

Supporting Information

Enzymatic turnover of macromolecules generates long lasting protein-water coupled motions beyond reaction steady-state

This file contains:

Supplementary Materials and methods

Supplementary Figures 1-8

Supplementary tables 1-2

Supplementary Materials and Methods

Circular dichroism (CD) measurements. CD spectroscopy and real-time CD have been performed at a Jasco J-815 (Jasco Co., Groß-Umstadt, Germany) CD spectrophotometer purged with nitrogen. The sample was temperature-stabilized at room-temperature by a Peltier element. Measurements were done in 10 mM Tris(H₂SO₄), 100 mM Na₂SO₄, 5 mM CaSO₄ buffer. CD data were collected from 190-260 nm at a scan rate of 20 nm min⁻¹ and with a time constant of 1 s. For each sample, five spectra were collected and averaged. Real-time CD was measured at fixed wavelengths of 205 nm and 228 nm with a bandwidth of 1 nm. The kinetic reaction of substrate hydrolysis was initialized by mixing the enzyme solution with the respective collagen substrate in a Biologic stopped-flow instrument SFM-300 (Biologic, Grenoble, France) in a 1:1 volume ratio with a reaction volume of 126 µL after mixing. The final concentrations were 20 µM substrate (collagen-like or gelatin-like) and 1 µM MT1-MMP enzyme solution. A 400 V pre-voltage alignment was used to detect substrate hydrolysis with a time constant of 2 ms and an instrument dead time of 15.2 ms. For analysis, 15 independent CD kinetics were averaged and evaluated.

Kinetic THz data collection and analysis. MT1-MMP activity was initiated by 1:6 (v:v) mixing of enzyme and substrate solutions in assay buffer using a stopped-flow cell (Unisoku, Japan) of 500 µm z-cut quartz windows and a volume of 40 µL. The final concentrations after mixing were 20 µM MT1-MMP and 400 µM substrate solution, keeping the [E:S] ratio at 1:20. For time-resolved terahertz absorption experiments, we combined a terahertz time domain spectrometer (THz-TDS) with a stopped flow mixer. This spectrometer generates terahertz pulses (0.3-1.0 THz) with a repetition rate of 1 GHz. For THz generation and detection, a mode locked 800 mW titanium sapphire femtosecond laser (Gigaoptics, model gigajet TWIN) is split into a pump and a probe beam. The beams have a dedicated difference in optical path length which is controlled by a translation stage. By changing the mutual delay of pump and probe beam, the THz pulse can be sampled in time. The pump beam triggers the generation of the THz pulse by photoconductive switching of a large area GaAs based antenna (Gigaoptics, Konstanz, Germany model Tera Sed). The THz pulse is detected by electro optical sampling in a 2 mm zinc telluride crystal. The transmitted probe beam is split into two linear components by a polarizing beam splitter. The intensity difference of the linear polarization components is measured by a autobalanced photodetector (Newport Co., Irvine, CA, model Nirvana, bandwidth: 125 kHz). Signal to noise ratio is optimized using a lock in amplifier (Signal Recovery, Oak Ridge, TN, model 7265, 250 kHz) with 6 dB/Oct and 640 µs integration time (respectively 24 dB/Oct and 20 ms for long time data). Data are acquired every 0.5 ms using a digitizer (GaGe, Lockport, IL, model Compuscope Octopus 14, Bandwidth 100 MHz, resolution: 14 bits). For kinetic terahertz measurements, the pump-probe delay of the THz-TDS was set to a constant value, such that kinetic measurements probe the time evolution of the THz pulse maximum E(t). Two sets of kinetic measurements were performed on different days. They complement each other with respect to time scales and sensitivity. In these measurements we focused on probing the long time relaxation of the THz signal instead of probing fast changes as in

our previous study, in order to study solvation effects beyond reaction steady-state. Each set consists of 20-30 single kinetic measurements. Each measurement was conducted as follows: 250 ms (resp. 1 s for the long time data) after the start of data acquisition, a signal was sent to the stopped flow device which triggered mixing. After 2250 ms (resp. 9 s for the long time data) the data acquisition was stopped. Before mixing, the sample from the prior measurement still is in the measurement cell. As a reference ($E_{Initial}$) served the equilibrated solution in the cell from the last measurement. The typical time between subsequent measurements was 30 s for the first data set and 45 s for the second data set (extending over 10 s) which was assumed to correspond to steady-state kinetics. The obtained time-dependent kinetic terahertz absorption (KITA(t)) signal was defined as:

$$KITA(t) = \frac{E_{THz}(0) - E_{THz}(t)}{E_{THz}(0)}$$

The terahertz transmission through this solution is used as a normalization factor for the data. Upon mixing, the solution is pushed into a reservoir on top of the measurement cell, from where it is removed with the help of a peristaltic pump.

Terahertz absorption measurements. Terahertz absorption spectroscopy measurements were carried out using our p-Germanium difference laser spectrometer as described before (1). Briefly, we recorded the change of the integrated absorption coefficient of the sample in the spectral range from 2.1 to 2.8 THz, at 20°C, in parallel to buffer measurements. Measurements were conducted in 50mM Tris pH 7.5, 100mM NaCl and 5mM CaCl₂.

Docking and MD simulations. The collagen-like substrate was modeled based on the canonical (Gly-X-Y)₅ triplets, and the three chains were assembled *in silico* to form a right-handed triple helix structure. To optimize the substrate model, we further conducted molecular dynamics (MD) simulations of the collagen-like substrate in explicit solvent for 120 ns to achieve full equilibration of the system. During the simulations, the sampled molecular conformations of the substrate clustered around three main representations: stretched, bended or ‘semi-bended’ while maintaining the triple helical configuration intact. These three different structures of the collagen-like substrate were separately docked into either the catalytic domain of MT1-MMP or to a homology model of full length MT1-MMP (see below for details) by rigid body docking (ref. for rigid body docking). Initial docking positions of the THP substrate around enzyme were generated by implementing Potential Smoothing and Search (PSS) algorithm (2, 3). However, the collagen-like substrate is taken in three distinct conformations identified in 120 ns molecular dynamics: stretched, semi-bended (S-shaped) and bended. Docking positions were refined by using all-atom energy minimization protocol (4) Docking of the collagen-like substrate was performed to the catalytic domain of MT1-MMP by using structural data available from PDB: 3MA2 (5), using a docking criterion in which the Gly22 carbonyl group proximate at the cleavage site of the collagen-like substrate is required to be within 5 Å distance from to the catalytic zinc ion and the active site structural water molecule. In this case we found different orientations and variety of bending angles of the collagen-like substrate when bound to MT1-MMP. To select a docking position of biological significance, we further modeled binding of the collagen-like substrate to full-

length MT1-MMP including its hemopexin domain known to be important for molecular recognition. A model of full length MT1-MMP was generated by homology modeling of the crystal structure of full length MMP-1 (PDB: 2CLT (6)), and the crystal structure of the hemopexin domain of MT1-MMP (PDB: 3C7X (7)). Refined docking structures were selected according to three criteria: a) distance from proximate Gly22 carbonyl group at the cleavage site of the collagen-like substrate to the catalytic zinc ion and the active site structural water molecule should be within 5 Å; b) bending angles of collagen-like substrate >120°; c) properly oriented collagen-like substrate interactions with the exosite area of our model of the full-length MT1-MMP. The complex between MT1-MMP and the gelatin-like substrate was obtained by removing the two collagen chains that do not directly interact with the zinc ion, followed by further energy minimization protocol (amber99 force field) (4).

Our docking results were compared to available NMR structures of structurally homologous MMP-collagen systems, demonstrating that the collagen-like substrates exist in either stretched or bended configurations. Our simulations demonstrate that the structure of the gelatin substrate did not remain in its extended conformation and rapidly folded after 2 ns of simulation to form a random coil structure, typical of gelatin, which has been in contrast to the collagen substrate (see Fig. S2).

Selected molecular conformations, shown in Fig. S1, have been chosen as starting configurations for additional molecular dynamics simulations. The selected systems are representative of two molecular scenarios as the substrate approaches the catalytic zinc ion characterized by the distance d between the zinc ion and the closest heavy atom of the substrate (see supplementary table S1): i) the free enzyme ($d \sim 7$ Å) and ii) the enzyme-substrate complex, i.e. the substrate approaching the enzyme active site (d is between 2 and 6 Å distance). d is calculated as an average over ten different simulations, as explained in the following.

The molecular dynamics simulations were performed using the GROMACS software package (8). For MT1-MMP enzyme and the substrates the CHARMM27 force field (9) with CMAP extensions was used (10). For water the CHARMM-specific TIP3P water model force field was applied (11). Bond lengths have been constrained, allowing for an integration time step of 2 fs. Electrostatic interactions were computed using the Particle-Mesh-Ewald method with a Fourier grid spacing of 1.35 Å and a short range interactions cutoff of 11 Å. Lennard-Jones interactions were smoothly switched to zero at 10 Å. The Berendsen thermostat (12) with a reference temperature of 300 K was applied with a time constant of 0.1 ps. For constant pressure simulations a similar algorithm was used for the barostat with a time constant of 1 ps, an external pressure of 1 bar and a system compressibility of $4.5 \cdot 10^{-5} \text{ bar}^{-1}$.

As a result, 7 different systems were then submitted to the same simulation protocol. After addition of ions and water molecules, an energy minimization was performed followed by a 50 ps equilibration in the NVT (constant number of particles, constant volume, constant temperature) ensemble with position restraints applied to heavy atoms of enzyme and peptide. Subsequently the box size was equilibrated in an NPT (constant number of particles, constant pressure, and constant temperature) simulation of 100 ps. The density of the system was then adjusted to the average density of the last 25 ps of this

simulation, before additional equilibration for 2 ns in the NVT ensemble. This was followed by a 20 ns simulation in the NVT ensemble. Ten independent snapshots (coordinates and velocities of the system) were taken every 2 ns from the last 20 ns of each of these simulations. These snapshots were used as starting points of ten microcanonical (NVE, constant number of particles, constant volume, constant energy) simulations, which were used for data evaluation. NVE simulations were extended to 100 ps length with a time step of 2 fs and coordinates and velocities were stored every 10 fs to be used for the analysis of water dynamics. For each system, the distance d between the zinc ion and the closest heavy atom of the substrate is calculated as an average over ten different simulations, i.e. the NVE simulations.

The analysis of water dynamics has been performed by calculating the water-water hydrogen bond auto-correlation functions $C_{HB}(t)$. Hydrogen bond rearrangement dynamics can be characterized via hydrogen bond time auto-correlation functions (12,13)

$$C_{HB}(t) = \left\langle \frac{h(0)h(t)}{h(0)h(0)} \right\rangle,$$

where the operator $h(t)$ gives 1 if a given hydrogen bond is intact at time t and 0 otherwise. A water-water hydrogen bond is considered to be intact, if the oxygen-oxygen distance is smaller than 3.5 Å and the O-H...O angle is larger than 150°. The brackets denote ensemble averaging. Hydrogen bond lifetimes τ_{HB} are defined as the time required for $C_{HB}(t)$ to decay to e^{-1} .


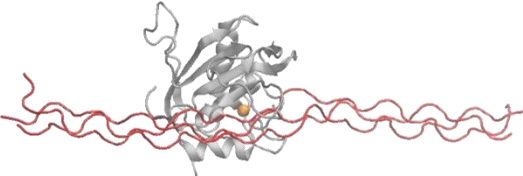

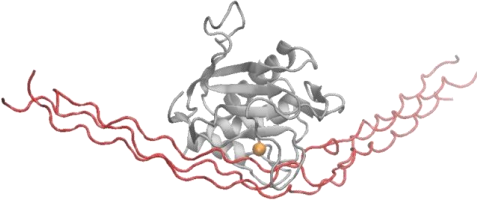
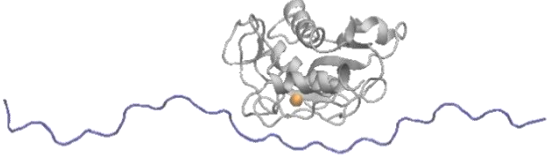
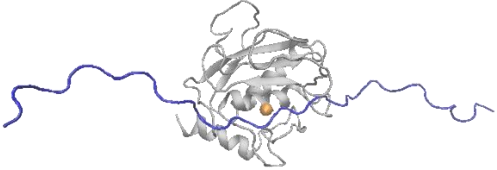
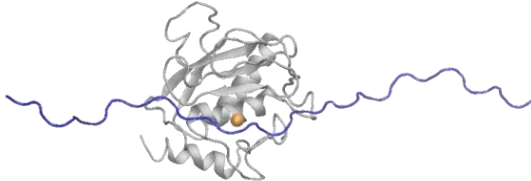
The analysis was carried out for water molecules in the bulk and water molecules solvating the enzyme and the substrate peptide, following a procedure similar to the one we previously performed (13). Water molecules are considered as bulk if their center of mass is more than 10 Å apart from the closest non-solvent atom. The bulk water properties were then compared to water molecules within the hydration shell of the enzyme (i.e. water molecules with less than 3 Å distance to the closest enzyme atom), the substrate peptide (water molecules with less than 3 Å to the closest substrate peptide atom) and the catalytic site of the enzyme (water molecules with less than 6 Å distance to the zinc ion of the catalytic site).

In Fig. S7 the hydrogen bond correlation functions of water in selected investigated systems are shown. As already found by Grossmann et al.(13) the analysis reveals a gradient of water dynamics from the enzyme active site to the other water shells and bulk. Notably, as shown in Fig. S7, for the enzyme-substrate complexes, the HB lifetime of the water molecules solvating the catalytic site does not depend on the enzyme-substrate distance (i.e. the distance d between the zinc ion and the closest heavy atom of the substrate) but only on the specific substrate.

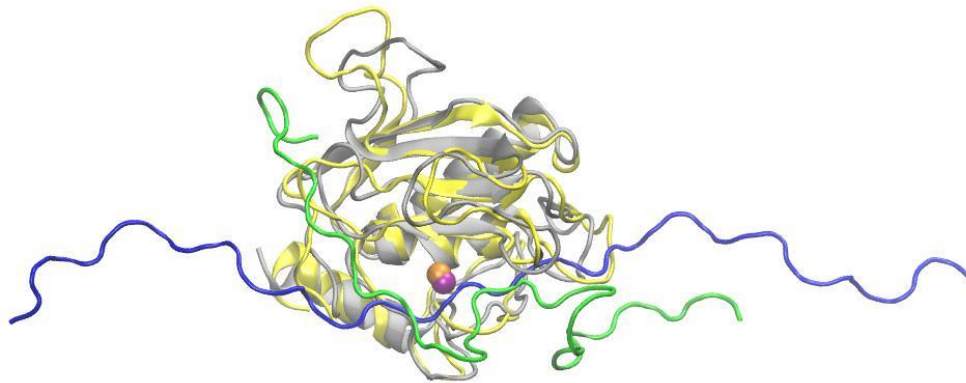
In particular, as also stated in the manuscript, the τ_{HB} of the water molecules solvating the catalytic site is 14.5 ps for the enzyme/gelatin complex and > 20 ps for the enzyme/collagen complex. Notably, $C(t)$ has been calculated as an average over all the

independent runs. In order to compare the errors associated to this analysis, it is convenient to refer to the standard deviation of $C(t)$ at its tail value (i.e. taking into account its average value with respect to the analyses performed over all the independent runs). In the case of the single chain peptide we have: 0.35 ± 0.11 ; while in the case of the triple chain peptide we have: 0.64 ± 0.14 .

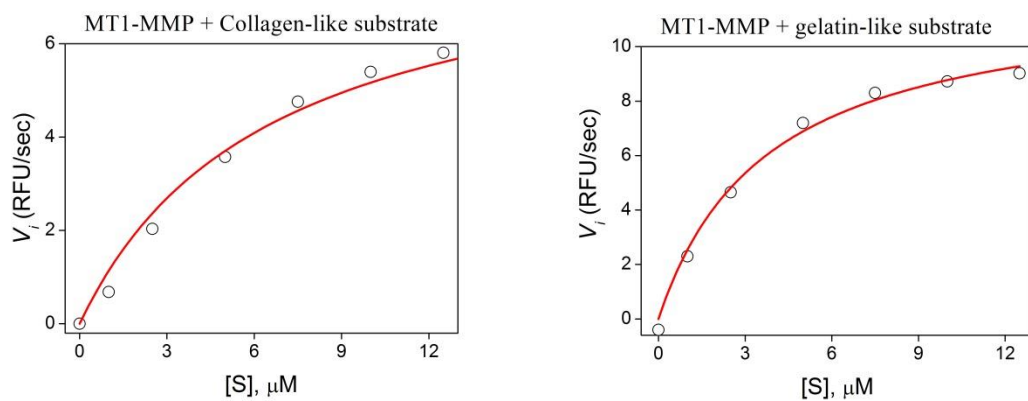
Finally we have also carried out MD simulations of the enzyme/products complex. The triple helical collagen substrate is known to be initially degraded into two triple helical chains: we have therefore modeled a product consisting of these two triple helical chains. The single chain peptide is directly degraded into separated chains. We have performed the water-water HB correlation analyses of the water hydrating the catalytic site in these two systems. Our results, shown in Fig. S8, reveal that product formation does not affect the observed water gradient.

# System	MMP – Collagen	Configuration and distance d (Å)
Initial		Unbound 6.91 ± 0.18
System 1		Stretched 4.87 ± 0.10
System 2		Stretched 2.14 ± 0.03
System 3		Bended 2.01 ± 0.02
MMP-Gelatin		
Initial		Unbound 7.03 ± 0.35
System 1		4.96 ± 0.12
System 2		6.01 ± 0.12

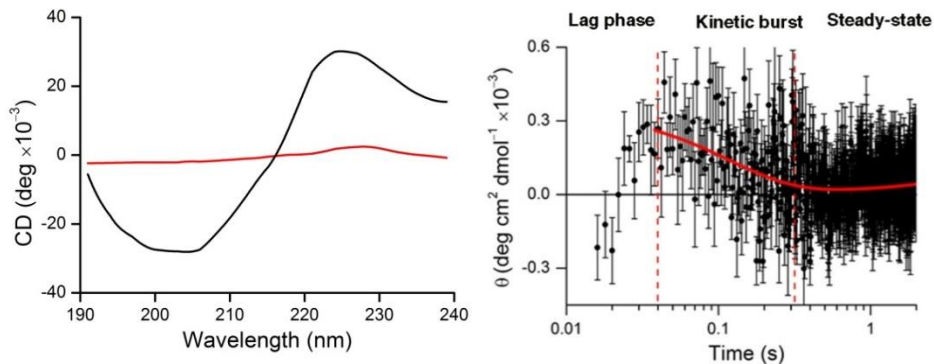
Supplementary Fig. 1. Initial configurations of the enzyme-substrate complexes from molecular dynamics simulations. The enzyme and the substrate are in cartoon representation (enzyme, gray; gelatin, blue; collagen, red; the orange sphere represents the catalytically active zinc ion). For the systems MMP-collagen (system 1,2,3) and MMP-gelatin (system 1,2), also the final configurations are shown (enzyme, yellow; gelatin, green; collagen, green; the purple sphere represents the catalytically active zinc ion). The displayed systems are representative of two molecular scenarios as the substrate approaches the catalytic zinc ion characterized by the distance d between the zinc ion and the closest heavy atom of the substrate: i) the free enzyme ($d \sim 7 \text{ \AA}$) and ii) the enzyme-substrate complex, i.e. the substrate approaching the enzyme active site (d is between 2 and 6 \AA distance). The investigated systems are listed together with d , the distance between the zinc ion and the closest heavy atom of the substrate (\AA). d is calculated as an average over ten different simulations, as explained in the text.



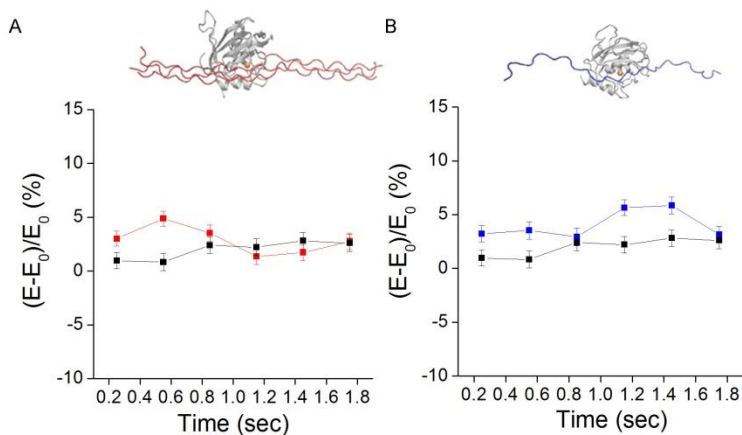
Supplementary Fig 2. The initial and final configurations of the MD simulations for MMP-gelatin (system 1 of Supplementary Fig. 1) are shown. For the initial configuration: enzyme, gray; gelatin, blue; the orange sphere represents the catalytically active zinc ion. For the final configuration: enzyme, yellow; gelatin, green; the purple sphere represents the catalytically active zinc ion. The structure of the gelatin substrate did not remain in its extended conformation and rapidly folded after 2 ns of simulation to form a random coil structure.



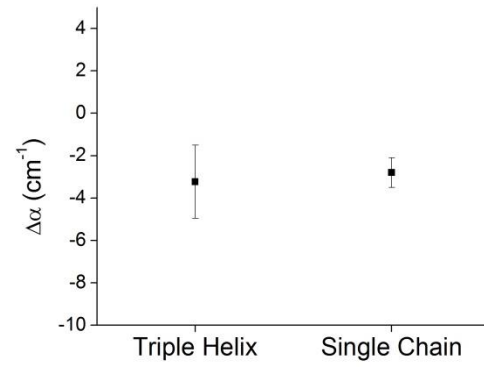
Supplementary Fig 3. Steady-state Michaelis-Menten kinetics of MT1-MMP with the collagen-like (left) and gelatin-like (right) substrates. Initial rates were obtained by fitting the data to the Michaelis-Menten equation: $V_i = V_{\max} \times [S] / ([S] + K_M)$ (RFU=Relative fluorescence units).



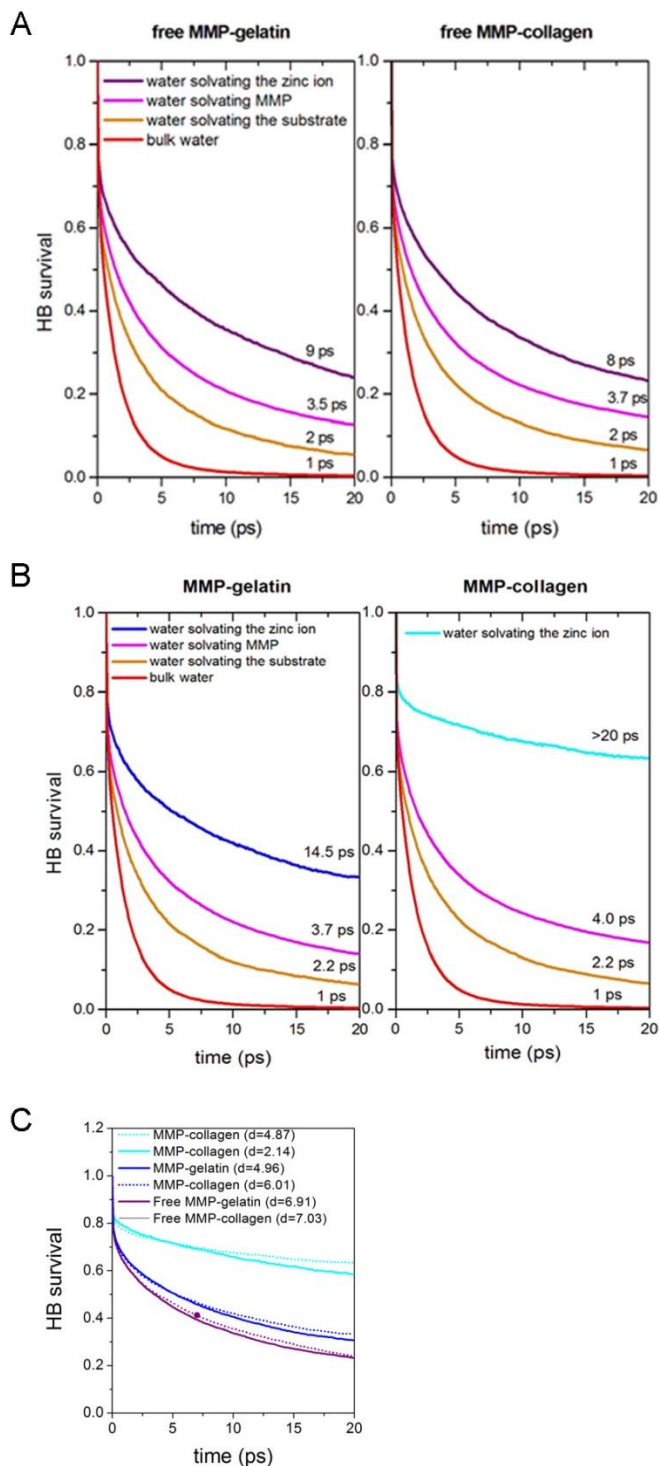
Supplementary Fig 4. Time resolved CD. Left: steady-state circular dichroism (CD) of the collagen-like (black) and gelatin-like (red) substrates. Right: Time-resolved CD analysis of the collagen-like substrate upon mixing with MT1-MMP (black). At 205 nm, the collagen-like substrate exhibits an increase in molar ellipticity between 15 to 30 ms (dead-time of the stopped-flow is 15 ms), which may indicate on substrate unwinding, followed by an exponential decay due to triple helix hydrolysis. Shown is the mean of 15 independent kinetic measurements in the notation Mean \pm SD.



Supplementary Fig. 5. Perturbation of the protein-water coupled motions beyond reaction steady-state within the first 2s of the reaction. Kinetic THz absorption data collected during hydrolysis of triple-helical substrate (A, red), single-stranded substrate (B, blue), and data of mixing only buffer (black). At 250ms, the absorption of the solution of both MMP-substrate solutions is decreased compared to the buffer. The THz absorption of the MMP-triple helical substrate reaction mixture reaches a transient minimum with a relative change of 3 – 4% between 400ms and 700ms, followed by an increase until 1.15 ± 0.15 s. For the gelatin-like single stranded substrate there are two absorption minima at 550 ± 150 ms and 1.30 ± 0.15 s.

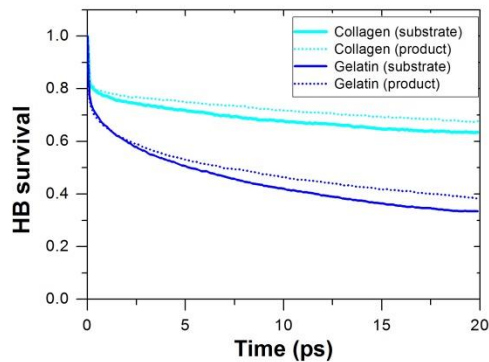


Supplementary Fig. 6: THz absorption measurement of 1mM triple helix collagen substrate or single chain gelatin substrate in the reaction assay buffer. The corresponding buffer absorbance was subtracted to yield $\Delta\alpha = \alpha_{\text{Protein+buffer}} - \alpha_{\text{Buffer}}$. Plotted are the averaged values of 5-10 independent measurements consisting of ~1000 data points each. No statistically significant differences were detected between the two substrates.



Supplementary Fig. 7. a) HB correlation functions of water in the free MMP-gelatin (left) and free MMP-collagen (right) system for bulk water (red; distance to closest protein atom $> 10 \text{ \AA}$), water molecules solvating the substrate (orange; distance to closest enzyme atom $< 3 \text{ \AA}$), water molecules solvating the enzyme (magenta; distance to closest substrate atom $< 3 \text{ \AA}$), and water molecules solvating the zinc ion of the catalytic site (purple; distance to zinc catalytic ion $< 6 \text{ \AA}$). b) HB correlation functions of water in the MMP-gelatin (left) and MMP-collagen (right) system for bulk water (red; distance to closest protein atom $> 10 \text{ \AA}$), water molecules solvating the substrate (orange; distance to closest enzyme atom $< 3 \text{ \AA}$), water molecules solvating the enzyme (magenta; distance to closest substrate atom $< 3 \text{ \AA}$), and water molecules

solvating the zinc ion of the catalytic site (blue for the MMP-gelatin and cyan for the MMP-collagen; distance to zinc catalytic ion $< 6 \text{ \AA}$). c) HB correlation functions of water solvating the catalytic site (distance to zinc catalytic ion $< 6 \text{ \AA}$) for selected simulated systems, characterized by different distances between the zinc ion and the closest heavy atom of the substrate (as listed in Supplementary Fig. 1).



Supplementary Fig. 8. **Product formation does not change water dynamics gradients.** HB correlation functions of water in the active site (distance to zinc catalytic ion $< 6 \text{ \AA}$) of the complexes formed between the enzyme and the collagen (cyan) or gelatin (blue) substrates (straight line). In comparison, displayed are the HB correlation functions of the complexes formed with the product (dot lines). No statistically significant differences are detected between the water dynamics in the presence of substrate or product for each of the substrates.

Supplementary tables

Time (ms)	ΔE_0 (eV)	N	R (Å)	σ^2 (Å ²)
0	1.30±0.40	4.0±0.5	1.99±0.01	1.9E-3±6.6E-4
15	0.54±0.55	4.84±0.38	2.00±0.02	5.1E-4±1.8E-3
60	2.64±0.63	4.89±0.44	2.02±0.02	5.3E-3±2.1E-3
80	2.39±0.83	4.98±0.59	2.01±0.03	6.0E-3±2.8E-3
100	1.61±0.62	4.92±0.44	2.01±0.02	4.4E-3±1.9E-3
150	-0.35±0.54	4.86±0.36	1.99±0.01	3.2E-3±1.5E-3
180	1.36±0.99	5.00±0.71	2.00±0.03	5.4E-3±3.2E-3

Supplementary Table 1. The raw X-ray absorption spectra during a single catalytic cycle of collagen-like substrate hydrolysis by MT1-MMP were processed and analyzed following reported procedures (14). Presented are the best-fit parameters of the first coordination shell of the residual spectra resulting from iterative subtractions of fractions of the starting phases ($t = 0$) where N is the coordination number, R is the zinc-ligand bond distance in Å and σ^2 is the Debye-Waller factor. The penta-coordinated intermediate is detected at 15 to 180 ms.

Time (ms)	ΔE_0 (eV)	N	R (Å)	σ^2 (Å ²)
0	1.30±0.40	4.0±0.5	1.99±0.01	1.9E-3±6.6E-4
30	-1.21±0.54	4.89±0.37	1.98±0.01	2.9E-3±1.4E-3
50	-0.57±0.66	3.89±0.38	2.00±0.01	1.8E-3±1.2E-3
60	-0.83±0.67	4.96±0.47	1.98±0.02	3.8E-3±1.9E-3
100	1.58±0.62	3.79±0.48	2.00±0.01	3.7E-3±1.8E-3
150	1.29±0.99	4.63±0.67	1.99±0.01	3.2E-3±1.9E-3
200	2.69±0.73	4.99±0.52	2.00±0.01	3.7E-3±1.7E-3

Supplementary Table 2. The raw X-ray absorption spectra during a single catalytic cycle of gelatin-like substrate hydrolysis by MT1-MMP were processed and analyzed following reported procedures (14). Fluctuations in zinc coordination number are detected, before a stable penta-coordinated intermediate is formed at ~200 ms.

References:

1. Meister K, *et al.* (2013) Long-range protein-water dynamics in hyperactive insect antifreeze proteins. *Proc Natl Acad Sci U S A* 110(5):1617-1622.
2. Pappu RV, Marshall GR, & Ponder JW (1999) A potential smoothing algorithm accurately predicts transmembrane helix packing. *Nat Struct Biol* 6(1):50-55.
3. Ponder JW (1998) TINKER: Software Tools for Molecular Design, Version 3.6. (School of Medicine, Washington University).
4. Wang JM, Cieplak P, & Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry* 21(12):1049-1074.
5. Grossman M, *et al.* (2010) The intrinsic protein flexibility of endogenous protease inhibitor TIMP-1 controls its binding interface and affects its function. *Biochemistry* 49(29):6184-6192.
6. Iyer S, Visse R, Nagase H, & Acharya KR (2006) Crystal structure of an active form of human MMP-1. *J Mol Biol* 362(1):78-88.
7. Tochowicz A, *et al.* (2011) The dimer interface of the membrane type 1 matrix metalloproteinase hemopexin domain: crystal structure and biological functions. *J Biol Chem* 286(9):7587-7600.
8. Van der Spoel D, *et al.* (2005) GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry* 26(16):1701-1718.
9. Mackerell AD, Feig M, & Brooks CL (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry* 25(11):1400-1415.
10. Buck M, Bouguet-Bonnet S, Pastor RW, & MacKerell AD, Jr. (2006) Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *Biophys J* 90(4):L36-38.
11. Bjelkmar P, Larsson P, Cuendet MA, Hess B, & Lindahl E (2010) Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *Journal of Chemical Theory and Computation* 6(2):459-466.
12. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, & Haak JR (1984) Molecular-Dynamics with Coupling to an External Bath. *Journal of Chemical Physics* 81(8):3684-3690.
13. Grossman M, *et al.* (2011) Correlated structural kinetics and retarded solvent dynamics at the metalloprotease active site. *Nat Struct Mol Biol* 18(10):1102-1108.
14. Solomon A, Akabayov B, Frenkel A, Milla ME, & Sagi I (2007) Key feature of the catalytic cycle of TNF-alpha converting enzyme involves communication between distal protein sites and the enzyme catalytic core. *Proc Natl Acad Sci U S A* 104(12):4931-4936.