The EMBO Journal vol.13 no.6 pp.1263–1269, 1994

The X-ray crystal structure of the catalytic domain of human neutrophil collagenase inhibited by a substrate analogue reveals the essentials for catalysis and specificity

Wolfram Bode[1], Peter Reinemer[1], Robert Huber, Thomas Kleine[2], Susanne Schnierer[2] and Harald Tschesche[2]

Max-Planck-Institut für Biochemie, Abteilung für Strukturforschung, D-82152 Martinsried, and [2]Universität Bielefeld, Fakultät für Chemie und Biochemie, D-33615 Bielefeld, Germany

[1]Corresponding author

Communicated by R.Huber

Matrix metalloproteinases are a family of zinc endopeptidases involved in tissue remodelling. They have been implicated in various disease processes including tumour invasion and joint destruction. These enzymes consist of several domains, which are responsible for latency, catalysis and substrate recognition. Human neutrophil collagenase (PMNL-CL, MMP-8) represents one of the two 'interstitial' collagenases that cleave triple helical collagens types I, II and III. Its 163 residue catalytic domain (Met80 to Gly242) has been expressed in *Escherichia coli* and crystallized as a non-covalent complex with the inhibitor Pro-Leu-Gly-hydroxylamine. The 2.0 Å crystal structure reveals a spherical molecule with a shallow active-site cleft separating a smaller C-terminal subdomain from a bigger N-terminal domain, composed of a five-stranded β-sheet, two α-helices, and bridging loops. The inhibitor mimics the unprimed (P1–P3) residues of a substrate; primed (P1′–P3′) peptide substrate residues should bind in an extended conformation, with the bulky P1′ side-chain fitting into the deep hydrophobic S1′ subsite. Modelling experiments with collagen show that the scissile strand of triple-helical collagen must be freed to fit the subsites. The catalytic zinc ion is situated at the bottom of the active-site cleft and is penta-coordinated by three histidines and by both hydroxamic acid oxygens of the inhibitor. In addition to the catalytic zinc, the catalytic domain harbours a second, non-exchangeable zinc ion and two calcium ions, which are packed against the top of the β-sheet and presumably function to stabilize the catalytic domain. The polypeptide folding and in particular the zinc environment of the collagenase catalytic domain bear a close resemblance to the astacins and the snake venom metalloproteinases.

Key words: collagenases/collagen digestion/matrix metalloproteinases/X-ray crystal structure/zinc endopeptidases

Introduction

Matrix metalloproteinases (MMPs, tissue collagenases, EC 3.4.24.7) are a family of zinc- and calcium-dependent endopeptidases responsible for degradation of extracellular matrix components such as interstitial and basement membrane collagens, fibronectin and laminin. They have been implicated in tissue remodelling processes associated with mammalian growth and development, ovulation and wound healing, but can also play a deleterious role in various disease processes such as rheumatoid joint destruction in rheumatoid and osteoarthritis, tumour invasion and periodontitis (for recent reviews see Woessner, 1991; Matrisian, 1992; Murphy and Docherty, 1992; Birkedal-Hansen *et al.*, 1993). At least nine different (yet highly homologous) MMP types have been characterized, including interstitial (FIB-CL, MMP-1) and neutrophil collagenase (PMNL-CL, MMP-8), two gelatinases [the 72 kDa (MMP-2) and the 92 kDa (MMP-9) gelatinases], the stromelysins [SL-1 (MMP-3), SL-2 (MMP-10) and matrilysin (pump, MMP-7)] and others (SL-3 (MMP-11), metalloelastase]. These proteinases are secreted as inactive multidomain-proenzymes, with the active-site zinc blocked by an unpaired cysteine (Springman *et al.*, 1990) within a strongly conserved PRCGVPD sequence motif of the ~80 residue propeptide which is removed in a stepwise process upon activation (Knäuper *et al.*, 1990; Nagase *et al.*, 1990). Following this propeptide is the zinc and calcium binding catalytic domain of ~170 residues, containing the conserved HEXX-HXXGXXH zinc signature (Bode *et al.*, 1992). This is followed in turn by a hemopexin-like domain of ~210 residues, which is important for the unique ability of FIB- and PMNL-CL to cleave triple-helical collagen (Hasty *et al.*, 1987; Murphy *et al.*, 1992; Knäuper *et al.*, 1993a; Sanchez-Lopez *et al.*, 1993; Schnierer *et al.*, 1993). The gelatinases possess an additional fibronectin-like domain which is inserted in the catalytic domain, and MMP-9 shows a unique extended hinge region between the catalytic and the hemopexin-like domain.

The interstitial collagenases, FIB-CL and PMNL-CL, are the most specific of these MMPs, exhibiting a unique triple helicase activity. Each of the three polypeptide chains of the native helical collagens of types I, II and III is cleaved at a single peptide bond, giving rise to fragments three quarters and one quarter of the native chain length. The neutrophil collagenase (PNML-CL, MMP-8) is synthesized during the early phases of cellular differentiation and is stored as a highly glycosylated protein in the specific granules of neutrophils which can be secreted as an inactive precursor after initiation by inflammatory mediators (Murphy *et al.*, 1977; Hasty *et al.*, 1986; Knäuper *et al.*, 1990). Upon activation cleavage of the proenzyme, active PMNL-CL species with N-terminal residues Phe79, Met80 or Leu81 have been observed (for sequence numbering see Figure 1), which differ significantly in activity (Grant *et al.*, 1987; Knäuper *et al.*, 1990; Mallya *et al.*, 1990; Suzuki *et al.*, 1990; Bläser *et al.*, 1991; Knäuper *et al.*, 1993b).

Here we present the X-ray crystal structure of the non-covalent complex formed between the catalytic domain (Met80–Gly242) of human PMNL-collagenase and the inhibitor Pro-Leu-Gly-hydroxylamine (NHOH), which

mimics binding of the N-terminal part of peptide substrates to the non-primed enzyme subsites (using the nomenclature of Schechter and Berger, 1967). Modelling experiments suggest that triple-helical collagen will not fit as a whole to the active-site region of collagenases but that single strands must be freed around their scissile peptide bond from the triple helix in order to be bound and cleaved.

Results and discussion

Overall structure

The catalytic domain of human PMNL-collagenase containing 163 residues from Met80 to Gly242 (using proenzyme sequence numbering, Figure 1) has been cloned, and the enzyme has been expressed in *Escherichia coli* and purified to homogeneity (Schnierer *et al.*, 1993). It has been co-crystallized with the substrate analogue Pro-Leu-Gly-NHOH, an inhibitor described for use in affinity column chromatography (Moore and Spilburg, 1986). The X-ray crystal structure of this non-covalent complex has been solved by multiple isomorphous replacement at 3.0 Å resolution and subsequently refined at 2.0 Å resolution (see Tables I–III). The structure shown in Figure 2 bears some similarity to crayfish astacin (Bode *et al.*, 1992; Gomis-Rüth *et al.*, 1993a) and snake venom adamalysin (Gomis-Rüth *et al.*, 1993b). It reveals that the PMNL-CL catalytic domain is a spherical molecule notched at the periphery to give a relatively flat active-site cleft, separating the small 'lower' subdomain from the 'upper' main body (Figure 2a). In this N-terminal Met80-form electron density is observed only from the 7th residue (Pro86) onwards. The first three quarters of the peptide chain (up to Gly204) form the regularly folded 'upper' domain. This domain consists mainly of a central, highly twisted five-stranded β-pleated sheet, flanked by an S-shaped 'double-loop' and two other bridging loops on its convex side, and by two long helices on its concave side (αA, αB in Figure 1). The sheet shields a large core built almost exclusively from hydrophobic residues. The 'lower' subdomain, comprising the last 38 residues, is organized in multiple turns, ending in a long C-terminal α-helix at the surface of the molecule (Figure 2b).

The 'catalytic' zinc ion (Zn999) is situated at the bottom of the active-site cleft. It is penta-coordinated by the three imidazole $N_{\epsilon 2}$ atoms of His197 (2.0 Å), His201 (2.2 Å) and His207 (1.9 Å) of the HEXXHXXGXXH zinc binding consensus motif, and by the carbonyl (2.0 Å) and the hydroxyl oxygen (2.2 Å) of the inhibitor's hydroxamic acid moiety (Figure 3). Besides these two oxygen–zinc contacts, the hydroxamate group is situated with its oxygen and nitrogen close to Glu198 $O_{\epsilon 1}$ and forms a hydrogen bond via its nitrogen to the carbonyl group of Ala161 of the extended edge strand Ala161-Phe164 (β4 in Figure 1) of the β-sheet. The active-site helix containing His197, Glu198 and His201 extends to Gly204, where the polypeptide chain turns sharply away from the helix axis towards His207, the third zinc ligand. There follows a wide right-handed loop, stabilized by a hydrogen bond between Ser208N and Met215O and a 1,4-tight turn Ala213-Leu214-Met215-Tyr216 known as the 'Met-turn' (Bode *et al.*, 1993; Gomis-Rüth *et al.*, 1993b). The side chain of Met215 provides a hydrophobic basement to the three zinc liganding histidine residues, a characteristic element for the 'metzincins' (Bode *et al.*, 1993).

The peptide chain then proceeds to the molecular surface at Pro217, where the chain is kinked, forming the outer wall of the putative S1′ pocket (Figure 3). The following C-terminal helix, situated at the bottom of the molecule

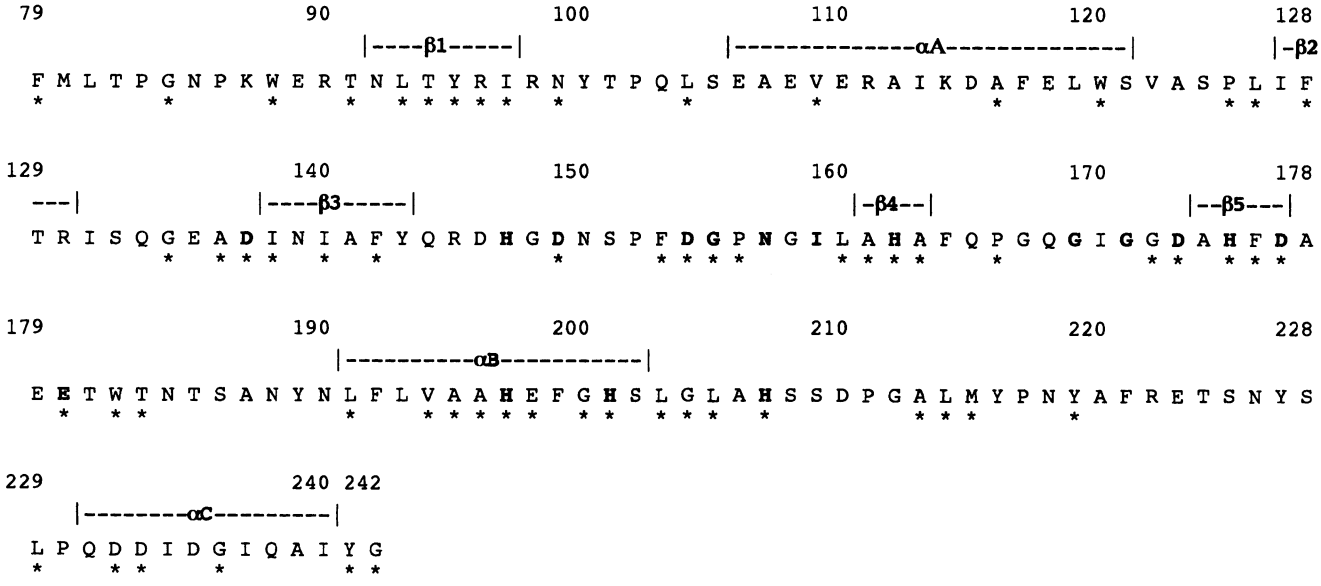

Fig. 1. Primary structure (given in one-letter code) for the catalytic domain of human PMNL-collagenase (Hasty *et al.*, 1990; Schnierer *et al.*, 1993). Sequence numbering corresponds to the short form procollagenase lacking the signal peptide and the hemopexin-like domain. Residues strictly conserved throughout the human collagenases and stromelysins [fibroblast collagenase (FIB-CL, MMP-1), neutrophil collagenase (PMNL-CL, MMP-8), stromelysin-1 (MMP-3), -2 (MMP-10) and -3 (MMP-11)] (for sequences and alignment see Birkedal-Hansen *et al.*, 1993) are marked with an asterisk (*). Secondary structure elements (αA, B, C: α-helices; β1, 2, 3, 4, 5: β-strands) calculated with DSSP (Kabsch and Sander, 1983) are overlayed over the primary structure. Residues involved in binding of the metal ions ['catalytic' zinc (Zn999): H197 ($N_{\epsilon 2}$), H201 ($N_{\epsilon 2}$), H207 ($N_{\epsilon 2}$), 'structural' zinc (Zn998): H147 ($N_{\epsilon 2}$), D149 ($O_{\delta 2}$), H162 ($N_{\epsilon 2}$), H175 ($N_{\delta 1}$), calcium site Ka997: D154 ($O_{\delta 2}$), G155 (O), N157 (O), I159 (O), D177 ($O_{\delta 1}$), E180 ($O_{\epsilon 1}$), calcium site Ka996: D137 (O), G169 (O), G171 (O), D173 ($O_{\delta 1}$)] are given in bold.

(Figure 2b), contains the strictly conserved residues Asp232 and Asp233. The buried side chain of the latter residue, which is essential for catalytic activity (Hirose *et al.*, 1993), forms a hydrogen bond to the Met-turn residues Leu214N and Met215N, thus stabilizing the active-site basement. Asp232 is positioned at the side of a hydrophobic groove formed by the C-terminal helix residues Gly236, Ala238 and Ile240 and closed by Leu205 and Met215. This region could stabilize the four N-terminal residues of the Phe79 'super-active' forms (Suzuki *et al.*, 1990; Knäuper *et al.*, 1993b), with the N-terminal ammonium group making a salt bridge to the carboxylate group of Asp232. We have recently solved the X-ray crystal structure of the Phe79 form, and have shown this to be the case (Reinemer *et al.*, 1994).

In addition to the 'catalytic' zinc ion, the catalytic domain of human PMNL-CL harbours another, hereafter referred to as 'structural' zinc ion (Zn998), and two calcium ions (Ka996, Ka997; see Figures 2 and 4). The second zinc (Zn998) and one of the two calcium ions (Ka997) are sandwiched between the surface S-shaped double loop Arg145-Leu160 (which connects strands β3 and β4 and bridges the spatially intervening strand β5) and the surface of the β-sheet (Figure 4). The N-terminal loop is formed by a hydrogen bond between Phe153N and Arg145O. The 'structural' zinc is coordinated in an almost tetrahedral manner by His147$N_{\epsilon 2}$ (2.0 Å) and Asp149$O_{\delta 2}$ (2.0 Å) of this loop and His162$N_{\epsilon 2}$ (1.9 Å) and His175$N_{\delta 1}$ (2.0 Å) of the opposing β-sheet. All four residues appear to be strictly conserved in all MMPs (except in SL-3, where His147 is replaced by an aspartate). We have been unable to remove or exchange this zinc in our crystals, suggesting an extremely tight binding of this zinc site, which is in agreement with exchange data in solution (Salowe *et al.*, 1993). The C-terminal part of this loop encircles one of the calcium ions (Ka997) and packs it against the side chains of Asp177 and Glu180, which protrude from the C-terminal end of strand β5 (see Figure 4). The carboxylate groups of these strictly conserved acidic residues, together with the three loop carbonyls and the Asp154 carboxylate, coordinate this calcium in a nearly octahedral manner; Gly155O, Asn157O,

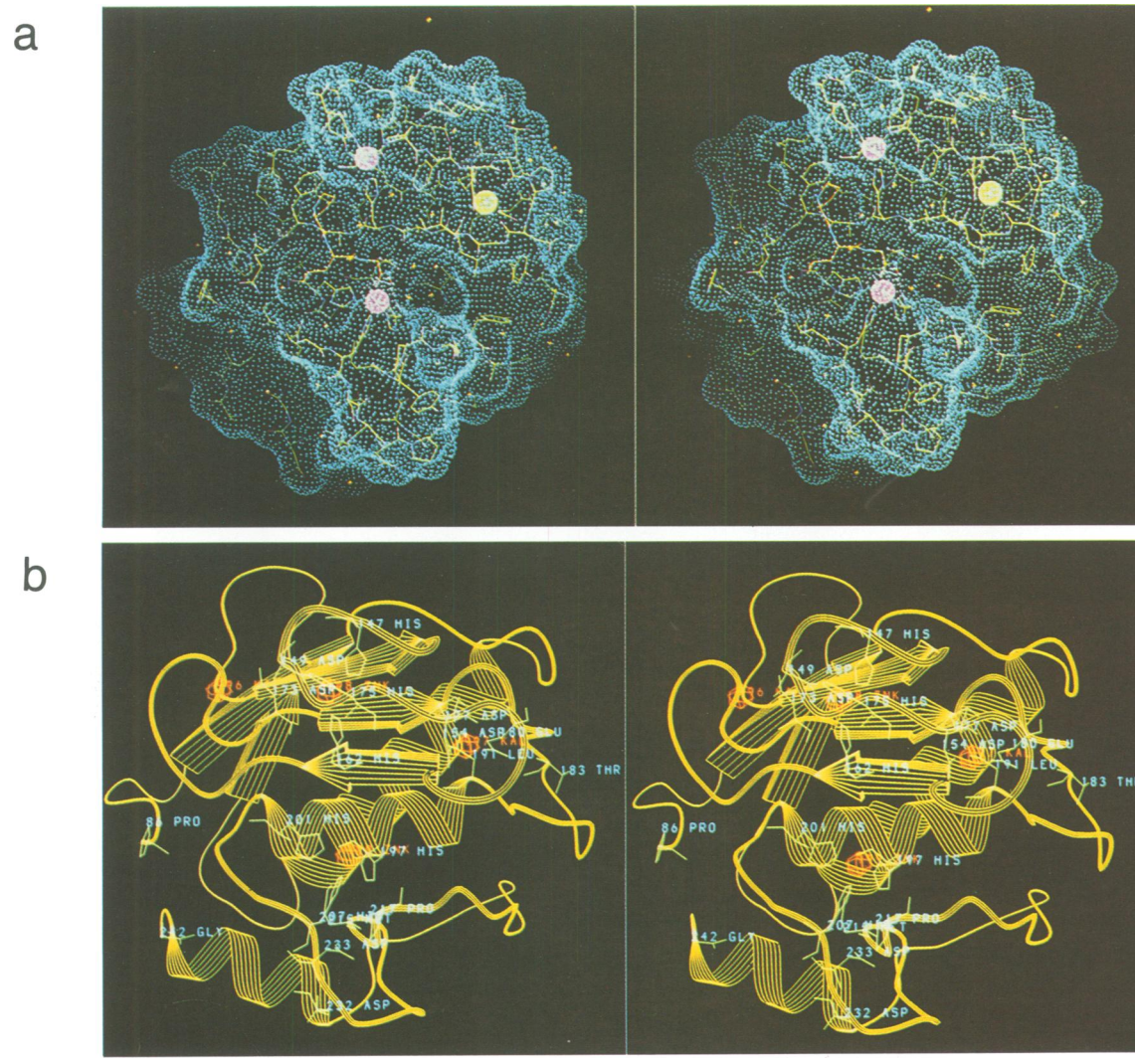

Fig. 2. Structure of the catalytic domain of human neutrophil collagenase (**a**) displayed together with a Connolly dot surface [using the program MAIN (Turk, 1992)] and (**b**) shown as a ribbon model [using the program RIBBON (Priestle, 1988; modified by A.Karshikoff)]. An upper (N-terminal) and a lower (C-terminal) domain are separated by a moderately deep active-site cleft, accommodating the catalytic zinc ion [displayed as a magenta-dotted sphere of radius 1 Å (a) or a red sphere (b)] near a very deep S1′ pocket (to the right of the catalytic zinc). The upper domain harbours the structural zinc ion (displayed in the same colours) and two calcium ions [displayed as yellow dot spheres (a) or red spheres (b)].

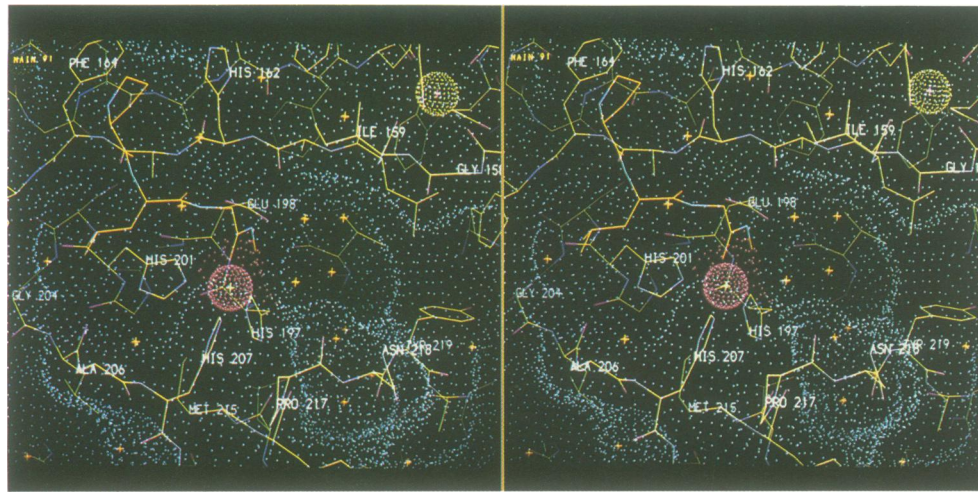

Fig. 3. A detailed view of the active site cleft of human PMNL-collagenase. Shown are the edge strand (β4), the zinc coordinating histidine residues (His197, His201) and the catalytically important glutamate residue (Glu198), which are part of the active-site helix (αB, background), the turn following the active-site helix containing the third zinc ligand (His207, bottom), the inhibitor Pro-Leu-Gly-NHOH (orange), the catalytic zinc ion (Zn999, magenta sphere), one of the calcium ions (Ka997, yellow sphere) and a few localized solvent molecules (crosses) together with an overlayed Connolly dot surface. The inhibitor probably mimics the unprimed residues of a productively bound peptide substrate, lying anti-parallel to the edge strand, and forming inter-main chain hydrogen bonds between the P2 residue LeuI2 and Ala163. The P3 residue ProI1 slots into a small cleft mainly formed by the side-chains of His162 and Phe164. The inhibitor's hydroxamate group is situated with its oxygen and nitrogen close to Glu198, forming a hydrogen bond to Ala161, while the carbonyl and the hydroxyl oxygen coordinate the zinc ion. A very deep S1′ cleft is visible to the right of the zinc ion, to which a P1′ side-chain of a productively bound substrate would extend.

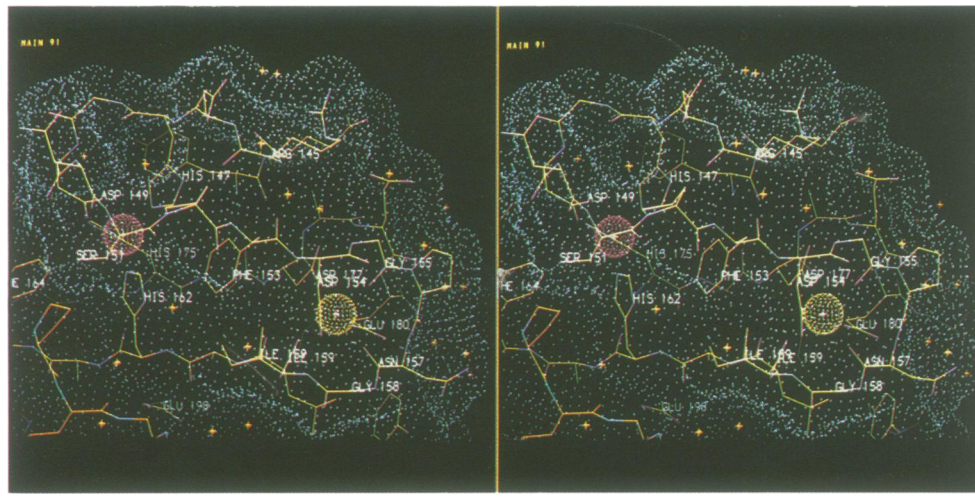

Fig. 4. A detailed view of the surface S-shaped double loop between Arg145 and His162 shown with the 'structural' zinc ion (Zn998, magenta sphere, top), a calcium ion (Ka997, yellow sphere, right), the inhibitor (orange) and a few localized solvent molecules (crosses) together with an overlaid Connolly dot surface. The zinc ion is tetrahedral coordinated by His147, Asp149, His162 and His175, while the calcium ion is octahedral coordinated by Asp154, Gly155, Asn157, Ile159, Asp177 and Glu180. All these residues with the exception of His147 are strictly conserved throughout the human collagenase family.

Ile159O, and Asp177$O_{\delta 2}$, 2.3, 2.2, 2.3 and 2.2 Å away, respectively, occupy the corners of a distorted square, while Asp154$O_{\delta 2}$ (2.3 Å) and Glu180$O_{\epsilon 1}$ (2.2 Å) are the axial ligands. Again, this loop structure is rather compact suggesting that this calcium ion helps stabilize the whole domain, also in agreement with results obtained in solution.

A second site, corresponding to a 6σ peak in an $F_o - F_c$ electron density omit map, is very probably occupied by a calcium (Ka996; for comparison Ka997 gives a 10σ peak in a similar $F_o - F_c$ omit electron density). It is found at the end of the glycine-rich open loop Phe164-Gly172, before the entrance to sheet strand β5 (to the left, Figure 2b). Ka996 is octahedrally coordinated, with the carbonyl groups of Gly169 (2.2 Å), Gly171 (2.1 Å), Asp137 (2.2 Å, strand β3) and a solvent molecule in equatorial position (2.1 Å), and the carboxylate group $O_{\delta 1}$ of the strictly conserved Asp173 (2.5 Å) and another solvent molecule (2.4 Å) providing the axial ligands.

Comparison with the sequences of other MMPs (Birkedal-Hansen *et al.*, 1993) shows that their catalytic domains can be modelled on our PMNL-CL model with reasonable confidence. All metal binding and catalytic residues and most internal residues are identical or at least homologous (see Figure 1); side-chains lining the (primed) substrate binding sites exhibit differences that probably confer individual specificities. The two gelatinases represent special cases, in that they contain a 182 residue insertion between Thr183 and Leu191. This insertion probably sits adjacent to the active-site cleft (to the right, Figure 2), with relatively few interdomainal contacts with the catalytic

domain, as only few of the PMNL-CL surface residues in this region differ drastically (Pro156, Asn157 in the zinc–calcium double loop, Asn188 and Asn190 at the entrance to the active-site helix).

Inhibitor binding

The Pro-Leu-Gly-NHOH, which probably mimics the unprimed residues of a productively bound peptide substrate, lies anti-parallel to the edge strand in a slightly twisted manner, forming two inter-main chain hydrogen bonds between the inhibitor's P2 residue Leu12 and Ala163 (Figure 3). The N-terminal proline residue slots into the hydrophobic cleft provided by the side chains of His162, Phe164 and Ser151, while the leucine residue nestles into a shallow groove lined by His201, Ala206 and His207, in agreement with earlier predictions (Odake *et al.*, 1991). Both proline and leucine are frequently found in the neighbourhood of collagenase cleavage sites of collagens, conferring beneficial effects to substrates and inhibitors if positioned at P3 and P2 (Moore and Spilburg, 1986; Netzel-Arnett *et al.*, 1991). The position of the glycine residue, which represents the P1 residue in all type I, II and III collagen cleavage sites (for references see Wu *et al.*, 1990 or Netzel Arnett *et al.*, 1991), might be slightly influenced by the zinc liganding of its terminating hydroxamate group. In a transiently bound substrate, the carbonyl group of the scissile peptide bond should be directed towards the 'catalytic' zinc (similar to the carbonyl of the hydroxamate), whilst the amido group should point towards the Ala161 carbonyl (as with the nitrogen of the hydroxamate), with a water molecule bound to Glu198 and Zn999 squeezed between, to attack the substrate carbonyl group (see Holmes and Matthews, 1981).

Preliminary model building studies show that substrate residues C-terminal to the scissile peptide bond could interact with the primed subsites of the enzyme in a sightly bent but essentially extended conformation. A P1′ side chain would extend into the deep, mainly hydrophobic S1′ pocket (to the right of the catalytic zinc in Figure 3), which is spacious enough to accommodate not only P1′ isoleucine and leucine residues (normally found in collagen substrates), but also aromatic side-chains up to the size of indole moieties, as observed experimentally (Netzel-Arnett *et al.*, 1991). The P2′-P3′ residues of a bound substrate would have to pass through an opening ~5 Å wide formed by the crossing-over peptide segments Asn157-Gly158 (part of the calcium loop) and Asn218-Tyr219 (forming the edge of the S1′ site), which would provide several amide groups for extensive main chain–main chain interactions (see Figure 3). Chemical compounds mimicking fragments of such a productively bound peptide substrate and/or filling out the S1′ pocket might represent powerful collagenase inhibitors and may be designed in a rational way on the basis of the receptor structure presented here.

Implications for triple helical collagen binding

The specific 'triple helicase' activity of FIB- and PMNL-CL requires the C-terminal (hemopexin-like) structures (Hasty *et al.*, 1987; Murphy *et al.*, 1992; Knäuper *et al.*, 1993a; Sanchez-Lopez *et al.*, 1993; Schnierer *et al.*, 1993). However, the catalytic domain also contains some collagen-recognizing determinants (Sanchez-Lopez *et al.*, 1993). The repetitive Pro-X-Gly segment of one strand of a regular triple-helical collagen-like structure could probably nestle against the non-primed collagenase subsites in such a manner that the glycyl carbonyl group would approach the catalytic zinc. In this case, however, the P1′ (proline) side-chain would not fill the S1′ pocket adequately, and the 15 Å diameter collagen triple helix (Yonath and Traub, 1969; Fraser *et al.*, 1979) could not fit through the opening at S2′-S3′. Thus, the proper fit of the scissile chain to the enzyme subsites almost certainly requires at least partial unfolding of the collagen triple helix around the cleavage site.

Mutagenesis experiments with murine α1(I) collagen and expression of mixed triple-helical collagen molecules show that the presence of hydrophobic P1′ residues (such as leucine and isoleucine) and the absence of proline residues in the primed positions of native helical type I collagen are required to make it susceptible to interstitial collagenases (Wu *et al.*, 1990). The non-cleavable mutants do not exhibit enhanced melting temperatures, indicating that cleavability is not a property of the isolated triple helix itself, but of the collagen–collagenase complex. Obviously, the residues flanking the cleavage site of the scissile collagen strand must be freed from the triple-helical structure by energetically favourable interactions with the collagenase subsites. In collagen complexes with full-length collagenases, the hemopexin-like domain might sandwich the bound triple helix to the enzyme, structurally and temporally stabilizing the initial encounter complex to result in a more efficient cleavage. This is not a new phenomenon in proteinases; e.g. the apparently rigid tryptic cleavage site of chymotrypsinogen must adapt considerably to the binding subsites of trypsin in order to be cleaved (Wang *et al.*, 1985).

Comparison with other 'metzincins'

The 'upper' domain of the collagenase catalytic domain, with its five-stranded β-sheet and the two long helices, is topologically similar to the corresponding domains first observed in crayfish astacin (Bode *et al.*, 1992; Gomis-Rüth *et al.*, 1993a) and now also in the snake venom metalloproteinase adamalysin II (Gomis-Rüth *et al.*, 1993) and the *Pseudomonas aeruginosa* alkaline protease (Baumann *et al.*, 1993); its subdomain bears a particularly strong homology to that observed in adamalysin. Moreover, the active-site environments composed of the 'active-site helix', the following strand and the Met-turn exhibit virtually identical conformations. A structural comparison (using OVRLAP, Rossmann and Argos, 1975) reveals that ~116 and 113 of the 157 ordered residues in the catalytic domain of human PMNL-CL can be considered as topologically equivalent to astacin and adamalysin, respectively, having r.m.s. deviations of 2.2 and 2.9 Å for their α-carbon atoms. Of these equivalent residues, however, only 16 (15) are identical, six (10) of which are located in the zinc binding region alone. The substrate binding site has much more in common with adamalysin, in agreement with the 'collagenolytic' properties of this metalloproteinase; however, the overall scaffold and the mode of interaction of the N-terminus are more reminiscent of astacin. Thus, the collagenase catalytic domain exhibits many characteristic structural elements in common with the astacins, the adamalysins and the serralysins, with which they should be grouped into a common superfamily, the metzincins (Bode *et al.*, 1993).

Table I. Statistics of data collection

Derivative	Native[a]	Apo[b]	Apo + Zn[b] (= native)	Apo + Hg[b]
No. of measurements	47920	30392	32306	30537
No. of observations	45439	29336	30154	28625
No. of unique reflections	9740	5652	5935	5515
Completeness of data (%)				
$\infty-2.03$ Å[a]/$\infty-3.00$ Å[b]	86.3	94.8	93.7	93.9
$2.07-2.03$ Å[a]/$3.05-3.00$ Å[b]	57.8	88.8	90.0	88.8
R_{Merge}[c]	0.113	0.094	0.128	0.137
R_{Sym}[d]	0.050	0.046	0.060	0.050

[a]High resolution native data set, collected on MAR image plate scanner.
[b]Derivative data sets, collected on FAST television area detector; Apo was prepared by soaking of a native crystal in 20 mM *o*-phenanthrolin for 1 day; Apo + Hg was prepared by soaking of a native crystal in 20 mM *o*-phenanthrolin for 1 day and subsequent transfer to 5 mM $Hg(OOCCH_3)_2$ for 2 h; Apo + Zn = native was used as a 'derivative' in the initial stages with Apo as the 'native' in order to improve phasing.
[c]$R_{Merge} = \Sigma_h\Sigma_i\,[|(I(h,i)-<I(h)>|]/\Sigma_h\Sigma_i\,I(h,i)$, where $I(h,i)$ is the intensity value of the i-th measurement of h and $<I(h)>$ is the corresponding mean value of h for all i measurements of h: the summation is over all measurements.
[d]$R_{Sym} = \Sigma(|I_F-<I_F>|)/\Sigma I_F$, where I_F is the averaged value of point group related reflections and $<I_F>$ is the averaged value of a Bijvoet pair.

Table II. Statistics of phasing

	No. of sites	Occ.[a]	Atomic coordinates (Å)	B_{iso}[c] (Å^2)	Phasing power[d]	Mean figure of merit[e]
(a) Apo as native						
Apo + Hg	1	59.76	23.3, 12.1, 18.5	19.7	1.93	0.70
Apo + Zn	1	37.33	23.4, 12.1, 18.6	16.0	1.18	
(b) Apo + Zn as native						
Apo + Hg	1	36.87	22.7, 11.8, 18.7	14.6	1.48	0.59
Apo	1	−30.87	24.3, 12.5, 17.4	15.8	0.90	

[a]Site occupancy factor in relative units on native scale; scale factors derived by Wilson plot (Wilson, 1942): $k = 10.195$ for Apo and $k = 11.450$ for Apo + Zn = native, respectively.
[b]Real coordinates *X, Y, Z.*
[c]Isotropic temperature factors.
[d]F_H/residual: r.m.s. mean heavy atom contribution/r.m.s. residual, defined as $[(F_{PHC}-F_{PH})^2/n]^{1/2}$ with the sum over all reflections, where F_{PHC} is the calculated structure factor of the heavy atom derivative and F_{PH} is the structure factor amplitude of the heavy atom derivative.
[e]$25.0-3.0$ Å resolution.

Materials and methods

Purification and crystallization of Met80-Gly242 catalytic domain of human PMNL-CL

The Met80-Gly242 catalytic domain of human PMNL-CL was expressed in *E.coli* and purified as previously described (Schnierer *et al.*, 1993). Briefly, the recombinant enzyme was renatured by dialysing the inclusion bodies, solubilized in 6 M urea, 100 mM β-mercaptoethanol, against a buffer containing 100 mM NaCl, 5 mM $CaCl_2$, 0.5 mM $ZnCl_2$ and 20 mM Tris−HCl, pH 7.5. The renatured enzyme was subsequently purified to apparent homogeneity as judged by SDS−PAGE by hydroxamate affinity chromatography. Crystallizations were performed at 22°C. Hanging droplets were made by mixing 1.5 μl protein solution (~12 mg/ml protein in 5 mM $CaCl_2$, 100 mM NaCl, 3 mM MES−NaOH, 0.02% NaN_3, pH 6.0), 2 μl inhibitor solution (50 mM Pro-Leu-Gly-NHOH (Moore and Spilburg, 1986) in 0.2 M MES−NaOH, 0.02% NaN_3, pH 6.0) and 6 μl PEG solution [10% (m/v) PEG 6000, 0.2 M MES−NaOH, 0.02% NaN_3, pH 6.0]. PMNL-CL catalytic domain was crystallized by vapour diffusion against 1 M potassium phosphate buffer, 0.02% NaN_3, pH 6.0. Small crystals of ~0.5 × 0.06 × 0.03 mm were obtained within 3 days. They were harvested in 20% (m/v) PEG 6000, 500 mM NaCl, 100 mM $CaCl_2$, 0.2 M MES−NaOH, 0.02% NaN_3, pH 6.0. The crystals diffract to at least 2.0 Å resolution, are orthorhombic and belong to space group $P2_12_12_1$ with lattice constants $a = 33.09$, $b = 69.37$, $c = 72.48$ Å, $\alpha = \beta = \gamma = 90°$, with the asymmetric unit containing a monomer.

Structure analysis

A low resolution native data set and derivative data sets were collected on a computer-controlled CAD-4 diffractometer equipped with a FAST area television detector (Enraf-Nonius, Delft) mounted on a Rigaku rotating anode generator operated at 5.4 kW ($\lambda = Cu_{K\alpha} = 1.5418$ Å). X-ray intensities were evaluated with the MADNES system (Messerschmidt and Pflugrath, 1987) and scaled, corrected for absorption effects and averaged (Messerschmidt *et al.*, 1990).

Table III. Current refinement parameters

Resolution range	8.0−2.03
No. of unique reflections	9600
Protein atoms (excluding H)	1266
Solvent atoms (excluding H)	111
R-factor[a]	0.182
R.m.s. deviations from target values	
bonds (Å)	0.012
angles (°)	1.8

[a]$R = (\Sigma|F_o-F_c|)/\Sigma F_o$.

X-ray measurements for the high resolution native data set were done on a MAR image plate area detector (MAR Research, Hamburg) also mounted on a Rigaku rotating anode X-ray generator operated at 5.4 kW. X-ray intensities were evaluated with the MOSFLM program package (Leslie, 1991). All X-ray data were loaded to the PROTEIN program package (Steigemann, 1991); data collection statistics are given in Table I.

Heavy atom derivatives were prepared by soaking under the conditions given in Table I and were analysed by difference Patterson methods using PROTEIN. Heavy atom sites (see Table II) and the phases were refined using MIRPH (J.Remington) as implemented in PROTEIN. A first 3 Å MIR Fourier map calculated for the native enzyme (i.e. based on phase set b in Table II) was very noisy but showed some secondary structure elements. A significantly improved electron density map was obtained by treating the zinc-apoenzyme ('Apo' in Table I) as the native and the

zinc-holoenzyme ('Apo + Zn' in Table I, the real native enzyme) as a 'heavy-atom' derivative (i.e. a Fourier calculated for the apoenzyme based on a phase set a in Table II). An apocatalytic domain starting model of human PMNL-CL (i.e. lacking the 'catalytic' zinc ion Zn999) was built using the interactive graphics program FRODO (Jones, 1978) implemented on an ESV-30 graphics workstation (Evans and Sutherland). The model was subsequently completed and refined in 10 rounds using X-PLOR (Brünger *et al.*, 1989). After each round a phase-combined Fourier map was calculated using phase set a (in Table II) and 'Apo' (in Table I) as native compound (up to round 5) or using phase set b (in Table II) and 'Apo + Zn' (in Table I) as the native compound, respectively. Starting with round 9 the high resolution native data set was used together with pure model phases. The current model, on which interpretation is based, has a crystallographic *R*-factor of 18.2% including all data (applying no systematic σ-cut-off) from 8.0 to 2.03 Å resolution and comprises residues 86–242 and 111 ordered solvent molecules; the current refinement parameters are given Table III. After completion of refinement the coordinates will be submitted to the PDB data base.

Acknowledgements

We thank P.Widawka and I.Mayr-Kröner for excellent help in cloning human PMNL-CL catalytic domain cDNA and crystallization, respectively. We are grateful to Dr M.T.Stubbs and Dr F.-X.Gomis-Rüth for helpful discussions. The financial support of the SFB 207 of the Universität München and of the Fonds der Chemischen Industrie to W.B., of the BAYER AG (PF-F/Biotechnology, Monheim, Germany) to P.R. and of the SFB 223 of the Universität Bielefeld and of the Fonds der Chemischen Industrie to H.T. is gratefully acknowledged.

References

Baumann,U., Wu,S., Flaherty,K.M. and MacKay,D.B. (1993) *EMBO J.*, **12**, 3357–3364.

Birkedal-Hansen,H., Moore,W.G.I., Bodden,M.K., Windsor,L.J., Birkedal-Hansen,B., DeCarlo,A. and Engler,J.A. (1993) *Crit. Rev. Oral Biol. Med.*, **4**, 197–250.

Bläser,J., Knäuper,V., Osthues,A., Reinke,H. and Tschesche,H. (1991) *Eur. J. Biochem.*, **202**, 1223–1230.

Bode,W., Gomis-Rüth,F.-X., Huber,R., Zwilling,R. and Stöcker,W. (1992) *Nature*, **358**, 164–167.

Bode,W., Gomis-Rüth,F.-X. and Stöcker,W. (1993) *FEBS Lett.*, **331**, 134–140.

Brünger,A.T., Karplus,M. and Petsko,G.A. (1989) *Acta Crystallogr. Sect. A*, **45**, 50–61.

Fraser,R.D.B., MacRae,T.P. and Suzuki,E. (1979) *J. Mol. Biol.*, **129**, 463–481.

Gomis-Rüth,F.-X., Stöcker,W., Huber,R., Zwilling,R. and Bode,W. (1993a) *J. Mol. Biol.*, **229**, 945–968.

Gomis-Rüth,F.-X., Kress,L.F. and Bode,W. (1993b) *EMBO J.*, **12**, 4151–4157.

Grant,G.A., Eisen,A.Z., Marmer,B.L., Roswit,W.T. and Goldberg,G.I. (1987) *J. Biol. Chem.*, **262**, 5886–5889.

Hasty,K.A., Hibbs,M.S., Kang,A.H. and Mainardi,C.L. (1986) *J. Biol. Chem.*, **261**, 5645–5650.

Hasty,K.A., Jeffrey,J.J., Hibbs,M.S. and Welgus,H.G. (1987) *J. Biol. Chem.*, **262**, 10048–10052.

Hasty,K.A., Pourmotabbed,T.F., Goldberg,G.I., Thompson,J.P., Spinella,D.G., Stevens,R.M. and Mainardi,C.L. (1990) *J. Biol. Chem.*, **265**, 11421–11424.

Hirose,T., Patterson,C., Pourrmotabbed,T., Mainardi,C.L. and Hasty,K.A. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 2569–2573.

Holmes,M.A. and Matthews,B.W. (1981) *Biochemistry*, **20**, 6912–6920.

Jones,T.A. (1978) *J. Appl. Crystallogr.*, **15**, 23–31.

Kabsch,W. and Sander,C. (1983) *Biopolymers*, **22**, 2577–2637.

Knäuper,V., Krämer,S., Reinke,H. and Tschesche,H. (1990) *Eur. J. Biochem.*, **189**, 295–300.

Knäuper,V., Osthues,A., DeClerk,Y.A., Langley,K.A., Bläser,J. and Tschesche,H. (1993a) *Biochem. J.*, **291**, 847–854.

Knäuper,V., Wilhelm,S.M., Seperack,P.K., DeClervk,Y.A., Langley,K.E., Osthues,A. and Tschesche,H. (1993b) *Biochem. J.*, **295**, 581–586.

Leslie,A.G.W. (1991) *Daresbury Lab. Inf. Q. Prog. Crystallogr.*, **26** (available from the Librarian, SERC Laboratory, Daresbury, Warrington, WA4 4AD, UK).

Mallya,S.K., Mookhtiar,K.A., Gao,Y., Brew,K., Dioszegi,M., Birkedahl-Hansen,H. and van Wart,H.E. (1990) *Biochemistry*, **29**, 10628–10634.

Matrisian,L.M. (1992) *BioEssays*, **14**, 455–463.

Messerschmidt,A. and Pflugrath,J.W. (1987) *J. Appl. Crystallogr.*, **20**, 306–315.

Messerschmidt,A., Schneider,M. and Huber,R. (1990) *J. Appl. Crystallogr.*, **23**, 436–439.

Moore,W.M. and Spilburg,C.A. (1986) *Biochemistry*, **25**, 5189–5195.

Murphy,G. and Docherty,A.J.P. (1992) *Am. J. Resp. Cell Mol. Biol.*, **7**, 120–125.

Murphy,G., Reynolds,J.J., Bretz,U. and Baggiolini,M. (1977) *Biochem. J.*, **162**, 195–197.

Murphy,G., Allan,J.A., Willenbrock,F., Cockett,M.I., O'Connel,J.P. and Docherty,A.J.P. (1992) *J. Biol. Chem.*, **267**, 9612–9618.

Nagase,H., Enghild,J.J., Suzuki,K. and Salveson,G. (1990) *Biochemistry*, **29**, 5783–5789.

Netzel-Arnett,S., Fields,G.B., Birkedahl-Hansen,H. and van Wart,H.E. (1991) *J. Biol. Chem.*, **266**, 6747–6755.

Odake,S., Okayama,T., Obata,M., Morikawa,T., Hattori,S., Hori,H. and Nagai,Y. (1991) *Chem. Pharm. Bull.*, **39**, 1489–1494.

Priestle,J.P. (1988) *J. Appl. Crystallogr.*, **21**, 572–576.

Reinemer,P., Grams,F., Huber,R., Kleine,T., Schnierer,S., Piper,M., Tschesche,H. and Bode,W. (1993) *FEBS Lett.*, **338**, 227–233.

Rossmann,M.G. and Argos,P. (1975) *J. Biol. Chem.*, **250**, 7525–7532.

Salowe,S.P., Marcy,A.I., Cuca,G.C., Smith,C.K., Kopka,I.E., Hagmann,W.K. and Hermes,J.D. (1993) *Biochemistry*, **31**, 4535–4550.

Sanchez-Lopez,R., Alexander,C.M., Behrendtsen,O., Breathanch,R. and Werb,Z. (1993) *J. Biol. Chem.*, **268**, 7238–7247.

Schnierer,S., Kleine,T., Gote,T., Hillemann,A., Knäuper,V. and Tschesche,H. (1993) *Biochem. Biophys. Res. Commun.*, **191**, 319–326.

Schechter,I. and Berger,A. (1967) *Biochem. Biophys. Res. Commun.*, **27**, 157–162.

Steigemann,W. (1991) In Moras,D., Podjarny,A.D. and Thierry,J.C. (eds), *From Chemistry to Biology*. Vol. 5, *Crystallographic Computing*, Oxford University Press, Oxford, UK, pp. 115–125.

Springman,E.B., Angleton,E.L., Birkedahl-Hansen,H. and van Wart,H.E. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 364–368.

Suzuki,K., Engfield,J.J., Morodomi,T., Salveson,G. and Nagase,H. (1990) *Biochemistry*, **29**, 10261–10270.

Turk,D. (1992) Ph.D. Thesis, Technische Universität München.

Wang,D., Bode,W. and Huber,R. (1985) *J. Mol. Biol.*, **185**, 595–624.

Wilson,A.J.C. (1942) *Nature*, **150**, 151–152.

Woesner,J.F. (1991) *FASEB J.*, **5**, 2145–2154.

Wu,H., Byrne,M.H., Stacey,A., Goldring,M.B., Birkhead,J.R., Jaenisch,R. and Krane,S.M. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 5888–5892.

Yonath,A. and Traub,W. (1969) *J. Mol. Biol.*, **43**, 461–477.

Received on December 8, 1993; revised on December 27, 1993