Crystal Structure and NMR Studies of the Apo SH2 Domains of ZAP-70: Two Bikes Rather than a Tandem‡

Rutger H. A. Folmer, Stefan Geschwindner, and Yafeng Xue*

Structural Chemistry Laboratory, AstraZeneca R&D Mölndal, S-431 83 Mölndal, Sweden

Received July 17, 2002; Revised Manuscript Received September 17, 2002

ABSTRACT: The protein kinase ZAP-70 is involved in T-cell activation, and interacts with tyrosine-phosphorylated peptide sequences known as immunoreceptor tyrosine activation motifs (ITAMs), which are present in three of the subunits of the T-cell receptor. We have studied the tandem SH2 (tSH2) domains of ZAP-70, by both X-ray and NMR. Here, we present the crystal structure of the apoprotein, i.e., the tSH2 domain in the absence of ITAM. Comparison with the previously reported complex structure reveals that binding to the ITAM peptide induces surprisingly large movements between the two SH2 domains and within the actual binding sites. The conformation of the ITAM-free protein is partly governed by a hydrophobic cluster between the linker region and the C-terminal SH2 domain. Our data suggest that the two SH2 domains are able to undergo large interdomain movements. The proposed relative flexibility of the SH2 domains is further supported by the finding that no NMR signals could be detected for the two helices connecting the SH2 domains; these are likely to be broadened beyond detection due to conformational exchange. It is likely that this conformational reorientation induced by ITAM binding is the main signaling event activating the kinase domain in ZAP-70. Another NMR observation was that the N-terminal SH2 domain could bind tetrapeptides derived from the ITAM sequence, apparently without the need to interact with the C-terminal domain. In contrast, the C-terminal domain has little affinity for tetrapeptides. The opposite situation is true for binding to plain phosphotyrosine, where the C-terminal domain has a higher affinity. Distinct features in the crystal structure, showing the interdependence of both domains, explain these binding data.

The ZAP-70 protein is one of the two members of the Syk family of nonreceptor protein tyrosine kinases (PTKs).[1] This 70 kDa protein contains two SH2 domains, arranged in tandem and separated by a linker region, and a C-terminal kinase domain. The tandem SH2 domain is connected to the kinase domain through two interdomains, denoted A and B. The protein is expressed in T-cells, where it associates with the ζ subunit of the T-cell receptor (TCR) in the early steps of T-cell activation. Association occurs through the tandem SH2 domain (tSH2) binding to the immunoreceptor tyrosine activation motifs [ITAMs, defined by the sequence $YXX(L/I)X_{6-8}YXX(L/I)$ (*1*, *2*)], after dual phosphorylation of the tyrosines by the regulatory kinase $p56^{lck}$ (*3*). The latter kinase is also implicated in ZAP-70 activation, through phosphorylation of Tyr493 on the catalytic domain (*4*). This event, in turn, leads to ZAP-70-mediated activation of its downstream targets in the T-lymphocyte signal transduction pathway. Similar mechanisms occur in B-cell receptor complexes, where src kinases phosphorylate ITAMs for Syk recruitment. However, Syk binding to the ITAM sequences may not require the involvement of both SH2 domains (*5*). Besides Tyr493, additional tyrosines within ZAP-70 undergo phosphorylation (e.g., Tyr292, -315, and -319), playing both positive and negative regulatory roles in TCR function (*6*, *7*). Tyr315 has furthermore been implicated in allosterically regulating the function of the SH2 domains (*8*).

While Syk is more ubiquitously expressed among hematopoietic cells, ZAP-70 is restricted to T cells. As such, ZAP-70 is an attractive target for pharmaceutical research; inhibitors of T-cell activation are of potential use for treatment of autoimmune diseases, inflammation, and transplant rejection. Studies on mice deficient in ZAP-70 indicate that there is an absolute requirement of ZAP-70 for αβ T-cells and epithelial γδ T-cells (*9*). In other tissues, though, Syk can partially compensate for the loss of ZAP-70 (*9*, *10*). Several studies on humans with disrupted ZAP-70 have shown that T-cell function is absent, leading to a severely compromised immune response (*11*−*13*). Several examples of ZAP-70 inhibitors have been published (see refs *14* and *15* for reviews).

The structures of both Syk (*16*) and ZAP-70 (*17*) in complex with dually phosphorylated ITAM peptides have been published. As expected on the basis of the overlapping roles the two kinases play in T-cell biology, their structures

‡ The coordinates of the tandem SH2 domain of ZAP-70 have been deposited in the RCSB Protein Data Bank as entry 1M61.

* To whom correspondence should be addressed. E-mail: Yafeng.Xue@astrazeneca.com. Telephone: +46-31-7762085. Fax: +46-31-7763792.

[1] Abbreviations: ASA, solvent accessible surface area; ITAM, immunoreceptor tyrosine activation motif; ITC, isothermal titration calorimetry; SH2, src homology domain 2; tSH2, tandem SH2 domain; ITAM−tSH2, tandem SH2 domain complexed to the ITAM peptide; MR, molecular replacement; rmsd, root-mean-square difference; PTK, protein tyrosine kinase; TCR, T-cell receptor; TROSY, transversal relaxation optimized spectroscopy; ZAP-70, 70 kDa ζ-associated protein; ZAP-N and ZAP-C, N- and C-terminal SH2 domains, respectively, of ZAP-70.

10.1021/bi026465e CCC: $22.00
Published on Web 11/08/2002

are fairly similar. An important difference, however, is that the N-terminal SH2 domain of ZAP-70 is rather unusual in that its phosphotyrosine (pTyr) pocket is formed by residues of both the C- and N-terminal domains. The SH2 domains in Syk are similar to each other and to other SH2 domains. It has therefore been concluded that, in contrast to the ZAP-70 protein, the domains in Syk can act independently.

So far, no structures have been reported of tSH2 domains of the Syk family of PTKs in the absence of substrate. Here, we present the crystal structure of the apo tSH2 domain of ZAP-70. In addition, we have studied substrate binding to the protein employing NMR, using mono- and diphosphorylated peptides derived from the ITAM sequence. Among others, these studies give new insight into the functionality and interdependency of the two SH2 domains, and on the overall mobility within the tandem domain.

MATERIALS AND METHODS

Protein Expression, Purification, and Characterization. The cDNA encoding the two SH2 domains (residues 1−256) of ZAP-70 was cloned into the pET28a expression vector (Novagen) at the *Nde*I and *Not*I sites, resulting in a His-tagged protein at the N-terminus. After the protein had been expressed in *Escherichia coli* strain BL21(DE3) at 22 °C, the protein was affinity-purified from the soluble fraction using Ni^{2+}−NTA agarose (Qiagen). After the His tag had been released by treatment with thrombin, the protein was finally purified via gel filtration on Superdex 75pg (Pharmacia) and concentrated by ultrafiltration (Centriprep 10, Amicon) to 20 mg/mL in 100 mM sodium phosphate buffer (pH 6.80), 50 mM NaCl, and 5 mM DTT.

NMR Studies. ^{15}N-labeled as well as triply ^{2}H-, ^{13}C-, and ^{15}N-labeled protein samples were generated by growing the expressing cells in Celtone medium (Martek Biosciences) enriched with these stable isotopes. All NMR experiments were performed at 296 K on a Bruker Avance 600 or 800 MHz system, equipped with triple-resonance (^{1}H, ^{13}C, and ^{15}N) single-gradient 5 mm probes. Samples were at pH 6.8 in a 100 mM phosphate buffer, containing 50 mM NaCl, 10 mM DTT, and 1 mM EDTA. Data sets were processed with nmrPipe (*18*). Analysis of the triple-resonance spectra was carried out with Xeasy (*19*), while NMRview (*20*) was used to follow titrations in HSQC spectra.

All triple-resonance experiments related to the sequential assignment were acquired on a 300 μL sample of 1.4 mM triply labeled ZAP-70 tSH2 (>98% ^{2}H). Initially, a set of conventional HNCO, HN(CA)CO, HNCA, HN(CO)CA, HN(CA)CB, and HN(COCA)CB experiments was carried out at 600 MHz and used for the assignment. Later, these experiments were repeated using TROSY versions (*21*, *22*), acquired at 800 MHz. The average measuring time for each of these experiments was 4 days.

Binding to peptides was studied using ^{15}N-labeled protein at concentrations between 0.2 and 0.5 mM, mainly by recording HSQC spectra. Typical evolution times were around 80 ms in both dimensions, with acquisition times between 30 min and 1 h.

Crystallization and Data Collection. The vapor diffusion method was used for crystallization. The reservoir was composed of (total of 0.5 mL) 20% PEG3350, 0.1 M HEPES (pH 8), 0.2 M NaCl, and 20 mM glutathione. Another solution with 15% PEG3350 but otherwise identical composition was used to mix the drop; 2 μL of this solution was mixed with 1.5 μL of the protein solution. Repeated microseeding was necessary to grow single crystals, and we needed fresh crystals (<5 days old) to obtain the best diffracting power. We believe these difficulties are related to the nine free cysteines in the protein, which requires fresh DTT so the solutions are kept reduced. The crystal has a unit cell of 49 Å × 52 Å × 55 Å, belonging to monoclinic space group $P21$, with a β of 89.9°.

A complete data set was collected at beamline 711 at the MAX lab in Lund, Sweden. The data were collected on a MAR345 detector at a crystal−detector distance of 175 mm, with 1° oscillation (wavelength of 1.0158 Å). The programs mosflm and scala (*23*) were used for data processing and scaling, respectively. In total, 39 581 reflections were accepted, corresponding to 9426 unique reflections, to a 2.5 Å resolution with an R_{merge} of 0.062 and an overall redundancy of 4.2 [0.456 and 4.2, respectively, in the outer shell (2.64−2.50 Å)].

Structure Determination. The published structure of the Syk tandem SH2 domains in complex with the ITAM peptide (*16*) (PDB entry 1a81) was used as a search model for molecular replacement (MR). The model was first trimmed to remove the flexible linker region between the two SH2 domains, based on visual inspection of the model using the graphics program O (*24*). After this deletion, the two remaining domains comprised residues 14−115 and 174−260. These two models were used to search for the individual SH2 domains in separate MR runs, and the solutions were brought into the same origin with a translation search for one, while fixing the other. The program AmoRe (*25*) was used to perform the MR.

Both rotation and translation were carried out with data of 15−3.3 Å resolution, followed by rigid body fitting using data between 15 and 2.8 Å. Before rigid body fitting, a checking for packing interaction of each solution was performed to filter out unreasonable solutions. In the MR run with the N-terminal SH2 domain (ZAP-N for short), the correct solution was the top one from the rotation search (4.56σ, the second highest peak was 3.90σ). The translation search ended with a correlation coefficient of 35.6% and an R-factor of 52.5%, which became 51.6 and 50.7%, respectively, after rigid body fitting. Similarly, the run with the C-terminal SH2 domain (ZAP-C) resulted in a correlation coefficient of 50.4% and an R-factor of 53.3% after rigid body fitting (33.5 and 53.9%, respectively, from translation). But the correct solution here corresponded to the 12th-highest peak from the rotation search (2.92σ for the 12th peak vs 3.48σ for the top solution). A translation search using data of 15−4 Å resolution was performed by fixing the first solution to bring them into the frame with the same origin. This resulted in a correlation coefficient of 46.1% and an R-factor of 44.4% (55.6 and 43.6%, respectively, after rigid body fitting). Then 20 cycles of rigid body refinement were carried out with the program refmac (*23*) using data between 20 and 2.5 Å resolution. This reduced the R-factor to 38.8%. The final verification of the MR solution was based on the electron density map after simulated annealing refinement with the program CNS (*26*).

The programs refmac and CNS (later CNX) were used for refinement, while O (*24*) was used for model (re-)-

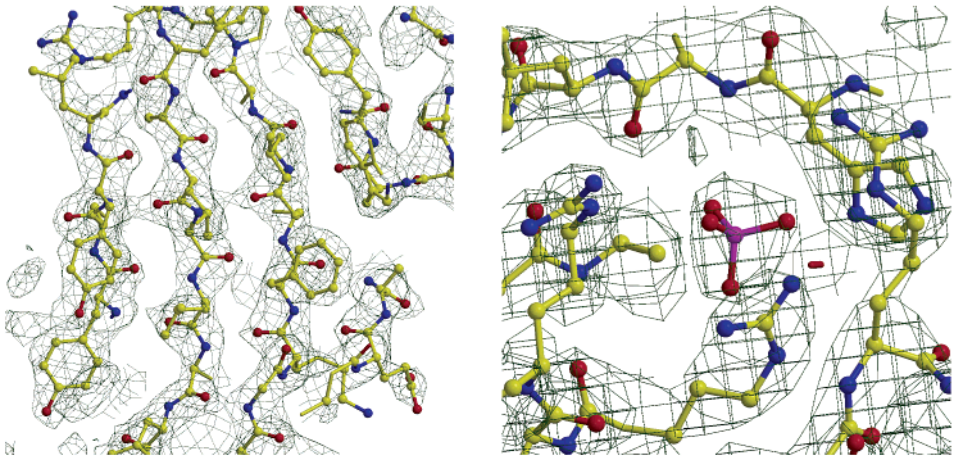

FIGURE 1: Electron density map ($2F_o - F_c$) contoured at 1σ, showing both an arbitrary part of the protein (the β-sheet in ZAP-C) and the phosphate molecule bound in the C-terminal phosphotyrosine pocket. This figure was generated using Molscript (*44*) and Raster3D (*45*).

Table 1: Data Collection and Refinement Statistics

data collection	
space group	$P2_1$
cell parameters (a, b, c, and β)	48.85 Å, 52.22 Å, 54.45 Å, and 89.91°
no. of molecules per asymmetric unit	1
solvent content (%)	47.4
resolution (last shell) (Å)	2.5 (2.63−2.5)
no. of unique reflections	9426
R_{merge} (%) (all/last shell)	6.2/45.6
$I/\sigma(I)$ (all/last shell)	10.2/1.7
completeness (%) (all/last shell)	98.1/97.8
redundancy	4.2
structure refinement	
resolution (last shell) (Å)	2.5 (2.50−2.61)
no. of reflections used (last shell, no cutoff)	9426 (1148)
fraction of test set for calculating R_{free} (%)	4
no. of reflections in the test set (last shell)	370 (47)
R_{work} (%)	24.6
R_{free} (last shell) (%)	28.8 (34.4)
rmsd for bond lengths (Å)	0.007
rmsd for bond angles (deg)	1.39
no. of protein atoms	2058
no. of waters	36
no. of ligand atoms (one phosphate group)	5

building. Repeated rebuilding was performed. For the linker region, several fragments were first built and then registered onto the sequence. Between rebuilding steps, a simulated annealed omit map was used to verify the model.

Solvent accessible surface areas (ASAs) (*27*, *28*) were calculated with the program AREAIMOL in the CCP4 package (*23*) using a 1.4 Å radius probe. The buried areas due to crystal packing were obtained directly from AREAIMOL, while the buried areas due to domain interaction were calculated from the following relation:

$$ASA_{buried} = ASA_{domain1} + ASA_{domain2} - ASA_{domain1+domain2}$$

To dissect the stabilization effect of buried ASA on individual domains, the structure was divided into three parts: ZAP-N, residues 1−111; linker domain, 112−162; and ZAP-C, 163−256. Since these domains are interconnected, such treatment will introduce a small overestimation of buried ASA around the connecting amino acid residues. However, the purpose of these calculations is to compare the domains in the ITAM-complexed (*29*) and uncomplexed

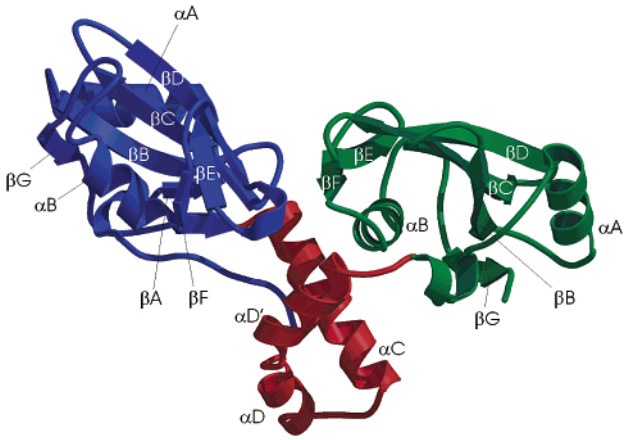

FIGURE 2: Ribbon diagram of the ZAP-70 tandem SH2 domain. The linker region is depicted in red, ZAP-N in blue, and ZAP-C in green. Secondary structure elements are labeled according to the notation of van Eck and co-workers (*32*). This figure was generated using Molscript (*44*) and Raster3D (*45*).

ZAP-70 structures; these differences will hardly be affected by this systematic overestimation.

RESULTS AND DISCUSSION

X-ray Structure Suggests Large Domain Reorientation upon ITAM Binding. The final refined model includes 2058 protein atoms, one phosphate group, and 36 solvent molecules. The current *R*-factor and R_{free} are 24.6 and 28.8%, respectively. The rmsds for bond lengths and bond angles are 0.007 Å and 1.39°, respectively. Figure 1 shows the electron density ($2F_o - F_c$, contoured at 1σ) calculated with the refined model. The overall quality of the current refined structure is good as judged by the electron density map and output from PROCHECK (*30*). More statistics are compiled in Table 1. The region linking the SH2 domains (residues 112−162) has a higher temperature factor than the individual domains (average *B* factor of 58.4 Å^2 vs a *B* factor of 49.3 Å^2), but the electron density is well-defined, except for some of the linker region side chain atoms (see also Figure 5).

Figure 2 shows a cartoon of the structure. The two SH2 domains feature the common fold observed in most SH2 domains, a central β-sheet flanked by two helices (*31*). The positioning of the secondary structure elements in both SH2 domains is very similar to that in the published structure

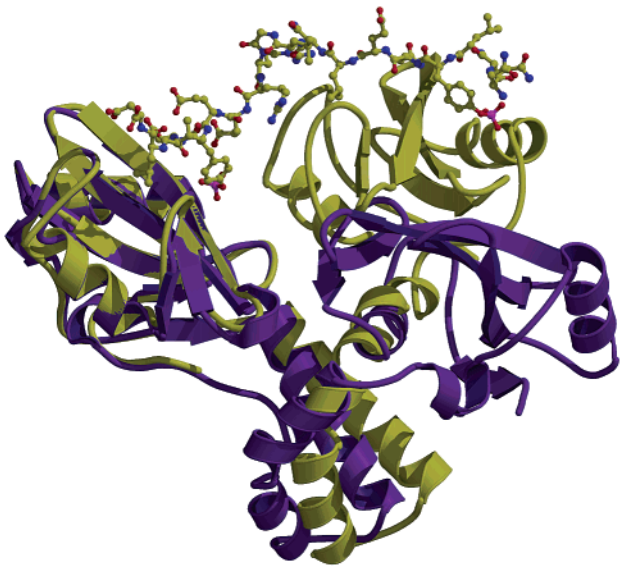

FIGURE 3: Comparison of the structures of the apo (purple) and ITAM-bound tSH2 domains of ZAP-70. The structures of the N-terminal domains (residues 9−105) were used to superimpose the molecules. The coordinates for the complex structure were published in PCT Application WO 97/08300 (*29*). This figure was generated using Molscript (*44*) and Raster3D (*45*).

(*17*) of the ITAM-complexed form of the protein (ITAM−tSH2 for short).

In contrast, as shown in Figure 3, the positioning of the domains themselves with respect to each other is dramatically different between the two forms of the protein. Consequently, the interfaces between the two SH2 domains are also very different. The Y-shaped structure of the complex has now become more L-shaped; the central β-sheet of ZAP-C has rotated ~50° from its position in the complex with the ITAM peptide, making an angle of ~110° with the central β-sheet of ZAP-N. The coiled-coil interaction of the inter-SH2 region in the complex has changed significantly, part of that interaction being replaced by helix−helix packing between the linker helix and helix αB from ZAP-C [secondary structure nomenclature of Eck and co-workers (*32*)]. Part of the coiled coil is unwound, with residues 155−157 being stretched to a random-coil conformation. Consequently, helix αB in ZAP-C no longer interacts with ZAP-N as it does in the ITAM-complexed structure. In the ITAM−tSH2 structure, Phe115 stacks with Trp233, whereas in the uncomplexed structure, Phe115 is close to a hydrophobic region that forms part of the phosphotyrosine pocket in the N-terminal domain of the ITAM−tSH2 complex. This hydrophobic interface includes Leu43 from ZAP-N, Val114, Phe115, and Leu118 from the linker, and Phe229, Tyr238, and Leu241 from the C-terminal SH2. This hydrophobic cluster likely dominates the interaction between the two SH2 domains in the uncomplexed structure.

When this large orientation difference between the N- and C-terminal domains in the apo and complex structures is considered, it is of interest to refer to the complex structure of the tSH2 domains of Syk reported by Fütterer et al. (*16*). These authors found six copies of the protein in the asymmetric unit, with a substantial variety in the relative orientation of the two SH2 domains. The two extreme orientations of ZAP-C, when keeping ZAP-N fixed, are related by a rotation of 18° and a translation of 2 Å. This shows that even with the ITAM peptide bound, remarkable flexibility may exist in the relative orientation of the two domains. Obviously, much larger mobility could be expected in the absence of the peptide, which is exactly what we observe for ZAP-70.

We performed a buried surface analysis to further analyze this large difference in orientation between the SH2 domains upon ITAM binding. Interestingly, ZAP-C has a markedly larger buried surface with the linker domain in the apo structure than in the ITAM complex (1525 vs 992 Å^2; see Table 2). This large buried surface in the ZAP-C linker interface in our structure is consistent with the aforementioned observation that extensive hydrophobic interactions were found between the ZAP-C and linker domains (notably, Val114, Phe115, and Leu118 from the linker and Phe229, Tyr238, and Leu241 from ZAP-C). This, in turn, is in line with the findings of Tsai et al. (*33*), who found that protein interfaces have much higher hydrophobicities than exposed surfaces. In the ZAP-70−ITAM complex, this lower-level interaction with the linker domain is compensated by a significant buried surface (2441 Å^2) of ZAP-C upon binding to the ITAM peptide. To reverse the argument, for apo ZAP-70 the lack of stabilization from peptide binding seems to have been compensated by an increase in buried surface between ZAP-C and the helical linker domain. The hydrophobic cluster holds the SH2 domains into a different conformation in the absence of the ITAM peptide, keeping the protein in the resting state where it cannot be activated. Only the large ITAM peptide has enough surface to offer to bring about the necessary conformational change that allows activation.

To investigate the role of crystal packing on the interdomain orientation, we compared buried surfaces due to packing to that due to domain interactions. It was found that the buried surface due to domain interactions (between ZAP-C and ZAP-N, and between the helical linker domain and both SH2 domains) is 3339 Å^2, which is ~20% of the total solvent accessible surface areas of the separate domains. This number matches quite well with studies by Tsai et al. (*33*), who analyzed a large number of stable protein interfaces and found that all the protein interfaces classified as stable constitute ~17% of the total accessible surface. As shown in Table 2, the buried surface at the ZAP-C−linker interface

Table 2: Buried Solvent Accessible Surface Area (Å^2)

	domain interface[a]					crystal packing[b]			
	total	ZAP-N and linker	linker and ZAP-C	ZAP-C and ZAP-N	ITAM binding	total	ZAP-N	linker	ZAP-C
apo	3339 (20%)	1188	1525	626		3587 (14)	1172 (10)	1018 (20)	1397 (15)
ITAM complex	4600 (25%)	486	992	681	2441	2908 (11)		860 (17)	

[a] Values in parentheses are the fractions of buried surface compared to total solvent accessible surface area. [b] Values in parentheses are the average buried surfaces per amino acid residue.

is 1525 Å^2, which is larger than the surface of ZAP-C buried in crystal packing (1397 Å^2). It is worth noting that the buried surface due to crystal packing (3587 Å^2) is comparable in size to that due to domain interactions (3339 Å^2). Finally, the ratio of buried surface to total accessible surface (13 047 Å^2) falls within the range commonly observed in protein crystals (*34*). All these calculations indicate that the packing interactions in our crystals are within normal ranges. Therefore, we believe that the observed orientation of the SH2 domains in the apo crystal is genuine and biologically relevant, and not caused by crystal packing.

Structural Changes in the C-Terminal Domain. Significant local differences between the free and bound state were also found within the C-terminal domain, whereas the N-terminal domain exhibited much less variation. Superimposition of the respective SH2 domains of the apo and complex structures shows an rms difference (Cα atoms) of only 0.35 Å for residues 9−105 in ZAP-N, while a much larger value of 0.83 Å is found for the corresponding residues in ZAP-C (163−256). Apparently, the N-terminal domain does not undergo large changes upon ITAM binding. In contrast, the C-terminal domain appears to adapt significantly to bind the peptide. The largest change is around residue Glu225 in the C-terminal SH2, the α-carbon atom of which differs by 3.3 Å in the free and bound positions. In particular, the main chain atoms around Glu225, a turn connecting βE and βF, have taken new conformations. This turn forms part of the leucine/isoleucine anchoring pocket for the ITAM peptide with the general pYXXI/L motif in the ITAM−tSH2 structure. This change in the conformation of Glu225 basically closes this pocket in the apoprotein. This is clearly seen from Figure 4, which shows a deep pocket for the leucine of the ligand in the complex structure. This pocket is almost completely absent in the structure of the apoprotein.

While the phosphotyrosine binding site of the N-terminal SH2 is empty in our structure, density was found at that site in ZAP-C. This density could only be interpreted as arising from a phosphate group, picked up from the crystallization medium.

ITAM Binding Stabilizes the Linker Region. Having access to an 800 MHz spectrometer, we anticipated that using TROSY-based triple-resonance experiments, it should be relatively straightforward to obtain a complete backbone assignment for this perdeuterated 259-amino acid construct. Nevertheless, we could unambiguously assign only 134 of the 242 backbone amide groups (55%). We were left with a fair number of peaks in the two-dimensional HSQC spectrum that simply did not give rise to cross-peaks in the more demanding three-dimensional experiments. These observations led us to conclude that the missing residues are subject to conformational exchange processes, to such an extent that their resonances are broadened beyond detection in most NMR experiments. This was supported by the fact that for these peaks, or amino acids, there was hardly any improvement in going from the conventional experiments to the TROSY versions recorded at a later stage. In contrast, significant signal-to-noise improvement was obtained for the resonances that had already been assigned in the conventional spectra. Clearly, the relaxation processes of the unassigned residues are dominated by exchange, which does not benefit from the relaxation optimization of TROSY (*35*).

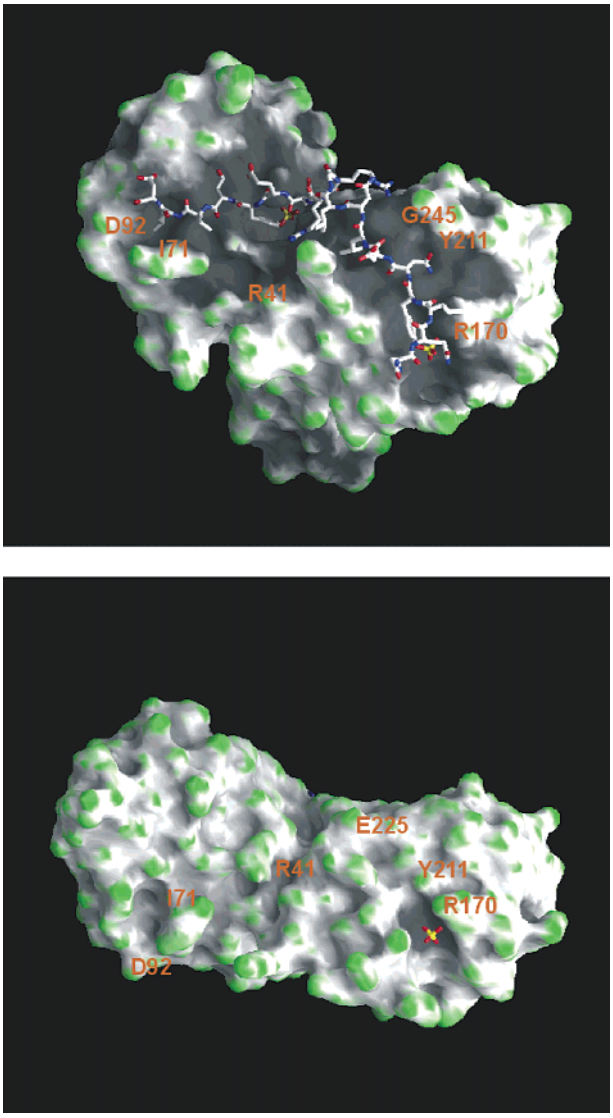

FIGURE 4: Comparison of the surfaces of the apoprotein (bottom) and ITAM-bound protein (top). Structures were oriented after superpositioning of the C-terminal domains; the N-terminal domains are hence in different orientations. Surfaces are colored according to curvature. Note how Glu225 in the apo structure nearly completely closes the deep pocket into which the leucine residue of the peptide protrudes in the complex. The figure was prepared using the program Grasp (*46*).

Figure 5 shows the distribution of assigned amino acids within the structure of the apoprotein. Strikingly, the assigned and nonassigned parts of the protein constitute truly distinct regions in the protein. Basically, the residues that could not be assigned constitute one single region of the protein: the two-helix coil and parts of the SH2 domains interacting with it. This can only lead to the conclusion that these helices are not completely stable in solution.

This conclusion is further supported by CD spectra we recorded in the absence and presence of the ITAM−peptide complex (Figure 6). The spectral changes at 210 and 196 nm upon addition of the ITAM−peptide complex are indicative of a gain in α-helical structure content.

While in solution there is significant flexibility in the connecting helices and in parts of ZAP-C, no direct evidence points to similar mobility in the crystal. Core regions for both SH2 domains are well-ordered on the basis of the electron density map and the distribution of temperature

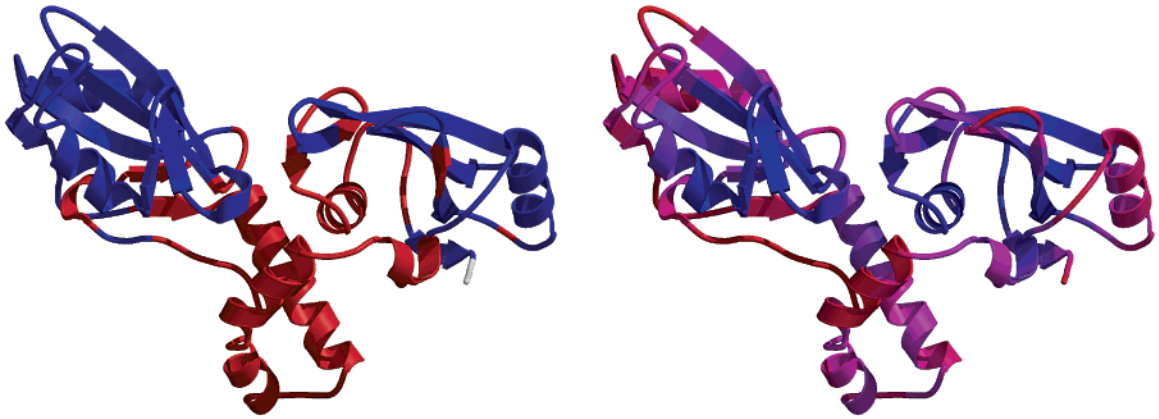

FIGURE 5: Ribbon diagrams of apo ZAP-70 tSH2, color-coded with respect to the presence of NMR assignments (left, blue for yes and red for no) and *B*-factor (right, color ramp from blue for low to red for high). This figure was generated using Molscript (*44*) and Raster3D (*45*).

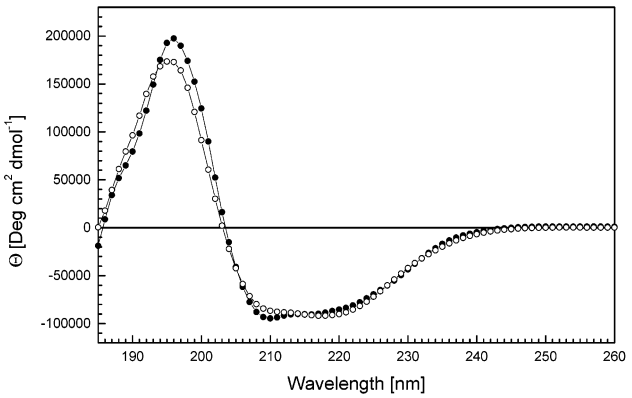

FIGURE 6: Far-UV circular dichroism spectra of the ZAP-70 tandem SH2 domain in the absence (○) and presence (●) of an equimolar amount of the ITAM−peptide complex. CD spectra of the ZAP-70 tandem SH2 domain [1.2 mg/mL in 20 mM sodium phosphate buffer (pH 6.8)] were recorded between 185 and 260 nm at 20 °C on a Jasco J-720 spectrometer equipped with a cell with a path length of 0.01 cm. The CD spectra are the averages over 16 scans and were converted into molar ellipticity.

factors. Buried solvent accessible surface areas were calculated to assess to what extent the stabilization of the observed conformation in the crystal is explained by crystal packing interactions. The various calculated buried surfaces are listed in Table 2. It shows that the helical linker domain is affected more by crystal packing than the other two domains as indicated by the larger buried surface due to crystal packing (an average of 20 Å^2 per amino acid versus 10 Å^2 for ZAP-N and 15 Å^2 for ZAP-C). The total buried surface of the linker due to packing in our apo structure is 1018 Å^2, which is also 18% larger than that in the structure of the complex (860 Å^2). Obviously, in solution the helices in the linker cannot pack to other structure elements, which explains their increased mobility as observed in the NMR spectra.

Figure 7A compares HSQC spectra of ZAP-70 tSH2 before and after addition of the diphosphorylated first ITAM motif of the TCR ζ subunit (ζITAM1, peptide used was LpYNELNLGRREEpYDVLD). Clearly, there are large differences between the spectra; only few peaks have the same chemical shifts in the two spectra. If the SH2 domains in solution had been in the same orientation as in the complex structure, binding of the peptide would only affect residues at the interface. Instead, we see changes all over the NMR spectrum, indicating significant structural changes upon

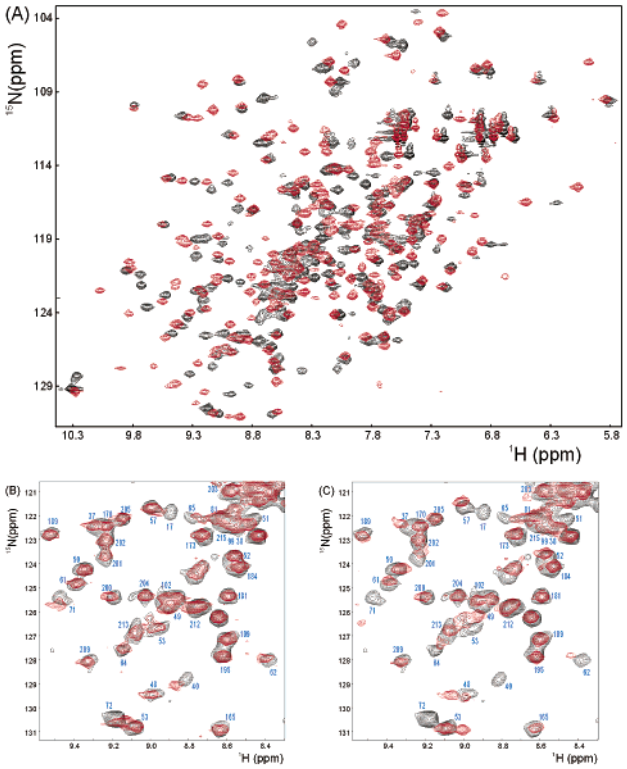

FIGURE 7: (A) Overlay of 800 MHz HSQC spectra of the apo (black) and ITAM-bound (red) ZAP-70 tSH2 domain. (B) Overlay of 600 MHz HSQC spectra of the apoprotein (black) and of that obtained after addition of a 4-fold excess of pYNEL (red). Shifts in this part of the spectrum occur predominantly at residues 17, 40, 71, and 72, which are residues directly interacting with the peptide. (C) Overlay of 600 MHz HSQC spectra of the apoprotein (black) and of that obtained after addition of a 4-fold excess of pYNEL (red), after addition of pYDVL. Here, shifts are much more pronounced for these residues in direct contact with the peptide, and (smaller) shifts are even observed for residues which are only allosterically affected by peptide binding (e.g., residues 53, 62, and 64).

peptide binding. This is of course in accordance with the difference in crystal structures that we described above.

A more detailed comparison reveals that the spectrum of the protein in complex with ITAM is significantly higher in quality than that of the apoprotein. The former has slightly better dispersion, but moreover, it shows more peaks with similarly narrow line widths. In the apo spectrum, on the

other hand, quite a number of peaks are broad and small, if not absent altogether, while others are narrow and intense. As discussed above, the NMR signals from the helices in the apoprotein could not be assigned and are broad due to conformational exchange processes. These signals become sharper upon binding to the ITAM peptide, indicating that the helices have taken up a more unique conformation. The tight binding to the ITAM peptide freezes the SH2 domains in a new orientation, which in turn apparently stabilizes the two-helix coil.

The results presented here agree well with an antibody study reported earlier (*36*). There, antibodies raised against two peptides from the inter-SH2 region had high reactivity toward the isolated tandem SH2 domains, while antibody binding was inhibited upon binding to the ITAM peptide. This clearly agrees with our observations that peptide binding induces a significant conformational change in the inter-SH2 region. A reduced level of antibody binding was also observed toward full-length ZAP-70, indicating that the inter-SH2 spacer is involved in intramolecular contacts with the C-terminal kinase domain (*36*). Interestingly, deletion of both SH2 domains increases ZAP-70 enzymatic activity and results in an increased level of tyrosine phosphorylation (*37*). These findings are compatible with a model in which deletion of this tSH2 region results in a relief of a conformational restraint, potentially leading to a conformational change that could expose regions for tyrosine phosphorylation (*37*). Increased enzymatic activity is also observed if the protein binds to the ITAM peptide, suggesting that ITAM binding also can relieve this conformational restraint in which the SH2 domains are implicated (*38*). This fits very well with the dramatic conformational change we observe upon ITAM binding, both from comparison of the crystal structures and from interpretation of the NMR data. Our data support the commonly accepted view in which ZAP-70 is activated through an allosteric mechanism after binding of the SH2 domains to the ζ chain of the TCR; the structural impact from ITAM binding is easily appreciated from Figure 3, and it is not difficult to imagine that such a large conformational rearrangement in the SH2 domains and their linker induces the necessary structural change in the kinase domain to allow phosphorylation and activation.

Finally, the NMR results show that the region around helix αB in ZAP-C (residues 222−256) undergoes some degree of flexibility as well. No clear NMR resonances could be detected for that region of the protein (intermediate exchange regime). As noted above, the largest local difference between the apo and complex crystal structures is exactly at the beginning of that region, namely, around Glu225. This turn region, around the Leu/Ile anchoring site, features drastically different main chain conformations between the two structures. In this respect, it is of interest to note that there is only one proline in the sequence following Glu225, while there are four in the equivalent region of ZAP-N, which is structurally clearly more stable.

Only ZAP-N Recruits Tetrapeptides pYXXL, and pY Binds to Only ZAP-C. An unexpected aspect of the crystal structure of ZAP-70 complexed to the doubly phosphorylated ζITAM1 sequence was that its C-terminal phosphotyrosine is bound in a pocket that is formed by contributions from both SH2 domains, rather than from the N-terminal domain only (*17*). This suggests that ZAP-N, if expressed in isolation, is incomplete. Nevertheless, binding assays have demonstrated that this isolated domain can bind phosphorylated ITAM sequences (*39*). In fact, this same study failed to show binding for the isolated C-terminal domain, the domain that ought to be complete and functional judging from the crystal structure of the complex. Furthermore, BIAcore analyses of single-point mutants of ZAP-70 show that its N-terminal SH2 domain can bind phosphopeptides in the absence of a functional C-domain (*40*).

Obviously, NMR spectroscopy is ideally suited to studying protein−ligand interactions even at lower affinities, and we have performed several titrations to study the binding affinity of phosphopeptides and phosphotyrosine. Panels B and C of Figure 7 show details of HSQC spectra recorded to study binding of the two tetrapeptides in ζITAM1 (pYNEL and pYDVL). An analysis of the chemical shift changes upon binding shows that both tetrapeptides preferably bind to ZAP-N, whereas little or no affinity is observed for ZAP-C. Furthermore, the effects measured for pYDVL are substantially larger than for pYNEL, indicating that ZAP-N has a higher affinity for the former. This is in agreement with BIAcore studies showing that ZAP-N prefers the RREEpYDVLDK sequence from ζITAM1 to its preceding GQNQLpYNELNL sequence (*40*). The crystal structure reported here clearly explains the low affinity of ZAP-C for these tetrapeptides; as was mentioned above and shown in Figure 4, the Ile/Leu anchoring site is virtually closed in the apoprotein. Our binding data are also in agreement with published ITC experiments performed on ZAP-70 and monophosphorylated peptides; although potentially both SH2 domains are available for binding, only one discrete binding event was observed (*41*).

While the C-terminal SH2 domain has less affinity for tetrapeptides than ZAP-N, it has higher affinity for plain phosphotyrosine than ZAP-N (data not shown). This is in line with the density for phosphate found in the crystal at the phosphotyrosine pocket of ZAP-C (shown in Figure 4), while this was absent at ZAP-N. This could have been anticipated, considering that the phosphotyrosine pocket on ZAP-N is formed through interaction with ZAP-C. In the apoprotein in solution, these interactions are probably not there, as discussed above. In that sense, it is important to note that the overall chemical shift changes upon pYDVL binding are relatively few (Figure 7B). Clearly, binding of the tetrapeptide does not involve a large conformational change, as would be necessary to recruit the C-terminal SH2 domain toward ZAP-N to form the complete pTyr binding pocket.

Binding Studies Support a Cooperative Model of Binding. The high specificity of binding of tandem SH2 domains to ITAM sequences is clearly linked to the well-described cooperative binding mode (*42*, *43*). A scenario can be envisaged where different stages of recognition occur between the ligand and the receptor. Both sides of the peptide contain the pYXXI/L sequence, and can therefore in principle bind to both domains. However, our structure and binding data indicate that binding will likely happen first to ZAP-N (not to the "closed" ZAP-C). Only in the (correct) case where the C-terminal half of the peptide is bound to ZAP-N will the other half have the necessary free energy potential to induce fit into the leucine binding site at ZAP-C and bring ZAP-C to the correct interface with ZAP-N. This domain in

turn completes its phosphotyrosine binding site. In such a binding process, variation in the length of the peptide between the two pYXXL/I motifs would modulate the specificity (*43*). Binding of the other half of the peptide to ZAP-N will not bring about the same cascade, and will remain a very weak interaction. While our data support such a model, more quantitative binding data would be required to verify it.

CONCLUSION

By a combined NMR and X-ray analysis, we have demonstrated the function and interdependencies of the two SH2 domains in ZAP-70. The crystal structure of the apoprotein shows a vastly different interface and orientation between the SH2 domains, compared to the ITAM-bound protein structure. In the absence of the ITAM peptide, mobility was observed for the helical region connecting the two SH2 domains. Only after ITAM binding did the domains become locked up into a more defined conformation, stabilizing the helical region. Another key difference between the apo and complex structures is the leucine anchoring site in ZAP-C, which is distorted, and basically closed, in the apoprotein.

The crystal structure reported here explains many findings reported in the literature related to peptide binding, and further helps in understanding the very high specificity with which this important protein class binds its targets. In particular, the fact that ITAM binding allows the protein to be activated is easily understood considering the large conformational change induced by binding.

Finally, the NMR titrations unambiguously show the preferences of the SH2 domains for tetrapeptides or plain phosphotyrosine. While ZAP-N is often considered to be noncomplete, it still can bind tetrapeptides, but not phosphotyrosine. The opposite is true for ZAP-C as explained by the crystal structure.

ACKNOWLEDGMENT

We thank Dr. Sara Botterell for providing the full-length ZAP-70 construct and Drs. Jerzy Schmidt and Nigel Darby for critically reading the manuscript. We also thank the anonymous reviewer for his suggestions on the use of buried surface.

REFERENCES

1. Reth, M. (1989) Antigen receptor tail clue, *Nature 338*, 383−384.
2. Cambier, J. C. (1995) Antigen and Fc receptor signaling. The awesome power of the immunoreceptor tyrosine-based activation motif (ITAM), *J. Immunol. 155*, 3281−3285.
3. van Oers, N. S., Killeen, N., and Weiss, A. (1996) Lck regulates the tyrosine phosphorylation of the T cell receptor subunits and ZAP-70 in murine thymocytes, *J. Exp. Med. 183*, 1053−1062.
4. Chan, A. C., Dalton, M., Johnson, R., Kong, G. H., Wang, T., Thoma, R., and Kurosaki, T. (1995) Activation of ZAP-70 kinase activity by phosphorylation of tyrosine 493 is required for lymphocyte antigen receptor function, *EMBO J. 14*, 2499−2508.
5. Grucza, R. A., Fütterer, K., Chan, A. C., and Waksman, G. (1999) Thermodynamic study of the binding of the tandem-SH2 domain of the Syk kinase to a dually phosphorylated ITAM peptide: Evidence for two conformers, *Biochemistry 38*, 5024−5033.
6. Gong, Q., Jin, X. H., Akk, A. M., Foger, N., White, M., Gong, Q. Q., Wardenburg, J. B., and Chan, A. C. (2001) Requirement for tyrosine residues 315 and 319 within zeta chain-associated protein 70 for T cell development, *J. Exp. Med. 194*, 507−518.
7. Magnan, A., Di Bartolo, V., Mura, A. M., Boyer, C., Richelme, M., Lin, Y. L., Roure, A., Gillet, A., Arrieumerlou, C., Acuto, O., Malissen, B., and Malissen, M. (2001) T cell development and T cell responses in mice with mutations affecting tyrosines 292 or 315 of the ZAP-70 protein tyrosine kinase, *J. Exp. Med. 194*, 491−505.
8. Di Bartolo, V., Malissen, M., Dufour, E., Sechet, E., Malissen, B., and Acuto, O. (2002) Tyrosine 315 determines optimal recruitment of ZAP-70 to the T cell antigen receptor, *Eur. J. Immunol. 32*, 568−575.
9. Kadlecek, T. A., van Oers, N. S., Lefrancois, L., Olson, S., Finlay, D., Chu, D. H., Connolly, K., Killeen, N., and Weiss, A. (1998) Differential requirements for ZAP-70 in TCR signaling and T cell development, *J. Immunol. 161*, 4688−4694.
10. Cheng, A. M., Negishi, I., Anderson, S. J., Chan, A. C., Bolen, J., Loh, D. Y., and Pawson, T. (1997) The Syk and ZAP-70 SH2-containing tyrosine kinases are implicated in pre-T cell receptor signaling, *Proc. Natl. Acad. Sci. U.S.A. 94*, 9797−9801.
11. Chan, A. C., Kadlecek, T. A., Elder, M. E., Filipovich, A. H., Kuo, W. L., Iwashima, M., Parslow, T. G., and Weiss, A. (1994) ZAP-70 deficiency in an autosomal recessive form of severe combined immunodeficiency, *Science 264*, 1599−1601.
12. Elder, M. E., Lin, D., Clever, J., Chan, A. C., Hope, T. J., Weiss, A., and Parslow, T. G. (1994) Human severe combined immunodeficiency due to a defect in ZAP-70, a T cell tyrosine kinase, *Science 264*, 1596−1599.
13. Arpaia, E., Shahar, M., Dadi, H., Cohen, A., and Roifman, C. M. (1994) Defective T cell receptor signaling and CD8+ thymic selection in humans lacking zap-70 kinase, *Cell 76*, 947−958.
14. Vu, C. B. (2000) Recent advances in the design and synthesis of SH2 inhibitors of Src, Grb2 and ZAP-70, *Curr. Med. Chem. 7*, 1081−1100.
15. Garcia-Echeverria, C. (2001) Antagonists of the Src homology 2 (SH2) domains of Grb2, Src, Lck and ZAP-70, *Curr. Med. Chem. 8*, 1589−1604.
16. Fütterer, K., Wong, J., Grucza, R. A., Chan, A. C., and Waksman, G. (1998) Structural basis for Syk tyrosine kinase ubiquity in signal transduction pathways revealed by the crystal structure of its regulatory SH2 domains bound to a dually phosphorylated ITAM peptide, *J. Mol. Biol. 281*, 523−537.
17. Hatada, M. H., Lu, X., Laird, E. R., Green, J., Morgenstern, J. P., Lou, M., Marr, C. S., Phillips, T. B., Ram, M. K., Theriault, K., et al. (1995) Molecular basis for interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor, *Nature 377*, 32−38.
18. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX Pipes, *J. Biomol. NMR 6*, 277−293.
19. Bartels, C., Xia, T., Billeter, M., Güntert, P., and Wüthrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules, *J. Biomol. NMR 5*, 1−10.
20. Johnson, B. A., and Blevins, R. A. (1994) NMRView: a computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603−614.
21. Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (1998) Trosy in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins, *Proc. Natl. Acad. Sci. U.S.A. 95*, 13585−13590.
22. Salzmann, M., Wider, G., Pervushin, K., Senn, H., and Wuthrich, K. (1999) TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins, *J. Am. Chem. Soc. 121*, 844−848.
23. Collaborative Computational Project No. 4 (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
24. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for binding protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
25. Navaza, J. (1994) AMoRe: an automated package for molecular replacement, *Acta Crystallogr. A50*, 157−163.
26. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
27. Chothia, C. (1974) Hydrophobic bonding and accessible surface area in proteins, *Nature 248*, 338−339.

28. Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: estimation of static accessibility, *J. Mol. Biol. 55*, 379−400.
29. Hatada, M. H., Lu, X., Laird, E. R., Karas, J. L., Zoller, M. J., and Holt, D. (1995) Crystalline ZAP family proteins, PCT Appl. WO 97/08300.
30. Laskowski, R., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
31. Kuriyan, J., and Cowburn, D. (1993) Structures of SH2 and SH3 domains, *Curr. Opin. Struct. Biol. 3*, 828−837.
32. Eck, M. J., Shoelson, S. E., and Harrison, S. C. (1993) Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56lck, *Nature 362*, 87−91.
33. Tsai, J., Lin, S. L., Wolfson, H. J., and Nussinov, R. (1997) Studies of protein−protein interfaces: A statistical analysis of the hydrophobic effect, *Protein Sci. 6*, 53−64.
34. Janin, J., and Rodier, F. (1995) Protein−protein interaction at crystal contacts, *Proteins 23*, 580−587.
35. Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) Attenuated T-2 relaxation by mutual cancellation of dipole−dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12366−12371.
36. Grazioli, L., Germain, V., Weiss, A., and Acuto, O. (1998) Anti-Peptide antibodies detect conformational changes of the inter-SH2 domain of Zap-70 due to binding to the zeta chain and to intramolecular interactions, *J. Biol. Chem. 273*, 8916−8921.
37. Magistrelli, G., Bosotti, R., Valsasina, B., Visco, C., Perego, R., Toma, S., Acuto, O., and Isacchi, A. (1999) Role of the Src homology 2 domains and interdomain regions in ZAP-70 phosphorylation and enzymatic activity, *Eur. J. Biochem. 266*, 1166−1173.
38. Visco, C., Magistrelli, G., Bosotti, R., Perego, R., Rusconi, L., Toma, S., Zamai, M., Acuto, O., and Isacchi, A. (2000) Activation of ZAP-70 tyrosine kinase due to a structural rearrangement induced by tyrosine phosphorylation and/or ITAM binding, *Biochemistry 39*, 2784−2791.
39. Isakov, N., Wange, R. L., Burgess, W. H., Watts, J. D., Aebersold, R., and Samelson, L. E. (1995) ZAP-70 binding specificity to T cell receptor tyrosine-based activation motifs: the tandem SH2 domains of ZAP-70 bind distinct tyrosine-based activation motifs with varying affinity, *J. Exp. Med. 181*, 375−380.
40. Labadia, M. E., Jakes, S., Grygon, C. A., Greenwood, D. J., Schembriking, J., Lukas, S. M., Warren, T. C., and Ingraham, R. H. (1997) Interaction between the SH2 domains of Zap-70 and the tyrosine-based activation motif 1 sequence of the zeta subunit of the T-cell receptor, *Arch. Biochem. Biophys. 342*, 117−125.
41. O'Brien, R., Rugman, P., Renzoni, D., Layton, M., Handa, R., Hilyard, K., Waterfield, M. D., Driscoll, P. C., and Ladbury, J. E. (2000) Alternative modes of binding of proteins with tandem SH2 domains, *Protein Sci. 9*, 570−579.
42. Ottinger, E. A., Botfield, M. C., and Shoelson, S. E. (1998) Tandem SH2 domains confer high specificity in tyrosine kinase signaling, *J. Biol. Chem. 273*, 729−735.
43. Vely, F., Nunes, J. A., Malissen, B., and Hedgecock, C. J. R. (1997) Analysis of immunoreceptor tyrosine-based activation motif (ITAM) binding to Zap-70 by surface plasmon resonance, *Eur. J. Immunol. 27*, 3010−3014.
44. Kraulis, P. J. (1991) MOLSCRIPT: a program to reproduce both detailed and schematic plots of protein structure, *J. Appl. Crystallogr. 24*, 946−950.
45. Merritt, E. A., and Bacon, D. J. (1997) Raster3D: photorealistic molecular graphics, *Methods Enzymol. 277*, 505−524.
46. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins 11*, 281−296.

BI026465E