THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 31, pp. 24099–24107, July 30, 2010

Crystal Structure of CCM3, a Cerebral Cavernous Malformation Protein Critical for Vascular Integrity*[S]

Received for publication, March 29, 2010, and in revised form, May 11, 2010 Published, JBC Papers in Press, May 19, 2010, DOI 10.1074/jbc.M110.128470

Xiaofeng Li[‡1], **Rong Zhang**[‡1], **Haifeng Zhang**[§], **Yun He**[§], **Weidong Ji**[§], **Wang Min**[§2], **and Titus J. Boggon**[‡3]

From the Departments of [‡]*Pharmacology and* [§]*Pathology, Yale University School of Medicine, New Haven, Connecticut 06520*

***CCM3* mutations are associated with cerebral cavernous malformation (CCM), a disease affecting 0.1–0.5% of the human population. CCM3 (PDCD10, TFAR15) is thought to form a CCM complex with CCM1 and CCM2; however, the molecular basis for these interactions is not known. We have determined the 2.5 Å crystal structure of CCM3. This structure shows an all α-helical protein containing two domains, an N-terminal dimerization domain with a fold not previously observed, and a C-terminal focal adhesion targeting (FAT)-homology domain. We show that CCM3 binds CCM2 via this FAT-homology domain and that mutation of a highly conserved FAK-like hydrophobic pocket (HP1) abrogates CCM3-CCM2 interaction. This CCM3 FAT-homology domain also interacts with paxillin LD motifs using the same surface, and partial CCM3 co-localization with paxillin in cells is lost on HP1 mutation. Disease-related CCM3 truncations affect the FAT-homology domain suggesting a role for the FAT-homology domain in the etiology of CCM.**

Cerebral cavernous malformation (CCM)[4] is a common vascular lesion that affects the central nervous system vasculature with a prevalence of 0.1–0.5% in the human population (1, 2) (OMIM 116860, 603284, 603285). CCMs manifest as thin-walled, dilated blood vessels lined by a monolayer of endothelial cells that lack tight junctions. The clinical effects of these lesions include seizures, headaches, and stroke in midlife and are often associated with focal hemorrhage (1, 2). These lesions can occur sporadically or as a familial form attributable to mutations in three different genes: *CCM1*, *CCM2*, and *CCM3*. A majority of mutations in CCM genes result in truncations of their protein products, CCM1 (Krev/Rap1 Interacting Trapped 1; KRIT1) (3, 4), CCM2 (malcavernin, MGC4607, osmosensing scaffold for mitogen-activated protein kinase kinase kinase-3; OSM) (5, 6), and CCM3 (programmed cell death 10; PDCD10, TF-1 cell apoptosis-related protein 15; TFAR15) (7, 8). These mutations are inherited in an autosomal dominant fashion (9) with acquisition of CCM lesions hypothesized to be due to a two-hit mechanism (10–12). Expression of the CCM proteins is required for both development and maintenance of endothelial cells in the vasculature; they are required for normal vasculogenesis (13) and global deletion of CCM1 renders mice non-viable (14), a result also seen in global or endothelial-specific deletion of CCM2 (15) and CCM3 (16). Overall, the clinical and *in vivo* data point to an essential role for the CCM proteins in endothelial cells that makes them critical for vasculogenesis and survival.

CCM1, CCM2, and CCM3 interact with one another and play roles in multiple signaling pathways with CCM2 acting as a hub that directly interacts with both CCM3 (17) and CCM1 (18, 19). The structural characteristics of the CCM proteins have not been directly assessed but have been inferred by molecular modeling and homology studies that suggest three regions in CCM1 (NP*X*Y-rich, ankyrin repeat, and FERM domains) (20) and a predicted phosphotyrosine binding (PTB) domain in CCM2 (5). The fold of CCM3 has not been identified, but has been predicted to contain a single domain (17). The regions of CCM3 and CCM2 that directly interact have not been mapped.

In the current study, we describe the crystal structure of CCM3 and show that it contains two domains: an N-terminal dimerization domain and a C-terminal focal adhesion targeting (FAT)-homology domain. We show that the role of the N-terminal domain as a dimerization domain and the C-terminal domain is a mediator of direct interaction with CCM2. We also show that the CCM3 C-terminal FAT-homology domain directly interacts with the CCM2 PTB domain, that this interaction is disrupted by point mutations in a FAK-like hydrophobic pocket 1 (HP1) on the surface of CCM3 and that co-localization of CCM2 and CCM3 within the cell is altered by introduction of CCM3 FAT-homology point mutations. Previously discovered FAT domains in proteins such as FAK and Pyk2 (21–23) have been shown to be critical for focal adhesion targeting by directly binding to leucine-rich LD-motifs in the focal adhesion protein, paxillin. Our crystal structure of CCM3 shows high structural homology to these domains, we therefore investigated whether CCM3 can bind paxillin. We show that CCM3 binds to the LD-motifs of paxillin, LD1, LD2, and LD4, and that mutations in HP1 and in paxillin LD1 abrogate this

* This work was supported, in whole or in part, by National Institutes of Health Grants R01 HL077357 and HL085789 (to W. M.) and by startup funds (to T. J. B.).

The atomic coordinates and structure factors (codes 3L8I and 3L8J) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental materials, Figs. S1–S4, and Table S1.

[1] Both authors contributed equally to this work.

[2] To whom correspondence may be addressed: Dept. of Pathology, Yale University School of Medicine, 10 Amistad St., 401B, New Haven, CT 06520-8023. Tel.: 203-785-6047; Fax: 203-737-2293; E-mail: wang.min@yale.edu.

[3] To whom correspondence may be addressed: Dept. of Pharmacology, Yale University School of Medicine, 333 Cedar St., SHM B-316A, New Haven, CT 06520. Tel.: 203-785-2943; Fax: 203-785-5494; E-mail: titus.boggon@yale.edu.

[4] The abbreviations used are: CCM, cerebral cavernous malformation; FAT, focal adhesion targeting; MES, 4-morpholineethanesulfonic acid; FITC, fluorescein isothiocyanate; TRITC, tetramethylrhodamine isothiocyanate; HA, hemagglutinin; PTB, phosphotyrosine binding; RMSD, root-mean-square deviation; GST, glutathione *S*-transferase; PDB, Protein Data Bank.

TABLE 1
Summary of crystallographic information

	CCM3	CCM3
Protein		
Space group	$P2_12_12_1$	$P4_122$
Cell dimensions, *a*, *b*, *c* (Å)	63.1, 116.0, 123.0	77.4, 77.4, 108.6
Resolution range (Å)[a]	25–2.50 (2.59–2.50)	25–3.05 (3.16–3.05)
Unique reflections[a]	31805	6699
Completeness (%)[a]	99.1 (97.4)	99.3 (98.0)
R_{sym} (%)[a]	6.1 (66.2)	8.4 (49.1)
Mn ($I/\sigma I$)[a]	16.3 (2.0)	19.3 (2.7)
R_{factor} (%)[a]	22.9 (30.2)	27.8 (34.9)
Free R_{factor} (%)[a]	28.8 (40.1)	29.3 (41.5)
Residue range built	A:15–212 B: 1–210 C:10–210 D: 8–210	14–210
Free *R* reflections (%)	5.1	4.9
Free *R* reflections, no.[a]	1600 (101)	316 (30)
No. non-hydrogen protein atoms	6655	1630
No. water molecules	70	0
Model quality		
RMSD bond lengths (Å)	0.008	0.006
RMSD bond angles (°)	1.106	0.881
Mean *B*-factors		
Overall (Å^2)	79.3	146.5
Protein atoms (Å^2)	79.5	146.5
Protein atoms NTD (N-69) (Å^2)	58.7	132.9
Protein atoms αE-EF linker (70–97) (Å^2)	80.0	148.5
Protein atoms FAT,(98-C) (Å^2)	90.1	152.5
Water (Å^2)	63.2	–
Ramachandran plot (%) favored/allowed/disallowed	92.5/7.5/0	90.4/9.5/0

[a] Parentheses indicate the highest resolution shell.

interaction. We then show that CCM2 and paxillin LD1 can compete for binding to CCM3. We also demonstrate that CCM3 partially co-localizes with paxillin and that this co-localization can be altered by mutations in HP1. This study provides the first insights into the structural dynamics of interactions between the CCM proteins, indicates direct interaction of the CCM3 FAT-homology domain with both CCM2 and LD-motif-containing proteins, and suggests further roles for the CCM proteins in regulation of focal adhesions.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—Full-length human CCM3 was produced in *Escherichia coli* and purified by Ni-affinity and ion-exchange chromatography. For crystallization CCM3 was concentrated to 10 mg/ml in 20 mM MES, pH 6.0, 200 mM NaCl. See supplemental materials for detailed experimental procedures.

Structure Solution—Two crystal forms of CCM3, tetragonal and orthorhombic, were grown in space group $P4_122$ and $P2_12_12_1$. $UO_2(C_2H_3O_2)_2$ was used to derivatize the orthorhombic crystal form. Details of the structure solution are described in supplemental materials, Table 1, supplemental Table S1, and supplemental Fig. S1. Final models have PDB codes 3L8I and 3L8J.

Pull-downs—In vitro pull-down assays were conducted by binding the N-terminal GST-tagged protein to glutathione-Sepharose beads and incubating with the purified soluble protein, followed by extensive washing. Quadruple Lys to Asp (K132D, K139D, K172D, K179D) CCM3 was generated (CCM3–4KD) as were single point mutants A135D and S175D. GST-LD repeat constructs provided by Christopher Turner encoded paxillin LD-repeats 1 through 5 (24, 25). Mutant GST-LD1 was made by mutation of L7/L8 to R7/R8.

N-terminal Paxillin Pull-down Assay—His-paxillin-(1–321) was provided by Joseph Schlessinger, and pull-down was performed in a similar fashion as for the GST fusion proteins.

Competition Assay—Increasing amounts of purified CCM2 PTB domain (51–251) were incubated with GST-LD1 and CCM3 prior to washing.

Co-immunoprecipitation—CCM3 dimerization. Full-length CCM3-HA was co-transfected with Flag-CCM3 variants into 293T cells. Association of CCM3-HA with CCM3 variants was determined by IP with anti-Flag followed by Western blot with anti-HA. Flag-CCM3 in the immunoprecipitates was determined by Western blot.

CCM3 Interactions with CCM2 or Paxillin—Co-immunoprecipitation was performed as described in supplemental materials.

Immunofluorescence—Full-length wild-type CCM3 or CCM3-4KD was transfected into bovine aortic endothelial cells (BAEC). Transfected CCM3 and endogenous paxillin were detected by indirect immunofluorescence microscopy with anti-Flag (rabbit; for CCM3) and anti-paxillin (mouse) followed by FITC-conjugated anti-rabbit and TRITC-conjugated anti-mouse secondary antibodies. CCM3 interactions with CCM2. Wild-type or 4 kDa Flag-tagged CCM3 FAT-homology domain (92–212) CCM3 and full-length GFP-CCM2 were co-transfected into BAECs and were detected by indirect immunofluorescence microscopy with anti-Flag (mouse; for CCM3) and anti-GFP (rabbit) followed by TRITC-conjugated anti-mouse and FITC-conjugated anti-rabbit secondary antibodies.

RESULTS

Overall Structure of CCM3—To determine the structure of CCM3, we purified and crystallized full-length human CCM3 (residues 1–212). We determined the structure in two crystal forms, to 3.05 Å resolution in space group $P4_122$ with unit cell dimensions of $a = b = 77.4$ Å, $c = 108.6$ Å and to 2.50 Å resolution in space group $P2_12_12_1$ with unit cell dimensions of $a = 63.1$ Å, $b = 116.0$ Å, $c = 123.0$ Å (Table 1). Initial phases were obtained using a uranium derivative and SIR/AS for the $P2_12_12_1$ crystal form and confirmed the register by analysis of phased anomalous difference Fourier maps of a selenomethionine derivative (supplemental Fig. S1). Structure validation was conducted using Procheck (26) and MolProbity (27).

There are four CCM3 molecules per asymmetric unit in the $P2_12_12_1$ crystal form arranged as two dimers, and one CCM3 molecule in the $P4_122$ crystal form arranged as a symmetry-related dimer around a 2-fold axis. In each of the five CCM3 peptide chains the protein contains two distinct domains folded from nine α-helices, which we term αA to αI (Fig. 1 and supplemental Fig. S1*E*). We observe clear electron density for residues 15 to 85, 101 to 145, and 160 to 207 in all five copies of CCM3. There is poor electron density for the loops between residues 86 and 100 and 146 and 159. The five copies of CCM3 clearly show a hinge between the N- and C-terminal domains focused around residues Lys-69 to Lys-70. This hinge allows the orientation of the N- and C-terminal domains to vary by ~25° with respect to one another (HELIXANG) (28) (Fig. 1*C*). In both crystal forms the N-terminal domain contains the strongest electron density and shows the lowest temperature factors,

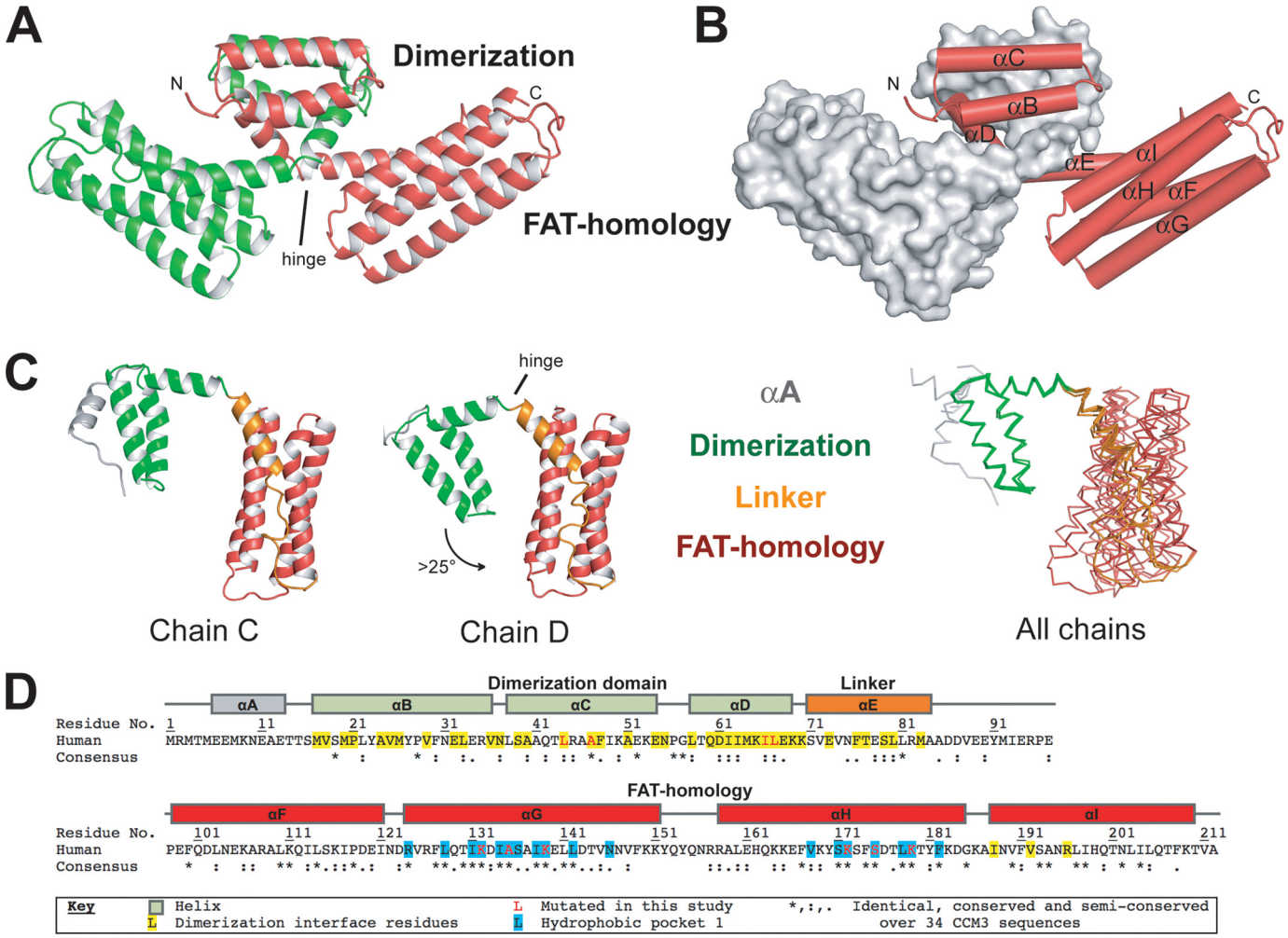

FIGURE 1. **Overall structure of CCM3.** *A*, *scheme representation* of the CCM3 dimer. The CCM3 dimerization (*Dimerization*) domain and FAT-homology (*FAT-homology*) domains are indicated. The location of the hinge between the domains is indicated. *B*, same orientation as *panel A* showing one chain as a *gray surface* and helices αB to αI indicated. *C*, five solved CCM3 chains hinge around residues Lys-69–Lys-70. Chains C and D of the $P2_12_12_1$ crystal form show the largest hinge angle difference and are shown. Backbone trace for all five chains from both crystal forms is superposed on the N-terminal dimerization domain. αA, *gray*; dimerization domain, *green*; αE and loop EF, *orange*; FAT-homology domain, *red*. *D*, sequence and secondary structure assignments of human CCM3. Human sequence (Swiss-Prot Q9BUL8) and residue numbers with secondary structure assignment are for helices αA through αI and colored per domain as in *panel B*. Residues involved in dimerization are colored *yellow* and the hydrophobic pocket 1 (HP1) in *blue*. The consensus sequence over 34 CCM3 sequences from human to *Drosophila* by ClustalW is shown. (*) indicates identical; (:), highly conserved; (.), semi-conserved (see supplemental Fig S4*B* for complete alignment). Residues mutated in this study are colored *red*. All structure figures are made with Pymol.

whereas the C-terminal domain has higher temperature factors with helices αG and αH showing the most flexibility. Residues 15–210 are built in all models and have equivalent secondary structure assignments (Fig. 1*D*). The largest root-mean-square deviation (RMSD) between the N-terminal domains (1–69) for the five CCM3 chains is 0.8 Å over 52 residues and for the C-terminal regions (70–212) is 1.2 Å over 131 residues (TOPP) (28).

CCM3 Contains an N-terminal Dimerization Domain with a Novel Fold—The N-terminal region of CCM3 contains four helices that we term helices αA through αD. Helix αA is seen in three copies of CCM3, is solvent exposed with high *B*-factors and is highly divergent between chains, therefore we do not consider this an intrinsic part of the N-terminal domain. In all copies of CCM3 in both crystal structures helices αB through αD mediate a 2-fold or pseudo 2-fold symmetric dimer for the $P2_12_12_1$ and $P4_122$ crystal forms respectively. This dimer folds as two intercalating Vs, with helices αB, αC as the down-stroke, and helix αD as the up-stroke. The two V-shaped N-terminal domains interlock and are followed by a hinge and helix αE, which also makes interchain contacts (Fig. 2). The CCM3 dimerization interface buries a total surface area of ~3700 Å^2 with an average of 1857 Å^2 from each molecule (Fig. 2*B*), is very well conserved (Fig. 2*C* and supplemental Fig S4*B*), and is almost entirely hydrophobic (Fig. 2*D*). Depending on the crystallographic dimer, 46–48 residues are involved in interchain interactions, fewer than 9 of which form inter-residue hydrogen bonds (PISA) (29). The largest change in surface accessibility on dimerization occurs for residues Ile-66, Leu-67, Phe-76, Met-17, Glu-73, and Leu-44 suggesting importance for these residues in maintaining a dimer interface (NACCESS, S. Hubbard and J. Thornton) (supplemental Fig. S2). The three crystallographically observed dimers in these crystal structures have protein-protein interface shape complementarity scores, a quantification of geometrical packing in protein

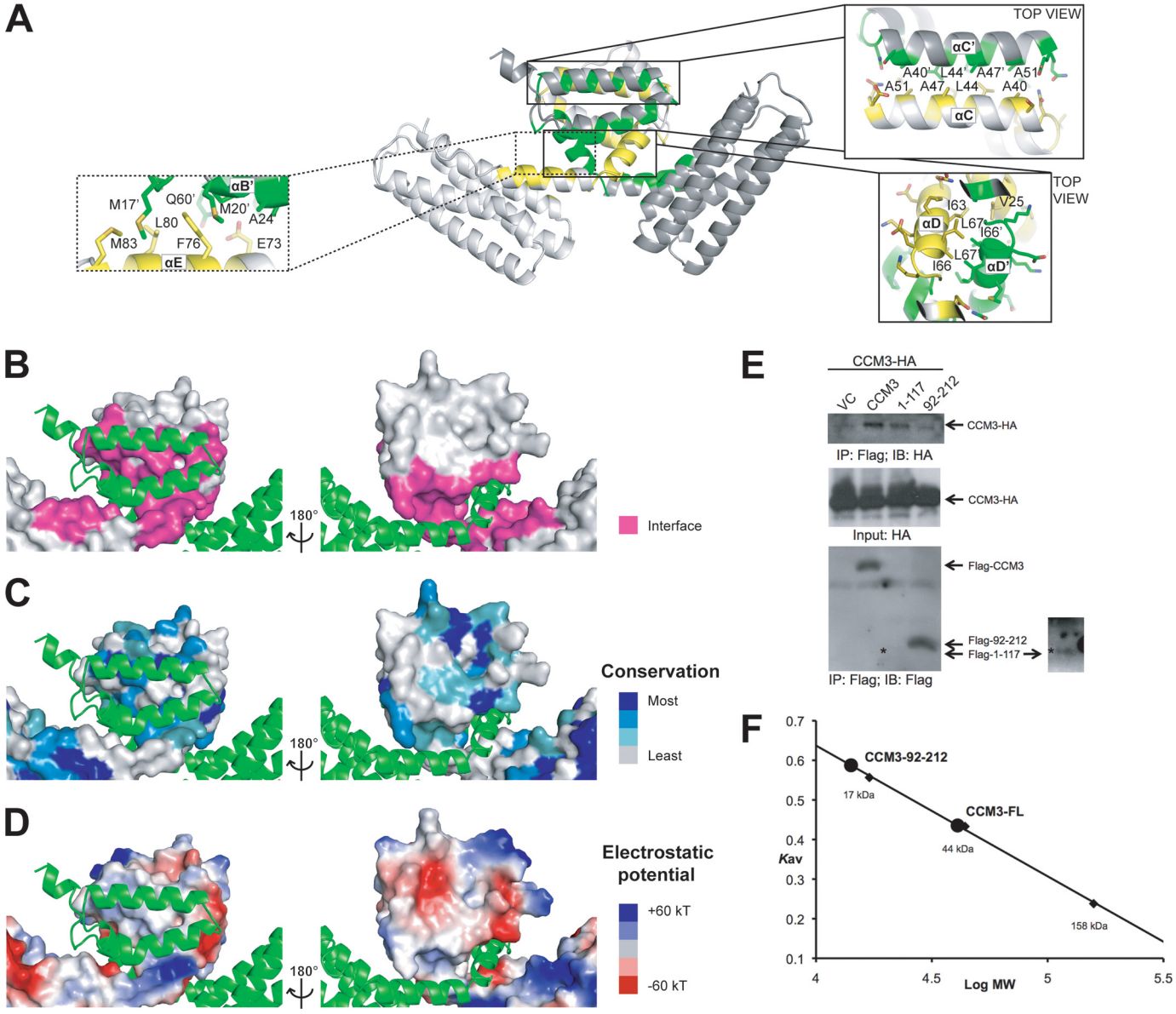

FIGURE 2. **CCM3 dimerization.** *A*, *scheme representation* of the CCM3 dimer. Exploded views show the anti-parallel hydrophobic stripe for helix αC and the symmetric hydrophobic focal point formed by Ile-66 and Leu-67. *B*, interacting residues. Surfaces are colored *pink* for residues that mediate the N-terminal dimerization interface. *C*, sequence conservation. The surface colored by sequence conservation is based on 34 CCM3 sequences. *Darker blue* indicates higher conservation. Generated using Consurf (36). *D*, electrostatic potential representation (+60 kT, *blue*; −60 kT, *red*). *E*, immunoprecipitation. CCM3-HA was co-transfected with Flag-CCM3 variants as indicated into 293T cells. Association of CCM3-HA with full-length CCM3, CCM3 N-terminal region (1–117), and CCM3 FAT domain (92–212) was determined by IP with anti-Flag (M2) followed by Western blot with anti-HA. Flag-CCM3 in the immunoprecipitates was determined by Western blot with anti-Flag. CCM3-HA in the input was also determined. *Arrows* indicate CCM3. Flag-1–117 is indicated with an (*) and shown in the overexposed panel. *VC*, vector control. *F*, size exclusion chromatography. A Superdex 200 analytical grade SEC column was used to analyze CCM3 constructs 1–212 and 92–212. Full-length (1–212) CCM3 elutes as a dimer (predicted mass: ~25 kDa) and construct 92–212 elutes as a monomer (predicted mass: ~14 kDa). Standards used to calibrate the column were bovine γ-globulin 158 kDa, chicken ovalbumin 44 kDa, and equine myoglobin 17 kDa. $K_{av} = (V_e - V_o)/(V_t - V_o)$ where V_e is the elution volume of the protein, V_o is the void volume, and V_t is the total bead volume.

interfaces, of 0.64, 0.62, and 0.58 (SC) (28), this is low for a homodimeric interface, usually observed the range 0.65 to 0.75 (30).

To form the CCM3 dimer, helices αB and αD′ interact through a highly conserved hydrophobic stripe, mediated by αB residues Val-25, Met-26, Val-29, Leu-33, and αD′ residues Leu-58′, Ile-62′, Ile-63′, Lys-65, and Ile-66′. Val-25 forms contacts with Ile-66′, Met-26 with Ile-66′, Val-29 with Ile-62′, and the aliphatic region of Lys-65′, and Leu-33 with Leu-58′. The C-terminal ends of helix αD and αD′ come together in a symmetric hydrophobic focal point formed between residues Ile-66, Leu-67, Ile-66′, and Leu-67′ (Fig. 2*A*). Helices αC and αC′ interact in an antiparallel fashion, mediated by residues Ala-40, Leu-44, Ala-47, Phe-48, and Ala-51. Residue Ala-40 packs against Ala-51′, Leu-44 makes van der Waals contacts with Ala-47′ and Phe-48′, Phe-48 packs against Phe-48′ and Leu-44′, and Ala-51 packs against residue Ala-40′ (Fig. 2*A*). After the hinge at Lys-69-Lys-70 separating helices αD and αE, residue Glu-73 hydrogen bonds to Gln-60′, and residue Phe-76 in helix αE interacts with Met-17′, Met-20′ and Pro-21′ from helix αB′ and Gln-60′ from helix αD. Additionally, an interaction is seen between helix αD residue Met-64 and helix αI′ residues Ile-

188′ and Val-192′. There are two invariant proline residues in the dimerization domain of CCM3; invariant residues Pro-56 and Gly-57 are important for the ~70° angle made between helices αC and helix αD that forms the elbow in the CCM3 dimerization interlocking V-shape. Pro-28 disfigures helix αB, and does not play a role in dimerization, but is a component of the broadly flat surface which allows stacking of CCM3 dimers in both crystal forms. We have not found a representative example of this architecture in previously determined macromolecular structures, searches using the CCM3 N-terminal domain as either a monomer or a dimer with Dali (31), HorA (32), VAST (33), and SSM (34) did not find a dimerization domain with this fold.

To ask whether CCM3 can dimerize *in vivo* we conducted cell-based assays using 293T cells co-transfected a C-terminal HA tagged full-length CCM3 and variants of CCM3 with N-terminal Flag tags. We then immunoprecipitated with anti-Flag followed by Western blot with anti-HA. We found that wild-type CCM3 co-immunoprecipitates with both full-length CCM3 and also with a construct that encodes residues 1–117 (Fig. 2*E*). CCM3 dimerization was confirmed *in vitro* by cross-linking (supplemental Fig. S2) and by size exclusion chromatography (SEC) showing that full-length CCM3 is a dimer in solution (Fig. 2*F*) and that CCM3–92-212 elutes from analytical SEC as a monomer, indicating that dimerization occurs prior to residue 92.

CCM3 Contains a C-terminal FAT-homology Domain—The C-terminal region of CCM3 contains five helices which we term helices αE through αI (Fig. 1*A*). Helix αE is immediately C-terminal to the CCM3 dimerization domain and both contributes to dimerization and packs against helices αF and αI. Helix αE is followed by a poorly structured loop from residues 84–97 in all of the traces (discussed below). Helices αF through αI form a four-helical up-and-down bundle that closely resembles the FAT domains of Pyk2 and FAK; Dali searches with helices αF to αI of each of the five copies of this domain yield Z-scores higher than 14 and RMS deviations over >110 residues of less than 2 Å for the focal adhesion targeting (FAT) domains of Pyk2 and FAK (Fig. 3*A*). As FAT domains fall into the structural category of four-helical up-and-down bundle that includes proteins such as talin VBS, vinculin head, α-catenin M-fragment, and apolipoprotein E, the CCM3 FAT-homology domain also shows high structurally homology to these proteins. While there is high structural similarity between CCM3 and canonical FAT domains, there is poor sequence homology, with ~14 and 10% identity between CCM3 FAT-homology domain and the FAT domains of FAK and Pyk2 respectively (Dali) (supplemental Fig S4*A*). The presence of a FAT-homology domain within CCM3 was not predicted.

The CCM3 FAT-homology domain folds into a right-turn up-down four-helix bundle with each of the helices rotated ~10° from the elongated axis of the FAT domain (HELIXANG) (28), as has been seen previously for FAT domain proteins (21). There is a well-defined highly conserved hydrophobic core in the CCM3 FAT-homology domain that includes residues from all four helices. At the C-terminal end of helix αF there is a kink at conserved proline, Pro-118, but CCM3 does not contain the conserved PxPP motif seen between helices α1 and α2 in the FAK FAT domain and thought to be important for conformational flexibility of helix α1 (35). Instead, this proline residue packs against helix αE with van der Waals interactions made with Leu-81 and aliphatic parts of Glu-78 and Arg-82. The kink in helix αF results in a hydrophobic cavity at the closed end of the FAT-homology domain reminiscent of the cavities found in previously determined FAT domains (21). The CCM3 dimerization domain is located at the closed end of the CCM3 FAT-homology domain.

CCM3 FAT-homology, but Not the CCM3 Dimer Domain, Directly Interacts with CCM2—To investigate whether the N-terminal dimerization domain or the C-terminal FAT-homology domain is responsible for direct CCM3 interaction with CCM2 we made constructs that encoded individual domains. We found that the CCM3 dimerization domain does not pull-down with full-length CCM2 but that the C-terminal FAT-homology domain (CCM3–92-212) interacts with CCM2 (Fig. 3*C* and supplemental Fig. S3). We then conducted sequence conservation analysis of CCM3 (36) and found that the FAT-homology domain contains a stunningly well-conserved surface on helices αG and αH (Fig. 3*B*). The center of this surface is aliphatic and bounded by multiple conserved lysine residues (Lys-132, Lys-139, Lys-172, and Lys-179). We have designated this patch HP1 (hydrophobic patch 1) after FAK nomenclature (21). To investigate whether this conserved surface mediates direct interaction with CCM2 we designed point mutations to alter the charges around the edge, or the hydrophobicity in the center of HP1. We introduced a quadruple mutant, mutating four of the conserved lysines (Lys-132, Lys-139, Lys-172, Lys-179) to aspartate (CCM3–4KD), and two single mutations of Ala-135 or Ser-175 to aspartate (CCM3-A135D and CCM3-S175D), and conducted pull-down analyses (Fig. 3*C*). Analogous mutations have previously been used to investigate FAT-homology domain intermolecular interactions (21). Introduction of these mutations prevents CCM3 pull-down with either full-length CCM2 or the predicted PTB domain of CCM2, but mutation in the N-terminal domain of CCM3 does not alter binding (supplemental Fig. S3). We then co-expressed CCM3 FAT-homology domain (92–212) and full-length CCM2 in bovine aortic endothelial cells (BAEC). We found that there is partial co-localization of CCM2 and the CCM3 FAT-homology domain, and that this occurs at the leading edge. Introduction of HP1 mutations to CCM3 (CCM3–4KD) reduces this partial co-localization at cell leading edges (Fig. 3*D*). We conclude that the CCM3 FAT-homology domain HP1 binding site mediates direct interaction with the PTB domain of CCM2. We are unaware of previously described intermolecular interactions between FAT and PTB domains; therefore we believe that this interaction represents a new class of protein-protein interaction.

CCM3 FAT-homology Domain Can Specifically Bind Paxillin LD-repeats—Strikingly, there is significant structural homology on the surface of CCM3 comprised of helices αG and αH when compared with FAK and Pyk2 (Fig. 4*A* and supplemental Fig. S4). This was a surprising result that was not predicted. We investigated this similarity further and found that on structural superposition of CCM3 with the FAK and Pyk2 FAT domain three of the four surface lysines (Lys-132, Lys-139, Lys-

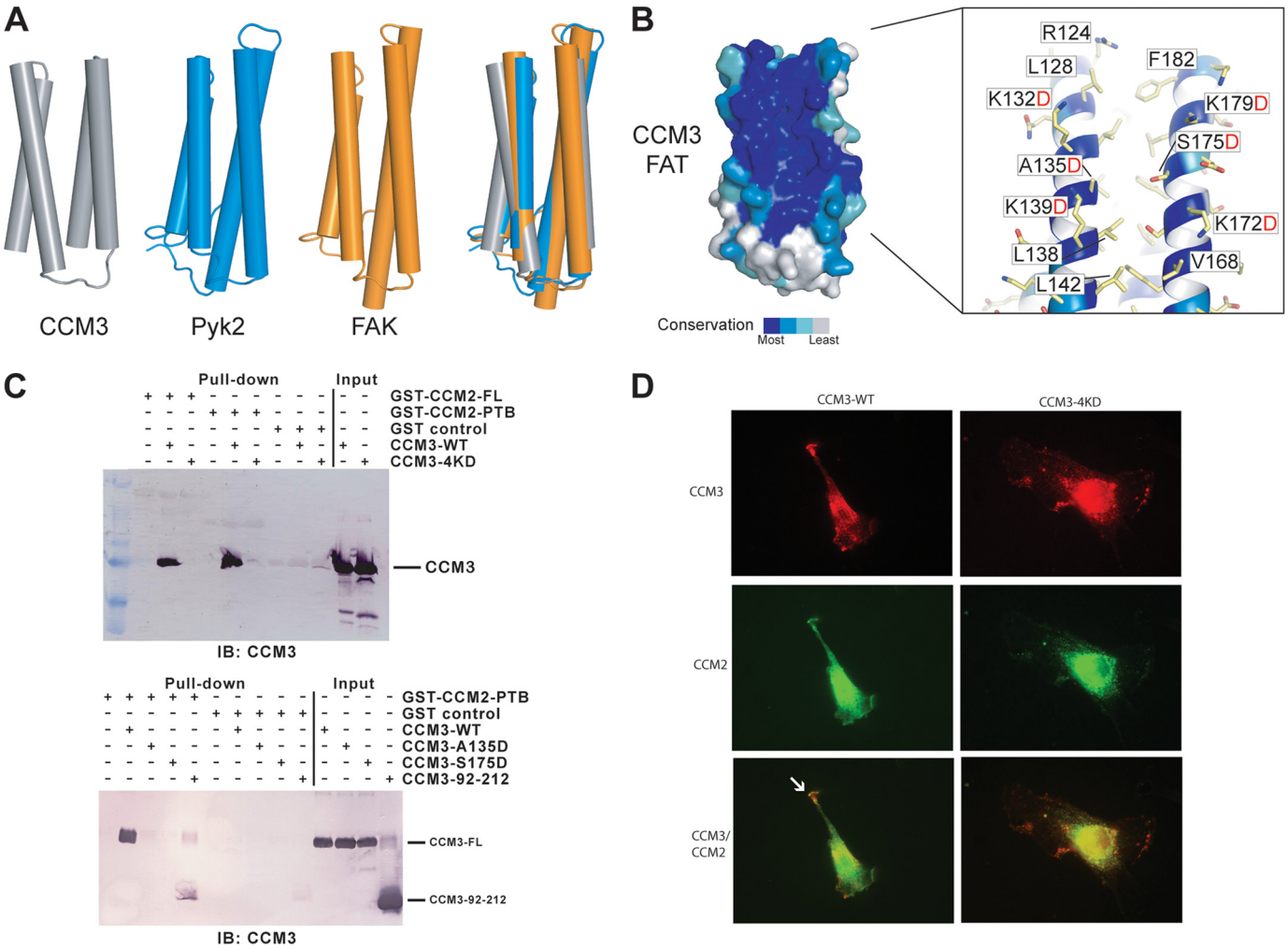

FIGURE 3. **CCM3 FAT-homology domain binds CCM2.** *A*, superposition of CCM3, Pyk2, and FAK. The four helices of Pyk2 FAT domain in *blue* (PDB ID: 3GM1) (22) and FAK FAT domain in *orange* (PDB ID: 1OW7) (23) are superposed with the FAT-homology domain of CCM3 (*gray*). *B*, surface conservation. The CCM3 FAT-homology domain shows very high conservation on the surfaces of helices αG and αH (see alignment in supplemental Fig. S3). The exploded view labels some of the highly conserved residues and those mutated in this study. *C*, pull-down. *Top*: GST fusions of both full-length CCM2 and the predicted CCM2-PTB domain pull-down full-length CCM3. Quadruple mutation of conserved lysines in CCM3, K132D, K139D, K172D, and K179D (CCM3–4KD) renders CCM3 unable to pull-down with CCM2. *Bottom*: CCM3 HP1 mutations A135D and S175D inhibit pull-down of CCM3 with CCM2-PTB domain. Both full-length CCM3 and CCM3–92-212 pull-down with CCM2-PTB. CCM3 constructs were purified by size exclusion chromatography and eluted as monodisperse peaks. *D*, immunofluorescence. Either CCM3 FAT-homology domain (lhs) or CCM3 FAT-homology-4KD (rhs) were co-transfected into BAECs with CCM2 and detected by indirect immunofluorescence. CCM2 and CCM3 FAT-homology partially co-localize (*white arrow*), but introduction of HP1 mutations reduce this co-localization. CCM3 constructs included a C-terminal Flag tag and CCM2, an N-terminal GFP tag.

172) on helices αG and αH of CCM3 correlate extremely well with basic residues in Pyk2 and FAK, and the surface exposed hydrophobic pocket between these helices is structurally conserved (Fig. 4*A* and supplemental Fig. S4). This region of FAK and Pyk2 mediates focal adhesion targeting by interaction with paxillin LD-motifs and structural alignment of FAK and Pyk2 crystal structures bound to LD-motifs (PDB codes: 3GMI, 1OW7, 1OW8) (22, 23) suggests that CCM3 could bind an LD-motif at the HP1 site in a manner analogous to focal adhesion-targeted kinases (Fig. 4*B*). We therefore decided to investigate whether CCM3 could bind paxillin a protein that contains five LD-motifs.

To test whether CCM3 can bind paxillin we initially conducted co-immunoprecipitations and found that CCM3 can co-immunoprecipitate with both paxillin and CCM2 (Fig. 4*C*). Next, we conducted pull-down assays using an N-terminal tagged paxillin LD-motif construct encoding the whole N-terminal region (1–321), which includes all five paxillin LD motifs. Using wild-type CCM3 and the quadruple lysine-to-aspartate mutant that alters the conserved HP1 binding site on the CCM3 FAT-homology domain, CCM3–4KD, we found that the FAT-homology domain of CCM3 can interact with paxillin, and that mutation of HP1 prevents CCM3 pull-down by paxillin (Fig. 4*D*). We then investigated whether CCM3 can bind specific LD-motifs of paxillin. To do this we conducted pull-down experiments using GST-LD-motif constructs (24, 25). We found that CCM3 binds to the LD1, LD2, and LD4 motifs of paxillin (Fig. 4*E*). This interaction is abrogated with the CCM3–4KD mutant (Fig. 4*F*). We then investigated whether mutation of the conserved leucine residues in LD1 could alter CCM3 binding and found that a double leucine to arginine mutation of LD1, LL7, 8RR, prevents binding of CCM3 (Fig. 4*G*), confirming a specific interaction between the CCM3 FAT-homology domain and the paxillin LD1-motif. These mutagenesis studies suggest that CCM3 binds both CCM2 and paxillin

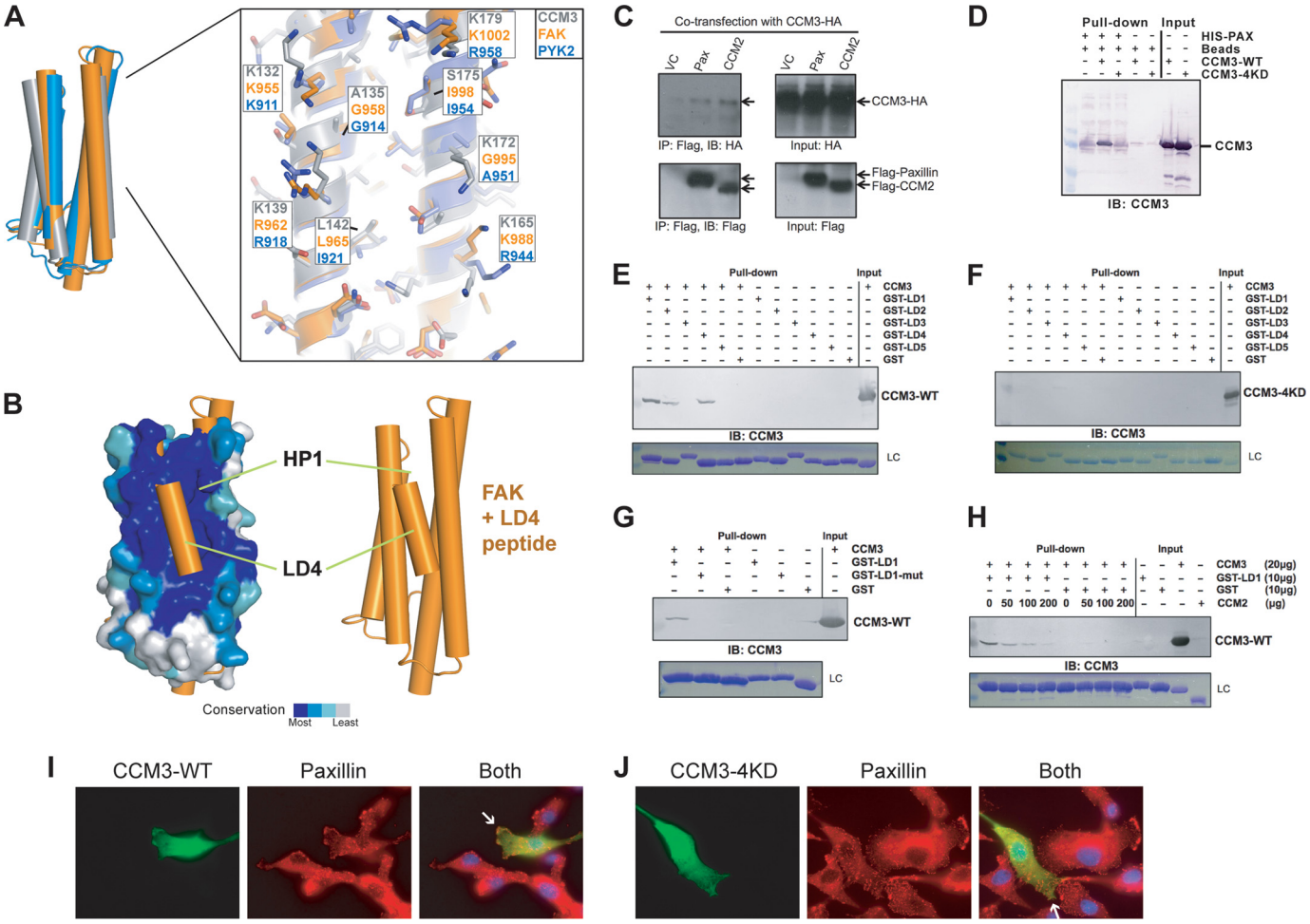

FIGURE 4. **CCM3 binds the LD-motif region of paxillin.** *A*, *left*, superposition of CCM3, Pyk2, and FAK as shown in Fig. 3*A*. *Right*, Pyk2 colored *blue* (PDB ID: 3GM1) (22), FAK in *orange* (PDB ID: 1OW7) (23), and CCM3 (*gray*). Residues are labeled for comparison. *B*, superposition of FAK FAT domain bound to paxillin LD4 peptide (PDB ID: 1OW7) (23) with CCM3 FAT-homology domain suggests that CCM3 can bind LD-motifs. *C*, CCM3 co-immunoprecipitation with CCM2 or paxillin. Full-length CCM3-HA and either full-length Flag-CCM2, full-length Flag-paxillin or vector control (*VC*) were co-transfected into 293T cells. CCM3 immunoprecipitates both Flag-CCM2 and Flag-paxillin. *D*, N-terminal 6×His-tagged paxillin-(1–321) pull-down CCM3, but not CCM3–4KD. *E–G*, CCM3 pull-downs with paxillin. *E*, GST fusions of paxillin LD-motifs (24, 25) pull-down full-length CCM3. *LC*, Coomassie-stained loading control for GST-LD motifs. *F*, quadruple mutation of conserved lysines in CCM3, K132D, K139D, K172D, and K179D (CCM3–4KD) renders CCM3 unable to pull-down with paxillin LD-motifs. *G*, mutation of paxillin LD1, LL7,8RR, prevents CCM3 pull-down. *H*, competition. CCM2 and paxillin compete to bind CCM3. Constant GST-LD1 bound to glutathione-Sepharose beads was incubated with constant CCM3 and increasing CCM2 PTB domain. CCM2 PTB domain competes with GST-LD1 for binding to CCM3. *I* and *J*, immunofluorescence. Full-length CCM3 (*I*) or CCM3–4KD (*J*) were transfected into BAECs. Transfected CCM3 and endogenous paxillin were detected by indirect immunofluorescence microscopy with anti-Flag (rabbit; for CCM3) and anti-paxillin (mouse), followed by FITC-conjugated anti-rabbit and TRITC-conjugated anti-mouse secondary antibodies. CCM3 partially co-localizes with endogenous paxillin at the leading edges of migrating cells. Localization of overexpressed CCM3–4KD to leading edges is reduced (indicated with *white arrows*).

using the same HP1 binding site within the CCM3 FAT-homology domain. We therefore investigated whether CCM2 and paxillin can compete to bind CCM3 using a pull-down competition assay. We bound GST-LD1 to glutathione-Sepharose beads and incubated with constant concentration of CCM3 and increasing concentrations of CCM2 PTB domain. We found that the CCM2 PTB domain can compete with GST-LD1 for binding to CCM3 (Fig. 4*H*), again suggesting an overlapping binding site.

CCM3 Partially Localizes with Paxillin to Leading Edges—We then overexpressed full-length CCM3 and CCM3–4KD in bovine aortic endothelial cells (BAEC). Indirect immunofluorescence microscopy was conducted. Wild-type CCM3 partially co-localizes with endogenous paxillin at leading edges (Fig. 4*I*) and localization of overexpressed CCM3–4KD to leading edges is reduced (Fig. 4*J*).

DISCUSSION

In this study we have determined the first crystal structure of full-length human CCM3. We show that this is a two-domain protein that contains a novel N-terminal dimerization domain and an unexpected C-terminal FAT-homology domain with poor sequence homology but high structural similarity to focal adhesion targeting FAT domains. We show that CCM3 dimerizes in cells and *in vitro*, and that the C-terminal FAT domain is responsible for direct interaction of CCM3 with CCM2. We also show that the CCM3 FAT-homology domain has a highly conserved surface that is structurally very similar to the LD-motif binding sites of FAK and Pyk2, and that this surface can bind the paxillin LD1-motif. We also show that paxillin LD1 and CCM2 can compete to bind the CCM3 FAT domain suggesting an overlapping binding site. Localization of CCM3 with CCM2 and paxillin within the cell was investigated and we

found that there is partial co-localization of CCM3 with CCM2 and with paxillin, but that this co-localization is reduced on introduction of mutations in the CCM3 FAT-homology HP1 binding site. Together these results provide a significant advance in the understanding of the roles of the CCM proteins.

The CCM3 dimerization domain described here is the first observation of this fold. We observe this interaction as a large hydrophobic interface (~3700 Å^2 total buried surface) in two crystal forms. Our *in vitro* studies using SEC suggest a dimer in solution, and co-immunoprecipitation of alternately tagged CCM3 from cells suggests CCM3 homo-oligomerization. We therefore conclude that CCM3 forms a dimer both *in vitro* and *in vivo*. Point mutations designed to disrupt the observed N-terminal dimerization interface (double mutations LA44/47DD, IL66/67DD, EF73/76RR, and quadruple mutation LAIL44/47/66/67DDDD) do not alter oligomerization with full-length protein by immunoprecipitation from cells and do not have an observable effect on elution from SEC (data not shown, EF73/76RR only with SEC), although *E. coli* expression of soluble protein for these constructs is reduced. From this we conclude that the large dimerization interface with low surface complementarity is able to tolerate charge insertions. Overall, our SEC and co-immunoprecipitation data show that CCM3 exists as a dimer *in vitro* and *in vivo* and that this dimerization is mediated by the N-terminal domain of CCM3 and not by the C-terminal FAT-homology domain.

The discovery of a FAT-homology domain at the C terminus of CCM3 was unexpected. Sequence homology, molecular modeling, and secondary structure analysis all failed to discover this domain (17). The CCM3 FAT-homology domain contains an extremely well conserved surface patch, HP1, for which many of the surface lysine and hydrophobic residues are invariant (supplemental Fig S4*B*). This suggests an evolutionarily critical role for this domain and for CCM3. We show (Fig. 3) that the CCM3 FAT-homology domain is responsible for direct interaction with the PTB domain of CCM2, and hypothesize that this direct interaction could be the reason for the stunning CCM3 sequence conservation in this region. We investigated the localization of overexpressed CCM3 FAT-homology domain and CCM2 in BAECs and found that wild-type CCM3 FAT-homology domain partially co-localizes with CCM2, but that co-localization is reduced on introduction of CCM3 HP1 mutations (Fig. 3*D*). Overall, these results along with structural alignment of the FAT-homology domain of CCM3 with the focal adhesion targeting FAT domains of both FAK and Pyk2 suggests that the conserved surface HP1 interacts with CCM2 (Fig. 4*B*).

The surprising structural homology of CCM3 to the focal adhesion targeting domains of FAK and Pyk2 lead us to investigate whether CCM3 could directly interact with the localization target for these kinases, the LD-motif containing protein, paxillin. To do this we first conducted co-immunoprecipitations to both confirm that CCM3 can interact with CCM2 *in vivo* and to investigate whether CCM3 and paxillin interact. We found that both CCM2 and paxillin can co-immunoprecipitate with CCM3 (Fig. 4*C*). We then investigated whether CCM3 can interact with paxillin *in vitro*. To do this we used the paxillin N-terminal region that contains five LD-motifs, and GST fusions of each of these LD-motifs. We found that CCM3 pulls down with paxillin. Incorporation of mutations into CCM3 (HP1 mutants) and paxillin LD1 (mutation of the two central leucine residues in LD1) prevents CCM3 pull-down (Fig. 4, *D–G*). We then used the pull-down assay to assess whether the interaction between CCM3 and both CCM2 and paxillin is competitive, and showed that increasing concentrations of CCM2 reduce CCM3 binding to paxillin LD1, again suggesting that an overlapping binding site is utilized for CCM3 binding to CCM2 and paxillin. We then found that overexpressed CCM3 partially co-localizes with endogeneous paxillin at leading edges and that this co-localization is reduced on introduction of mutations in the CCM3 HP1 surface.

Overall, the highly conserved HP1 surface on the FAT-homology domain of CCM3 seems to be important for direct interactions with both CCM2 and an LD-motif containing protein, maybe paxillin. We suggest that our data indicates a direct competition between CCM2 and this LD-motif containing protein to bind the HP1 site of CCM3. As the spatial and temporal localization of CCM2 and CCM3 have previously been shown not to be identical (18) we hypothesize that these CCM3 FAT-homology domain interactions could be regulated in the cell, by for example post-translational modification.

Cerebral cavernous malformations are sometimes-fatal dysplasias that can result in hemorrhagic stroke, epilepsy and focal neurological outcomes. The discovered mutations in *CCM3* that are associated with CCM in the human disease result in frame shifts and truncations of the expressed protein, and almost always affect the FAT-homology domain (7–9), indicating the importance of this domain for normal function of endothelial cells. An example truncation mutation, CCM3–1-117, was tested in this study and found to be poorly expressed, implying that inclusion of a fully folded CCM3 FAT-homology domain is important for stabilization of the expressed protein in the *in vivo* setting. Therefore, truncation of CCM3 results in a loss of critical protein-protein interactions with CCM2 and LD-motif-containing proteins, and in the destabilization of endothelial cell-cell interactions seen in CCM. These truncations invariably affect the CCM3 FAT-homology domain, suggesting functional importance for this domain in the regulation of endothelial cell stability. Overall, the discovery that CCM3 is a multidomain dimeric protein that contains a FAT-homology domain which can bind both CCM2 and LD-motifs of paxillin provides a significant advance that will allow for improved understanding of the signaling pathways involved in acquisition of cerebral cavernous malformation, and the functions of the CCM proteins in development and maintenance of endothelial cells in the vasculature.

Acknowledgments—We thank David Calderwood, Yuhang Chen, Joseph Schlessinger, Valsan Mandiyan, Karin Reinisch, Ya Ha, Demetrios Braddock, Amy Stiegler, Nilda Alicea-Velazquez, and Ewa Folta-Stogniew. The paxillin-(1–321) construct was kindly provided by Joseph Schlessinger. Paxillin GST-LD-motif plasmids were kindly provided by Christopher Turner. Modified pET-32 was designed by F. Poy. We acknowledge the Beamlines X6A and X29 at NSLS and NE-CAT at APS.

REFERENCES

1. Revencu, N., and Vikkula, M. (2006) *J. Med. Genet.* **43,** 716–721
2. Labauge, P., Denier, C., Bergametti, F., and Tournier-Lasserve, E. (2007) *Lancet Neurol.* **6,** 237–244
3. Sahoo, T., Johnson, E. W., Thomas, J. W., Kuehl, P. M., Jones, T. L., Dokken, C. G., Touchman, J. W., Gallione, C. J., Lee-Lin, S. Q., Kosofsky, B., Kurth, J. H., Louis, D. N., Mettler, G., Morrison, L., Gil-Nagel, A., Rich, S. S., Zabramski, J. M., Boguski, M. S., Green, E. D., and Marchuk, D. A. (1999) *Hum. Mol. Genet.* **8,** 2325–2333
4. Günel, M., Awad, I. A., Anson, J., and Lifton, R. P. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92,** 6620–6624
5. Liquori, C. L., Berg, M. J., Siegel, A. M., Huang, E., Zawistowski, J. S., Stoffer, T., Verlaan, D., Balogun, F., Hughes, L., Leedom, T. P., Plummer, N. W., Cannella, M., Maglione, V., Squitieri, F., Johnson, E. W., Rouleau, G. A., Ptacek, L., and Marchuk, D. A. (2003) *Am. J. Hum. Genet.* **73,** 1459–1464
6. Denier, C., Goutagny, S., Labauge, P., Krivosic, V., Arnoult, M., Cousin, A., Benabid, A. L., Comoy, J., Frerebeau, P., Gilbert, B., Houtteville, J. P., Jan, M., Lapierre, F., Loiseau, H., Menei, P., Mercier, P., Moreau, J. J., Nivelon-Chevallier, A., Parker, F., Redondo, A. M., Scarabin, J. M., Tremoulet, M., Zerah, M., Maciazek, J., and Tournier-Lasserve, E. (2004) *Am. J. Hum. Genet.* **74,** 326–337
7. Bergametti, F., Denier, C., Labauge, P., Arnoult, M., Boetto, S., Clanet, M., Coubes, P., Echenne, B., Ibrahim, R., Irthum, B., Jacquet, G., Lonjon, M., Moreau, J. J., Neau, J. P., Parker, F., Tremoulet, M., and Tournier-Lasserve, E. (2005) *Am. J. Hum. Genet.* **76,** 42–51
8. Guclu, B., Ozturk, A. K., Pricola, K. L., Bilguvar, K., Shin, D., O'Roak, B. J., and Gunel, M. (2005) *Neurosurgery* **57,** 1008–1013
9. Stahl, S., Gaetzner, S., Voss, K., Brackertz, B., Schleider, E., Sürücü, O., Kunze, E., Netzer, C., Korenke, C., Finckh, U., Habek, M., Poljakovic, Z., Elbracht, M., Rudnik-Schöneborn, S., Bertalanffy, H., Sure, U., and Felbor, U. (2008) *Hum. Mutat.* **29,** 709–717
10. Akers, A. L., Johnson, E., Steinberg, G. K., Zabramski, J. M., and Marchuk, D. A. (2009) *Hum. Mol. Genet.* **18,** 919–930
11. Pagenstecher, A., Stahl, S., Sure, U., and Felbor, U. (2009) *Hum. Mol. Genet.* **18,** 911–918
12. Gault, J., Shenkar, R., Recksiek, P., and Awad, I. A. (2005) *Stroke* **36,** 872–874
13. Boettner, B., and Van Aelst, L. (2009) *Curr. Opin. Cell Biol.* **21,** 684–693
14. Whitehead, K. J., Plummer, N. W., Adams, J. A., Marchuk, D. A., and Li, D. Y. (2004) *Development* **131,** 1437–1448
15. Whitehead, K. J., Chan, A. C., Navankasattusas, S., Koh, W., London, N. R., Ling, J., Mayo, A. H., Drakos, S. G., Jones, C. A., Zhu, W., Marchuk, D. A., Davis, G. E., and Li, D. Y. (2009) *Nat. Med.* **15,** 177–184
16. He, Y., Zhang, H., Yu, L., Gunel, M., Boggon, T. J., Chen, H., and Min, W. (2010) *Sci. Signal* **3,** ra26
17. Voss, K., Stahl, S., Schleider, E., Ullrich, S., Nickel, J., Mueller, T. D., and Felbor, U. (2007) *Neurogenetics* **8,** 249–256
18. Zawistowski, J. S., Stalheim, L., Uhlik, M. T., Abell, A. N., Ancrile, B. B., Johnson, G. L., and Marchuk, D. A. (2005) *Hum. Mol. Genet.* **14,** 2521–2531
19. Zhang, J., Rigamonti, D., Dietz, H. C., and Clatterbuck, R. E. (2007) *Neurosurgery* **60,** 353–359
20. Serebriiskii, I., Estojak, J., Sonoda, G., Testa, J. R., and Golemis, E. A. (1997) *Oncogene* **15,** 1043–1049
21. Hayashi, I., Vuori, K., and Liddington, R. C. (2002) *Nat. Struct. Biol.* **9,** 101–106
22. Lulo, J., Yuzawa, S., and Schlessinger, J. (2009) *Biochem. Biophys. Res. Commun.* **383,** 347–352
23. Hoellerer, M. K., Noble, M. E., Labesse, G., Campbell, I. D., Werner, J. M., and Arold, S. T. (2003) *Structure* **11,** 1207–1217
24. Nikolopoulos, S. N., and Turner, C. E. (2000) *J. Cell Biol.* **151,** 1435–1448
25. Nikolopoulos, S. N., and Turner, C. E. (2002) *J. Biol. Chem.* **277,** 1568–1575
26. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
27. Davis, I. W., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2004) *Nucleic Acids Res.* **32,** W615–619
28. CCP4 (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
29. Krissinel, E., and Henrick, K. (2007) *J. Mol. Biol.* **372,** 774–797
30. Bahadur, R. P., Chakrabarti, P., Rodier, F., and Janin, J. (2004) *J. Mol. Biol.* **336,** 943–955
31. Holm, L., and Sander, C. (1995) *Trends Biochem. Sci.* **20,** 478–480
32. Kim, B. H., Cheng, H., and Grishin, N. V. (2009) *Nucleic Acids Res.* **37,** W532–538
33. Gibrat, J. F., Madej, T., and Bryant, S. H. (1996) *Curr. Opin. Struct. Biol.* **6,** 377–385
34. Krissinel, E., and Henrick, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2256–2268
35. Arold, S. T., Hoellerer, M. K., and Noble, M. E. (2002) *Structure* **10,** 319–327
36. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005) *Nucleic Acids Res.* **33,** W299–302