Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Atomic resolution structure of a mutant of the spectrin SH3 domain

**Rita Berisio,[a]* Anarosa Viguera,[b] Luis Serrano[b] and Matthias Wilmanns[c]**

[a]Centro di Studio di Biocristallografia, CNR and Dipartimento di Chimica, Università di Napoli 'Federico II', Via Mezzocannone 4, I-80134 Napoli, Italy, [b]EMBL, Meyerhofstrasse 1, D-69117 Heidelberg, Germany, and [c]EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany

Correspondence e-mail: rberisio@chemistry.unina.it

© 2001 International Union of Crystallography
Printed in Denmark – all rights reserved

The crystal structure of an $\alpha$-spectrin Src-homology 3 (SH3) domain mutant has been refined at 1.12 Å resolution. This X-ray structure is at the highest resolution achieved so far for an SH3 domain. The structure allows the identification of a complete set of specific interactions and is useful for the elucidation of relations between structure and pH-dependent thermodynamic stability in a series of SH3 domain mutants.

Received 30 August 2000
Accepted 2 November 2000

**PDB Reference:** SH3 domain mutant, 1g2b.

## 1. Introduction

The Src-homology 3 (SH3) region is a modular protein domain of about 60 residues occurring in a variety of eukaryotic proteins involved in signal transduction, cell polarization and membrane–cytoskeleton interactions. Its three-dimensional structure consists of a five-stranded orthogonal $\beta$-sheet sandwich and contains three distinct $\beta$-hairpins. A common denominator in SH3 function is its capability to bind to polyproline-rich segments of a wide variety of ligand proteins. To date, about 30 structures of various SH3 domains in the absence of any ligands or in the presence of either peptide ligands or interacting protein domains have been determined by X-ray crystallography or NMR spectroscopy. Of particular interest were recent structures of tyrosine kinases from different families, bearing an N-terminal SH3 domain, that reveal how their catalytic activities are regulated by the presence of SH2 and SH3 domains (Xu *et al.*, 1997; Schindler *et al.*, 1999).

The SH3 domain of the cytoskeletal protein spectrin has been subject to extensive investigations on its crystal structure (Musacchio *et al.*, 1992), its structure in solution by NMR spectroscopy (Viguera *et al.*, 1996) and structural investigations on complexes with proline-rich ligand motifs (Musacchio *et al.*, 1994; Pisabarro *et al.*, 1998). A number of kinetic and thermodynamic studies have aimed to elucidate its folding pathway (Viguera *et al.*, 1996; Martinez *et al.*, 1999). As part of these studies, it has been shown that the refolding rate of the SH3 domain can be increased sevenfold by circular permutation when the native termini are fused and new termini are created by cutting the long RT loop (S19–P20s mutant), but not when the cut is situated in the regular distal $\beta$-hairpin (N47–D48s mutant). Moreover, the crystal structures of both the S19–P20s and the N47–D48s circular permutants have been determined (Viguera *et al.*, 1996). Crystals of one of the circular permutants (Viguera *et al.*, 1996) have been further improved, allowing X-ray data collection at atomic resolution. Here, we report the structure of the N47–D48s circular permutant of the $\alpha$-spectrin SH3 domain at 1.12 Å resolution. This structure provides the most accurate coordinate set of an SH3 domain determined so far and will be useful as a reference model for future studies on SH3 that require accurate structural data.

## 2. Materials and methods

### 2.1. Crystallization and data collection

Crystallization was performed by the hanging-drop technique using magnesium sulfate as a precipitant agent, as described by Viguera *et al.* (1996). More recent crystallization trials have produced larger crystals of about 0.5 mm in all three dimensions. A number of crystals were tested for optimized data collection. Synchrotron data were collected from the most optimum crystal at 100 K at the X11 beamline, EMBL Outstation, DESY, Hamburg using a 30 cm MAR Research image plate as detector. Prior to mounting in the nitrogen stream, the crystals were transferred into mother liquor containing 10%(*v*/*v*) glycerol. The data were processed using the *XDS* package (Kabsch, 1988). The crystals belong to the $P2_12_12_1$ space group and are isomorphous to those of Viguera *et al.* (1996), with unit-cell parameters $a$ = 31.10, $b$ = 42.97, $c$ = 53.05 Å. The Wilson plot, calculated in the resolution range 4.5–1.12 Å using the *CCP*4 package (Collaborative Computational Project, Number 4, 1994), yielded an overall isotropic temperature factor of 7.0 Å$^2$. Statistics of the data processing are reported in Table 1.

**Table 1**
Data-processing statistics.

Values in parentheses are for the highest resolution shell (1.2–1.12 Å).

| | |
|---|---|
| Resolution range (Å) | 14.0–1.12 |
| No. of observations | 102998 |
| No. of unique reflections | 25803 |
| Completeness (%) | 93.4 (91.2) |
| $R_{\rm merge}$ (%) | 7.1 (16.0) |

### 2.2. Structure refinement

A first model was calculated after molecular replacement (*AMoRe*; Navaza, 1994) using the 1.77 Å resolution structure (Viguera *et al.*, 1996; PDB entry code 1tud). The structure was refined using *SHELXL* (Sheldrick & Schneider, 1997) after refinement in *REFMAC* (Murshudov *et al.*, 1997). The solvent model was built and was updated at each cycle using *ARP* (Lamzin & Wilson, 1993). In *SHELXL*, bond distances and bond angles were initially restrained with target $\sigma$ values of 0.02 and 0.04 Å, respectively. The introduction of anisotropy resulted in a decrease in the $R$ factor and $R_{\rm free}$ from 23.4 and 25.8% to 18.8 and 22.8%, respectively. During anisotropic refinement, rigid-body restraints were applied to the differences in the mean-square atomic displacement amplitudes along 1,2 and 1,3 distances (standard deviation 0.01 Å$^2$) and similarity restraints were applied to the $U_{ij}$ components of atoms that were close in space (e.s.d. 0.05 Å$^2$), whereas solvent sites were restrained to be approximately isotropic. H atoms were included as 'riding' H atoms, their positions being calculated stereochemically at each cycle. Based on the omit ($F_o - F_c$) and the ($3F_o - F_c$) electron-density maps, Asp48 and the backbone of the N-terminal Met0 were modelled, whereas the side chain of Met0 was not included in the structure-factor calculation as it was unstructured. In addition, three sulfate ions were identified. Solvent sites in the vicinity of residues having double conformation were considered as partially occupied, their occupancies being restrained to those of the appropriate side chain using the SUMP option in *SHELXL* (Sheldrick & Schneider, 1997). During the entire refinement, the working set consisted of 95% of the data, while the remaining 5% was used to monitor the progress of refinement by $R_{\rm free}$ calculation (Brünger, 1992). Most of the restraints were relaxed in the last steps of the *SHELXL* refinement, on the indication of $R_{\rm free}$ changes. To follow these changes through the whole refinement, the same reference set of reflections was used for both *REFMAC* and *SHELXL*.

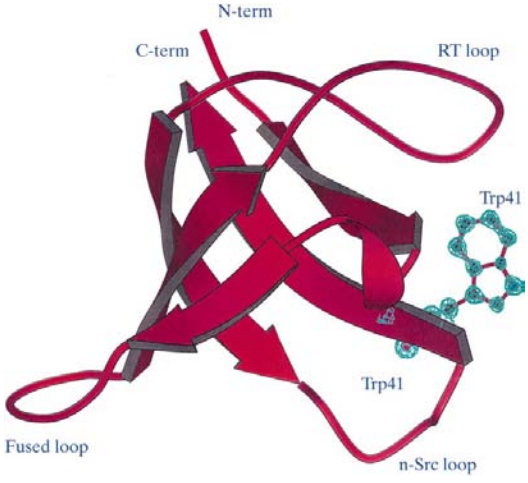


**Figure 1**
Ribbon-diagram representation of the N47–D48s permutant of the $\alpha$-spectrin SH3 domain. The ($3F_o - 2F_c$) map, contoured at 3$\sigma$, is also shown for Trp41.

**Table 2**
Summary of refinement statistics.

| | |
|---|---|
| No. of protein residues | 61 |
| Total number of water sites (partially occupied) | 158 (14) |
| Number of sulfate ions | 3 |
| R.m.s. deviations from ideal values (Å) | |
| Bond distances | 0.026 |
| Angle distances | 0.039 |
| Average isotropic atomic displacement factors (Å$^2$) | |
| Main-chain protein atoms | 8.5 |
| Side-chain protein atoms | 14.4 |
| Sulfate ions | 38.8 |
| Water molecules | 28.4 |
| Ramachandran plot | |
| Residues in most favoured region (%) | 98.2 |
| Residues in additionally allowed region (%) | 1.8 |
| Residues in generously allowed or disallowed regions (%) | 0 |

Estimated uncertainties in the atomic positions were derived using the block-matrix least-squares refinement implemented in *SHELXL* (Scheldrick & Schneider, 1997). For block-matrix refinement, the model was split into two overlapping blocks containing approximately the same number of atoms and most of the restraints were removed during refinement. All data were used in the final refinement and map calculations and in the final block-refinement round. The model converged to an $R$ factor of 14.8% (13.6% for $F > 4\sigma$) and $R_{\rm free}$ of 20.2% (18.4% for $F > 4\sigma$).

The programs *PROCHECK* (Laskowski *et al.*, 1993) and *WHAT-CHECK* (Hooft *et al.*, 1996) were used to validate the structure.

## 3. Results and discussion

The crystal structure of the N47–D48s circular permutant presented here constitutes the highest resolution structure of an SH3 domain. Refinement of N47–D48s resulted in good quality of the electron-density maps, as shown for the representative Trp41 in Fig. 1. The final deviations of the refinement parameters from target values are given in Table 2. The mean estimated uncertainties of the atomic positions, as derived from block-matrix refinement, were as low as 0.03 Å for the main-chain atoms and 0.04 Å for the side-chain atoms. Many H atoms (about 30% of the total number of H atoms) were visible in the electron-density maps before their contribution was included in the structure-factor calculation. These H atoms were mainly located on the backbone (about 50% of the backbone H atoms) and on the C$^\beta$ atoms (about 39% of the total number of H$^\beta$).

To validate the crystal structure, the correlation between the N—C$^\alpha$—C backbone angles and the conformation of the residues was investigated. This analysis was based on the recent observation that in ultrahigh-resolution protein structures (Esposito *et al.*, 2000) residues in the $\beta$-strand conformation show a N—C$^\alpha$—C angle smaller than the target value used for refinement (111.2°). Our analysis was restricted to the 21 residues in the extended conformation, as defined by Kabsch & Sander (1983). An average value of 109.7° was calculated for this angle, although the structure was refined in a restrained mode. The difference of 1.5° from the target is significantly larger than the estimated uncertainty associated with this angle (0.5°) and therefore accounts for the high accuracy of the model.

The structure superimposes on the previously determined structure at 1.77 Å resolution (PDB entry code 1tud) with a root-mean-square deviation (r.m.s.d.), averaged over the main-chain atoms, of 0.37 Å (for 240 backbone atom pairs). The structural differences from the previous model are a consequence of the better quality of the electron-density maps. The overall structure is also similar both to the native SH3 domain structure (Viguera *et al.*, 1996; PDB entry code 1shg) and to the S19–P20s permutant (PDB entry code 1tuc), although all three domains crystallize in different space groups. The r.m.s.d. between our structure and the native one, calculated for

224 backbone atom pairs, is as low as 0.54 Å. Larger differences from the S19–P20s permutant (PDB entry code 1tuc), with an r.m.s.d. of 1.53 Å for 232 backbone atom pairs, are mainly located in the fused loop and the RT loop, which is cleaved in the S19–P20s mutant.

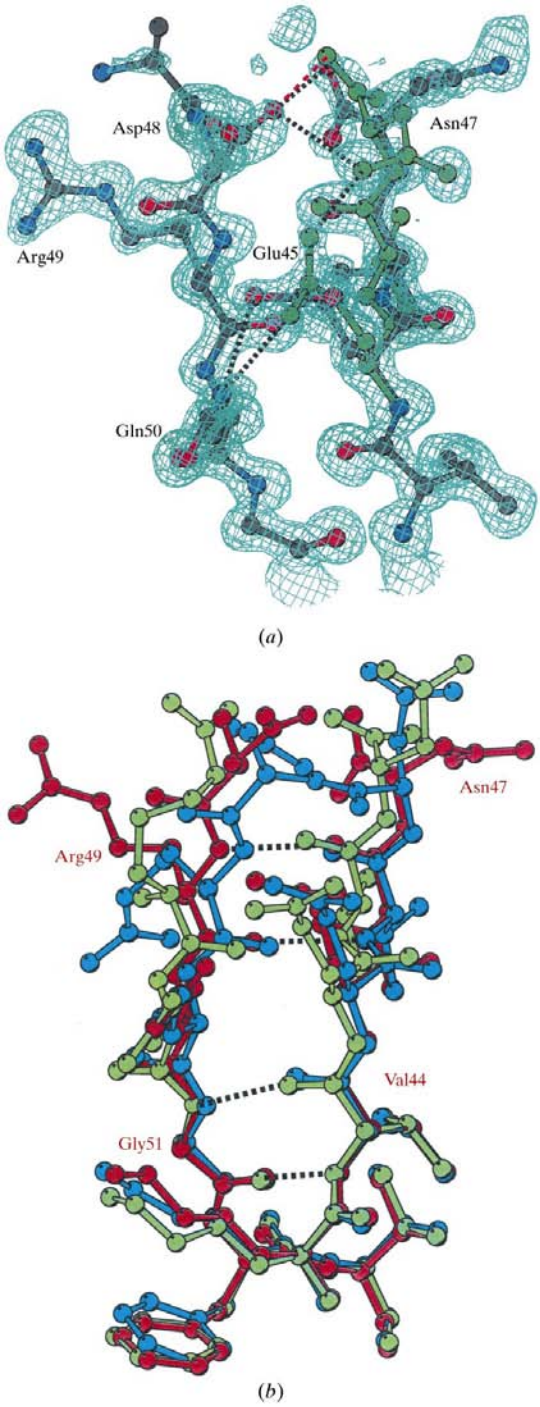


**Figure 2**
Description of the cleaved distal loop. (*a*) $3F_o - 2F_c$ electron-density map contoured at 1.6$\sigma$ of the cleaved site. The *A* conformation of the C-terminus is shown in green and the unfavourable charge–charge interaction between its carboxyl-terminus and Asp48 is shown in red. (*b*) Superposition of the cleaved distal loop of the structure here presented (red; PDB code entry 1g2b) with the previously refined N47–D48s (green; PDB code entry 1tud) and the native structure (cyan; PDB code entry 1shg). The main-chain hydrogen bonds are indicated by dashed lines. Only the most populated conformation of the C-terminus is shown for clarity in 1g2b. Figures were drawn using the program *BOBSCRIPT* (Esnouf, 1997).

Conformational differences from the native structure, further discussed in this contribution, are mainly localized in the two engineered regions: the loop corresponding to the fused native-like termini (fused loop) and the cleaved distal loop. The fused loop is well ordered and characterized by relatively low atomic displacement parameters in both our model and the previously determined model (PDB entry code 1tud). However, no specific direct intramolecular interactions of this loop with other parts of the protein contribute to its stabilization and interactions within this loop are mediated by two water molecules. Moreover, the fused loop is exposed to crystal contacts and its conformation is different in all permutants determined so far (Viguera *et al.*, 1995). It is therefore likely that packing forces strongly contribute to its stabilization. Indeed, previous NMR studies (Viguera *et al.*, 1996) have indicated that this loop is not structured in solution.

More interesting results arose from the analysis of the cleaved distal loop. Interactions among residues in this cleaved loop were invoked previously (Viguera *et al.*, 1996; Martinez *et al.*, 1999) to explain the higher enthalpy of unfolding of the N47–D48s permutant when compared with the wild-type protein. Viguera *et al.* (1996) observed, using the previously refined model at 1.77 Å, that the cleavage of the distal loop (residues 41–54) did not disrupt the protein structure. Indeed, all hydrogen bonds stabilizing the cleaved distal loop were maintained except for the hydrogen-bonding interaction between the carbonyl O atom of Val46 and the backbone N atom of Arg49. This interaction was substituted by a hydrogen bond between Val46 and the side-chain $N^{\epsilon}$ of Arg49.

Our atomic resolution structure provides more accurate as well as new information about the cleaved distal loop. In fact, not only could Asp48 and the backbone of Met0 be modelled, but also two alternative conformations were identified for the C-terminal Glu45, Val46 and Asn47 (Fig. 2*a*). The *A* and *B* conformations of the C-terminal residues, shown in Fig. 2(*a*), occur with occupancy factors of 0.4 and 0.6, respectively. As shown in Fig. 2(*b*), the conformation of Arg49 differs from that found in the previously refined N47–D48s permutant and its backbone conformation is more similar to that adopted in the native structure. Indeed, in our atomic resolution structure the backbone N atom of Arg49 maintains the hydrogen bond with the carbonyl O atom of Val46 (Fig. 2*b*), which in the native structure stabilizes the $\beta$-turn. The side chain of Arg49 is further stabilized by electrostatic interactions with two sulfate molecules. Therefore, not only is the structure of the domain substantially unaffected by the cleavage of the distal loop, but also none of the main-chain hydrogen bonds stabilizing the $\beta$-turn in the wild-type protein are lost. Despite the differences from the previously refined model (Viguera *et al.*, 1996), the hypotheses formulated to explain the calorimetric data from the N47–D48s permutant are strengthened by our results.

The analysis of the C-terminus of the SH3 mutant has revealed another interesting feature. As reported by Martinez *et al.* (1999), the N47–D48s permutant shows an enhanced stability at pH 3.5, leading to an increase in the melting temperature of 4 K compared with neutral pH. In contrast, this acidic pH induced stabilization has not been observed in the wild-type domain. No structural finding was available to explain this result.

In our atomic resolution structure, two conformations were modelled for the C-terminus. In the *A* conformation, the $N^{\delta 2}$ atom of Asn47 is hydrogen bonded to the $O^{\delta 2}$ atom of Asp48 and the $O^{\epsilon 2}$ atom of Glu45, the latter residue also showing two conformations (Fig. 2*a*). However, this conformation leads to an unfavourable charge–charge interaction between the side chain of Asp48 and the carboxyl group of Asn47 (3.6 Å from the $O^{\delta 2}$ atom of Asp48 to one of the carboxy-terminal O atoms of Asn47). In contrast, in conformation *B* this charge–charge interaction is not present. The two hydrogen bonds of Asn47 to Asp48 and with Glu45 are also lost and Glu45 rotates, keeping the hydrogen bond with the $N^{\epsilon 2}$ atom of Gln50 (Fig. 2*a*). Therefore, two

energetically equivalent conformations, which are a consequence of stabilizing hydrogen bonds and destabilizing charge–charge interactions, exist inside the crystal. At pH 3.5, it is likely that the negatively charged carboxyl groups involved in the charge–charge interaction are protonated, thus favouring an *A*-like conformation. In this conformation, stabilizing hydrogen bonds can be formed without any destabilizing effect, thus accounting for the enhanced stability of the N47–D48s permutant at acidic pH.

## 4. Conclusions

Atomic resolution data have here provided a detailed model for the entire structure of the N47–D48s permutant of the $\alpha$-spectrin SH3 domain. The low e.s.d. values as well as the analysis of the correlation between conformation and geometry indicate that the structure is precise and accurate. Therefore, this structure may be well suited for use in further structure-based studies of SH3 domains.

Moreover, the data were employed to improve several features of the model, such as the conformation of the residues in the cleaved region, some of which were not structured in the previously refined model (Viguera *et al.*, 1996). The structural results were used to reinvestigate, on a molecular level, the available thermodynamic data. Moreover, they have provided a structural tool to explain the enhanced thermal stability of the N47–D48 mutant at low pH.

RB was supported by the TMR/LSF Grant (HPRI-CT-1999-00017) for her visits to the EMBL Hamburg Outstation.

## References

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.
Esposito, L., Vitagliano, L., Sica, F., Sorrentino, G., Zagari, A. & Mazzarella, L. (2000). *J. Mol. Biol.* **297**, 713–32.
Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. (1996). *Nature (London)*, **381**, 272.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**(1), 67–71.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Martinez, J. C., Viguera, A. R., Berisio, R., Willmanns, M., Mateo, P. L., Filinov, V. & Serrano, L. (1999). *Biochemistry*, **38**. 549–559.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. & Saraste, M. (1992). *Nature (London)*, **359**, 851–855.
Musacchio, A., Saraste, M. & Wilmanns, M. (1994). *Nature Struct. Biol.* **1**, 546–551.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Pisabarro, M. T., Serrano, L. & Wilmanns, M. (1998). *J. Mol. Biol.* **281**, 513–521.
Schindler, T., Sicheri, F., Pico, A., Gazit, A., Levitzki, A. & Kuriyan, J. (1999). *Mol. Cell*, **3**, 639–648.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Viguera, A. R., Blanco, F. J. & Serrano, L. J. (1995). *J. Mol. Biol.* **247**, 670–681.
Viguera, A R., Serrano, L. & Wilmanns, M. (1996). *Nature Struct. Biol.* **3**, 874–880.
Xu, W., Harrison, S. C. & Eck, M. J. (1997). *Nature (London)*, **385**, 595–609.