

The crystal structure of human CskSH3: structural diversity near the RT-Src and n-Src loop

T.V. Borchert**, M. Mathieu, J.Ph. Zeelen, S.A. Courtneidge, R.K. Wierenga*

EMBL, Postfach 102209, D-69012 Heidelberg, Germany

Received 19 January 1994

Abstract

SH3 domains are modules occurring in diverse proteins, ranging from cytoskeletal proteins to signaling proteins, such as tyrosine kinases. The crystal structure of the SH3 domain of Csk (c-Src specific tyrosine kinase) has been refined at a resolution of 2.5 Å, with an R-factor of 22.4%. The structure is very similar to the FynSH3 crystal structure. When comparing CskSH3 and FynSH3 it is seen that the structural and charge differences of the RT-Src loop and the n-Src loop, near the conserved Trp⁴⁷, correlate with different binding properties of these SH3 domains. The structure comparison suggests that those glycines and acid residues which are very well conserved in the SH3 sequences are important for the stability of the SH3 fold.

Key words: Fyn; Csk; Src; Crystal structure; Averaging; Tyrosine kinase

1. Introduction

The Src family of tyrosine kinases are negatively regulated by phosphorylation of a tyrosine residue near their carboxy-termini (Tyr⁵²⁷ in the case of Src). To date, only one enzyme has been described which is able to phosphorylate this site. This ubiquitously expressed enzyme [1] has been called Csk, for c-Src kinase, although it is now clear that it can phosphorylate several different members of the Src family. The importance of this enzyme is demonstrated by the observation that mice lacking *csk* die in utero [2,3].

Molecular cloning of the gene encoding Csk [4,5] revealed that topographically it was remarkably similar to the Src family of kinases themselves, in that it also has one SH3 domain and one SH2 domain amino terminal to the catalytic domain. However, Csk lacks myristylation sequences, a unique domain, and carboxy terminal sequences of the sort that regulate Src kinases. Csk has a very restricted substrate specificity: phosphorylating only native Src family kinases, and not denatured protein or peptides modelled on the carboxy terminus of Src [1]. The molecular basis of its substrate specificity is not

known, although it is possible that its SH2 and SH3 domains are involved.

SH3 domains (60–100 residues) occur in many proteins involved in signal transduction, but also in cytoskeletal proteins, such as spectrin [6]. SH3 domains are probably involved in protein–protein interactions [7]; it has been found that some SH3 domains can bind proline rich peptides [8–10]. In tyrosine kinases of the Src family the SH3 domain is important for the regulation of the tyrosine kinase activity via intramolecular interactions [11–13] and for cytoplasmic signaling via intermolecular protein–protein interactions [9,10].

Several structures of SH3 domains have been determined (reviewed by Kuriyan and Cowburn [14] and Musacchio et al. [15]). Crystal structures are known of SH3 from human Fyn [16], and chicken spectrin [17], and NMR structures are known of SH3 from chicken Src [18], human PLC- γ [19], and of the p85 α subunits of bovine PI-3 kinase [20] and human PI-3 kinase [21]. The SH3 molecule consists of two sheets of three antiparallel β -strands, wrapped around a hydrophobic core. There are three extensive loops in the structure (Fig. 1). It has been found that the conserved residues [17] are on one side of the molecule [17,19,20]. This patch of conserved residues overlaps with residues identified in NMR experiments with SrcSH3 [18] and the SH3 domain of p85 α [20] as being involved in the binding of proline-rich peptides. Particular attention has been drawn to the occurrence of aromatic residues within this patch of conserved residues [16,21], as well as to the importance of the loops flanking this conserved patch [16,18,20]. Here we report

*Corresponding author. Fax: (49) (6221) 387 306.

**Present address: Novo-Nordisk A/S, Industrial Biotechnology, Novo Allé, DK-2880 Bagsvaerd, Denmark. Fax: (45) (44) 980 246.

Abbreviations: DTT, 1,4-dithiothreitol; EDTA, ethylenediamine-tetraacetic acid; TEA, triethanolamine.

the expression (in *E. coli*), purification and structural characterisation of human CskSH3. We discuss its structure and compare it with the FynSH3 crystal structure.

2. Materials and methods

By multiple sequence alignment of human Csk with other members of the Src-kinase family we concluded that the domain boundaries of the human CskSH3 were residues Ser¹⁰ and Val⁷¹. The gene encoding the human Csk gene was provided by M. Bergman. A DNA fragment encoding the sequence spanning from Ser¹⁰ to Val⁷¹ and an additional N-terminal methionine residue was synthesised by PCR using the human Csk gene as a template and two oligonucleotides as primers. The 5' primer introduced an *Nde*I restriction site overlapping the N-terminal methionine codon. The 3' primer introduces a stop codon after residue 71 and a downstream *Bam*III restriction site. *E. coli* strain XL1-blue [22] was used for the genetic manipulations. The CskSH3 domain was cloned into the *E. coli* plasmid pET3A [23] using the above-mentioned restriction sites. Plasmid pET3A was provided by F.W. Studier. The resulting plasmid was denoted pCSKSH3. The DNA sequence of the coding region was verified. Expression strain BL21(DE3) [24] was transformed with pCSKSH3.

The expression strain was grown at 25°C in M9 medium [25] supplemented with 0.2% casamino acids and 100 µg/ml ampicillin. The culture was induced with 0.4 mM isopropyl-β-D-thiogalactopyranoside at OD₆₀₀ = 0.6. Protein expression was carried out for 17 h. The pellet from 1 l culture was resuspended in 30 ml of buffer A (100 mM TEA, pH 7.8, 50 mM NaCl, 1 mM of each of DTT, EDTA and Na₂S₂O₃ and 0.5 mM phenylmethylsulfonyl fluoride). Cells were lysed by sonication. All purification steps were done at 4°C. 0.25% polyethylene imine was added to the supernatant from a centrifugation step (20 min at 15,000 × g). Precipitated proteins and DNA were discarded after another centrifugation step. The 60–90% ammonium sulphate fraction of the supernatant was resuspended in buffer B (25 mM TEA, pH 7.8, 25 mM NaCl, 1 mM DTT, 1 mM EDTA, 1 mM Na₂S₂O₃). The protein was loaded on an S-Sepharose column, pre-equilibrated with buffer B. The column was eluted with buffer B. The sample was concentrated on an Amicon concentrator to 3 ml and loaded on a Sephadex G75 column equilibrated with buffer C (100 mM TEA, pH 7.5, 500 mM NaCl, 1 mM DTT, 1 mM EDTA and 1 mM Na₂S₂O₃). CskSH3-containing fractions were pooled. 15 mg of pure CskSH3 was obtained from 1 l culture.

Well diffracting crystals were obtained with the following conditions, using the hanging drop method. The protein was dialysed against 10 mM TEA/HCl, pH 7.2, 20 mM NaCl, 1 mM DTT, 1 mM EDTA, 1 mM Na₂S₂O₃ and concentrated to 10 mg/ml. The well solution was 100 mM acetic acid/NaOH, pH 4.8, 3.0 M ammonium sulphate, 10% ethanol, 1 mM DTT, 1 mM EDTA, 1 mM Na₂S₂O₃. The hanging drops were made by mixing 2 µl protein solution with 2 µl well solution. Crystals grow in 3 days at room temperature. The crystals have the symmetry of space group C222₁ (Table 1). From the volume of the cell and the size of the protein it can be expected that the asymmetric unit consists most likely of 4 ($V_m = 2.7 \text{ Å}^3/\text{Da}$) or 5 ($V_m = 2.2 \text{ Å}^3/\text{Da}$) SH3 domains per asymmetric unit.

Data were collected of native and heavy atom derivative data sets with the FAST area detector. With the native data set the self rotation function was calculated with the program ALMN [26] and the program GLRF [27]. The most prominent peaks were found for $Kappa = 180^\circ$. Cross rotation functions were calculated but no peaks above background were found. An extensive heavy atom search resulted in one good derivative and several weak derivatives as shown in Table 1. The heavy atom parameters were refined using the native data set with the program mlphare [28] at 3.2 Å resolution, and MIRAS-phases with an average figure-of-merit of 0.6 for acentric reflections were calculated. In the resulting map the spherical boundaries of four SH3 molecules could be recognised. The 'bones' tracing [29] was used for manually finding the best superposition of a FynSH3 [16] model with the region of the electron density map which could most clearly be interpreted. This position was refined using the real-space 6D search program NCS6D provided in the RAVE package [30], and the first CskSH3 molecule (molecule A) was then built and a mask constructed around it. The other three SH3 molecules (molecules B, C and D) were also

manually fitted into their corresponding densities and the transformation relationships with respect to the first reference molecule were improved with the a_rt_improve program of the O-package [29]. The four SH3-molecules are packed in the cell as a dimer of dimers; the orientations of the local twofold axes agree with the peaks in the self rotation function. Fourfold averaging was then carried out using the DEMON package [31], and the four CskSH3 molecules were built in the improved electron density map. The structure was refined with X-PLOR [32] at 2.5 Å resolution, using the native2 data set, first applying non-crystallographic symmetry restraints. Several simulated annealing runs, followed by energy minimisation, were interleaved with model-building sessions to check the fit of the model with the electron density map. In the final stages of refinement the non-crystallographic symmetry restraints were removed and eight solvent peaks present in the 2Fo-Fc map and in the Fo-Fc map were interpreted as water molecules and included in the refinement. The final refinement statistics are tabulated in Table 1. The heavy atom sites are on the surface of the molecules: for example the four mercury sites are near the SG atoms of the exposed residue Cys¹⁴. The structure has been analysed with the O and WHAT IF [33] programs.

For comparison of the different CskSH3 molecules with themselves or with the FynSH3 crystal structure (molecule A of 1SHF.PRE in the Brookhaven Protein Data Bank) 21 CA atoms of the 6 different strands were used for the superposition. In CskSH3 this concerns residues 13–16 (strand a), 35–40 (strand b1, strand b2), 47–51 (strand c), 57–60 (strand d), and 65–66 (strand e). The alignments are shown in Fig. 1. The three major loops, between the a/b strands, the b/c strands and the c/d strands (Fig. 1) will be referred to as the RT-Src loop, the n-Src loop and the distal loop, respectively [16]. The 14 conserved residues identified by Musacchio et al. [17] are tabulated in Fig. 1.

3. Results and discussion

The final model has a low R-factor and good geometry (Table 1). All main chain dihedrals are within the allowed region of the Ramachandran plot.

There is continuous density for all four molecules in the final map at the 1.2 sigma level, and there are no regions with average main chain B-factors above 30 Å², except near the N- and C-terminus. Two residues at the N-terminus and 3 residues at the C-terminus are mobile; as is documented in Fig. 1, residues 11–68 have a well defined structure. The size of this fragment, 58 residues, is only one residue shorter than that observed for FynSH3 (Fig. 1). According to the sequence alignment of Fig. 1, there is 33% sequence identity between CskSH3 and FynSH3.

Despite the different packing arrangements of the four molecules, there are only minor differences in the four structures. For comparison purposes we will use molecule A. The secondary structure (Fig. 1) is very similar to that seen in FynSH3, including the presence of a ₃₁₀-helix between strand d and strand e.

In Fig. 2 the CA traces of CskSH3 and FynSH3 have been superimposed. The 21 CA atoms used for superposition, superimpose with an rms difference of 0.5 Å.

In Fig. 3 the structural differences between CskSH3 and FynSH3 have been quantified in a CA distance plot. Except for the N-terminus and C-terminus there are three loops where equivalent CA atoms deviate more than 1.5 Å: the RT-Src loop (near Ala²⁴), the n-Src loop (near Lys⁴³), and the distal loop (near Lys⁵³). In the n-Src

loop and in the distal loop there is a one-residue insertion and deletion, respectively.

The structural change in the RT-Src loop is possibly

| Fyn ¹ | Csk ² | Secondary structure nomenclature | Src ³ |
|------------------|------------------|----------------------------------|------------------|
| Val84 | Gly11 | | |
| Thr85 | Thr12 | | |
| Leu86 | Glu13 | β | βa |
| Phe87 | Cys14 | β | βa |
| Val88 | Ile15 | β | βa |
| Ala89* | Ala16 | β | βa |
| Leu90* | Lys17 | | Leu89 |
| Tyr91* | Tyr18 | | Tyr90 |
| Asp92* | Asn19 | | Asp91 |
| Tyr93* | Phe20 | | Tyr92* |
| Glu94 | His21 | | Glu93 |
| Ala95 | Gly22 | | Ser94 |
| <u>Arg96</u> | Thr23 | RT-Src loop | <u>Arg95*</u> |
| Thr97 | Ala24 | | <u>Thr96*</u> |
| <u>Glu98</u> | <u>Glu25</u> | | <u>Glu97</u> |
| <u>Asp99</u> | Gln26 | | Thr98* |
| <u>Asp100*</u> | <u>Asp27</u> | | Asp99 |
| Leu101 | Leu28 | | Leu100 |
| Ser102 | Pro29 | | Ser101 |
| Phe103 | Phe30 | | Phe102 |
| His104 | Cys31 | | Lys103 |
| Lys105 | Lys32 | | Lys104 |
| Gly106* | Gly33 | | Gly105 |
| Glu107* | Asp34 | | Glu106 |
| Lys108 | Val35 | β | βb1 |
| Phe109 | Leu36 | β | βb1 |
| Gln110 | Thr37 | β | βb2 |
| Ile111 | Ile38 | β | βb2 |
| Leu112 | Val39 | β | βb2 |
| Asn113 | Ala40 | β | βb2 |
| Ser114 | Val41 | β | βb2 |
| Ser115 | Thr42 | | Thr114* |
| <u>Glu116</u> | <u>Lys43</u> | n-Src loop | <u>Glu115*</u> |
| Gly117 | <u>Asp44</u> | | Gly116 |
| | Pro45 | | |
| <u>Asp118</u> | Asn46 | | Asp117 |
| <u>Trp119*</u> | Trp47 | β | βc |
| Trp120* | Tyr48 | β | βc |
| <u>Glu121</u> | <u>Lys49</u> | β | βc |
| Ala122 | Ala50 | β | Ala121 |
| Arg123 | Lys51 | β | His122 |
| Ser124 | Asn52 | | Ser123 |
| Leu125 | | | Leu124 |
| Thr126 | Lys53 | distal loop | Thr125 |
| Thr127 | Val54 | | Thr126 |
| Gly128 | Gly55 | | Gly127 |
| Glu129 | Arg56 | | Gln128 |
| Thr130 | Glu57 | β | βd |
| Gly131* | Gly58 | β | βd |
| <u>Tyr132*</u> | Ile59 | β | βd |
| Ile133 | Ile60 | β | βd |
| Pro134* | Pro61 | β | Pro133 |
| Ser135 | Ala62 | | Ser134 |
| Asn136 | Asn63 | 3 ₁₀ -helix | Asn135* |
| <u>Tyr137*</u> | Tyr64 | | <u>Tyr136*</u> |
| Val138 | Val65 | β | Val137 |
| Ala139 | Gln66 | β | Ala138 |
| Pro140 | Lys67 | | Pro139 |
| Val141 | Arg68 | | Ser140 |
| Asp142 | | | |

Fig. 1. The sequences of the SH3 domains of human Fyn, human Csk and chicken Src. ¹The listed sequence covers the structured part of FynSH3 [16]. The fourteen conserved residues [17] are marked by an asterisk (*). The four conserved aromatic residues are in bold. The charged residues near the conserved Trp¹¹⁹ are underlined. ²The sequence of CskSH3. The tabulated sequence has been built in the electron density map. The charged residues near the conserved Trp⁴⁷ have been underlined. ³The sequence of the SrcSH3 domain, for which the NMR structure has been determined by Yu et al. [18]. The marked residues (*) have been identified by NMR to be close to the binding site of proline-rich peptides. The secondary structures of FynSH3 (molecule A) and CskSH3 (molecule A) have been calculated by the DSSP program [35]; the secondary structure of SrcSH3 is according to Yu et al. [18].

related to the presence of a phenylalanine in position 20. In virtually all other sequences of SH3 domains of tyrosine kinases a tyrosine is found at this site, which is the second tyrosine of the ALYDY motif. An exception is the SH3 domain of the tyrosine kinase Abl [6], where a phenylalanine is also found at this position. In FynSH3 this tyrosine is hydrogen-bonded to a conserved [17] aspartate side chain. This is Asp¹⁰⁰ in FynSH3 and Asp²⁷ in CskSH3, which are subsequently hydrogen-bonded to the main chain nitrogens of Arg⁹⁶ and Thr⁹⁷ in FynSH3 and of Thr²³ in CskSH3, respectively. In CskSH3 the Asp²⁷ side chain has adopted a somewhat different position, and the RT-Src loop residues have moved in a concerted manner. In addition in CskSH3, the Asp²⁷ side chain forms a hydrogen bond with the OG1 atom of Thr²³. It is interesting to point out that in AblSH3, which does have the conserved aspartate, there is a serine at the equivalent position of Thr²³. This suggests that the RT-Src loop of the AblSH3 domain could adopt a similar conformation to that seen in CskSH3.

The importance of the Asp²⁷ residue for positioning the RT-src loop might also explain the presence of a conserved (Fig. 1) glycine at position 58 (in strand d). Any side chain atom at this position would clash with main chain atoms of the residues near Asp²⁷, both in FynSH3 and CskSH3.

It has been speculated that a patch of conserved residues is important for the function of SH3. Within this patch two pairs of aromatic residues are observed [16]. In FynSH3 the first pair concerns residues Tyr⁹¹ and Tyr¹³⁷ (equivalent to Tyr¹⁸ and Tyr⁶⁴ in CskSH3). These two residues are almost universally conserved in the SH3 domains. The first tyrosine is the first tyrosine of the ALYDY motif and the second tyrosine is in the PxxY motif which is folded like the 3₁₀-helix. The second pair concerns Trp¹¹⁹ and Tyr¹³² in FynSH3 (equivalent to Trp⁴⁷ and Ile⁵⁹ in CskSH3). The tryptophan is also universally conserved, but at the position of the tyrosine other residues are observed, such as an isoleucine in CskSH3 or a phenylalanine in LckSH3 [6].

First we will discuss some structural properties of the region near the second pair. This pair is positioned between the RT-Src loop and the n-Src loop (see Figs. 2 and 5). Therefore, the surface properties of CskSH3 and FynSH3 near the conserved Trp⁴⁷ (Trp¹¹⁹ in FynSH3) are rather different: in addition to the mutation at position 59 and the different conformations of the RT-Src loop and the n-Src loop (Fig. 2), the charge distribution in this region is also very different. In FynSH3 it is predominantly negative (overall charge −5, Fig. 1), and in CskSH3 it is almost neutral (overall charge is −1, Fig. 1). The potential importance of charged residues in this region of the SH3 domain of p85α has been discussed by Booker et al. [20], and Koyama et al. [21]. If this region is indeed of importance for protein–protein interactions, then the functional properties of CskSH3 and

FynSH3 will be very different. This is actually in agreement with recent comparative binding studies of SH3 domains ([10]; Fumagalli and Courtneidge, unpublished observations).

The structural features near the first aromatic pair of CskSH3 and FynSH3 are more similar. Not only do the two tyrosine side chains of the FynSH3 and CskSH3

superimpose quite well (Fig. 5), but also the charge distribution on this side of the SH3 domain is similar. This tyrosine pair is at the edge of a region completely lacking in charged residues in CskSH3 and FynSH3, as is documented for CskSH3 in Fig. 4; the centre of this region is near the side chains of Ala⁶² and Tyr⁴⁸. This tyrosine pair is also close to the three conserved charged residues

Table 1: Crystallographic Data

Space group C2221, a=64.90 Å, b=74.69 Å, c=125.03 Å, α=β=γ=90
4 molecules per asymmetric unit

AUCN : KAu(CN)₂
HG : Methyl Mercury Chloride

PTEDA : Pt Ethylenediamine dichloride
URAC : Uranyl Acetate

data collection :

| dataset | native 1 | native 2 | AUCN | PTEDA | HG | URAC |
|----------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| total obs. | 21116 | 60252 | 21552 | 19749 | 19978 | 17027 |
| unique refl. | 7006 | 10807 | 6668 | 6750 | 7120 | 5838 |
| R-merge ^δ | 5.5 % | 8.3 % | 8.9 % | 9.2 % | 7.7 % | 10.0 % |
| overall compl. | 87.7 % | 94.8 % | 76.6 % | 77.2 % | 82.7 % | 74.5 % |
| resolution | 25 - 2.77 Å | 25 - 2.50 Å | 25 - 2.70 Å | 25 - 2.70 Å | 25 - 2.70 Å | 25 - 2.80 Å |

Phasing statistics (to 3.2 Å):

mean figure of merit (acentric) : 0.60 for 4289 reflections
mean figure of merit (centric) : 0.74 for 720 reflections

| | AUCN | PTEDA | HG | URAC |
|--|-----------|-----------|-----------------------------|-----------|
| R-difference [†] | 27.1 % | 14.9 % | 39.6 % | 46.2 % |
| Number of sites | 2 | 1 | 4 | 1 |
| R-cullis [‡] | 0.65 | 0.92 | 0.69 (7.0 to 3.2 Å) | 0.88 |
| Phasing power [‡] (acentric/centric) | 1.2 / 1.0 | 0.3 / 0.3 | 1.4 / 1.1 (7.0 to 3.2 Å) | 0.8 / 0.6 |
| cell dimensions a (in Å) | 65.19 | 65.22 | 64.75 | 65.14 |
| b | 74.27 | 74.75 | 74.49 | 74.26 |
| c | 125.87 | 125.41 | 125.04 | 126.12 |

Refinement statistics :

number of reflections in refinement : 9913 (8.0 to 2.5 Å)

completeness (8.0 - 2.6 Å) : 96 %
(2.6 - 2.5 Å) : 85 %

protein atoms : 1824

solvent atoms : 8

R-factor [♦] : 22.4 %

free R-factor : 27.8 %

geometric parameters :

rms deviation from ideal bond lengths : 0.013 Å

rms deviation from ideal bond angles : 1.9°

χ₁-χ₂ imperfection values[‡] : 33°

residues with bad φ/ψ : 0

rms deltaB for bonded atoms : 4.3 Å²

average B-factor (main chain atoms, side chain atoms) 20.0 Å², 23.5 Å²

$$\diamond R_{\text{merge}} = \frac{\sum_h \sum_i |I_h - I_{h,i}|}{\sum_h \sum_i I_h} \quad \dagger R_{\text{difference}} = \frac{\sum |F_{\text{PH}} - F_{\text{P}}|}{\sum F_{\text{P}}} \quad \star R_{\text{factor}} = \frac{\sum |F_{\text{OBS}} - F_{\text{CALC}}|}{\sum |F_{\text{OBS}}|}$$

$$\ddagger R_{\text{cullis}} = \frac{\sum |F_{\text{PH}} \pm F_{\text{P}} - F_{\text{H}}|}{\sum |F_{\text{PH}} \pm F_{\text{P}}|} \quad \text{for centric reflections} \quad \ddagger \text{ Phasing power} = \frac{\text{r.m.s.}(F_{\text{H}})}{\text{r.m.s.}(E)}$$

where F_P, F_{PH}, F_H are the native, derivative and heavy atom structure factor amplitudes respectively, E the lack of closure, and F_{OBS} and F_{CALC} the observed and calculated structure factor amplitudes.

‡ The χ₁-χ₂ imperfection value(*) is the r.m.s. difference between observed χ₁-χ₂ values and the nearest preferred cluster values, as observed in a database of well refined structures.

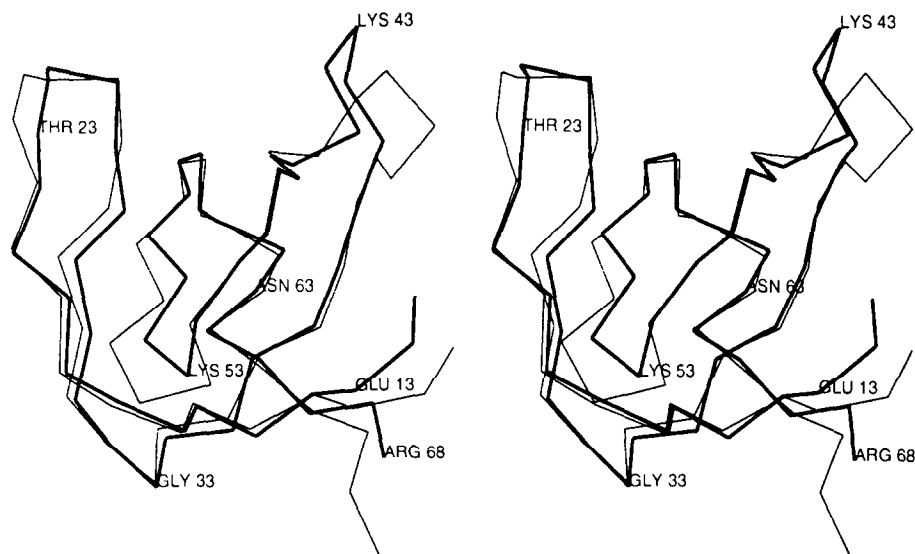


Fig. 2. The CA traces of the superposed CskSH3 and FynSH3. The CA trace of CskSH3 is in thick lines. The CA trace of FynSH3 is in thin lines. The CskSH3 CA trace is labelled. The RT-Src loop is near residue Thr²³, the n-Src loop is near Lys⁴³, the distal loop is near Lys⁵³, and the 3_{10} -helix is near Asn⁶³.

[17]; in FynSH3 these residues are Asp⁹², Asp¹⁰⁰, and Glu¹⁰⁷, in CskSH3 it concerns residues Asn¹⁹, Asp²⁷, and Asp³⁴. Only the side chain of residue Asn¹⁹ is pointing into the solvent, the residues Asp²⁷ (see above) and Asp³⁴ are more buried and seem to be of importance for the stability of the SH3 domain. Asp³⁴ is the last residue of a type II β -turn (of residues Cys³¹-Lys³²-Gly³³-Asp³⁴), just before strand b. A conserved hydrogen bond is observed between OD2(Asp³⁴) and N(Cys³¹). Asp³⁴ is preceded by a conserved glycine [17], which is very common

at position $i+2$ of a type II β -turn. The phi and psi angles of Gly³³ are 98° and -25° , respectively; such values are not allowed for non-glycine residues.

SH3 molecules are believed to be involved in protein-protein interactions. Therefore it is of interest to study the crystal contacts. The four molecules occur as a dimer of dimers; they are related to each other by three local twofold axes, but there are only two independent local twofold axes: one within a dimer and one between the two dimers. Molecules D,A and B,C form the two

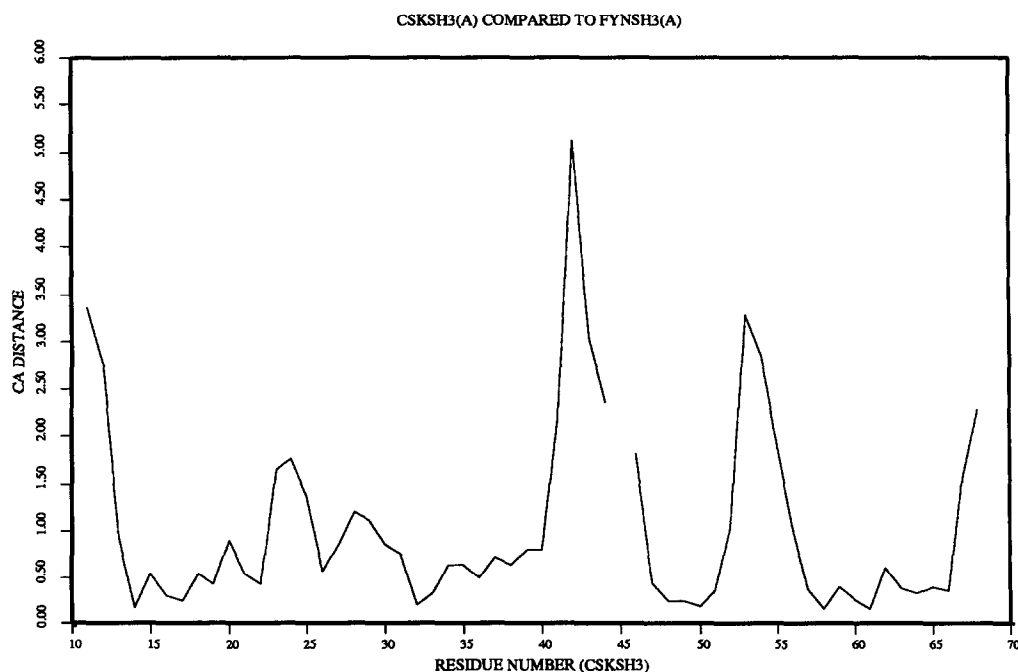


Fig. 3. The CA-distance plot of CskSH3 and FynSH3. On the y-axis are shown the distances between equivalent CA atoms after superposition. On the horizontal axis are plotted the residue numbers of CskSH3. The curve is interrupted at position 45 due to the insertion in CskSH3.

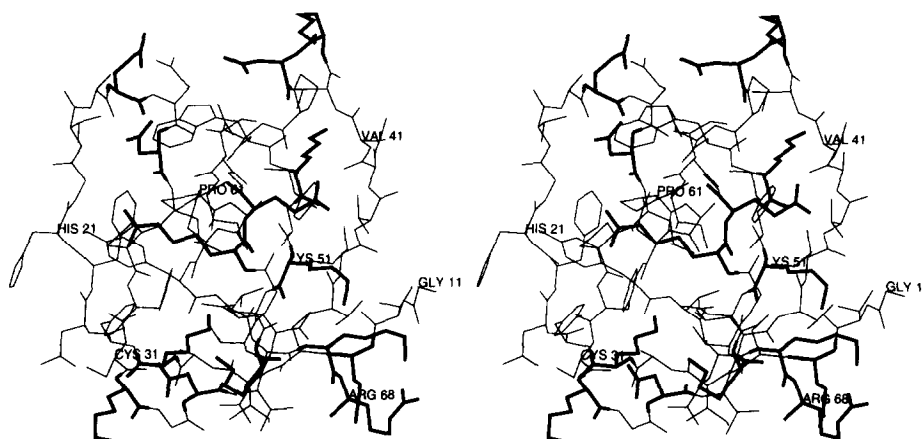


Fig. 4. An all atom representation of CskSH3. The charged residues are in bold.

dimers. There are approximately 80 atom–atom contacts (within 4 Å) between the molecules of the dimer and approximately 30 atom–atom contacts across the local twofold relating the two dimers. All other crystal contacts are less extensive. The interaction at the D,A interface (within the dimer) is predominantly an interaction between the RT-Src loops, whereas the interaction at the A,B interface (between the dimers) is between the n-Src loops. A prominent feature of the D,A interface is the stacking of a histidine side chain and a tyrosine side chain; this concerns the stacking interactions, at a distance of approximately 3.7 Å, between HisD²¹, TyrA¹⁸ and TyrD¹⁸, HisA²¹ in the first dimer. A statistical analysis of side chain–side chain interactions in 62 high resolution protein structures [34] shows that this stacking is seen more frequently than expected from a random distribution, suggesting that it is an energetically favourable interaction.

The crystal packing of the CskSH3-molecules is very different from the packing of the FynSH3-molecules in the FynSH3 crystals, however there is one common fea-

ture: in both crystals a tight stacking interaction between a histidine and a tyrosine is observed. In FynSH3 this concerns His¹⁰⁴ and Tyr⁹¹ (Fig. 5). There are no other common interactions at this crystal contact but the approximate positions of the two interacting molecules are the same and both in the FynSH3 and in the CskSH3 crystals the interactions at this region are rather extensive (more than 50 atom–atom contacts within 4.0 Å).

Summarizing, the structural diversity near the conserved Trp⁴⁷ correlates with the diversity of binding properties of the SH3 domains [10], which supports the hypothesis that this region is involved in interactions with other proteins. Possibly the aromatic residues of the first pair (Tyr¹⁸ and Tyr⁶⁴) are important for the interactions with the other domains of the same enzyme. Further biochemical and structural studies are required to establish the mode of interactions.

Acknowledgements: The coordinates of CskSH3 will be deposited in the Brookhaven Protein Data Bank. We thank Manfred Koegl, Andrea Musacchio and Matti Saraste for their advice and interest.

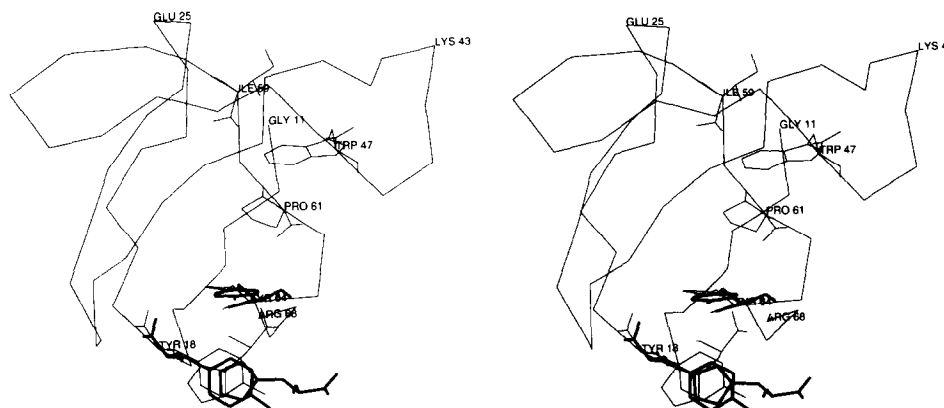


Fig. 5. The stacking interactions between the side chains of a histidine and a tyrosine as observed in crystal contacts in the crystals of CskSH3 and FynSH3. In thin lines are shown the CA trace of CskSH3, as well as the residues Tyr¹⁸, Tyr⁶⁴ (the first pair) and His²¹ of a neighbouring molecule. The residues Trp⁴⁷ and Ile⁵⁹ of CskSH3 mark the positions of the second pair; Glu²⁵ is in the RT-Src loop and Lys⁴³ is in the n-Src loop. In thick lines are shown Tyr⁹¹, Tyr¹³⁷ as well as His¹⁰⁴ of the neighbouring molecule, as observed in the crystals of FynSH3.

References

- [1] Okada, M., Nada, S., Yamanashi, Y., Yamamoto, T. and Nakagawa, H. (1991) *J. Biol. Chem.* 266, 24249–24252.
- [2] Imamoto, A. and Soriano, P. (1993) *Cell* 73, 1117–1124.
- [3] Nada, S., Yagi, T., Takeda, H., Tokunaga, T., Nakagawa, H., Ikawa, Y., Okada, M. and Aizawa, S. (1993) *Cell* 73, 1125–1135.
- [4] Nada, S., Okada, M., MacAuley, A., Cooper, J.A. and Nakagawa, H. (1991) *Nature* 351, 69–72.
- [5] Partanen, J., Armstrong, E., Bergman, M., Makela, T.P., Hirvonen, H., Huebner, K. and Alitalo, K. (1991) *Oncogene* 6, 2013–2018.
- [6] Musacchio, A., Gibson, T., Lehto, V.-P. and Saraste, M. (1992) *FEBS Lett.* 307, 55–61.
- [7] Pawson, T. and Schlessinger, J. (1993) *Curr. Biol.* 3, 434–442.
- [8] Cicchetti, P., Mayer, B.J., Thiel, G. and Baltimore, D. (1992) *Science* 257, 803–806.
- [9] Ren, K., Mayer, B.J., Cicchetti, P. and Baltimore, D. (1993) *Science* 259, 1157–1161.
- [10] Gout, I., Dhand, R., Hiles, I.D., Fry, M.J., Panayotou, G., Das, P., Truong, O., Totty, N.F., Hsuan, J., Booker, G.W., Campbell, I.D. and Waterfield, M.D. (1993) *Cell* 75, 25–36.
- [11] Superti-Furga, G., Fumagalli, S., Koegl, M., Courtneidge, S.A. and Draetta, G. (1993) *EMBO J.* 12, 2625–2634.
- [12] Okada, M., Howell, B., Broome, M.A. and Cooper, J.A. (1993) *J. Biol. Chem.* 268, 18070–18075.
- [13] Murphy, S.M., Bergman, M. and Morgan, D.O. (1993) *Mol. Cell. Biol.* 13, 5290–5300.
- [14] Kuriyan, J. and Cowburn, D. (1993) *Curr. Opin. Struct. Biol.* 3, 828–837.
- [15] Musacchio, A., Wilmanns, M. and Saraste, M. (1994) *Prog. Biophys. Mol. Biol.* (in press).
- [16] Noble, M.E.M., Musacchio, A., Saraste, M., Courtneidge, S.A. and Wierenga, R.K. (1993) *EMBO J.* 12, 2617–2624.
- [17] Musacchio, A., Noble, M., Pauptit, R., Wierenga, R.K. and Saraste, M. (1992) *Nature* 359, 851–855.
- [18] Yu, H., Rosen, M.K., Shin, T.B., Seidel-Dugan, C., Brugge, J.S. and Schreiber, S.L. (1992) *Science* 258, 1665–1668.
- [19] Kohda, D., Hatanaka, H., Odaka, M., Mandiyan, V., Ullrich, A., Schlessinger, J. and Inagaki, F. (1993) *Cell* 72, 953–960.
- [20] Booker, G.W., Gout, I., Downing, A.K., Driscoll, P.C., Boyd, J., Waterfield, M.D. and Campbell, I.D. (1993) *Cell* 73, 813–822.
- [21] Koyama, S., Yu, H., Dalgarno, D.C., Shin, T.B., Zydowsky, L.D. and Schreiber, S.L. (1993) *Cell* 72, 945–952.
- [22] Bullock, W.O., Fernandez, J.M. and Short, J.M. (1987) *BioTechniques* 5, 376–378.
- [23] Rosenberg, A.H., Lade, B.N., Chui, D.-S., Lin, S.-W., Dunn, J.J. and Studier, F.W. (1987) *Gene* 56, 125–135.
- [24] Studier, F.W. and Moffatt, B.A. (1986) *J. Mol. Biol.* 189, 113–130.
- [25] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: *Molecular Cloning*, Cold Spring Harbor Laboratory Press, NY.
- [26] Crowther, R.A. (1972) in: *The Fast Rotation Function*, (Rossmann, M. ed.) pp. 173–177, Gordon and Breach, New York.
- [27] Tong, L. and Rossmann, M.G. (1990) *Acta Crystallogr.* A46, 783–792.
- [28] CCP4 (1979) The S.E.R.C. (U.K.) Collaborative computing project no. 4, a suite of programs for Protein Crystallography, distributed from Daresbury Laboratory, Warrington, WA4 4AD, U.K.
- [29] Jones, T.A., Zou, J.-Y., Cowan, S.W. and Kjeldgaard, M. (1991) *Acta Crystallogr.* A47, 110–119.
- [30] Cowan, S.W., Newcomer, M.E. and Jones, T.A. (1993) *J. Mol. Biol.* 230, 1225–1246.
- [31] Vellieux, F.M.D., Hajdu, J., Verlinde, C.L.M.J., Groendijk, H., Read, R.J., Greenhough, T.J., Campbell, J.W., Kalk, K.H., Littlechild, J.A., Watson, H.C. and Hol, W.G.J. (1993) *Proc. Natl. Acad. Sci. USA* 90, 2355–2359.
- [32] Brünger, A.T. (1992) in: *X-Plor: Version 3.1, A System for X-ray Crystallography and NMR* Yale University Press, New Haven and London.
- [33] Vriend, G. (1990) *J. Mol. Graph.* 8, 52–56.
- [34] Singh, J. and Thornton, J.M. (1992) *Atlas of Protein Side-Chain Interactions*, pp. 748–787, Oxford University Press, Oxford.
- [35] Kabsch, W. and Sander, C. (1983) *Biopolymers* 22, 2577–2637.