of BTK-KD needs to be rotated toward the ATP binding cleft by 20° to be superimposed onto the helix αC of the LCK-KD. The helix αC in phosphorylated LCK-KD adopts a closed conformation consistent with a catalytically active state (34). By comparison, the helix αC of the BTK-KD adopts a more open conformation than that of the LCK-KD. The conformation of helix αC of the BTK-KD is different from the open conformation of helix C in c-SRC as well (see Fig. 2a), in accordance with a unique conformation of the A loop in BTK-KD. Significant
structural differences between BTK-KD and LCK-KD were also found in helices H9251, H9251, H9251, and H9251 (which differed in location by 2 Å), and in the glycine loop (H9252–H9252). The H9252 strand of the glycine loop is highly flexible and was observed in two distinct alternative conformations in the BTK-KD crystal structure. One conformation of the H9252 strand is similar to the conformation of the corresponding H9252 strand in cAPK-KD, whereas the other conformation places residues 410–412 in a position that allows the triphosphate of ATP to bind BTK molecule A. A portion of the glycine loop is disordered in BTK molecule B, which is related to molecule A by a noncrystallographic 2-fold axis. The adopted conformation of the invariant PTK residues 439DFG441 in BTK-KD is consistent with the conformation of the same residues in the apo-IRK structure (37).

Unique Noninhibitory Conformation of the Activation Loop in the BTK-KD—Although it is highly mobile (B factors of 30–40 Å²), the A loop of BTK-KD, except for residues 553–556, is visible within the electron density map. The A loop of the BTK-KD is structurally very similar to the A loop in the phosphorylated LCK-KD and the ligated IRK-KD structures, which contain a phosphorylated tyrosine residue (see Fig. 2a; IRK is not shown). Arg-544, Glu-445, and Tyr-551 are well defined in electron densities of the 2Fo-Fc map (Fig. 1B). Based on the distance and geometry in the refined structure, the hydroxyl group of Tyr-551 interacts with Arg-544, Ser-553, and a water molecule via hydrogen bonds, and it possibly interacts electrostatically with the nearby Arg-520 residue. The aromatic ring of Tyr-551 has van der Waals contacts with Val-546 and Phe-574. Tyr-551 of the A loop of the BTK-KD is not phosphorylated, but it interacts with Arg-544 as is the case for the phospho-IRK and phospho-LCK structures. The structural difference is that the unphosphorylated Tyr-551 in the A loop of the BTK-KD interacts with Arg-544 via a hydroxyl group rather than through a phosphate (see Fig. 2a). The crystal structure of the BTK-KD indicates that the A loop is essentially in a closed noninhibitory conformation (Fig. 2, a and b). Hence, only minor structural adjustments would be expected for Tyr-551 and the surrounding residues upon Tyr-551 phosphorylation. Similarly, only minor structural changes in the orientation of the loop residues 556–560 would be expected upon substrate binding. A similar active conformation of the A loop has also been observed in the crystal structures of the recently reported checkpoint kinase (38), and checkpoint-like phosphatase kinase (39, 40), in which the catalytic core is constitutively active. This unique conformation of the A loop of the
unphosphorylated BTK-KD supports a novel molecular mechanism for its initial activation (Fig. 3).

The A loop of kinase domains of PTKs and protein-serine kinases usually serves as a negative regulator of kinase activity by blocking ATP binding and/or substrate peptide binding. Several protein kinases, including IRK, calmodulin-dependent protein kinase II, myosin light chain kinase, and protein kinase C, have a pseudosubstrate sequence within the A loop which sterically blocks the access to the catalytic cleft by a substrate peptide (for review, see Ref. 41). In the apo-IRK structure, the A loop involving the Tyr-551-equivalent tyrosine residue behaves as a substrate peptide mimic and sterically blocks access to the active site (Fig. 2b) (36, 37). Similarly, the A loop in the inactive c-SRC, although different from that in apo-IRK, also hinders peptide binding and blocks access to the active site (Fig. 2a) (42, 43). The autoinhibition mechanism illustrated in the apo-IRK structure was thought to be applicable to BTK as well (12). However, a close examination of the BTK-KD crystal structure reveals that Tyr-551 is nowhere near the active site residue D521 (Fig. 2). Contrary to the expectations, the catalytic cleft of the BTK-KD is not occluded by the A loop or by any other portion of the KD. The conformation of the A loop in the unphosphorylated BTK-KD structure is very similar to the active conformation of phosphorylated LCK-KD and hence does not limit substrate access to the active site.

**Activation of BTK by Transphosphorylation of the Activation Loop Tyrosine Residue Tyr-551**—In a typical protein kinase structure, the beginning and end of the loop from helix αC to β4 and the linker loop between the lobes act as hinge points (44). Correspondingly, two critical structural components that are associated with the active state conformation include the closure of the two lobes and the position of helix αC relative to the N-terminal lobe. As indicated earlier, the two lobes in the BTK structure adopt a closed conformation. A nearly identical conformation was observed for the two BTK-KD molecules that are related by a 2-fold noncrystallographic symmetry and have different molecular packing. This suggests that the apo-BTK kinase domain favors a closed/active conformation. However, the distance of helix αC from the active site is larger in BTK-KD than it is in IRK and c-APK ternary complex structures; whereas the distance between Glu-445 and Lys-430 is 10.2 Å in BTK-KD, the corresponding distances in IRK and c-APK ternary complexes are ~3 Å.

Lys-430 and Glu-445 are two invariant residues in the structural superfamily of protein kinases (32). Even very conservative mutations of these residues in BTK-KD such as K430R and E445D have been associated with severe XLA (45). Mutations of the less conserved Arg-544 residue are also associated with severe XLA (46). The location of the Glu-445-equivalent residue relative to the Lys-430-equivalent residue and the location of the Glu-445-equivalent residue relative to the ATP triphosphate serve as indicators of whether helix αC is in a favorable position for catalysis (43) (Fig. 3). In the active KD structures such as those of the ligated cAPK and IRK or the phosphorylated LCK-KD (Fig. 3A), the Glu-445-equivalent residues are associated with the Lys-430-equivalent residue and the ATP triphosphate either directly by hydrogen bonding or indirectly through a medium of molecules such as water or magnesium. The C-terminal oxygen atoms of Glu-445 in the BTK structure are 10.2 Å away from the Lys-430 terminal atom. Glu-445 is hydrogen-bonded to Arg-544, suggesting that Arg-544 may play a regulatory role in preventing Glu-445 from relocating to the active site and may hinder hydrogen bond formation with Lys-430. This unique structural arrangement strongly suggests a novel molecular mechanism of activation by Tyr-551 phosphorylation (Fig. 3b). Upon phosphorylation of Tyr-551, Arg-544 may be more engaged in the interaction with the phosphate group of P*Tyrr551, and as a consequence, Glu-445 could be released to relocate to the active site and form a hydrogen bond with Lys-430. Helix αC in the BTK-KD structure has to be rotated by 20° from the pivot point at residue 452 (Fig. 3A). Therefore, the phosphorylation of Tyr-551 may activate BTK by disengaging Arg-544 and Glu-445, which consequently enables Glu-445 to take part in the catalytic reaction, as illustrated in Fig. 3b. This unique regulatory inhibition of BTK by Arg-544 differs from the regulatory inhibition of c-SRC by SH3, in which the salt bridge formation between Glu-310 (equivalent to...
Glu-445) and Lys-295 (equivalent to Lys-430) is prevented by the binding of the c-SRC SH3 domain to the proline-rich linker region between SH2 and catalytic domains (47, 48); it also differs from the mechanism of CDK2, in which the relocation of helix C is stabilized with the help of binding with cyclin. In summary, the activation of BTK by Tyr-551 phosphorylation likely involves an exchange of hydrogen-bonded pairs from Glu-445/Arg-544 to Glu-445/Lys-430, which can occur in concert with the phosphorylation of Tyr-551 and subsequent relocalization of helix C.

The BTK Y551F mutant was reported to abrogate BTK autophosphorylation (19). Others observed that Y551F mutation causes a 90% reduction of LYN-mediated enhancement of both BTK tyrosine phosphorylation and kinase activity (12). A phenylalanine residue cannot engage in hydrogen bonding interactions that link Tyr-551 to Arg-544, and thus the conformation of the A loop bearing this mutation may only partially resemble the internally bound inhibitory configuration. On the other hand, the Y551F mutant loses the ability to be phosphorylated, and based on our proposed mechanism it cannot release Glu-445 to the active site. In contrast, the Tyr-551-equivalent residue of c-SRC when mutated to phenylalanine was predicted to unlock the inhibitory “A helix” and actually activate the kinase activity (43; see also 49 and 50). These observations point to an interesting difference between BTK and the kinase proteins of the SRC family. The x-ray crystal structure of the BTK-KD revealed that neither Tyr-551 nor Tyr-545, the other tyrosine residue of the A loop, resides in the catalytic site. This finding strongly supports the notion that BTK is activated by a trans-phosphorylation mechanism (via intermolecular interaction) at Tyr-551 as suggested in previous experiments with LYN (12, 13, 17, 51), rather than a cis-phosphorylation mechanism.

Although a homology model of BTK is useful for providing the structural basis of XLA-causing mutations, a crystal structure provides the most accurate interaction, especially considering that the two-lobe conformation among kinases and the geometry of active site could vary significantly. Most of known XLA-causing missense mutations are listed in Table II, together with our predictions (or justification) of structural consequences of these mutations. However, observations involving R544K/G and K430R mutations are more intriguing. An arginine residue is different from a lysine residue in side chain length as well as hydrogen bonding capability and interactions with phosphate and/or a glutamic acid residue. These differences may adversely affect the binding of ATP to the catalytic domain of BTK R544K/G. The R544G mutant may not be able to stabilize the phosphorylated Tyr-551 and may consequently destabilize the A loop, the po-

**Fig. 3.** Unique activation mechanism proposed based on the crystal structures of BTK-KD and phospho-LCK. Panel A, comparison of the superimposed BTK-KD (white backbones and multiple color side chains in stick model) and the phospho-LCK structures (gold backbones and red stick model side chains). The black lines indicate hydrogen bonds. Panel B, we propose that critical changes may occur when Arg-544 solely interacts with phosphotyrosine 551 upon phosphorylation and Arg-544 releases Glu-445, which subsequently takes part in ATP binding, which may be a critical component of catalysis. N-lobe, N-terminal lobe; C-lobe, C-terminal lobe.
Tables:

| BTK residue mutation | Structural position and function of involved residue | Structural consequences |
|----------------------|---------------------------------------------------|------------------------|
| Residue no. | Amino acid | | |
| 408 | L → P | Maintain the β conformation of the β1 strand | A, indirectly |
| 414 | G → R | Highly conserved residue, backbone forms a hydrogen bond with the β phosphate of ATP | A |
| 418 | Y → H | Maintain the local hydrophobic structural stability near the domain hinge region | A, indirectly |
| 429 | I → N | In a hydrophobic cluster and maintain the natural curvature of the N-terminal lobe β sheet | B |
| 430 | K → E → R | Invariant residue of PTK family, align ATP via interaction with the α phosphate | C |
| 445 | E → D | Invariant residue of PTK family, align ATP via interaction with the β phosphate | C |
| 462 | G → D → V | On the β4 strand and the interface with β sheet and helix C | E |
| 476 | Y → D | Close to adenine base of ATP | D |
| 477 | M → R | Close to adenine base of ATP | D |
| 502 | C → F → W | On helix E, interact with F583 and M587 in a compact hydrophobic cluster | B |
| 504 | D → V | Ion pair with K456, near the hinge region of helix C | D |
| 506 | C → R → Y | On helix E, interact with F583 (helix F) and the carbonyl groups of 644 and 502 | B |
| 508 | A → D | On helix E, near the hinge region of helix C | D |
| 509 | M → I → V | Interact with 1522 and V537 in a compact hydrophobic cluster | B |
| 512 | I → P → Q | Near the side-chains of L457, L452, and F517, M509 in a compact hydrophobic cluster | B |
| 518 | L → R | Close to V546, F574, and Y551 in a hydrophobic cluster | B + E |
| 520 | R → Q | Conserved in all RD kinases, interacts with the nearby activation Y551, stabilizing the phosphotyrosine | B + E |
| 521 | D → G → H → N | Invariant residue of PTK family, activation of the nucleophilic attacking hydroxyl group of the target residue (tyrosine in BTK) by deprotonation | C |
| 525 | R → P → Q | Alignment and stabilization of the γ phosphate during phosphoryl transfer process | A |
| 526 | N → K | Invariant residue of PTK family, coupled with D539 to align the γ phosphate of ATP | A |
| 535 | V → F | Interact with L665 | D |
| 542 | L → P | Near the highly conserved DFG motif which adopts two different conformations during catalytic process | D |
| 544 | R → K → G | Form a hydrogen bond with E445 in a nonproductive conformation | See text |
| 559 | F → S | In the P + 1 substrate binding pocket | F, P + 1 |
| 562 | R → P → W | On a helix turn starting from F560 to S564 in the protein substrate binding site | F |
| 563 | W → L | Hydrophobic interaction with A523 and is part of protein substrate binding site | B |
| 567 | E → K | Ion pair with R641 and thus contributes to the interaction between helix EF and helix H or I | B |
| 569 | L → P | On helix EF | B |
| 578 | S → Y | On helix F, hydrogen bonded with the 575 amide near the activation loop | B |
| 581 | W → R | On helix F, hydrophobic interaction with W634 in a compact hydrophobic cluster | B |
| 582 | A → V | Near the 522 carbonyl group and D521 | A + E |
| 583 | F → S | See 506 | B |
| 587 | M → L | On helix F, see 502 | B |
| 589 | E → D → G | On helix F, see 506 | B |
| 592 | S → P → Y | Forms a hydrogen bond with the 595 amide | B |
| 594 | G → E → R | Near R487, which forms part of the P → 1 binding pocket | F, P + 1 |
| 597 | P → T | Buried near M596 which forms part of the P → 1 binding pocket | F, P + 1 |
| 598 | Y → C | Interact with F601, L616, and T606 in a compact hydrophobic cluster | B |
| 606 | T → P | On the inward face of helix G, see 598 | B |
| 607 | A → D | On helix G, mostly exposed | X |
| 613 | G → D | Near R615 | B |
| 616 | L → F | On helix G, interact with W634, Y598, and W588 in a compact hydrophobic cluster | B |
| 619 | P → A → /S/T | On the inward face of helix G and close to W588, S592, and A622 | B |
| 622 | A → P | Interact with 619, see 619 | B |
| 626 | V → G | Form a hydrophobic cluster on the inward face of helix H and close to Y591 | B |
| 630 | M → K → T | On the inward face of helix H, interacts with W634 and W588 | B |
| 633 | C → Y | On the inward face of helix H in a compact hydrophobic cluster | B |
| 641 | R → C → H | On helix H and pair with E567, see 567 | B |
| 644 | F → L → S | On the inward face of helix I and in a hydrophobic cluster | B |
| 647 | L → R | On the inward face of helix I and close to C633 and F583 | B |
| 652 | L → P | Exposed | X |

sition of which is sensitive for correct alignment of peptide substrate binding. On the other hand, R544K would be expected to unlock Glu-445 and would probably trigger part of the activation process as discussed earlier in this paper. The effect of the R544K mutation on BTK kinase activity is less certain because the Arg-544-equivalent residues vary in different PTKs. The BTK structure suggests that R544K is unlikely to abolish the kinase activity entirely.

Arg-520 is not entirely conserved in the protein kinase superfamily but is present in all “RD” kinases that require activation by phosphorylation (41). Notably, the Arg-520-equivalent residue in IRK was found to be mutated (R1131N) in patients with non-insulin-dependent diabetes mellitus (52). The side chain of Arg-520 is close to Tyr-551 in the BTK-KD crystal structure. A survey of the equivalent residue in the RD kinase structures revealed that the Arg-520-equivalent residue is close to a phosphate or a carboxylate group and apparently plays a role in stabilizing the phosphotyrosine/Ser/Thr. The hot spot mutation, R520Q, certainly changes the interaction pattern with P*Tyr-551, and a glutamine residue is much less likely to be associated with a phosphate group than an arginine residue (53). Thus, the R520Q mutant probably would have a destabilized A loop.

XLA-associated BTK mutations involving the N-terminal
lobe of the kinase domain are less frequent than those involving the C-terminal lobe (Fig. 4 and Table II). Gly-414, Leu-408, Tyr-418, and Ile-429 were identified as “mutation hot spots” in XLA patients. Notably, residue Gly-414, which is highly conserved as a glycine (or less likely, small residues like alanine), is located at the beginning of the β2 strand and is right on top of the triphosphate group of ATP where its backbone forms a hydrogen bond with the oxygen atom of the β phosphate. At this position, a large side chain substitution such as the G414R mutation would dramatically limit the loop flexibility that may be required to accommodate ATP and subsequently release ADP. With the arginine substitution, neither the hydrogen bonding nor the ATP binding conformation would be optimal.

Therefore, the functional implications of this mutation should be interpreted with caution. Our modeling studies indicate that the XLA-associated mutations W563L, P597T, F583S, and Arg-562 are directly or indirectly involved in the peptide substrate binding. Trp-563 is situated between Pro-597 and Ala-523, the latter of which is near the center of the active site. The W563L mutation may alter the conformation of the peptide substrate and has been identified in patients with XLA. Pro-597 is relatively distant from the central region of the active site, but the side chain of Pro-597 is totally buried behind the nearby residues including Met-596, which forms part of the P−1 binding pocket. The nearest atom pairs between Trp-563 and Pro-597 or Ala-523 are 3.7 Å away. The three residues are packed against each other as the core part of the substrate peptide binding site. Clearly, the P597T mutation would impair substrate binding. By comparison, the F583S mutation may change the selection preference of the binding region for the P−1 position. The side chain of Arg-562 forms a network of hydrogen bonds with Asn-603, which is a part of the P+1 binding pocket, and Thr-606, which is connected to the main chain carbonyl group. The R562P mutation can be expected to alter the helical turn because of the rigid proline residue and thereby change the local conformation, including that of the important PTK invariant P-site residue Pro-560. The R1174N mutant (corresponds to a mutation of Arg-562) in IRK has been identified in non-insulin-dependent diabetes mellitus patients.

Remarkable insights were gained by mapping the XLA-associated missense mutations of the C-terminal lobe onto the BTK-KD crystal structure. Most of these mutations on the C-terminal lobe are concentrated on the inward face of helices and loops (see blue residues in Fig. 4). According to our analysis, the majority of the side chains are at least half-buried, suggesting that these mutations may destabilize the C-terminal lobe of the kinase domain by altering its interactions with neighboring residues (see also Table II). The mutations include S578Y, W581R, A582V, F583S, M587L, E589D/G, and S592P on both sides of helix F, which is flanked by other helices. The inward sides of helices E, G, I, and loops appear to be most susceptible to XLA-causing mutations (see Fig. 4). L652P is an example in that its side chain occurs on the protein surface. Therefore, the functional implications of this mutation should be interpreted with caution.

Acknowledgments—We thank Steven Bartell for technical support. X-ray data were collected at the F2 station of the Cornell High Energy Synchrotron Source (CHESS). Assistance during data collection at the CHESS synchrotron was provided by Dr. Igor Kurinov and MacCHESS staff Drs. Marian Szebenyi, Michael Cook, and Daniel Thiel.

REFERENCES

1. Uckun, F. M. (1998) Biochem. Pharmacol. 56, 683–691
2. Rawlings, D. J., and Witte, O. N. (1994) Immunol. Rev. 138, 105–119
3. Kurosaki, T., and Kurosaki, M. (1997) J. Biol. Chem. 272, 15595–15598
4. Vassilen, A., Ozer, Z., Navara, C., Mahajan, S., and Uckun, F. M. (1999) J. Biol. Chem. 274, 1644–1656
5. Mahajan, S., Ghosh, S., Sadbeck, E. A., Zheng, Y., Downs, S., Hupke, M., and Uckun, F. M. (1999) J. Biol. Chem. 274, 9587–9599
6. Tsukada, S., and Witte, O. N. (1994) Adv. Exp. Med. Biol. 365, 233–238
7. Vetrie, D., Vorechovsky, I., Sideras, P., Poland, D., Davies, A., Finster, F., Hammarstrom, L., Kinnunen, C., Levinson, R., Bohr, M., Smith, C. I. E., and Bentley, D. (1993) Nature 363, 226–233
8. Rawlings, D. J., Sufran, D. C., Tsukada, S., Lloyd, G., Gilmartin, J. C., Cohen, L., Mohr, R., Bazan, J. F., Howard, M., Copeland, N. G., Jenkins, N. A., and Witte, O. N. (1995) Science 269, 358–361
9. Vihinen, M., Cooper, M. D., de Saint Basile, G., Fischer, A., Good, R. A., Hendriks, R. W., Kinnunen, C., Kwan, S. P., Litman, G. W., Notarangelo, L. D., and Zhu, Q. (1995) Immunity 3, 301–312
10. Kerner, J. D., Appleby, M. W., Mohr, R. N., Schur, S., Rawlings, D. J., Maliszewski, C. R., Witte, O. N., and Perlmuter, M. R. (1995) Immunity 3, 301–312
11. Hyverna, M., and Saraste, M. (1997) EMBO J. 16, 3396–3404
12. Rawlings, D. J., Scharenberg, A. M., Park, H., Wahl, M. I., Lin, S., Kato, R. M., Fuchik, A. C., Witte, O. N., and Kinet, J. P. (1996) Science 271, 822–825
13. Nisitani, S., Kato, R. M., Rawlings, D. J., Witte, O. N., and Wahl, M. I. (1999) J. Biol. Chem. 274, 400–405
14. Nisitani, S., Satterthwaite, A. B., Akashi, K., Weissman, I. L., Witte, O. N., and Wahl, M. I. (2000) J. Biol. Chem. 275, 2226–2226
Crystall Structure of BTK

15. Oda, A., Ikeda, Y., Ochs, H. D., Druker, B. J., Ozaki, K., Handa, M., Ariga, T., Sakiyama, Y., Witte, O. N., and Wahl, M. I. (2000) *Blood* **95**, 1663–1670
16. Park, H., Wahl, M. I., Afar, D. E., Turek, C. W., Rawlings, D. J., Tam, C., Scharenberg, A. M., Kinet, J. P., and Witte, O. N. (1996) *Immunity* **4**, 515–525
17. Wahl, M. I., Fluckiger, A. C., Kato, R. M., Park, H., Witte, O. N., and Rawlings, D. J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 11526–11533
18. Afar, D. E., Park, H., Howell, B. W., Rawlings, D. J., Cooper, J., and Witte, O. N. (1996) *Mol. Cell. Biol.* **16**, 3465–3471
19. Mahajan, S., Parghioli, J., Burkhardt, A. L., Kott, S. A., Saouaf, S. J., and Bolen, J. B. (1995) *Mol. Cell. Biol.* **15**, 5304–5311
20. Otwinowski, Z., and Minor, W. (1998) *Methods Enzymol.* **276**, 307–325
21. Collaborative Computational Project 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–776
22. Narayana, N., Cox, S., Shaltiel, S., Taylor, S. S., and Xuong, N. (1997) *Biochemistry* **36**, 4438–4448
23. Kleywegt, G. J., and Jones, T. A. (1994) in *First Map to Final Model* (Bailey, S., Hubbard, R., and Waller, D., eds) pp. 59–66, SERC Daresbury Laboratory, Daresbury, UK
24. Kleywegt, G. J., and Jones, T. A. (1994) *Acta Crystallogr. Sect D Biol. Crystallogr.* **50**, 178–185
25. Jones, T. A. (1992) in *Molecular Replacement* (Dodson, E. J., Grover, S., and Wolf, W., eds) pp. 91–105, SERC Daresbury Laboratory, Daresbury, UK
26. Kleywegt, G. J., and Jones, T. A. (1999) *Acta Crystallogr. Sect D Biol. Crystallogr.* **55**, 941–944
27. Brunger, A. T. (1992) *XPLOR*, Version 3.1, Yale University Press, New Haven, CT
28. Sack, J. S. (1988) *J. Mol. Graphics* **6**, 244–245
29. Jones, T. A., Bergdoll, M., and Kjeldgaard, M. (1989) in *Crystallographic Computing and Modeling Methods in Molecular Design* (Bugg, C., and Ealick, S., eds) pp. 129–139, Springer, New York
30. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–291
31. Hanks, S. K., and Hunter, T. (1988) *Science* **241**, 42–52
32. Hanks, S. K., and Quinn, A. M. (1991) *Methods Enzymol.* **200**, 38–62
33. Hanks, S. K., and Hunter, T. (1995) *FASEB J.* **9**, 576–596
34. Yamasaki, F., Funahashi, T., and Hendrickson, W. A. (1996) *Nature* **384**, 484–489
35. Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S., and Sowadski, J. M. (1991) *Science* **253**, 407–414
36. Hubbard, S. R. (1997) *EMBO J.* **16**, 5572–5581
37. Hubbard, S. R., Wei, L., Ellis, L., and Hendrickson, W. A. (1994) *Nature* **372**, 746–754
38. Chen, P., Luo, C., Deng, Y., Ryan, K., Register, J., Margoski, S., Tempczyk, Russell, A., Nguyen, B., Myers, P., Lundgren, K., Kan, C. C., and O’Connor, P. M. (2000) *Cell* **100**, 681–692
39. Skamnaki, V. T., Owen, D. J., Noble, M. E., Lowe, E. D., Lowe, G., Oikonomakos, N. G., and Johnson, L. N. (1999) *Biochemistry* **38**, 14718–14730
40. Lowe, E. D., Noble, M. E., Skamnaki, V. T., Oikonomakos, N. G., Owen, D. J., and Johnson, L. N. (1997) *EMBO J.* **16**, 6646–6658
41. Johnson, L. N., Noble, M. E., and Owen, D. J. (1996) *Cell* **85**, 149–158
42. Schindler, T., Sicheti, F., Fico, A., Ganz, A., Levitaki, A., and Kuriyan, J. (1999) *Mol. Cell* **3**, 639–648
43. Xu, W., Doshi, A., Lei, M., Eck, M. J., and Harrison, S. C. (1999) *Mol. Cell* **3**, 629–638
44. Johnson, L. N., Lowe, E. D., Noble, M. E., and Owen, D. J. (1998) *FEBS Lett.* **430**, 1–11
45. Vihinen, M., Brandau, O., Branden, L. J., Kwan, S. P., Lappalainen, I., Lester, T., Noordzij, J. O., Ochs, H. D., Olilla, J., Pienaar, S. M., Bikkonen, P., Saha, B. K., and Smith, C. I. E. (1998) *Nucleic Acids Res.* **26**, 242–247
46. Orlandi, P., Ritiri, K., Moschese, V., Angelini, F., Arvanitidis, K., Spelatlas, M., Sideras, P., Plebani, A., and Rossi, P. (2000) *Hum. Mutat.* **15**, 117
47. Hubbard, S. R. (1999) *Nat. Struct. Biol.* **6**, 711–714
48. Xu, W., Harrison, S. C., and Eck, M. J. (1997) *Nature* **385**, 595–602
49. Brown, M. T., and Cooper, J. A. (1996) *Biochim. Biophys. Acta* **1287**, 121–149
50. Ferracini, R., and Brugge, J. (1998) *Oncogene Res.* **5**, 205–219
51. Schmitz, R., Baumann, G., and Gram, H. (1996) *J. Mol. Biol.* **260**, 664–677
52. Accili, D., Cama, A., Barbeti, F., Kadowaki, H., Kadowaki, T., and Taylor, S. I. (1999) *J. Endocrinol. Invest.* **18**, 857–864
53. Copley, R. R., and Barton, G. J. (1994) *J. Mol. Biol.* **242**, 321–329
54. Kraulis, P. (1991) *J. Appl. Crystallogr.* **24**, 946–950
55. Merritt, E. A., and Bruennola, S. (1991) *Mol. Cell. Biol.* **11**, 869–873
56. Nicholls, A. (1991) *Proteins Struct. Funct. Genet.* **28**, 296–302
57. Hubbard, S. R. (1997) *Insight Version II*, Molecular Simulations Inc., San Diego, CA
