THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 50, pp. 38845–38853, December 15, 2006

Atypical Polyproline Recognition by the CMS N-terminal Src Homology 3 Domain*[S]

Received for publication, July 5, 2006, and in revised form, September 27, 2006 Published, JBC Papers in Press, October 3, 2006, DOI 10.1074/jbc.M606411200

Gabriel Moncalián[‡1], Nayra Cárdenes[‡2], Yonathan Lissanu Deribe[§], Mercedes Spínola-Amilibia[‡], Ivan Dikic[§], and Jerónimo Bravo[‡3]

From the [‡]Signal Transduction Group, Structural Biology and Biocomputing Programme, Centro Nacional de Investigaciones Oncológicas, Melchor Fernández Almagro 3, E-28029 Madrid, Spain and [§]Institute of Biochemistry II, Goethe University Medical School, 60590 Frankfurt, Germany

The CIN85/CMS (human homologs of mouse SH3KBP1/CD2AP) family of endocytic adaptor proteins has the ability to engage multiple effectors and couple cargo trafficking with the cytoskeleton. CIN85 and CMS (Cas ligand with multiple Src homology 3 (SH3) domains) facilitate the formation of large multiprotein complexes required for an efficient internalization of cell surface receptors. It has recently been shown that c-Cbl/Cbl-b could mediate the formation of a ternary complex between one c-Cbl/Cbl-b molecule and two SH3 domains of CIN85, important for the ability of Cbl to promote epidermal growth factor receptor down-regulation. To further investigate whether multimerization is conserved within the family of adaptor proteins, we have solved the crystal structures of the CMS N-terminal SH3 domain-forming complexes with Cbl-b- and CD2-derived peptides. Together with biochemical evidence, the structures support the notion that, despite clear differences in the interaction surface, both Cbl-b and CD2 can mediate multimerization of N-terminal CMS SH3 domains. Detailed analyses on the interacting surfaces also provide the basis for a differential Cbl-b molecular recognition of CMS and CIN85.

Binding of growth factors to receptor tyrosine kinases (RTKs)[4] promotes receptor activation and initiates intracellular signaling cascades that ultimately control biological responses such as cell proliferation, cellular adhesion, or apoptosis (1). Down-regulation of RTKs is a critical step in modulating their activity and is mediated by endocytosis, receptor ubiquitination, and subsequent lysosomal degradation (2). Key to this process is the formation of large protein complexes (3) capable of removing activated receptors from the plasma membrane via clathrin-dependent or -independent pathways (4–7).

The CIN85/CMS (Cas ligand with multiple SH3 domains) family of adaptor proteins participates in the formation of such protein complexes and orchestrates multiple steps in RTK signaling and endocytosis (8). Both CMS (or its mouse ortholog CD2AP) and CIN85 (mouse ortholog SH3KBP1) contain a coiled-coil domain at the C-terminal end, a proline-rich region, and three Src homology 3 (SH3) domains in the N-terminal end. All three domains are responsible for the multiprotein complex formation (8, 9). CIN85 and CMS were suggested to promote clustering of multiple Cbl molecules together with their ability to promote oligomerization through their C-terminal coiled-coil regions. These interactions appear to be important for proper internalization and endocytosis of activated receptor tyrosine kinases (10, 11).

SH3 domains bind proline-rich regions and are essential in the assembly and regulation of many intracellular signaling processes (reviewed in Ref. 12). They are generally composed of ~60 amino acids in a β-barrel structure made up of five β-strands. This arrangement forms one acidic and two hydrophobic pockets between the two β-sheets involved in the recognition and binding of proline-rich motifs (for a review see Refs. 13–15). The majority of SH3 domains characterized to date bind to a proline-rich sequence with a φP$X\varphi$P (PXXP) motif, where φ is usually a hydrophobic residue (16). Each motif adopts a left-handed type II polyproline (PPII) helix. Proline residues are in *trans* conformation, and the helical structure shows precisely three residues/turn (17). Peptide ligands potentially bind a given SH3 domain in either one of two opposite directions governed by the location of a positively charged residue, usually Arg, which often precedes or follows the PXXP core element (18, 19). According to this, the PXXP motif is usually further classified into +$X\varphi$P$X\varphi$P (class I) and φP$X\varphi p$ + (class II) (where + is usually an arginine residue).

Conversely, all CIN85/CMS SH3 domains, can only recognize a novel proline-arginine motif (PXXXPR) present in the C-terminal region of Cbl and Cbl-b (20, 21). Additional cellular proteins, such as CD2 receptor (CD2), Pak, Disabled 1 and 2, SETA-binding protein 1 (SB-1), SLP-65/BLNK, or Alg2-interacting protein (AIP1), and ASAP, contain the consensus sequence PXXXPR and are also bound to CMS/CIN85 in mammalian cells (8, 20). CD2 is a transmembrane cell surface glycoprotein expressed on T lymphocytes, thymocytes, and natural

* This work was supported by Research Grant SAF2003-03860 of the "Ministerio de Ciencia y Tecnología," Spain. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental data.

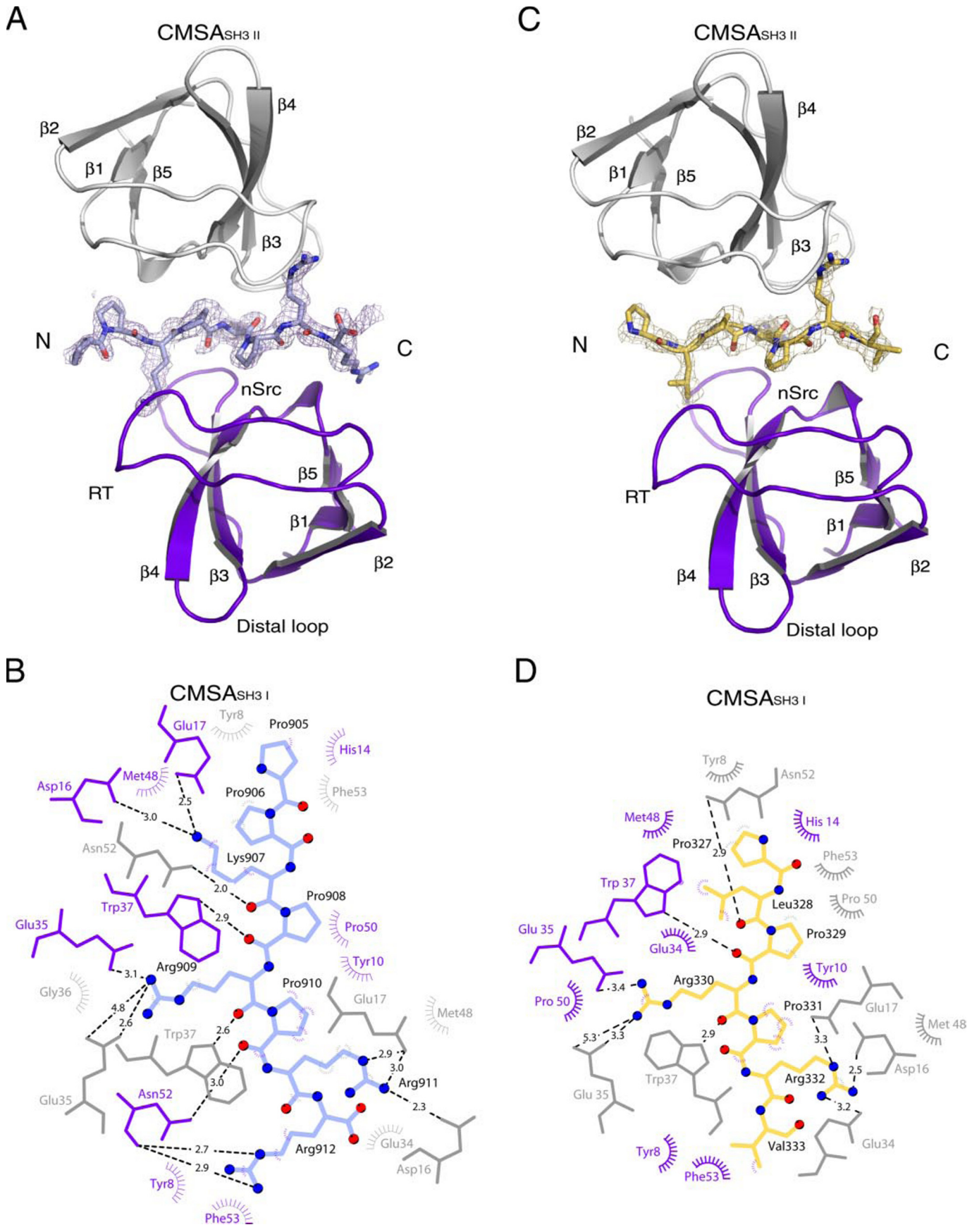
The atomic coordinates and structure factors (code 2j6f, 2j6o, 2j7i, and 2j6k) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] Supported by the "Comunidad de Madrid," Spain. Present address: Departamento de Biología Molecular, Universidad de Cantabria, C/Herrera Oria s/n, 39011 Santander, Spain.

[2] Holder of a fellowship from the "Fundación Ramón Areces," Spain.

[3] To whom correspondence should be addressed. Tel.: 34-912246900 (ext. 2125); Fax: 34-912246976; E-mail: jbravo@cnio.es.

[4] The abbreviations used are: RTK, receptor tyrosine kinase; SH, Src homology; GST, glutathione *S*-transferase; HA, hemagglutinin; wt, wild type; HEK, human embryonic kidney; RT, Arg-Thr.


A
CMSASH3 II
β2
β4
β1
β5
β3
N
C
nSrc
RT
β5
β1
β4
β3
β2
Distal loop
C
CMSASH3 II
β2
β4
β1
β5
β3
N
C
nSrc
RT
β5
β1
β4
β3
β2
Distal loop
B
CMSASH3 I
Glu17
Tyr8
Pro905
His14
Asp16
Met48
Phe53
Pro906
Asn52
Lys907
Trp37
Pro908
Glu35
Pro50
Tyr10
Arg909
Pro910
Gly36
Glu17
Met48
Trp37
Arg911
Glu35
Asn52
Glu34
Arg912
Asp16
Tyr8
Phe53
D
CMSASH3 I
Tyr8
Asn52
Met48
His 14
Pro327
Phe53
Trp 37
Leu328
Pro 50
Glu 35
Pro329
Glu34
Tyr10
Arg330
Pro 50
Pro331
Glu17
Met 48
Trp37
Asp16
Arg332
Glu 35
Glu34
Val333
Tyr8
Phe53

killer cells, which mediates low affinity cell/cell interactions by binding to related immunoglobulin superfamily proteins (CD58 in humans). It functions to promote T cell adhesion, immune recognition/activation, and motility (9, 22). Central clustering of CD2 has been proposed to have an important role in maintaining a stable domain for sustained antigen receptor engagement and signaling (23). This CD2 clustering as well as T-cell polarization and recruitment of CD2 into the central contact zone during formation of an immunological synapse has been shown to be dependent on binding of the CMS mouse ortholog CD2AP to the cytoplasmic domain of CD2 (22, 23). The N-terminal SH3 domain of CD2AP appears to have a relatively high specificity and avidity for the conserved proline-rich region of the CD2 cytoplasmic end (22). Both N-terminal CMS and CIN85 SH3 domains (CMSA and CIN85A) have a stronger affinity for CD2 with similar dissociation constants ($K_D = \sim 100\ \mu M$) (22) than the rest of CMS and CIN85 SH3 domains (22, 23).

It has recently been shown that Cbl can simultaneously accommodate two SH3 domains from separate CIN85 molecules, promoting the formation of higher order Cbl/CIN85 complexes (24). The study led to the identification of crucial additional residues in the proline-arginine motif of Cbl involved in the formation of a ternary complex with CIN85 N-terminal SH3 and the assembly of the endocytic machinery that governs down-regulation of activated receptor tyrosine kinases. Cbl-b arginines 904 and 911 make equivalent hydrogen bonds with acidic residues in each SH3 domain, allowing the peptide to bind in class I and II simultaneously. The presence of Arg-904 (class I-like binding) defines the multimerization state of the CIN85 SH3-peptide complexes, indicating why other ligands that also contain the P*XXX*PR motif but lack the N-terminal arginine, such as PAK, bind with a 1:1 stoichiometry (24–26). Furthermore, it has also been suggested that other targets of CIN85 that lack the N-terminal arginine, such as BLNK or CD2, should similarly bind in a 1:1 stoichiometry.

CMS is closely related to CIN85; they both interact with Cbl and CD2 and are involved in down-regulation of RTKs (8). To elucidate whether this unique SH3 recognition mechanism is conserved within the CMS/CIN85 family of adaptors and further investigate the role of the N-terminal arginine in the R*XP*P*XXX*PR motif, we have solved the crystal structures of the CMS N-terminal SH3 domain (CMSA) in complex with Cbl-b- and CD2-derived peptides. A differential recognition pattern with respect to CIN85 has been characterized by combining the analysis of structural information and mutagenesis in *in vivo* data. We also aimed at setting up the basis for the atypical proline-rich sequences of Cbl recognition of CMS with respect to CIN85.

EXPERIMENTAL PROCEDURES

Protein Overexpression and Purification—The N-terminal SH3 domain from human CMS (CMSA, residues 1–62) was cloned into pET21 and overexpressed without the His tag in the *Escherichia coli* bv. *rosetta* (DE3) pLys strain at 37 °C. Purification was performed as previously described for CIN85A (24).

Crystallization, Crystallographic Data Collection, Structure Determination, and Refinement—Prior to crystallization, 3 μmol of corresponding peptide were added to 1.2 μmol of purified CMSA protein (70 mg/ml final concentration). All crystals were obtained at 20 °C using the hanging drop vapor diffusion method by mixing 1 μl of protein solution with 1 μl of reservoir solution. Crystals were then transferred to stabilization buffer (10% (v/v) ethylene glycol, 90% of reservoir solution) and flash-cooled to 110 °K. CMSA/Cbl-b (Cbl-b PARPPKPRPRR; Genosys) was crystallized in the space group I422 with cell dimensions $a = 66.51$, $b = 66.51$, and $c = 67.78$ Å of CMSA:Cbl-b (Table 1) with reservoir solution containing 20% polyethylene glycol 3000, 0.1 M acetate at pH 5.5.

The CMSA/CD2 heterodimer (CD2 KGPPLPRPRV; Genosys) was crystallized with reservoir solution of 30% polyethylene glycol 8000, 0.2 M NaCl, 0.1 M Tris-HCl, pH 8.0, in the space group $P2_12_12_1$ with cell dimensions $a = 36.37$, $b = 59.80$, and $c = 71.12$ Å and the CMSA/CD2 heterotrimer crystallized in the I422 space group with cell dimensions $a = 66.394$, $b = 66.448$, and $c = 68.739$ Å (crystallization solution: 20% polyethylene glycol 3000, 0.1 M acetate, pH 5.5). All data sets were collected in a Mar 345 imaging plate and a Bruker rotating anode source and were processed with DENZO and SCALEPACK software (27) for CMSA/Cbl-b or MOSFLM and SCALA software (28) for CMSA/CD2 complexes.

The CMSA:Cbl-b complex structure was solved by molecular replacement using AMoRe software (29) and the coordinates of an isolated CIN85A as the search model.[5] Similarly, the CMSA/Cbl-b structure was used as a template for solving CMSA/CD2 structures. The structures were refined by cycles of positional and restrained individual B-value refinement with CNS (30) and Refmac5 software (28) and by manual model building using Xfit (31) or Coot software (32).

Biochemical Assays —Cell transfection with the indicated cDNAs, cell lysis, glutathione *S*-transferase (GST) pulldowns, immunoprecipitations, and immunoblotting were essentially performed as described previously (20).

RESULTS

CMSA/Cbl-b Overall Crystal Structure—To understand the molecular details of the different members of the CIN85/CMS family of adaptor molecules, we have crystallized the N-termi-

[5] N. Cárdenes, G. Moncalián, and J. Bravo, unpublished data.

FIGURE 1. **Overall structure of the CMSA·Cbl-b and CMSA·CD2 heterotrimeric complexes and protein/peptide interaction details.** *A* and *C*, $CMSA_{SH3I}$ binds the peptides in a class I orientation (shown in *purple*), and $CMSA_{SH3II}$ recognizes the peptide in class II orientation (shown in *gray*). The Cbl-b peptide is shown in *light blue* (*A*) and CD2 peptide in yellow (*B*). The σA-weighted electron density map around them is contoured at 1.0 σ. Elements of secondary structure and the positions of the RT, *n*-Src, and distal loops are indicated. *B* and *D*, Schematic representation of contacts between CMSA and the Cbl-b (*B*) and CD2 (*D*) peptides (*light blue* and *yellow*, respectively). Residues 902, 903, and 904 from the Cbl-b peptide were not visible at the electron density map. Residues 324, 325, and 326 from the CD2 peptide are also disordered. Interactions and residues from SH3I are shown in *purple* and in *gray* for SH3II. *Dashed lines* show hydrogen bonds (labeled with peptide-protein distances in Å), and *purple* and *gray rays* designate hydrophobic interactions.

nal SH3 domain of CMS (CMSA) in complex with the same Cbl-b-derived peptide used for the crystallization of the CIN85A/Cbl-b complex (24).

The refined structure at 1.7 Å of CMSA/Cbl-b shows a ternary complex composed of two CMSA molecules binding, in opposite orientations, to a single Cbl-b peptide molecule (Fig. 1*A*). One SH3 domain recognizes Lys-907 in the N terminus of the peptide, resembling a classical class I orientation, and the second SH3 domain recognizes Arg-911 in the C terminus, resembling class II. The asymmetric unit contains one CMSA molecule bound to a single Cbl-b peptide, which lies along a crystallographic 2-fold axis. Thus, a crystallographically related CMSA molecule binds the peptide on its opposite side, and electron density maps show an average of two orientations for the peptide (see supplemental figure). Final refined B factors for the peptide, considering occupancy of 0.5 for each orientation, do not differ significantly from the B factors of the CMSA-interacting residues (~25 Å^2). The double orientation of the peptide is possible, because crystallographic contacts are exclusively established by CMSA molecules. CMSA residues 1 and 59–62 are not visible in the electron density maps. Cbl-b residues 902, 903, and 904 were also disordered.

The structure of the CMSA molecule displays the typical SH3 domain folding, adopting the characteristic five-stranded β-barrel structure. The peptide, on the other hand, does not display a typical polyproline helix but adopts an extended conformation instead. The structure shows that CMS also has the ability to form heterotrimeric complexes with a Cbl-b derived peptide, as it has recently been reported for CIN85 (24).

CMSA/Cbl-b Molecular Interface—In the heterotrimer arrangement, the interactions of the $CMSA_{SH3I}$ (class I-like orientation) and $CMSA_{SH3II}$ (class II-like orientation) with the Cbl-b peptide are almost identical with 941 and 827 Å^2 of buried surface, respectively, and are comparable with values observed for other SH3 domain/peptide structures. However, the interacting surface of the two CMSA molecules is negligible. Only 54 Å^2 of the 3918 Å^2 total surface of each SH3 is shared, corresponding mainly to both side chains of Glu-35 (from $CMSA_{SH3I}$ and $CMSA_{SH3II}$).

Hydrophilic specific interactions via side chains are only made by Lys-907 and Arg-911 at the N and C termini of the Cbl-b peptide with both Asp-16 and Glu-17 from $CMSA_{SH3I}$ and $CMSA_{SH3II}$ respectively and Arg909 with Glu-35 from both CMSA molecules (Fig. 1*B*). Hydrophobic contacts in the class I orientation are made by Cbl-b Pro-910 with conserved residues of hydrophobic pocket II. In turn, Cbl-b Pro-906 and -908 are embedded within hydrophobic pockets I and II, respectively. Thus, Cbl-b Pro-908 and -910 occupy equivalent positions in the opposed SH3 domains of the heterotrimer. Additionally, there are sequence-unspecific contacts with peptide main chain atoms, such as Pro-910 and Lys-907 carboxylic oxygens interacting, respectively, with Asn-52 from each $CMSA_{SH3I}$ and $CMSA_{SH3II}$. Similarly, Pro-908 and Arg-909 carboxylic oxygens interact with Trp-37 from both $CMSA_{SH3I}$ and $CMSA_{SH3II}$ (Fig. 1*B*).

In Vivo Interactions between Cbl-b and CMS—The structures of the CMSA domain and the peptide derived from Cbl-b containing the P*XXX*PR motif shows that Cbl-b Lys-907 and Arg-911 residues are important for the interaction between the SH3 domain and its ligand (Fig. 1*B*).

We set out to see the importance of these residues for the interaction between Cbl-b and CMS in the context of full-length proteins expressed in cells. To this end, expression vectors encoding HA-tagged Cbl-b wt and Cbl-b K907A, Cbl-b R911A, and Cbl-b K907A/R911A point mutants were transfected and overexpressed in HEK293T cells. Lysates from these cells were then used to test binding to GST-CMSA in GST pulldown assays. Fig. 2*A* shows that both Cbl-b Lys-907 and Arg-911 play important roles in the interaction, although the Cbl-b K907A mutation abrogated binding to a much lesser extent than Arg-911. Double mutation of both Cbl-b Lys-907 and Arg-911 abolished the binding almost completely.

Because of the overall similarity in the P*XXX*PR motif in Cbl-b and c-Cbl, we then checked whether the corresponding residues in c-Cbl also mediate the interaction with CMSA. Thus, we mutated the equivalent residues Lys-825 and Arg-829 to alanines in c-Cbl and proceeded as for GST pulldown assays with Cbl-b transfections. Interestingly, a very similar profile of binding to Cbl-b was observed. Again both c-Cbl Lys-825 and c-Cbl Arg-829 appear to be important for binding to the CMSA domain. The c-Cbl K825A mutation impaired binding to the SH3 domain slightly, whereas R829A led to a much greater loss in binding, and the double point mutations largely abrogated the interaction of c-Cbl with the N-terminal SH3 domain of CMS (Fig. 2*B*). Both experiments reproducibly showed the relative importance of these residues to the binding between the PK*XX*PR motif in Cbl-b/c-Cbl and CMSA.

Further on, we were interested to see the roles of Cbl-b Lys-907 and Cbl-b Arg-911, and the corresponding residues in c-Cbl in the binding of full-length Cbl-b/c-Cbl and CMS. To this end, expression vectors encoding HA-tagged Cbl-b wt and K907A, R911A, and K907A/R911A mutants along with FLAG-tagged CMS were transfected in HEK293T cells. Cell lysates were subjected to immunoprecipitation with an anti-Cbl-RF antibody (recognizing the ring finger region of both Cbl-b and c-Cbl) and immunoblotted with anti-FLAG antibody.

As shown in Fig. 2*C*, FLAG-CMS is co-precipitated quite strongly with wt Cbl-b. This is in sharp contrast to Cbl-b K907A and R911A mutants, where a minimal amount of FLAG-CMS is co-immunoprecipitated with the mutants. This result demonstrates the vital importance of these residues in binding to CMS. The residual binding seen in the case of the Cbl-b K907A/R911A double mutant is most likely because of FLAG-CMS being co-precipitated with endogenous Cbl-b/c-Cbl that is immunoprecipitated by the Cbl-RF antibody.

The same assay was performed with the transfection of HA-c-Cbl wt, K825A, R829A, and K825A/R829A point mutants together with FLAG-CMS. It can be clearly seen that FLAG-CMS is co-immunoprecipitated with wt c-Cbl; however, the c-Cbl K825A and/or R829A mutation markedly abolished the co-immunoprecipitation of CMS (Fig. 2*C*, *lower panel*). Thus, in a similar fashion to Cbl-b, c-Cbl K825A and R829A mutants seem to be important for binding to CMS.

The crystal structure of CMSA:Cbl-b suggests a different interaction pattern in class I recognition compared with CIN85A/Cbl-b. Cbl-b Lys-907 in CMSA/Cbl-b occupies an

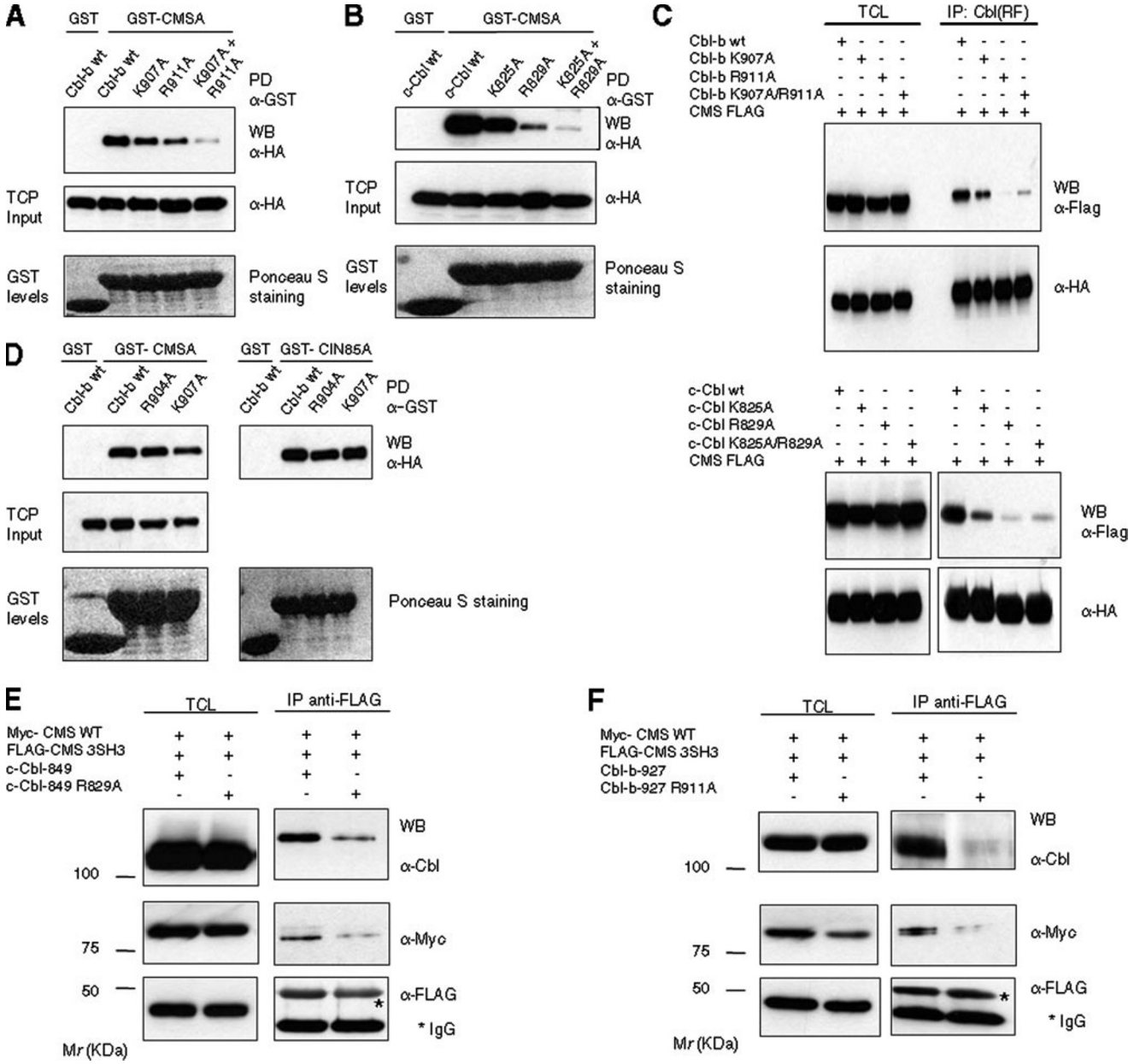


FIGURE 2. **Lys-907 and Arg-911 from Cbl-b play a crucial role in CMSA/Cbl-b binding and heterotrimerization.** *A*, HEK293T cells were transfected with HA-Cbl-b wt, HA-Cbl-b K907A, HA-Cbl-b R911A, or HA-Cbl-b K907A/R911A. Lysates were incubated with GST alone or GST-CMSA, and the inputs and pulldowns (*PD*) were then subjected to immunoblotting (*WB*) with anti-HA antibody. *B*, GST pulldowns and WB were performed also with lysates of HEK293T cells transfected with HA-c-Cbl K825A, R829A, or K825A/R829A point mutants. *C*, HA-tagged full-length Cbl-b wt, K907A, R911A, K907A/R911A mutants and c-Cbl wt, K825A, R829A, and R825A/R829A mutants were transfected in HEK293T cells along with FLAG-tagged full-length CMS. Cell lysates were subjected to immunoprecipitation with an anti-Cbl-RF antibody (recognizing the ring finger region of both Cbl-b and c-Cbl) and immunoblotted with anti-FLAG antibody. *D*, comparison of Cbl-b Arg-904 *versus* Lys-907 in CMSA/Cbl-b and CIN85A/Cbl-b binding. HEK293T cells were transfected with HA-Cbl-b wt, HA-Cbl-b R904A, or HA-Cbl-b K907A. Lysates were incubated with GST alone, GST-CMSA, or GST-CIN85A. The inputs and pulldowns were then subjected to immunoblotting (*WB*) with anti-HA antibody. Shown are HEK293T cells transfected with FLAG-CMSA-3SH3, full-length Myc-tagged CMS, and c-Cbl-849 (*E*) or Cbl-b-927 (*F*) constructs lysed and immunoprecipitated by anti-FLAG affinity agarose beads. Total cell lysates (*TCL*) and immunoprecipitates (*IP*) were subjected to immunoblotting (*WB*) with corresponding antibodies.

equivalent position as Arg-904 in the CIN85A/Cbl-b structure (see Figs. 1, *A* and *B*; and 5). Therefore, we wanted to compare the relative importance of Cbl-b Arg-904 and Cbl-b Lys-907 residues in binding to N-terminal SH3 domains from CMS and CIN85. Fig. 2*D* shows that GST-CMSA was able to bring down Cbl-b R904A as potently as wild type but pulled down Cbl-b K907A only weakly. In sharp contrast, GST-CIN85A was able to pull down both wild type Cbl-b and Cbl-b K907A mutant equally potently, with a slight reduction with respect to Cbl-b R904A in accordance with a previous observation (24).

The structure of CMSA/Cbl-b shows the formation of heterotrimers involving two CMS N-terminal SH3 domains and one Cbl peptide *in vitro*. To check whether the formation of CMSA/Cbl heterotrimers can also take place *in vivo*, HEK293T cells were transfected with a FLAG-tagged construct of CMS containing only the three SH3 domain regions, Myc-tagged

full-length CMS, and either c-Cbl-849 or Cbl-b-927 constructs that lack the ability to homodimerize. These Cbl mutants were used to ensure that co-precipitation between CMS molecules was a consequence of direct binding of two independent CMS molecules to a single Cbl molecule rather than due to the interaction with homodimeric Cbl molecules (24). Cells overexpressing c-Cbl-849 or Cbl-b-927 have the ability to co-precipitate full-length CMS using a FLAG antibody (Fig. 2, *E* and *F*), indicating that Cbl is mediating the formation of the heterotrimer in cells. We have shown that point mutants of c-Cbl or Cbl-b are significantly impaired in binding to CMS (Fig. 2, *A–D*). When checking for heterotrimer formation in cells using the same point mutations in the constructs that lack the ability to homodimerize (Cbl-b-927 R911A and c-Cbl-849 R829A), the amount of co-precipitated full-length CMS is markedly diminished (Fig. 2, *E* and *F*, respectively). It is important to note that, because we cannot avoid the endogenous Cbl proteins that can mediate the formation of a heterotrimeric complex, we have detected some full-length CMS co-precipitated even in the absence of overexpressed Cbl proteins. Taken together, we can conclude that a complex containing Cbl proteins and two intermolecular SH3 domains of CMS can be formed in cells in the context of full-length CMS proteins.

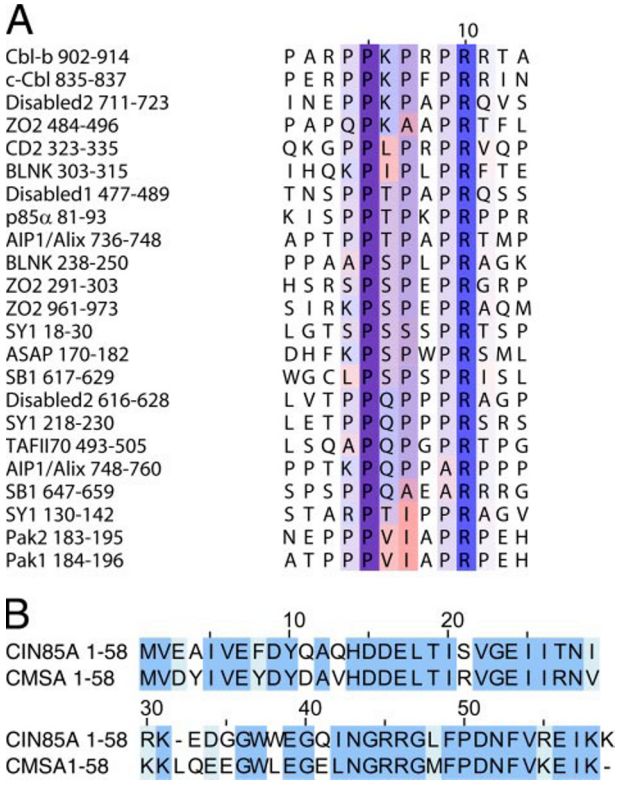


FIGURE 3. **Alignment of interacting sequences from CIN85/CMS partners and of CMSA and CIN85A.** *A*, several polyproline-rich sequences containing the P*XXX*PR motif from CIN85/CMS interacting molecules were aligned. Hydrophobic residues are shown in conservation gradient from *blue* to *red*. *B*, sequence alignment of the N-terminal SH3 domains from CMS and CIN85 was performed. Identical residues are shown in *blue* and conserved hydrophobic residues in *light blue*.

CMSA/CD2 Crystal Structures—Structural and *in vivo* CMSA/Cbl-b results demonstrate the ability of CMS to form heterotrimeric complexes with Cbl-b-derived peptides in a similar fashion as was reported for CIN85. We wanted to check whether a similar CMS recognition was possible for other atypical polyproline-containing molecules. We observed, in the CMSA/Cbl-b structure, that Lys-907 is involved in the interaction with SH3I (Fig. 1*B*). However, this residue is only conserved in Cbl-b or c-Cbl proteins, and it is not present in some other CMS/CIN85-interacting proteins such as CD2, SB-1, SLP-65/BLNK, or AIP1 (Fig. 3*A*). In CD2, the position of the lysine is occupied by a leucine (PLPRPR). The fact that the Cbl-b Lys-907 side chain interacts with the RT loop of SH3I in the CMSA/Cbl-b crystal structure (Fig. 1*D*) could suggest that a CD2-derived peptide could be predominantly a class II-like interaction. To check this possibility, the crystal structure of the CMSA/CD2-derived peptide complex was solved at 2.9 and 2.2 Å (see "Experimental Procedures" and Table 1). Two different crystals were obtained. Unexpectedly, one of the structures was a heterotrimer of two SH3 domains binding to the peptide in the same mode that Cbl-b-derived peptide is being recognized (Fig. 1*C*), yet the other structure is of a single SH3 domain recognizing the peptide in a class II orientation. Both CMSA/CD2 and CMSA/Cbl-b heterotrimeric structures display similar peptide recognition modes, with the exception of contacts established by Cbl-b Lys-907 (Leu-328 in CD2) (Fig. 1*D*).

TABLE 1
Crystallographic data collection and analysis

	CMSA/Cbl-b	CMSA/CD2 dimer	CMSA/CD2 trimer
Data collection			
Space group	I422	$P2_12_12_1$	I422
Unit cell dimensions			
a, *b*, *c* (A)	66.51, 66.51, 67.78	36.32, 59.72, 70.87	66.43, 66.43, 68.72
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Data range (Å)	20–1.7	36.3–2.9	34.4–2.5
Observations (unique)	107,249 (8,696)	139,722 (3,729)	66,722 (4,006)
Completeness (%) (last shell)	99.9 (100.0)	100.0 (100.0)	98.0 (86.6)
R_{sym}[a] (last shell)	0.034 (0.269)	0.095 (0.307)	0.110 (0.408)
Mean I/σ (last shell)	20.9 (6.9)	22.0 (6.2)	24.5 (7.5)
Refinement			
Non-hydrogen atoms (solvent molecules)	641 (60)	1104 (32)	595 (69)
R_{work}[b] (R_{free})[c]	0.189 (0.218)	0.245 (0.293)	0.241 (0.286)
Root mean square bond length (Å)	0.012	0.0080	0.010
Root mean square bond angles (°)	1.379	1.400	1.306

[a] R_{sym} is the unweighted *R* value on *I* between symmetry mates.
[b] $R_{work} = \Sigma_{hkl}[[F_{obs}(hkl)] - [F_{calc}(hkl)]]/\Sigma_{hkl}[F_{obs}(hkl)]$.
[c] R_{free} is the cross-validation *R*-factor for 5% of reflections against which the model was not refined.

DISCUSSION

CMSA/Cbl-b Forms a Heterotrimeric Complex with a Differential Recognition from CIN85A/Cbl-b—Both CMS and CIN85 share the same domain organization with a 39% sequence identity (~70%) within the conserved domains (Fig. 3*B*). This similarity suggests that they may have overlapping binding specificities. Consequently, they have been assumed to share similar biological functions (8). Superposition of CMSA/Cbl-b and CIN85A/Cbl-b structures reveal that the overall structure of CMSA/Cbl-b resembles the heterotrimer reported for CIN85A/Cbl-b (24). One SH3 domain binds the Cbl-b peptide in a class I-oriented fashion, whereas the second SH3 recognizes the peptide in the opposite orientation (class II). Such similarity in the overall structures of the heterotrimeric complexes provides a molecular explanation for a common role of Cbl in the molecular mechanism of CIN85/CMS-mediated down-regulation of RTKs. Cbl might also participate in the clustering of CMS as reported for CIN85 and β-PIX (24). The C-α of the SH3 individual domains overlap well with one another (root mean square deviation, 0.70 Å) (Fig. 4). Main differences are located within the *n*-Src and distal loops. Nevertheless, the overall CMSA/Cbl-b heterotrimer shows some important differences with respect to CIN85A/Cbl-b. In particular, the orientation of the SH3 domains in CMSA/Cbl-b differs from that of the SH3 orientations in CIN85A/Cbl-b, revealing a distinctive recognition of the Cbl-b-derived peptide for CMS or CIN85 (Fig. 4).

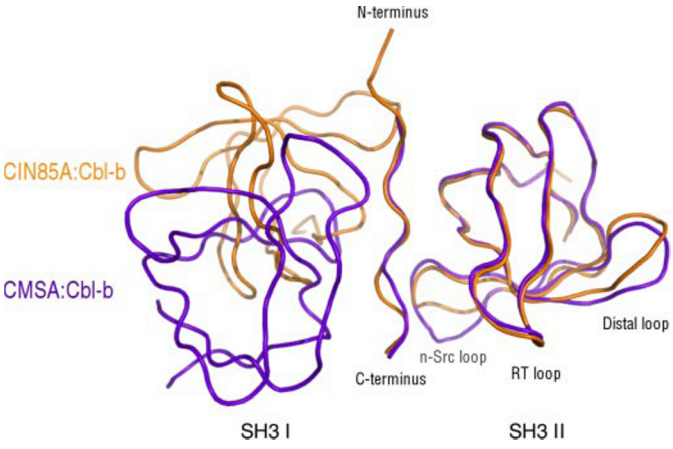


FIGURE 4. **Comparison of the ternary CMSA/Cbl-b complex with the reported CIN85A/Cbl-b complex (2BZ8).** SH3II backbones from the CMSA/Cbl-b (*purple*) and CIN85A/Cbl-b (*orange*) complexes are overlaid. *n*-Src, RT, and distal loops are indicated.

The interactions corresponding to the class II orientation of the peptide ($CMSA_{SH3II}$/Cbl-b) are conserved with respect to CIN85 ($CIN85A_{SH3II}$/Cbl-b) (Fig. 1*B*). Arg-911 determines the class II orientation, and its importance in (CMS/CIN85)/Cbl interaction has already been described and proven for CIN85 (20, 21). The conserved heterotrimeric arrangement suggests it also plays a critical role in CMS/Cbl-b class II interaction. The only difference in class II recognition between both structures concerns Cbl-b Arg-909, which hydrogen bonds with Asp-33 from CIN85 or Glu-35 from CMS (Fig. 1*B*).

However, significant differences in the interaction surface between Cbl-b and CMSA or CIN85A are observed in the class I orientation. Residues forming the two hydrophobic binding pockets are highly conserved between CMSA and CIN85A, yet the peptide residues interacting with their SH3I domains are different. Most differences are found in the acidic pocket formed by the *n*-Src and RT loops. Comparison of the uncomplexed CMSA (supplemental data) and CIN85A crystal structures show no conformational changes in the *n*-Src and RT loops upon peptide binding (root mean square deviations, 0.42 and 0.64 Å, respectively), and therefore, differences observed between CMSA/Cbl-b and CIN85/Cbl-b are not due to conformational changes upon peptide binding.

When compared with the CIN85A/Cbl-b-reported structure, the N terminus of the Cbl-b peptide adopts a different conformation that mostly reflects changes in the class I binding orientation. A significant three-residue shift of the Cbl-b peptide is observed when comparing $CMSA_{SH3I}$ and $CIN85_{SH3I}$ (Fig. 5). Resultant $CMSA_{SH3I}$/Cbl-b interactions include hydrogen bonds of Asp-16 and Glu-17 side chains to Lys-907, contacts of the Pro-908 main chain with the conserved Trp-37, and hydrophobic interactions of the Pro-910 side chain with residues from hydrophobic pocket II.

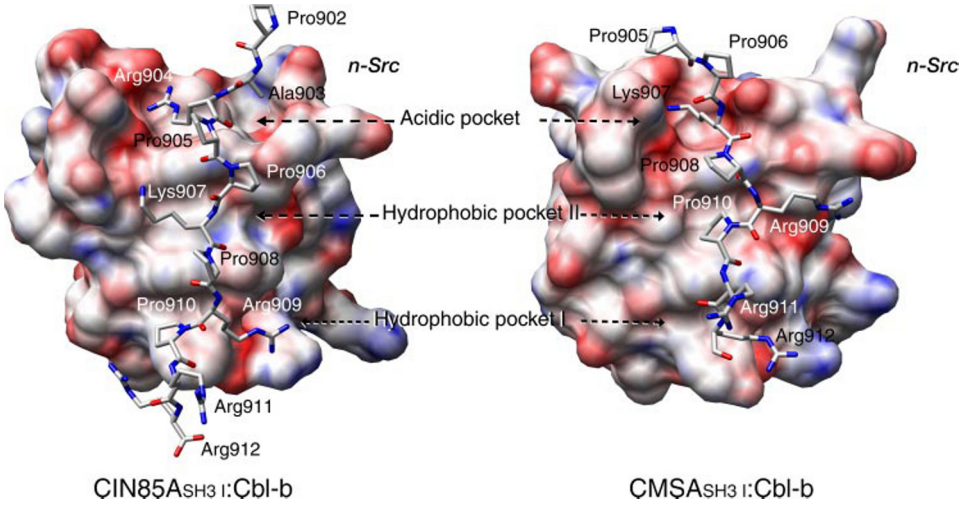


FIGURE 5. **Comparison of the SH3I/Cbl-b interface from CIN85A and CMSA.** Electrostatic potential surfaces of SH3 domains from CIN85A/Cbl-b (*left*) and CMSA/Cbl-b (*right*) complexes that bind the peptide in a class I-like orientation are shown. The Cbl-b peptides are represented in *sticks*, and their residues are indicated. *n*-Src loops are labeled in both structures, and binding pockets are indicated with *arrows*.

In CIN85A, Asp-16 and Glu-17 from the RT loop hydrogen bond with Cbl-b Arg-904, which plays a critical role in the formation of the CIN85A/Cbl-b heterotrimeric complex. However, in CMSA/Cbl-b, Arg-904 does not participate in the interaction, and it is the residue Lys-907 that interacts with the RT loop in the acidic pocket (Fig. 5).

In addition, Cbl-b Arg-909 in CIN85A/Cbl-b occupies a similar position to Cbl-b Arg-912 in the CMSA/Cbl-b structure, and hydrophobic pocket I is empty in $CMSA_{SH3I}$, because the synthesized peptide ends at Arg-912. The next

residue in the Cbl-b sequence is Thr-913, and there is certainly room for its accommodation in CMSA/Cbl-b. Overall, the three-residue shift observed for class I recognition leads to the conclusion that the Cbl-b sequence interacting with CIN85 might differ from the optimal Cbl-b sequence interacting with CMS, affecting the affinity for both adaptor molecules. This could have severe consequences in the multiprotein complex formation responsible for RTK down-regulation and other cellular processes.

This three-residue shift in the class I orientation can be rationalized by sequence and structural differences within the respective *n*-Src loops (Figs. 3*B* and 4), which have been reported to account for binding selectivity (16). With respect to the CIN85 sequence, a leucine is inserted in CMS at position 32 that significantly changes the overall conformation of the *n*-Src loop. In addition, Glu-35 (Gly-34 in CIN85) hydrogen bonds Arg-909 from the peptide. The N-terminal end of the peptide faces opposite directions in each structure, and such displacement might be due to the change in the conformation of the *n*-Src loop, given that Glu-34 in CMSA does not allow the peptide to adopt a conformation similar to the one displayed in the CIN85A·Cbl-b complex (Fig. 5).

CMSA Interacts with CD2 Forming Two Types of Molecular Arrangements—We have shown that residue Lys-907 from Cbl-b plays an important role on CMSA binding (Figs. 1*B* and 2, *A* and *C*). Cbl-b Lys-907 is not conserved for other CMS ligands, and CD2 contains Leu-328 at the equivalent position. This would suggest that at least the class I type of interaction could be compromised upon CD2 binding to CMSA. In fact, we have obtained crystals of a dimeric form of a 1:1 CMSA·CD2 complex on a class II type of interaction. Unexpectedly, an additional heterotrimeric crystal has also been obtained in an independent crystal form. The existence of the two molecular arrangements suggests a minimal energy difference between the heterotrimeric and dimeric states that might have some biological consequences in contrast to the predominant heterotrimeric complex formed by CMSA/Cbl-b.

In the CMSA·Cbl-b complex, Cbl-b Lys-907 hydrogen bonds to CMS Asp-16 and Glu-17. This interaction cannot take place with CD2 because of a substitution of the equivalent Cbl-b lysine to leucine. The total non-polar buried surface upon CMSA binding is similar for Cbl-b Lys-907 and CD2 Leu-328 (101 and 120 $Å^2$, respectively); thus, the main difference is the absence of the Lys-907-mediated hydrogen bond. However, the absence of this hydrogen bond is not enough to prevent the formation of the CMSA/CD2 heterotrimer.

In addition, CD2 peptides obtained in the heterotrimer and dimeric forms perfectly superimpose (root mean square deviation, 0.48 Å). The two independent CMSA/CD2 structures present entirely different crystal packing, indicating that the extended conformation of the peptide observed in CMSA·peptide complex structures cannot be due to packing forces.

CIN85 and CMS SH3 Binding Consensus Motifs—Biochemical experiments shown here provide *in vivo* confirmations for the observed structures and the residues involved in the binding between the SH3 domains and their ligands. GST pulldown and co-immunoprecipitation experiments shown in Fig. 2 clearly indicate that Cbl-b Lys-907 and Cbl-b Arg-911 and the corresponding residues in c-Cbl play an important role in binding to CMS in the context of full-length proteins in cells. Furthermore, we have shown a differential binding of SH3 domains of CMS and CIN85 to Cbl-b. This is in such a way that Cbl-b Arg-904 is more important for binding to CIN85A than to CMSA, whereas the opposite holds true for Cbl-b Lys-907, which is observed to be more important for binding to CMSA. In addition, we have also shown that Cbl mediates the formation of heterotrimers in the context of full-length CMS. Cbl-mediated multimerization of CMS proteins occurring *in vivo* suggests similar roles between CMS and CIN85. Despite the differential Cbl molecular recognition between the two adaptor molecules, they both contribute to the Cbl-mediated clustering and formation of large protein complexes required for receptor down-regulation.

The structures presented here unambiguously indicate that residues involved in class I recognition are different from those involved in class II. Cbl-b Lys-907 and Pro-910 are responsible for CMSA/Cbl-b class I orientation. Specific key residues such as Cbl-b Arg-904 might only be essential for CIN85A heterotrimerization and not for CMSA heterotrimerization. Besides, we have shown, from the CMSA/CD2 interaction, that Cbl-b Lys-907 can be replaced by a Leu and still form a heterotrimer. Given that hydrophobic pocket I is empty in $CMSA_{SH3I}$/Cbl-b and $CMSA_{SH3I}$/CD2, the motif could be extended to $\varphi XXPXX$ for CMS class I and P*X*P*XX*R for both CIN85 and CMS class II. Taken together, our results indicate that consensus sequences should indeed be considered independent for class I and class II recognition types. Nevertheless, the overall motif is similar to the one described by Kowanetz *et al.* (20) (P*X*P/A*X*PR), and a structure of CMSA/Cbl-b P908A at 1.9 Å resolution confirms that the same heterotrimerization is possible with Pro-908 replaced by an alanine (supplemental data). It has been shown that some other hydrophobic residues could be present at this position without significantly affecting the interaction, such as isoleucine as found in the Pak polyproline motif (Fig. 3*A*) (26). Thus, a more specific CMS consensus motif could be narrowed down to $P\varphi$(P/A)*X*PR*X* for the heterotrimer formation and R*X*P*X*(P/A)*X*PR in the case of CIN85.

The process of endocytosis requires the formation of large multiprotein complexes for efficient internalization of cell surface proteins. To a large extent, this is facilitated by the so-called endocytic adaptor proteins that could engage multiple effectors and couple cargo trafficking with the cytoskeletal apparatus (3). The CIN85/CMS family of proteins represents an example of such adaptors. These proteins enhance endocytosis by clustering Cbl proteins through their three SH3 domains (20). The existence of yet another level of organization is also suggested by the recent observation that c-Cbl/Cbl-b proteins can mediate the formation of a ternary complex between one c-Cbl/Cbl-b molecule and two SH3 domains of CIN85 (24). The structural and biochemical study provided here further supports the notion of additional clustering via SH3 domain dimerization through a c-Cbl/Cbl-b molecule and extends a similar possibility for CD2.

Acknowledgments—We thank Miguel Ortiz for help in the refinement and critical review of the manuscript and José Rivera and Carme Fábrega for useful discussion.

REFERENCES

1. Schlessinger, J. (2002) *Cell* **110,** 669–672
2. Waterman, H., and Yarden, Y. (2001) *FEBS Lett.* **490,** 142–152
3. Dikic, I., and Giordano, S. (2003) *Curr. Opin. Cell Biol.* **15,** 128–135
4. Bonifacino, J. S., and Traub, L. M. (2003) *Annu. Rev. Biochem.* **72,** 395–447
5. Szymkiewicz, I., Shupliakov, O., and Dikic, I. (2004) *Biochem. J.* **383,** 1–11
6. Sigismund, S., Woelk, T., Puri, C., Maspero, E., Tacchetti, C., Transidico, P., Di Fiore, P. P., and Polo, S. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102,** 2760–2765
7. Le Roy, C., and Wrana, J. L. (2005) *Nat. Rev. Mol. Cell Biol.* **6,** 112–126
8. Dikic, I. (2002) *FEBS Lett.* **529,** 110–115
9. Tibaldi, E. V., and Reinherz, E. L. (2003) *Int. Immunol.* **15,** 313–329
10. Soubeyran, P., Kowanetz, K., Szymkiewicz, I., Langdon, W. Y., and Dikic, I. (2002) *Nature* **416,** 183–187
11. Szymkiewicz, I., Kowanetz, K., Soubeyran, P., Dinarina, A., Lipkowitz, S., and Dikic, I. (2002) *J. Biol. Chem.* **277,** 39666–39672
12. Zarrinpar, A., Bhattacharyya, R. P., and Lim, W. A. (2003) *Science's STKE* 2003, RE8
13. Cesareni, G., Panni, S., Nardelli, G., and Castagnoli, L. (2002) *FEBS Lett.* **513,** 38–44
14. Larson, S. M., and Davidson, A. R. (2000) *Protein Sci.* **9,** 2170–2180
15. Ferreon, J. C., and Hilser, V. J. (2004) *Biochemistry* **43,** 7787–7797
16. Mayer, B. J. (2001) *J. Cell Sci.* **114,** 1253–1263
17. Adzhubei, A. A., and Sternberg, M. J. (1993) *J. Mol. Biol.* **229,** 472–493
18. Feng, S., Chen, J. K., Yu, H., Simon, J. A., and Schreiber, S. L. (1994) *Science* **266,** 1241–1247
19. Lim, W. A., Richards, F. M., and Fox, R. O. (1994) *Nature* **372,** 375–379
20. Kowanetz, K., Szymkiewicz, I., Haglund, K., Kowanetz, M., Husnjak, K., Taylor, J. D., Soubeyran, P., Engstrom, U., Ladbury, J. E., and Dikic, I. (2003) *J. Biol. Chem.* **278,** 39735–39746
21. Kurakin, A. V., Wu, S., and Bredesen, D. E. (2003) *J. Biol. Chem.* **278,** 34102–34109
22. Hutchings, N. J., Clarkson, N., Chalkley, R., Barclay, A. N., and Brown, M. H. (2003) *J. Biol. Chem.* **278,** 22396–22403
23. Dustin, M. L., Olszowy, M. W., Holdorf, A. D., Li, J., Bromley, S., Desai, N., Widder, P., Rosenberger, F., van der Merwe, P. A., Allen, P. M., and Shaw, A. S. (1998) *Cell* **94,** 667–677
24. Jozic, D., Cardenes, N., Deribe, Y. L., Moncalian, G., Hoeller, D., Groemping, Y., Dikic, I., Rittinger, K., and Bravo, J. (2005) *Nat. Struct. Mol. Biol.* **12,** 972–979
25. Mott, H. R., Nietlispach, D., Evetts, K. A., and Owen, D. (2005) *Biochemistry* **44,** 10977–10983
26. Hoelz, A., Janz, J. M., Lawrie, S. D., Corwin, B., Lee, A., and Sakmar, T. P. (2006) *J. Mol. Biol.* **358,** 509–522
27. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
28. Collaborative Computational Project (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–763
29. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50,** 157–163
30. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905–921
31. McRee, D. E. (1999) *J. Struct. Biol.* **125,** 156–165
32. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 2126–2132