Structures of human Bruton's tyrosine kinase in active and inactive conformations suggest a mechanism of activation for TEC family kinases

Douglas J. Marcotte,[1] Yu-Ting Liu,[1] Robert M. Arduini,[1] Catherine A. Hession,[1] Konrad Miatkowski,[1] Craig P. Wildes,[1] Patrick F. Cullen,[1] Victor Hong,[1] Brian T. Hopkins,[1] Elisabeth Mertsching,[2] Tracy J. Jenkins,[1] Michael J. Romanowski,[3] Darren P. Baker,[1] and Laura F. Silvian[1]*

[1]Biogen Idec, Inc., Drug Discovery Department, 12 Cambridge Center, Cambridge, Massachusetts 02142

[2]Biogen Idec, Inc., 5200 Research Place, San Diego, California 92122

[3]Sunesis Pharmaceuticals, Inc., 395 Oyster Point Boulevard, South San Francisco, California 94080

Received 12 October 2009; Revised 11 December 2009; Accepted 14 December 2009
DOI: 10.1002/pro.321
Published online 5 January 2010 proteinscience.org

Abstract: Bruton's tyrosine kinase (BTK), a member of the TEC family of kinases, plays a crucial role in B-cell maturation and mast cell activation. Although the structures of the unphosphorylated mouse BTK kinase domain and the unphosphorylated and phosphorylated kinase domains of human ITK are known, understanding the kinase selectivity profiles of BTK inhibitors has been hampered by the lack of availability of a high resolution, ligand-bound BTK structure. Here, we report the crystal structures of the human BTK kinase domain bound to either Dasatinib (BMS-354825) at 1.9 Å resolution or to 4-amino-5-(4-phenoxyphenyl)-7H-pyrrolospyrimidin-7-yl-cyclopentane at 1.6 Å resolution. This data provides information relevant to the development of small molecule inhibitors targeting BTK and the TEC family of nonreceptor tyrosine kinases. Analysis of the structural differences between the TEC and Src families of kinases near the Trp-Glu-Ile motif in the N-terminal region of the kinase domain suggests a mechanism of regulation of the TEC family members.

Keywords: Bruton's tyrosine kinase; BTK; Dasatinib; Celera compound; TEC-family; crystal structure

Introduction

Bruton's tyrosine kinase (BTK) is a member of the TEC kinase family, nonreceptor tyrosine kinases that play important roles in T-cell receptor-, B-cell receptor-, and Fcγ-receptor-mediated signaling. BTK participates in signal transduction from B-cell antigen receptors resulting in phospholipase C-γ2-mediated calcium mobilization[1] which, in turn, affects pre-B-cell functional maturation and expansion. Since BTK is required for B-cell function, it is an important target for the potential treatment of inflammatory diseases that involve B-cell activation. Mutations in the human *BTK* gene are responsible for

Abbreviations: BTK, Bruton's tyrosine kinase; BTK-KD, BTK kinase domain; ITK, interleukin-2 inducible T-cell kinase; PH, pleckstrin homology; SH, Src homology; TEC, tyrosine kinase expressed in hepatocellular carcinoma; XLA, X-linked agammaglobulinemia.

Disclosure: All authors (with the exception of M. J. Romanowski) are employees of Biogen Idec and own company stock.

Michael J. Romanowski's current address is Novartis Institutes for Biomedical Research, Cambridge, MA 02139.

**Correspondence to:* Laura F. Silvian, Biogen Idec, Inc., 12 Cambridge Center, Cambridge, MA 02142.
E-mail: laura.silvian@biogenidec.com

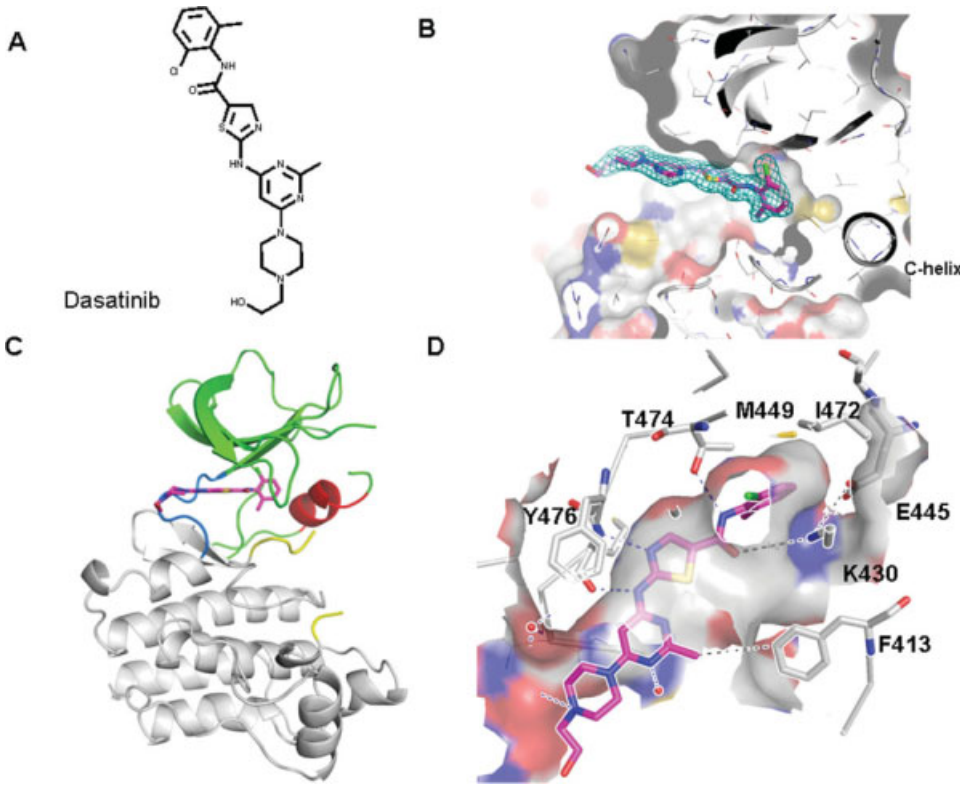

Figure 1. BTK-KD Y551E/Dasatinib crystal structure. A: Chemical structure of Dasatinib. B: Electron density (2Fo-Fc map at 1 sigma) for Dasatinib within a surface representation of the BTK protein in the human BTK-KD-Y551E/Dasatinib complex. C: Overall view of the BTK kinase domain bound to Dasatinib. Inhibitor (magenta), amino-terminal lobe (green), carboxy-terminal lobe (gray), C-helix (red), hinge (blue), and activation loop (yellow). D: Close-up of the active site and residues within 5 Å of the bound Dasatinib. The surface is colored according to the convention of base (blue) and acid (red).

X-linked agammaglobulinemia (XLA), a male immunodeficiency that results in a deficit of mature B cells and serum immunoglobulin.[2,3]

Several compounds that inhibit BTK kinase activity in biochemical assays have been described in the literature and differ in their kinase selectivity profiles. One weak compound, LFM-A13 (α-cyano-β-hydroxy-β-methyl-*N*-(2,5-dibromophenyl)-propenamide) is a BTK inhibitor with an IC_{50} of 2.5 μ*M* in a biochemical assay, but also inhibits PLK3 and JAK2.[4–6] However, it was found to be somewhat specific for BTK, exhibiting 100-fold higher IC_{50} values for related tyrosine kinases such as JAK1, HCK, EGFR, and insulin-receptor kinase (IRK).[7] Another compound, Dasatinib ([*N*-(2-chloro-6-methylphenyl)-2-(6-(4-(2-hydroxyethyl)piperazin-1-yl)-2-methylpyrimidin-4-ylamino)thiazole-5-carboxamide] or BMS-354825) [Fig. 1(A)], originally used to target BCR-Abl, has been shown to bind to BTK with an IC_{50} of 5 n*M*[8] but also binds to other kinases such as SRC family members (HCK, SRC, and CSK), and ephrin receptors, FGR, PDGFRα, and YES.[9] BTK was identified as a target of Dasatinib through pull-down experiments in the CML cell line K562.[8] The reversible Celera compound, 3-cyclopentyl-1-(4-phenoxyphenyl)-1H-pyrazolo pyrimidin-7-amine,[10] was recently described by Pan *et al.*[11] as a potent inhibitor of unphosphorylated BTK (8.2 n*M* IC_{50} in a biochemical assay). However, it also inhibits Lck and Src with IC_{50} values of 2 and 70 n*M*, respectively.[10] It is chemically similar to the commercially available 4-amino-5-(4-phenoxyphenyl)-7H-pyrrolo[2,3-d]pyrimidin-7-yl-cyclopentane [B43; Fig. 2(A)] described as a potent inhibitor of Lck.[12] Finally, an irreversible inhibitor from Pharmacyclics[11] is currently in Phase I for B-cell lymphomas. It is expected to bind irreversibly to Cys481 in the BTK kinase domain active site (0.72 n*M* IC_{50} in a biochemical assay) and its selectivity profile is better than the reversible binder because it exhibits greater selectivity against Lck, which lacks this cysteine (>1000-fold selectivity in a biochemical assay). Future design of potent, specific BTK inhibitors would be facilitated by the structures of these compounds bound to BTK, to discern whether there are regions surrounding the ligand that are unique to this kinase.

BTK is composed of several domains: an N-terminal pleckstrin homology (PH) domain, a proline-rich TEC homology domain, two SRC homology domains (SH3 followed by SH2), and a C-terminal kinase domain (BTK-KD). Mutations in all domains of human BTK have been found to lead to XLA and missense mutations have been found in all domains except for the SH3 domain.[13] Structures have been

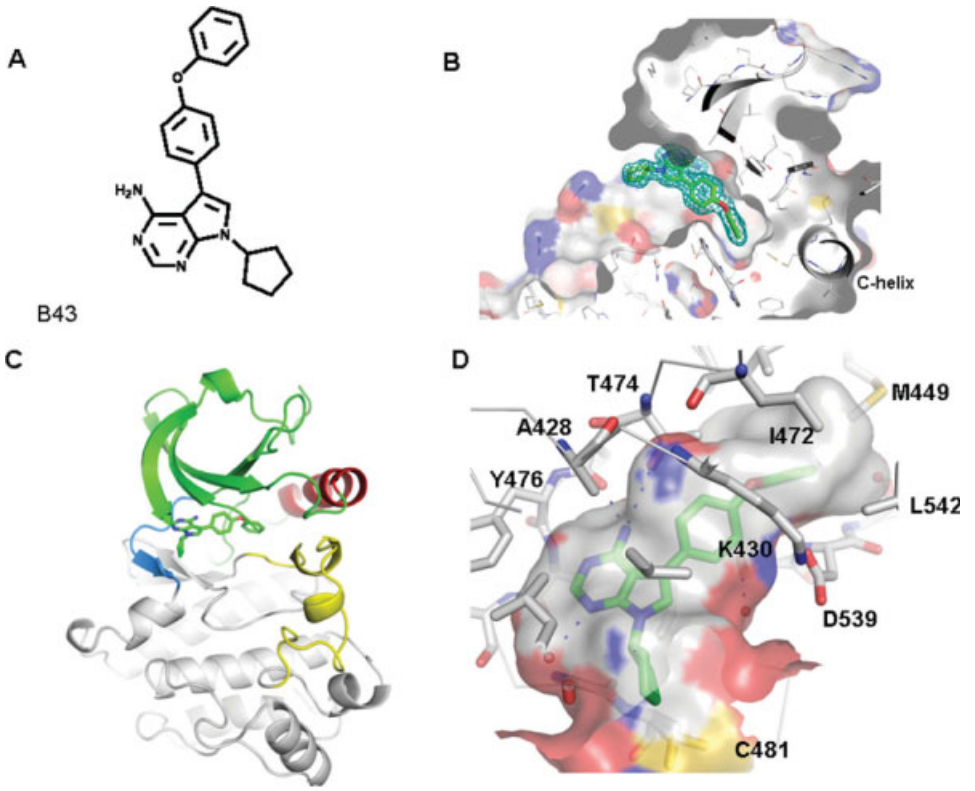

Figure 2. BTK-KD/B43 crystal structure. A: Chemical structure of B43. B: Electron density (2Fo-Fc map at 1 sigma) for B43 within a surface representation of the BTK protein in the human BTK-KD-B43 complex. C: Overall view of the BTK kinase domain bound to B43. Inhibitor (green), amino-terminal lobe (green), carboxy-terminal lobe (gray), C-helix (red), hinge (blue), and activation loop (yellow). D: Close-up of the active site and residues within 5 Å of the bound B43. The surface is colored according to the convention of base (blue) and acid (red). The active site cysteine is shown with a yellow surface.

solved for the kinase domains of apo-murine BTK[7] and human ITK,[14] but a high-resolution structure of a full-length protein with regulatory domains is not available. Low-resolution structures of BTK solved by small angle X-ray scattering have revealed an extended, linear arrangement of the SH3, SH2, and kinase domains, which contrasts with structures of autoinhibited full-length Src and Abl kinases in which a more compact arrangement of the SH2 and SH3 domains allows for the SH2 domain to bind near the C-terminal tail of the kinase domain.[15] Structural studies of the Src family of tyrosine kinases have revealed that these proteins can adapt two conformations: an autoinhibitory state of the protein, referred to as an "assembled regulatory domain" conformation, and an active, more open, structure, where the SH2 domain does not interact with the unphosphorylated C-terminal tail.[16]

Here, we describe the 1.94 Å resolution crystal structure of the human BTK-KD Y551E mutant bound to Dasatinib and a 1.6 Å resolution crystal structure of the unphosphorylated human BTK-KD bound to B43. We observe that the two structures differ in the orientation of the C-helix, similar to conformational changes observed in Src kinase family members that are locked into active or inactive states. Both BTK-KD structures reveal ordered density for the WEX motif at the N-terminus of the kinase domain, where X is a hydrophobic residue. The location of the tryptophan side chain at the base of the C-helix provides an explanation for how the WEX motif acts as an important regulatory element for the TEC family of kinases, similar to its role in regulation of the Src family of kinases, and suggests that the two families have a similar mechanism of regulation.

Results

Protein purification and characterization

BTK-KD and BTK-KD Y551E were purified to ~95% purity using a simple, three-step process utilizing two successive glutathione-Sepharose chromatography steps followed by size exclusion chromatography. Mass spectrometry indicated that the majority of the wild-type BTK-KD and BTK-KD Y551E was intact and unphosphorylated (wild-type calculated mass = 32,641.6 Da, measured mass = 32,643 Da; mutant calculated mass = 32,607.5 Da, measured mass = 32,609 Da), although 2 and 8%, respectively, were missing the first 4 *N*-terminal residues (wild-type calculated mass = 32,317.2 Da, measured mass

= 32,319 Da and mutant calculated mass = 32,283.2 Da, measured mass = 32,285 Da).

Crystal structures

The crystal structures of BTK-KD Y551E/Dasatinib and BTK-KD/B43 complexes were determined using 1.94 Å diffraction data (R_{free} = 25.8%) and 1.6 Å diffraction data (R_{free} = 23.1%), respectively (Table I). The electron density maps clearly revealed the positions of the ligands. Figure 1(B) depicts the electron density for Dasatinib and Figure 2(B) reveals the electron density for B43.

The human BTK-KD-Y551E/Dasatinib complex structure exhibits a "C-helix in," active conformation in which the catalytic lysine forms H-bonds to Glu445 of the C-helix [Fig. 1(C,D)]. The bilobal fold of the BTK KD is similar to that reported for other tyrosine kinase structures: the N-terminal lobe (residues 392–473) consists of five antiparallel β-sheets and two α-helices and contains the glycine-rich loop that covers the active site. The C-terminal lobe (residues 482–659) consists of nine α-helices and two β-strands and the activation loop. Finally, the hinge region (residues 474–481) connects the two lobes of the kinase and contains the active site cysteine (Cys481). The ordered region of the Dasatinib cocrystal structure encompasses residues 392–658 but electron density for parts of the activation loop (residues 542–558) and residues 435–441 is not visible. However, the DFG-portion of the activation loop is visible and is in the so-called "DFG-in" conformation.

The unphosphorylated BTK-KD/B43 cocrystal structure, in contrast, displays a "C-helix-out" inactive conformation [Figs. 2(C) and 3(B)]. Most of the kinase is ordered except the tip of the glycine rich loop (residues 410–414). The activation loop displays a "DFG-in" conformation and is completely ordered. It is composed of two alpha helices, in which Arg544 within the first helix of the activation loop forms a salt bridge to Glu445 of the C-helix.

Table I. *Data Collection and Refinement Statistics*

	BTK-KD/B43 complex	BTK-KD Y551E/Dasatinib complex
Data collection		
Space group	$P2_12_12$	$P2_12_12$
Cell dimensions		
a (Å)	72.5	73.4
b (Å)	104.3	104.6
c (Å)	38.0	38.1
Wavelength (Å)	1.54	0.98
Resolution (Å)	50–1.6	50–1.94
R_{sym} (%)[a]	8.0 (36.3)[a]	8.3 (61.9)
$I/\sigma(I)$[a]	18.5 (4.6)[a]	32.1 (3.3)
Multiplicity	3.8 (4.1)[a]	6.8 (6.6)
Total no. reflections/no. unique reflections	405,732/39,020	490,862/22,496
Completeness (%)[a]	98.1 (98.5)[a]	99.2 (99.2)
Refinement		
Resolution (Å)	24.5–1.6	28.4–1.94
No. reflections	36,293	21,141
R_{work} (R_{free})[b] (%)	19.2 (23.1)	19.9 (25.8)
No. molecules per asymmetric unit	1	1
No. atoms		
Protein	2161	2016
Ligand	28	33
Water	495	164
Average *B*-factors (Å^c)		
Protein atoms	25.8	23.4
Ligand	30.5	41.7
Water	53.5	35.7
R.m.s.d		
Bond lengths (Å)	0.007	0.024
Bond angles (°)	1.14	2.05
Ramachandron plot		
% Allowed	98.5	98.2
% Generously allowed	1.5	1.4
% Not allowed	0	0.4

$R_{sym} = \Sigma|I_{hkl} - <I_{hkl}>|/\Sigma I_{hkl}$.

[a] The value in parentheses is for the highest resolution bin (approximate interval, 0.1 Å).

[b] $R_{work} = \Sigma I_{hkl}|\ |F_o - |F_c|\ |/\Sigma I_{hkl}|F_o|$ for all data except 5%, which is used for the R_{free} calculation.

Dasatinib interactions

Dasatinib makes several hydrogen bonds to the hinge and occupies a hydrophobic pocket behind the Thr474 gatekeeper residue [Fig. 1(B,D)]; similar to the previously reported structures of Dasatinib-bound Abl (PDB ID: 2GQG),[17] Lyn (PDB ID: 2ZVA),[18] and cSrc (PDB ID: 3G5D).[19] Its H-bond interactions to the hinge are described in Figure 1(D), including interactions with the backbone carbonyl and amide of 477, water-mediated interactions with the Tyr476 side chain, and an interaction between the Thr474 side chain and the compound amide nitrogen. The glycine-rich loop of BTK-KD curls toward Dasatinib to form a hydrophobic interaction with Phe413 [Fig. 1(D)]. The ortho-chloro, methyl phenyl substituent of Dasatinib is twisted to enter a hydrophobic pocket, composed of both hydrophobic and hydrophilic residues Met449, Val458, Leu460, Ile472, Lys430, Glu445, and Ser538. Finally, Glu445 of the C-helix forms a salt bridge with the catalytic Lys430; the epsilon amino group of this lysine is poised to make a pi-cation interaction with the Dasatinib ortho-chloro, methyl phenyl ring [Fig. 1(D)].

B43 interactions

The B43 compound makes several hydrogen bonds to the hinge and occupies a hydrophobic pocket behind the Thr474 gatekeeper residue [Fig. 2(D)]. The 4-amino pyrrolopyrimidine of B43 occupies the position of the adenine ring of ATP and the cyclopentyl ring occupies the space generally occupied by the ATP ribose in typical protein kinase structures. The 4-amino pyrrolopyrimidine makes many interactions with the hinge; the exocyclic amine directly interacts

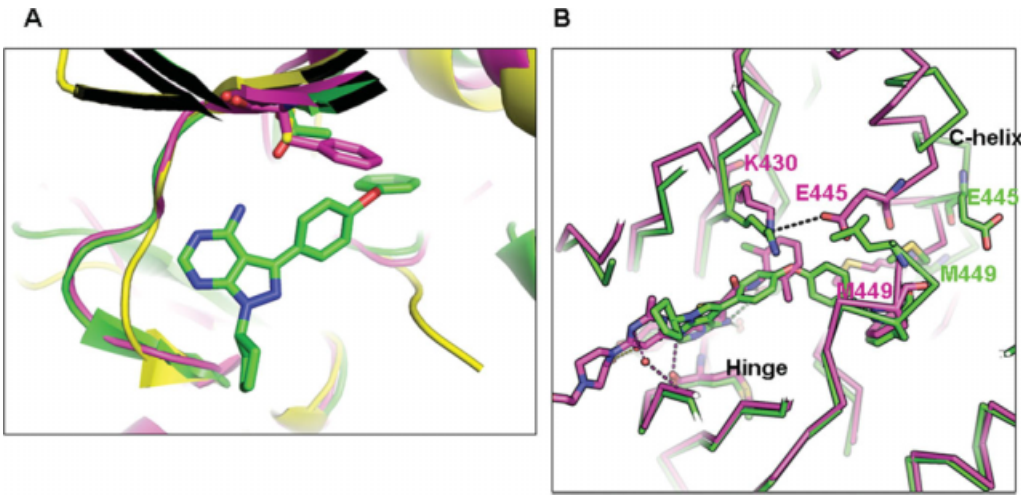

Figure 3. Differences between the "inactive" B43 and "active" Dasatinib-bound conformation of BTK. A: Overlay of BTK-KD bound to B43 (green), ITK bound to staurosporin (magenta, PDB ID: 1SNU) and P38α (yellow, PDB ID: 1P38). This indicates that the larger phenylalanine gatekeeper of ITK compared to the Thr474 gatekeeper of BTK prevents B43 from binding. It also indicates that the shorter hinge of P38α blocks the B43 binding site. B: Overlay of B43 structure (green) and Dasatinib structure (magenta) indicates that they superimpose well near the hinge but differ near the C-helix, resulting in a shift of >10 Å in the conformation of the Glu445 side chain and the residues that surround the hydrophobic back pocket.

with the gatekeeper Thr474 hydroxyl and the backbone carbonyl of 475, the N-3 of the pyrimidine accepts a hydrogen bond from the backbone amine of Met477, and the N-1 of the pyrimidine forms a water-mediated hydrogen bond network to the hydroxyl of Tyr476 and the backbone carbonyl of Ala 478 (Fig. 2). The distal phenyl group of the phenoxyphenyl is twisted 38° out of plane of the phenylether, such that it enters a hydrophobic pocket composed of only hydrophobic residues, Phe442, Met449, Leu460, Ile472, Phe540, and Leu542, and is in proximity to Asp539 of the "DFG" motif. Phe540 of this motif forms a face-to-edge pi-stacking interaction with the phenoxyphenyl group of B43. Asp539 of the DFG-motif forms a salt bridge with the catalytic Lys430 but does not form direct hydrogen bond interactions with the compound.

Discussion

Relevance of structures for drug discovery

The structures of the human BTK-KD Y551E/Dasatinib and BTK-KD/B43 complexes we report here differ from the publicly available structure of apo-murine BTK-KD (PDB ID: 1K2P) and are arguably more relevant for drug discovery for diseases in which inhibition of BTK may be desired. When the apo-mouse BTK structure is superimposed on the human BTK-KD/B43 structure (RMSD of 1.19 Å over 230 aligned α-carbons), the biggest differences are observed in the activation loop and in the glycine-rich loop. The activation loop of the mouse apo-BTK-KD structure adapts an extended configuration with Tyr551 pointed toward solvent. In the mouse apo-BTK structure, the glycine loop also caves into the active site and occludes the ATP binding pocket. Because the mouse and human BTK-KD's are 98.3% identical, and only four amino acids are replaced in the mouse sequence (Q379K, K433R, K625R, and T653S), it is likely that the kinase domain flexibility observed in the apo-murine BTK-KD structure is due to a lack of occupancy of a compound in the active site, rather than due to an intrinsic structural difference between the mouse and human species.

Selectivity of compound inhibition

For both Dasatinib and the reversible Celera compound, the size and hydrogen bonding nature of the gatekeeper residue of a given kinase generally correlates with its degree of biochemical inhibition[9] (Table II). Most of the kinases that are inhibited by 10 μ*M* Dasatinib with a K_d < 1 n*M*, or that are inhibited by 10 μ*M* Celera compound with less than 5% residual activity, have a threonine gatekeeper. A valine residue in this gatekeeper position is tolerated for the Celera compound binding, but is not as well tolerated for Dasatinib binding to the Ret and KDR kinases (Table II). Because the threonine gatekeeper forms H-bond interactions with both compounds, it is possible that the H-bonding binding energy plays a greater role in binding Dasatinib compared to the Celera compound. An alternative explanation for the poor binding of Dasatinib to valine gatekeeper containing kinases KDR and Ret is that there are differences in side chains within 5 Å of the compound (Table III). In particular, one residue in the back pocket that forms close hydrophobic interactions with Dasatinib in BTK is Met449, which is replaced by a leucine in KDR and Ret. Because the back pocket in the Dasatinib cocrystal structure

Table II. *Gatekeeper Residue in Various Kinases, % Activity for Celera Compound and K_d Values for Dasatinib*

Kinase	Gatekeeper	% Activity in the presence of 10 μ*M* Celera[a]	Dasatinib K_d (n*M*)[b]
Ack1	Ser	−3	6
Abl	Thr	−3	0.5
Bmx	Thr	−1	0.0
BTK	Thr	1	1.0
CSK	Thr	0	0.2
EGFR	Thr	20	100
FGR	Thr	−3	0.5
FYN	Thr	3	0.7
Hck	Thr	−5	0.3
Kit	Thr	4	0.6
Lck	Thr	−13	0.2
Lyn	Thr	−4	0.6
P38α	Thr	57	30
Src	Thr	−1	0.2
Yes	Thr	1	0.3
SIK	Thr	17	ND
RIPK2	Thr	4	30
cRAF	Thr	11	600
EphB1	Thr	65	0.4
NEK11	Thr	64	7000
RET	Val	0	200
KDR/VEGFR2	Val	6	3000
RSK1	Leu	80	ND
RSK2	Leu	83	ND
AurA	Leu	101	2000
IGF1R	Met	87	ND
JAK2	Met	88	ND
JNK3	Met	76	ND
Syk	Met	50	3000
Itk	Phe	18	ND

ND, not determined.
[a] Upstate kinase profiler.
[b] Taken from Ref. 9.

is composed of mixed hydrophobic and hydrophilic residues, Dasatinib may have a greater reliance on Met449 compared to B43, whose back pocket is completely surrounded by hydrophobic residues. Either explanation (i.e., different energy considerations for H-bonding to a threonine gatekeeper for different compounds or differences in the makeup of side chains that line Dasatinib's unique hydrophobic pocket), could explain why Dasatinib does not bind as well to Ret and KDR.

The exception to the rule of requiring a small gatekeeper for compound binding is p38α, EGFR, and NIMA (never in mitosis gene A)-related kinase 11 (NEK11) kinases, which have threonine gatekeepers, but are only moderately inhibited by both small molecules. P38α kinase has a shorter hinge, and thus its reduced affinity can be ascribed to a smaller binding site [Fig. 3(A)]. Similarly, there are differences in the other residues within 5 Å of the two small molecules (Table III), which could account for the differences in affinity for NEK11 and EGFR. Both differ in the residues that make up the hydrophobic pocket behind the threonine gatekeeper, with EGFR substituting a cysteine for Val458 and NEK11 substituting a larger phenylalanine for Leu460 and a leucine for Met449. Furthermore, NEK11 contains a glycine substitution for Ser538, a residue which is likely to enhance the flexibility of the DFG motif which follows (Table III). In contrast, Abl, Lyn, and Src, kinases which are inhibited well by both small molecules, show no significant variation in the residues that make up the hydrophobic pocket. In these kinases Ser538 is replaced with alanine, a residue with similar conformational flexibility, and is not likely to effect the flexibility of the activation loop (Tables II and III). As the size of the gatekeeper residue increases from threonine to phenylalanine [as in ITK, Fig. 3(A)], methionine (as in Igf1R, JAK1, JNK3, and SYK) (Table II), or leucine (as in Aurora A) (Table II) neither inhibitor reduces kinase activity to the full extent because both binding modes require insertion of the compound past the bulky gatekeeper into the hydrophobic back pocket.

Table III. *Residues Within 5 Å of Compounds in the BTK-KD/B43 and BTK-KD Y551E/Dasatinib Structures*

	NTD		GRL				UPH					LWH			HYP			AVL			
	428[b]	430[c]	408[b]	409[d]	413[d]	416[b]	472[d]	474[a,c]	475[b]	476[b]	477[b]	479[d]	480[b]	481[c]	458[c]	460[b]	449[b]	538[b]	539[b]	540[b]	542[b]
BTK	A	K	L	G	F	V	I	T	E	Y	M	N	G	C	V	L	M	S	D	F	L
NEK11	V[e]	K	L	G	F	V	I	T	E	Y	C[e]	G[e]	R[e]	D[e]	V	F[e]	L[e]	G[e]	D	F	V[e]
EGRF	A	K	L	G	F	V	L[e]	T	Q[e]	L[e]	M	F[e]	G	C	C[e]	L	M	T[e]	D	F	L
ABL	A	K	L	G	Y[e]	V	I	T	E	F	M	Y[e]	G	N[e]	V	L	M	A[e]	D	F	L
LYN	A	K	L	G	F	V	I	T	E	Y	M	K[e]	G	S[e]	V	L	M	A[e]	D	F	L
SRC	A	K	L	G	F	V	I	T	E	Y	M	K[e]	G	S[e]	V	L	M	A[e]	D	F	L
RET	A	K	L	G	F	V	L[e]	V[e]	E	Y	A[e]	Y[e]	G	S[e]	I[e]	L	L[e]	S	D	F	L
KDR	A	K	L	G	F	V	V[e]	V[e]	E	F[e]	C[e]	F[e]	G	N[e]	V	L	L[e]	S	D	F	L

NTD, N-terminal domain; GRL, glycine rich loop; UPH, upper hinge; LWH, lower hinge; HYP, hydrophobic pocket; AVL, activation loop.
[a] Gatekeeper residue.
[b] Residue interacts with both B43 and Dasatinib in BTK.
[c] Residue only interacts with Dasatinib in BTK.
[d] Residue only interacts with B43 compound in BTK.
[e] Differences in sequence from BTK sequence.

Overall conformational changes

In the BTK-KD Y551E/Dasatinib complex structure, the Gly-rich loop of BTK partially shields Dasatinib from solvent, as found in the Abl kinase structure (data not shown). Because of the curled-in glycine-rich loop in the BTK structure, a hydrophobic interaction is formed between Phe413 and the exocyclic methyl of the Dasatinib pyridine ring [Fig. 1(D)], and between Gln412 and Asp539 of the "DFG" motif. In every other respect, however, the overall conformation of the BTK-KD/Dasatinib structure is similar to the known Dasatinib-bound conformations of Lyn,[18] cSrc,[19] and Abl,[17] including the compound's H-bond interactions with the hinge and the position of its C-helix.

The nonvisible residues differ between the two BTK structures, which can be correlated with the different interactions made with different compounds or by the mutation in the activation loop. The tip of the glycine-rich loop is disordered in the B43 structure but is ordered in the Dasatinib structure, while the activation loop is disordered in the Dasatinib structure but is well ordered in the B43 structure. The curled glycine-rich loop forms van der Waals contacts with Dasatinib, between a methyl substituent and the Phe413 side chain, whereas no similar interactions take place between B43 and the BTK-KD protein residues and the glycine-rich loop is disordered. One might expect that the Y551E mutation in the activation loop of the BTK-KD Y551E/Dasatinib structure is responsible for the activation loop disorder; the mutated residue is electrostatically incompatible with the conformation of the activation loop seen in the BTK-KD/B43 structure. In the BTK-KD/B43 structure, the Y551 residue is in close proximity to an Asp521 side chain; this is likely to be electrostatically repelled by mutation of the tyrosine to a glutamate. While it is often hard to pinpoint why flexible regions of crystal structures are disordered, it appears that formation of critical molecular interactions produces ordered electron density for the more flexible regions of BTK.

Comparison of the structures of the human BTK-KD-Y551E/Dasatinib complex and the BTK-KD/B43 complex reveals a change of conformation from catalytically "active" to "inactive." The Dasatinib complex is more similar to the ATP-bound conformation of most kinases, in which a conserved glutamate from the C-helix forms a salt bridge to the catalytic lysine. In fact, no crystals could be formed with the unphosphorylated, wild-type BTK kinase construct, prompting us to make the Y551E mutant as a mimic of the phosphorylated wild-type protein. In contrast, the BTK-KD/B43 complex shows an outward shift of the C-helix [Fig. 3(B)] relative to its position in the Dasatinib structure, the conserved salt bridge from the glutamate to the catalytic lysine breaks, and a large hydrophobic pocket opens behind the gatekeeper residue. The ability of different kinases to adapt a C-helix out conformation might enable the design of specific inhibitors that targets this larger hydrophobic pocket. Furthermore, Cys481 in the active site of BTK-KD could also be exploited to gain kinase selectivity in which a small molecule may be irreversibly bound to this cysteine via a covalent bond.[11]

To determine the overall similarity of the BTK-KD/B43 structure to other kinases, the B43 complex structure was submitted to the Dali-lite server for structure alignment and scoring[20] (http://ekhidna.biocenter.helsinki.fi/dali_server/). The top hits, inactive Hck (PDB ID: 1QCF), inactive SRC (PDB ID: 2SRC), inactive ABL (PDB ID: 2G1T), ITK (PDB ID: 1SM2), and mouse BTK (PDB ID: 1K2P), could be aligned with the human BTK over more than 260 α-carbons and with an rmsd of 2.0 Å or better. The highest scoring hits, excluding the TEC family of kinases, were all inactive conformations of tyrosine kinases from the Src and Abl families, consistent with their overall sequence similarities to human BTK. The conformation of the activation loop and C-helix in the human BTK-KD/B43 structure is very similar to the inactive Src structure (PDB ID: 2SRC) with an rmsd 1.64 Å over 257 α-carbons; in Src the activation loop forms two alpha helices and occludes access of the substrate peptide. The overall conformation of the BTK-KD Y551E/Dasatinib structure is similar to the active c-Src (CSK) structure (PDB ID: 1Y57) where the activation loop is swung out and the C-helix moves toward the active site (Fig. 4).

Comparison of Src and TEC family kinases

The phosphorylation-triggered regulation of BTK and Src differ. Unlike the Src family, the TEC family of nonreceptor tyrosine kinases lacks a conserved tyrosine in the C-terminus that could be phosphorylated to then bind to the SH2 domain.[16] BTK is regulated by the phosphorylation of two tyrosine residues, Tyr223 in the SH3 domain and Tyr551 in the activation loop of the kinase domain, both of which participate in kinase activation.[21] In a recent study of BTK autophosphorylation, the Y551F mutant was shown to have a 5 to 10-fold lower enzymatic activity than the wild-type protein, indicating that this tyrosine plays an important role in BTK activation.[22] Furthermore, mutation of a conserved tryptophan in the N-terminal W-E-X motif, in which X is a hydrophobic residue, also appears to effect the activities of the two kinase families differently. In Src, mutation of the Trp to Ala increases kinase activity while in BTK, mutation of the Trp to Ala reduces kinase activity.[23]

The human BTK structures described here include ordered density for the WEI motif (residues 392–395), an area which was disordered in the

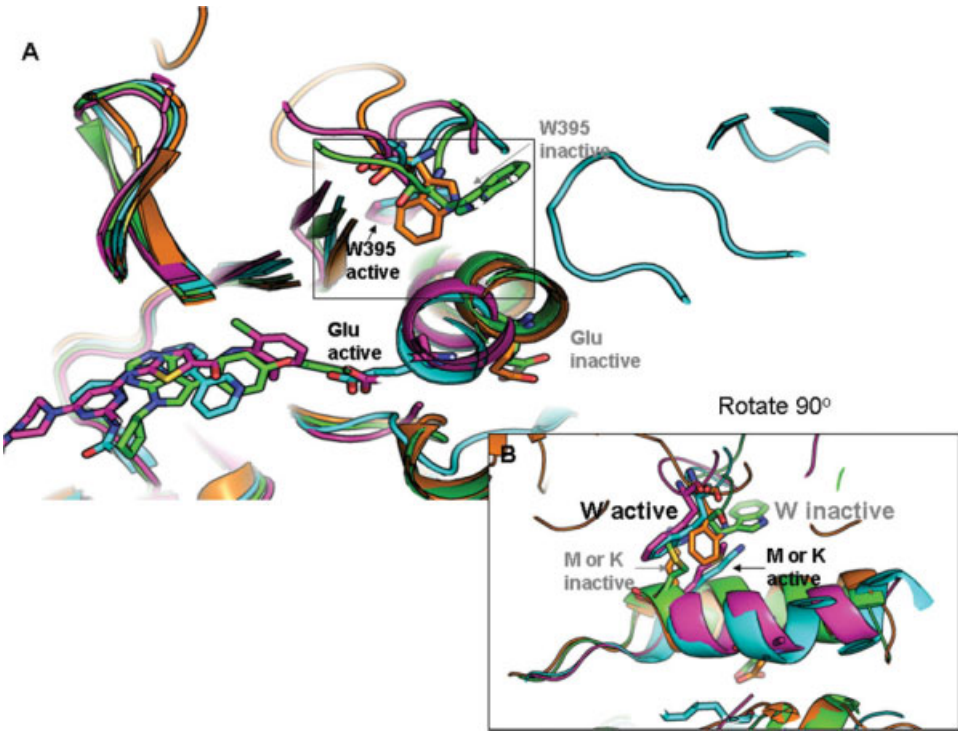

Figure 4. Mechanism of kinase activation. A: Structural superposition of human BTK-KD bound to B43 (green), BTK-KD Y551E bound to Dasatinib (magenta), inactive SRC (orange, PDB ID: 2SRC), and active C-SRC (cyan, PDB ID: 1Y57). Close-up of conformational changes in C-helix and WEI motif behind the C-helix. B: Close-up of the WEI motif rotated 90° relative to (A) demonstrating that "active" BTK and c-SRC structures have a superimposable Trp side chain and that its conformation is incompatible with the "inactive" BTK and SRC conformation due to steric clashes with a methionine or lysine side chain on the C-helix.

available murine BTK structure (PDB ID: 1K2P) and the human ITK structure (PDB ID: 1SM2). This enables a structural comparison of the TEC family and the Src family kinases in this conserved region. The Trp side chain shifts from being solvent exposed in the "inactive" BTK-KD/B43 complex structure to being wedged into a pocket behind the inward C-helix in the "active" BTK-KD Y551E/Dasatinib complex structure [Fig. 4(A,B)]. A structural superposition of the two BTK structures with the inactive SRC (PDB ID: 2SRC)[24] and an active CSK structure (PDB ID: 1Y57)[16] show that the side chain of Trp395 superimposes in the active structures of both kinase families. In the "inactive" conformations, a lysine or methionine side chain from the rotated C-helix sterically occludes the tryptophan side chain, and the Trp side chains are not superimposable [Fig. 4(B)]. The similar structural shift observed suggests that the Src and Tec families utilize a similar means of activating the conformation of the kinase domain. However, the inactive conformations of the Trp side chain differ and the differential energy required in switching from inactive to active conformations may differ; this could explain why mutation of this residue to alanine produces different regulatory effects on the two proteins.

This study reports the structures of human BTK-KD with Dasatinib or 4-amino-5-(4-phenoxyphenyl)-7H-pyrrolo[2,3-d]pyrimidin-7-yl-cyclopentane bound at the active site. Analysis of the binding modes reveals that the size and H-bonding potential of the gatekeeper residue in different kinases and the nature of the residues that make up the hydrophobic pocket behind the gatekeeper residue are critical in dictating if either compound binds. We propose that the presence of a distinct hydrophobic pocket in the B43 structure suggests that one could design compounds with limited selectivity for kinases that can adapt the "C-helix-out" conformation. Movements in the C-helix, presumably induced by phosphorylation of the activation loop residue Tyr551, can shift BTK into an active conformation and could generate a second metal binding site containing Glu445 that Lin *et al*. have shown is important for optimal catalytic activity of the Tyr551-phosphorylated BTK.[25] Finally, we show that, like the Src-family of kinases, BTK can adapt a similar conformational rotamer of Trp395 in its "active" conformation, which is correlated with a similar movement of the C-helix. With a growing recognition that BTK plays a key role in many B cell lymphomas and autoimmune diseases, these structures will assist with selective drug design.

Materials and Methods

Reagents

Glutathione Sepharose 4 Fast Flow resin was from GE Healthcare (Piscataway, NJ); benzamidine, bestatin, E-64, leupeptin, aprotinin, pepstatin, PMSF, DTT, and glutathione were from Sigma Chemical Co. (St. Louis, MO); GST-His-tagged Turbo 3C protease was from Accelagen (San Diego, CA); 4-amino-5-(4-phenoxyphenyl)-7H-pyrrolo[2,3-d]pyrimidin-7-yl-cyclopentane was from Calbiochem (San Diego, CA).

Cloning, expression, and purification

The kinase domain of human BTK (residues 382–659, genebank accession number: AAB65205, or residues 382–659 a Y551E mutation) was inserted into the baculovirus expression vector pDest20 (Invitrogen), which was modified to include a GST tag followed by a 3C protease cleavage site preceding the N-terminus of BTK. Baculovirus was generated using standard procedures by Blue Sky Biotech, Inc. (Worcester, MA). Protein was expressed in baculovirus-infected sf9 insect cells grown in SF-900 II serum-free medium (Invitrogen) in a 20 L fermentor controlled for agitation, aeration, and temperature. Cells were harvested by centrifugation 48 h postinfection, and stored at −70°C. All protein purification steps were carried out at 4°C unless stated otherwise. Frozen cells were thawed and suspended (4 mL/g wet weight of cells) in 50 m*M* Tris-HCl pH 7.5, 400 m*M* NaCl, 3 m*M* DTT, containing the following protease inhibitors: 1 m*M* benzamidine, 10 μ*M* bestatin, 10 μ*M* E-64, 10 μg/mL leupeptin, 10 μg/mL aprotinin, 5 μ*M* pepstatin, and 2 m*M* PMSF. Cells were disrupted by passing twice through a M-110L microfluidizer (Microfluidics, Newton, MA) at 11,000 psi, then the cell lysate centrifuged at 30,000*g* for 1 h to remove cell debris. The BTK kinase domain was purified on glutathione Sepharose 4 Fast Flow resin, eluting with 200 m*M* Tris-HCl pH 7.5, 400 m*M* NaCl, 10 m*M* glutathione, 3 m*M* DTT, containing the same protease inhibitors as mentioned earlier. BTK-containing fractions were pooled, and the protein incubated overnight with GST-tagged Turbo 3C protease (12 units/mg of BTK) while concomitantly dialyzing against 50 m*M* Tris-HCl pH 7.5, 400 m*M* NaCl, 3 m*M* DTT. The liberated BTK was purified from uncleaved protein, the GST tag, and the GST-tagged protease on glutathione Sepharose 4 Fast Flow resin. The flow through fractions, containing the tag-free BTK, were pooled, concentrated using a Pellicon-XL concentrator fitted with a 10,000 molecular weight cut off membrane (Millipore, Bedford, MA), then purified further by size exclusion chromatography at ambient temperature using a Superdex 200 50/100 column (GE Healthcare) with 20 m*M* Tris-HCl pH 8.0, 50 m*M* NaCl, 3 m*M* DTT as the mobile phase. BTK-containing fractions were pooled, the protein concentrated to ∼ 12 mg/mL, flash-frozen with liquid nitrogen, and stored at −70°C.

Protein concentration determination

The concentration of the wild-type and Y551E mutant of BTK-KD was determined by absorbance measurements at 280 nm using the predicted extinction coefficients of 55,350 and 53,860 L mol^{-1} cm^{-1}, respectively, based on the tryptophan and tyrosine content.[26]

Mass spectrometry

Approximately 150 pmol of BTK kinase domain was incubated in PBS pH 7.6, 4*M* urea, 40 m*M* DTT, and the reduced protein analyzed on an LC-MS system composed of an HPLC solvent delivery system (2695 Alliance Separations Module), a 2487 dual wavelength UV detector, and an LCT mass spectrometer (Waters Corp., Milford, MA). The sample was desalted on-line on a Mass-PREP cartridge (2.1-mm internal diameter × 10 mm). Molecular masses were obtained by deconvolution of raw mass spectral data using the MaxEnt 1 program embedded within the MaxLynx 4.0 software (Waters Corp.).

Kinase inhibition assays

Upstate Kinase Profiler data measuring the inhibition of the Celera compound against a kinase panel of 265 kinases at 10 μ*M* compound concentration of the Celera concentration and ATP concentration at K_m values were derived as per the provider (Upstate, Charlottesville, VA). Data are presented in Table II as the percent of kinase activity remaining.

Crystallization of BTK-KD Y551E/Dasatinib

Crystals were grown in a similar manner as the BTK-KD/B43 complex but cocrystals only appeared with the BTK-KD Y551E mutant and could not be grown with the wild-type BTK-KD construct. BTK-KD Y551E was incubated with Dasatinib at a ratio of 1 m*M* inhibitor to 150 μ*M* (5 mg/mL) BTK-KD Y551E in the presence of 10% DMSO. The complex was mixed 1:1 (v/v) with a well solution of 0.1*M* Bis TRIS pH 6.5, 0.2*M* ammonium acetate and 20% PEG5000 MME and crystals formed by multiple rounds of seeding. Rectangular, block-shaped, single crystals of the BTK-KD Y551E/Dasatinib complex were cryoprotected by transferring to 0.1*M* Bis-TRIS pH 6.5, 0.2*M* ammonium acetate, 20% PEG5000 MME, 25% PEG200, and flash-frozen with liquid nitrogen.

Crystallization of BTK-KD/B43

Crystals were grown at 4°C using the sitting-drop, vapor-diffusion technique. The BTK-KD was mixed with B43 at a ratio of 1 m*M* inhibitor to 180 μ*M* (5.9 mg/mL) BTK in the presence of 10% DMSO. The complex was mixed 1:1 (v/v) with well solution (0.1*M*

MES pH 6.5, 0.2*M* ammonium sulfate, 30% (w/v) Peg5000 MME). Rectangular, block-shaped, single crystals of the BTK-KD/B43 complex were cryoprotected by transferring to 85 m*M* MES pH 6.5, 170 m*M* ammonium sulfate, 25.5% (w/v) Peg MME5000, 15% ethylene glycol, and flash-frozen with liquid nitrogen.

Data collection and structural refinement

X-ray diffraction data was collected using a Rigaku (The Woodlands, TX) FRE for the B43 complex and at LRLcat at the Argonne Photon Source for the Dasatinib complex, and was processed with HKL-2000.[27] Both crystals belong to space group $P2_12_12$ with one molecule per asymmetric unit (Table II). The B43 structure was solved by molecular replacement with MOLREP[28] using the publicly available mouse BTK-KD structure (PDB ID: 1K2P) as a search model, in which the glycine-rich loop and activation loop were removed. The best solution had an R_{free} of 53.0% and a correlation coefficient of 0.332. This was then subjected to rigid body refinement in which the amino terminal lobe of the kinase (residues 391–473) was refined separately from the carboxy-terminal lobe (residues 474–659) in REFMAC5,[29] resulting in an R_{free} of 47.7% to 3.5 Å resolution. Subsequent model building in COOT 0.4,[30] and restrained refinement in REFMAC5 with Babinet scaling and fixed TLS parameters led to a model with R_{free} of 23.1% and R-factor of 19.2% to 1.6 Å resolution with good geometry. In this structure, residues that were disordered included 382–391 at the amino terminus and 410–414 in the glycine-rich loop. For the BTK-KD Y551E/Dasatinib structure, molecular replacement with the B43 structure in MOLREP followed by model building and subsequent refinement led to the final structure with R_{free} of 25.8% and R-factor 19.9% to 1.94 Å resolution. The residues missing in the Dasatinib cocrystal structure include residues 382–391, 435–441, and 542–558. A summary of the data collection and refinement statistics is described in Table I and electron density for Dasatinib and B43 is shown in Figures 1(B) and 2(B), respectively.

Accession numbers

The atomic coordinates and structure factors for BTK-KD Y551E/Dasatinib (Research Collaboratory for Structural Bioinformatics Protein Databank = PDB ID: 3K54) and for the BTK-KD/B43 (Research Collaboratory for Structural Bioinformatics Protein Databank = PDB ID: 3GEN) have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Databank, Rutgers University, New Brunswick, NJ (http://www.rcsb.org).

Acknowledgment

The authors thank Arthi Kumaravel for help in seeding crystals for data collection.

References

1. Kurosaki T, Kurosaki M (1997) Transphosphorylation of Bruton's tyrosine kinase on tyrosine 551 is critical for B cell antigen receptor function. J Biol Chem 272: 15595–15598.
2. Tsukada S, Witte ON (1994) X-linked agammaglobulinemia and Bruton's tyrosine kinase. Adv Exp Med Biol 365:233–238.
3. Vetrie D, Vorechovsky I, Sideras P, Holland J, Davies A, Flinter F, Hammarström L, Kinnon C, Levinsky R, Bobrow M, Smith CIE, Bentley DR (1993) The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. Nature 361:226–233.
4. Mahajan S, Ghosh S, Sudbeck EA, Zheng Y, Downs S, Hupke M, Uckun FM (1999) Rational design and synthesis of a novel anti-leukemic agent targeting Bruton's tyrosine kinase (BTK), LFM-A13 [alpha-cyano-beta-hydroxy-beta-methyl-*N*-(2,5-dibromophenyl)propenamide]. J Biol Chem 274:9587–9599.
5. van den Akker E, van Dijk TB, Schmidt U, Felida L, Beug H, Lowenberg B, von Lindern M (2004) The Btk inhibitor LFM-A13 is a potent inhibitor of Jak2 kinase activity. Biol Chem 385:409–413.
6. Uckun FM, Dibirdik I, Qazi S, Vassilev A, Ma H, Mao C, Benyumov A, Emami KH (2007) Anti-breast cancer activity of LFM-A13, a potent inhibitor of Polo-like kinase (PLK). Bioorg Med Chem 15:800–814.
7. 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.
8. Hantschel O, Rix U, Schmidt U, Burckstummer T, Kneidinger M, Schutze G, Colinge J, Bennett KL, Ellmeier W, Valent P, Superti-Furga G (2007) The Btk tyrosine kinase is a major target of the Bcr-Abl inhibitor dasatinib. Proc Natl Acad Sci USA 104:13283–13288.
9. Carter TA, Wodicka LM, Shah NP, Velasco AM, Fabian MA, Treiber DK, Milanov ZV, Atteridge CE, Biggs WH, III, Edeen PT, Floyd M, Ford JM, Grotzfeld RM, Herrgard S, Insko DE, Mehta SA, Patel HK, Pao W, Sawyers CL, Varmus H, Zarrinkar PP, Lockhart DJ (2005) Inhibition of drug-resistant mutants of ABL, KIT, and EGF receptor kinases. Proc Natl Acad Sci USA 102: 11011–11016.
10. Arnold LD, Calderwood DJ, Dixon RW, Johnston DN, Kamens JS, Munschauer R, Rafferty P, Ratnofsky SE (2000) Pyrrolo[2,3-d]pyrimidines containing an extended 5-substituent as potent and selective inhibitors of Lck. Bioorg Med Chem Lett 10:2167–2170.
11. Pan Z, Scheerens H, Li SJ, Schultz BE, Sprengeler PA, Burrill LC, Mendonca RV, Sweeney MD, Scott KC, Grothaus PG, Jeffery DA, Spoerke JM, Honigberg LA, Young PR, Dalrymple SA, Palmer JT (2007) Discovery of selective irreversible inhibitors for Bruton's tyrosine kinase. ChemMedChem 2:58–61.
12. Calderwood DJ, Johnston DN, Munschauer R, Rafferty P (2002) Pyrrolo[2,3-d]pyrimidines containing diverse N-7 substituents as potent inhibitors of Lck. Bioorg Med Chem Lett 12:1683–1686.
13. Valiaho J, Smith CI, Vihinen M (2006) BTKbase: the mutation database for X-linked agammaglobulinemia. Hum Mutat 27:1209–1217.
14. Brown K, Long JM, Vial SC, Dedi N, Dunster NJ, Renwick SB, Tanner AJ, Frantz JD, Fleming MA, Cheetham GM (2004) Crystal structures of interleukin-2 tyrosine kinase and their implications for the design of selective inhibitors. J Biol Chem 279:18727–18732.

15. Marquez JA, Smith CI, Petoukhov MV, Lo Surdo P, Mattsson PT, Knekt M, Westlund A, Scheffzek K, Saraste M, Svergun DI (2003) Conformation of full-length Bruton tyrosine kinase (Btk) from synchrotron X-ray solution scattering. EMBO J 22:4616–4624.
16. Cowan-Jacob SW, Fendrich G, Manley PW, Jahnke W, Fabbro D, Liebetanz J, Meyer T (2005) The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. Structure 13:861–871.
17. Tokarski JS, Newitt JA, Chang CY, Cheng JD, Wittekind M, Kiefer SE, Kish K, Lee FY, Borzillerri R, Lombardo LJ, Xie D, Zhang Y, Klei HE (2006) The structure of Dasatinib (BMS-354825) bound to activated ABL kinase domain elucidates its inhibitory activity against imatinib-resistant ABL mutants. Cancer Res 66:5790–5797.
18. Williams NK, Lucet IS, Klinken SP, Ingley E, Rossjohn J (2009) Crystal Structures of the Lyn protein tyrosine kinase domain in its apo- and inhibitor-bound state. J Biol Chem 284:284–291.
19. Getlik M, Grutter C, Simard JR, Kluter S, Rabiller M, Rode HB, Robubi A, Rauh D (2009) Hybrid compound design to overcome the gatekeeper T338M mutation in cSrc. J Med Chem 52:3915–3926.
20. Holm L, Kaariainen S, Rosenstrom P, Schenkel A (2008) Searching protein structure databases with DaliLite v.3. Bioinformatics 24:2780–2781.
21. Rawlings DJ, Scharenberg AM, Park H, Wahl MI, Lin S, Kato RM, Fluckiger AC, Witte ON, Kinet JP (1996) Activation of BTK by a phosphorylation mechanism initiated by SRC family kinases. Science 271: 822–825.
22. Dinh M, Grunberger D, Ho H, Tsing SY, Shaw D, Lee S, Barnett J, Hill RJ, Swinney DC, Bradshaw JM (2007) Activation mechanism and steady state kinetics of Bruton's tyrosine kinase. J Biol Chem 282: 8768–8776.
23. Guo S, Wahl MI, Witte ON (2006) Mutational analysis of the SH2-kinase linker region of Bruton's tyrosine kinase defines alternative modes of regulation for cytoplasmic tyrosine kinase families. Int Immunol 18: 79–87.
24. Xu W, Doshi A, Lei M, Eck MJ, Harrison SC (1999) Crystal structures of c-Src reveal features of its autoinhibitory mechanism. Mol Cell 3:629–638.
25. Lin L, Czerwinski R, Kelleher K, Siegel MM, Wu P, Kriz R, Aulabaugh A, Stahl M (in press) Activation loop phosphorylation modulates Bruton's tyrosine kinase (Btk) kinase domain activity. Biochemistry.
26. Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 182:319–326.
27. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. Macromolecular Crystallography (Part A) 276:307–326.
28. Vagin A, Teplyakov A (2000) An approach to multi-copy search in molecular replacement. Acta Crystallogr D Biol Crystallogr 56:1622–1624.
29. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53:240–255.
30. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126–2132.