Available online at www.sciencedirect.com
ScienceDirect
Biochimica et Biophysica Acta 1774 (2007) 35–43
Biochimica et Biophysica Acta BBA
www.elsevier.com/locate/bbapap

Solution structure of the second SH3 domain of human CMS and a newly identified binding site at the C-terminus of c-Cbl☆

Bo Yao, Jiahai Zhang, Haiming Dai, Jianping Sun, Yuanyuan Jiao, Yajun Tang, Jihui Wu*, Yunyu Shi*

School of Life Science, University of Science and Technology of China, Hefei, Anhui 230027, Peoples Republic of China
Hefei National Laboratory for Physical Sciences at Microscale, University of Science and Technology of China, Hefei, Anhui 230027, Peoples Republic of China

Received 8 April 2006; received in revised form 12 September 2006; accepted 12 September 2006
Available online 27 October 2006

Abstract

CMS, cas ligand with multiple Src homology 3 (SH3) domains, belongs to a family of ubiquitously expressed adaptor proteins. Among the CMS binding proteins, c-Cbl has been mostly extensively studied. It was reported that the motif PKPFPR (residues 824–829) of c-Cbl can bind to the N-terminus SH3 domains of CMS. Here we report the solution structure of the second SH3 domain of CMS (CMS_SH3_B), furthermore, we have identified that a peptide from residues 701 to 714 of c-Cbl (Cbl-p), i.e. MTPSSRPLRPLDTS, can specially bind to CMS_SH3_B using NMR chemical shift perturbation, suggesting that the peptide is a new potential CMS binding site. Among the peptide, TPSSRPLR is the core binding motif and Arg709 plays a key role in the interaction. Cbl-p binding interface on CMS_SH3_B along a hydrophobic channel is composed of RT loop, n-Src loop and *β4* strand and divided into three pockets. This work indicates the solution structure of CMS_SH3_B bears the canonical *β*–*β*–*β*–*β*–α–*β* fold and a new binding site in c-Cbl involved in its interaction with CMS, which probably contributes to the clustering of CMS. All the information provided here should be beneficial for the future functional study of CMS.

Keywords: CMS; SH3 domain; C-Cbl; Binding site; NMR

1. Introduction

Adaptor molecules are usually scaffolding proteins containing one or more non-catalytic protein–protein interaction domains. They can bind to other proteins to form many multiprotein complexes, which are involved in intracellular signal transmitters that regulate cell growth, differentiation, proliferation, migration and survival [1,2]. Human CMS is just such an adaptor molecule. CMS contains three SH3 domains in its N-terminal region, also a proline-rich region, and a coiled-coil domain and several actin-binding sites in the C-terminus. As a scaffolding protein, CMS can bind to many proteins via its domains and motifs [3–5]. The mouse CMS, CD2-associated protein (CD2AP), can bind to the adhesion molecule CD2 via its first SH3 domain, which enhances CD2 clustering at the T-cell antigen-presenting cell contact area, which in turn affects the cytoskeleton polarity of the cell. In the kidney, CMS/CD2AP assembles a complex with nephrons and podocin. This is essential for maintaining the integrity of the nephrons. CD2AP physically interacts with polycystin-2, the product of autosomal dominant polycystic kidney disease (ADPKD) related gene PKD2. CD2AP mediates the association of polycystin-2 with multimeric intracellular and membrane bound complexes in renal tubular epithelia [6–8]. Besides, the proline-rich region of human CMS can also bind to the other SH3 domains, such as the one in the N-terminus of $p130^{cas}$, which contributes to the dynamic regulation of actin cytoskeleton, inducing membrane

Abbreviations: 2D and 3D, two- and three-dimensional; NMR, nuclear magnetic resonance; HSQC, heteronuclear single quantum correlation; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlated spectroscopy; RMSD, root-mean-square deviation; CMS, cas ligand with multiple Src homology 3 (SH3) domains

☆ The atomic coordinates have been deposited in the RCSB Protein Data Bank with accession code 2FEI.

* Corresponding authors. School of Life Science, University of Science and Technology of China, Hefei, Anhui 230027, Peoples Republic of China. Tel.: +86 551 3607464; fax: +86 551 3601443.
E-mail addresses: wujihui@ustc.edu.cn (J. Wu), yyshi@ustc.edu.cn (Y. Shi).

1570-9639/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2006.09.018

ruffling, positively regulating cell migration [4,9,10]. The coiled-coil domain in the C-terminus of CMS contains a putative leucine zipper motif that mediates homodimerization of CMS. There are also five putative actin-binding sites in its coiled-coil domain, which hinted that CMS might be involved in the regulation of actin cytoskeleton [4].

Among the CMS binding proteins, c-Cbl has been extensively investigated [11–13]. Like CMS, c-Cbl is also an adaptor protein functioning as a negative regulator of many signaling pathways that starts from receptors at cell surface [14–16]. The interaction of CMS and c-Cbl mostly involves the N-terminus SH3 domains of CMS, especially CMS_SH3_B, and the C-terminal Proline–Arginine motifs of c-Cbl [11,13]. The SH3 domain is ubiquitous in numerous proteins, serves widely as a protein-interaction module. As one of the first modular protein interaction domains identified, the SH3 domain provides an efficient method for protein-assembling to form more complicated multiprotein complexes [18]. Early studies indicated that the SH3 domain bound to PxxP motif through its proline-rich segment. Considering the location of the basic residue (usually arginine) to the PxxP core in the ligand, there are two classes of the SH3 bound ligands. The class I ligands have the general consensus+xXPxXP and the class II ligands have the general consensus XPxXPx+(where x is any residue, X is in most cases a hydrophobic residue, and +is usually a basic residue). The consensus sequences determine the ligand binding orientations to the SH3 domains [17–19]. However, an atypical Proline–Arginine motif (PKPFPR) at the C-terminus of c-Cbl was also identified to interact with the N-terminus SH3 domains of CIN85, an adaptor protein in the CIN85/CMS family. This is a novel SH3 domains binding motif with PxxxPR core [20,21].

Even though the PKPFPR (residues 824–829) of c-Cbl has been identified to be involved in the interaction between c-Cbl and the N-terminus SH3 domains of CMS [21], there is still the question of whether it is the only binding site in the C-terminus of c-Cbl. To find the answer, a potential SH3 binding segment MTPSSRPLRPLDTS (Cbl-p) located between c-Cbl residues 701–714 was selected for study, which contains a core motif TPSSRPLR. Since CMS_SH3_B plays an important role in the interaction of CMS and c-Cbl [11,13], we focused on studying the binding characteristics between CMS_SH3_B and Cbl-p based on the 3D solution structure of CMS_SH3_B. In this paper, we presented the solution structure of CMS_SH3_B by NMR method and studied the interaction between CMS_SH3_B and Cbl-p by NMR chemical shift perturbation experiment. For the first time, we unveiled the structural details of CMS_SH3_B and provided that the Proline-rich motif (residues 701–714) of the c-Cbl can bind to the second SH3 domain of CMS directly with a low binding affinity. Chemical shift perturbation experiments indicate that Gln13, Asn14, Glu15, Asp16, Val32, Trp36, Leu47, Phe48, and Asn51 on CMS_SH3_B all located in a hydrophobic channel formed by RT loop, n-Src loop and *β4*, especially Trp36, may play important roles in the interaction between CMS and Cbl-p. Hydrophobic and electrostatic interaction may play an important role during molecular recognition between these two proteins.

2. Materials and methods

2.1. Cloning, expression and purification of CMS_SH3_B

The DNA fragment encoding the second SH3 domain (residues 111–166) of CMS (Genebank accession No. AF146277) was amplified from human brain cDNA library (Clontech) by a polymerase chain reaction. And the constructed plasmid was transformed into *Escherichia coli* BL21 (DE3) for expression. The CMS_SH3_B protein product was purified using Ni-chelating column (Qiagen) and then sephadex™ 75 gel filtration (Pharmacia). Uniformly labeled recombinant SH3 domain was produced using minimal medium containing 0.5 g/l 99% ^{15}N-ammonium chloride and 2.5 g/l 99% ^{13}C-glucose as the sole nitrogen and carbon source. Taq polymerase, Ex Taq polymerase, T-vector, DNA ligase and the relevant restriction enzymes were obtained from Takara. The Sequence of constructed plasmid was confirmed by DNA sequencing (Takara).

2.2. Mass spectroscopy and concentration measurement

The purity of CMS_SH3_B sample was checked by SDS-PAGE and confirmed by Mass Spectroscopy (Bruker). Its molecular weight measured is 7765.87 (unlabeled), close to the expected value of 7765.67. Sample concentration was measured with BCA kits (Pierce). The ^{15}N-labeled and ^{13}C, ^{15}N-labeled CMS_SH3_B were about 1.0 mM. All the samples for NMR contained 20 mM phosphate buffer (pH 6.85), 50 mM sodium chloride in 90% H_2O/10% D_2O.

2.3. NMR spectroscopy and data processing

All NMR experiments were performed on a Bruker DMX600 spectrometer with self-shielded z-axis gradients. The following spectra were recorded at 295 K to obtain backbone and side chain resonance assignments: 2D ^{1}H, ^{15}N-HSQC, 3D triple-resonance spectra HNCO, HN(CA)CO, CBCA(CO)NH, CBCANH, C(CO)NH-TOCSY, H(CCO)NH-TOCSY, ^{15}N-TOCSY, HCCH-TOCSY, HCCH-COSY, HBHA(CBCACO)NH, 3D ^{15}N-separated and ^{13}C-separated NOESY were acquired with mixing time of 100 ms and 130 ms respectively.

The ^{15}N, ^{13}C-labeled CMS_SH3_B protein was lyophilized and dissolved in 99.96% D_2O followed immediately with HSQC experiments to monitor the disappearance of NH signals at 295 K. After all of the corresponding peaks had vanished, 2D homonuclear TOCSY and NOESY were recorded on this sample at 295 K, which exhibited proton resonance from aromatic rings exclusively in the region beyond 6.0 ppm.

NMR data processing was carried out using NMRPipe software [22], and the data were analyzed with Sparky software. All the software was run on redhat 9.0 system. Linear prediction was used to improve spectral resolution in the indirect dimensions where constant-time acquisition was used, for example, the ^{15}N evolution dimension in all the triple-resonance experiments mentioned above.

2.4. Experimental restraints and NMR structure calculation

NMR distance restraints were collected from two different NOESY spectra: 3D ^{15}N-separated NOESY in H_2O for amide protons and 3D ^{13}C-separated NOESY in D_2O for aliphatic protons. NOE restraints were grouped into four distance ranges: strong, 1.8–3.0 Å and 1.8–3.5 Å; medium, 1.8–4.0 Å and 1.8–4.5 Å; weak, 1.8–5.0 Å; and very weak, 1.8–6.0 Å. The 1.8 Å lowest limits were imposed only implicitly by the van der Waals repulsion force. For methyl protons, nonstereospecifically assigned methylene protons, and aromatic ring protons, r^{-6} summation averages were applied [23]. Besides, the dihedral angle restraints were forecasted by TALOS with the chemical shifts of Ca, Cb, C′, N and HA [22,24]. The calculation results included φ and ψ angles and angle deviation on every residue of CMS_SH3_B.