SUPPLEMENTAL MATERIALS FOR

CRYSTAL STRUCTURE OF CCM3, A CEREBRAL CAVERNOUS MALFORMATION PROTEIN CRITICAL FOR VASCULAR INTEGRITY

Xiaofeng Li*[1], Rong Zhang*[1], Haifeng Zhang[2], Yun He[2], Weidong Ji[2], Wang Min[+2] and Titus J. Boggon[+1]
Depts of Pharmacology[1] and Pathology[2], Yale University School of Medicine, New Haven, CT 06520
* These authors contributed equally. **Running head:** Crystal structure of CCM3
+ Address correspondence to: Wang Min, Ph.D., Department of Pathology, Yale University School of Medicine, 10 Amistad St., 401B, New Haven, CT 06520-8023, Phone: 203-785-6047, Fax: 203-737-2293, E-mail: wang.min@yale.edu and Titus J. Boggon, Ph.D., Department of Pharmacology, Yale University School of Medicine, 333 Cedar St., SHM B-316A, New Haven, CT 06520. Phone: 203-785-2943. Fax: 203-785-5494. E-mail: titus.boggon@yale.edu

Protein expression and purification - Full-length human CCM3 (SwissProt: Q9BUL8) was sub-cloned into a modified pET-32 vector with an N-terminal polyhistidine-tag. Recombinant CCM3 was produced in *Escherichia coli* ROSETTA (Novagen) by induction with 0.2 mM IPTG at 16°C, OD_{600} 0.6. Cell pellets were resuspended in lysis buffer (20 mM Tris, 500 mM NaCl, 20 mM imidazole, 10% glycerol, pH 8.0) and immediately lysed by 3 cycles of freeze-thaw using dry ice with ethanol and water at room temperature followed by 3 cycles of sonication for 5 seconds. The supernatant was applied to HisTrap HP chelating affinity column (GE Healthcare) initially and eluted with 400 mM imidazole in lysis buffer. The fusion protein was cleaved by TEV protease at 4°C overnight, dialyzed against 20 mM Tris, 500 mM NaCl, 10% glycerol, pH 8.0, and loaded onto a Mono-S (GE Heathcare) ion-exchange chromatography column with 20 mM MES (pH 6.0) and eluted over a 5 to 500 mM NaCl gradient. Protein purity was estimated by SDS-PAGE to be greater than 98%. For crystallization CCM3 was concentrated to 10 mg/mL in 20 mM MES, pH 6.0, 200 mM NaCl.

Structure solution - Two crystal forms of CCM3, tetragonal and orthorhombic, were grown by hanging-drop vapor diffusion using 1 µL volumes of protein and precipitant solutions. The tetragonal crystal form (space group $P4_122$, unit cell dimensions a=b=77.4 Å, c=108.6 Å), grew in 1 month at 4 °C from a reservoir solution of 0.16 M calcium acetate, 0.08 M sodium cacodylate pH 6.5, 20% glycerol, 13.5% PEG 8000, to crystal sizes up to 600 x 100 x 100 µm. Following dehydration by equilibration against 18% PEG 8000 for two days at 4 °C, the best tetragonal crystal diffracted anisotropically to a maximum of 3.05 Å resolution. Selenomethionine derivatives grew from similar conditions. The orthorhombic crystal form (space group $P2_12_12_1$, unit cell dimensions a=63.1 Å, b=116.0 Å, c=123.0 Å) grew in 7 days at room temperature. Precipitant conditions of 0.1 M potassium fluoride and 13% PEG 3350 were used to obtain crystals with dimensions up to 300 x 120 x 50 µm. For derivatization, the orthorhombic crystal form was soaked in precipitant solution substituted with 1 mM $UO_2(C_2H_3O_2)_2$ for 2 hours (a generous gift from Yuhang Chen, W. Hendrickson lab) and data collected at λ = 0.9551. Anisotropic diffraction for the orthorhombic crystal form to 2.50 Å for the native and 3.10 Å for the uranium derivative was observed. Crystals were flash frozen in liquid nitrogen using crystallization conditions for the tetragonal crystals and by adding 20% glycerol to the mother liquor for the orthorhombic crystals. The best resolution data for both crystal forms were obtained at the beamline X6A at the National Synchrotron Light Source (NSLS) (**Table 1**). All data were processed using HKL2000 (37).

Initial phases for CCM3 were obtained by single isomorphous replacement with anomalous scattering (SIR/AS) for the native and uranium derivative orthorhombic crystals (**Fig S1**). Automated phasing using autoSHARP found two uranium atoms and calculated initial phases (38). An initial Cα trace was manually built for the orthorhombic crystal form using the program COOT (39). Model refinement was conducted in REFMAC5 (28) using a maximum-likelihood target with TLS and NCS to 2.50 Å resolution

(Table 1). The refined orthorhombic model comprises a total of 814 residues in four chains and 69 water molecules. Following refinement of the orthorhombic crystal form of CCM3, molecular replacement for the tetragonal crystal form was conducted with PHASER (28) using the N-terminal domain of chain B and the C-terminal domain of chain C. The best solution gave a translation factor Z-score of 22.5 and an *R*-factor of 47.1%. Model refinement was conducted in REFMAC5 (28) using a maximum-likelihood target with TLS, and included all data to 3.05 Å resolution. The register of this model of CCM3 was confirmed by phased anomalous difference Fourier maps for a 4.2 Å selenomethionine dataset, 6 of 10 selenium atoms show anomalous signal, (residues Met1, Met3, Met5, Met8 are disordered in the tetragonal crystal form). The refined tetragonal model comprises a total of 197 residues in one chain and no water molecules (Table 1). Atomic coordinates for both models have been deposited into the Protein Data Bank with the accession numbers 3L8I and 3L8J.

Phased anomalous difference Fourier maps show that the $P4_122$ crystal form is centered around a 6-methionine zipper comprising Met83'', Met17', Met20, Met20', Met17, Met83''' (**Fig S1**) allowing a compact crystal lattice in one direction, however, the two orthogonal directions have very poor crystal contacts mediated solely by the loop between helices αG and αH (residues 153 to 158). Only upon dehydration of our best $P4_122$ crystal were we able to obtain the anisotropic 3.05 Å resolution diffraction data. Interestingly, the $P2_12_12_1$ crystal form also contains a methionine zipper of these same residues. Crystal packing between CCM3 dimers is similar for both crystal forms where the dimerized CCM3 stacks by means of a hydrophobic interaction that is over 1200 $Å^2$ in size for each molecule (PISA) (29). This interaction is broadly flat, and mediated by the N-terminal dimerization domains and helix αE, centered around the 6-methionine zipper (**Fig S1**). Pro28 in this surface is invariant, but other residues in this interface are not as well conserved as other parts of CCM3. This interface may represent a crystal packing artifact, but there is also the potential that it could represent a functionally important surface for CCM3 oligomerization or protein-protein interactions.

Pull-downs - In vitro pull-down assays were conducted by incubating 5 µg of purified N-terminal GST-tagged CCM2 (full-length, 1-438; PTB, 51-251) bound to glutathione-Sepharose beads (GE). A quadruple lysine to aspartate (K132D, K139D, K172D, K179D) mutant CCM3 was generated by site-directed mutagenesis (CCM3-4KD) as were single point mutants A135D (CCM3-A135D) and S175D (CCM3-S175D). Both wild-type CCM3, CCM3-4KD, CCM3-A135D and CCM3-S175D were purified as above by affinity, ion exchange and size exclusion chromatography, all four constructs eluted with an indistinguishable profile from wild-type by size exclusion chromatography. 10 µg of purified CCM3 was added to a total volume of 400 µL of lysis buffer, 150 mM NaCl, 50 mM Tris pH 7.5, 0.1% Triton X100. The samples were incubated with rocking at 4 °C for 2 hours, spun down and washed three times with 1 mL lysis buffer. 30 µL of beads and 2x sample buffer were analyzed by Western blot using a CCM3 antibody. CCM3 antibody was generated using full-length recombinant human CCM3 protein expressed and purified from *E. coli*. The antigen was injected into two rabbits and the antibody affinity purified and stored at -80 °C. Immunization and purification were conducted by Invitrogen (Carlsbad, CA).

LD-motif pull-down assays. GST-LD-motif constructs were provided by Christopher Turner and encoded paxillin LD-motifs 1 through 5 (24,25). Mutant GST-LD1 was made by QuikChange mutation of L7/L8 to R7/R8. 10 µg of purified paxillin GST-LD-motifs were bound on 20 µL of beads and incubated with 20 µg of purified CCM3 or CCM3-4KD for 2 hrs at 4°C in 500 µL pull-down buffer; 50 mM Tris pH 7.5, 150 mM NaCl, 0.1% Triton X100, protease inhibitors (Roche). The beads were then washed 3 times with 1 ml of pull-down buffer and analyzed by Western blot using anti-CCM3.

N-terminal paxillin pull-down assay. His-paxillin (1-321) provided by Joseph Schlessinger was bound to 20 µL of beads and incubated with 20 µg of purified CCM3 or CCM3-4KD for 2 hrs at 4°C in 500 µL pull-down buffer supplemented with 20 mM imidazole. The beads were then washed 3 times with 1 ml of pull-down buffer supplemented

with 20 mM imidazole and analyzed by Western blot using anti-CCM3.

Competition assay. 10 μg of purified paxillin GST-LD1 or GST was bound on 20 μL of beads and incubated with 20 μg of purified CCM3 and increasing amounts of purified CCM2 PTB domain (51-251) for 2 hrs at 4°C in 500 μL pull-down buffer, 50 mM Tris pH 7.5, 150 mM NaCl, 0.1% Triton X100, protease inhibitors (Roche). The beads were then washed 3 times with 1 ml of pull-down buffer and analyzed by Western blot using anti-CCM3.

Co-immunoprecipitation - CCM3 dimerization. Full-length CCM3-HA was co-transfected with Flag-CCM3 variants into 293T cells. Flag-CCM3 constructs included residues full-length CCM3, 1-117, and 92-212. Association of CCM3-HA with CCM3 variants 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.

CCM3 interactions with CCM2 or paxillin. Full-length CCM3-HA and either full-length Flag-CCM2 or full-length Flag-paxillin were co-transfected into 293T cells. 40 μL of cell lysates were diluted into 400 μL of NP-40 buffer (50 mM Tris pH 7.7, 150 mM NaCl, 1% NP-40, protease inhibitors). 1.5 μL of antibody were added and incubated at 4°C for 30 min. 10 μL of protein A/G agarose were added, rocked for 2 hr at 4°C and then washed 3 times with 1 mL of NP-40 buffer.

Immunofluorescence - Full-length wild-type CCM3 or full-length CCM3 with quadruple mutation of Lys132, Lys139, Lys172 and Lys179 to Asp (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 4KD 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.

REFERENCES

37. Otwinowski, Z., and Minor, W. (1997) *Methods in Enzymology* **276, Part A**, 307-326
38. Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007) *Methods Mol Biol* **364**, 215-230
39. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1), 2126-2132

Figure S1. Experimental data, crystal packing and topology of CCM3. **A)** Initial SIR/AS experimental map of CCM3 in the $P2_12_12_1$ crystal form obtained using SHARP. Map is contoured at 1.5 σ. Cα backbone trace for the refined model is shown. **B)** Phased anomalous difference Fourier map for the $P4_122$ crystal form selenomethionine dataset to 4.2 Å resolution confirms the register of the methionine residues. The selenomethionine data also confirms the location of a methionine zipper formed with six methionines in a linear arrangement. Map is contoured at 4.0 σ. **C)** Packing of the $P4_122$ crystal form showing poor crystal contacts that contribute to highly anisotropic diffraction in two directions. Panel colored by rainbow from blue to red indicating increasing temperature factor. **D)** The six-methionine zipper for CCM3 is conserved in one crystallographic interaction within the $P2_12_12_1$ crystal form. $2F_{obs}$-F_{calc} map in blue contoured at 1.0 σ. **E)** Topology schematic for CCM3. The two molecules of the CCM3 dimer are shown. Helices for one molecule are colored as in Fig 1C, with helix αA in grey, αB-αD in green, αE and loop EF in orange and αF-αI in red. The locations of residues that mediate the dimerization interface are colored purple. Start and end residues for all helices are shown.

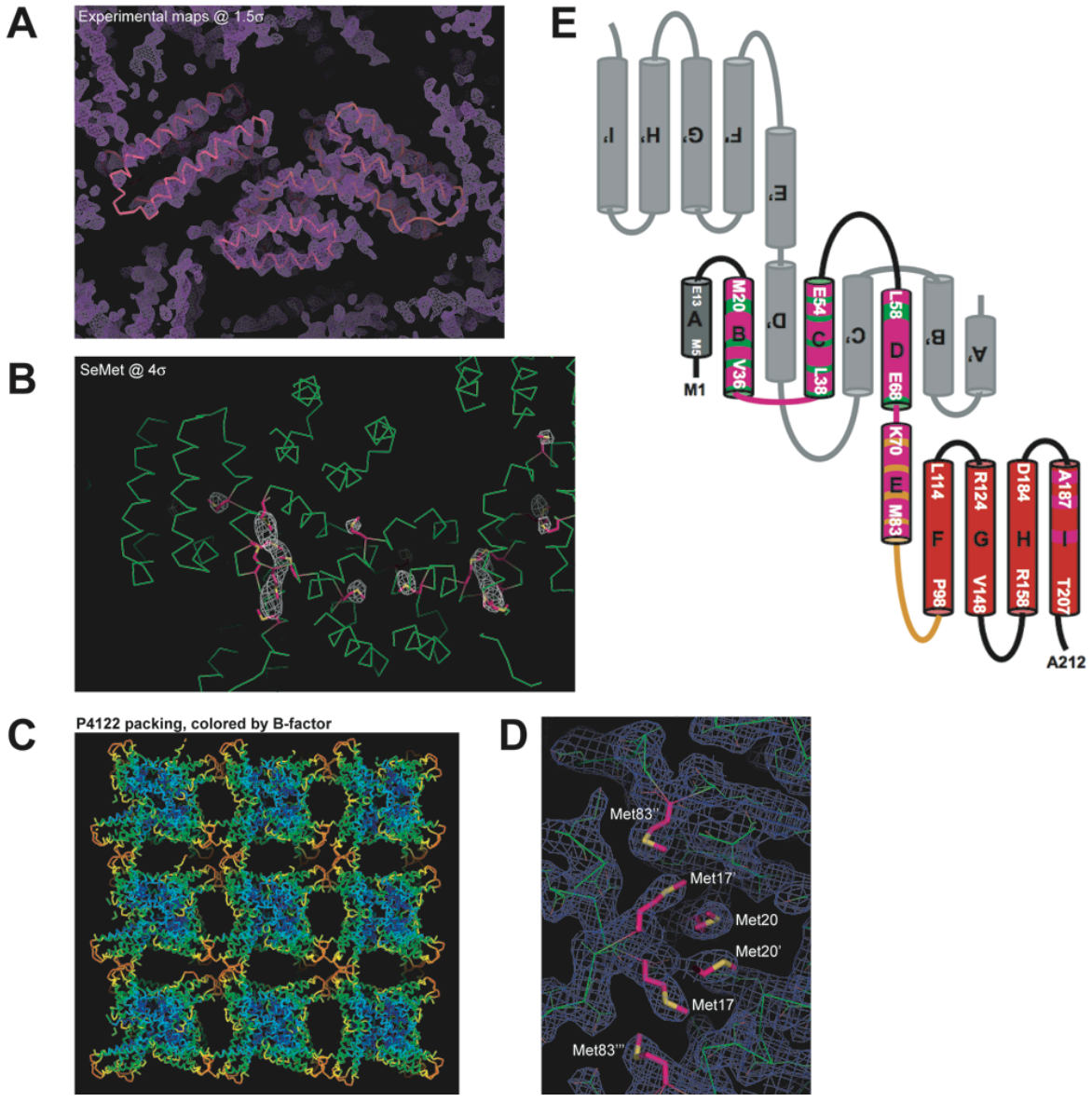


Figure S2. CCM3 dimerization. A and B) Change in surface accessibility was measured using NACCESS (S. Hubbard and J. Thornton, www.bioinf.manchester.ac.uk/naccess/) and plotted against residue number (**A**) and on the surface of molecule C (**B**) using the same orientations as in Fig 2. (**B**) Colored according to average changes in surface accessibility on dimerization, red through green. **C)** Crosslinking. Chemical cross-linking of native CCM3 (3 mg/mL buffer-exchanged to 20mM HEPES, pH 7.5) with increasing final concentrations of chemical cross-linker, BS^3 (Thermo Scientific) (0.25 mM, 0.50 mM, 1 mM, 2.5 mM, 5 mM). Quencing with 50 mM, Tris pH 7.5 for 15 minutes followed incubation for 2 hours on ice. Cross-linked products were separated on a 12% SDS–PAGE followed by Coomassie brilliant blue staining. Protein markers (Marker) are shown in kDa. Bands corresponding to monomer and dimer are indicated.

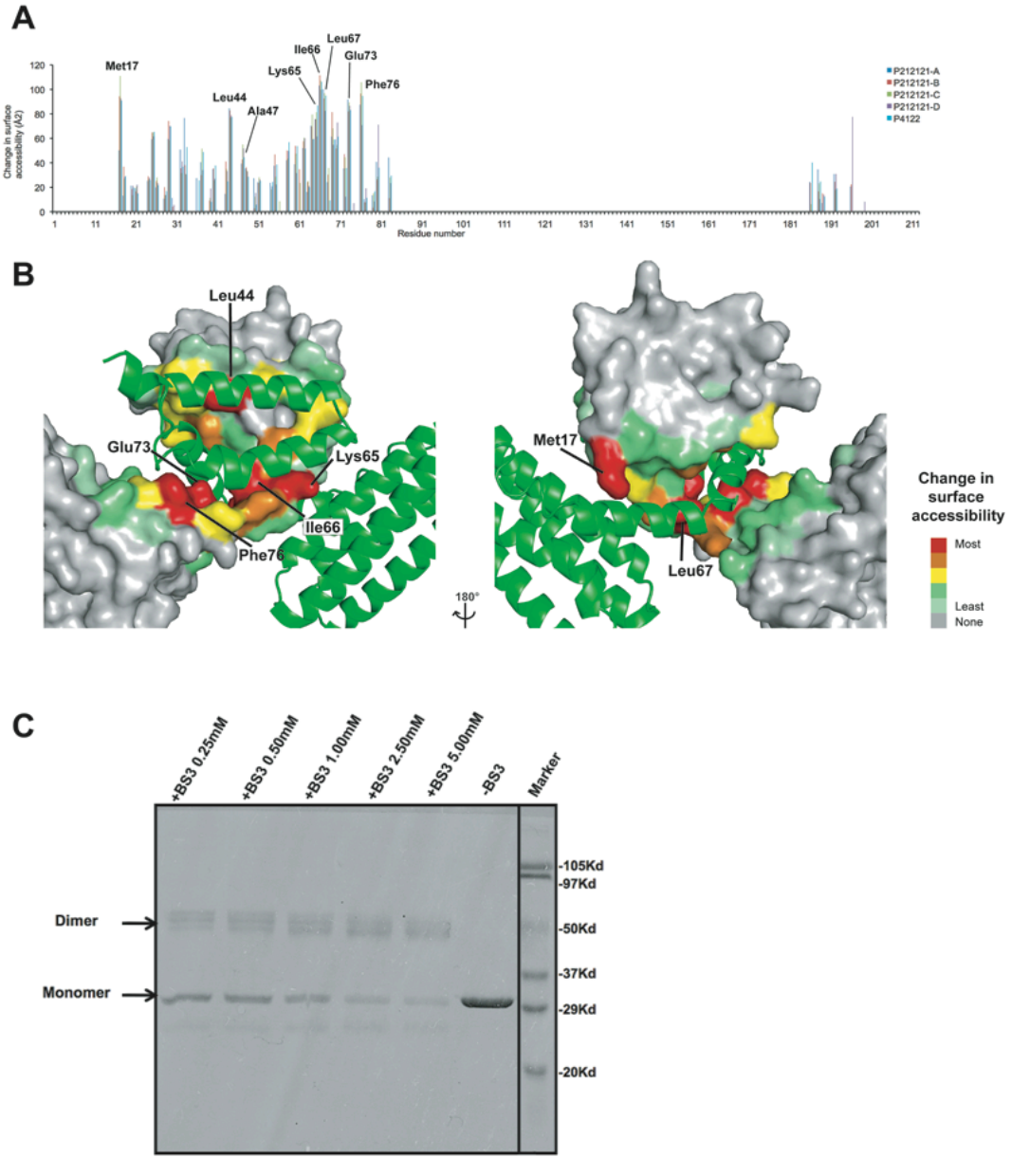
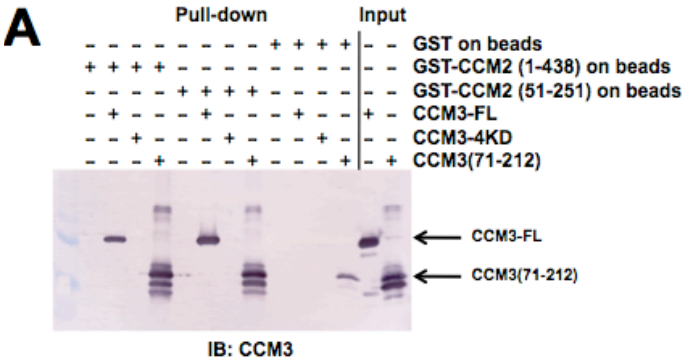
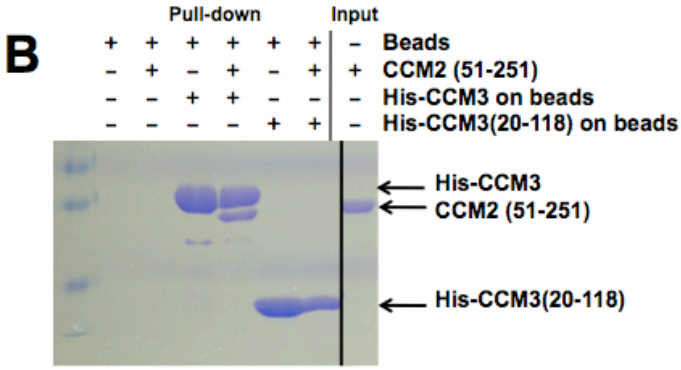


Figure S3. Pull-downs for CCM3 and CCM2. A) The C-terminal CCM3 construct 71-212 that includes helix αE, loop EF and the FAT domain pulls-down with GST fusions of both full-length CCM2 (1-438) and the CCM2 PTB domain (51-251). As shown in Fig 3, full-length CCM3 but not the quadruple mutant CCM3-4KD pulls-down with CCM2. **B)** Coomassie stained gel showing that the CCM2 PTB domain pulls-down with His-tagged full-length CCM3 on Ni-beads. CCM2 PTB domain fails to pull-down with a His-tagged N-terminal construct of CCM3 encoding residues 20 to 118. **C)** Double mutation of residues Ile66 to Asp and Leu67 to Asp (CCM3-IL/DD) does not prevent CCM3 pulling down with GST fusions of full-length or PTB domain of CCM2.





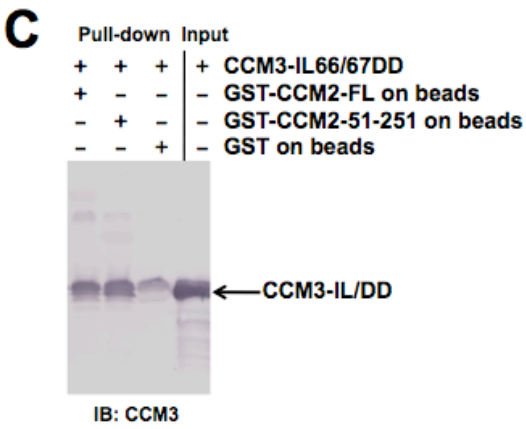
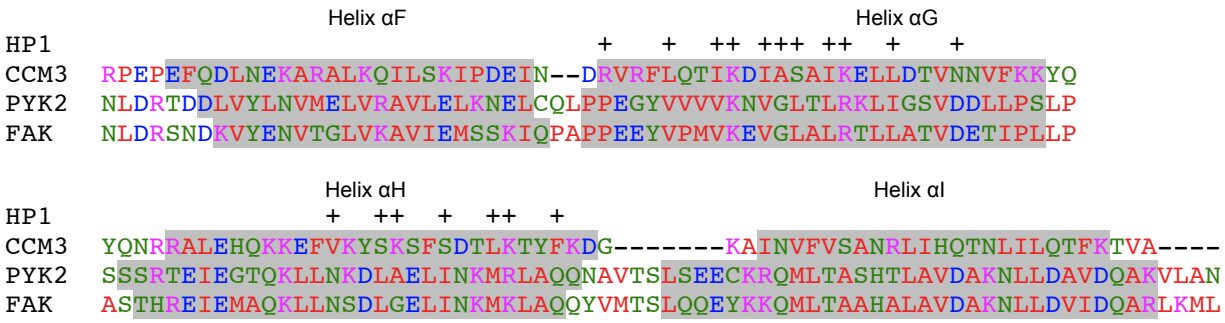
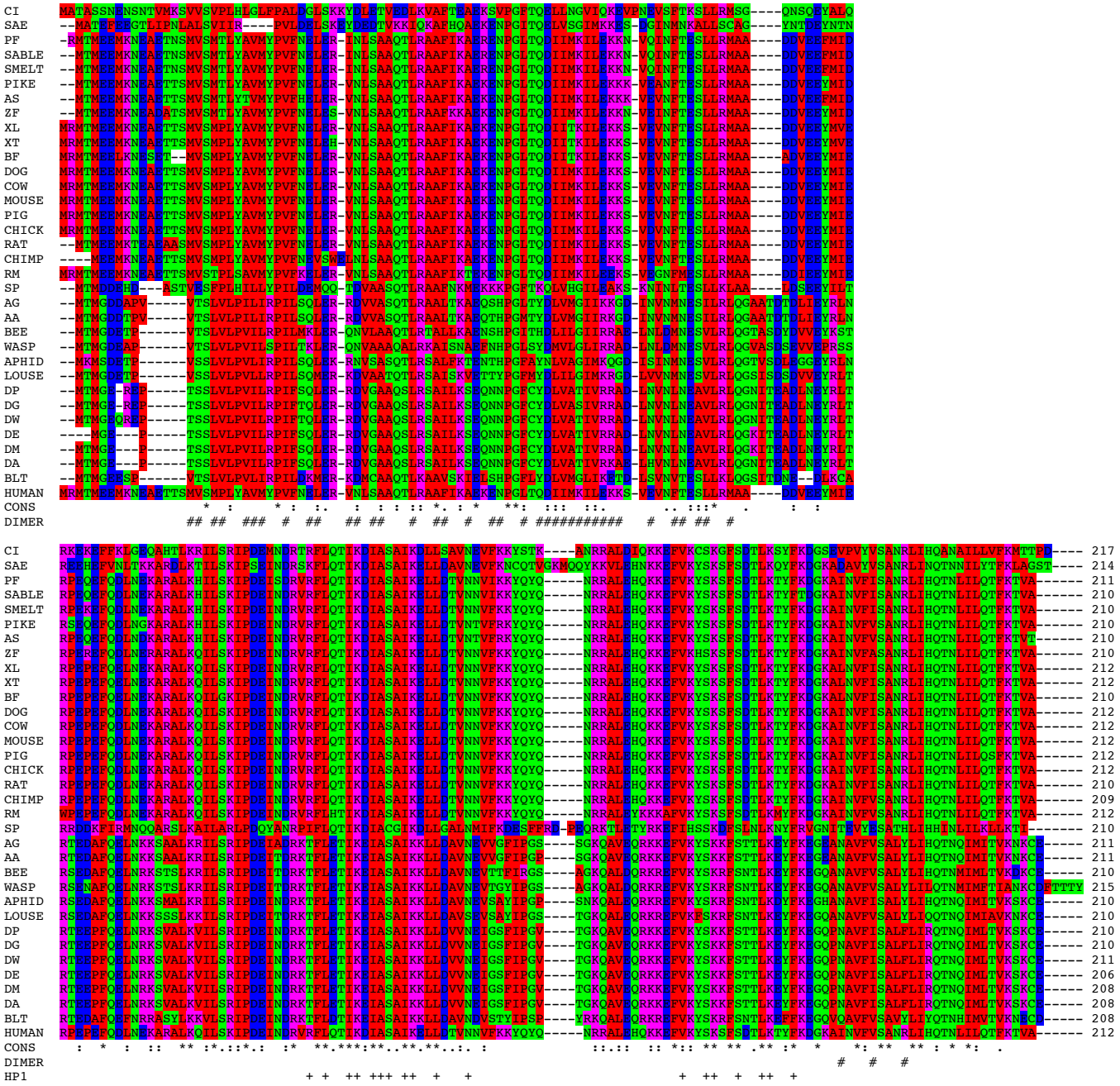


Figure S4. CCM3 sequence alignments. A) Comparison of FAT domains for PYK2, FAK and CCM3. Structure-based sequence alignment of the FAT domains of PYK2 (SwissProt: Q14289, residues 868 to 1009), FAK (SwissProt: Q05397, residues 912 to 1052) and CCM3 (SwissProt: Q9BUL8, residues, 98 to 212). CCM3 conserved HP1 residues are indicted with a + and helices for each structure are shaded grey. Colored as per CLUSTALW; red, AVFPMILW; blue, DE; purple, RK; green, STYHCNGQ. **B)** Sequence alignment of CCM3 from 34 species. Database entries and abbreviations as follows: HUMAN, Q9BUL8; CI, Ciona-intestinalis, XP_00212923; SAE, Starlet sea anemomae, XP_001626938; PF, Pufferfish, CAG10329; SABLE, Sablefish, ACQ58009; SMELT, Rainbow smelt, ACO09496; PIKE, Northern pike,ACO13672l; AS, Atlantic salmon, ACM08980; ZF, Zebrafish, Q6PHH3; XL, Xenopus laevis, Q8AVR4; XT, Xenopus tropicalis, Q6DF07; BF, Bullfrog, ACO51593; DOG, XP_545269; Cow, Q0VCQ6; MOUSE, Q8VE70; PIG NP_001132945; CHICK, Chicken, Q5ZIV5; RAT, Q6NX65; CHIMP, XP_001160181; RM, Rhesus monkey, XP_001099410; SP, Strongylocentrotus p, XP_797062; AG, Anopheles gambiae, XP_001689007; AA, Aedes aegypti, XP_001657782; BEE, Honey bee, XP_395623; WASP, Jewel wasp, XP_001602326; APHID, Pea aphid, XP_001949296; LOUSE, Human Body louse, XP_002432420; DP, Drosophila pseudo, XP_001358383; DG, Drosophila grimsh, XP_001989945; DW, Drosophila willis, XP_002069696; DE, Drosophila erecta, XP_001980057; DM, Drosophila melanogas, NP_650459; DA, Drosophila ananas, XP_001954505; BLT, Black legged tick, XP_002408528. Alignment conducted using CLUSTALW. "*", indicates identical; ":" highly conserved, "." semi-conserved. Colored as per CLUSTALW; red, AVFPMILW; blue, DE; purple, RK; green, STYHCNGQ. CCM3 dimerization residues indicated with a #, HP1 residues with a +.

A


Helix αF
Helix αG
HP1
CCM3 RPEPEFQDLNEKARALKQILSKIPDEIN--DRVRFLQTIKDIASAIKELLDTVNNVFKKYQ
PYK2 NLDRTDDLVYLNVMELVRAVLELKNELCQLPPEGYVVVVKNVGLTLRKLIGSVDDLLPSLP
FAK NLDRSNDKVYENVTGLVKAVIEMSSKIQPAPPEEYVPMVKEVGLALRTLLATVDETIPLLP
Helix αH
Helix αI
HP1
CCM3 YQNRRALEHQKKEFVKYSKSFSDTLKTYFKDG-------KAINVFVSANRLIHQTNLILQTFKTVA----
PYK2 SSSRTEIEGTQKLLNKDLAELINKMRLAQQNAVTSLSEECKRQMLTASHTLAVDAKNLLDAVDQAKVLAN
FAK ASTHREIEMAQKLLNSDLGELINKMKLAQQYVMTSLQQEYKKQMLTAAHALAVDAKNLLDVIDQARLKML

B


CI
SAE
PF
SABLE
SMELT
PIKE
AS
ZF
XL
XT
BF
DOG
COW
MOUSE
PIG
CHICK
RAT
CHIMP
RM
SP
AG
AA
BEE
WASP
APHID
LOUSE
DP
DG
DW
DE
DM
DA
BLT
HUMAN
CONS
DIMER
HP1