Supporting Information

Bourhis et al. 10.1073/pnas.1300480110

SI Materials and Methods

Molecular Biology. Mutations were introduced using the QuikChange XL site-directed mutagenesis kit from Stratagene. For expression of human procollagen III C-propeptide (CPIII)-His and CPIII-Long using the baculovirus system, DNA sequences were amplified by PCR from the N306Q miniprocollagen III construct (i.e., mutated at the single *N*-glycosylation site) in pCEP4 (1). The resulting fragments were inserted into the SacII/XhoI sites of the pBAC3 (Novagen) vector, in frame with an N-terminal His_6-tag. For expression of CUB1CUB2 (CUB, Complement C1r/C1s, Uegf, BMP-1) in insect cells, first the T6A mutation (numbered from the start of the mature protein) was introduced into full-length procollagen C-proteinase enhancer protein (PCPE)-1 in pBlueScriptSK$^+$ (2) to eliminate the *N*-glycosylation site. After amplification, the region encoding CUB1CUB2 was inserted into the BamHI/XhoI sites of the pBAC1 (Novagen) vector, including a C-terminal His_6-tag. Recombinant baculoviruses were obtained using the BacMagic-2 kit (Novagen) according to the manufacturer's protocol. CPIII-Long, CUB1CUB2, and their mutated forms were also produced by transient transfection in 293T cells using the pHLsec plasmid (3). For this, the CPIII-Long encoding region was amplified from the CPIII-Long-pBAC3 plasmid and inserted into the AgeI/XhoI sites of pHLsec. Similarly, the CUB1CUB2 encoding region was amplified from the CUB1CUB2-pBAC1 plasmid and inserted into the Age1/Kpn1 sites of pHLsec. Mutations in CPIII-Long and CUB1CUB2 were generated from these pHLsec constructs, including a subcloning step in pJet 1.2. For production of inactive bone morphogenetic protein-1 (BMP-1), the E94A mutation was introduced into the pCEP4-BMP1 plasmid (2). All primers and amino acid sequences are shown in Fig. S7.

Protein Production and Purification. For expression of CPIII-His, CPIII-Long, and CUB1CUB2 using the baculovirus system, *Trichoplusia ni* (High Five BTI-TN-5B1-4; Invitrogen) insect cells were cultured as previously described (1). When the cell density reached 1×10^6 cells/mL, cells were infected with recombinant baculovirus using multiplicities of infection of 1 for CPIII-His and CPIII-Long or 5 for CUB1CUB2. After 3 d (for CPIII-His and CPIII-Long) or 4 d (for CUB1CUB2) of infection, the culture medium was collected and the pH adjusted to 6.5. Subsequent purification was as described for CPIII-His (1). Production of CPIII-Long and its mutants by transient transfection in HEK 293T cells, and subsequent purification, are as described for CPIII-His produced in the same expression system (4). Production of CUB1CUB2 and its D191A mutant in HEK 293T cells was similar, except that a Superdex 75 16/60 column was used instead of a Superdex 200 16/60 column for the gel filtration step. All protein concentrations were determined using a Nanodrop spectrophotometer, with extinction coefficients calculated from amino acid sequences.

Characterization, Interactions, and Kinetics. Circular dichroism measurements were carried out as previously described (5). To investigate possible conformational changes on mixing, proteins were analyzed separately then mixed in a 1:1 molar ratio and the measured mean residue weight molar ellipticities for the mixtures compared with those calculated from the sum of the individual spectra weighted according to molecular mass.

To study the CPIII/PCPE-1 complex by gel filtration, 250-μL aliquots of PCPE-1 and CPIII (each at 833 nM) were first analyzed separately on a Superdex 200 HR10/30 column (GE Healthcare), at room temperature, in 20 mM Hepes (pH 7.4), 0.15 M NaCl, and 5 mM $CaCl_2$. CPIII was then mixed with an excess of PCPE-1 and, after incubation for 30 min at room temperature, analyzed on the same column in the same conditions. Fractions were analyzed by SDS/PAGE.

Surface plasmon resonance experiments were performed using a Biacore T100 (GE Healthcare) at the Protein Production and Analysis facility of the UMS3444 (Lyon). For interaction studies using WT and mutated forms of both CPIII-Long and the CUB1CUB2 fragment of PCPE-1, these proteins were expressed by transient transfection in 293T cells. Immobilization of PCPE-1 or CPIII-Long, regeneration of sensor chips, and analysis of kinetics were as previously described (6, 7). Sensorgrams were recorded at 25 °C using 10 mM Hepes (pH 7.4), 0.15 M NaCl, 5 mM $CaCl_2$, and 0.05% P-20 as running buffer.

For Biacore studies on the interaction of the inactive E94A mutant of BMP-1, miniprocollagen III (840 RU) was first immobilized on a CM5 sensorchip (series S, noncertified). To measure interactions in the absence of CUB1CUB2, increasing concentrations of BMP-1 E94A (9.4, 18.8, 37.5, 75, and 300 nM) were injected, for 120 s, at a flow rate of 30 μL/min. For interactions in the presence of CUB1CUB2 (produced by limited proteolysis of full-length PCPE-1), the same surface was first saturated with an excess of CUB1CUB2 (50 nM), injected for 120 s using the same running buffer, followed by coinjection (for a further 120 s, always at 50 nM CUB1CUB2) using the same range of BMP-1 E94A concentrations as used earlier. Between injections, the sensorchip was regenerated using 0.5 M EDTA (pH 8.0) (30 s) followed by 2 M GuHCl (30 s). In all cases the dissociation step was done using running buffer. Model fitting was carried out using BiaEvaluation software version 4.1. Before fitting, for the coinject experiments, the sensorgram for CUB1-CUB2 alone was subtracted from the overall data.

Best fits to the Biacore data were obtained using the heterogeneous ligand model. This often arises as a consequence of immobilization of the ligand in distinct orientations, by amine-coupling chemistry, such that strong binding sites are inaccessible in one subpopulation, leading to an apparent second binding site with a higher K_D. Such heterogeneity is a result of the immobilization process and can arise no matter which partner is immobilized, as observed in the experiments reported here. In addition, heterogeneous binding sites could also be an intrinsic property of one or both partners. In this regard, the intrinsic asymmetry observed in the procollagen III C-propeptide trimer (8) could be a contributing factor.

For enzyme kinetics, BMP-1 activity, in the presence and absence of PCPE-1, was measured at 37 °C in a reaction volume of 120 μL, in 50 mM Hepes (pH 7.4), 0.15 M NaCl, 5 mM $CaCl_2$, and 0.02% Brij-35, using CPIII-Long (produced in the baculovirus system), for a range of substrate concentrations from 0 to 3 μM. When PCPE-1 was present, it was used at a fixed concentration equal to the highest substrate concentration used (9). In each case (i.e., with and without PCPE-1), the BMP-1 concentration was adjusted to remain within the linear region of the plot of amount of product released as a function of time. Reactions were started by addition of enzyme (57 nM in the absence or 6 nM in the presence of PCPE-1) and stopped by the addition of EDTA (25 mM final concentration) and unstained Laemmli sample buffer. Samples (20 μL) were taken at fixed time intervals throughout the reaction. These were then analyzed by SDS/PAGE on 4–20% Criterion gradient gels (Bio-Rad), in reducing conditions, followed by Sypro Ruby staining and

quantitation using a Molecular Dynamics Storm 860 imager (blue laser) and ImageQuant software (5). Sample loading on the gels was adjusted to approximately 500 ng per lane to remain within the dynamic range of Sypro Ruby. Kinetic constants k_{cat} and K_m were determined by fitting to the experimental data using Prism 5.04 (GraphPad Software).

Structural Analysis. For the small-angle X-ray scattering (SAXS) studies, all proteins were expressed using the baculovirus system. As a final step in purification, individual proteins (CPIII-His, CPIII-Long, CUB1CUB2) were isolated by gel filtration on a Superdex 200 HR16/60 prep grade column (GE Healthcare), equilibrated in 20 mM Hepes (pH 7.4), 150 mM NaCl, 5 mM $CaCl_2$, and 5% (vol/vol) glycerol, followed by concentration using Vivaspin 500 (Sartorius) 10K molecular weight cut-off polyethersulfone (MWCO PES) centrifugal ultrafiltration devices. Complexes (CPIII-His:CUB1CUB2 and CPIII-Long:CUB1CUB2) were prepared by mixing purified proteins using an excess of CUB1CUB2, followed by incubation for at least 1 h at 4 °C and then further concentration as above. Samples were analyzed on the SWING beamline at SOLEIL which includes a further in-line gel filtration column for complex isolation and optimal sample homogeneity. Data were collected in the Q range 0.01–0.50 $Å^{-1}$, where $Q = 4\pi\sin(\theta)/\lambda$, 2θ is the scattering angle, and λ =1.03 Å is the X-ray wavelength. Scattering patterns were recorded every 1.5 s. After data reduction and subtraction of buffer scattering, sample homogeneity within each peak was first checked by Guinier analysis (R_g constant), then data sets were pooled. Particle mass was calculated from the scattering intensity extrapolated to zero angle, using water as a reference (10) and a partial specific volume (11) of 0.7425 $cm^3\ g^{-1}$. Distance distribution functions p(r) were calculated using the program GNOM (12). The models generated by MONSA for the individual molecules were very similar to those generated by DAMMIN. In Fig. 1, the resolution of CPIII-His and CPIII-Long in *B* is less than in *A*, because overall resolution is limited by that for CUB1CUB2. The structure of the CUB1CUB2 fragment of PCPE-1 was modeled as previously described (7). For the molecular superpositions, the crystal structure of CPIII-His (Protein Data Bank code 4AK3) was built manually into the averaged SAXS envelopes, as well as the CUB domain models for the MONSA envelopes, and then optimized by rigid body refinement using CHIMERA (13). The quality of the fits was good, as assessed by the correlation coefficients (0.9124 for CPIII and 0.9022 for CUB1CUB2, where the maximum is 1). Situs (14) was used to generate maps from the bead-model shapes. Figures were drawn using PyMOL, version 1.4.1 (Schrödinger, LLC). Sequence alignments were obtained with MUSCLE (15) and rendered using ESPript (16).

1. Vadon-Le Goff S, et al. (2011) Procollagen C-proteinase enhancer stimulates procollagen processing by binding to the C-propeptide region only. *J Biol Chem* 286(45):38932–38938.
2. Moali C, et al. (2005) Substrate-specific modulation of a multisubstrate proteinase. C-terminal processing of fibrillar procollagens is the only BMP-1-dependent activity to be enhanced by PCPE-1. *J Biol Chem* 280(25):24188–24194.
3. Aricescu AR, Lu W, Jones EY (2006) A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr D Biol Crystallogr* 62(Pt 10): 1243–1250.
4. Bourhis JM, et al. (2012) Production and crystallization of the C-propeptide trimer from human procollagen III. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 68(Pt 10):1209–1213.
5. Bekhouche M, et al. (2010) Role of the netrin-like domain of procollagen C-proteinase enhancer-1 in the control of metalloproteinase activity. *J Biol Chem* 285(21): 15950–15959.
6. Blanc G, et al. (2007) Insights into how CUB domains can exert specific functions while sharing a common fold: conserved and specific features of the CUB1 domain contribute to the molecular basis of procollagen C-proteinase enhancer-1 activity. *J Biol Chem* 282(23):16924–16933.
7. Kronenberg D, et al. (2009) Strong cooperativity and loose geometry between CUB domains are the basis for procollagen c-proteinase enhancer activity. *J Biol Chem* 284(48):33437–33446.
8. Bourhis JM, et al. (2012) Structural basis of fibrillar collagen trimerization and related genetic disorders. *Nat Struct Mol Biol* 19(10):1031–1036.
9. Cornish-Bowden A (2004) *Fundamentals of Enzyme Kinetics* (Portland Press, London).
10. Orthaber D, Bergmann A, Glatter O (2000) SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J Appl Cryst* 33:218–225.
11. Mylonas E, Svergun DI (2007) Accuracy of molecular mass determination of proteins in solution by small angle X-ray scattering. *J Appl Cryst* 40:s245–s249.
12. Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Cryst* 25:495–503.
13. Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
14. Wriggers W (2010) Using Situs for the integration of multi-resolution structures. *Biophys Rev* 2(1):21–27.
15. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797.
16. Gouet P, Courcelle E, Stuart DI, Métoz F (1999) ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15(4):305–308.

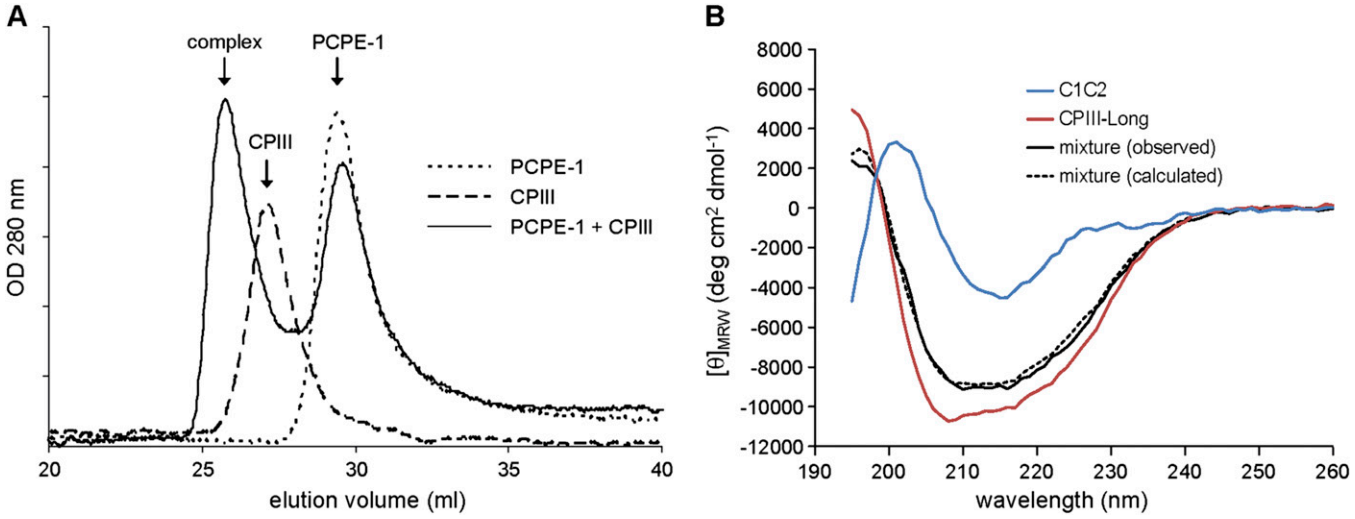

Fig. S1. Characterization of complexes. (*A*) Gel filtration chromatography of isolated PCPE-1 and CPIII as well as a mixture of both proteins (containing an excess of PCPE-1). The PCPE-1:CPIII complex is larger than CPIII and could be separated from excess free PCPE-1. (*B*) Absence of change in secondary structure on mixing CPIII-Long and CUB1CUB2. Circular dichroism spectra were first recorded for CPIII-Long, CUB1CUB2, and a 1:1 molar mixture (all at 200 μg/mL total protein concentration) at 25 °C in 20 mM Tris (pH 7.4), 0.15 M NaCl, and 5 mM $CaCl_2$, then mean residue weight molar ellipticity was calculated for the mixture and compared with that observed.

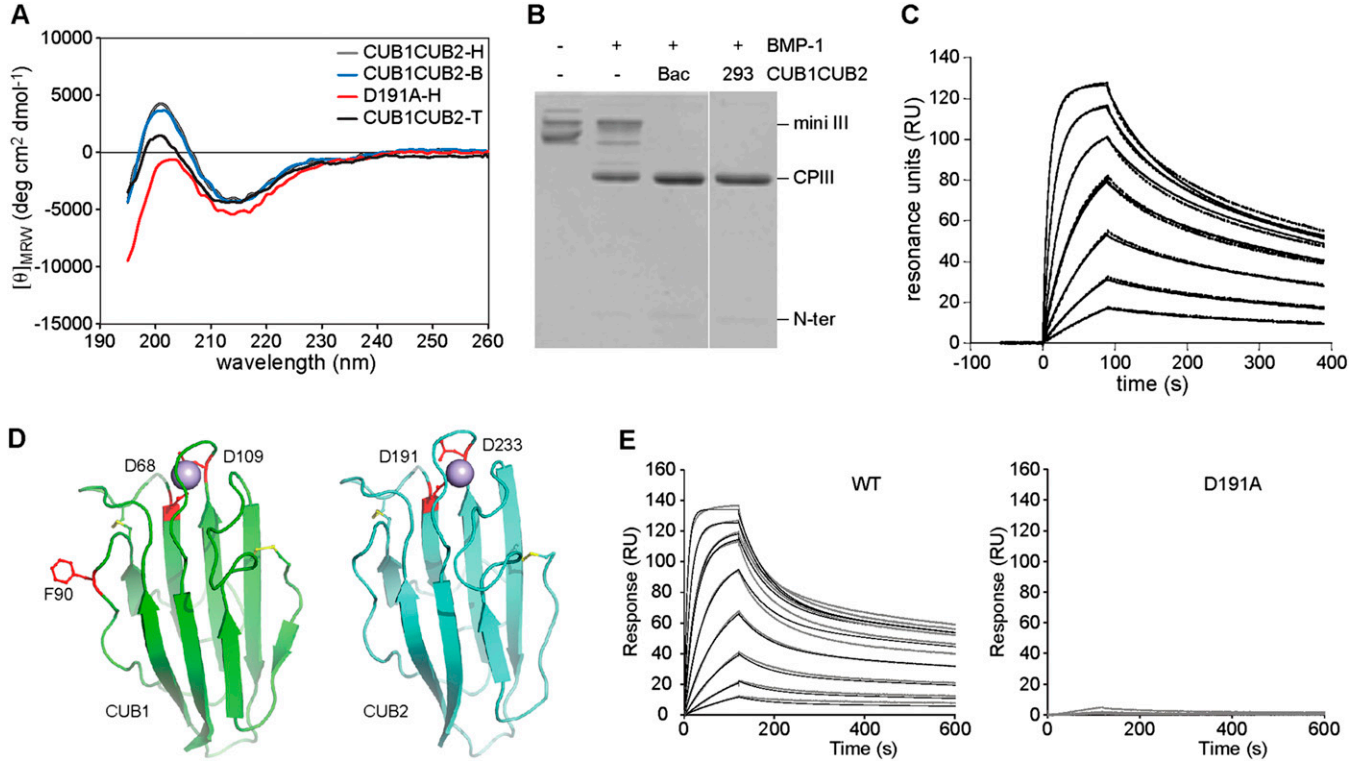

Fig. S2. Characterization of the CUB1CUB2 fragment of PCPE-1. (*A*) Circular dichroism spectra of CUB1CUB2 obtained by limited trypsin digestion of PCPE-1 (CUB1CUB2-T; 236 μg/mL), CUB1CUB2 produced in insect cells (CUB1CUB2-B; 166 μg/mL), and CUB1CUB2 (236 μg/mL) and its D191A mutant (142 μg/mL) produced in HEK 293T cells (CUB1CUB2-H and D191A-H, respectively) analyzed at 25 °C in 10 mM Tris·HCl (pH 7.4), 150 mM NaCl, and 5 mM $CaCl_2$. Circular dichroism spectra for CUB1CUB2-B for CUB1CUB2-H are indistinguishable. Small differences between CUB1CUB2-B and CUB1CUB2-T could be due to the His_6-tag in CUB1CUB2-B and both the *N*-glycosylation site and part of the CUB2-NTR linker in CUB1CUB2-T. (*B*) CUB1CUB2-B is as active as CUB1CUB2-T. Equimolar (350 nM) concentrations of miniprocollagen III and the two forms of CUB1CUB2 were incubated in the presence of 17 nM BMP-1 for 2 h at 37 °C; SDS/PAGE (10% acrylamide), reducing conditions, Coomassie blue staining. (*C*) CUB1CUB2-B binds to immobilized miniprocollagen III (840 RU) with K_Ds of 0.54 and 17 nM (heterogeneous ligand model), similar to those for CUB1CUB2-T (0.54 and 9.1 nM) reported elsewhere (1). Increasing concentrations of CUB1CUB2-B (0.78–50 nM, 6.25 nM concentration injected twice) were injected on immobilized miniprocollagen III at 30 μL/min. Best fits are also shown (black lines; $\chi^2 = 1.36$). (*D*) Cartoon representations of 3D structural models of the CUB1 and CUB2 domains, showing Ca^{2+} ions as light blue spheres, and residues in CUB1 shown previously to be required for CPIII binding (2), as well as probable Ca^{2+} coordinating residues in CUB2, all shown in red. (*E*) Surface plasmon resonance (Biacore) analysis of the binding of CUB1CUB2 (WT or its D191A mutant; 0–256 nM) to immobilized (595 RU) CPIII-Long. Best fits are also shown for WT (black lines), obtained using the heterogeneous ligand model, giving dissociation constants (K_D) of 0.34 and 7.5 nM ($\chi^2 = 12$).

1. Vadon-Le Goff S, et al. (2011) Procollagen C-proteinase enhancer stimulates procollagen processing by binding to the C-propeptide region only. *J Biol Chem* 286(45):38932–38938.
2. Blanc G, et al. (2007) Insights into how CUB domains can exert specific functions while sharing a common fold: Conserved and specific features of the CUB1 domain contribute to the molecular basis of procollagen C-proteinase enhancer-1 activity. *J Biol Chem* 282(23):16924–16933.

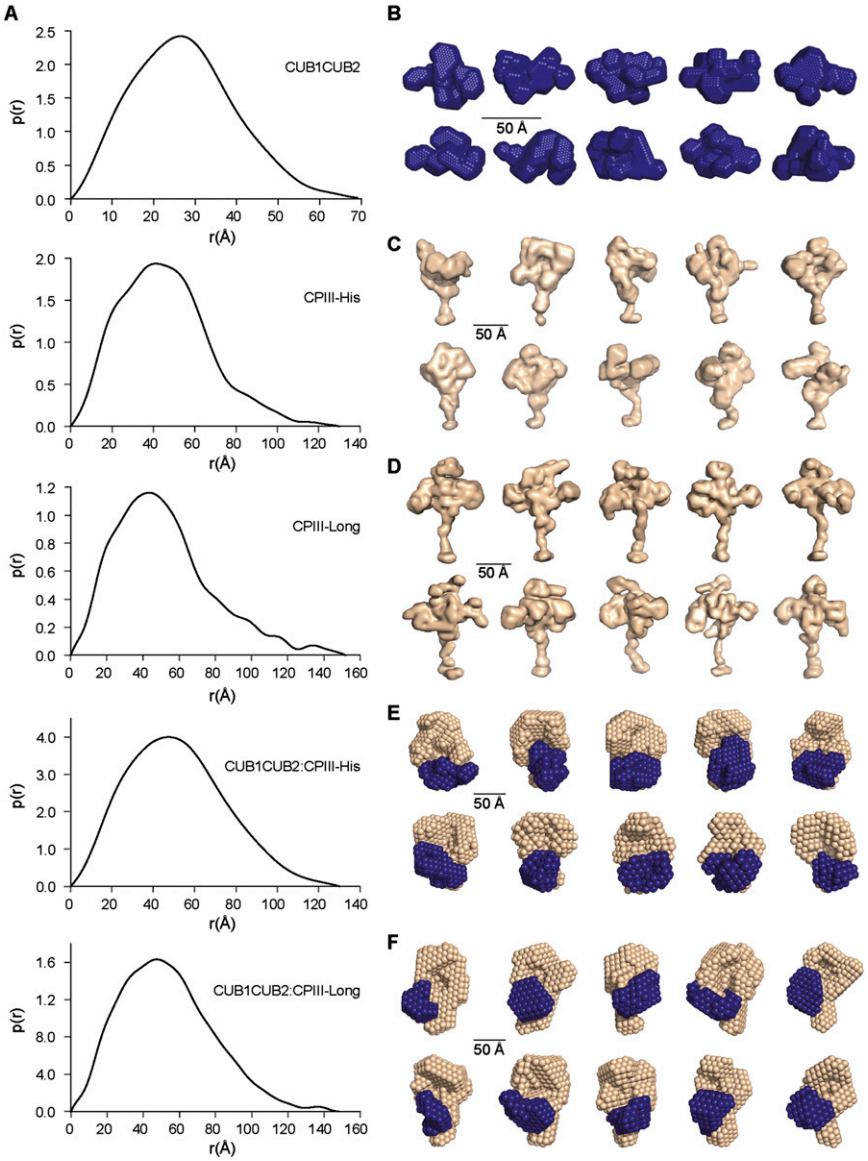

Fig. S3. Distance distribution functions and ab initio shapes calculated from the SAXS data. (*A*) Distribution of intraparticle distances p(r) and maximum dimensions D_{max} (indicated by the right hand intercept with the horizontal axis) for CUB1CUB2, CPIII-His, CPIII-Long, the CUB1CUB2:CPIII-His complex, and the CUB1CUB2:CPIII-Long complex. (*B–D*) Gallery of low-resolution structures determined by DAMMIF for (*B*) CUB1CUB2, (*C*) CPIII-His, and (*D*) CPIII-Long. (*E* and *F*) Gallery of low-resolution structures determined by MONSA for (*E*) the CUB1CUB2:CPIII-His complex and (*F*) the CUB1CUB2:CPIII-Long complex. Each molecule is represented as a collection of "dummy atoms," with CUB1CUB2 shown in blue and CPIII-His and CPIII-Long shown in wheat. Note that CUB1CUB2-Bac and CPIII-His are found to be aligned at the base of the stalk region, whereas CPIII-Long projects beyond the CUB1CUB2 binding site at the base of the stalk region.

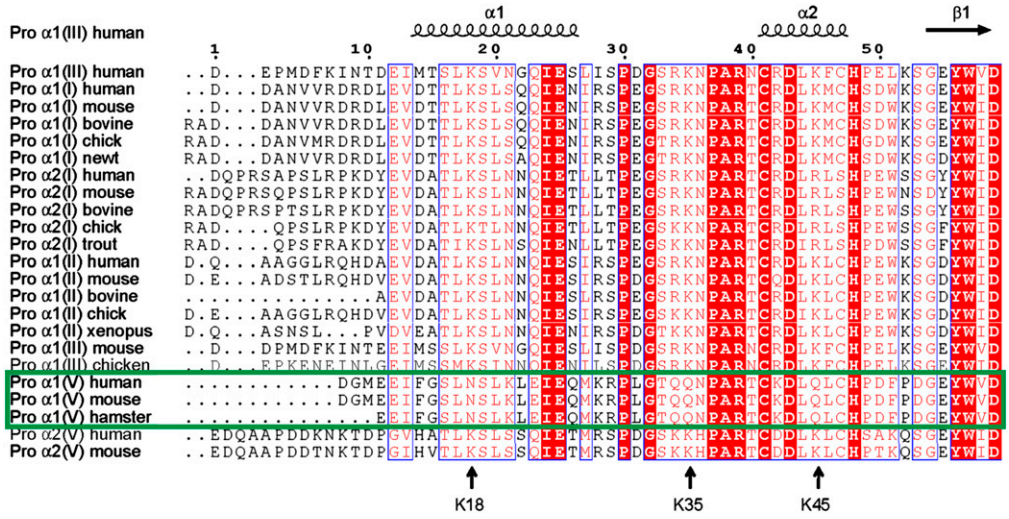

Fig. S4. Amino acid sequence alignment in the stalk/base regions of selected fibrillar procollagen C-propeptides (types I, II, III, and V). Highly conserved regions are shown in red. Secondary structures for the known structure of the C-propeptide of human procollagen III (1) are shown at the top. Locations of the highly conserved lysine residues at positions 18, 35, and 45 are indicated. Note that these positions are not occupied by lysines in the proα1(V) chain of procollagen V (boxed in green).

1. Bourhis JM, et al. (2012) Structural basis of fibrillar collagen trimerization and related genetic disorders. *Nat Struct Mol Biol* 19(10):1031–1036.

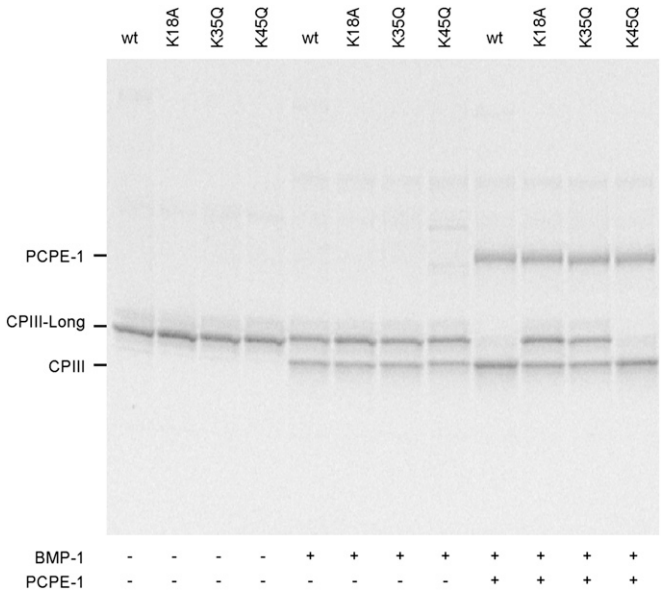

Fig. S5. Representative SDS/PAGE data corresponding to the results shown in Fig. 2*A*. CPIII-Long (340 nM; both WT and the K18A, K35Q, and K45Q mutants) was incubated either alone, with BMP-1 (16 nM) or with BMP-1 (16 nM) and PCPE-1 (340 nM) for 1 h at 37 °C. Analysis by SDS/PAGE (reducing conditions) and Sypro Ruby staining.

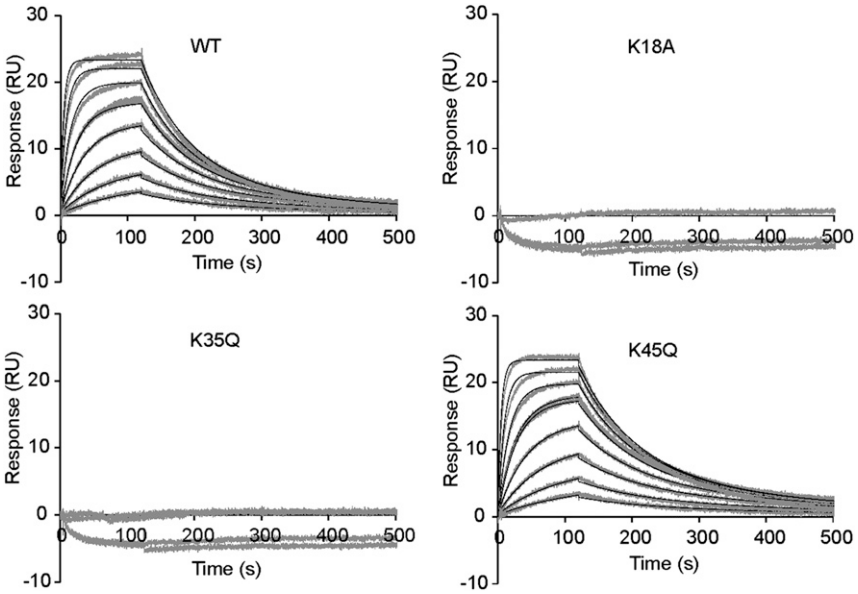

Fig. S6. Surface plasmon resonance (Biacore) analysis for CPIII-Long and its site-directed mutants. Binding of CPIII-Long (0-256 nM twofold serial dilutions for both WT and the K45Q mutant; 0–1 μM for the K18A and K35Q mutants) to immobilized (74 RU) PCPE-1. Best fits are also shown (black lines), obtained using the heterogeneous ligand model, giving dissociation constants (K_D) of 0.77 and 10 nM for WT and 0.76 and 8.5 nM for K45Q ($\chi^2 = 0.13$ and 0.20, respectively).

CUB1CUB2-His-pBAC1

MLPAATASLLGPLLTACALLPFAQGQTPNYARPVFLCGGDVKGESGYVASEGFPNLYPPNKECIWTITVPEGQTVSLSFRVFDLELHPACRYDALEVFAGSGTSGQR
LGRFCGTFRPAPLVAPGNQVTLRMTTDEGTGGRGFLLWYSGRATSGTEHQFCGGRLEKAQGTLTTPNWPESDYPPGISCSWHIIAPPDQVIALTFEKFDLEPDTYCR
YDSVSVFNGAVSDDSRRLGKFCGDAVPGSISSEGNELLVQFVSDLSVTADGFSASYKTLPRGTLEHHHHHH
For: GGATCCATGCTGCCTGCAGCCACA
Rev: CTCGAGTTAAGTGCCCCGCGGCAG

CUB1CUB2-His-pHLsec

MGILPSPGMPALLSLVSLLSVLLMGCVAETGQTPNYARPVFLCGGDVKGESGYVASEGFPNLYPPNKECIWTITVPEGQTVSLSFRVFDLELHPACRYDALEVFAGS
GTSGQRLGRFCGTFRPAPLVAPGNQVTLRMTTDEGTGGRGFLLWYSGRATSGTEHQFCGGRLEKAQGTLTTPNWPESDYPPGISCSWHIIAPPDQVIALTFEKFDLE
PDTYCRYDSVSVFNGAVSDDSRRLGKFCGDAVPGSISSEGNELLVQFVSDLSVTADGFSASYKTLPRGTLGTKHHHHHH
For: GTCAACCGGTCAGACCCCCAACTAC
Rev: GATTGGTACCGAGAGTGCCCCG

CUB1CUB2-His-D191A-pHLsec

MGILPSPGMPALLSLVSLLSVLLMGCVAETGQTPNYARPVFLCGGDVKGESGYVASEGFPNLYPPNKECIWTITVPEGQTVSLSFRVFDLELHPACRYDALEVFAGS
GTSGQRLGRFCGTFRPAPLVAPGNQVTLRMTTDEGTGGRGFLLWYSGRATSGTEHQFCGGRLEKAQGTLTTPNWPESDYPPGISCSWHIIAPPDQVIALTFEKFDLE
PDTYCRYASVSVFNGAVSDDSRRLGKFCGDAVPGSISSEGNELLVQFVSDLSVTADGFSASYKTLPRGTLGTKHHHHHH
For: CTACTGCCGCTATGCCTCGGTCAGCGTGTTC
Rev: GAACACGCTGACCGAGGCATAGCGGCAGTAG

CUB1CUB2-His-D233A-pHLsec

MGILPSPGMPALLSLVSLLSVLLMGCVAETGQTPNYARPVFLCGGDVKGESGYVASEGFPNLYPPNKECIWTITVPEGQTVSLSFRVFDLELHPACRYDALEVFAGS
GTSGQRLGRFCGTFRPAPLVAPGNQVTLRMTTDEGTGGRGFLLWYSGRATSGTEHQFCGGRLEKAQGTLTTPNWPESDYPPGISCSWHIIAPPDQVIALTFEKFDLE
PDTYCRYDSVSVFNGAVSDDSRRLGKFCGDAVPGSISSEGNELLVQFVSALSVTADGFSASYKTLPRGTLGTKHHHHHH
For: CAGTTCGTCTCAGCTCTCAGTGTCAC
Rev: GTGACACTGAGAGCTGAGACGAACTG

CPIII-His-pBAC3

MPMLSAIVLYVLLAAAAHSAFAAMVHHHHHHSADEPMDFKINTDEIMTSLKSVNGQIESLISPDGSRKNPARNCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCN
METGETCISANPLNVPRKHWWTDSSAEKKHVWFGESMDGGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQQITYHCKNSIAYMDQASGNVKKALKLMGSNEGEFKA
EGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL
For: CCCGCGGATGAACCAATGGATTTC
Rev: TCTCGAGTTATAAAAAGCAAACAGGG

CPIII-Long-pBAC3

MPMLSAIVLYVLLAAAAHSAFAAMVHHHHHHSAGPPGPPGAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPDGSRKNP
ARNCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGESMDGGFQFSYGNPELPEDVLDVQLAFLRLLSSRAS
QQITYHCKNSIAYMDQASGNVKKALKLMGSNEGEFKAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL
For: CCGCGGGCCCTCCTGGACCTCC
Rev: TCTCGAGTTATAAAAAGCAAACAGGG

CPIII-Long-pHLsec

MGILPSPGMPALLSLVSLLSVLLMGCVAE**TG**HHHHHHSAGPPGPPGAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPD
GSRKNPARNCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGESMDGGFQFSYGNPELPEDVLDVQLAFLRL
LSSRASQQITYHCKNSIAYMDQASGNVKKALKLMGSNEGEFKAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL
For: ACCGGTCATCATCACCACCATCACTCC
Rev: TCTAGACTCGAGTTATAAAAAGCAAACAGG

CPIII-Long-K18A-pHLsec

MGILPSPGMPALLSLVSLLSVLLMGCVAE**TG**HHHHHHSAGPPGPPGAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSL**A**SVNGQIESLISPD
GSRKNPARNCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGESMDGGFQFSYGNPELPEDVLDVQLAFLRL
LSSRASQQITYHCKNSIAYMDQASGNVKKALKLMGSNEGEFKAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL
For: GACTTCACTCGCGTCTGTTAATGG
Rev: CCATTAACAGACGCGAGTGAAGTC

CPIII-Long-K35Q-pHLsec

MGILPSPGMPALLSLVSLLSVLLMGCVAE**TG**HHHHHHSAGPPGPPGAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPD
GSR**Q**NPARNCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGESMDGGFQFSYGNPELPEDVLDVQLAFLRL
LSSRASQQITYHCKNSIAYMDQASGNVKKALKLMGSNEGEFKAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL
For: GATGGTTCTCGTCAAAACCCCGCTAG
Rev: CTAGCGGGGTTTTGACGAGAACCATC

CPIII-Long-K45Q-pHLsec

MGILPSPGMPALLSLVSLLSVLLMGCVAE**TG**HHHHHHSAGPPGPPGAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPD
GSRKNPARNCRDL**Q**FCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGESMDGGFQFSYGNPELPEDVLDVQLAFLRL
LSSRASQQITYHCKNSIAYMDQASGNVKKALKLMGSNEGEFKAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL
For: CTGCAGAGACCTGCAATTCTGCCATCCTG
Rev: CAGGATGGCAGAATTGCAGGTCTCTGCAG

BMP1-E94A-pCEP4

MPGVARLPLLLGLLLLPRPGRPLDLADYTYDLAEEDDSEPLNYKDPCKAAAFLGDIALDEEDLRAFQVQQAVDLRRHTARKSSIKAAVPGNTSTPSCQSTNGQPQRG
ACGRWRGRSRSRRAATSRPERVWPDGVIPFVIGGNFTGSQRAVFRQAMRHWEKHTCVTFLERTDEDSYIVFTYRPCGCCSYVGRRGGGPQAISIGKNCDKFGIVVH**E**
LGHVVGFWHEHTRPDRDRHVSIVRENIQPGQEYNFLKMEPQEVESLGETYDFDSIMHYARNTFSRGIFLDTIVPKYEVNGVKPPIGQRTRLSKGDIAQARKLYKCPA
CGETLQDSTGNFSSPEYPNGYSAHMHCVWRISVTPGEKIILNFTSLDLYRSRLCWYDYVEVRDGFWRKAPLRGRFCGSKLPEPIVSTDSRLWVEFRSSSNWVGKGFF
AVYEAICGGDVKKDYGHIQSPNYPDDYRPSKVCIWRIQVSEGFHVGLTFQSFEIERHDSCAYDYLEVRDGHSESSTLIGRYCGYEKPDDIKSTSSRLWLKFVSDGSI
NKAGFAVNFFKEVDECSRPNRGGCEQRCLNTLGSYKCSCDPGYELAPDKRRCEAACGGFLTKLNGSITSPGWPKEYPPNKNCIWQLVAPTQYRISLQFDFFETEGND
VCKYDFVEVRSGLTADSKLHGKFCGSEKPEVITSQYNNMRVEFKSDNTVSKKGFKAHFFSEKRPALQPPRGRPHQLKFRVQKRNRTPQDYKDDDDK
For: CGGCATTGTGGTCCACGCGCTGGGCCACGTCGTC
Rev: GACGACGTGGCCCAGCGCGTGGACCACAATGCCG

Fig. S7. Protein sequences and primers used to generate the different constructs. Site-directed mutations at *N*-glycosylation sites are underlined. Other site-directed mutations are highlighted in blue, where the numbering begins at the N terminus of the mature protein (highlighted in green). For CPIII and it variants, the latter corresponds to the BMP-1 cleavage site.

Table S1. Analysis of circular dichroism spectra for WT and mutated CUB1CUB2

Protein	Regular α-helix	Distorted α-helix	Regular β-strand	Distorted β-strand	β-turns	Unordered	Total
WT	0.00	0.02	0.23	0.13	0.23	0.38	0.99
D191A	−0.01	0.01	0.22	0.12	0.22	0.42	0.98

Spectra from Fig. S2*A* were deconvoluted using the DICHROWEB server (1) using the program CDSSTR with reference set 7.

1. Whitmore L, Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* 89(5):392–400.