The EMBO Journal Vol.16 No.12 pp.3396–3404, 1997

Structure of the PH domain and Btk motif from Bruton's tyrosine kinase: molecular explanations for X-linked agammaglobulinaemia

Marko Hyvönen and Matti Saraste[1]

European Molecular Biology Laboratory, Postfach 10.2209, 69012 Heidelberg, Germany

[1]Corresponding author
e-mail: saraste@embl-heidelberg.de

Bruton's tyrosine kinase (Btk) is an enzyme which is involved in maturation of B cells. It is a target for mutations causing X-linked agammaglobulinaemia (XLA) in man. We have determined the structure of the N-terminal part of Btk by X-ray crystallography at 1.6 Å resolution. This part of the kinase contains a pleckstrin homology (PH) domain and a Btk motif. The structure of the PH domain is similar to those published previously: a seven-stranded bent β-sheet with a C-terminal α-helix. Individual point mutations within the Btk PH domain which cause XLA can be classified as either structural or functional in the light of the three-dimensional structure and biochemical data. All functional mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids. It is likely that these mutations inactivate the Btk pathway in cell signalling by reducing its affinity for inositol phosphates, which causes a failure in translocation of the kinase to the cell membrane. A small number of signalling proteins contain a Btk motif that always follows a PH domain in the sequence. This small module has a novel fold which is held together by a zinc ion bound by three conserved cysteines and a histidine. The Btk motif packs against the second half of the β-sheet of the PH domain, forming a close contact with it. Our structure opens up new ways to study the role of the PH domain and Btk motif in the cellular function of Btk and the molecular basis of its dysfunction in XLA patients.

Keywords: Bruton's tyrosine kinase/Btk motif/inositol phosphates/PH domain/ X-linked agammaglobulinaemia

Introduction

X-linked agammaglobulinaemia (XLA) is a hereditary immune disease which is estimated to affect one in every 150 000 human males (Smith *et al*., 1994). XLA patients lack circulating B cells and subsequently have very low levels of immunoglobulins in their serum. They are very vulnerable to infectious diseases and require continuous medical treatment. The cause of this disease has recently been mapped to the gene encoding for Bruton's tyrosine kinase (Btk) (Tsukada *et al*., 1993; Vetrie *et al*., 1993). A large number of mutations in Btk have since been found in patients suffering from XLA (Vihinen *et al*., 1996). The XLA-causing mutations inactivate signal transduction pathways involving Btk and arrest development of B cells at the pre-B cell stage. Interestingly, a similar but milder disease in mice, X-linked immunodeficiency (*xid*), is also caused by a mutation within the Btk PH domain (Rawlings *et al*., 1993; Thomas *et al*., 1993).

Btk is a cytoplasmic protein tyrosine kinase (PTK) which belongs to the Tec family (Bolen, 1995). In addition to the Src homology 2 and 3 (SH2 and SH3) domains and the catalytic kinase domain, which are found in proteins belonging to the Src superfamily of PTKs, the members of the Tec family contain an N-terminal pleckstrin homology (PH) domain and a so-called Tec homology domain that includes a Btk motif and a proline-rich region. They occupy the place of the unique domain and myristylation site of Src. The Btk motif comprises only 26 residues which characteristically include three fully conserved cysteines and a histidine (Vihinen *et al*., 1994). This motif is found in only a few proteins outside the Tec family of PTKs and is always adjacent to a PH domain.

Btk can be activated by different extracellular signals such as cross-linking of the B-cell antigen receptor or binding of interleukin 5 (IL-5) (de Weers *et al*., 1994; Sato *et al*., 1994). Upon activation of the B-cell receptor, Btk is phosphorylated on Tyr551 by Src family tyrosine kinases and subsequently localized to the plasma membrane (Rawlings *et al*., 1996). This transphosphorylation of Btk leads to autophosphorylation of Tyr223 in the Btk SH3 domain (Park *et al*., 1996). Conversely, cells simulated with IL-5 do not show increased tyrosine phosphorylation but rather an enhanced activity of Btk (Sato *et al*., 1994).

The PH domains are small signal transduction domains which have been found in close to 100 different eukaryotic proteins up to now (Gibson *et al*., 1994). Although sequence conservation between different domains is very low, their three-dimensional structures appear to share a common fold and a remarkable electrostatic polarization. The positively charged end of several PH domains bind inositol phosphates *in vitro* with K_ds ranging from 210 nM to 1.23 mM for phospholipase Cδ1 (PLCδ1) and dynamin domains, respectively (Harlan *et al*., 1994; Lemmon *et al*., 1995; Salim *et al*., 1996). This binding enhances recruitment of PLCδ1 to the cell membrane *in vivo* (Paterson *et al*., 1995). Two three-dimensional structures of PH domains have been determined in complex with inositol(1,4,5)trisphosphate [Ins(1,4,5)P_3, herein we use the inositol phosphate nomenclature of Divecha and Irvine (1995)] (Ferguson *et al*., 1995; Hyvönen *et al*., 1995).

In most cases, the PH domains bind preferentially to Ins(1,4,5)P_3. However, a recent study by Salim *et al.* (1996) showed that the Btk PH domain specifically binds PtdIns(3,4,5)P_3 rather than PtdIns(4,5)P_2. Interestingly, this specificity for PtdIns(3,4,5)P_3 appears to be regulated by divalent ions, as it was only observed in the presence

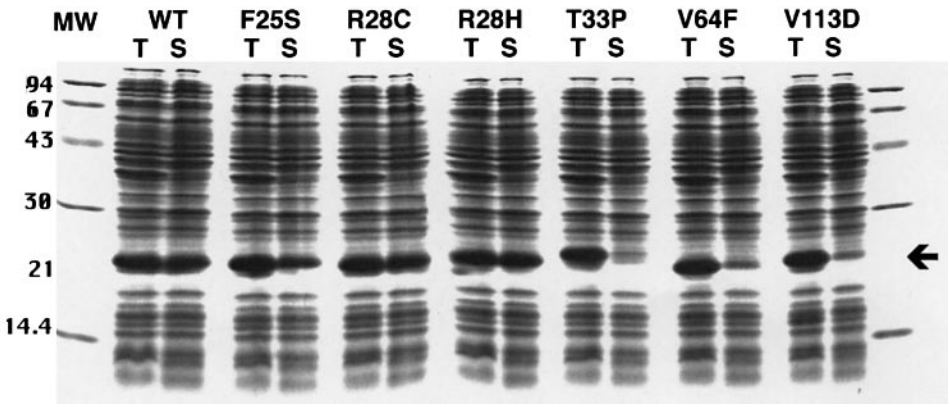

Fig. 1. Solubility of the mutants. The figure shows the SDS–PAGE of the wild-type and mutant proteins expressed in *E.coli* as described in Materials and methods. Total cell lysates (T) and high-speed supernatants (S) were mixed with Laemmli sample buffer, boiled and run in a 15% polyacrylamide gel. The wild-type protein is shown as control (WT), and the mutants are marked on top of the gel (F25S, R28H, R28C, T33P, V64F, V113D). The position of the expressed proteins is marked with an arrow. Molecular weight markers (in kDa) are indicated on the side of the gel.

of 2 mM Mg^{2+} or Ca^{2+}. Similar results have also been obtained with soluble inositol phosphates (Fukuda *et al*., 1996). The Btk PH domain showed tightest binding to Ins(1,3,4,5)P_4 compared with the other tested compounds, with 40 nM affinity.

Most of the structural and functional studies on the interaction of PH domains with their ligands have been carried out using soluble inositol phosphates, but they are thought to be relevant for binding of phosphatidylinositol lipids to the same sites. Consequently, in the following, we shall regard the ligand as a (soluble) phospholipid head group.

Several XLA-causing point mutations in Btk are located in the PH domain and one in the Btk motif. This is the only case in which mutations in a PH domain are known to cause a human disease. We have studied the roles of individual mutations in the PH domain of human Btk and determined the structure of the R28C mutant corresponding to the *xid* mutation in mouse by X-ray crystallography at 1.6 Å resolution. This structure allows us to discuss how the point mutations may inactivate Btk and subsequently cause XLA. Our structure also contains the Btk motif, which is shown to have a novel, zinc-binding fold. The function of this motif is still unknown and the structure will provide a rational basis for functional studies.

Results

The segment containing the N-terminal 170 amino acid residues of human Btk, including both the PH domain and the Btk motif, was expressed in *Escherichia coli* in a soluble form and purified to homogeneity. Earlier attempts to express the PH domain alone resulted in insoluble protein which resisted all refolding trials. Only the inclusion of the Btk motif in the construct allowed expression of a soluble, folded protein. Several XLA mutations in the PH domain (F25S, R28H, T33P, V64F, V113D) were introduced into similar expression constructs and the corresponding proteins were produced in *E.coli*. In addition, the protein carrying the mouse *xid* mutation (R28C) (Rawlings *et al.*, 1993; Thomas *et al.*, 1993) was produced.

Using the solubility of individual mutants in *E.coli* as a criterion, we could classify the mutations into two groups: (i) folding mutations, which prevent formation of a stable native-like structure, and (ii) functional mutations which do not affect the overall fold, but which can be expected to disrupt ligand-binding site(s) on the domain. As shown in Figure 1, wild-type protein and mutants F25S, R28C and R28H are soluble in *E.coli*, whereas mutants T33P, V64F and V113D fail to yield soluble protein under the same conditions. Structural reasons for this behaviour are discussed later. The far-UV circular dichroism (CD) spectra of the purified wild-type protein and soluble mutants are virtually identical (data not shown).

The wild-type protein and all the soluble mutants were used for crystallization trials, but only the mouse *xid* mutant R28C gave high quality crystals. Crystals belong to space group P21 (unit cell a = 49.15, b = 59.87, c = 55.94 Å, β = 98.2°), with two molecules in the asymmetric unit. Crystals diffracted beyond 1.5 Å on a synchrotron source. The structure was solved with a single heavy metal derivative (trimethyl lead acetate) using both isomorphous and anomalous differences (see Materials and methods). The model is refined at 1.6 Å resolution; refinement statistics are shown in Table I. In addition to the two protein molecules in the asymmetric unit, 220 water molecules, two zinc ions and two sodium ions are included in the model. Metal analysis showed that the protein contains 0.93 mol of zinc per 1 mol of protein. Accordingly, the ions that where found to be present in the crystal structure of the Btk motif were assigned as zinc. Sodium chloride was required for crystallization of the protein and, in the final stages of the refinement, two ions were found with a typical octahedral coordination characteristic of sodium. The sodium ions are bound between the two molecules in the asymmetric unit.

Structure of the Btk PH domain and the Btk motif

Similarly to the other PH domains, the structure is composed of a strongly bent seven-stranded antiparallel β-sheet which packs against a C-terminal α-helix (Figure 2). The Btk PH domain contains a long insertion in the loop between β-strands 5 and 6 which includes a short 1.5-turn α-helix (α1). It points away from the core of the domain and its hydrophilic middle part is disordered and could

Table I. Data collection, phasing and refinement statistics

	In-house native	$(CH_3)_3PbAc$	BW7B at EMBL-Hamburg c/o DESY
Resolution	2.6	2.6	1.6
Wavelength (Å)	1.514	1.514	0.882
Completeness (highest shell) (%)	99.8 (98.2)	97.2 (95.8)	99.9 (99.5)
R_{merge} (high shell)[a]	0.060 (0.120)	0.061 (0.149)	0.054 (0.236)
No. of unique reflections	9817	9319	42 840
Phasing (2.6 Å)[b] isomorphous/anomalous		2.77/2.85	
Figure of merit		0.602	
No. of sites		1	

Refinement statistics (XPLOR)						
Resolution (Å)	40.5–2.74	2.74–2.17	2.17–1.98	1.98–1.72	1.72–1.60	40.5–1.60
R-factor (%)	18.6	24.1	26.1	29.8	31.5	23.1
R_{free}(%)	25.8	29.9	27.6	29.5	33.1	28.0
Geometrical statistics (WHAT CHECK)						
R.m.s. deviation from ideal bond lenghts (Å)	0.11					
R.m.s. deviation from ideal bond angles (°)	1.989					

[a]$R_{merge} = \Sigma|I_i - \langle I \rangle|/\Sigma I_i$, where I_i is the intensity of an individual reflection and $\langle I \rangle$ is the mean intensity of the reflection. The value in parenthesis is for the highest resolution shell.
[b]Phasing power is the ratio between the root mean square of the heavy atom scattering amplitude and lack of closure error.

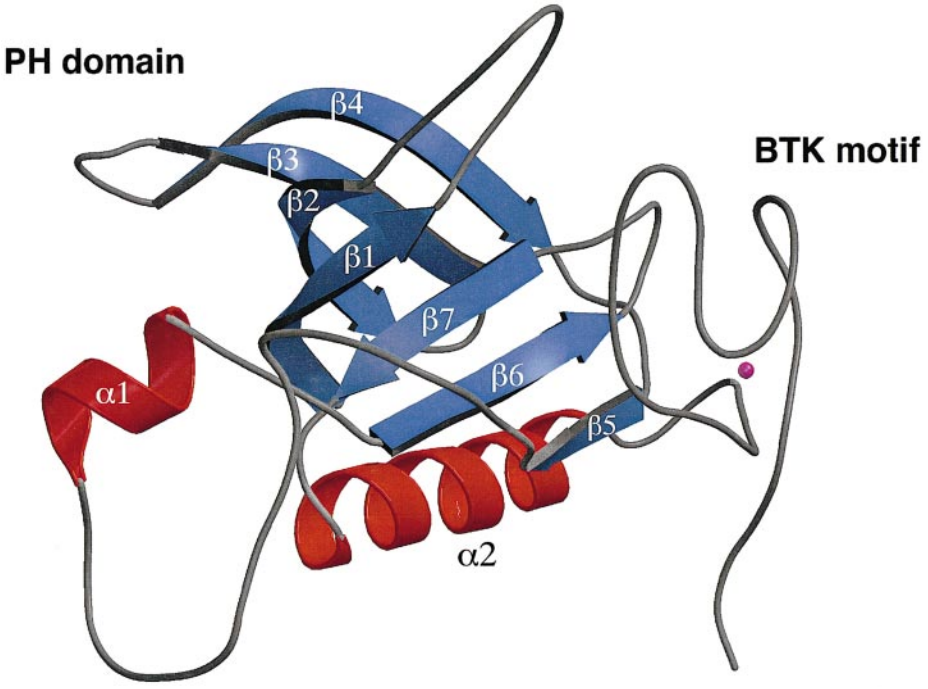

Fig. 2. The structure of the Btk PH domain and Btk motif. The figure shows a ribbon representation of the R28C mutant of the human Btk PH domain and Btk motif. The β-strands and α-helixes are numbered β1–β7 and α1–α2. The zinc ion in the Btk motif is shown in purple. This figure and Figures 5A and B, 6A and B were prepared using MOLSCRIPT (Kraulis, 1991) and rendered in Raster3d (Bacon and Anderson, 1988; Merrit and Murphy, 1994).

not be modelled properly. Also, the loop between β-strands 1 and 2 has weak electron density and was partly modelled as polyalanine. There will naturally be differences in our structure around the mutated residue R28C compared with the wild-type protein, but these are not thought to be significant to our analysis of the structure.

The Btk motif has a globular core which packs against β-strands 5–7 of the PH domain. F146 in the Btk motif is inserted into a hydrophobic pocket formed by residues V67, P101, Y112 and F114 (labelled in Figure 4). The core of the Btk motif is connected to the C-terminus of the PH domain with a 7 or 8 residue linker. The structure could be described as a long loop which folds back on itself and is held together by a zinc ion. This zinc is coordinated with a distorted tetrahedral geometry to H143 and C154, C155 and C165 (Figure 3), which are fully conserved in Btk motifs (Vihinen *et al.*, 1994).

Database searches using the program DALI (Holm and Sander, 1993) have revealed no structural homology for the Btk motif in the Protein Data Bank (PDB). The diacylglycerol/phorbol ester-binding domain (PDB entry 1PTR), LIM repeats (1IMR) and GATA-1 zinc finger (1PTR) have a similar zinc coordination with three cysteines and histidine, and a similar tight turn (formed by C154 and C155 in the Btk motif) at the zinc-binding site, but otherwise the structures seem to be unrelated.

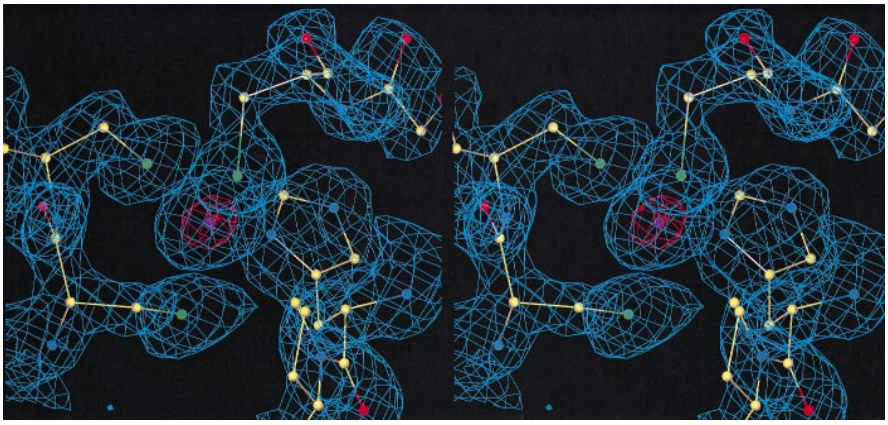

Fig. 3. A stereo view of the electron density in the zinc-binding site of the Btk motif. The $2F_o–F_c$ omit electron density map around the zinc ion in the Btk motif is shown. Prior to the map calculation, the zinc ion and residues inside a 5 Å sphere around the zinc ion were removed, and the model was subjected to a conventional positional refinement in program X-PLOR. A $2F_o–F_c$ electron density map was calculated with the resulting model and contoured at 1σ (cyan) and 10σ (orange). The final protein model is shown as a ball-and-stick model and coloured according to the atom type (C yellow, N blue, O red, S green, Zn purple). The residues shown in the figure are C154 and C155 on the left side of the zinc, C165 on the top and H143 on the right side of the ion. Bonding distances between the zinc and its ligands are 2.24–2.36 Å for zinc–sulfur and 2.05–2.06 Å for zinc–nitrogen bonds.

Inositol phosphate-binding site

The three-dimensional structures of Ins(1,4,5)P_3 complexes have been solved for the β-spectrin and PLCδ1 PH domains. These two structures show variant binding sites around loop 1–2. Ins(1,4,5)P_3 binds to the β-spectrin PH domain between loops 1–2 and 5–6 (Hyvönen *et al*., 1995), whereas the PLCδ1 PH domain binds the ligand between loops 1–2 and 3–4 (Ferguson *et al*., 1995). Previous sequence- and structure-based alignments and three-dimensional models of the Btk PH domain made prior to the knowledge of the experimental structure have an error which will position the residues involved in the binding of inositol phosphates in the first β-strand incorrectly for Btk (Musacchio *et al*., 1993; Ferguson *et al*., 1995; Vihinen *et al*., 1995). Also, an insertion of a gap in the beginning of the second β-strand has led to incorrect interpretations (Ferguson *et al*., 1995).

The revised alignment is shown in Figure 4, in which the residues involved in ligand binding are highlighted. This comparison reveals that the Btk PH domain contains key residues for both inositol-binding sites. On one side of loop 1–2, residues R13, F25 and K77 correspond to the Ins(1,4,5)P_3-binding site in the β-spectrin domain, and on the other side of the loop, residues K12, N24, R28 and K53 are equivalent to those binding Ins(1,4,5)P_3 in the PLCδ1 domain (Figure 5A).

Although many of the residues that bind Ins(1,4,5)P_3 in the β-spectrin PH domain are conserved in Btk, it is unlikely that the latter would bind a ligand in this region. R13 of the Btk domain is hydrogen-bonded to the carbonyl group of W147 in the Btk motif and is not available for ligand binding. This site is also partly covered by the Btk motif and loop 1–2. It is more likely that the binding site of Ins(1,3,4,5)P_4 in the Btk domain is similar to the site in the PLCδ1 domain. Many of the ligand-binding residues are present in Btk, and most of the XLA mutations point to this site. In the complex of the PLCδ1 PH domain and Ins(1,4,5)P_3, additional contacts come from residues 54–57 in loop 3–4. In Btk, this loop has positive residues (R46, R48, R49) but it points away from the binding site, and a rather large conformational change would be required to bring these residues close enough to interact with a ligand.

E41 of the Btk PH domain is close to the predicted binding site in loop 3–4. In a random mutagenesis study, Li *et al.* (1995) recently found a constitutively active form of Btk which is caused by a substitution of this residue by lysine (E41K). This mutant can transform NIH3T3 cells and shows enhanced tyrosine phosphorylation and membrane localization as compared with the wild-type protein. These results suggest that the E41K mutant has increased affinity for a membrane-bound (lipid) ligand, supporting our assignment of the binding site.

All PH domain structures show a clear polarization of charges, and the binding sites for Ins(1,4,5)P_3 are located in the positive ends of the domains. In the Btk PH domain, the residues within the proposed Ins(1,3,4,5)P_4-binding site are similarly in the most positively charged area of the domain (Figure 5B).

XLA mutations

The database of XLA-causing mutations in Btk, the BTKbase, contains nine point mutations within the PH domain and one in the Btk motif (Vihinen *et al*., 1996). We have constructed and expressed five of them: F25S, R28H, T33P, V64F and V113D. Mutations L11P, K12R, S14F and R28P in the PH domain and C155G in the Btk motif have been reported while our work was in progress.

Structural analysis of the mutations (T33P, V64F and V113D) which affect folding and/or stability of the domain shows the following. T33 is in loop 2–3, and mutation to a rigid proline would disrupt the structure of the loop. V64 is part of the hydrophobic core and is fully conserved in all proteins containing both the PH domain and the Btk motif. Introduction of a phenylalanine in this position cannot be accommodated. V113 is also a conserved, buried hydrophobic residue, and its substitution by an aspartate will introduce an unfavourable polar residue into the core of the domain.

All other XLA-causing mutations in the Btk PH domain

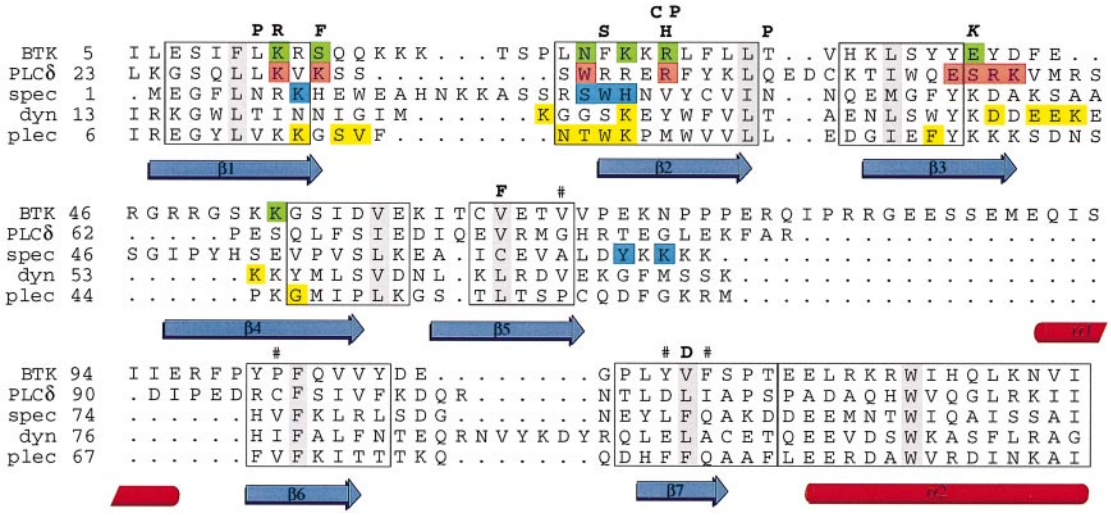

Fig. 4. Structural alignment of PH domains. Structural alignment of PH domains from Btk, PLCδ1 (Ferguson *et al*., 1995) (PDB entry 1MAI), β-spectrin (Hyvönen *et al*., 1995) (1BTN), dynamin (Ferguson *et al*., 1994) (1DYN) and pleckstrin (Yoon *et al*., 1994) (1PLS, the first of the 10 NMR structures in the PDB entry was used for the superpositioning). Superpositioning and the alignment were generated with the MODELLER program (Sali and Blundell, 1993). Secondary structure elements of the Btk PH domain are shown below the alignment, with blue arrows for β-sheets and red cylinders for α-helices, and numbered as in Figure 2. Residues which are assigned to the Ins(1,3,4,5)P_4-binding site of Btk are coloured green. Ins(1,4,5)P_3-binding residues are coloured blue in β-spectrin and red in the PLCδ1 domain. Residues showing chemical shifts upon binding to Ins(1,4,5)P_3 in pleckstrin and dynamin PH domains are coloured yellow. The boxed areas of the alignment correspond to the well-defined and superimposable parts of the PH domains. The most conserved hydrophobic residue in each of these blocks appears on a grey background. The XLA-causing point mutations are marked on top of the Btk PH domain sequence, and the activating mutation E41K is shown in italics. A hash-mark shows the residues forming the hydrophobic pocket for F146 of the Btk motif. The numbering corresponds to that of the PDB entries and is used in the text.

are located around loop 1–2 in the area where the predicted binding site for inositol phosphates resides (Figure 6A and B). K12 of Btk and K30 of PLCδ1 are in equivalent positions, and point to the inositol-binding site of PLCδ1. In a Japanese XLA patient, this residue is mutated to arginine (Hashimoto *et al*., 1996). In the PLCδ1 PH domain, K30 binds to both 4- and 5-phosphates of Ins(1,4,5)P_3, and a guanido group of arginine could not be accommodated in this position. Similar mutation of a conserved, phosphate-binding lysine to arginine in the ATP-binding sites of protein kinases are known to inactivate these enzymes (Hanks *et al*., 1988).

S14 points in the same direction as R12, and the S14F mutation would sterically block the predicted binding site. Another mutated residue of the Btk domain is F25 in the beginning of the second β-strand. It is equivalent to W23 in the β-spectrin PH domain, one of the Ins(1,4,5)P_3-binding residues (Hyvönen *et al*., 1995). F25S mutation would disrupt this binding site if Btk bound an inositol phosphate in the same site as β-spectrin. The other possible consequence of this mutation, which we favour, is destabilization of loop 1–2 and an indirect effect on the other putative binding site. This is supported by the reduced solubility of this mutant during expression and purification.

R28 is substituted with a histidine or a proline in human XLA and with a cysteine in mouse *xid*. It corresponds to R40 in the PLCδ1 domain, a residue in contact with the 5-phosphate of Ins(1,4,5)P_3. All of these mutations (R28H, R28P and R28C) remove a positive charge, and would abolish binding of the ligand.

L11P mutation could affect the folding of the domain, although we have not tested its expression in *E.coli*. This residue is at the end of the first β-strand where a proline might be structurally tolerated. It could, however, affect the following residue K12 by positioning it unfavourably for interaction with an inositol phosphate ligand.

In accordance with our interpretations, the mutations F25S, R28C and R28H have been shown recently to reduce the affinity for Ins(1,3,4,5)P_4 compared with the wild-type protein *in vitro* (Fukuda *et al*., 1996). In the same study, the mutants T33P, V64F and V113D showed no binding to the ligand, as could be expected if the protein cannot acquire its native fold.

The single known XLA mutation in the Btk motif is C155G. This cysteine is one of the conserved zinc ligands, and its substitution will disrupt the metal centre and prevent folding of the motif.

Discussion

All mutations in the PH domain of Btk cause severe forms of XLA and are likely to abolish completely the binding of a natural ligand. The structure of the Btk PH domain has revealed that, in most cases, the mutations in the PH domain are located around a putative inositol phosphate-binding site. Previous attempts to explain these mutations by sequence analysis and modelling have been inaccurate, mainly because of a misalignment of the first β-strand.

All functional mutations in the Btk PH domain cluster onto one end of the domain. Is this the only functionally important area on the domain? Apart from inositol phosphates, only a few other interaction partners for PH domains have been suggested, and the data supporting these findings are often rather weak. The βγ-subunits of trimeric G-proteins are believed to interact with some PH domains (Touhara *et al*., 1994). There is evidence that this is also the case for the Btk PH domain (Tsukada *et al*., 1994). In the β-adrenergic receptor kinase (βARK), the binding site for the βγ-subunits has been mapped to

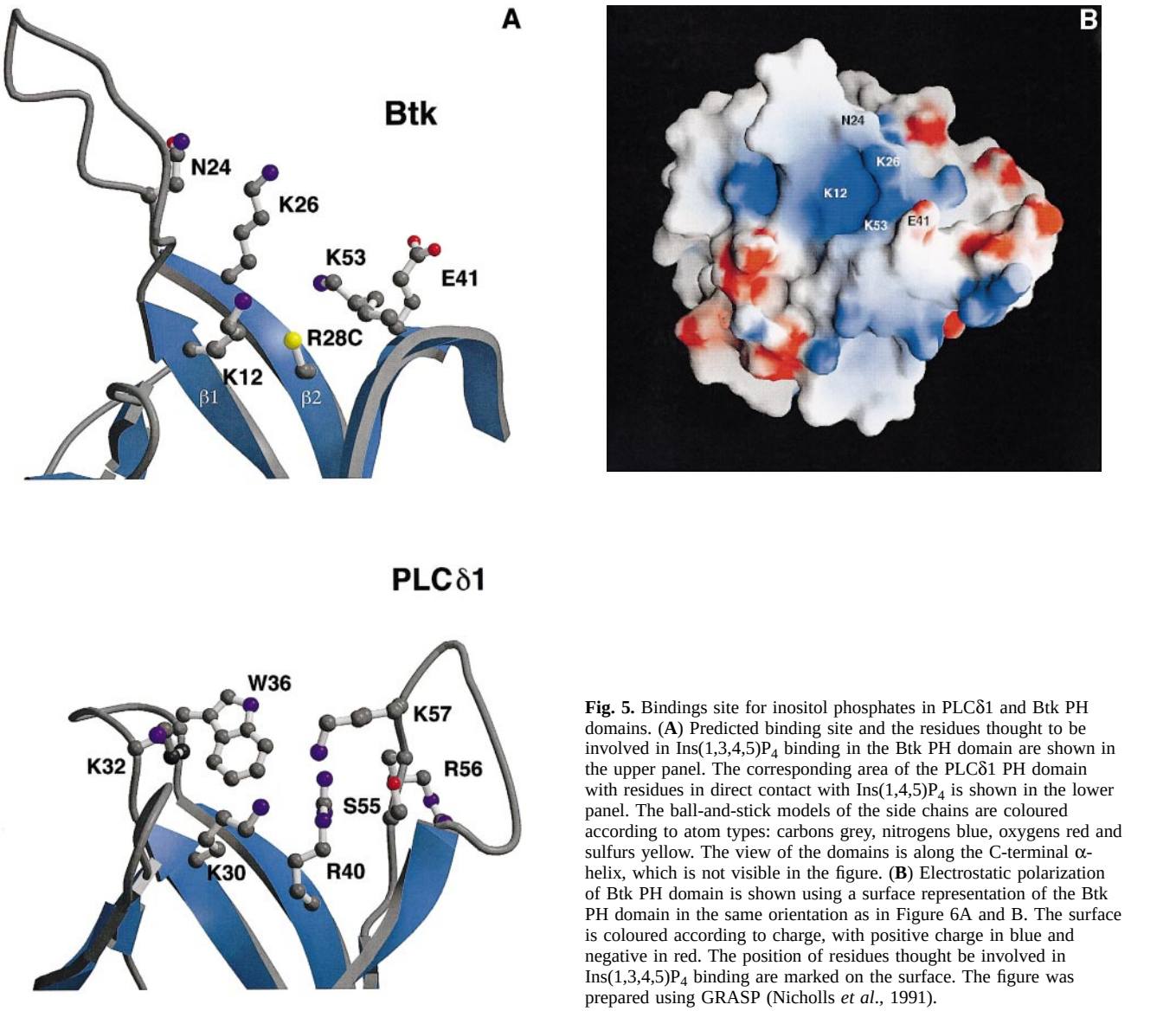

Fig. 5. Bindings site for inositol phosphates in PLCδ1 and Btk PH domains. (**A**) Predicted binding site and the residues thought to be involved in Ins(1,3,4,5)P_4 binding in the Btk PH domain are shown in the upper panel. The corresponding area of the PLCδ1 PH domain with residues in direct contact with Ins(1,4,5)P_4 is shown in the lower panel. The ball-and-stick models of the side chains are coloured according to atom types: carbons grey, nitrogens blue, oxygens red and sulfurs yellow. The view of the domains is along the C-terminal α-helix, which is not visible in the figure. (**B**) Electrostatic polarization of Btk PH domain is shown using a surface representation of the Btk PH domain in the same orientation as in Figure 6A and B. The surface is coloured according to charge, with positive charge in blue and negative in red. The position of residues thought be involved in Ins(1,3,4,5)P_4 binding are marked on the surface. The figure was prepared using GRASP (Nicholls *et al*., 1991).

the very C-terminus of the PH domain, and residues beyond the domain are needed for this interaction. In Btk, the binding site has not been mapped precisely. The site-directed mutagenesis and deletions of the possible binding site in Btk were done without structural considerations, and most of these mutations would certainly affect folding and stability of the domain (Tsukada *et al*., 1994). Co-expression of several different βγ-subunit combinations with Btk has been shown to activate the kinase in a PH domain-dependent manner, but evidence for a direct interaction is missing (Langhans Rajasekaran *et al*., 1995).

Yao and co-workers (1994) have shown that the Btk PH domain interacts with various forms of protein kinase C (PKC). The PH domain was found to be phosphorylated on serine residue(s) and interacted with PKCs both *in vitro* and *in vivo*. This interaction was not very strong, but it may nevertheless point to a possible role for PKCs in regulating Btk. PKC-dependent serine phosphorylation of Btk was shown to down-regulate its kinase activity. PKCβ seems to be involved in the same pathway as Btk in development of B cells as it has been recently described that a PKCβ knock-out mouse has a similar phenotype to the *xid* mouse (Leitges *et al*., 1996). Activation of Btk in this mouse is, however, not affected and PKCβ is likely to act downstream of Btk.

The binding site for inositol phosphates is predicted to be similar to that in PLCδ1, but the ligand for Btk does not seem to be Ins(1,4,5)P_3, but rather Ins(1,3,4,5)P_4 (Fukuda *et al*., 1996; Salim *et al*., 1996). The same ligand has also been described for Ras-GAP proteins which contain PH domains, namely human IP4BP and murine GAP1^m (Cullen *et al*., 1995; Fukuda and Mikoshiba, 1996). In the case of GAP1^m, ligand binding requires additional elements outside the PH domain. Both of these proteins also contain a Btk motif.

Most of the binding studies between PH domains and inositol phosphates have been conducted with soluble head groups of inositol phospholipids, but more careful studies using phospholipid vesicles are needed to determine the role of other lipids in membrane interaction of PH domains. A 3-fold increase in membrane affinity has

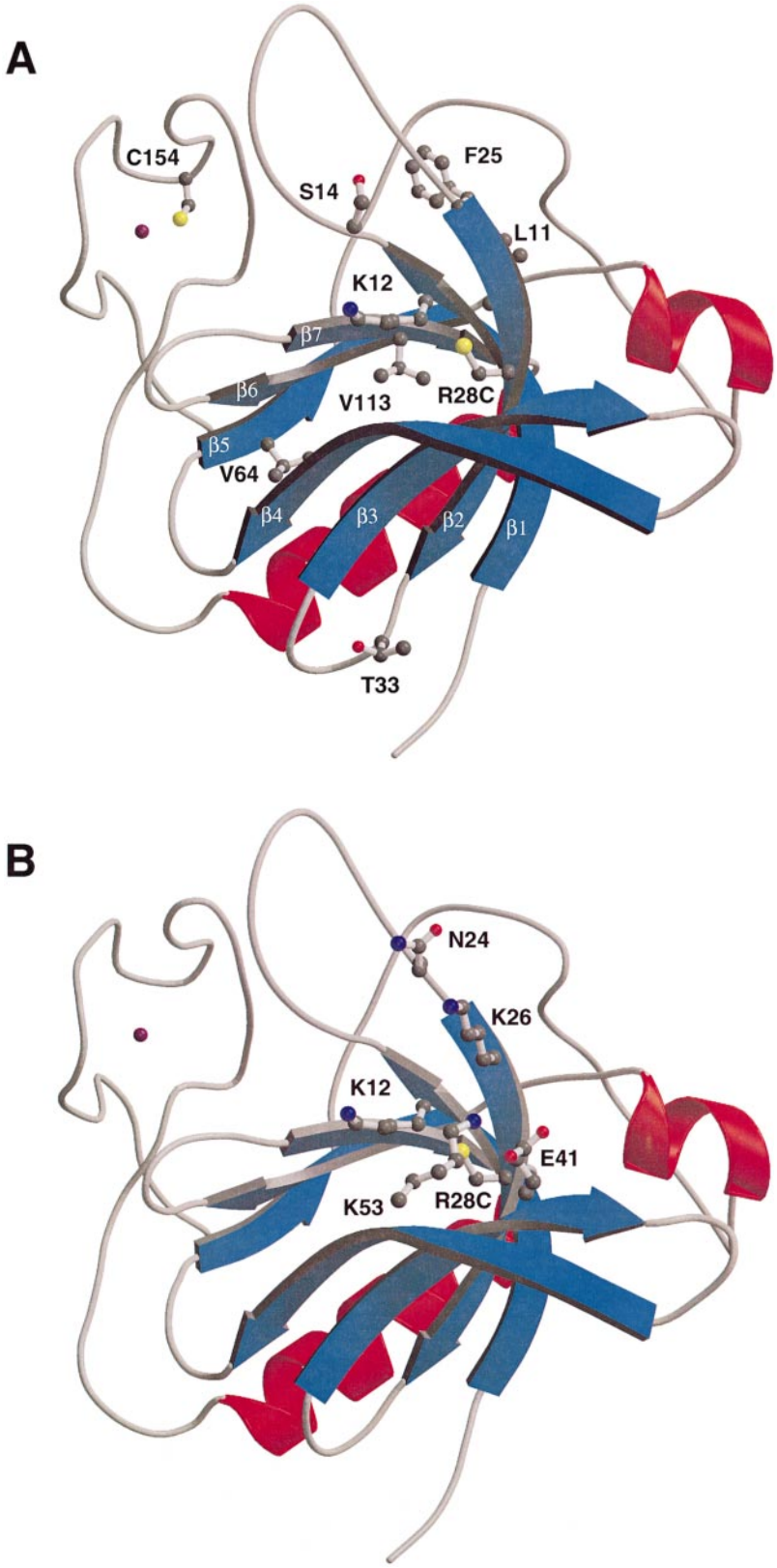

Fig. 6. Mutations and the predicted inositol-binding site in the Btk PH domain. (**A**) A ribbon representation of the Btk PH domain and Btk motif showing residues mutated in XLA patients as ball-and-stick models. Colouring is as in Figures 2 and 5A. (**B**) Same view of the molecule as in (A), showing the residues thought to form the Ins(1,3,4,5)P_4-binding site as ball-and-stick models.

been demonstrated for PLCδ1 using vesicles containing both PtdIns(4,5)P_2 and phosphatidylserine (Rebecchi *et al*., 1992). A similar co-operative role for other negative phospholipids in the membrane is thought to affect the membrane association of other PH domains (Hyvönen *et al*., 1995).

Materials and methods

Cloning, expression and purification of the domains

A DNA fragment encoding amino acids 1–170 of human Btk (Swiss-Prot: Q06187) was amplified by PCR using human Btk cDNA as template (the cDNA was kindly provided by Dr C.I.E.Smith) and cloned into *E.coli* expression vector pBAT4 (Peränen *et al*., 1996). Mutants of the Btk PH domain were created by PCR-directed site-specific mutagenesis using the wild-type expression construct as a template. All expression constructs were verified by dideoxy sequencing.

Expression of proteins was carried out in the *E.coli* strain BL21(DE3) carrying plasmid pUBS520 (Brinkmann *et al*., 1989). In order to obtain soluble protein, the expression was carried out at 15°C for 20 h after induction. The cells were collected by centrifugation and resuspended in 50 mM sodium phosphate buffer pH 7.2. They were lysed by passing them twice through a French Press, and the lysate was centrifuged for 30 min at 100 000 *g* to separate soluble and insoluble fractions. The solubility of the mutant proteins was analysed by SDS–PAGE (Laemmli, 1970). Soluble wild-type and mutant proteins were purified using three chromatographic steps. Cation-exchange chromatography on an S-Sepharose FF column (all columns are from Pharmacia, Sweden) was run with the lysis buffer, and protein was eluted with a linear salt gradient at 250–300 mM NaCl. Gel filtration in a Superdex 75 16/60 preparation grade column was run with 20 mM EPPS buffer (Sigma, USA), 100 mM NaCl, 2 mM dithiothreitol (DTT), pH 8.0. Protein from gel filtration was diluted 3-fold, and loaded onto a MonoS HR 10/10 cation-exchange column equilibrated with 20 mM EPPS, 2 mM DTT, pH 8.0. Protein eluted at ~250 mM NaCl. Purified protein was concentrated to 6–8 mg/ml in 10 mM Tris, pH 8.0, 100 mM NaCl, 2 mM DTT, and stored in small aliquots at –80°C. The molecular weights of purified proteins were verified by mass spectroscopy, which showed that the N-terminal methionine was cleaved off. The protein used in all studies thus corresponds to residues 2–170 of human Btk.

Metal analysis was done by atomic absorption spectroscopy at the University of Saarland, (Saarbrücken, Germany) in the laboratory of Dr Michael Zeppezauer. Protein concentration was determined spectrophotometrically in 6 M guanididinium–HCl using a calculated molar absorption coefficient of 23 200/M (Gill and von Hippel, 1989).

Crystallization

Crystallization trials were conducted at room temperature in hanging drops. Diffraction quality crystals were obtained only with the mutant R28C with 32.5% (w/v) PEG 3350 (Sigma, USA), 100 mM Tris pH 8.5, 200 mM $MgCl_2$, 500 mM NaCl as the well solution. Plate-like crystals appeared after 24–48 h and were stable for several weeks.

Data collection and heavy metal derivative search

All datasets were collected at 110 K using a cryo-stream cooler from Oxford Cryosystems. In-house data collection was done using a MAR Research imaging plate detector and a Siemens GX21 generator running at 40 kV and 70 mA. Synchrotron data was collected at the EMBL Hamburg Outstation c/o DESY on the BW7B beamline using a MAR Research imaging plate. Data were collected in two passes to obtain reliable low resolution data also. All datasets were processed with DENZO, and scaled and merged in SCALEPACK (Z.Otwinowski and W.Minor, in preparation).

During the search for the heavy metal derivatives, it became evident that crystals were often non-isomorphous. An the end, only one heavy metal derivative, trimethyl lead acetate [$(CH_3)_3$PbAc], was found, but the non-isomorphous nature of the crystals lowered the quality of the data. To overcome this problem, both native and derivative data sets were collected from the same single crystal. After collecting a high quality native dataset, the crystal was thawed, soaked in 1 mM $(CH_3)_3$PbAc in the mother liquor for 24 h, and a derivative data set was then collected under cryo conditions. The resulting datasets were isomorphous and the derivative proved to have very good phasing power (Table I).

Phasing and solvent flattening

Phasing was done using the PHASES package (Furey and Swaminathan, 1996). The position of the heavy metal was determined by Patterson methods. Phasing and phase refinement was done using both isomorphous and anomalous differences. Resulting SIRAS phases were already good enough to allow partial model building and determination of the non-crystallographic symmetry (NCS). These phases were then used in the program DM of the CCP4 package for solvent flattening, histogram matching and two-fold NCS averaging. Over 80% of the molecule was built into the resulting electron density maps.

As high-resolution data became available, original SIRAS phases were extended to 1.6 Å with DM using NCS averaging, histogram matching and solvent flattening. Electron density maps calculated with these experimental phases confirmed the correctness of the model and revealed the positions of several previously missing residues.

Refinement and structure analysis

Refinement was carried out using the programs X-PLOR (Brünger, 1988), TNT (Tronrud *et al*., 1987) and ARP (Lamzin and Wilson, 1993). Five percent of the data (2164 reflections in the high-resolution dataset) was set aside for the calculation of the free *R*-factor (R_{free}), which was monitored during the refinement. The model was fixed between the refinement cycles on a graphics terminal using the program O. In the beginning of the refinement, the NCS constraints were applied, but were released gradually upon inclusion of high-resolution data. The structure was analysed using PROCHECK (Laskowski *et al*., 1993) and WHAT CHECK of the WHAT IF program (Vriend, 1990).

In the final map, two loops of the PH domain have very weak electron density. The loops between β-strands 1 and 2 are modelled partly as polyalanine, as is loop 5–6 in the second molecule within the asymmetric unit. In the first molecule, residues 80–88 of this loop are left out completely. Both of these loops point away from the core of the domain and do not seem to have contacts with it. Several different refinement protocols were tried, but none of them was able to improve the quality of the electron density for these two loops.

The coordinates will be deposited to the Brookhaven Protein Data Bank. Before that, the coordinates will be available from the authors by e-mail.

Acknowledgements

We wish to thank people at the EMBL-Hamburg Outstation c/o DESY, in particular Victor Lamzin, for helping with the data collection, Kristina Djinovic-Carugo and Matthias Wilmanns for their generous advice during structure determination, Jaime Pascual and Kristina Djinovic-Carugo for critically reading the manuscript, Liisa Holm for helping with DALI searches, Toby Gibson for discussions, Edvard Smith for providing the cDNA of human Btk, and Michael Zeppezauer and his colleagues for performing the metal analysis.

References

Bacon,D.J. and Anderson,W.F. (1988) A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graphics*, **6**, 219–220.

Bolen,J.B. (1995) Protein tyrosine kinases in the initiation of antigen receptor signaling. *Curr. Opin. Immunol.*, **7**, 306–311.

Brinkmann,U., Mattes,R.E. and Buckel,P. (1989) High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product. *Gene*, **85**, 109–114.

Brünger,A.T. (1988) *X-PLOR Version 3.1 A System for X-ray Crystallography and NMR*. Yale University Press, New Haven and London.

Cullen,P.J., Hsuan,J.J., Truong,O., Letcher,A.J., Jackson,T.R., Dawson, A.P. and Irvine,R.F. (1995) Identification of a specific Ins(1,3,4,5)P_4-binding protein as a member of the GAP1 family. *Nature*, **376**, 527–530.

de Weers,M., Brouns,G.S., Hinshelwood,S., Kinnon,C., Schuurman,R.K., Hendriks,R.W. and Borst,J. (1994) B-cell antigen receptor stimulation activates the human Bruton's tyrosine kinase, which is deficient in X-linked agammaglobulinemia. *J. Biol. Chem.*, **269**, 23857–23860.

Divecha,N. and Irvine,R.F. (1995) Phospholipid signaling. *Cell*, **80**, 269–278.

Ferguson,K.M., Lemmon,M.A., Schlessinger,J. and Sigler,P. (1994) Crystal structure at 2.2 Å resolution of the pleckstrin homology domain from human dynamin. *Cell*, **79**, 199–209.

Ferguson,K.M., Lemmon,M.A., Schlessinger,J. and Sigler,P.B. (1995) Structure of the high affinity complex of inositol trisphosphate with a phospholipase C pleckstrin homology domain. *Cell*, **83**, 1037–1046.

Fukuda,M. and Mikoshiba,K. (1996) Structure–function relationships of the mouse Gap1m. Determination of the inositol 1,3,4,5-tetrakisphosphate-binding domain. *J. Biol. Chem.*, **271**, 18838–18842.

Fukuda,M., Kojima,T., Kabayama,H. and Mikoshiba,K. (1996) Mutation of the pleckstrin homology domain of Bruton's tyrosine kinase in immunodeficiency impaired inositol 1,3,4,5-tetrakisphosphate binding capacity. *J. Biol. Chem.*, **271**, 30303–30306.

Furey,W. and Swaminathan,S. (1996) PHASES-95: a program package for the processing and analysis of diffraction data from macromolecules. *Methods Enzymol.*, in press.

Gibson,T.J., Hyvönen,M., Musacchio,A., Saraste,M. and Birney,E. (1994) PH domain: the first anniversary. *Trends Biochem. Sci.*, **19**, 349–353.

Gill,S.C. and von Hippel,P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.*, **182**, 319–326.

Hanks,S.K., Quinn,A.M. and Hunter,T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science*, **241**, 41–52.

Harlan,J.E., Hajduk,P.J., Yoon,H.S. and Fesik,S.W. (1994) Pleckstrin homology domains bind to phosphatidylinositol-4,5-bisphosphate. *Nature*, **371**, 168–170.

Hashimoto,S. *et al*. (1996) Identification of Bruton's tyrosine kinase (Btk) gene mutations and characterization of the derived proteins in 35 X-linked agammaglobulinemia families: a nationwide study of Btk deficiency in Japan. *Blood*, **88**, 561–573.

Holm,L. and Sander,C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.*, **233**, 123–138.

Hyvönen,M., Macias,M.J., Nilges,M., Oschkinat,H., Saraste,M. and Wilmanns,M. (1995) Structure of the binding site for inositol phosphates in a PH domain. *EMBO J.*, **14**, 4676–4785.

Kraulis,P. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of proteins. *J. Appl. Crystallogr.*, **24**, 946–950.

Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.

Lamzin,V. and Wilson,K. (1993) Automated refinement of protein models. *Acta Crystallogr.*, **D49**, 129–147.

Langhans Rajasekaran,S.A., Wan,Y. and Huang,X.Y. (1995) Activation of Tsk and Btk tyrosine kinases by G protein βγ subunits. *Proc. Natl Acad. Sci. USA*, **92**, 8601–8605.

Laskowski,R.A., MacArthur,M.W., Hutchinson,E.G. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.

Leitges,M., Schmedt,C., Guinamard,R., Davoust,J., Schaal,S., Stabel,S. and Tarakhovsky,A. (1996) Immunodeficiency in protein kinase Cβ-deficient mice. *Science*, **273**, 788–791.

Lemmon,M.A., Ferguson,K.M., O'Brien,R., Sigler,P.B. and Schlessinger,J. (1995) Specific and high-affinity binding of inositol phosphates to an isolated pleckstrin homology domain. *Proc. Natl Acad. Sci. USA*, **92**, 10472–10476.

Li,T., Tsukada,S., Satterthwaite,A., Havlik,M.H., Park,H., Takatsu,K. and Witte,O.N. (1995) Activation of Bruton's tyrosine kinase (BTK) by a point mutation in its pleckstrin homology (PH) domain. *Immunity*, **2**, 451–460.

Merrit,E.A. and Murphy,M.E.P. (1994) Raster3D version 2.0—a program for photorealistic molecular graphics. *Acta Crystallogr.*, **D50**, 869–873.

Musacchio,A., Gibson,T., Rice,P., Thompson,J. and Saraste,M. (1993) The PH domain: a common piece in the structural patchwork of signalling proteins. *Trends Biochem. Sci.*, **18**, 343–348.

Nicholls,A., Sharp,K.A. and Honig,B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.*, **11**, 281–296.

Park,H., Wahl,M.I., Afar,D.E., Turck,C.W., Rawlings,D.J., Tam,C., Scharenberg,A.M., Kinet,J.P. and Witte,O.N. (1996) Regulation of Btk function by a major autophosphorylation site within the SH3 domain. *Immunity*, **4**, 515–525.

Paterson,H.F., Savopoulos,J.W., Perisic,O., Cheung,R., Ellis,M.V., Williams,R.L. and Katan,M. (1995) Phospholipase Cδ1 requires a pleckstrin homology domain for interaction with the plasma membrane. *Biochem. J.*, **312**, 661–666.

Peränen,J., Rikkonen,M., Hyvönen,M. and Kääriäinen,L. (1996) T7 vectors with modified T7lac promoter for expression of proteins in *Escherichia coli*. *Anal. Biochem.*, **236**, 371–373.

Rawlings,D.J. *et al*. (1993) Mutation of unique region of Bruton's tyrosine kinase in immunodeficient XID mice. *Science*, **261**, 358–361.

Rawlings,D.J., Scharenberg,A.M., Park,H., Wahl,M.I., Lin,S., Kato,R.M., Fluckiger,A.C., Witte,O.N. and Kinet,J.P. (1996) Activation of BTK by a phosphorylation mechanism initiated by SRC family kinases. *Science*, **271**, 822–825.

Rebecchi,M., Peterson,A. and McLaughlin,S. (1992) Phosphoinositide-specific phospholipase C-δ1 binds with high affinity to phospholipid vesicles containing phosphatidylinositol 4,5-bisphosphate. *Biochemistry*, **31**, 12742–12747.

Sali,A. and Blundell,T. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, **234**, 779–815.

Salim,K. *et al*. (1996) Distinct specificity in the recognition of phosphoinositides by the pleckstrin homology domains of dynamin and Bruton's tyrosine kinase. *EMBO J.*, **15**, 6241–6250.

Sato,S. *et al*. (1994) IL-5 receptor-mediated tyrosine phosphorylation of SH2/SH3-containing proteins and activation of Bruton's tyrosine and Janus 2 kinases. *J. Exp. Med.*, **180**, 2101–2111.

Smith,C.I., Islam,K.B., Vorechovsky,I., Olerup,O., Wallin,E., Rabbani,H., Baskin,B. and Hammarstrom,L. (1994) X-linked agammaglobulinemia and other immunoglobulin deficiencies. *Immunol. Rev.*, **138**, 159–183.

Thomas,J.D., Sideras,P., Smith,C.I., Vorechovsky,I., Chapman,V. and Paul,W.E. (1993) Colocalization of X-linked agammaglobulinemia and X-linked immunodeficiency genes. *Science*, **261**, 355–358.

Touhara,K., Inglese,J., Pitcher,J.A., Shaw,G. and Lefkowitz,R.J. (1994) Binding of G protein βγ-subunits to pleckstrin homology domains. *J. Biol. Chem.*, **269**, 10217–10220.

Tronrud,D.E., Ten Eyck,L.F. and Matthews,B.W. (1987) An efficient general-purpose least-square refinement program for macromolecular structures. *Acta Crystallogr.*, **A43**, 489–501.

Tsukada,S. *et al*. (1993) Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia. *Cell*, **72**, 279–290.

Tsukada,S., Simon,M.I., Witte,O. and Katz,A. (1994) Binding of βγ subunits of trimeric G proteins to the PH domain of Bruton tyrosine kinase. *Proc. Natl Acad. Sci. USA*, **91**, 11256–11260.

Vetrie,D. *et al*. (1993) The gene involved in X-linked agammaglobulinemia is a member of the src family of protein-tyrosine kinases. *Nature*, **361**, 226–233.

Vihinen,M., Nilsson,L. and Smith,C.I. (1994) Tec homology (TH) adjacent to the PH domain. *FEBS Lett.*, **350**, 263–265.

Vihinen,M., Zvelebil,M.J., Zhu,Q., Brooimans,R.A., Ochs,H.D., Zegers,B.J., Nilsson,L., Waterfield,M.D. and Smith,C.I. (1995) Structural basis for pleckstrin homology domain mutations in X-linked agammaglobulinemia. *Biochemistry*, **34**, 1475–1481.

Vihinen,M., Iwata,T., Kinnon,C., Kwan,S.P., Ochs,H.D., Vorechovsky,I. and Smith,C.I. (1996) BTKbase, mutation database for X-linked agammaglobulinemia (XLA). *Nucleic Acids Res.*, **24**, 160–165.

Vriend,G. (1990) WHAT IF—a molecular modelling and drug design program. *J. Mol. Graphics*, **8**, 52–56.

Yao,L., Kawakami,Y. and Kawakami,T. (1994) The pleckstrin homology domain of Bruton tyrosine kinase interacts with protein kinase C. *Proc. Natl Acad. Sci. USA*, **91**, 9175–9179.

Yoon,H.S., Hajduk,P.J., Petros,A.M., Olejniczak,E.T., Meadows,R.P. and Fesik,S.W. (1994) Solution structure of a pleckstrin-homology domain. *Nature*, **369**, 672–675.

Received on January 21, 1997; revised on February 25, 1997