Classification: BIOLOGICAL SCIENCES: Biochemistry

Structural insights into triple-helical collagen cleavage by matrix metalloproteinase 1

Szymon W. Manka[a], Federico Carafoli[b], Robert Visse[a], Dominique Bihan[c], Nicolas Raynal[c], Richard W. Farndale[c], Gillian Murphy[d], Jan J. Enghild[e], Erhard Hohenester[b1] and Hideaki Nagase[a1]

[a]Kennedy Institute of Rheumatology, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, London, W6 8LH, United Kingdom; [b]Department of Life Sciences, Imperial College London, London, SW7 2AZ, United Kingdom; [c]Department of Biochemistry, University of Cambridge, Downing Site, Cambridge, CB2 1QW United Kingdom; [d]Department of Oncology, University of Cambridge, Cancer Research UK, Cambridge Institute, Li Ka Shing Centre, Cambridge, CB2 0RE, United Kingdom; and [e]Center for Insoluble Protein Structure and Interdisciplinary Nanoscience Center at Department of Molecular Biology, University of Aarhus, Gustav Wieds Vej 10C, Aarhus, DK-8000, Denmark

[1]To whom correspondence should be addressed: e.hohenester@imperial.ac.uk; hideaki.nagase@kennedy.ox.ac.uk

Corresponding author: Hideaki Nagase, Kennedy Institute of Rheumatology, University of Oxford, 65 Aspenlea Road, London, W6 8LH, UK; Tel, 44-20-8383-4488; e-mail, hideaki.nagase@kennedy.ox.ac.uk

Abstract

Collagenases of the matrix metalloproteinase (MMP) family play major roles in morphogenesis, tissue repair and human diseases, but how they interact and cleave the collagen triple helix is not fully understood. Here we report that human MMP-1 (collagenase 1) binds to collagen through the cooperative action of the catalytic and hemopexin domains in a temperature-dependent manner. The contact sites of the two molecules were mapped by screening the Collagen Toolkit peptide library and by hydrogen/deuterium exchange. The crystal structure of MMP-1 bound to a triple-helical collagen peptide revealed that all three collagen chains interact extensively with the two domains of MMP-1. An exosite in the hemopexin domain, which binds the leucine ten residues C-terminal to the scissile bond, is critical for collagenolysis and represents a novel target for inhibitor development. Modeling shows that a localized excursion of the collagen chain close to the active site, facilitated by thermal loosening and an inter-domain flexing of the enzyme, leads to the first transition state of collagenolysis.

Introduction

The interstitial collagens I, II and III are the major structural proteins in connective tissues such as skin, bone, cartilage, tendon and blood vessels (1). They consist of three α chains with repeating Gly-X-Y triplets (X and Y are often proline and hydroxyproline, respectively) that intertwine each other to form a triple helix of ~300 nm length (2). Interstitial collagens are resistant to most proteolytic enzymes, but vertebrate collagenases cleave them at a single site approximately ¾ of the way from the N-terminus of the triple helix, thus initiating collagenolysis (3). Owing to this unique activity, collagenases play important roles in embryo development, morphogenesis, tissue remodeling, wound healing, and human diseases, such as arthritis, cancer and atherosclerosis (4, 5).

Matrix metalloproteinase 1 (MMP-1) is a typical vertebrate collagenase (3). It consists of an N-terminal catalytic (Cat) domain containing an active-site zinc ion and a C-terminal hemopexin (Hpx) domain comprised of a four-bladed β-propeller, which are connected by a linker region (6, 7). While the Cat domain can cleave a number of non-collagenous proteins including heat-denatured collagen (gelatin), its activity on native triple-helical collagen is negligible. The combination of the Cat and Hpx domains is required for MMP-1 to be able to degrade native collagen, and this is true for all other collagenolytic MMPs, namely MMP-2, MMP-8, MMP-13 and MMP-14 (3). How collagenases interact with collagen and how the Hpx domain endows these enzymes with collagenolytic activity is not clearly understood. Another enigma of collagenolysis became apparent when the crystal structures of MMP-1Cat (8-10), MMP-8Cat (11, 12), and full-length pig MMP-1 (6) were solved: The catalytic cleft of these enzymes is too narrow to accommodate collagen in its native triple-helical conformation. A number of hypotheses have been proposed to explain how collagenases may destabilize and cleave triple-helical collagen (13-16), and we have experimentally demonstrated that MMP-1 unwinds triple-helical collagen locally prior to peptide bond hydrolysis (17, 18). A recent study mapped the interaction sites in MMP-1 and in triple-helical collagen using NMR and proposed a model of collagen binding and unwinding (19). However, because the interactions between MMP-1 and collagen were not observed directly, a number of assumptions had to be made to derive a mechanism of collagenolysis.

In this study, we have combined biochemical experiments with crystallographic structure determination to show how MMP-1 recognizes its unique cleavage site in interstitial

collagens. Our results suggest that collagenolysis relies on multiple exosite interactions, which not only serve to position the scissile bond near the active site, but also assist in the local unfolding of the collagen triple helix that is required for cleavage to occur.

Results

The Cat and Hpx domains of MMP-1 cooperate in collagen binding. We first examined a catalytically inactive MMP-1(E200A) mutant (residue numbering based on the proMMP-1 sequence) and its individual domains for binding to native collagen I. Full-length MMP-1(E200A) showed saturable collagen binding with an apparent K_D of 0.4 μM, whereas the catalytic domain Cat(E200A) showed no detectable binding and the Hpx domain bound only weakly (Fig. 1*A*). The Cat(E200A) domain did not bind to collagen I even when it was added together with increasing concentrations of the Hpx domain (Fig. S1*A*), suggesting that the Cat and Hpx domains bind collagen cooperatively only when they are linked together.

Fig. 1 near here

The binding of MMP-1(E200A) to collagen increased with temperature until collagen denatured above 37 °C (Fig. 1*B*), indicating that MMP-1 prefers a looser triple helix, but not denatured collagen. The temperature-dependent enhancement of MMP-1(E200A) binding to collagen was substantially reduced by the active-site inhibitor GM6001 (Fig. 1*C*). This effect was not observed below 10 °C, where little collagen cleavage occurs (17). No significant temperature-dependent increase in collagen binding was observed with either the Cat(E200A) or Hpx domain alone (Fig. S1*B* and *C*). These results suggest that the active site cleft of MMP-1 participates in collagen binding at body temperature by stabilizing and/or inducing a looser collagen conformation.

The MMP-1-binding motif in interstitial collagens. To identify the collagen residues interacting with MMP-1, we screened the Collagen Toolkit library of triple-helical peptides encompassing the entire collagen II sequence (20) for binding to MMP-1(E200A). Only one peptide, Col II-44, starting at the collagenase cleavage site G~LAGQR… (~ indicates the scissile bond) was recognized by the enzyme (Fig. 2). Col II-43, which also contains the cleavage site

but with fewer C-terminal residues, did not bind MMP-1(E200A). Further experiments with truncated and alanine-substituted peptides identified the leucines at the P1' and P10' positions (nomenclature after Schechter and Berger (21)) of collagen as important for MMP-1(E200A) binding (Fig. S2*A*). A sequence alignment showed that the leucines at these two positions are conserved in all α chains of collagens I, II and III from different species (Fig. S2*B*). The consensus sequence for collagenase binding is G~(**L/I**)(A/L)-GXY-GXY-G**L**(O/A) (P1' and P10' positions are bold; O, hydroxyproline). This motif is unique and not found elsewhere in collagens I, II and III.

Fig. 2 near here

Mapping of the collagen-binding regions of MMP-1 in solution. We then used hydrogen/deuterium exchange mass spectrometry (H/DXMS) to map the regions in MMP-1 involved in collagen binding. Five MMP-1 regions showed delayed deuterium incorporation in the presence of collagen I (Fig. 3*A*, Fig. S3 and Table S1): Residues 145-161 (site 1) and 205-215 (site 3) in the Cat domain, and residues 266-277 (site 4), 283-297 (site 5) and 330-346 (site 6) in the Hpx domain. Sites 4 and 5 were previously reported to reduce deuterium exchange upon binding to a triple-helical peptide (22). Residues 179-193 (site 2) showed increased deuterium incorporation in the presence of collagen, suggesting that this region, which includes the loop between the fifth β-strand and the second α-ahelix of the Cat domain, is more exposed to the solvent in the collagen-bound state. This loop region was previously characterized as important for the collagenolytic activity of MMP-1 (23), and our data suggest that it participates in collagenolysis dynamically. The combined H/DXMS results indicate the general orientation of the collagen triple helix binding to MMP-1 (Fig. 3, dashed line).

Fig. 3 near here

The S10' exosite in the Hpx domain. Considering the H/DXMS results and the importance for MMP-1(E200A) binding of the two leucines at the P1' and P10' positions of collagen, we searched for hydrophobic sites in MMP-1 that could accept these residues. For MMP-1 to cleave the Gly(P1)-Leu(P1') bond, the P1' Leu must be located close to the S1' pocket of the Cat

domain. Using this constraint, we docked a modeled triple-helical peptide to MMP-1 and identified candidate residues in the Hpx domain that might be involved in binding the P10' Leu (Fig. 3*A*). They were mutated and the mutant proteins tested for their collagenolytic activity (Fig. 3*B*). The most dramatic reduction of collagenase activity was observed with the F301Y single mutant and the I271A/R272A double mutant, which exhibited 10 % and 4 % of wild-type activity, respectively. Together, Phe301, Ile271 and Arg272 in the Hpx domain participate in forming the hydrophobic S10' binding pocket, or exosite. Ile271 and Arg272 were identified in the H/DXMS experiment as part of site 4. Phe301 is located between H/DXMS sites 5 and 6. It was not detected by H/DXMS analysis most likely because its backbone amide is buried in the Hpx domain.

Crystal structure of an MMP-1(E200A)-collagen peptide complex. Based on the results of the collagen peptide binding studies, we designed a triple-helical collagen peptide for co-crystallization with MMP-1(E200A). The peptide spans positions P7-P17' of the collagen II cleavage site and is capped by two and three non-native Gly-Pro-Hyp triplets at the N- and C-termini, respectively, to enhance triple-helicity (Fig. 4*A*). The melting temperature of this peptide was 31 °C (Fig. S4). The peptide formed a stable complex with MMP-1(E200A) in solution (Fig. S5). Crystals of the complex were grown at room temperature and used to determine the structure at 3.0 Å resolution (Table S2).

Fig. 4 near here

The overall structure of human MMP-1(E200A) in complex with collagen is similar to those reported for full-length pig (6) and human (7) MMP-1 without collagen. The collagen peptide in the complex with MMP-1(E200A) is an uninterrupted triple helix ~115 Å in length and bent by ~10° near the Cat-Hpx junction (Fig. 4*B* and Fig. S6*A*). The three chains of the collagen triple helix are arranged with the characteristic one residue stagger, resulting in a leading (L), middle (M) and trailing (T) chain (24). All three chains contribute to MMP-1(E200A) binding, creating an extensive interaction surface spanning nearly 60 Å in length. Complex formation buries 480 Å^2 of solvent-accessible surface of the L chain, 440 Å^2 of the M chain, and 370 Å^2 of the T chain.

Interactions between MMP-1(E200A) and the collagen peptide. The majority of the interactions between the collagen peptide and the Cat domain are of a polar nature (Fig. 4*C*, Fig. S6*B* and Fig. S7*B*). Of the three collagen chains, the L chain is closest to the active site of MMP-1, but the S1' pocket remains empty. The residue closest to the S1' pocket is Gln(P4'$_L$) (residues are identified by their position relative to the scissile bond and a subscript designating the collagen chain). The side chain of Gln(P4'$_L$) forms a hydrogen bond with the backbone amide of Tyr221 at the entrance of the S1' pocket, and the backbone amide of Gln(P4'$_L$) forms a hydrogen bond with the side chain of Asn161. Leu(P1'$_L$) is 9 Å away from the S1' pocket and stabilized by a hydrophobic contact with the zinc ligand His203. The Arg(P5'$_L$) side chain and backbone amide form hydrogen bonds with the side chain of Tyr218 and the carbonyl oxygen of Pro219, respectively. The M chain makes only a single contact with the Cat domain, between the side chain of Gln(P4'$_M$) and the side chain of Tyr218. The T chain, on the other hand, makes extensive interactions with the upper part of the substrate binding cleft, including five hydrogen bonds (Fig. 4*C*, Fig. S6*B* and Fig. S7*B*). Some of these hydrogen bonds are identical to those formed with a cysteine-switch sequence (25-28) in proMMPs and with a predicted peptide substrate (29).

The interactions between the collagen peptide and the Hpx domain are a combination of polar and apolar contacts, and are mediated mainly by the M chain (Fig. 4*D*, Fig. S6*B* and Fig. S7*B*). Consistent with our mutagenesis results, the side chain of Leu(P10'$_M$) is bound in a hydrophobic S10' exosite pocket delineated by Ile271, Met276, Phe301, Trp302 and the alkyl portion of Arg272 (Fig. 4*D* and Fig. S6*B*). The S10' pocket is surrounded by polar residues (Arg272, Glu274, Arg285 and Gln335) that make a total of five hydrogen bonds with the collagen peptide (Fig. S7*B*). The interactions at the S10' exosite are critical for collagen recognition and collagenolysis by MMP-1 (Fig. 3*B*). In addition, Ile(P7'$_M$) forms hydrophobic interactions with Val300 and Phe301. The M and T chains make van der Waals contacts with Phe289 and Tyr290, but these interactions evidently are not critical, since the triple mutant MMP-1(F289A/Y290A/P291A) retains 76% of wild-type collagenase activity (Fig. 3*B*).

Fig. 5 near here

Implications for collagenolysis. The way in which the collagen peptide is bound to the active-site cleft of MMP-1(E200A) is more compatible with cleavage of the Gly(P3'$_L$)-Gln(P4'$_L$) bond instead of the canonical Gly(P1)-Leu(P1') bond (30). To test whether MMP-1 can cleave also before Gln(P4'), we performed N-terminal sequencing of the products generated by active MMP-1 at 25 °C and 37 °C. We found that only the Gly(P1)-Leu(P1') bond was cleaved in the peptide, as in native collagen II, and concluded that the mode of collagen binding in the crystallized complex is unproductive. However, the helical symmetry of homotrimeric collagen allows the productive binding mode to be readily achieved after axial rotation of the peptide and it differs from the observed complex only in the register of the chain passing through the active site cleft (Fig. 5 and Fig. S7). If Leu(P10'$_L$) instead of Leu(P10'$_M$) is made to occupy the hydrophobic S10' exosite, the T chain becomes the strand that interacts most closely with the active site, and now Leu(P1'$_T$) is situated near the S1' pocket, i.e. the register of the chain in the active site cleft is changed by one triplet. The third possible arrangement with Leu(P10'$_T$) in the S10' exosite is also unproductive and not further considered. Notably, the productive chain arrangement with Leu(P1'$_T$) near the S1' pocket does not alter any of the extensive interactions with the upper part of the substrate binding cleft (now made by the M chain) or with the S10' exosite in the Hpx domain (now made by the L chain) (Fig. 5 and Fig. S7). By contrast, the T chain loses several contacts around the active site that were observed with the L chain in the crystal structure, in particular the hydrogen bonds between the Gln(P4'$_L$) side chain and the Tyr221 backbone amide and between the Arg(P5'$_L$) side chain and the Tyr218 side chain (Fig. S7). The fewer contacts of the T chain may make it more prone to local unfolding and insertion of Leu(P1'$_T$) into the S1' pocket. The energy difference between the productive and unproductive complex is likely to be small and we think that the two forms co-exist in solution. The unproductive binding mode may be favored by the E200A mutant of MMP-1, or the unproductive complex may have crystallized preferentially in our experiment.

Starting from the modeled productive complex (Fig. S7), it proved straightforward to achieve the first transition state of collagenolysis, in which one collagen chain is inserted into the active site (Fig. 6). By stretching the P1'-P6' region of the T chain and reorienting three residues preceding the scissile bond, it was possible to insert Leu(P1'$_T$) into the S1' pocket in a manner that fully agrees with the observed binding mode of single-chain peptidomimetic MMP inhibitors (8, 31). The Cα atom of Leu(P1'$_T$) moves by less than 4 Å during this localized

excursion from the regular triple-helical structure, and only four collagen inter-chain hydrogen bonds are broken in the transition state. We propose that the energy penalty for this local unfolding event is compensated by the favorable interactions that the enzyme forms with the looped-out collagen chain.

Fig. 6 near here

Discussion

Our study has revealed an unprecedented proteinase-substrate interaction. The most striking feature of the MMP-1-collagen complex is the extensive interaction of all three collagen chains with the two domains of MMP-1, which explains the strict requirement of the Hpx domain for efficient collagenolysis (3). The individual Cat or Hpx domains of MMP-1 showed very little collagen binding in our experiments, but when linked together, they acted cooperatively in a temperature-dependent manner. The affinity for collagen I increased up to 37 °C, a temperature that would not cause collagen to denature within the time frame of the experiment (~2 h) (32). Above this temperature the binding sharply decreased. This temperature-dependence, which is also unusually pronounced in the rate of collagen cleavage by MMP-1 (33), can be explained by the relative thermoflexibility of the triple helix at the collagenase cleavage site (34-37) and the contribution of inter-domain flexibility of MMP-1. The cooperative binding is therefore energetically most favored close to body temperature, when the triple helix is locally looser but the overall fold of the substrate is still maintained. A complete loss of triple-helicity makes collagen bind less tightly to MMP-1, as evidenced from a K_m value of ~1 μM for collagen I compared with 4-7 μM for gelatin (38). The small active-site inhibitor GM6001 abolished the temperature-dependent increase of collagen binding, indicating that the catalytic site of MMP-1 is not only involved in the final hydrolysis step, but also in the unfolding machinery, which uses the active site to extract one collagen chain to form the first transition state of collagenolysis (Fig. 6).

Apart from providing a favorable environment for the loose triple helix at the cleavage site, what other mechanism(s) might contribute to collagenolysis? Comparison of the crystal structure of proMMP-1 (28) with the structure of the activated form of MMP-1(E200A) (6, 7) suggested that the Cat-Hpx linker region confers inter-domain flexibility. Such flexing was further demonstrated for MMP-1 by NMR and small angle X-ray scattering studies (39, 40),

and seems to be a common property of MMPs (41, 42). Our crystal structure of MMP-1 bound to collagen has a more closed conformation than free MMP-1 (Fig. S8). This change in inter-domain angle is accompanied by the bending of the collagen peptide at a point half-way between the Cat and Hpx domains. Thus, domain motions in MMP-1 may facilitate structural changes in the triple helix. Our structural results show how the two domains of MMP-1 anchor two of the collagen chains (middle and leading in the productive complex) and position the third (trailing) chain near the active site. The paucity of interactions made by the trailing chain, combined with its stretching by inter-domain flexing of the enzyme, may assist in local collagen unfolding. Because the P1' residue of the trailing chain is poised above the S1' pocket in the productive complex with MMP-1, transient excursions of the scissile bond can be efficiently captured by the enzyme, leading to the stabilization of the first transition state of collagenolysis (Fig. 6). While thermal relaxation of the cleavage site is important for MMP-1 binding, this alone is not sufficient for efficient collagenolysis (17, 36). We propose that mammalian collagenases do not serve merely as passive acceptors of spontaneously occurring unfolded collagen states (43, 44), but promote a local perturbation of the triple helix that facilitates collagenolysis. The recently reported NMR study support this notion by demonstrating a weakening of the collagen interchain contacts in the complex with MMP-1 (19).

Our mechanism of collagenolysis (Fig. 6) differs from that derived from NMR experiments (19) in that it does not require a separation of the Cat and Hpx domains. Bertini et al. (19), who also used the E200A mutant of MMP-1, found that the Cat domain interacted mainly with the L chain, and the Hpx domain mainly with the L and M chains. These results agree with our crystal structure of an unproductive complex. In their modeling, Bertini et al. (19) imposed a productive interaction of the L chain with the S1' pocket, which forced the breaking of the Cat-Hpx interface. Our interpretation requires only small conformational changes in the enzyme, but postulates the co-existence of productive and unproductive binding modes.

The crystal structure of a bacterial collagenase comprising of a N-terminal activator domain, a glycine-rich linker and a peptidase domain has been reported recently (45). Like MMPs, this bacterial collagenase requires an ancillary activator domain (unrelated to the Hpx domain of MMPs) to cleave triple-helical collagen. The authors proposed that the enzyme clamps collagen between the activator and peptidase domains of the saddle-shaped structure.

Our MMP-collagen structure shows that MMP-1 binds collagen in an extended manner, rather than surrounding it. These differences notwithstanding, inter-domain plasticity seems to be a common feature of enzymes that can cleave the intractable collagen triple helix.

Most MMPs are believed to have specific substrate recognition sites distant from the catalytic site, which are referred to as exosites. Apart from the Hpx domains of collagenases, examples include the fibronectin type II repeats in MMP-2 and MMP-9, which contribute to the binding and degradation of elastin, gelatin and type IV collagen (46). Recent NMR and mutagenesis studies of the Cat domain of MMP-12 characterized several exosites for elastin (47, 48). Similar studies were carried out with the isolated Hpx domain of MMP-1 and a triple-helical collagen peptide (40), and suggested a role for Phe282 (Phe301 in the numbering used in ref. 40) in binding of collagen. However, the exact modes of interaction between the enzyme and the target substrate have not been elucidated in these studies. We show that Phe282 remains buried in the Cat-Hpx interface in the collagen–MMP-1 complex. Our crystal structure has revealed in detail the interactions that occur between collagen and full-length MMP-1, using the exosites in both the Cat and Hpx domains and a triple-helical peptide long enough to define all relevant contacts. Defining the exosites in proteolytic enzymes has important implications for the development of specific inhibitors. The hydrophobic S10' exosite in MMP-1 is a particularly attractive target as it is essential for collagenolysis. Recent studies of Robichaud et al. (49) showed the importance of Leu at the P10’ subsite of collagen III for the cleavage by MMP-1. Non-collagenous extracellular matrix and non-matrix substrates are cleaved by the Cat domain of MMP-1 (16). Therefore, typical inhibitors targeting the S1' pocket block all MMP-1 activities, whereas inhibitors directed against the S10' exosite are predicted to selectively inhibit collagenolysis. This exosite of MMP-1 is largely conserved in other MMPs, but to our knowledge it is specific for collagenolytic activity. We propose that molecules specifically binding the S10' exosite be therapeutically useful in diseases associated with accelerated collagen breakdown such as arthritis, cancer, aneurysm and atherosclerosis.

Materials and Methods

Proteins and Peptides, Binding Assays, H/DXMS Analysis and Enzyme Activity Assay. Recombinant proMMP-1(E200A), its Cat(E200A) and Hpx domains and MMP-1 variants were produced in *E. coli*. Collagen I was purified from Guinea pig skin and treated with pepsin.

The library of triple-helical peptides of human collagen II (Toolkit II) and all other collagen peptides were synthesized by Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry. Collagen and peptide binding assays were carried out in ELISA format using biotinylated proteins and detected by streptavidin-horseradish peroxidase system. Local hydrogen/deuterium exchange kinetics of MMP-1(E200A) free and bound to collagen I were analyzed by digestion with pepsin and mass spectrometry and quantified using HX-Express software. The collagenolytic activity of MMP-1 variants was determined by SDS-PAGE and densitometric analyses of collagen I cleavage products. Detailed methods are described in *SI Materials and Methods*.

Crystallization, Data Collection and Structure Determination. MMP-1(E200A) and the collagen peptide were mixed in a 1:1.5 molar ratio and the complex purified by size exclusion chromatography. Crystals of the complex were obtained at pH 8.5 and diffracted to 3.0 Å resolution. The structure was solved by molecular replacement and refined to a free R-factor of 0.273. Detailed methods are described in *SI Materials and Methods.* Data processing and refinement statistics are listed in Table S2.

ACKNOWLEDGMENTS. We thank Noriko Ito for construction of MMP-1 mutants, Ida B. Thøgersen for peptide sequencing, the staff at Diamond beamline I24 for help with X-ray data collection, and Peter Brick for many helpful discussions. Protein visualization and modeling was done in PyMOL, W.L. DeLano, DeLano Scientific, San Carlos, Ca, USA (http:/www.pymol.org). The work was supported by grants from Wellcome Trust (refs 057473, 068724, 094470 and 083942), Medical Research Council, Arthritis Research UK, and Cancer Research UK. E.H. is a Wellcome Trust Senior Research Fellow.

Author contributions: S.W.M., R.W.F., G.M., J.J.E., E.H., and H.N. designed the research; S.W.M, F.C., R.V., D.B., N.R., and E.H. performed experiments; and S.W.M., E.H., and H.N wrote the paper.

The authors declare no conflict of interest

References

1. Kadler KE, Baldock C, Bella J, Boot-Handford RP (2007) Collagens at a glance. *J Cell Sci* 120:1955-1958.
2. Brodsky B, Persikov AV (2005) Molecular structure of the collagen triple helix. *Adv Protein Chem* 70:301-339.
3. Nagase H, Visse R (2011) Triple helicase activity and the structural basis of collagenolysis in Extracellular Matrix Degradation, Biology of Extracellular Matrix, eds Parks WC, Mecham RP (Springer-Verlag, Heidelberg), pp 95-122.
4. Brinckerhoff CE, Matrisian LM (2002) Matrix metalloproteinases: a tail of a frog that became a prince. *Nat Rev Mol Cell Biol* 3:207-214.
5. Page-McCaw A, Ewald AJ, Werb Z (2007) Matrix metalloproteinases and the regulation of tissue remodelling. *Nat Rev Mol Cell Biol* 8:221-233.
6. Li J, et al. (1995) Structure of full-length porcine synovial collagenase reveals a C-terminal domain containing a calcium-linked, four-bladed beta-propeller. *Structure* 3:541-549.
7. Iyer S, Visse R, Nagase H, Acharya KR (2006) Crystal Structure of an Active Form of Human MMP-1. *J Mol Biol* 362:78-88
8. Borkakoti N, et al. (1994) Structure of the catalytic domain of human fibroblast collagenase complexed with an inhibitor. *Nature Struct Biol* 1:106-110.
9. Lovejoy B, et al. (1994) Structure of the catalytic domain of fibroblast collagenase complexed with an inhibitor. *Science* 263:375-377.
10. Spurlino JC, et al. (1994) 1.56 Å structure of mature truncated human fibroblast collagenase. *Proteins* 19:98-109.
11. Bode W, et al. (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. *EMBO J* 13:1263-1269.
12. Stams T, et al. (1994) Structure of human neutrophil collagenase reveals large S1' specificity pocket. *Nature Struct Biol* 1:119-123.
13. Bode W (1995) A helping hand for collagenases: the haemopexin-like domain. *Structure* 3:527-530.
14. de Souza SJ, Pereira HM, Jacchieri S, Brentani RR (1996) Collagen/collagenase interaction: does the enzyme mimic the conformation of its own substrate? *FASEB J* 10:927-930.
15. Ottl J, et al. (2000) Recognition and catabolism of synthetic heterotrimeric collagen peptides by matrix metalloproteinases. *Chem Biol* 7:119-132.
16. Overall CM (2002) Molecular determinants of metalloproteinase substrate specificity: matrix metalloproteinase substrate binding domains, modules, and exosites. *Mol Biotechnol* 22:51-86.
17. Chung L, et al. (2004) Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. *EMBO J* 23:3020-3030.
18. Han S, et al. (2010) Molecular mechanism of type I collagen homotrimer resistance to mammalian collagenases. *J Biol Chem* 285:22276-22281.
19. Bertini I, et al. (2012) Structural basis for matrix metalloproteinase 1-catalyzed collagenolysis. *J Am Chem Soc* 134:2100-2110.
20. Farndale RW, et al. (2008) Cell-collagen interactions: the use of peptide Toolkits to investigate collagen-receptor interactions. *Biochem Soc Trans* 36:241-250.

21. Schechter I, Berger A (1968) On the active site of proteases. 3. Mapping the active site of papain; specific peptide inhibitors of papain. *Biochem Biophys Res Commun* 32:898-902.
22. Lauer-Fields JL, et al. (2009) Identification of specific hemopexin-like domain residues that facilitate matrix metalloproteinase collagenolytic activity. *J Biol Chem* 284:24017-24024.
23. Chung L, et al. (2000) Identification of the (RWTNNFREY191)-R-183 region as a critical segment of matrix metalloproteinase 1 for the expression of collagenolytic activity. *J Biol Chem* 275:29610-29617.
24. Emsley J, Knight CG, Farndale RW, Barnes MJ, Liddington RC (2000) Structural basis of collagen recognition by integrin α2β1. *Cell* 101:47-56.
25. Becker JW, et al. (1995) Stromelysin-1: three-dimensional structure of the inhibited catalytic domain and of the C-truncated proenzyme. *Protein Sci* 4:1966-1976.
26. Morgunova E, et al. (1999) Structure of human pro-matrix metalloproteinase-2: Activation mechanism revealed. *Science* 284:1667-1670.
27. Elkins PA, et al. (2002) Structure of the C-terminally truncated human ProMMP9, a gelatin-binding matrix metalloproteinase. *Acta Crystallogr D Biol Crystallogr* 58:1182-1192.
28. Jozic D, et al. (2005) X-ray structure of human proMMP-1: new insights into procollagenase activation and collagen binding. *J Biol Chem* 280:9578-9585.
29. Grams F, et al. (1995) X-ray structures of human neutrophil collagenase complexed with peptide hydroxamate and peptide thiol inhibitors. Implications for substrate binding and rational drug design. *Eur J Biochem* 228:830-841.
30. Billinghurst RC, et al. (1997) Enhanced cleavage of type II collagen by collagenases in osteoarthritic articular cartilage. *J Clin Invest* 99:1534-1545.
31. Lovejoy B, et al. (1999) Crystal structures of MMP-1 and -13 reveal the structural basis for selectivity of collagenase inhibitors. *Nature Struct Biol* 6:217-221.
32. Leikina E, Mertts MV, Kuznetsova N, Leikin S (2002) Type I collagen is thermally unstable at body temperature. *Proc Natl Acad Sci USA* 99:1314-1318.
33. Welgus HG, Jeffrey JJ, Eisen AZ (1981) Human skin fibroblast collagenase. Assessment of activation energy and deuterium isotope effect with collagenous substrates. *J Biol Chem* 256:9516-9521.
34. Brown RA, Hukins DW, Weiss JB, Twose TM (1977) Do mammalian collagenases and DNA restriction endonucleases share a similar mechanism for cleavage site recognition? *Biochem Biophys Res Commun* 74:1102-1108.
35. Fields GB (1991) A model for interstitial collagen catabolism by mammalian collagenases. *J Theor Biol* 153:585-602.
36. Fiori S, Sacca B, Moroder L (2002) Structural properties of a collagenous heterotrimer that mimics the collagenase cleavage site of collagen type I. *J Mol Biol* 319:1235-1242.
37. Stultz CM (2002) Localized unfolding of collagen explains collagenase cleavage near imino-poor sites. *J Mol Biol* 319:997-1003.
38. Welgus HG, Jeffrey JJ, Stricklin GP, Eisen AZ (1982) The gelatinolytic activity of human skin fibroblast collagenase. *J Biol Chem* 257:11534-11539.
39. Bertini I, et al. (2009) Interdomain flexibility in full-length matrix metalloproteinase-1 (MMP-1). *J Biol Chem* 284:12821-12828.
40. Arnold LH, et al. (2011) The interface between catalytic and hemopexin domains in matrix metalloproteinase-1 (MMP-1) conceals a collagen binding exosite. *J Biol Chem* 286:45073-45082.

41. Rosenblum G, et al. (2007) Insights into the structure and domain flexibility of full-length pro-matrix metalloproteinase-9/gelatinase B. *Structure* 15:1227-1236.
42. Bertini I, et al. (2008) Evidence of reciprocal reorientation of the catalytic and hemopexin-like domains of full-length MMP-12. *J Am Chem Soc* 130:7011-7021.
43. 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* 70:1154-1161.
44. Salsas-Escat R, Nerenberg PS, Stultz CM (2010) Cleavage site specificity and conformational selection in type I collagen degradation. *Biochemistry* 49:4147-4158.
45. Eckhard U, Schonauer E, Nuss D, Brandstetter H (2011) Structure of collagenase G reveals a chew-and-digest mechanism of bacterial collagenolysis. *Nat Struct Mol Biol* 18:1109-1114.
46. Visse R, Nagase H (2003) Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res* 92:827-839.
47. Palmier MO, et al. (2010) NMR and bioinformatics discovery of exosites that tune metalloelastase specificity for solubilized elastin and collagen triple helices. *J Biol Chem* 285:30918-30930.
48. Fulcher YG, Van Doren SR (2011) Remote exosites of the catalytic domain of matrix metalloproteinase-12 enhance elastin degradation. *Biochemistry* 50:9488-9499.
49. Robichaud TK, Steffensen B, Fields GB (2011) Exosite interactions impact matrix metalloproteinase collagen specificities. *J Biol Chem* 286:37535-37542.

Figure legends

Fig. 1. Binding of MMP-1(E200A) to immobilized collagen I. (*A*) Binding of full-length MMP-1(E200A) and its individual Cat(E200A) and Hpx domains at 20 °C.
(*B*) Binding of MMP-1(E200A) at different temperatures. (*C*) Binding of 1 μM MMP-1(E200A) in the presence of the active-site inhibitor GM6001 at different temperatures.

Fig. 2. Screening of a triple-helical peptide library of collagen II with proMMP-1(E200A), MMP-1(E200A), Hpx domain and Cat(E200A) domain. Error bars show standard deviations from three repeats. Critical Leu and Ile residues at P1' and P10' subsites are highlighted in the collagen II portion of the peptide sequences surrounding the collagenase cleavage site; ~, bond cleaved by collagenases.

Fig. 3. Collagen I footprint on MMP-1(E200A) determined by H/DXMS and mutagenesis. (*A*) Sites 1 and 3-6 are protected from deuterium incorporation and site 2 showing enhanced deuterium incorporation upon collagen binding. The sites are mapped onto the crystal structure of MMP-1(E200A) (7). Residues potentially involved in collagenolysis that were

mutated are indicated. Dashed line, predicted collagen binding direction. (*B*) Relative initial velocities of collagen I cleavage by MMP-1 and its variants normalized to WT.

Fig. 4. Crystal structure of the MMP-1(E200A)-triple helical collagen peptide complex. (*A*) Sequence of the collagen peptide used for co-crystallization with MMP-1(E200A). Residues within 4 Å distance of the enzyme in the complex are colored. The subsite designation is indicated for the leading chain. (*B*) Stereo view of the MMP-1(E200A)-collagen peptide complex. The collagen chains are colored cyan (L), green (M) and red (T) and the enzyme is shown as a grey surface with areas within 4 Å distance of the L, M and T chains colored correspondingly; magenta sphere, active-site zinc ion. (*C*) Interactions of the collagen chains (colored as in *B*) with the active site cleft of MMP-1(E200A). Hydrogen bonds are indicated by dashed lines. (*D*) Interactions of the collagen chains with the Hpx domain. Selected residues making enzyme-substrate contacts are labeled with respective colors and shown in stick representation (N, dark blue, O, red).

Fig. 5. Modes of collagen binding to MMP-1. Shown is relationship between the unproductive and productive collagen binding modes (see text). In the sequence alignment on the left, the trailing chain is shown twice to emphasize the circular nature of the chain arrangement. The P1' and P10' residues are boxed and residues interacting with MMP-1 are in bold. The structural models on the right show that the unproductive and productive binding modes are related by a simple rotation/translation of the collagen triple helix. Note that only the interactions with one collagen chain are changed (L in the unproductive mode, T in the productive mode); the interactions with the other two chains are the same in both modes.

Fig. 6. Model of the first transition state of collagenolysis based on the productive complex. The upper rim of the catalytic site cleft anchors the M chain (green) and the Hpx domain anchors the L chain (cyan). These interactions position the T chain above the active site. Inter-domain flexing of the enzyme bends the collagen triple helix and facilitates the insertion of Leu(P1') of the T chain into the S1' pocket.

Figure 1

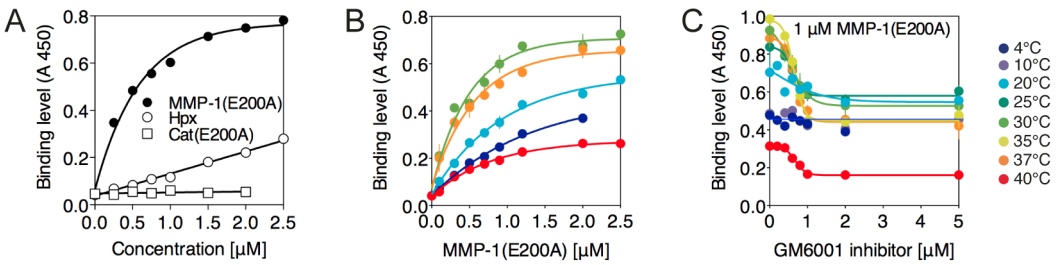


Figure 2

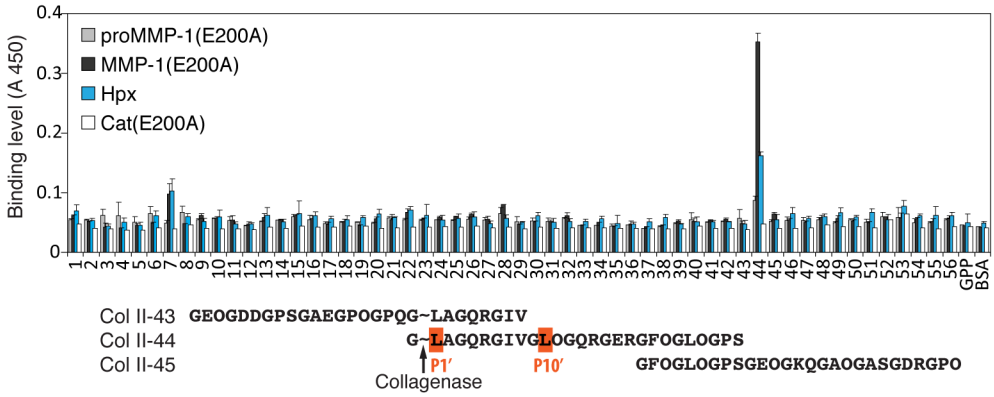


Figure 3

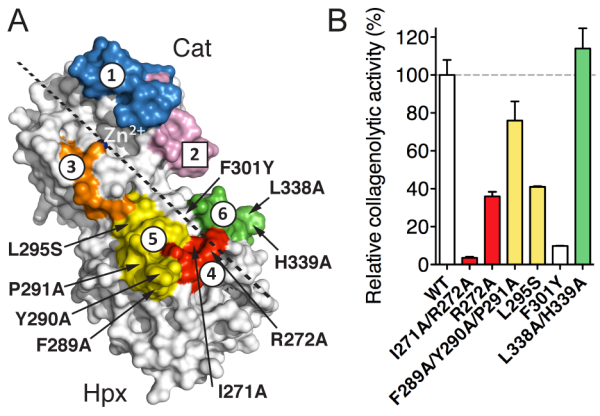


Figure 4

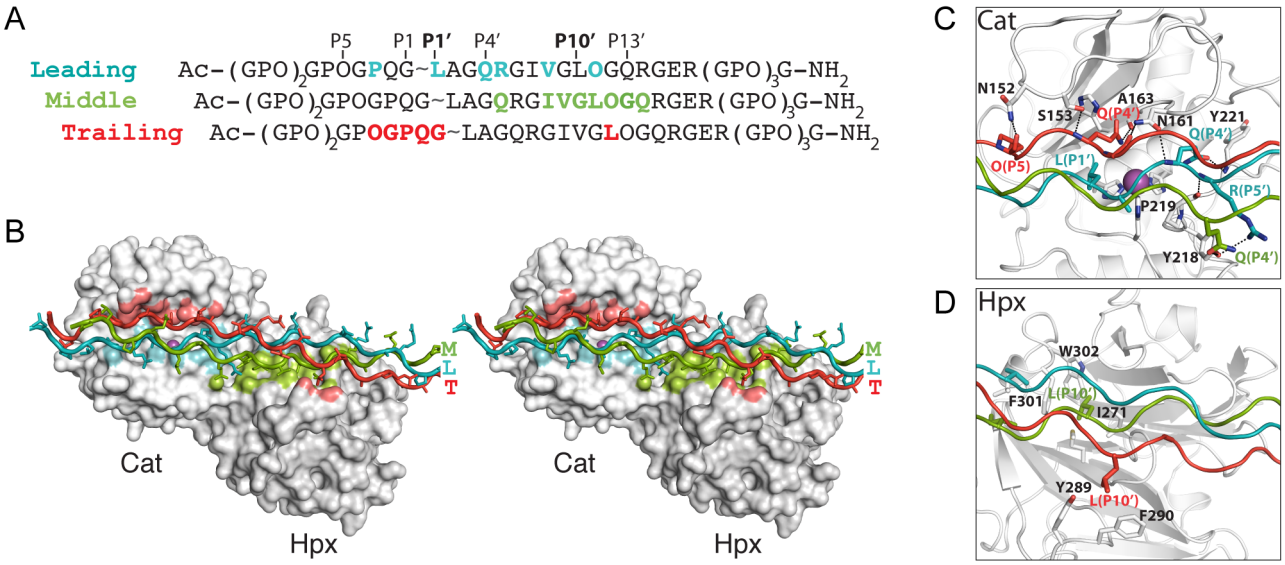


Figure 5

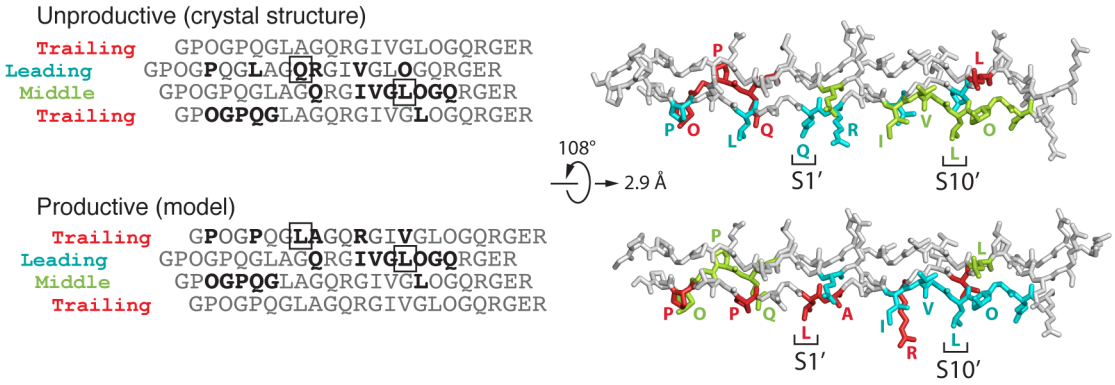
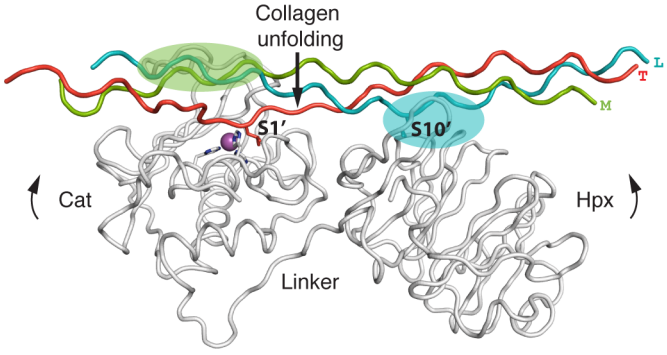


Figure 6



Supporting Information

SI Materials and Methods

Preparation of MMP-1 derivatives. ProMMP-1(E200A) (N-terminus of prodomain is number 1), the Hpx domain of MMP-1 (252-447), proMMP-1Cat(E200A) (1-242), and the proforms of the active MMP-1 variants were constructed using standard PCR methods, cloned in the pET3a vector, overexpressed in *E. coli* BL21 DE3, refolded from inclusion bodies and purified as described previously (1). ProMMPs were activated with a catalytic domain of MMP-3 in a 50:1 molar ratio and 1 mM 4-aminophenyl mercuric acetate in TNC buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 0.02 % NaN_3) for 60-120 min at 37 °C. The mature forms were finally purified by Sephacryl S-200 gel filtration (GE Healthcare).

Purification of collagen I. Type I collagen was extracted from Guinea pig dermis. The skin was extensively scraped, cut into pieces, washed with saline, extracted with 0.5 M acetic acid with pepsin added to 1/50 of the total wet weight for 24 h at 4 °C, and collagen was purified as described (2). Yield was determined after freeze-drying.

Synthesis of collagen peptides. A series of triple-helical collagen peptides that collectively represent the whole molecule of human collagen II (Toolkit II) as well as mutant peptides (Fig. S2*A*) and the peptide used for co-crystallization with MMP-1(E200A) were synthesized by Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry as C-terminal amides on TentaGel R RAM resin in an Applied Biosystems Pioneer automated synthesizer and purified as described (3). Every peptide in the Toolkit contains 27 amino acids of collagen II (so called "guest" sequence), with 9 amino acid overlap between the neighbouring peptides, flanked by 5 GPP repeats and a GPC knot ("host" sequences) that impart triple-helical conformation on the whole peptide. For exact sequences see (4). All peptides were verified by mass spectrometry and shown to adopt triple-helical conformation by polarimetry.

Collagen I and collagen peptide binding assays. Proteins in 50 mM (CHES) buffer pH 8.8 containing 200 mM NaCl and 10 mM $CaCl_2$ were reacted with EZ link LC-NHS biotin (Sigma) in water at a final 1:2 protein:biotin molar ratio for 1 h at room temperature. Excess biotin was removed by Sephadex G-25M PD-10 column (GE Healthcare) equilibrated in TNC buffer. The biotinylation did not affect collagenase activity of active MMP-1 and collagen unfolding activity(5) of MMP-1(E200A). For collagen binding assay, Costar® High Binding 96-well microtiter plates (Corning, UK) were coated with 50 µl of 20 µg/ml collagen I in TNC buffer, overnight at room temperature. They were washed with TNC buffer containing 0.05 % Tween 20 (TNC-T) and blocked with 3 % bovine serum albumin (Sigma) in TNC-T buffer. Biotinylated proteins at increasing concentrations were added in TNC buffer and incubated for 2 h at 4-40 °C. The wells were then washed in TNC-T buffer at the temperature of incubation and in temperature-dependent binding experiments they were then fixed with 3 % p-formaldehyde

for 30 min. The fixing step was omitted in room temperature binding experiments. Plates were developed using streptavidin-horseradish peroxidase conjugate (R&D, UK) and 3,3',5,5'-tetramethylbenzidine 2-Component Microwell Peroxidase Substrate Kit™ (KPL, UK) for a fixed time. For the collagen peptide library (Toolkit II) screening the Costar plates were coated with a 5 μg/ml collagen peptide solution in 10 mM acetic acid, incubated overnight at 4 °C, washed and blocked as described above. Biotinylated proteins at 1 μM concentration were added and incubated 1-2 h at room temperature. Plates were developed as in the collagen binding assay. All assays were carried out in a triplicate and compared analyses were always developed simultaneously.

H/DXMS experiment. We first established a peptic peptide mass mapping of MMP-1(E200A) which was essential for H/D exchange localization. MMP-1(E200A) was digested with pepsin by passing through a homemade 2.1 x 50 mm pepsin column based on POROS AL20 material (Applied Biosystems) and the peptic fragments were separated on a 2.1 x 250 mm C18 column (Grace/Vydac). The C18 fractions were subjected to nano-LC-ESI and MALDI MS/MS analyses (Waters). The combined peptide mapping covered 84% of MMP-1(E200A) sequence (Table S1). For H/D exchange analyses 15 μM MMP-1(E200A) was incubated with or without 20 μM collagen I at 25 °C for 1 h in 25 mM MOPS buffer containing 150 mM NaCl and 10 mM $CaCl_2$ (H_2O-MOPS-NC) pH7.5 and diluted 5-fold into 25 mM D_2O-MOPS-NC pD7.5, so that the final D_2O content was 80 %. Samples were collected from the H/D exchange reaction mix at various timepoints (5 s – 5 h), cooled to 0 °C, acidified to pH 2.5 by addition of concentrated formic acid, and immediately passed through the pepsin column, followed by a 2.1 x 20 mm POROS R1 (Applied Biosystems) trap column. The resultant peptides were then immediately separated on a 2.1 x 50 mm C18 column (Grace/Vydac) at 1 °C and eluted directly into the Micromass Q-TOF mass spectrometer (Waters). Deuterium exchanged spectra were assigned using the catalogue of MMP-1 peptic peptides (Table S1). Natural isotopic patterns of assigned MMP-1(E200A) peptides as well as those of each time-point of H/D exchange reaction were found in total ion chromatograms and identified in MS spectra by manual inspection. Centroid (average) mass of each isotopic pattern was computed using HX-Express software (6), which enabled determination of a relative deuterium level in a peptide at given H/D exchange time-point (Fig. S3*B*).

Collagenase activity assay. The wild-type MMP-1 and its mutants were used at a concentration of 10 nM to digest 4.5 μM collagen I in TNC buffer at 25 °C. Reactions were stopped at different time-points by the addition of SDS-PAGE loading buffer containing 20 mM EDTA. Products were analyzed by SDS-PAGE with 7.5 % total acrylamide under reducing conditions and the gels were stained with Coomassie Brilliant Blue R-250. The percentage cleavage was measured by densitometry using ImageScanner III (GE Healthcare) and Phoretics 1D (TotalLab) quantification software from which the initial rate of cleavage was calculated.

Complex formation and crystallization. The MMP-1(E200A)-collagen peptide complex for crystallization was formed by dissolving 2.2 mg of lyophilized peptide in 230 µl of a concentrated MMP-1(E200A) solution (23.5 mg/ml protein in TNC buffer). After incubation for 30 minutes at room temperature, the mixture was subjected to size exclusion chromatography on a Superdex 75 HR10/30 column (GE Healthcare) at room temperature (Fig. S5). The MMP-1(E200A)-collagen peptide complex was concentrated to 8 mg/ml and screened for crystallization in vapor diffusion sitting drops using a Mosquito nanolitre robot (TTP LabTech). Clusters of thin needles grew after 2-3 days using 0.1 M Tris-HCl pH 8.5, 2% (v/v) Tacsimate pH 8.0, 16% (w/v) PEG 3350 as precipitant. These crystals were used to generate a seed stock using the Seed Bead kit (Hampton Research) according to the manufacturer's protocol. Single needle-shaped crystals of a few µm thickness were obtained by microseeding crystallization drops pre-equilibrated in 0.1 M Tris-HCl pH 8.5, 8-11% (v/v) Tacsimate pH 8.0, 10-12.5% (w/v) PEG 3350. The crystals were flash-frozen in liquid nitrogen after a brief soak in mother liquor supplemented with 20% glycerol.

Crystal structure determination. Diffraction data were collected at 100 K on the microfocus beam line I24 at the Diamond Light Source (Oxfordshire, UK). Many crystals were screened to identify a single good crystal, from which all data were collected. To minimize radiation damage, a fresh crystal volume was exposed after each 20° of data collected, yielding a total of 140° of usable data. The data were processed with MOSFLM (www.mrc-lmb.cam.ac.uk/harry/mosflm) and programs of the CCP4 suite (7). The MMP-1-collagen structure was solved by molecular replacement with PHASER (8) using the human MMP-1 structure (PDB entry 2CLT) as a search model. After refinement with CNS (9), the electron density map calculated from the correctly positioned search model showed weak density for the collagen triple helix. A collagen peptide with poly-Gly-Ala-Ala sequence was constructed based on PDB entry 2WUH and positioned with PHASER. The model was completed using O (10) and refined with CNS (Table S2). Omit maps were used extensively to assign the correct register of the three collagen chains. The register of the leading chain passing through the active site cleft is defined unambiguously by the strong electron density of the Arg($P5'_L$) side chain. The two MMP-1(E200A)-collagen complexes in the asymmetric unit are very similar and tight non-crystallographic symmetry restraints were applied throughout the refinement. The only deviations from non-crystallographic symmetry are observed for the first three collagen triplets, which are not part of the collagen II sequence and do not interact with MMP-1(E200A).

1. Chung L, et al. (2000) Identification of the (183)RWTNNFREY(191) region as a critical segment of matrix metalloproteinase 1 for the expression of collagenolytic activity. *J Biol Chem* 275:29610-29617.
2. Seyer JM, Hutcheson ET, Kang AH (1976) Collagen polymorphism in idiopathic chronic pulmonary fibrosis. *J Clin Invest* 57:1498-1507.
3. Raynal N, et al. (2006) Use of synthetic peptides to locate novel integrin alpha2beta1-binding motifs in human collagen III. *J Biol Chem* 281:3821-3831.

4. Leo JC, et al. (2010) First analysis of a bacterial collagen-binding protein with collagen Toolkits: promiscuous binding of YadA to collagens may explain how YadA interferes with host processes. *Infect Immun* 78:3226-3236.
5. Chung L, et al. (2004) Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. *EMBO J* 23:3020-3030.
6. Weis DD, Engen JR, Kass IJ (2006) Semi-automated data processing of hydrogen exchange mass spectra using HX-Express. *J Am Soc Mass Spectrom* 17:1700-1703.
7. The CCP4 suite: programs for protein crystallography (1994). *Acta Crystallogr D Biol Crystallogr* 50(5):760-763.
8. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658-674.
9. Brunger AT, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54:905-921.
10. Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47:110-119.
11. Iyer S, Visse R, Nagase H, Acharya KR (2006) Crystal structure of an active form of human MMP-1. *J Mol Biol* 362:78-88.

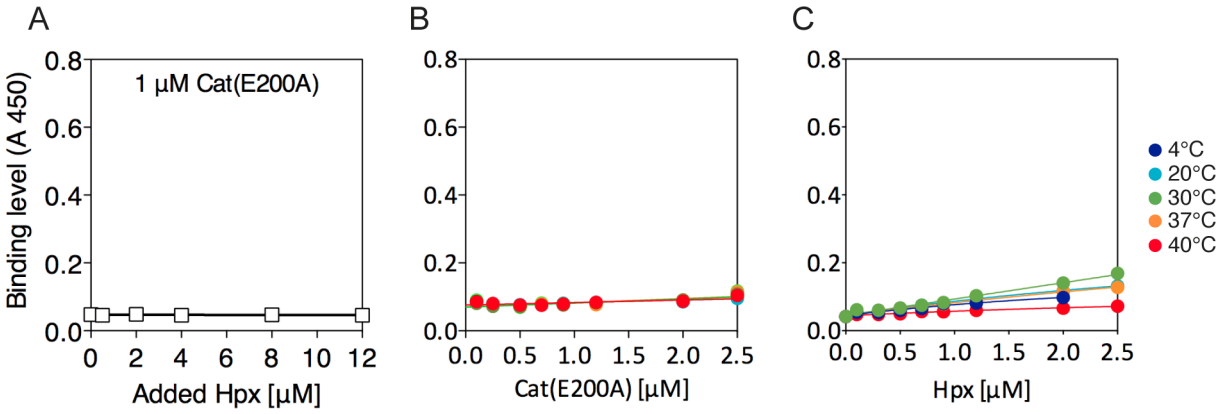


Fig. S1. Binding of the individual domains of MMP-1(E200A) to immobilized collagen I**.** (*A*) Cat(E200A) domain in the presence of increasing concentrations of the Hpx domain. (*B*) Cat(E200A) domain at increasing temperatures. (*C*) Hpx domain at increasing temperatures. All data points are triplicate average with SD smaller than a point size.

A

Sequence	MMP-1(E200A) binding
Col II-44 G~LAGQRGIVGLOGQRGERGFOGLOGPS	0.35 ± 0.01
G~LAGQRGIVGLOGQRGER	0.23 ± 0.04
GFOGLOGPS	0.06 ± 0.01
GQRGIVGLOGQRGER	0.08 ± 0.01
GIVGLOGQRGER	0.08 ± 0.01
G~LAGQRGIVGLOGQR	0.25 ± 0.03
G~LAGQRGIVGLO	0.23 ± 0.04
GQRGIVGLO	0.09 ± 0.02
GIVGLOGQR	0.10 ± 0.02
GLOGQRGERGFO	0.09 ± 0.02
GPQG~LAGQRGIVGLOGQRGER	0.29 ± 0.02
GAQG~LAGQRGIVGLOGQRGER	0.36 ± 0.04
GPAG~LAGQRGIVGLOGQRGER	0.45 ± 0.03
GPQG~AAGQRGIVGLOGQRGER	0.14 ± 0.01
GPQG~LAGARGIVGLOGQRGER	0.27 ± 0.01
GPQG~LAGQAGIVGLOGQRGER	0.55 ± 0.05
GPQG~LAGQRGAVGLOGQRGER	0.39 ± 0.01
GPQG~LAGQRGIAGLOGQRGER	0.36 ± 0.02
GPQG~LAGQRGIVGAOGQRGER	0.09 ± 0.01
GPQG~LAGQRGIVGLAGQRGER	0.26 ± 0.02
GPQG~LAGQRGIVGLOGARGER	0.65 ± 0.04
GPQG~LAGQRGIVGLOGQAGER	0.22 ± 0.01
GPQG~LAGQRGIVGLOGQRGAR	0.56 ± 0.02
GPQG~LAGQRGIVGLOGQRGEA	0.25 ± 0.02

B

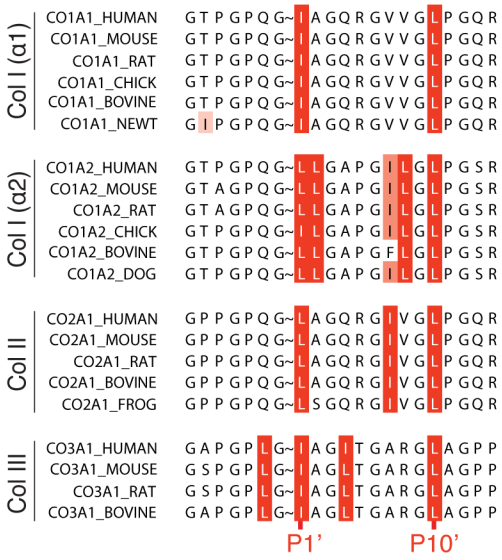


Fig. S2. Identification of the MMP-1(E200A) binding site in collagen II. (*A*) Binding of MMP-1(E200A) to truncated and alanine-substituted (shown in blue) derivatives of peptide Col II-44. Binding results are expressed as mean A_{450} values ± standard deviations from three independent experiments. Low binding data are shown in red text; O, hydroxyproline. (*B*) Sequence alignment of interstitial collagens I, II and III from different species in the proximity of the collagenase-cleavage site. Sequences are derived from cDNA. Leu and Ile residues are highlighted in red, and residues at P1' and P10' subsites are indicated. ~, Bond cleaved by collagenases.

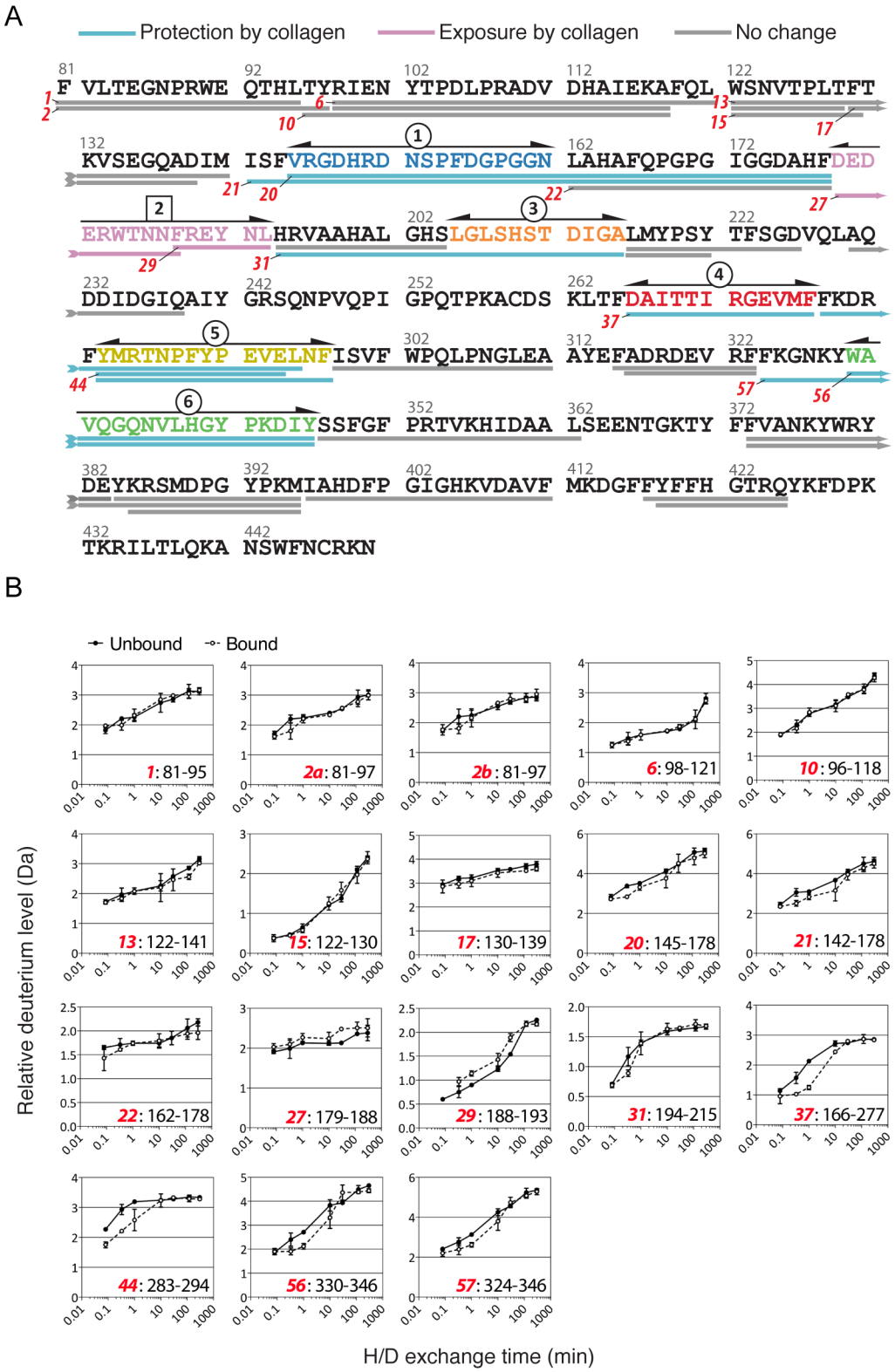


Fig. S3. Collagen I footprint on MMP-1(E200A) determined by H/DXMS. (*A*) Sites protected from deuterium incorporation (1 and 3-6, circled) or showing enhanced deuterium incorporation (site 2, square) upon collagen binding are indicated in the MMP-1(E200A) sequence. Site numbering and color coding corresponds with Fig. 2*A*, where the sites are mapped onto the crystal structure of MMP-1(E200A). All analyzed peptides are shown as bars below the sequence. Numbers in red refer to peptide numbers listed in Table S1. (*B*) H/D exchange time-course analyses of selected MMP-1(E200A) peptic peptides; Unbound, free MMP-1(E200A); Bound, MMP-1(E200A) bound to collagen.

Quantification of the exchange in MS spectra was performed using HX-Express software (6). Centroid (average) mass of each isotopic pattern was computed, which allowed determination of a relative deuterium level (*rdl*) in a peptide at given H/D exchange time-point by the following formula: $rdl = m - m_0$, where m is the average mass after H/D exchange and m_0 is the unexchanged average mass. Data are plotted by averaging results of two or three independent experiments. The numbers in red shown with sequence delimiters correspond to the peptide numbers in *A*; 2a and 2b denote different ions of the same peptide.

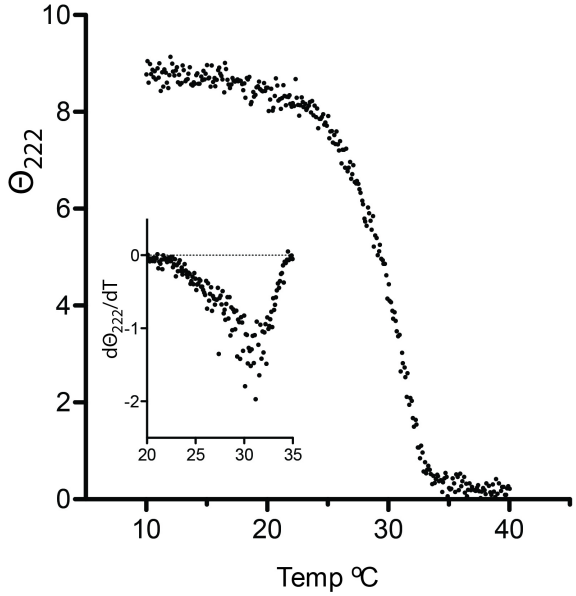


Fig. S4. Melting curve of collagen peptide used for co-crystallization with MMP-1(E200A) determined by circular dichroism (CD). The thermal transition of collagen peptide (10 µM) in TNC buffer was determined in a Jasco 815 CD instrument at 222 nm with a 0.1 cm pathlength. Temperature was increased at 0.1 °C/min. The first derivative of the smoothed melting curve is shown in the inset. Θ, elipticity.

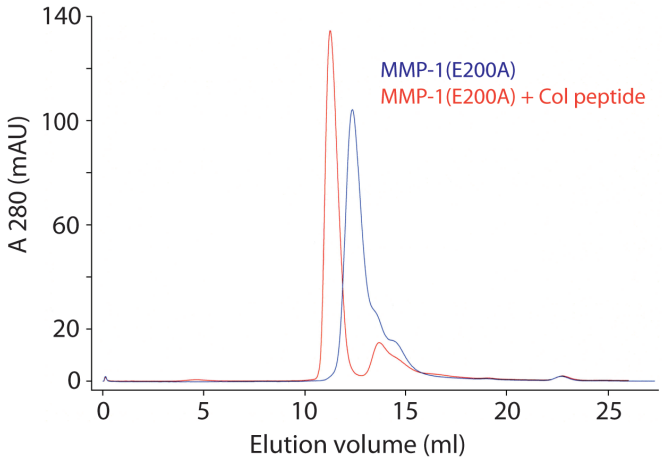


Fig. S5. Solution binding of MMP-1(E200A) to the collagen peptide. Analytical size exclusion (Superdex 75) chromatograms of the free MMP-1(E200A) and its complex with the collagen peptide.

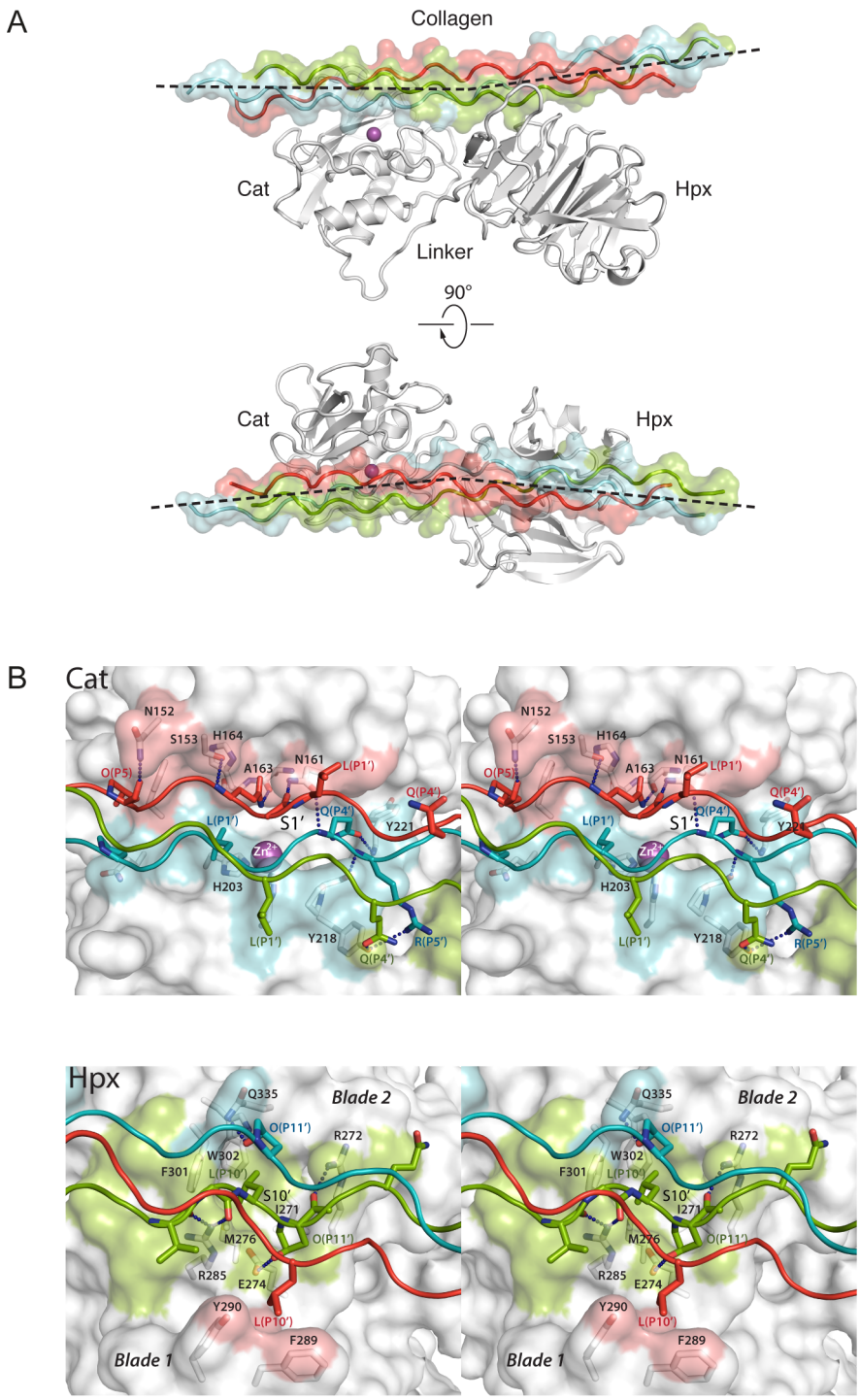


Fig. S6. Overall structure of the MMP-1(E200A)-collagen peptide complex and details of the enzyme-substrate interactions. (*A*) Collagen bending. The collagen chains are colored cyan (L), green (M) and red (T). The dashed lines indicate the helix axes in the N- and C-terminal regions of the collagen peptide, which is bent by ~10° halfway between the Cat and Hpx domains. A magenta sphere represents the active-site zinc ion. (*B*) Stereo views of the Cat and the Hpx domains in the complex. The collagen chains are colored as in *A*. The MMP-1(E200A) surface areas within 4 Å distance of the collagen peptide are colored correspondingly. Selected residues making enzyme-substrate contacts are labeled and shown in stick representation (N, dark blue, O, red). Dotted lines denote hydrogen bonds. The S1' and S10' pockets are indicated.

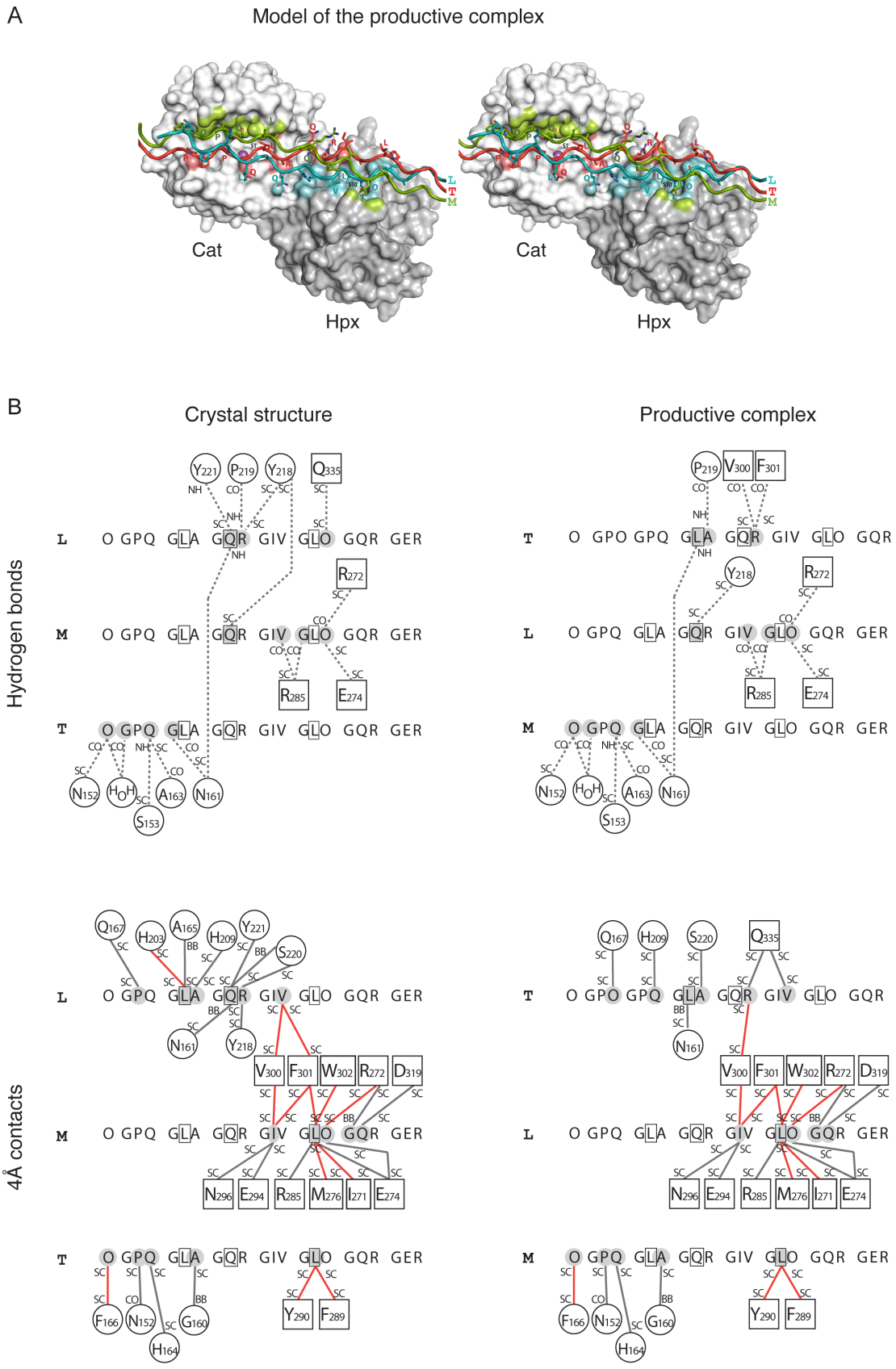


Fig. S7. Model of the productive MMP-1-collagen complex and schematic representation of MMP-1-collagen interactions. (*A*) Stereo image of the modeled productive complex. The collagen chains are in cyan (L), green (M) and red (T), respectively. The MMP-1(E200A) surface areas within 4 Å distance of the L, M and T chains are colored

correspondingly. Collagen side chains are shown as sticks and have been labeled. (*B*) Schematic representation of MMP-1-collagen interactions in the crystallized unproductive complex (left panels) and modeled productive complex (right panels). MMP-1 residues are shown as circles (Cat domain) or squares (Hpx domain). Collagen subsites P1', P4' and P10' are boxed. Dashed lines indicate hydrogen bonds (upper panels) and solid lines indicate contacts of 4 Å distance or less (lower panels). Hydrophobic contacts are emphasized by red lines. BB, backbone; CO, main-chain carbonyl oxygen; NH, main-chain amide nitrogen; SC, side chain.

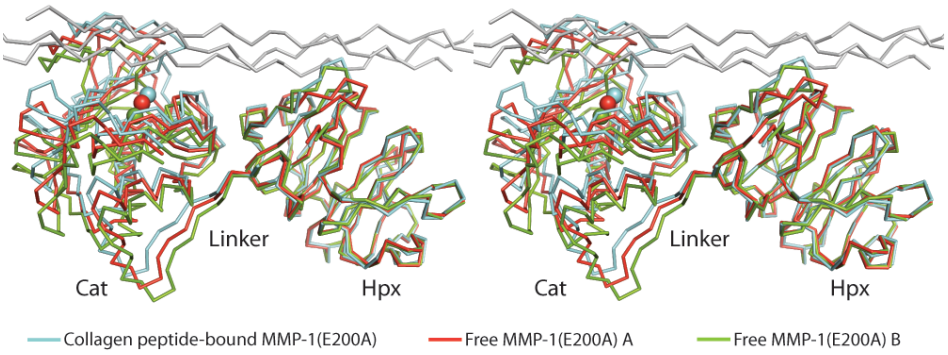


Fig. S8. Inter-domain flexibility in MMP-1. Shown is a stereoview of a superposition of the two crystallographically independent molecules of free MMP-1(E200A) (11) (molecule A, red; molecule B, green) and the MMP-1(E200A)-collagen peptide complex (MMP-1, cyan; collagen, gray). The structures were superimposed on the Hpx domains using Pymol.

Table S1. MMP-1(E200A) peptic peptides identified by nano-LC-Q-TOF and/or MALDI-Q-TOF MS/MS for H/D exchange analyses.

			Observed m/z = (M+z)/z		M	M
No.	Start-End	Sequence	ESI	MALDI	(experimental)	(theor.)
1	81-95	FVLTEGNPRWEQTHL	609.75 - 609.79; 914.16	1826.93	1826.23 - 1826.34	1825.91
2	81-97	FVLTEGNPRWEQTHLTY	1046.25	2091.03	2090.49	2090.02
3	82-97	VLTEGNPRWEQTHLTY		1943.9	1942.89	1942.95
4	98-116	RIENYTPDLPRADVDHAIE	742.19; 556.9 - 556.92		2223.54 - 2223.65	2223.09
5	98-118	RIENYTPDLPRADVDHAIEKA	808.56 - 808.57	2423.24	2422.67 - 2422.69	2422.23
6	98-121	RIENYTPDLPRADVDHAIEKAFQL	703.58-703.6; 703.75	2811.43	2810.38; 2810.98	2810.44
7	104-118	PDLPRADVDHAIEKA	549.74		1646.2	1645.84
8	105-115	DLPRADVDHAI	407.57		1219.69	1220.61
9	96-116	TYRIENYTPDLPRADVDHAIE	-	2488	2487	2487.2
10	96-118	TYRIENYTPDLPRADVDHAIEKA	672.75		2686.99	2686.34
11	131-141	TKVSEGQADIM	589.43 - 589.93		1176.84 - 1177.84	1177.56
12	127-141	PLTFTKVSEGQADIM	-	1636.81	1635.8	1635.82
13	122-141	WSNVTPLTFTKVSEGQADIM	1112.77 - 1112.78	2224.06	2223.53	2223.09
14	122-129	WSNVTPLT	917.62 - 917.7		916.61 - 916.7	916.47
15	122-130	WSNVTPLTF	532.75 - 532.78		1063.49 - 1063.55	1063.53
16	130-138	FTKVSEGQA	483.73 - 483.74		965.45 - 965.47	965.48
17	130-139	FTKVSEGQAD	541.25 - 541.26; 541.39		1080.51; 1080.76	1080.51
18	130-141	FTKVSEGQADIM	663.05 - 663.49	1325.62	1324.09 - 1324.96	1324.63
19	145-161	VRGDHRDNSPFDGPGGN	599.59 - 599.60	1796.79	1795.78	1795.8
20	145-178	VRGDHRDNSPFDGPGGNLAHAFQPGPGIGGDAHF	868.1 - 868.12	3469.67	3468.47 - 3468.66	3468.61
21	142-178	ISFVRGDHRDNSPFDGPGGNLAHAFQPGPGIGGDAHF	-	3816.99; 3817.10	3815.98; 3816.09	3815.79
22	162-178	LAHAFQPGPGIGGDAHF	564.59 - 564.61	-	1690.75 - 1690.81	1690.82
23	179-187	DEDERWTNN	-	1178.47	1177.46	1177.46
24	179-190	DEDERWTNNFRE	-	1610.67	1609.66	1609.68
25	179-193	DEDERWTNNFREYNL	667.62	2000.85	1999.84	1999.87
26	188-198	FREYNLHRVAA	-	1375.71	1374.7	1374.72
27	179-188	DEDERWTNNF	663.25 - 663.27	-	1324.48 - 1324.52	1324.53
28	189-198	REYNLHRVAA	-	1228.64	1227.63	1227.65
29	188-193	FREYNL	421.20 - 421.22	-	840.38 - 840.42	840.41
30	194-214	HRVAAHELGHSLGLSHSTDIG	-	2136.09 Glu -> Ala (E) [-58.01]	2135.08	2135.1
31	194-215	HRVAAHELGHSLGLSHSTDIGA	736.53 Glu -> Ala (E) [-58.01]	2207.13 Glu -> Ala (E) [-58.01]	2206.58	2206.14
32	194-204	HRVAAHELGHS	385.87 Glu -> Ala (E) [-58.01]	-	1154.59	1154.61
33	208-216	SHSTDIGAL	450.72	-	899.42	899.43
34	216-226	LMYPSYTFSGD	640.75 - 640.76	-	1279.49 - 1279.50	1279.54
35	230-238	AQDDIDGIQ	487.71; 974.61 - 974.66	-	973.4; 973.6 - 973.65	973.47
36	266-276	DAITTIRGEVM	603.29 - 603.31; 603.46	1205.59	1204.56 - 1204.61	1204.61
37	266-277	DAITTIRGEVMF	676.81 - 677.0	1352.67; 1368 ox(M) [+16]	1351.6 - 1351.99	1351.68
38	266-273	DAITTIRG	423.73 - 423.74	-	845.44 - 845.47	845.46
39	267-277	AITTIRGEVMF	619.33 - 619.34	-	1236.65	1236.65
40	270-277	TIRGEVMF	476.73 - 476.75	-	951.44 - 951.48	951.48
41	277-282	FFKDRF	430.23	-	858.44	858.44
42	277-281	FFKDR	356.77 - 356.78	-	711.53 - 711.55	711.33
43	278-295	FKDRFYMRTNPFYPEVEL	784.68	-	2351.03	2351.14
44	283-294	YMRTNPFYPEVE	773.30 - 773.35	1545.7; 1561 ox(M) [+16]	1544.69	1544.7
45	283-295	YMRTNPFYPEVEL	829.86 - 829.88	1658.79	1657.78	1657.78
46	283-297	YMRTNPFYPEVELNF	960.63 - 960.64	1919.88	1919.25 - 1919.27	1918.89
47	283-296	YMRTNPFYPEVELN	886.88 - 886.9	-	1771.74 - 1771.79	1771.82
48	298-311	ISVFWPQLPNGLEA	-	1602.79 dox(W) [+31.99]	1601.78	1601.81
49	313-323	YEFADRDEVRF	-	1446.66	1445.65	1445.66
50	315-323	FADRDEVRF	358.52 - 358.53 ; 577.78 - 577.93	1154.52; 1154.55	1153.53 - 1153.84	1153.55
51	316-323	ADRDEVRF	-	1007.48	1006.47	1006.48
52	316-329	ADRDEVRFFKGNKY	-	1744.86	1743.85	1743.87
53	315-329	FADRDEVRFFKGNKY	-	1891.84	1890.83	1890.94
54	315-322	FADRDEVR	336.5; 504.24	-	1006.47; 1006.48	1006.48
55	347-362	SSFGFPRTVKHIDAAL	582.77	-	1745.28	1744.93
56	330-346	WAVQGQNVLHGYPKDIY	-	1987.98	1986.97	1987
57	324-346	FKGNKYWAVQGQNVLHGYPKDIY	682.09	-	2724.33	2724.38
58	330-338	WAVQGQNVL	507.76	-	1013.5	1013.53
59	352-362	PRTVKHIDAAL	610.85	-	1219.68	1219.7
60	363-372	SEENTGKTYF	588.22 - 588.41	1175.52	1174.42 - 1174.8	1174.51
61	363-371	SEENTGKTY	514.71	-	1027.41	1027.45
62	373-383	FVANKYWRYDE	497.53 - 497.55; 745.81 - 745.83	1490.67; 1522.69 dox(W) [+31.99]	1489.58 - 1489.66	1489.7
63	373-384	FVANKYWRYDEY	551.91	1653.76	1652.69 - 1652.75	1652.76
64	373-395	FVANKYWRYDEYKRSMDPGYPKM	-	2944.37; 2944.42	2943.36; 2943.41	2943.38
65	373-385	FVANKYWRYDEYK	-	1653.76 Lys-loss (C-term) [-128.09]	1652.75	1652.76
66	382-395	DEYKRSMDPGYPKM	572.91	-	1715.71	1715.76
67	384-395	YKRSMDPGYPKM	491.54 - 491.57; 491.69	1488 ox(M) [+16]	1471.6 - 1471.68; 1472.04	1471.7
68	385-395	KRSMDPGYPKM	437.2 - 437.21	1309.63	1308.59 - 1308.62	1308.63
69	396-409	IAHDFPGIGHKVDA	492.9 - 492.91	1476.76	1475.69 - 1475.75	1475.75
70	396-411	IAHDFPGIGHKVDAVF	431.46; 574.93 - 575.09; 861.92	1722.89	1721.78 - 1722.28	1721.89
71	417-425	FYFFHGTRQ	-	1202.57	1201.56	1201.57
72	418-425	YFFHGTRQ	528.25 - 528.26	1055.5	1054.48 - 1054.5	1054.5
73	419-423	FFHGT	608.43	-	607.43	607.28

Numbers in red refer to the peptides for which H/D exchange kinetics is shown in Figure S2. Many peptides were identified in differently charged forms, especially with the electrospray ionization (ESI) method. Several ions were detected multiple times, and the ranges of their observed mass-over-charge (m/z) values are shown. M, peptide mass; M(experimental) was calculated as follows: M = ((m/z) z) - z.

Table S2. Crystallographic statistics of the MMP-1(E200A)-collagen peptide structure

Data collection	
Space group	P2
Unit cell dimensions	
a, b, c (Å)	76.67, 102.24, 80.73
α, β, γ (°)	90, 103.76, 90
Asymmetric unit content	2 MMP-1(E200A)-collagen complexes
Solvent content (%)	57
Resolution (Å)	20-3.0 (3.16-3.00)[a]
R_{merge}	0.128 (0.400)
$\langle I/\sigma(I) \rangle$	5.0 (2.2)
Completeness (%)	86.8 (77.9)
Multiplicity	1.7 (1.6)
Refinement	
Resolution (Å)	20-3.0
Reflections	20982
Protein atoms	7368
Solvent atoms	8 Ca^{2+} + 4 Zn^{2+} + 7 H_2O
R_{work}/R_{free}	0.211/0.273
R.m.s. deviation bonds (Å)	0.008
R.m.s. deviation angles (°)	1.4
Ramachandran plot (%)[b]	79.2, 18.5, 2.2, 0.1

[a]Values in parentheses are for the highest resolution shell.

[b]Percentage of residues in core, allowed, generously allowed and disallowed regions (program PROCHECK).