JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT°

ELSEVIER

COMMUNICATION

Crystal Structure of Human Coactosin-like Protein

Lin Liu[1,2]†, Zhiyi Wei[1,2]†, Yanli Wang[1,2], Mao Wan[2], Zhongjun Cheng[1,2] and Weimin Gong[1,2]*

[1]*National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101 People's Republic of China*

[2]*School of Life Sciences, Key Laboratory of Structural Biology, University of Science and Technology of China Hefei, Anhui 230026, People's Republic of China*

*Corresponding author

Human coactosin-like protein is an actin filament binding protein but does not bind to globular actin. It associates with 5-Lipoxygenase both *in vivo* and *in vitro*, playing important roles in modulating the activities of actin and 5-Lipoxygenase. Coactosin counteracts the capping activity of capping protein which inhibits the actin polymerization. We determined the crystal structures of human coactosin-like protein by multi-wavelength anomalous dispersion method. The structure showed a high level of similarity to ADF-H domain, although their amino acid sequences share low degree of homology. A few conserved hydrophobic residues that may contribute to the folding were identified. This structure suggests coactosin-like protein bind to F-actin in a different way from ADF/Cofilin family. Combined with the information from previous mutagenesis studies, the binding sites for F-actin and 5-Lipoxygenase were analyzed, respectively. These two sites are quite close, which might prevent F-actin and 5-Lipoxygenase from binding to coactosin simultaneously.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: human coactosin-like protein; crystal structure; F-actin; 5-Lipoxygenase; capping protein

Actin plays important roles in cell architecture, motility, phagocytosis, endocytosis and cytoplasmic streaming.[1] Its functions are modulated by a large number of actin-binding proteins (ABPs). The structures of actin and many of its binding proteins have been determined to high resolution using X-ray crystallography and NMR spectroscopy, whereas electron microscopy and image processing have established the interaction sites on F-actin for myosin and a range of actin-binding proteins.[2]

Coactosin is a 17 kDa actin-binding protein originally isolated from *Dictyostelium discoideum*.[3] It is revealed that coactosin is able to counteract the activity of capping proteins that retard actin polymerization, while coactosin itself has no effect on actin polymerization.[4]

The human version of coactosin named coactosin-like protein (CLP) shows a significant homology to coactosin with 33.3% identity and 75% homology in amino acid sequence. Human CLP nucleotide sequence was initially found as a sequence flanking a deletion on chromosome 17 characterizing the Smith–Magenis syndrome (SMS). The SMS critical region overlaps with a breakpoint cluster region associated with primitive neuro-ectodermal tumors, suggesting that the CLP gene is involved in DNA rearrangements of somatic cells.[5] CLP is also reported as a human pancreatic cancer antigen by SEREX method.[6]

CLP binds directly to filamentous-actin (F-actin) but does not form a stable complex with globular actin (G-actin). CLP binds to actin filaments with a stoichiometry of 1 : 2 (CLP : actin subunits), but could be cross-linked to only one subunit of actin.[7]

Human CLP was first obtained in a yeast two-hybrid screen using 5-Lipoxygenase (5LO), also an actin-binding protein, as a bait.[8] 5LO is a 78 kDa enzyme, the first enzyme in cellular leukotriene biosynthesis. 5LO catalyzes two-step conversion of arachidonic acids to leukotrienes, which are potent mediator of inflammation of allergy disorders including arthritis, asthma, and allergic reactions. In resting cells, 5LO is localized in soluble compartments, in the cytosol and/or within the nucleus. Upon activation, 5LO becomes associated with the nuclear membrane. The migration of 5LO is probably of the most importance for regulation of the cellular 5LO activity.[9] Modulation of

translocation and activation of 5LO may involve interactions with other proteins.

Structure overview

The crystal structure of human CLP is the first structure of coactosin/CLP family. The structure was solved by multi-wavelength anomalous dispersion (MAD) method using a seleno-methionine substituted protein crystal, and refined at 2.8 Å resolution (Table 1). There is only one methionine (Met70) in the recombinant human CLP protein. Mass spectroscopy experiment proved that the first Met was removed when expressed in *Escherichia coli* (data not shown). To our knowledge, this is one of the structures solved using MAD method with the most number of residues (148 residues tagged with six His residues in the C-terminal end) per selenium atom. The final models contain two monomers in the asymmetric unit. The C-terminal region seems very disordered, with no electron density fitting for residues 138–148 in chain A, and residues 132–148 in chain B, respectively.

The monomer structure consists of a six-stranded mixed β-sheet in which the four central strands (β2-β5) are anti-parallel and the two edge strands (β1 and β6) run parallel with the neighboring strands. The sheet is surrounded by two α-helices on each side (Figure 1). Such a structure is a typical ADF-H domain,[10] which is the core structure of many actin-binding proteins. The two monomers are almost structurally identical with an root-mean-square deviation (rmsd) of 0.65 Å for 130 residues (Figure 2(B)), related by a non-crystallographic

Table 1. Data collection, MAD phasing and refinement statistics

Data collection			
Space group	C2		
Unit cell parameters	$a = 124.0$ Å, $b = 37.0$ Å, $c = 60.3$ Å, $\beta = 106.8°$		
Data sets	Edge (0.9816 Å)	Peak (0.9813 Å)	Remote (0.9000 Å)
Resolution range (Å)	50–2.8 (2.87–2.8)	50–2.8 (2.87–2.8)	50–2.8 (2.87–2.8)
No. of total reflections	52027	77113	92275
No. of unique reflections	5856 (387)	5756 (377)	6039 (403)
I/σ	17.7 (5.9)	22.6 (9.2)	25.0 (9.8)
Completeness (%)	88.8 (88.6)	87.4 (85.7)	91.5 (91.6)
R_{merge} (%)[a]	5.6 (13.5)	5.0 (11.5)	4.9 (13.1)
MAD phasing (20–3Å)			
Mean FOM[b]	0.44		
Mean FOM[b] after density modification	0.70		
Structure refinement			
Resolution (Å)	50–2.8 (2.98–2.8)		
R_{cryst}/R_{free} (%)[c]	19.8 (29.8)/26.7 (37.5)		
No. of reflections			
Working set	5382		
Test set	657		
Average B-factor (Å²)			
Main-chain	34.4/37.3[d]		
Side-chain	33.5/37.2[d]		
Water	31.0		
No. of atoms			
Protein atoms	2051		
Water molecules	33		
Selenium atoms	2		
rmsd			
Bond distances (Å)	0.008		
Bond angles (deg.)	2.1		

The recombined human CLP was expressed in *E. coli* bacterial strain BL21(*DE3*) using expression vector pET22b(+) (Novagen), and was purified using affinity chromatography as a standard procedure. For phase determination, the recombinant plasmid was transferred into Met-auxotrophic strain B834(*DE3*) to obtain the seleno-methionyl derivative of human CLP protein. Crystals of CLP were obtained at 277 K by the hanging-drop vapor diffusion method from the protein sample (20 mg/ml protein in 50 mM Tris–HCl (pH 8.0), 5 mM NaCl, 2 mM imidazole) combined in a 1 : 1 ratio with a reservoir solution consisting of 30% (w/v) PEG2K, and 0.1 M Hepes (pH 7.0). A MAD data set was collected from a single seleno-methionine substituted CLP crystal at 100 K on Beijing Synchrotron Radiation Facility (BSRF) beamline 3W1A at the Institute of High Energy Physics, Chinese Academy of Sciences. All data were processed and scaled with the DENZO and SCALEPACK.[21] Two expected selenium positions were found in an asymmetric unit by SOLVE.[22] RESOLVE[23,24] also built a fragmented model of modest quality containing backbone atoms for about 50% of all residues. The program O[25] was used to rebuild and connect the fragments manually for the initial model. The model was refined against the high-energy remote data set in 50–2.8 Å resolution range by using CNS.[26] NCS restraints were applied through all stages of refinement aside from last cycle. The stereochemical quality of the final model was checked by PROCHECK.[27]

  [a] $R_{merge} = \Sigma|I_i - I_m|/\Sigma I_i$, where I_i is the intensity of the measured reflection and I_m is the mean intensity of all symmetry-related reflections.

  [b] Mean *FOM (figure of merit)* $= <|\Sigma P(\alpha)e^{i\alpha} \alpha/\Sigma P(\alpha)|$., where α is the phase and $P(\alpha)$ is the phase probability distribution. Numbers in parentheses represent the value for the highest resolution shell.

  [c] $R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure factors. $R_{free} = \Sigma_T||F_{obs}| - |F_{calc}||/\Sigma_T|F_{obs}|$, where T is a test data set of about 10% of the total reflections randomly chosen and set aside prior to refinement.

  [d] Values for the two different monomers (A and B) respectively in an asymmetric unit. Numbers in parentheses represent the value for the highest resolution shell.
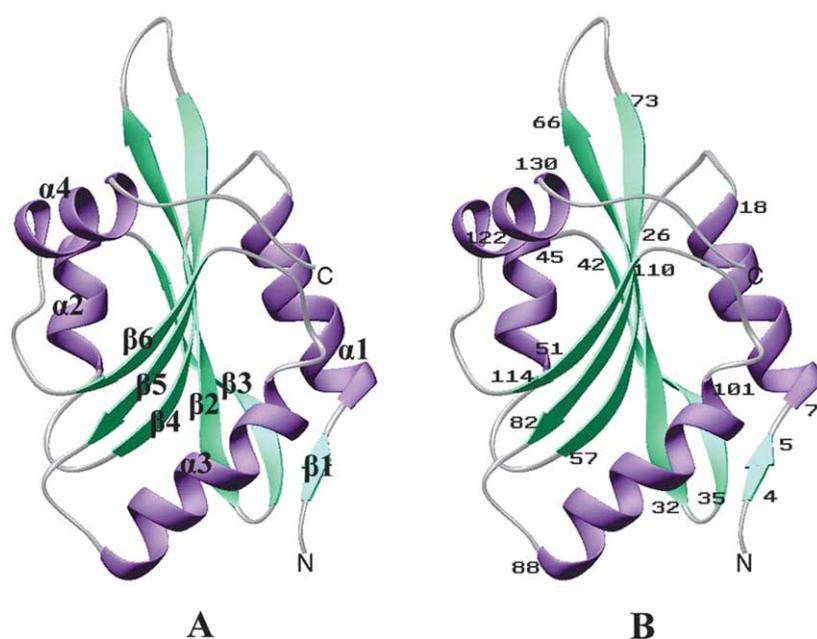
Figure 1. The stereo view of ribbon diagram of human CLP. The six-stranded mixed β-sheet is in purple; the helices are in light-green; and the connecting loops are in gray. Structural elements are labeled in the left diagram. Figures 1–3(B) were prepared using Ribbons.[17]

2-fold screw axis with a 32 Å translation along the axis (Figure 2(A)). Recently, the secondary structure of CLP determined by NMR has been reported.[11] NMR structures of human CLP (PDB code: 1WNJ) and mouse CLP (PDB code: 1UDM) were released in PDB. Superposition of these two solution structures and our crystal structure in backbone (Figure 2(B)) shows that the loop between β4 and β5 are obviously flexible, while the other regions keep highly similar. It has been reported that human CLP could exist both as a monomer and a dimer.[7] During the purification, we also observed human CLP always showed two bands with the molecular masses corresponding to the monomer and the dimer on SDS-PAGE, but gave a sharp single peak in mass spectroscopy with the monomer mass. The interactions between the two monomers in the crystal are only two hydrogen bonds, (carbonyl oxygen of Gly67-A to nitrogen of Gly33-B, and N^ε of Lys126-A to carbonyl oxygen of Asp32-B,) the biological significance of this packing needs further investigation.

Comparison with ADF-H proteins

A structural similarity search in the Protein Data Bank (PDB) with program DALI[12] indicates that human CLP shares highly homology with ADF/cofilin family (A/Cs) in secondary structures arrangement and peptide folding, with the conformational variations most occurring in the loops (Figure 3(B)), although the amino acid sequence identity between human CLP and A/Cs is as low as less than 15%. Another group of actin-binding proteins structurally similar to human CLP is gelsolin/villin family, which contains a common β-sheet and a long helix as CLP (Figure 3(B)).

Sequence alignment of human CLP with other CLP or coactosin, A/C family and gelsolin/villin family shows some conserved hydrophobic residues throughout the amino acid sequence (Figure 3(A)). These conserved hydrophobic residues interdigitate to form two hydrophobic cores (Phe29, Tyr31, Phe59, Phe61, Val101 and Val105 forming the first hydrophobic core, and Cys53, Trp81 and Leu120 forming the second hydrophobic core, Figure 3(B)), which are essential for stabilizing the similar folds of CLP, cofilin and gelsolin. This result provides key information to explain why CLP and ADF-H domains have low similarity in sequence but are highly conserved in three-dimensional structure.

Although CLP is similar to ADF-H domains in folding, the actin-binding models revealed by the putative cofilin-actin complex UNC-60B[13] could not be simply applied to CLP. Cofilin binds to both G-actin and F-actin, while CLP binds to F-actin only. Systematic mutagenesis studies suggest that the residues (except Lys75) whose counterparts involved in cofilin–actin binding have no effects on CLP–F-actin interactions. Our CLP structure may provide explanation on why CLP does not bind to G-actin. Compared with yeast cofilin, the delegate of AC family, the N terminus [1]MSRSG[5] of which Ser4 and Gly5 are highly conserved for both G-actin and F-actin binding is not at all conserved in human CLP with the N-terminal sequence [1]MATKI.[5] Another important residue for G-actin binding to cofilin, Arg96 located in the "kinked helix", is replaced with Leu89 in human CLP.

The "kinked" α-helix (α3 in CLP), which is considered to be the F-actin-binding region in both cofilin and gelsolin, is also kinked in CLP as in cofilin. Although residue G95 is in the middle of α3 in CLP, it would not be the fact causing the bend because it is not conserved in coactosin and AC family, and its dihedral angles are in the most favored regions in Ramachandran plot. In order to characterize the geometry of the "kinked" α-helices
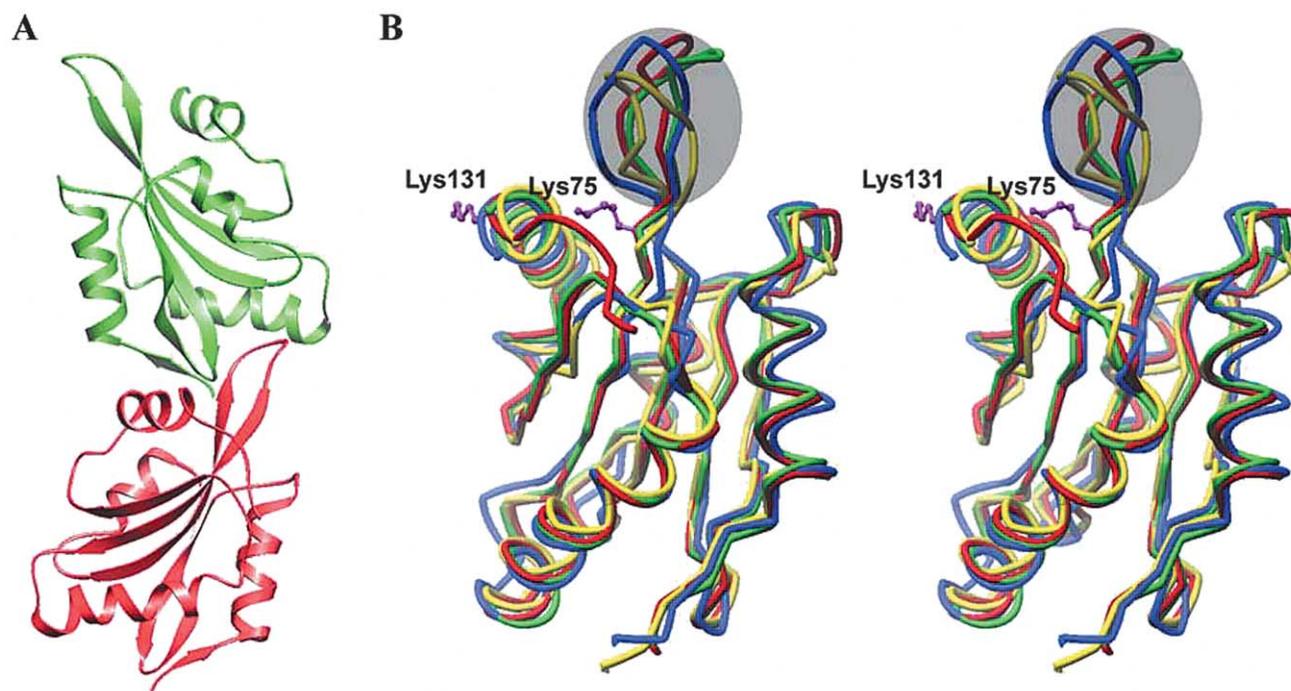
Figure 2. The two monomers of human CLP. (A) The ribbons diagram of two human CLP monomers packing in an asymmetric unit. The two monomers are related by a non-crystallographic 2-fold screw axis. (B) The stereo view of the superposition of the two monomers (green and red), NMR structure of human CLP (blue), and NMR structure of mouse CLP (yellow). The structural difference of these four structures, the loop connecting β4 and β5, are shadowed. Actin-binding site Lys75 and 5LO binding site Lys131 are showed in stick and ball model (purple). The each model of the two NMR structures used for superposition is the first model of 20 models. The N termini (residues before initial Met) and the C termini (residues after Ala132) of the selected models are cut before superposition, because the N termini do not exist in nature structure and the N- and C termini are highly disordered in the two NMR structures.
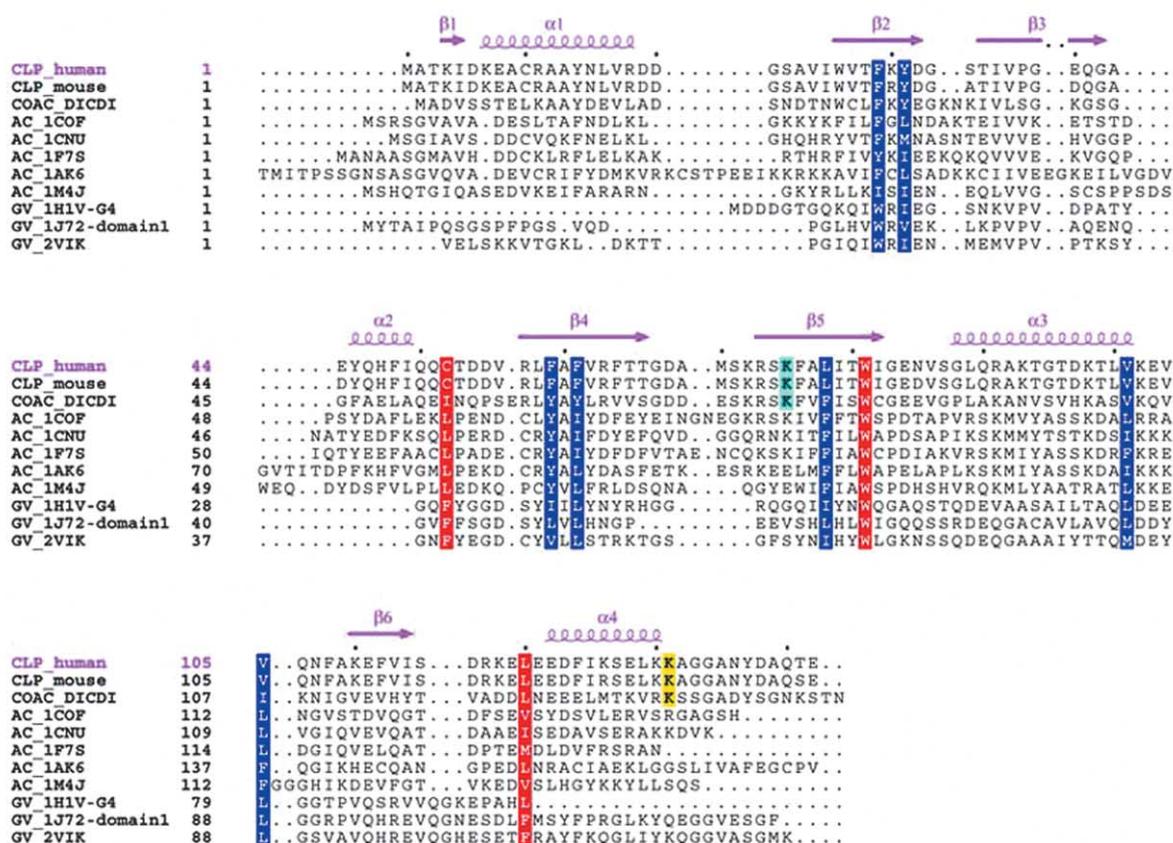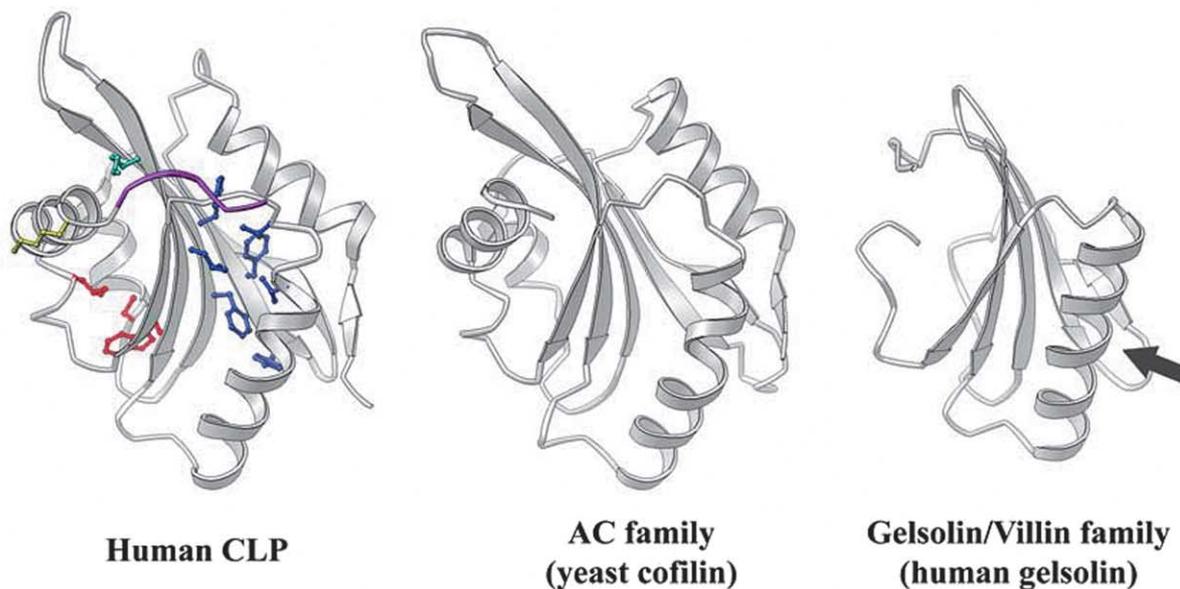
Figure 3. (A) Sequence alignment of human CLP and the proteins in ADF-H and gelsolin/villin families. The sequences of CLP from *Homo sapiens* (Human), CLP from *Mus musculus* (House Mouse), and coactosin from *D. discoideum* were aligned using ClustalW. The sequence of human CLP is aligned with AC family members: yeast cofilin (PDB code: 1COF), *Acanthamoeba* actophorin (PDB code: 1CNU), *Arabidopsis thaliana* ADF1 (PDB code: 1F7S), human destrin (PDB code: 1AK6), mouse ADF-H domain (PDB code: 1M4J), and gesolin family members: human gelsolin (PDB code: 1H1V) domain G4, human Cap-G (PDB code: 1J72) domain1, chicken villin (PDB code: 2VIK) mainly based on the structural element alignment gave by DALI. The secondary structure of human CLP, which is defined by the analysis of the structure using DSSP program, is indicated above the alignment. The conserved hydrophobic residues forming the two hydrophobic cores are boxed in blue, and red respectively. The essential actin and 5LO-binding sites are

in human CLP and AC family members, program HELANAL[16] was used to calculate the bending angle. The results show that the maximum bending angle of the "kinked" α-helix in human CLP (21.3°) is smaller than those in AC family members (∼40–50°). A critical role of maize ADF residues Tyr67 (strictly conserved in AC family) in proper protein folding has been demonstrated by mutating Tyr67 (equivalent to yeast cofilin Tyr64) to phenyl-alanine, which implicates that the strong hydrogen bond between the hydroxyl of Tyr64 (in yeast cofilin) and the carbonyl oxygen of Tyr101 (in yeast cofilin) is necessary for F-actin binding. The strong hydrogen bond would be helpful for stabilizing the kinked α-helix of AC family members. However, the corresponding position to yeast cofilin Tyr64 is Phe59 in human CLP (Figure 3(A)), which indicates that the smaller bending angle of human CLP is properly caused by missing the strong hydrogen bond conserved in AC family. In addition, the "kinked" α-helix of human CLP is shorter (14 residues) than that of structures of AC family and gelsolin/villin family (18–19 residues). Arg96 and Lys98 of yeast cofilin in this helix are involved in cofilin–actin interactions,[14] but the corresponding residues in human CLP are Leu89 and Arg91. There is no evidence indicating they are related to actin binding. These structural variations imply the actin-binding mode of human CLP is different from that in AC family.

F-actin and 5LO binding sites in CLP

Polar residues cover most area of the CLP surface, with few hydrophobic residues exposed, suggesting that CLP binds to other proteins with hydrogen bonds and/or salt bridges. The critical binding residues for F-actin and 5LO are Lys75 and Lys131, respectively.[7,8] Mutation of Arg73 also affects CLP binding to β-actin.[8] In the current structure, Lys75 is located in the bottom of a cleft formed by β5 and the C-terminal helix (α4) (Figure 2(B)). The C-terminal residues 133–137 of chain A, which could not be seen in chain B, cover the cleft and bury Lys75. If Lys75 interacts with F-actin directly, the relocation of the C-terminal peptide should be necessary. Besides Lys75 in the bottom of the cleft, Arg73 and Lys130 are on each side of the cleft. Based on this basic cleft, F-actin should interact with CLP with an acidic protrusion. By searching the Holmes–Lorenz model of F-actin,[19,20] two potential CLP binding regions, with their negative charged surface and protruding shape, were found in the F-actin molecule. The first is the N terminus of actin subdomain 1, which is a stretched peptide with the sequence of ^{1}DEDE.[4] The

second is the region of subdomain 1, which is consisted of ^{361}EYED^{364}. To be consistent, the biochemical studies on Lys75 also suggested the N-terminal peptide of actin subdomain 1 could act as a "fishing rod" to attracting the positive charged surfaces of actin-binding proteins.[7]

Lys131 lies on the surface of helix 4. The distance between the NZ atom of Lys75 to the C^α atom of Lys131 is 7.0 Å only. Since F-actin is an elongated structure and 5LO is a 78kD large protein, Lys75 and Lys131 are so spatially close that the steric hindrance could make the CLP–5LO–F-actin ternary complex impossible. The current structure gives support to the experimental result that the 5LO–CLP and CLP–F-actin interactions are mutually exclusive, suggesting a modulation in actin dynamics. This is also consistent with the fact that no F-actin–CLP–5LO ternary complex has ever been observed experimentally.

Interestingly, these two critical binding residues are both close to the C-terminal region in structure and the C-terminal structure is quite disordered, suggesting the flexibility of the C-terminal region of CLP is important for actin and 5LO binding.

In addition, actin-binding proteins compose a complex system of different kinds of proteins, regulating in different ways and competing with each other in binding with monomeric or polymeric actin. Coactosin interferes with the activity of capping proteins in this complicated system. The hydrophobic region of subdomains 1 and 3 are supposed to be the binding surfaces for capping proteins.[15] According to the F-actin–CLP models suggested by us, these surfaces would not be exposed when human CLP binds to F-actin, so that CLP could efficiently counteract capping proteins association with F-actin. But because human CLP only binds to the "outer side" of F-actin, it would not have any effect on actin polymerization.

Coordinates

The atomic coordinates for human CLP have been deposited with the Protein Data Bank (PDB accession code: 1T2L).

boxed in cyan and yellow, which conserved in CLPs and coactosin. This figure was prepared using ESPript.[18] (B) The ribbon representation of human CLP, AC family (rendering with yeast cofilin) and gelsolin/villin family (rendering with human gelsolin) structures. Residues showed with ball-stick correspond to the ones explained above with the same color. The C terminus of human CLP are colored in purple, which buries the actin-binding site Lys75. The black arrow denotes the "long helix" that interacts with G-actin directly in gelsolin domain G4.[15]

References

1. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J. D. (1994). *The Molecular Biology of the Cell* (3rd edit.). Garland Press, New York.

2. McGough, A. (1998). F-actin binding proteins. *Curr. Opin. Struct. Biol.* **8**, 166–176.

3. de Hostos, E. L., Bradtke, B., Lottspeich, F. & Gerisch, G. (1993). Coactosin, a 17 kDa F-actin binding protein from *Dictyostelium discoideum. Cell Motil. Cytoskeleton*, **26**, 181–191.

4. Rohrig, U., Gerisch, G., Morozova, L., Schleicher, M. & Wegner, A. (1995). Coactosin interferes with the capping of actin filaments. *FEBS Letters*, **374**, 284–286.

5. Chen, K. S., Manian, P., Koeuth, T., Potocki, L., Zhao, Q., Chinault, A. C. *et al.* (1997). Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nature Genet.* **17**, 154–163.

6. Nakatsura, T., Senju, S., Ito, M., Nishimura, Y. & Itoh, K. (2002). Cellular and humoral immune responses to a human pancreatic cancer antigen, coactosin-like protein, originally defined by the SEREX method. *Eur J Immunol.* **32**, 826–836.

7. Provost, P., Doucet, J., Stock, A., Gerisch, G., Samuelsson, B. & Radmark, O. (2001). Coactosin-like protein, a human F-actin-bining protein: critical role of lysine-75. *Biochem. J.* **359**, 255–263.

8. Provost, P., Doucet, J., Hammarberg, T., Gerisch, G., Samuelsson, B. & Radmark, O. (2001). 5-Lipoxygenase interacts with coactosin-like protein. *J. Biol. Chem.* **276**, 16520–16527.

9. Peters-Golden, M. & Brock, T. G. (2000). Intracellular compartmentalization of leukotriene biosynthesis. *Am. J. Respir. Crit. Care Med.* **161**, S36–S40.

10. Lappalainen, P., Kessels, M. M., Cope, M. J. & Drubin, D. G. (1998). The ADF homology (ADF-H) domain: a highly exploited actin-binding module. *Mol. Biol. Cell*, **9**, 1951–1959.

11. Dai, H., Wu, J., Xu, Y., Tang, Y., Ding, H. & Shi, Y. (2004). ^1H, ^{13}C and ^{15}N resonance assignments and the secondary structures of human coactosin like protein (hCLP) D123N. *J. Biomol. NMR*, **29**, 455–456.

12. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.

13. Ono, S., McGough, A., Pope, B. J., Tolbert, V. T., Bui, A., Pohl, J. *et al.* (2001). The C-terminal tail of UNC-60B (actin depolymerizing factor/cofilin) is critical for maintaining its stable association with F-actin and is implicated in the second actin-binding site. *J. Biol. Chem.* **276**, 5952–5958.

15. McLaughlin, P. J., Gooch, J. T., Mannherz, H. G. & Weeds, A. G. (1993). Structure of gelsolin segment 1-actin complex and the mechanism of filament severing. *Nature*, **364**, 685–692.

14. Lappalainen, P., Fedorov, E. V., Fedorov, A. A., Almo, S. C. & Drubin, D. G. (1997). Essential functions and actin-binding surfaces of yeast cofilin revealed by systematic mutagenesis. *EMBO J.* **16**, 5520–5530.

16. Bansal, M., Kumar, S. & Velavan, R. (2000). HELANAL: a program to characterize helix geometry in proteins. *J. Biomol. Struct. Dynam.* **17**, 811–819.

17. Carson, M. (1997). Ribbons. *Methods Enzymol.* **277**, 493–505.

18. Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). ESPript: multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305–308.

19. Lorenz, M., Popp, D. & Holmes, K. C. (1993). Refinement of the F-actin model against X-ray fiber diffraction data by the use of a directed mutation algorithm. *J. Mol. Biol.* **234**, 826–836.

20. Holmes, K. C., Popp, D., Gebhard, W. & Kabsch, W. (1990). Atomic model of the actin filament. *Nature*, **347**, 44–49.

21. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Macromolecular Crystallography* (Carter, C.W.a.S.R.M., ed) *Methods in Enzymology*, vol. 276.

22. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849–861.

23. Terwilliger, T. C. (2000). Maximum likelihood density modification. *Acta Crystallog. sect. D*, **56**, 965–972.

24. Terwilliger, T. C. (2002). Automated main-chain model-building by template-matching and iterative fragment extension. *Acta Crystallog. sect. D*, **59**, 34–44.

25. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of error in these models. *Acta Crystallog. sect. A*, **47**, 110–119.

26. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system (CNS): a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.

27. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.

Edited by R. Huber