Journal of Biomolecular NMR (2005) 31: 357–361
DOI 10.1007/s10858-005-0946-6

NMR structure note

Solution structure of the Src homology 2 domain from the human feline sarcoma oncogene Fes

Anna Scott[a,f], David Pantoja-Uceda[a], Seizo Koshiba[b], Makoto Inoue[b], Takanori Kigawa[b], Takaho Terada[b,c], Mikako Shirouzu[b,c], Akiko Tanaka[b], Sumio Sugano[d], Shigeyuki Yokoyama[b,c,e] & Peter Güntert[a,*]

[a]*Tatsuo Miyazawa Memorial Program, RIKEN Genomic Sciences Center, Tsurumi, Yokohama, 230-0045, Japan;* [b]*RIKEN Genomic Sciences Center, Tsurumi, Yokohama, 230-0045, Japan;* [c]*RIKEN Harima Institute at SPring-8, Sayo-gun, Hyogo, 679-5148, Japan;* [d]*Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, 108-8639, Japan;* [e]*Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, Tokyo, 113-0033, Japan;* [f]*Present address: Department of Biochemistry, University of Utah, Salt Lake City, UT 84103-3201, USA*

Received 29 December 2004; Accepted 11 January 2005

Key words: feline sarcoma oncogene, Fes, fes/fps, protein structure, SH2 domain, structural genomics

Biological context

The human feline sarcoma oncogene (Fes) protein is a non-receptor tyrosine kinase and a multi-domain protein that comprises Fps/Fes/Fer/CIP4 homology, Src homology 2 (SH2), coiled coil and tyrosine kinase domains. Fes was originally isolated as a retroviral oncogene in avian and feline retroviruses. Genetic analysis identified its cellular homologs, fes/fps. Fes is involved in the growth and differentiation of myeloid hematopoietic cells, vascular endothelial cells and neurons. It is also implicated in the regulation of cytoskeletal rearrangement (Takashima et al., 2003).

Fes adopts an inactive conformation *in vitro*. Its active conformation is oligomeric, requiring the N-terminal coiled-coil regions for self-association. Deletion of the SH2 domain in Fes greatly reduces the kinase domain activity *in vitro* and inhibits Fes autophosphorylation (Rogers et al., 2000). Thus, the Fes SH2 domain acts in a negative regulatory fashion by binding the Fes kinase domain.

SH2 domains bind phosphotyrosine-containing peptides. The Fes SH2 domain was found to interact most tightly with peptides containing the sequence pattern pYEX(V/I) (X = any residue). The binding of the peptides is described in terms of +1, +2 and +3 positions relative to phosphotyrosine. The strongest selectivity was observed for Glu at the +1 position and a slight preference for Val or Ile in the +3 position. This sequence pattern is very similar to the consensus sequence, pYEEI, found for the Src SH2 domain (Songyang et al., 1994).

SH2 domains typically regulate the kinase activity of their full-length proteins by phosphotyrosine-dependent interactions with kinase substrates and/or with kinase domains of their own protein. To better understand how the SH2 domain regulates Fes activity, a chimera was constructed in which the native Fes SH2 domain was replaced by that of Src (Rogers et al., 2000; Takashima et al., 2003). With the Src SH2 domain, the Fes kinase is constitutively active showing both strong kinase activity and a growth-inhibitory phenotype. This indicates that even though both the Fes SH2 domain and the Src SH2 domain bind a similar phosphotyrosine peptide, these domains behave differently in the context of the full-length Fes protein. The chimera was also tested in the yeast *Saccharomyces cerevisiae* to investigate the possibility that the regulation of the kinase activity is due to interactions of the SH2 domain with cellular proteins (Rogers et al., 2000; Takashima et al., 2003). The

*To whom correspondence should be addressed. E-mail: guentert@gsc.riken.jp

chimera exhibited the same phenotypes in yeast as it did in human cells, indicating that the regulatory interaction is within the full-length Fes. The structures of both the Fes and Src SH2 domains provide the structural basis for understanding how these two domains operate differently in the context of the full length Fes protein.

Methods and results

The Fes SH2 domain was chosen as a target for solution structure determination by the RIKEN Structural Genomics/Proteomics Initiative (Yokoyama et al., 2000). A single 1.2 mM uniformly ^{13}C and ^{15}N- labeled sample was prepared in 20 mM Tris-HCl buffer, pH 7.0, 100 mM NaCl, 1 mM dithiothreitol, 0.02% NaN_3 with the addition of D_2O to 10% v/v as described earlier (Scott et al., 2004). The protein sample used for the NMR measurements comprised 114 amino acid residues, numbered 1–114 throughout this paper. Residues 8–108 constitute the Fes SH2 domain and correspond to residues 450–550 in the full-length Fes protein. The non-native flanking sequences of residues 1–7 and 109–114 are related to the expression and purification system.

NMR experiments were performed at 25 °C on Bruker DRX 600 or Bruker AV 800 spectrometers operating at 600 and 800 MHz proton frequency, respectively (Figure 1). Two two-dimensional and 11 three-dimensional NMR experiments were recorded with the uniformly $^{13}C/^{15}N$-labeled sample for the determination of the sequence-specific polypeptide backbone and side-chain chemical shift assignments (Scott et al., 2004). For the collection of conformational constraints a 3D ^{15}N-edited NOESY-HSQC and two ^{13}C-edited NOESY-HSQC spectra (Muhandiram et al., 1993), one for aliphatic and one for aromatic ^{13}C, were recorded with 80 ms mixing time at 800 MHz proton frequency. The time-domain data were multiplied by a squared sine-bell window function with the corresponding shift optimized for every spectrum and zero-filled to at least twice the acquired number of data points prior to Fourier transformation. Baseline correction was applied in all dimensions. The raw NMR data were processed on a Linux computer using the program NMRPipe (Delaglio et al., 1995). The program NMRView (Johnson and Blevins, 1994) was used for interactive spectrum analysis.

Excluding the unassigned purification tag residues 1–6, all non-labile ^{1}H chemical shifts and all

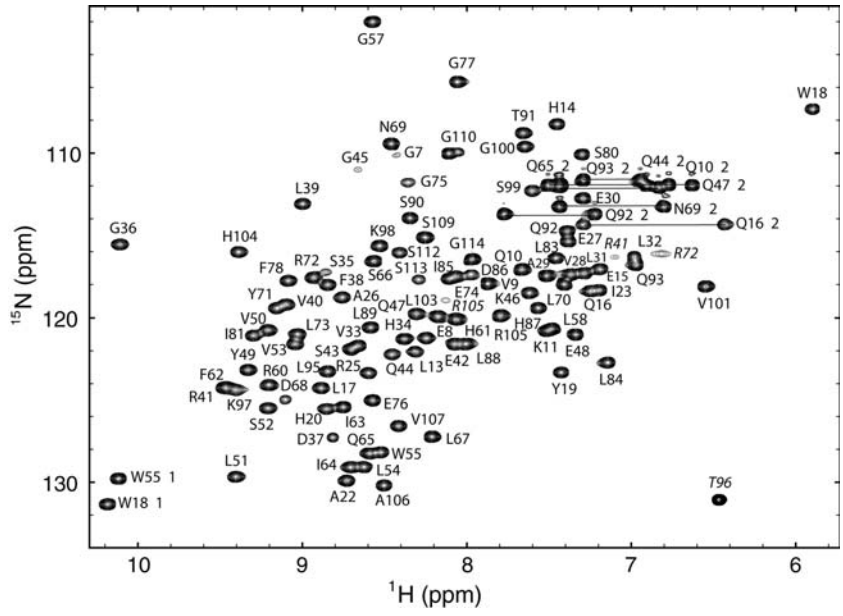

Figure 1. [^{1}H, ^{15}N]-HSQC spectrum of the human Fes SH2 domain recorded at 600 MHz and 25 °C using 1.2 mM uniformly ^{13}C and ^{15}N-labeled protein, 20 mM Tris-HCl, pH 7.0, 100 mM NaCl, 1 mM dithiothreitol, 0.02% NaN_3, 10% D_2O (v/v). Folded peaks are labeled in italics and cross-peaks connected by horizontal lines correspond to side chain NH_2 groups of Asn and Gln residues.

backbone amide ^{1}H chemical shifts except H^{N} of Asp56 and H^{N} of Val102 could be assigned (Scott et al., 2004). All hydrogen-bound ^{15}N resonances are assigned except for N^{ε}/N^{η} of Arg25, N^{ε}/N^{η} of Arg60 and N^{ζ} of the 4 Lys residues. All hydrogen-bound ^{13}C resonances and all backbone carbonyl carbon resonances are assigned. The hydroxyl proton resonance of Ser99 could be observed and assigned. Overall, the assignments cover more than 97% of the chemical shifts of the non-labile protons and backbone amide protons. The chemical shift assignments are available from the BioMagResBank (BMRB entry 6331).

Peak lists for the NOESY spectra were generated by interactive peak picking and peak volumes determined by the automatic integration function of NMRView (Johnson and Blevins, 1994). For all Xxx-Pro bonds, the *trans* conformation was confirmed independently by strong sequential H^{α}–H^{δ} NOESY cross peaks, and by the $^{13}C^{\beta}$ and $^{13}C^{\gamma}$ chemical shift differences (Schubert et al., 2002). The three-dimensional structure was determined by combined automated NOESY cross peak assignment (Herrmann et al., 2002) and structure calculation with torsion angle dynamics (Güntert et al., 1997) implemented in the program CYANA (Güntert, 2003). The standard CYANA protocol of seven iterative cycles of NOE assignment and structure calculation, followed by a final structure calculation was applied. Stereo-specific assignments for 55 isopropyl and methylene groups were determined by the GLOMSA method (Güntert et al., 1991) prior to the final structure calculation by analyzing the structures obtained in the preceding, seventh NOE assignment/structure calculation cycle. Weak constraints on (ϕ,ψ) torsion angle pairs and on side-chain torsion angles between tetrahedral carbon atoms were used temporarily during the NOE assignment/structure calculation cycles in order to favor allowed regions of the Ramachandran plot and staggered rotamer positions, respectively. In each cycle, the structure calculation started from 100 randomized conformers and the standard CYANA simulated annealing schedule (Güntert et al., 1997) was used with 10,000 torsion angle dynamics steps per conformer. The 20 conformers with the lowest final CYANA target function values were subjected to restrained energy-minimization in a water shell with the program OPALp (Koradi et al., 2000), using the AMBER force field (Cornell et al., 1995). PROCHECK-NMR (Laskowski et al., 1996) was used for validation, and MOLMOL (Koradi et al., 1996) and PyMOL (DeLano Scientific) for visualization of the structures.

Statistics about the quality and precision of the 20 energy-minimized conformers (Figure 2a) that represent the solution structure of SH2 domain of the human Fes protein are summarized in Table 1. The structures are well defined and show excellent agreement with the experimental data. About 97.3% of the 4732 cross peaks that had been identified in the three 3D NOESY spectra could be assigned be the program CYANA (Herrmann et al., 2002; Güntert, 2003). About 21 NOE distances constraints per residue, including 905 long-range distance constraints between protons five or more residues apart in the sequence, were used in the final structure calculations with CYANA, and in the subsequent restrained energy-refinement with the program OPALp (Koradi et al., 2000). The 20 energy-minimized conformers of the SH2 domain from human Fes with the lowest CYANA target function values in the final structure calculation were deposited in the Protein Data Bank (PDB entry 1WQU).

The solution structure of the SH2 domain from human Fes (Figure 2b) consists of a central anti-parallel β-sheet, which constitutes the core of the protein and is comprised of β-strands βB (Asp37–Glu42), βC (Tyr49–Trp55) and βD (Leu58–Ile63), flanked on one side by helix αA (Arg25–Leu31) and on the other side by helix αB (Ile81–Thr91). Three smaller anti-parallel β-strands βD′ (Gln65–Leu67), βE (Leu70–Arg72) and βF (Gly77–Phe78) pack on the αB helix. The N- and C-termini are located close to βB, where two short strands βA (Tyr19–Gly21) and βG (Arg105–Ala106) form hydrogen bonds with βB and cap the hydrophobic core of the domain. The standard nomenclature for SH2 domain secondary structure elements is used (Eck et al., 1993).

Discussion and conclusions

The Fes SH2 domain overlays with that of Src (Figure 2c) with a RMSD of 2.0 Å and a DALI Z score of 11.9 (Holm and Sander, 1996). The two largest differences between the backbone folds are that the Src SH2 domain has extended βC and βD strands compared to the Fes domain and a shorter

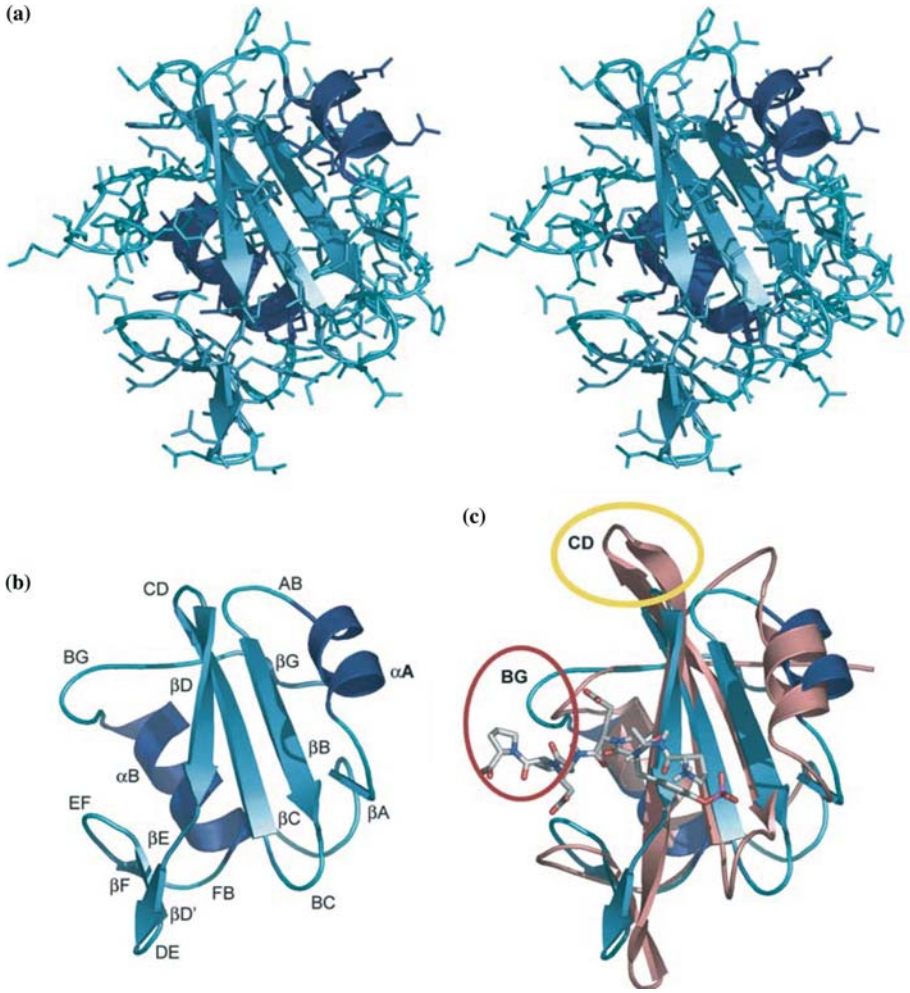

Figure 2. (a) Stereo-view of the solution structure of the Fes SH2 domain. (b) Ribbon diagram of the Fes SH2 domain indicating the secondary structure elements (Eck et al., 1993). (c) Superposition of the structures of the Fes SH2 domain (light blue and dark blue) and the Src SH2 domain (PDB code 1SPS; light orange) with the pYEEI peptide bound to the Src SH2 domain (Waksman et al., 1993). The yellow circle highlights the extended βC and βD strands in Src. The red circle indicates the longer BG loop in Fes and the area of a possible steric clash between the bound peptide and the BG loop of Fes.

BG loop. Structural studies on Src with and without bound peptide show that the SH2 domain undergoes only minor conformational changes upon the binding of the peptide (Waksman et al., 1993). The most pronounced movement is of the BC loop, also referred to as the phosphate binding loop; however, there are also small changes in the hydrophobic binding pockets. A comparison with the structure of the Src SH2 domain–peptide complex (Figure 2c) shows that the BG loop of the Fes SH2 domain would interfere strongly with the +3 peptide position in the Src complex suggesting that either the BG loop must undergo a conformational change upon binding of the peptide or that another binding mode exists. Future binding studies with the Fes SH2 domain and other SH2 domains for which NMR structures have been solved recently at RIKEN will be able to shed light on these possibilities.

Src and Fes show a high affinity for the same phosphotyrosine-containing peptide, pYEX(V/I). However, the Src SH2 domain is not able to functionally substitute for the Fes SH2 domain in the full-length Fes protein (Rogers et al., 2000; Takashima et al., 2003). Given the similar peptide specificity, it is unlikely that the difference in

Table 1. Statistics for the NMR solution structure of human Fes SH2 domain

NOE distance constraints	
Short-range, $\|i-j\| \leq 1$	1007
Medium-range, $1 < \|i-j\| < 5$	379
Long-range, $\|i-j\|5$	905
Total	2291
Maximal violation	0.15 Å
Number of violations >0.2 Å	0
Final CYANA target function value[a]	0.47 Å^2
AMBER energy[b]	−4047 kcal/mol
RMS deviations from ideal geometry	
Bond lengths	0.014 Å
Bond angles	1.81°
RMSD to mean coordinates	
Backbone N, C^{α}, C′ of residues 8–108	0.44 Å
All heavy atoms of residues 8–108	0.93 Å
Ramachandran plot statistics[c]	
Most favorable regions	81.4%
Additional allowed regions	17.6%
Generously allowed regions	1.0%
Disallowed regions	0.0%

[a]The final CYANA target function value was computed for the structure before energy minimization with OPALp.
[b]Average values over the 20 final energy-minimized CYANA conformers.
[c]PROCHECK-NMR (Laskowski et al., 1996).

function is due to the phosphotyrosine binding site, however, there are clearly differences in the binding pocket. The electrostatic surface potential of the Fes and Src SH2 domains are similar in the peptide binding pocket but overall Fes has a predominantly negative surface potential, whereas that of Src is mainly positive (data not shown). This difference in the electrostatic surface potentials could account for the inability of the Src SH2 domain to negatively regulate Fes kinase activity. How the Fes SH2 domain interacts with the other Fes domains is not yet known. Nevertheless, a large surface charge difference between the Fes and Src SH2 domains could prevent the Src SH2 domain from binding the Fes kinase domain. This would force Fes into a more 'open' and active conformation or, possibly, energetically favor oligomerization of the Fes protein via the coiled coil regions, thus causing its activation. The present work provides the structural basis for further biochemical experiments to test various hypotheses about how Fes regulation differs from that of Src.

Acknowledgements

A.S. was supported by the National Science Foundation under Grant Number OISE-0413490, and by the Japan Society for the Promotion of Science. This work was supported by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT), and by the Tatsuo Miyazawa Memorial Program of RIKEN Genomic Sciences Center.

References

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995) *J. Am. Chem. Soc.*, **117**, 5179–5197.
Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.
Eck, M.J., Shoelson, S.E., and Harrison, S.C. (1993) *Nature*, **362**, 87–91.
Güntert, P. (2003) *Prog. NMR Spectrosc.*, **43**, 105–125.
Güntert, P., Braun, W., and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 517–530.
Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.
Herrmann, T., Güntert, P., and Wüthrich, K. (2002) *J. Mol. Biol.*, **319**, 209–227.
Holm, L., and Sander, C. (1996) *Science*, **273**, 595–603.
Johnson, B.A., and Blevins, R.A. (1994) *J. Biomol. NMR*, **4**, 603–614.
Koradi, R., Billeter, M., and Güntert, P. (2000) *Comput. Phys. Commun.*, **124**, 139–147.
Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graph.*, **14**, 51–55.
Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996) *J. Biomol. NMR*, **8**, 477–486.
Muhandiram, D.R., Farrow, N., Xu, G.Y., Smallcombe, S.H., and Kay, L.E. (1993) *J. Magn. Reson. Ser. B*, **102**, 317–321.
Rogers, J.A., Cheng, H.Y., and Smithgall, T.E. (2000) *Cell Growth Differ.*, **11**, 581–592.
Schubert, M., Labudde, D., Oschkinat, H., and Schmieder, P. (2002) *J. Biomol. NMR*, **24**, 149–154.
Scott, A., Pantoja-Uceda, D., Koshiba, S., Inoue, M., Kigawa, T., Terada, T., Shirouzu, M., Tanaka, A., Sugano, S., Yokoyama, S., and Güntert, P. (2004) *J. Biomol. NMR*, **30**, 463–464.
Songyang, Z., Shoelson, S.E., McGlade, J., Olivier, P., Pawson, T., Bustelo, R., Barbacid, M., Sabe, H., Hanafusa, H., Yi, T., Ren, R., Baltimore, D., Ratnofsky, S., Feldman, R.A., and Cantley, L.C. (1994) *Mol. Cell. Biol.*, **14**, 2777–2785.
Takashima, Y., Delfino, F.J., Engen, J.R., Superti-Furga, G., and Smithgall, T.E. (2003) *Biochemistry*, **42**, 3567–3574.
Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D., and Kuriyan, J. (1993) *Cell*, **72**, 779–790.
Yokoyama, S., Hirota, H., Kigawa, T., Yabuki, T., Shirouzu, M., Terada, T., Ito, Y., Matsuo, Y., Kuroda, Y., Nishimura, Y., Kyogoku, Y., Miki, K., Masui, R., Kuramitsu, S., Kurumizaka, H., Kawaguchi, S., Shibata, T., Kainosho, M., and Inoue, Y. (2000) *Nat. Struct. Biol.*, **7**, 943–994.