

Procollagen C-proteinase enhancer grasps the stalk of the C-propeptide trimer to boost collagen precursor maturation

Jean-Marie Bourhis^{a,1,2}, Sandrine Vadon-Le Goff^{a,1}, Hassnae Afrache^{a,3}, Natacha Mariano^a, Daniel Kronenberg^{a,4}, Nicole Thielens^b, Catherine Moali^a, and David J. S. Hulmes^{a,5}

^aUnité Mixte de Recherche 5305, Institut de Biologie et Chimie des Protéines, Centre National de la Recherche Scientifique, Université Lyon 1, 69367 Lyon cedex 7, France; and ^bInstitut de Biologie Structurale Jean-Pierre Ebel, Unité Mixte de Recherche 5075, Commissariat à l'Énergie Atomique et aux Énergies Alternatives, Centre National de la Recherche Scientifique, Université Joseph Fourier, 38027 Grenoble cedex 1, France

Edited* by Darwin J. Prockop, Texas A&M Health Science Center, Temple, TX, and approved March 8, 2013 (received for review January 9, 2013)

Tight regulation of collagen fibril deposition in the extracellular matrix is essential for normal tissue homeostasis and repair, defects in which are associated with several degenerative or fibrotic disorders. A key regulatory step in collagen fibril assembly is the C-terminal proteolytic processing of soluble procollagen precursors. This step, carried out mainly by bone morphogenetic protein-1/tolloid-like proteinases, is itself subject to regulation by procollagen C-proteinase enhancer proteins (PCPEs) which can dramatically increase bone morphogenetic protein-1/tolloid-like proteinase activity, in a substrate-specific manner. Although it is known that this enhancing activity requires binding of PCPE to the procollagen C-propeptide trimer, identification of the precise binding site has so far remained elusive. Here, use of small-angle X-ray scattering provides structural data on this protein complex indicating that PCPE binds to the stalk region of the procollagen C-propeptide trimer, where the three polypeptide chains associate together, at the junction with the base region. This is supported by site-directed mutagenesis, which identifies two highly conserved, surface-exposed lysine residues in this region of the trimer that are essential for binding, thus revealing structural parallels with the interactions of Complement C1r/C1s, Uegf, BMP-1 (CUB) domain-containing proteins in diverse biological systems such as complement activation, receptor signaling, and transport. Together with detailed kinetics and interaction analysis, these results provide insights into the mechanism of action of PCPEs and suggest clear strategies for the development of novel antifibrotic therapies.

proteolysis | structural proteins | fibrosis

Regulation of collagen deposition in the extracellular matrix is essential for normal tissue homeostasis and repair, defects in which are associated with numerous, often lethal, disorders, including chronic wounds and the many different forms of fibrosis (affecting heart, lung, liver, kidney, skin, etc.) (1, 2). At the protein level, a key regulatory step is the proteolytic conversion of soluble precursor molecules, procollagens, by removal of N- and C-propeptides (3), resulting in spontaneous assembly of collagen fibrils. Although collagen deposition *in vivo* is also controlled by numerous interactions with cell-surface and extracellular matrix proteins (4, 5), it is generally acknowledged that a key rate-limiting step is the proteolytic removal of the procollagen C-propeptides. The principal proteinases involved here are zinc metalloproteinases, called bone morphogenetic protein-1 (BMP-1)/tolloid-like proteinases (collectively known as BTPs), which cleave the C-propeptides from the major fibrillar procollagens (types I, II, and III) (3, 6–8).

Although BTPs cleave a number of extracellular substrates (including structural proproteins, proenzymes, and latent growth factors or their antagonists) (8, 9), their activity on fibrillar procollagens is specifically increased, several-fold, by procollagen C-proteinase enhancers (PCPE-1 and -2) (10–13). There is increasing interest in PCPEs as possible therapeutic targets. PCPE-1,

for example, has been shown to be up-regulated in rat models of liver (14) and cardiac (15) fibrosis. More recently, cardiac fibrosis after chronic pressure overload was found to be diminished in mice lacking PCPE-2 (16). Although the precise mechanism of enhancement of proteinase activity remains to be elucidated, it is clear that this requires binding of PCPEs, or more precisely their CUB (Complement C1r/C1s, Uegf, BMP-1) domain region, to the procollagen C-propeptide trimer (17, 18). Interestingly, the stoichiometry of binding is one molecule of PCPE to one molecule of procollagen (19), suggesting that the PCPE binding site on the C-propeptide trimer involves more than one polypeptide chain.

We recently reported the high-resolution structure of the C-propeptide trimer from human procollagen III (CPIII) (20). This has the overall shape of a flower, consisting of a stalk (in the form of a triple-stranded coiled coil), a highly conserved base region (including three Ca²⁺ binding sites), and three petals (involved in chain recognition specificity). Here we report the low-resolution structure of the complex between CPIII and the CUB1CUB2 region of human PCPE-1, which indicates that the PCPE binding site encompasses the stalk region of CPIII up to the base. This is supported by the identification of two highly conserved lysine residues in the CPIII stalk and base that are essential for PCPE-1 binding and enhancing activity. Binding also requires conserved aspartate residues involved in Ca²⁺ coordination in both the CUB1 and CUB2 domains of PCPE-1. These interactions reveal structural parallels with those of other CUB domain-containing proteins involved in such diverse biological systems as complement activation, receptor signaling, and transport. Finally, we present quantitative data on the effects of PCPE-1 (or its CUB1CUB2 region) on both substrate binding and enzymatic activity of BMP-1, which together with the structural data give clues about the mechanism of this striking increase in catalytic efficiency of BTPs.

Author contributions: J.-M.B., S.V.-L.G., H.A., N.M., D.K., C.M., and D.J.S.H. designed research; J.-M.B., S.V.-L.G., H.A., N.M., D.K., C.M., and D.J.S.H. performed research; J.-M.B., S.V.-L.G., N.T., C.M., and D.J.S.H. analyzed data; and D.J.S.H. wrote the paper.

The authors declare a conflict of interest. This work forms part of a patent application by J.-M.B., S.V.-L.G., N.M., C.M. and D.J.S.H.

*This Direct Submission article had a prearranged editor.

¹J.-M.B. and S.V.-L.G. contributed equally to this work.

²Present address: Unit for Virus Host Cell Interactions, Unité Mixte Internationale 3265, Centre National de la Recherche Scientifique, Université Joseph Fourier, European Molecular Biology Laboratory, 38042 Grenoble cedex 9, France.

³Present address: Immunology and Cancer Laboratory, Centre de Recherche en Cancérologie de Marseille, Institut National de la Santé et de la Recherche Médicale U1068, 13273 Marseille cedex 9, France.

⁴Present address: Institute for Experimental Musculoskeletal Medicine, University of Münster, 48149 Münster, Germany.

⁵To whom correspondence should be addressed. E-mail: d.hulmes@ibcp.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1300480110/-DCSupplemental.

Results

Structure of the PCPE-1/CPIII Complex. We recently showed that the action of PCPE-1 relies solely on its binding to the C-propeptide region of the procollagen molecule (18). Because the corresponding dissociation constant (K_D) is in the nanomolar range, we reasoned that it should be possible to isolate the PCPE-1: C-propeptide trimer complex. As expected (Fig. S14), when CPIII was first incubated with an excess of PCPE-1 then analyzed by gel filtration, the complex migrated ahead of the elution position of CPIII and could be separated from free PCPE-1.

To determine the low-resolution structure of the complex, and hence localize the PCPE-1 binding site on the CPIII molecule, we used small-angle X-ray scattering (SAXS). Proteins analyzed (Fig. 1) included CPIII-His (18), which consists of the human procollagen III C-propeptide with an N-terminal His₆-tag, as well as a longer construct called CPIII-Long (which also includes the last three GXY triplets of the triple-helical region, as well as the entire C-telopeptide, the short nontriple helical region that remains after removal of the C-propeptide by BTPs). We also produced the CUB1CUB2 region of PCPE-1 mutated at the single *N*-glycosylation site in CUB1.

In preliminary experiments, we showed (Fig. S24) that CUB1-CUB2 produced in the baculovirus system had a similar circular dichroism (CD) spectrum to CUB1CUB2 obtained by trypsin digestion of full length PCPE-1 (nonmutated at the *N*-glycosylation site) expressed in human embryonic kidney 293 EBNA cells (17). In addition, we found both forms of CUB1CUB2 to be equally effective in enhancing BMP-1 activity on a miniprocollagen substrate, with a similar binding affinity as determined by Biacore (Fig. S2 B and C). These results show that *N*-glycosylation of the CUB1CUB2 region is not involved in PCPE-1 activity.

In the SAXS experiments, CUB1CUB2, CPIII-His, and CPIII-Long were first analyzed individually and then in the form of complexes (CUB1CUB2 with either CPIII-His or CPIII-Long). Immediately before SAXS analysis on the SWING beamline (21) at the SOLEIL synchrotron, all samples (proteins and complexes) were passed through an in-line gel filtration column. As shown in Table 1, by extrapolation of the scattering intensity to zero angle, all molecules (CUB1CUB2, CPIII-His, and CPIII-Long) behaved as monomers in solution, with the expected molecular masses. Table 1 also shows the observed values for the radius of gyration R_g , determined from Guinier plots, as well as maximum particle diameters D_{max} , determined from the distance distribution functions $p(r)$ (Fig. S34).

Shown in Fig. 1A are the averaged shapes determined for the different molecules from the SAXS data. For each molecule, first

10 individual ab initio shape determinations were made (without symmetry restrictions) using the program DAMMIF (22), and then averaged shapes were calculated using DAMAVER (23). For CPIII-His, the ab initio shape consisted of a globular trimeric region connected to a short stalk. This is consistent with the recently determined high-resolution structure of CPIII (20), shown superimposed on the low-resolution structure in Fig. 1A. Note that the first 21 residues of CPIII-His, at the tip of the stalk, were not resolved in the high-resolution structure. CPIII-Long had a similar shape to CPIII-His, except that the molecule was some 22 Å longer, as reflected in the D_{max} values (Table 1), owing to the increased length of the stalk region. This arises from the presence of the C-telopeptide and the last three triplets of the triple-helical region in CPIII-Long that are absent in CPIII-His. This greater size of CPIII-Long was also reflected in the observed increase in the radius of gyration R_g , compared with CPIII-His (Table 1). Individual shape determinations for CUB1CUB2, CPIII-His, and CPIII-Long are shown in Fig. S3 B–D. Within each group, individual shapes were very similar, and the fits to the scattering data were good, as illustrated in Fig. 1A, with corresponding χ^2 values given in Table 1.

When the different complexes were examined by SAXS, in both cases estimates of the molecular mass were consistent with the formation of complexes with a 1:1 stoichiometry (Table 1). In addition, comparison of R_g and D_{max} values gave early indications about the localization of CUB1CUB2 on its binding partner (CPIII-His or CPIII-Long; Table 1). For both complexes, there was no change in D_{max} on CUB1CUB2 binding compared with CPIII-His or CPIII-Long alone [$p(r)$ curves shown in Fig. S3A]. In addition, there was little change in R_g . These data suggest that CUB1CUB2 binds to a site on its binding partner that is relatively central (i.e., changing neither the overall length of the particle nor the radial distribution of mass).

To further localize the CUB1CUB2 binding site, we used the program MONSA to build ab initio shapes using the scattering curves for the complexes together with those for their individual components (24). As shown in Fig. 1B, this resulted in shapes in which CUB1CUB2 was bound along the stalk of the CPIII trimer, with one end in contact with the globular domain, in both CPIII-His and CPIII-Long. For CPIII-His, CUB1CUB2 seemed to be bound in almost all (>90%) shape determinations in register with the foot of the CPIII stalk (i.e., at the BMP-1 cleavage site) (Fig. 1B and Fig. S3E). In contrast, for CPIII-Long, the tip of the stalk was always visible (Fig. 1B and Fig. S3F). This is consistent with the additional presence of the C-telopeptide and the short triple-helical region in CPIII-Long. For the complexes,

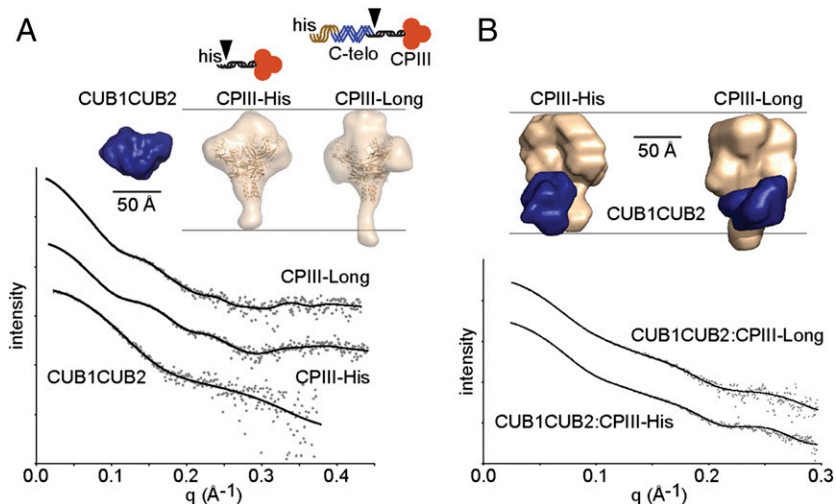


Fig. 1. SAXS. (A) Observed scattering profiles for individual proteins (dots) with DAMMIF-generated fits (solid lines). Proteins injected at 2.8 mg/mL (CUB1CUB2), 26 mg/mL (CPIII-His), and 11 mg/mL (CPIII-Long). Also shown, averaged low-resolution structures for CUB1CUB2 (blue) and CPIII-His/CPIII-Long (wheat color; CPIII-His crystal structure superimposed) with corresponding diagrams (his, His-tag; triple helix, brown; C-telopeptide, blue; C-propeptide trimer, red with black stalk region; arrowheads, BMP-1 cleavage sites). (B) Observed scattering profiles for complexes (dots) with MONSA-generated fits (solid lines). Samples injected at 11–12 mg/mL total protein (2:1 CUB1CUB2 molar excess), followed by complex isolation by on-line gel filtration. Also shown, averaged low-resolution structures, same color code as in A. Scattering curves shifted arbitrarily in the vertical direction (logarithmic scale) to improve clarity.

Table 1. Structural parameters for individual proteins and complexes

Protein or Complex	R_g (Å)	D_{max} (Å)	kDa (obs)	kDa (calc)	χ^2
CUB1CUB2	20.5 ± 0.17	69 ± 5	25 ± 0.7	28	1.7
CPIII-His	35.3 ± 0.22	130 ± 5	78 ± 0.9	86	1.5
CPIII-Long	39.5 ± 0.12	152 ± 5	88 ± 1.2	96	1.6
CUB1CUB2: CPIII-His	40 ± 0.5	130 ± 5	107 ± 0.9	114	2.7 (1.6, 2.3)
CUB1CUB2: CPIII-Long	41.8 ± 0.16	148 ± 5	112 ± 0.9	124	2.3 (1.7, 1.8)

Errors for radius of gyration R_g and kDa (obs) are standard deviations (N at least 10) based on estimates derived from Guinier plots. Errors in maximum diameter D_{max} are based on multiple GNOM runs. Typical χ^2 values are shown for the ab initio modeling. For individual proteins, fits were calculated using DAMMIF. For complexes, fits were calculated using MONSA, where the first figure shows the χ^2 value for the complex, and the figures in parentheses refer to each separate component (CUB1CUB2 followed by CPIII-His or CPIII-Long). Obs, observed; calc, calculated.

again the individual shape determinations were very similar, with good fits to the scattering data (Fig. 1B and Table 1). Use of the program MONSA assumes that there were no major conformational changes in either partner on formation of the complex. To check this, we measured the CD spectra of equimolar concentrations of CUB1CUB2 with CPIII-Long both before and after mixing. The data for the summed individual spectra and for the mixture were found to be similar, showing the absence of major changes in secondary structure on binding (Fig. S1B).

In conclusion, the SAXS data suggest that the CUB1CUB2 fragment of PCPE-1 binds to the stalk region of the procollagen III C-propeptide trimer, at the junction with the globular region. This zone corresponds to the N-terminal ~80 residues of the amino acid sequence of each chain. Fig. 2A and B shows the high-resolution structure of CPIII-His (20), as well as the predicted structure of CUB1CUB2, superimposed on the low-resolution shape for the CPIII-His:CUB1CUB2 complex determined by SAXS.

Identification of Critical Residues. There is an emerging paradigm in the field concerning the interactions of CUB domain-containing proteins (such as the serine proteinases of the complement system or the protein cubilin involved in vitamin B₁₂ uptake) with their cognate binding partners (25–27). Interactions occur between surface-located Ca²⁺-coordinating aspartate residues in the CUB domains and exposed basic residues (lysine or arginine) on the partner. Prompted by these observations, we looked for putative interacting residues in the region of CPIII identified by SAXS as the PCPE-1 binding zone. Sequence alignment among different procollagen types and species using

the UniProt-SwissProt database (Fig. S4) revealed three highly conserved lysine residues, at position 18 (in the stalk region; numbering begins at the BMP-1 cleavage site) and at positions 35 and 45 (in the base region) as possible CUB1CUB2 interaction sites. These residues were then individually mutated and their effects on PCPE-1 binding and enhancing activity determined. As shown in Fig. 3A (and in Fig. S5), all three mutations had no effect on basal cleavage by BMP-1 (leading to release of the C-propeptides) in the absence of PCPE-1. In contrast, for the K18A and K35Q variants, these mutations completely abrogated PCPE-1 enhancing activity. As for the K45Q mutation, however, PCPE-1 enhancing activity was completely unaffected. These observations were perfectly mirrored in the binding assays (Fig. S6). By surface plasmon resonance (Biacore), the K45Q mutation had no effect on PCPE-1 binding to CPIII-Long, compared with WT. In contrast, there was no binding for both the K18A and K35Q mutants. To check for possible changes in conformation in CPIII-Long introduced by the mutations, we carried out CD analysis. In all cases, the CD spectra were indistinguishable from WT (Fig. 3B). We conclude that lysines 18 and 35 are required for the interaction of PCPE-1 with CPIII, thus reinforcing the SAXS data indicating binding of the CUB1CUB2 region of PCPE-1 to the CPIII stalk region at the junction with the base. Fig. 2C shows the positions of the mutated lysines in the 3D structure of CPIII.

We previously identified critical residues in the CUB1 domain of PCPE-1 required for binding to CPIII (28). These include a surface exposed phenylalanine residue that is unique to the CUB1 domains of PCPEs, as well as two aspartate residues that are likely, by analogy with CUB domain-containing proteins of known structure (26), to be involved in Ca²⁺ coordination. Here we modeled the structure of the CUB2 domain of PCPE-1 and identified two corresponding aspartate residues, D191 and D233 (Fig. S2D; numbering begins at the N terminus of the mature protein). We then mutated these residues, individually, to alanines. Attempts to express the D233A mutant were unsuccessful, probably owing to poor folding or lack of stability. We were, however, able to express the D191A mutant. As shown in Fig. 3C, when tested for enhancing activity in the presence of BMP-1 using CPIII-Long as substrate, CUB1CUB2 containing the D191A mutation was no longer able to enhance BMP-1 activity. In addition, by surface plasmon resonance, the D191A mutant was unable to bind to CPIII-Long, unlike WT CUB1CUB2, which bound in a concentration-dependent manner (Fig. S2E). By CD, the spectrum for the D191A mutant was on the whole similar to WT (Fig. S2A), although it differed somewhat in the region below 210 nm, perhaps reflecting local unfolding in the region of the mutation. This is supported by deconvolution of the spectra, which showed only small differences in secondary structure between WT CUB1CUB2 and the D191A mutant (Table S1). Taken together, these data show that an aspartate

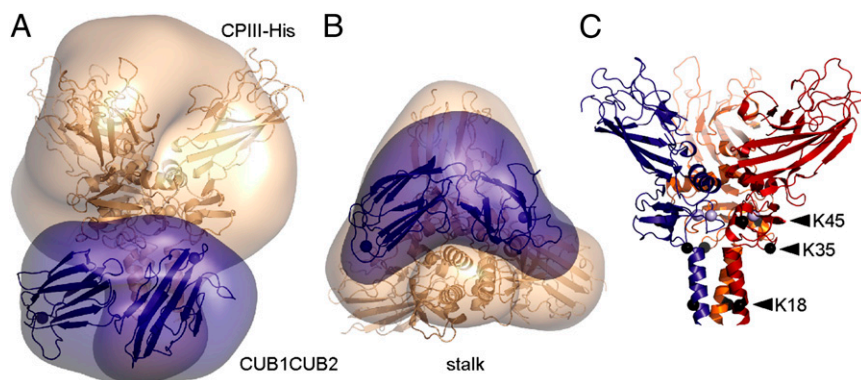


Fig. 2. Structure of the complex and positions of conserved lysines. (A) Superposition of the crystal structure of CPIII (20) (PDB code 4AK3), as well as the predicted structure of CUB1CUB2, onto the low-resolution shape of the CPIII-His:CUB1CUB2 complex. CPIII is in wheat color and CUB1CUB2 in blue. (B) Same as A but rotated about the horizontal axis such that the α -helical coiled-coil stalk region is at the bottom. (C) Positions of conserved lysine residues (shown by black spheres located at the C- α atoms) in the 3D structure of CPIII-His. Each chain is represented in a different color, with bound Ca²⁺ ions shown as light blue spheres.

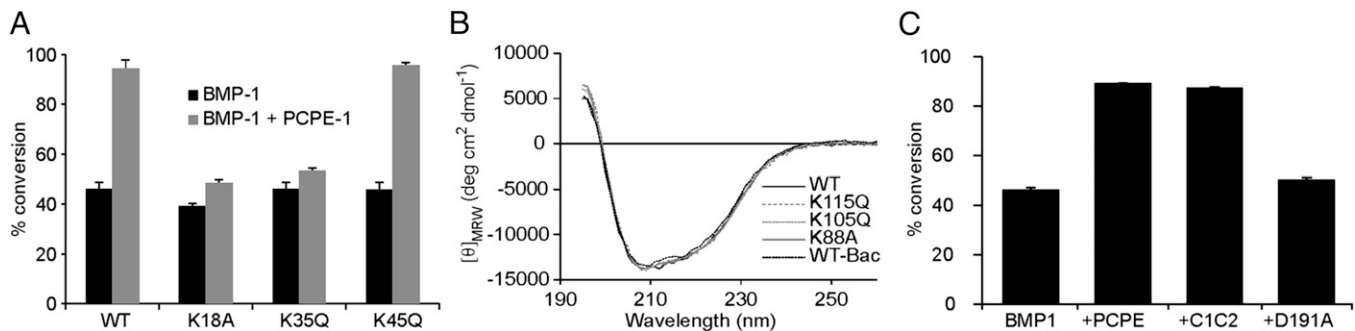


Fig. 3. Localization of interacting residues. (A) Percent conversion of CPIII-Long [340 nM; both WT and the K18A, K35Q, and K45Q mutants] by BMP-1 (16 nM) in the absence and presence of PCPE-1 (340 nM), incubated for 1 h at 37 °C. Analysis by SDS/PAGE (reducing conditions) and Sypro Ruby staining, followed by fluorescence quantitation. Error bars correspond to SDs ($n = 3$). (B) CD spectra of WT CPIII-Long and its K115Q, K105Q, K88A, and WT-Bac mutants, all produced in HEK 293T cells, analyzed at 25 °C in 20 mM Tris-HCl (pH 7.4), 2.5 mM CaCl_2 . Also shown is the CD spectrum for CPIII-Long produced in insect cells (WT-Bac), which is indistinguishable from that for protein produced in 293T cells. Protein concentrations 219–400 $\mu\text{g}/\text{mL}$. (C) Percent conversion of CPIII-Long (350 nM) by BMP-1 (16 nM) in the absence and presence of full-length PCPE-1 or its CUB1CUB2 fragment [both WT (C1C2) and the D191A mutant], all at 350 nM, incubated for 1 h at 37 °C. Analysis by SDS/PAGE (reducing conditions) and Sypro Ruby staining, then fluorescence quantitation. Error bars correspond to SDs ($n = 4$).

residue most likely to be involved in Ca^{2+} coordination in CUB2 is also required for PCPE-1 enhancing activity, as previously shown for Ca^{2+} -coordinating aspartate residues in the CUB1 domain (28).

Interactions and Enzyme Kinetics. We also used surface plasmon resonance to study the interaction of BMP-1 with the substrate, in this case miniprocollagen III, in the presence and absence of CUB1CUB2 (Fig. 4A). For these experiments we immobilized miniprocollagen III and used the catalytically inactive E94A mutant of BMP-1. When different concentrations of BMP-1 E94A were first injected over immobilized miniprocollagen, a heterogeneous ligand model could be fitted to the sensorgrams ($\chi^2 = 0.90$) with dissociation constants K_D of 72 nM and 117 nM. After regeneration, the surface was then saturated with 50 nM CUB1CUB2, followed by coinjection using the same range of BMP-1 concentrations but in the continued presence of 50 nM CUB1CUB2. This resulted in a second series of sensorgrams, to which again the heterogeneous ligand model was fitted ($\chi^2 =$

10.7) but this time giving K_D s of 1.7 nM and 42 nM. These data show that the binding affinity of BMP-1 E94A for miniprocollagen III is increased when CUB1CUB2 is also bound to the substrate.

Surprisingly, early studies on the kinetics of BMP-1 activity, using a procollagen I substrate, reported increases in both kinetic constants k_{cat} and K_m in the presence of PCPE (10), indicating that PCPE somehow diminishes the affinity of the enzyme for its substrate while at the same time increasing the turnover rate. To further investigate this question, we carried out detailed analysis of the cleavage of CPIII-Long by BMP-1 in the presence and absence of PCPE-1. Apparent kinetic parameters were calculated by fitting to the data shown in Fig. 4B. In the presence of PCPE-1, a k_{cat} of $110 \pm 10 \text{ min}^{-1}$ and a K_m of $2,452 \pm 379 \text{ nM}$ were obtained ($n = 3$). In the absence of PCPE-1 it was difficult to obtain accurate kinetic constants owing to the very high K_m and hence the need to work at prohibitively high substrate concentrations. We obtained a k_{cat} of $26 \pm 10 \text{ min}^{-1}$ and a K_m of $6,877 \pm 3,285 \text{ nM}$ ($n = 3$). These data indicate therefore that there was

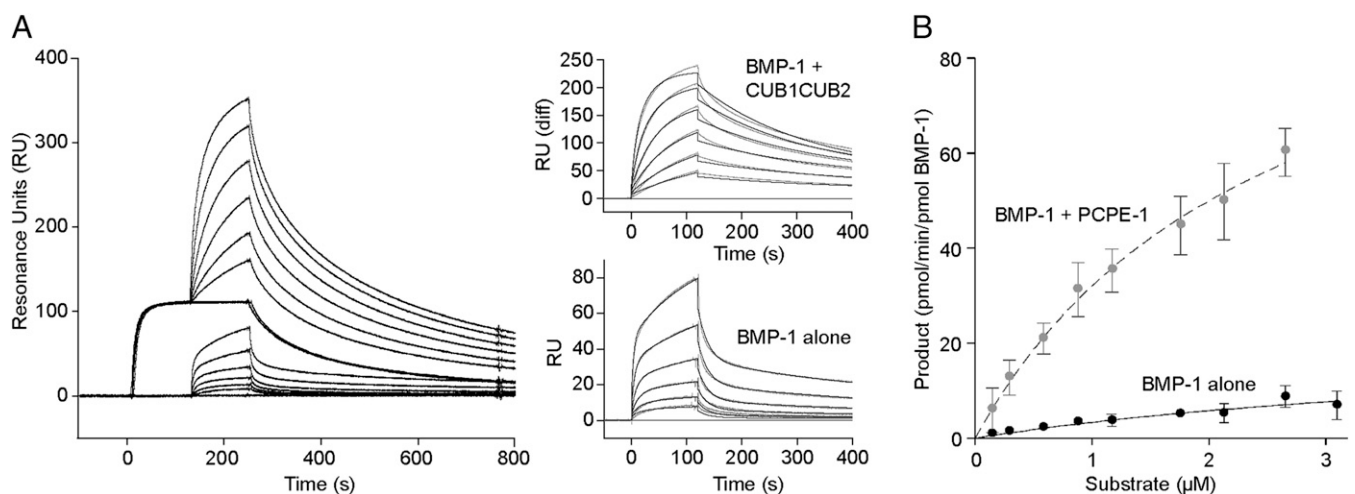


Fig. 4. Interactions and enzyme kinetics. (A) Interaction of the inactive E94A mutant of BMP-1 (9.4–300 nM) with immobilized miniprocollagen III (840 RU), in the absence or presence of CUB1CUB2. For the latter, the surface was first saturated for 120 s with 50 nM CUB1CUB2, followed by coinjection for a further 120 s with BMP-1 (9.4–300 nM) and CUB1CUB2 (50 nM). In all cases, dissociation was with running buffer. Flow rate 30 $\mu\text{L}/\text{min}$, temperature 25 °C. (Insets) Fits (black lines) to the experimental data, those for the coinject experiment after subtraction of data for CUB1CUB2 alone. (B) Kinetics of PCPE-1 enhancement of the cleavage of CPIII-Long by BMP-1. Substrate (CPIII-Long; 0–3 μM) was incubated with BMP-1 in the absence and presence of PCPE-1 (used at the maximum substrate concentration used), the concentration of BMP-1 being adjusted to remain within the linear range of the reaction. Analysis by SDS/PAGE (reducing conditions), then Sypro Ruby staining and fluorescence quantitation. Data show means from three independent experiments with error bars corresponding to SDs.

a large increase in k_{cat} in the presence of PCPE-1, and a small decrease in K_m . Values of k_{cat}/K_m increased 10-fold, from $0.0037 \text{ min}^{-1} \text{ nM}^{-1}$ in the absence of PCPE-1 to $0.044 \text{ min}^{-1} \text{ nM}^{-1}$ in its presence.

Discussion

Here, by a combination of SAXS and site-directed mutagenesis, we localize the binding region for PCPE-1 on the C-propeptide region of CPIII. The data show that the CUB1CUB2 fragment of PCPE-1 binds to the stalk/base region of CPIII (20), and this involves two conserved lysine residues in the latter (i.e., two per chain or six per trimer), one in the coiled-coil part of the stalk and the other at the junction with the base region. Both lysines are conserved in the C-propeptide regions of all of the main fibrillar procollagens, with the notable exception of the $\text{pro}\alpha 1(\text{V})$ chain (Fig. S4), for which C-propeptide cleavage seems to be unaffected by PCPE-1 (29). In contrast, these lysines are present in the $\text{pro}\alpha 2(\text{V})$ C-propeptide, for which cleavage is enhanced by PCPE-1 (29), thus reinforcing the results presented here.

We previously showed that both the binding and enhancing activities of PCPE-1 require a pair of aspartate residues in the CUB1 domain that are universally conserved in all of the CUB domains of human PCPEs and other Ca^{2+} -binding CUB domain-containing proteins (26, 28). Here we show that one of the corresponding aspartates in the CUB2 domain is also essential for binding and enhancing activity. In CUB domains of known structure, including those found in tumor necrosis factor stimulated gene 6 protein (TSG-6) [Protein Data Bank (PDB) code 2WNO], cubilin (25), and the complement initiating proteinases C1s (30), mannose binding lectin associated serine protease (MASP)-2 (31), and MASP-1 (32), these aspartate residues, as well as conserved glutamate and tyrosine residues, are involved in Ca^{2+} coordination (26). It is therefore likely that these are also Ca^{2+} ligands in PCPEs. In addition, two of these conserved residues (the glutamate and one of the aspartates) have surface exposed carboxy oxygens, which in the case of the cubilin/intrinsic factor/cobalamin complex have been shown to participate in interactions with surface lysine or arginine residues on their cognate binding partner (25). An identical scenario occurs in the binding of MASP-1 to the exposed lysine residue on mannan binding lectin (27, 33, 34), and this is also likely to occur in MASP binding to ficolins (35, 36) and in C1r/C1s binding to C1q (37). Here we show that binding of the CUB1CUB2 region of PCPE-1 to CPIII also requires residues involved in Ca^{2+} coordination in both CUB domains, as well as surface exposed lysine residues in CPIII. These results therefore extend the known repertoire of such CUB domain interactions, previously seen in such diverse biological systems as complement activation, receptor signaling and transport, to extracellular matrix assembly.

The observation that the stoichiometry of the PCPE-1:CPIII interaction is 1:1 (i.e., one molecule of PCPE-1 per C-propeptide trimer) is intriguing. We show here that binding of the CUB1CUB2 region of PCPE-1 to CPIII involves sites on both the CUB1 and CUB2 domains, as well as lysines K18 and K35 in the stalk/base region of CPIII, these residues being present in each of the three chains. It should be noted that strong cooperativity between the CUB1 and CUB2 domains in binding to CPIII has already been observed using single- and two-domain constructs (17). As for the sites on CPIII, we cannot determine from the low-resolution structure precisely which of the three chains are involved in binding to CUB1CUB2. From the data shown in Fig. 2A and B, CUB1CUB2 seems to straddle the stalk of CPIII, suggesting interactions with multiple chains. The observation that PCPE-1 does not enhance BMP-1 cleavage of individual procollagen chains (12) further supports this hypothesis. Whatever the topology of the interaction, however, owing to the trimeric nature of CPIII, with each chain carrying a pair of lysines

K18 and K35, a single CUB1CUB2 molecule should theoretically have a choice of three ways in which to bind to CPIII. We suggest that only one CUB1CUB2 molecule is bound because of steric exclusion, such that binding of one molecule of CUB1CUB2 to CPIII blocks binding of additional molecules, as suggested by Fig. 2B. In this way, binding of one CUB1CUB2 molecule to one C-propeptide trimer would introduce marked asymmetry in the complex, in addition to that already detected in CPIII itself (20). The presence of such asymmetry may provide clues to the mechanism of enhancing activity (see below).

What then is the mechanism of action of PCPE? Two possible consequences of PCPE-1 binding to the C-propeptide trimer can be envisaged (not mutually exclusive), leading to changes in the affinity of the enzyme for its substrate or changes in turnover rate. Concerning the first possibility, complex formation might create a new interaction surface, involving both substrate and enhancer, that increases the affinity of BMP-1 for its substrate. This is supported by the decrease in K_D measured by Biacore in the presence of CUB1CUB2, which indicates a higher affinity of BMP-1 for its substrate. Further support comes from the decrease in K_m in the kinetics data, in the presence of PCPE-1. Concerning the second possibility, PCPE-1 binding might lead to a change in structure of the enzyme:substrate complex, giving rise to the observed increase in catalytic activity. The proposed conformational change could occur in the substrate, in the enzyme, or both. Comparison with other proteolytic mechanisms provides possible clues. For example, there is evidence that cleavage of collagen molecules by at least some matrix metalloproteinases (MMPs) requires a conformational change in the substrate, the triple-helical structure being too large to enter the active site cleft (38, 39). Noncatalytic domains such as the linker/haemopexin domain in MMP-1 or MMP-14, or the fibronectin II-like collagen binding domain in MMP-2, seem to stabilize partially unwound forms of the substrate, thereby exposing individual polypeptide chains (40, 41). This also could be a role for PCPEs, to interact with the trimeric substrate (in this case the stalk of the C-propeptide trimer) and present individual chains for cleavage by BTPs. This is consistent with the observed binding of PCPE-1 to the stalk region of CPIII, close to the BMP-1 cleavage site. In contrast, we found no evidence of a change in secondary structure of the substrate on PCPE binding, albeit small changes would not be detectable by CD. Alternatively, interaction with the procollagen:PCPE complex might lead to a conformational change in BMP-1, thereby increasing catalytic efficiency. In this regard, recent observations on the low-resolution structures of BMP-1, mammalian tolloid, and mammalian tolloid like-1 suggest a substrate exclusion mechanism of enzyme regulation involving the noncatalytic CUB and epidermal growth factor domains (42, 43).

Identification of the site of interaction of PCPE-1 on the procollagen substrate could have important implications for the development of novel strategies to prevent the excessive accumulation of collagen seen in fibrotic diseases. Several strategies have been proposed to prevent fibrosis, including reducing the deposition of extracellular matrix (2). Different steps along the collagen biosynthetic pathway have been targeted, notably cleavage of the procollagen C-propeptides, the rate-limiting step in collagen fibril assembly. In view of the increasing number of other BTP substrates, however (currently approximately 20 in total) (8, 9), using inhibitors targeted to the active site of the enzyme could have undesirable side effects. In contrast, given the substrate-specific nature of PCPE-1 enhancing activity, which seems to be limited to C-terminal processing of the major fibrillar procollagens, strategies aimed at blocking the action of PCPEs could be a promising therapeutic approach, especially because there is evidence of PCPE up-regulation in animal models of liver and cardiac fibrosis (14–16).

Materials and Methods

Recombinant Proteins. All recombinant proteins were human forms. Proteins CP111-His, CP111-Long, and CUB1CUB2, all mutated at their N-glycosylation sites, were expressed in insect cells (baculovirus system) and also in human embryonic kidney 293T cells (including site-directed mutants), as described in *SI Materials and Methods* and Fig. S7. Recombinant CP111, PCPE-1, miniprocollagen III, BMP-1-FLAG, and its inactive E94A mutant were produced in 293-EBNA cells and purified as previously described (12, 28, 44). CUB1CUB2 was also produced by limited proteolysis of full-length recombinant PCPE-1 (not mutated at the N-glycosylation site) with trypsin (17).

Characterization, Interactions, and Kinetics. CD, isolation of complexes, interaction analysis by surface plasmon resonance and also enzyme kinetics of BMP-1 in the presence and absence of PCPE-1 were carried out as described in *SI Materials and Methods*.

Structural Analysis. To determine low-resolution structures, proteins and complexes were analyzed by SAXS, on the SWING beamline at the SOLEIL synchrotron (St. Aubin, France) (21). For ab initio shape determination, the programs DAMMIF (22), DAMMIN (45), GASBOR (46), and MONSA (24) were used. Further information is provided in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank the following colleagues for help and advice: A. Chaboud, I. Grosjean, and Y. Tauran [Protein Production and Analysis Facility, Unité Mixte de Service (UMS) 3444]; F. Delolme (Protein Microanalysis Facility, UMS 3444); N. Aghajari, X. Robert, and U. Sharma [Unité Mixte de Recherche (UMR) 5086]; P. Robin (SOLEIL synchrotron); Y. Zhao (Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford); and E. Kessler (Sackler Faculty of Medicine, Tel Aviv University). The work was funded by the Région Rhône-Alpes (D.J.S.H. and N.T.), Fondation de France Grant 11878 (to D.J.S.H.), Agence Nationale de la Recherche Grants ANR 07 PHYSIO 022 01 (to D.J.S.H.) and ANR 2010 BLAN 1526 (to C.M.), the Centre National de la Recherche Scientifique, the Université Lyon 1, and Lyonbiopôle.

- Kisseleva T, Brenner DA (2008) Mechanisms of fibrogenesis. *Exp Biol Med* (Maywood) 233(2):109–122.
- Wynn TA, Ramalingam TR (2012) Mechanisms of fibrosis: Therapeutic translation for fibrotic disease. *Nat Med* 18(7):1028–1040.
- Moali C, Hulmes DJS (2009) Extracellular and cell surface proteases in wound healing: new players are still emerging. *Eur J Dermatol* 19(6):552–564.
- Kadler EG, Kadler KE (2005) Procollagen trafficking, processing and fibrillogenesis. *J Cell Sci* 118(Pt 7):1341–1353.
- Kadler KE, Hill A, Canty-Laird EG (2008) Collagen fibrillogenesis: Fibronectin, integrins, and minor collagens as organizers and nucleators. *Curr Opin Cell Biol* 20(5):495–501.
- Kessler E, Takahara K, Biniaminov L, Brusel M, Greenspan DS (1996) Bone morphogenetic protein-1: The type I procollagen C-proteinase. *Science* 271(5247):360–362.
- Li SW, et al. (1996) The C-proteinase that processes procollagens to fibrillar collagens is identical to the protein previously identified as bone morphogenetic protein-1. *Proc Natl Acad Sci USA* 93(10):5127–5130.
- Muir A, Greenspan DS (2011) Metalloproteinases in *Drosophila* to humans that are central players in developmental processes. *J Biol Chem* 286(49):41905–41911.
- Moali C, Hulmes DJS (2012) *Extracellular Matrix: Pathobiology and Signaling*, ed Karamanos NK (De Gruyter, Berlin), pp 539–561.
- Adar R, Kessler E, Goldberg B (1986) Evidence for a protein that enhances the activity of type I procollagen C-proteinase. *Coll Relat Res* 6(3):267–277.
- Steiglitz BM, Keene DR, Greenspan DS (2002) PCOLCE2 encodes a functional procollagen C-proteinase enhancer (PCPE2) that is a collagen-binding protein differing in distribution of expression and post-translational modification from the previously described PCPE1. *J Biol Chem* 277(51):49820–49830.
- 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.
- Petropoulou V, Garrigue-Antar L, Kadler KE (2005) Identification of the minimal domain structure of bone morphogenetic protein-1 (BMP-1) for chordinase activity: Chordinase activity is not enhanced by procollagen C-proteinase enhancer-1 (PCPE-1). *J Biol Chem* 280(24):22616–22623.
- Ogata I, et al. (1997) Up-regulation of type I procollagen C-proteinase enhancer protein messenger RNA in rats with CCl₄-induced liver fibrosis. *Hepatology* 26(3):611–617.
- Kessler-Icekson G, Schlesinger H, Freimann S, Kessler E (2006) Expression of procollagen C-proteinase enhancer-1 in the remodeling rat heart is stimulated by aldosterone. *Int J Biochem Cell Biol* 38(3):358–365.
- Baicu CF, et al. (2012) Effects of the absence of procollagen C-endopeptidase enhancer-2 on myocardial collagen accumulation in chronic pressure overload. *Am J Physiol Heart Circ Physiol* 303(2):H234–H240.
- 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.
- 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.
- Moschovich L, et al. (2001) Folding and activity of recombinant human procollagen C-proteinase enhancer. *Eur J Biochem* 268(10):2991–2996.
- Bourhis JM, et al. (2012) Structural basis of fibrillar collagen trimerization and related genetic disorders. *Nat Struct Mol Biol* 19(10):1031–1036.
- David G, Perez J (2009) Combined sampler robot and high-performance liquid chromatography: A fully automated system for biological small angle X-ray scattering experiments at the synchrotron SOLEIL SWING beamline. *J Appl Cryst* 42:892–900.
- Franke D, Svergun DI (2009) DAMMIF, a program for rapid ab initio shape determination in small-angle scattering. *J Appl Cryst* 42:342–346.
- Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Cryst* 36:860–864.
- Petoukhov MV, Svergun DI (2006) Joint use of small-angle X-ray and neutron scattering to study biological macromolecules in solution. *Eur Biophys J* 35(7):567–576.
- Andersen CBF, Madsen M, Storm T, Moestrup SK, Andersen GR (2010) Structural basis for receptor recognition of vitamin-B(12)-intrinsic factor complexes. *Nature* 464(7287):445–448.
- Gaboriaud C, et al. (2011) Structure and properties of the Ca²⁺-binding CUB domain, a widespread ligand-recognition unit involved in major biological functions. *Biochem J* 439(2):185–193.
- Gingras AR, et al. (2011) Structural basis of mannan-binding lectin recognition by its associated serine protease MASP-1: Implications for complement activation. *Structure* 19(11):1635–1643.
- 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.
- Steiglitz BM, et al. (2006) Procollagen C proteinase enhancer 1 genes are important determinants of the mechanical properties and geometry of bone and the ultrastructure of connective tissues. *Mol Cell Biol* 26(1):238–249.
- Gregory LA, Thielens NM, Arlaud GJ, Fontecilla-Camps JC, Gaboriaud C (2003) X-ray structure of the Ca²⁺-binding interaction domain of C1s. Insights into the assembly of the C1 complex of complement. *J Biol Chem* 278(34):32157–32164.
- Gregory LA, et al. (2004) The X-ray structure of human mannan-binding lectin-associated protein 19 (MAP19) and its interaction site with mannan-binding lectin and L-ficolin. *J Biol Chem* 279(28):29391–29397.
- Teillet F, et al. (2008) Crystal structure of the CUB1-EGF-CUB2 domain of human MASP-1/3 and identification of its interaction sites with mannan-binding lectin and ficolins. *J Biol Chem* 283(37):25715–25724.
- Teillet F, et al. (2007) Identification of the site of human mannan-binding lectin involved in the interaction with its partner serine proteases: The essential role of Lys55. *J Immunol* 178(9):5710–5716.
- Wallis R, et al. (2004) Localization of the serine protease-binding sites in the collagen-like domain of mannose-binding protein: Indirect effects of naturally occurring mutations on protease binding and activation. *J Biol Chem* 279(14):14065–14073.
- Girija UV, Dodds AW, Roscher S, Reid KB, Wallis R (2007) Localization and characterization of the mannose-binding lectin (MBL)-associated-serine protease-2 binding site in rat ficolin-A: Equivalent binding sites within the collagenous domains of MBLs and ficolins. *J Immunol* 179(1):455–462.
- Lacroix M, et al. (2009) Residue Lys57 in the collagen-like region of human L-ficolin and its counterpart Lys47 in H-ficolin play a key role in the interaction with the mannan-binding lectin-associated serine proteases and the collectin receptor calreticulin. *J Immunol* 182(1):456–465.
- Bally I, et al. (2009) Identification of the C1q-binding Sites of Human C1r and C1s: A refined three-dimensional model of the C1 complex of complement. *J Biol Chem* 284(29):19340–19348.
- Chung L, et al. (2004) Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. *EMBO J* 23(15):3020–3030.
- Tam EM, Moore TR, Butler GS, Overall CM (2004) Characterization of the distinct collagen binding, helicase and cleavage mechanisms of matrix metalloproteinase 2 and 14 (gelatinase A and MT1-MMP): The differential roles of the MMP hemopexin c domains and the MMP-2 fibronectin type II modules in collagen triple helicase activities. *J Biol Chem* 279(41):43336–43344.
- Nerenberg PS, Salsas-Escat R, Stultz CM (2008) Do collagenases unwind triple-helical collagen before peptide bond hydrolysis? Reinterpreting experimental observations with mathematical models. *Proteins* 70(4):1154–1161.
- Manka SW, et al. (2012) Structural insights into triple-helical collagen cleavage by matrix metalloproteinase 1. *Proc Natl Acad Sci USA* 109(31):12461–12466.
- Berry R, et al. (2009) Role of dimerization and substrate exclusion in the regulation of bone morphogenetic protein-1 and mammalian tollid. *Proc Natl Acad Sci USA* 106(21):8561–8566.
- Berry R, Jowitz TA, Garrigue-Antar L, Kadler KE, Baldock C (2010) Structural and functional evidence for a substrate exclusion mechanism in mammalian tollid like-1 (TLL-1) proteinase. *FEBS Lett* 584(4):657–661.
- Bernocco S, et al. (2001) Biophysical characterization of the C-propeptide trimer from human procollagen III reveals a tri-lobed structure. *J Biol Chem* 276(52):48930–48936.
- Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76(6):2879–2886.
- Svergun DI, Petoukhov MV, Koch MH (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* 80(6):2946–2953.