Journal of Biomolecular NMR (2006) 36: 73–78
DOI 10.1007/s10858-006-9064-3

NMR Structure Note

Solution structure and phosphopeptide binding of the SH2 domain from the human Bruton's tyrosine kinase

Kuo-Chun Huang, Hsi-Tsung Cheng, Ming-Tao Pai, Shiou-Ru Tzeng & Jya-Wei Cheng*
Institute of Biotechnology and Department of Life Science, National Tsing Hua University, Hsinchu 300, Taiwan, R.O.C

Received 10 April 2006; Accepted 12 July 2006

Biological context

X-linked agammaglobulinemia (XLA) is caused by mutations in the Bruton's tyrosine kinase (Btk) (Vetrie et al., 1993). The kinase is expressed in most haematopoietic cells, but is selectively down regulated in plasma cells and T-lymphocytes. Mutations or deletions in the Btk gene were detected in unrelated XLA patients (Tsukada et al., 1993), and it strongly suggests that the kinase is directly involved in the disease and, therefore, in the process of B cell development.

Btk, along with Tec, Itk and Bmx, belongs to a small family of tyrosine kinases (the Tec family) which share common structural features (Vetrie et al., 1993). *In vivo* and *in vitro* studies have shown that Btk can bind to the B-cell linker protein, the protein product of the *c-cbl* protooncogene, the βγ subunits of heterotrimeric G proteins, Fyn, Lyn, Hck, and protein kinase C (Tsukada et al., 2001). The SH3 and SH2 domains are small protein modules that mediate protein–protein interactions and occur in many proteins involved in intracellular signal transduction (Pawson, 1995). It was shown that the Btk SH2 domain is essential for phospholipase C-γ phosphorylation (Tsukada et al., 2001) and mutations in this domain lead to XLA. Recently, the B-cell linker protein (BLNK, also called SLP-65) was found to interact with the SH2 domain of Btk and this association is required for the activation of phospholipase C-γ (Tsukada et al., 2001).

The binding specificity of SH2 domains is provided by the phosphotyrosine, and by the residues carboxyterminal to it (Songyang et al., 1993). Recently, it has been shown that Btk SH2 domain binds with phosphopeptides in the order pYEEI > pYDEP > pYMEM > pYLDL > pYIIP (Tzeng et al., 2000). In this paper, we have determined the solution structure of the Btk SH2 domain. In addition, we have identified the binding sites between human BLNK phosphopeptides and Btk SH2 domain by comparing NMR chemical shifts of the free and the peptide-bound forms. Our results provide a structural basis for understanding the role of the Btk SH2 domain in XLA.

Methods and results

The Btk cDNA encoding the SH2 domain (residues 270–387) was subcloned into the pET15b (Novagen) vector and expressed in *E. coli* BL21 (DE3). Uniformly ^{15}N or $^{15}N/^{13}C$ labeled proteins were purified from bacteria in M9 minimal medium supplemented with ^{13}C glucose, $^{15}NH_4Cl$ and trace vital vitamins. The samples were purified by Ni-NTA column (Qiagen) and then the fractions containing Btk-SH2 were pooled and concentrated by Amicon Concentrator (Millipore). Sample concentrations for NMR experiments were typically 0.6 to 2.0 mM in phosphate buffer (50 mM

*To whom correspondence should be addressed. E-mail: jwcheng@life.nthu.edu.tw

K_2HPO_4, 100 mM NaCl, 10 mM NaN_3, and 0.1 mM EDTA, pH 6.5) in either 90% H_2O/10% D_2O or 100% D_2O.

All NMR data were acquired with Bruker Avance 600 spectrometer equipped with triple-resonance probe at 25°C. The NMR experiments performed included 2D ^{1}H-^{15}N HSQC, ^{1}H-^{13}C HSQC, 3D ^{15}N-NOESY-HSQC, HNCO, HN(CA)CO, HN(CO)CA, HNCA, CBCA (CO)NH, and HNCACB for backbone assignments, ^{15}N-TOCSY-HSQC, HCC(CO)NH-TOCSY, HCCH-TOCSY, HCCH-COSY, HBHA(CO)NH for side chain assignments (Ferentz and Wagner, 2000). All spectra were processed with the program XWIN-NMR 2.6 and analyzed using Aurelia Amix 2.1.3 (Bruker).

NOE based distance restraints were collected from analysis of 3D ^{15}N-edited NOESY-HSQC and ^{13}C-edited NOESY-HSQC spectra recorded with mixing time of 150 and 80 ms, respectively. The ϕ and ψ angles were obtained based on $^3J_{HNH\alpha}$ coupling constants derived from HNHA experiment and predicted from TALOS (Cornilescu et al., 1999). Hydrogen bond restraints were included in calculations only if the amide protons were slowly exchanging and if the β-strand inter-strand NOE cross-peaks were observed. Each hydrogen bond was enforced by two distance restraints of 1.8–2.3 Å (amide proton to carbonyl oxygen) and 2.8–3.3 Å (amide nitrogen to carbonyl oxygen).

The structure calculations were carried out using X-PLOR 3.851 (Brunger, 1993) program on a SuSE Linux 7.3 PC equipped with an Intel Pentium 4 1.8 GHz CPU and 512 MB DDR-SDRAM. The structures were calculated and refined with *ab inito* simulated annealing protocol, "sa.inp", and SA refinement protocol, "refine.inp" in X-PLOR 3.851. All force constants and molecular parameters, etc. were set to their default values, as in the original sa.inp and refine.inp protocols, with the exception of the timestep (ps), which was decreased to 0.001 ps throughout the calculations. Simulated annealing was performed using 9,000 steps at 1000 K and 10,000 steps by gradually cooling to 100 K. Finally, these structures were energy minimized using 500 steps of Powell energy minimization. The best 20 lowest energy structures were further analyzed with MOLMOL (Koradi et al., 1996) and PROCHECK-NMR (Laskowski et al., 1996).

Table 1. Summary of structural constrains and structure statistics

NOE restraints	1108
Intraresidues ($\|i-j\| = 0$)	498
Sequential ($\|i-j\| = 1$)	274
Medium range ($2 \leqq \|\{i-j\| \leqq 4$)	92
Long range ($\|i-j\| > 4$)	192
Hydrogen bond constraints	52
Dihedral angles[a]	106
Final energy (kcal mol^{-1})	
E (total)	275.5479 ± 13.0389
E (bond)	9.398048 ± 0.8718
E (angle)	170.7720 ± 3.61895
E (improper)	30.34438 ± 1.17958
E (van der Waals)	25.83243 ± 6.35718
E (NOE)	37.93673 ± 7.82073
E (cdih)	1.264251 ± 0.40831
Structural Statistics (15 Structures)	
NOE violations, number > 0.3 Å	0
Dihedral angel violations, number > 5°	0
RMSD for geometrical analysis	
Bond lengths (Å)	0.002 ± 0.0001
Bond angles (degree)	0.576 ± 0.0061
Impropers (degree)	0.448 ± 0.0088
Atomic RMSD for protein[b]	
All heavy atoms	1.30 ± 0.15
Backbone	0.64 ± 0.18
Ramachandran statistics	
Most favoured region (%)	72.4
Additionally allowed (%)	25.4
Generously allowed (%)	1.9
Disallowed (%)	0.3

[a] Dihedral angles were predicted from the program TALOS.
[b] For residues 279–362 and 374–379.

The backbone ^{1}H, ^{15}N, ^{13}C, and ^{13}CO assignments of the 118 residue Btk SH2 domain have been published previously (Pai et al., 2002). A total of 1108 NOE-derived distance constraints, including 498 intraresidue, 274 sequential, 92 medium ($1 < i-j < 5$), 192 long-range ($i-j \geq 5$) and 52 hydrogen bond distance restraints in conjunction with 106 backbone dihedral angles are used in the structure calculations. The energetic and structural statistics are listed in Table 1. The rmsd calculated from the averaged coordinates for the Btk SH2 domain is 0.64 Å for the backbone heavy atoms (N, C, and C_α) and 1.30 Å for all heavy atoms. Analysis of ϕ and ψ backbone angles is facilitated by using the program

Procheck_NMR (Laskowski et al., 1996). Only 0.3% of the residues are in disallowed regions and 98% of the residues are within most favored or additional allowed regions of the Ramachandran plot. The Btk SH2 domain forms three long anti-parallel β-strands, flanked by a few smaller β-strands and an α-helix at each end which resembles the mixed α/β topology of typical SH2 domains (Figure 1A, B) (Kuriyan and Cowburn, 1993). The coordinates of both the representative structure and the family of structures have been deposited at the Brookhaven Protein Data Bank (access number 2GE9).

Discussion and conclusions

Although the global fold of the Btk SH2 domain is similar to that of the Itk SH2 domain of the Tec family, there are localized differences between these two proteins. For example, formation of a *cis-trans* isomerization of a single prolyl imide bond within the CD loop has been observed in the Itk SH2 domain (Mallis et al., 2002; Pletneva et al., 2006). The CD loop of the *cis* conformer of the Itk SH2 domain was suggested to mimic a polyproline ligand binding to the SH3-binding pocket and to be responsible for the SH3-SH2

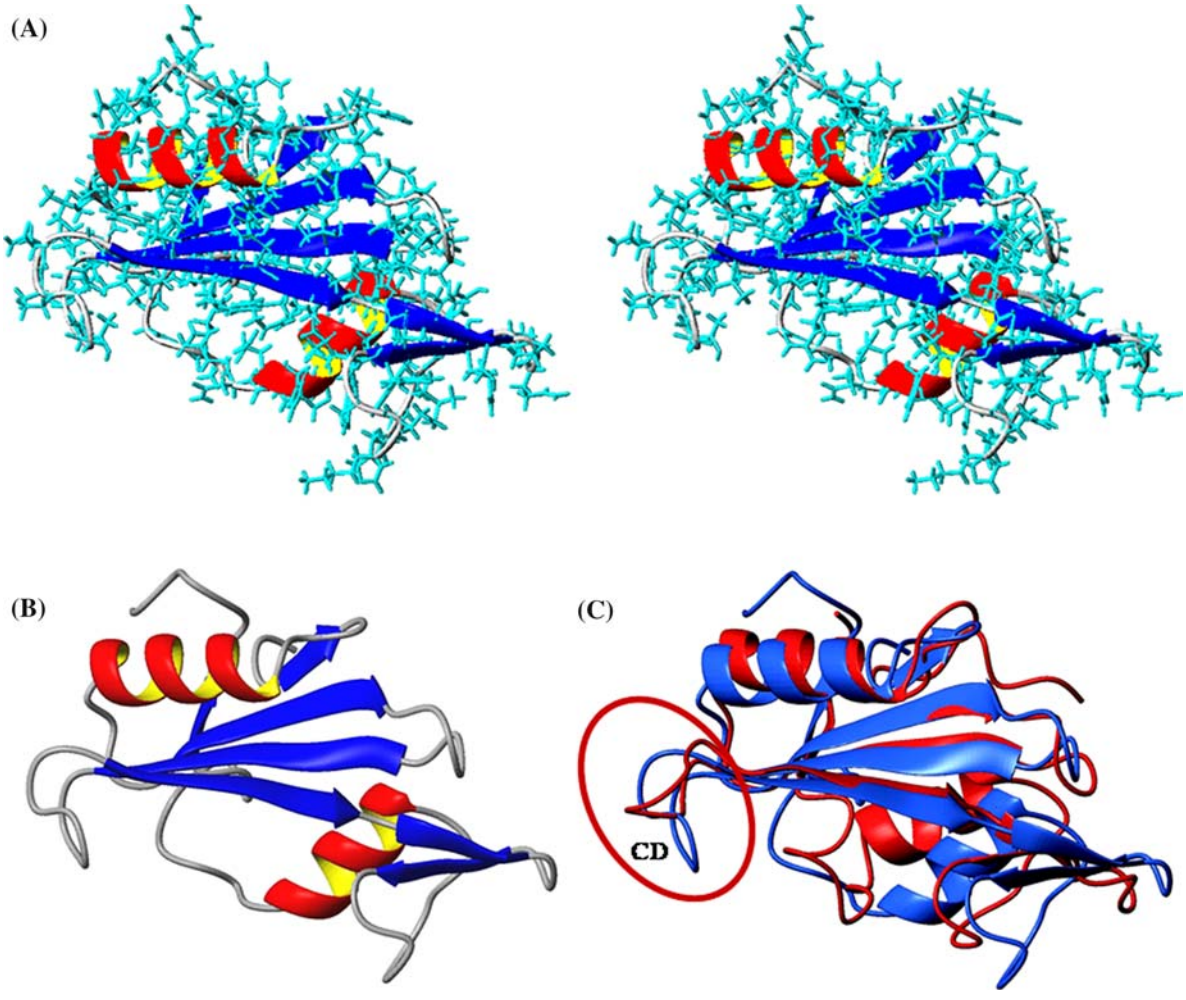

Figure 1. (A) A stereoview of the solution structure of Btk SH2 domain. (B) Ribbon diagram of the Btk SH2 domain indicating the secondary structure elements. (C) Superposition of the structures of the Btk SH2 domain (blue) and the Itk SH2 domain (PDB code 1LUK; red). The red circle indicates the different conformation of the CD loops between the Btk SH2 domain (*trans*-proline) and the Itk SH2 domain (*cis*-proline).

domain-domain interactions within Itk (Mallis et al., 2002). In contrast, the *trans* conformer of the Itk SH2 domain is not likely to fit into the SH3-binding pocket (Mallis et al., 2002). In addition, recent studies indicate that the proline-dependent conformational switch within the CD loop of Itk SH2 domain regulates substrate recognition and mediates regulatory interactions with cyclophilin A (CypA). CypA and Itk form a stable complex in Jurkat T cells and the phosphorylation levels of Itk and a downstream substrate, PLCγ1, are affected by the complex formation (Brazin et al., 2002). In our present studies, the Btk SH2 domain which also contains a Pro residue in the CD loop, only adopts a single *trans* conformation in solution (Figure 1C). Moreover, there is no intra- or inter-molecular association through the SH3 and SH2 domain–domain interactions within Btk (K.C.H. and J.W.C., unpublished results). Our results confirm the suggestion that the position of a Pro residue within the SH2 domain is important in determining the conformational flexibility of the prolyl imide bond and hence affects its structure and function (Mallis et al., 2002). We have also used reduced spectral density functions to analyze the NMR relaxation rates and ^{1}H-^{15}N heteronuclear NOEs of the Btk SH2 domain (Farrow et al., 1995). Our results indicate that the loop regions, especially the αB-βG loop, undergo picosecond fast motions as evident from the larger J(H) values. However, more ordered conformations were found in the Itk SH2 domain (Mallis et al., 2002).

It has been shown that Btk SH2 domain binds with phosphopeptides in the order pYEEI > pYDEP > pYMEM > pYLDL > pYIIP (Tzeng et al., 2000). Based on these results and the amino acid sequences of BLNK, the most possible Btk SH2 domain binding sites in BLNK are predicted to be pYENP (Y72) and pYEPP (Y96). We have synthesized two phosphotyrosine-containing peptides, DFDSDpYENPDEHSD (Y72) and

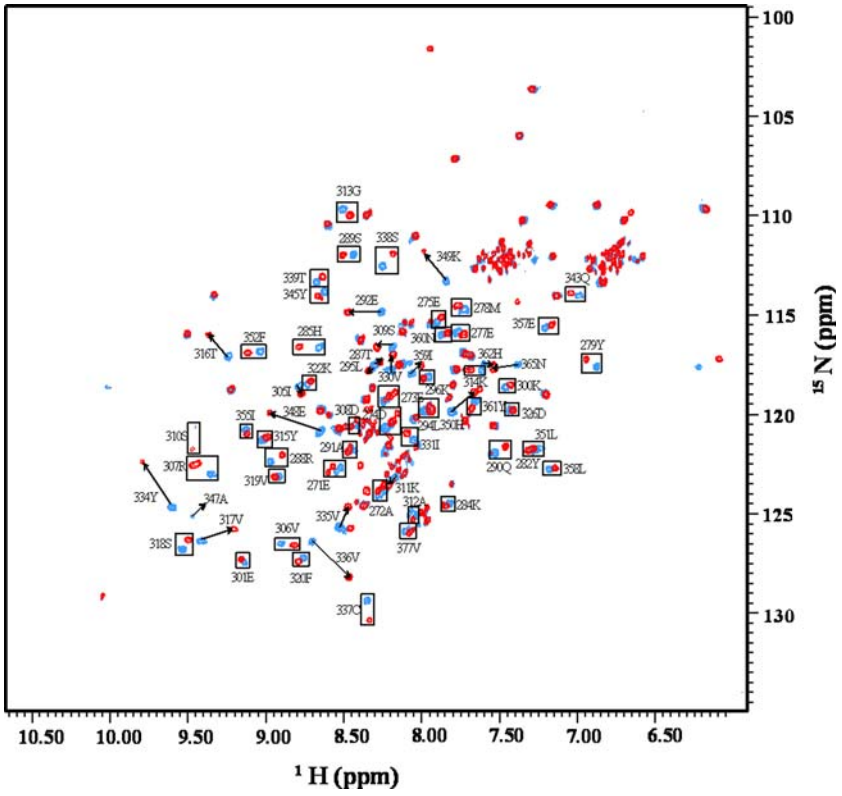

Figure 2. Overlay of the ^{1}H-^{15}N HSQC spectrum of the Btk SH2 domain (blue) and the ^{1}H-^{15}N HSQC spectrum of the Btk SH2 domain/Y96 phosphopeptide complex (red). Cross-peaks showing the shifts are labeled according to the residue types and numbers.

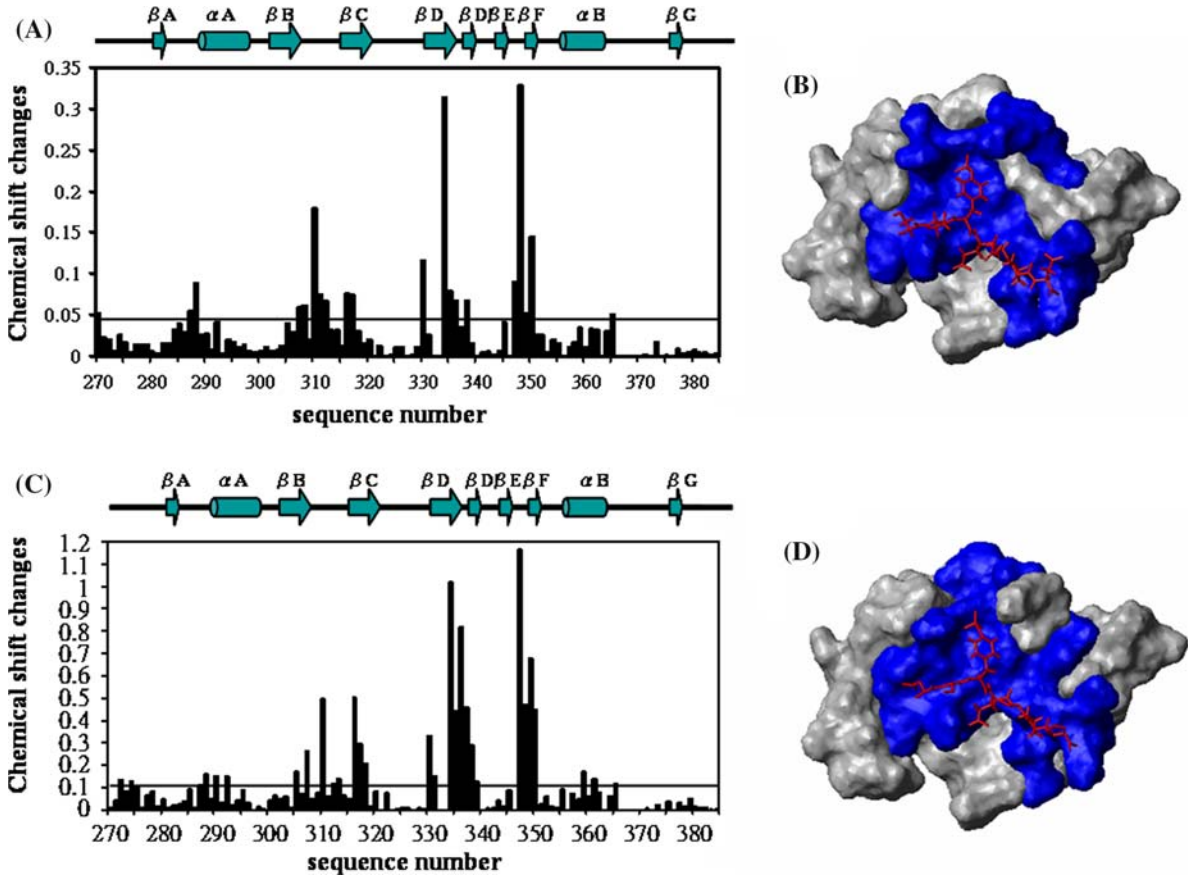

Figure 3. (A) Chemical shift change versus residue number of the Btk SH2 domain upon Y72 binding. The chemical shift changes are calculated according to the equation: $\Delta\delta = [[\Delta\delta(^{1}H)^{2}] + [[0.2 \times \Delta\delta(^{15}N)]^{2}]^{1/2}$. (B) Surface rendering of the Btk SH2 domain bound to Y72. The blue shading corresponds to the residues that give rise to a change in $\Delta\delta$ chemical shift great than 0.05 ppm. The peptide is shown as a stick model in red. (C) Chemical shift change versus residue number of the Btk SH2 domain upon Y96 binding. (D) Surface rendering of the Btk SH2 domain bound to Y96. The blue shading corresponds to the residues that give rise to a change in $\Delta\delta$ chemical shift great than 0.1 ppm. The peptide is shown as a stick model in red.

NADDSpYEPPPVEQE (Y96), according to the specific recognition sequences of BLNK. Comparison of the ^{1}H-^{15}N cross-peaks in the free and the phosphopeptide bound forms of Btk SH2 domain (Figure 2) indicates that complex formation causes significant chemical shift changes for a discrete set of resonances (Figure 3A, C). The residues with significant chemical shift changes upon binding with Y72 and Y96 are mapped onto the structure of the Btk SH2 domain as shown in Figure 3(B, D). In addition, we have found that the Y96 phosphopeptide induces larger chemical shift changes than the Y72 phosphopeptide does (Figures 3A, C). However, both the Y96 and the Y72 phosphopeptides have similar calculated binding affinities ($\sim$1 μM). Previous studies have shown that the K_D values for binding of various phosphopeptides to SH2 domain are within approximately an order of magnitude indicating that binding specificity of SH2 domains is limited (Ladbury and Arold, 2000). Thus, the binding specificity between Btk and BLNK must not only come from the Btk SH2 domain alone. Interestingly, there is also a putative SH3 domain binding site in both BLNK (residues 138–149) and SLP-76 (residues 184–195) which is adjacent to the possible Itk and Btk SH2 domain binding site, pY145 in SLP-76 and pY96 in BLNK. The presence of adjacent binding sites for Btk SH3 and SH2 domains suggested that BLNK could bind synergistically to the Btk SH3 and SH2 domains and give a higher level of specificity due to the

favourable free energies of binding and the relative positions of the domains (Bunnell et al., 2000). To further understand the mechanism of such an interaction, NMR structural and ligand binding studies between Btk SH3-SH2 domain and BLNK are currently undergoing in our laboratory.

Mutations in the Btk SH2 domain were found in several patient families with XLA (Conly et al., 1998). Previously, we expressed the R288Q, R288W, R307G, R307T, Y334S, Y361C, L369F, and I370M mutants of the Btk SH2 domain identified from XLA patients and measured their binding affinity with the phosphopeptides (Tzeng et al., 2000). Our studies revealed that mutations of R288, R307, Y334, Y361, L369, and I370 resulted in a 3 to 200 fold decrease in the peptide binding. Based on the present structure of the Btk SH2 domain and the NMR chemical shift mapped binding sites, the R288 and R307 mutations are involved in the phosphotyrosine binding and the Y334, Y361, L369 and I370 mutations are involved in the hydrophobic binding pocket at the pY + 3 position (Bradshaw et al., 1999). Previously, the three-dimensional structure of Btk SH2 domain was modeled based on v-Src (Vihinen et al., 1994). The model was used to predict the pY + 1 and pY + 3 binding sites within Btk SH2 domain. There are some localized differences between the predicted structural model and the present NMR studies. For example, residues L346 and A347 were predicted to be involved in the pY + 3 binding sites but no NMR chemical shifts were observed. In stead, residues E348, K349, and H350 were found to have significant shifts. Nevertheless, most of the predicted binding residues are within the NMR chemical shift mapped binding sites.

In summary, we have determined the solution structure and the phosphopeptide binding sites of the Btk SH2 domain. We have also found that mutation sites of the Btk SH2 domain identified from XLA patients are involved in the phosphopeptide binding. It is likely that the point-mutated Btk SH2 domains fail to present to the ligand the crucial residue in the correct context, thus leading to a weaker binding. The altered binding behavior likely affects the kinase function, and possibly causes XLA.

Acknowledgments

We thank Dr. Shi-Han Chen, Dr. Feng-Di Lung for helpful discussions. This work is supported by grants from the National Science Council, ROC (NSC95-2311-B-007-003 and NSC94-2311-B-007-020).

References

Bradshaw, J.M., Mitaxov, V. and Waksman, G. (1999) *J. Mol. Biol.*, **293**, 971–985.
Brazin, K.N., Mallis, R.J., Fulton, D.B. and Andreotti, A.H. (2002) *Proc. Natl. Acad. Sci.*, **99**, 1899–1904.
Brunger, A.T. (1993) *X-PLOR Version 3.1: A system for X-Ray Crystallography and NMR* Yale University Press, New Haven.
Bunnell, S.C., Diehn, M., Yaffe, M.B., Findell, P.R., Cantley, L.C. and Berg, L.J. (2000) *J. Biol. Chem.*, **275**, 2219–2230.
Conly, M.E., Mathias, D., Treadaway, J., Minegishi, Y. and Rohrer, J. (1998) *Am. J. Hum. Genet.*, **62**, 1034–1043.
Cornilescu, G., Delaglio, F. and Bax, A. (1999) *J. Biomol. NMR*, **13**, 289–302.
Farrow, N.A., Zhang, O., Szabo, A., Torchia, D.A. and Kay, L.E. (1995) *J. Biomol. NMR*, **6**, 153–162.
Ferentz, A.E. and Wagner, G. (2000) *Q. Rev. Biophys.* **33**, 29–65.
Koradi, R., Billeter, M. and Wuthrich, K. (1996) *J. Mol. Graph.*, **14**, 51–55.
Kuriyan, J. and Cowburn, D. (1993) *Curr. Opin. Struc. Biol.*, **3**, 828–837.
Ladbury, J.E. and Arold, S. (2000) *Chem. Biol.*, **7**, 3–8.
Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) *J. Biomol. NMR*, **8**, 477–486.
Mallis, R.J., Brazin, K.N., Fulton, D.B. and Andreotti, A.H. (2002) *Nat. Struct. Biol.*, **9**, 900–905.
Pai, M.T., Huang, K.C., Tzeng, S.R. and Cheng, J.W. (2002) *J. Biomol. NMR*, **24**, 163–164.
Pawson, T. (1995) *Nature*, **373**, 573–580.
Pletneva, E.V., Sundd, M., Fulton, D.B. and Andreotti, A.H. (2006) *J. Mol. Biol.*, **357**, 550–561.
Songyang, Z., Shoelson, S.E., Chaudhari, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., Neel, B.G., Birge, R.B., Fajardo, E.J., Chou, C.M., Hanafusa, H., Schaffhausen, B. and Cantley, L.C. (1993) *Cell*, **72**, 767–778.
Tsukada, S., Baba, Y. and Watanabe, D. (2001) *Adv. Immunol.*, **77**, 123–162.
Tsukada, S., Saffran, D.C., Rawlings, D.J., Parolini, O., Allen, R.C., Klisak, I., Sparkes, R.S., Kubagawa, H., Mohandas, T. and Quan, S. (1993) *Cell*, **72**, 279–290.
Tzeng, S.R., Pai, M.T., Lung, F.D., Wu, C.W., Roller, P.P., Lei, B., Wei, C.J., Tu, S.C., Chen, S.H., Soong, W.J. and Cheng, J.W. (2000) *Protein Sci.*, **9**, 2377–2385.
Vetrie, D., Vorechovsky, I., Sideras, P., Holland, J., Davies, A., Flinter, F., Hammarstrom, L., Kinnon, C., Levinsky, R., Bobrow, M., Smith, C.I.E. and Bentley, D.R. (1993) *Nature*, **361**, 226–233.
Vihinen, M., Nilsson, L. and Smith, C.I.E. (1994) *Biochem. Biophys. Res. Comm.*, **205**, 1270–1277.