

Mapping intermolecular interactions and active site conformations: from human MMP-1 crystal structure to molecular dynamics free energy calculations

Anthony Nash, Helen L. Birch & Nora H. de Leeuw

To cite this article: Anthony Nash, Helen L. Birch & Nora H. de Leeuw (2017) Mapping intermolecular interactions and active site conformations: from human MMP-1 crystal structure to molecular dynamics free energy calculations, Journal of Biomolecular Structure and Dynamics, 35:3, 564-573, DOI: [10.1080/07391102.2016.1153521](https://doi.org/10.1080/07391102.2016.1153521)

To link to this article: <https://doi.org/10.1080/07391102.2016.1153521>



© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Accepted author version posted online: 17 Feb 2016.
Published online: 25 Apr 2016.



[Submit your article to this journal](#)



Article views: 1388



[View related articles](#)





[View Crossmark data](#)



Citing articles: 7 [View citing articles](#)

Mapping intermolecular interactions and active site conformations: from human MMP-1 crystal structure to molecular dynamics free energy calculations

Anthony Nash^{a*}, Helen L. Birch^b  and Nora H. de Leeuw^{a,c} 

^aDepartment of Chemistry, University College London, London, UK; ^bInstitute of Orthopaedics & Musculoskeletal Science, University College London, London, UK; ^cSchool of Chemistry, Cardiff University, Cardiff, UK

Communicated by Ramaswamy H. Sarma

(Received 19 December 2015; accepted 9 February 2016)

The zinc-dependent Matrix Metalloproteinases (MMPs) found within the extracellular matrix (ECM) of vertebrates are linked to pathological processes such as arthritis, skin ulceration and cancer. Although a general backbone proteolytic mechanism is understood, crystallographic data continue to suggest an active site that is too narrow to encompass the respective substrate. We present a fully parameterised Molecular Dynamics (MD) study of the structural properties of an MMP-1-collagen crystallographic structure (Protein Data Bank – 4AUO), followed by an exploration of the free energy surface of a collagen polypeptide chain entering the active site, using a combined meta-dynamics and umbrella sampling (MDUS) approach. We conclude that the interactions between MMP-1 and the collagen substrate are in good agreement with a number of experimental studies. As such, our unrestrained MD simulations and our MDUS results, which indicate an energetic barrier for a local uncoiling and insertion event, can inform future investigations of the collagen-peptide non-bonded association steps with the active site prior to proteolytic mechanisms. The elucidation of such free energy barriers provides a better understanding of the role of the enzyme in the ECM and is important in the design of future MMP inhibitors.

Keywords: collagenase; proteinase; metzincin; matrix-metalloproteinase; zinc; free energy

Introduction

Matrix metalloproteinases (MMPs) are a family of 24 zinc-dependent endoproteinases found within the extracellular matrix (ECM) of vertebrates of which 23 types are present in humans (Chaudhary et al., 2010; Overall & López-Otín, 2002; Sekhon, 2010; Vihinen & Kähäri, 2002). Owing to their physiological ability to degrade major structural molecules in the ECM, MMPs are known to play important roles in mechanisms such as wound healing, tissue remodelling, morphogenesis and embryogenesis (Duffield, Lupher, Thannickal, & Wynn, 2013) and are implicated in a number of cellular functions, such as cell migration, proliferation and apoptosis (Thiennu & Werb, 2000). They have also been linked to pathological processes such as arthritis (Mitchell et al., 1996; Reboul, Pelletier, Tardif, Cloutier, & Martel-Pelletier, 1996; Wernicke, Seyfert, Hinzmann, & Gromnica-Ihle, 1996), skin ulceration (Sekhon, 2010) and cancer progression (Bernhard, Gruber, & Muschel, 1994; Deryugina & Quigley, 2006; Liotta, Abe, Robey, & Martin, 1979; Liotta et al., 1980). As cancer is a leading cause of mortality and morbidity, (Jemal et al., 2004) and MMPs are implicated in several stages of tumour progression, MMPs are considered promising targets for

drug delivery mechanisms and therapeutics, such as engineered-selective tissue inhibitors (Brew & Nagase, 2010).

The general structure of MMPs consists of several distinct domains conserved between types. All types, except MT-MMPs (membrane type – MMPs) which are membrane-bound, begin with a cell secretion leader or prodomain at the N-terminus, followed by a highly conserved propeptide, approximately 80 amino acids in length. Cleavage of the propeptide domain results in the separation of a conserved cysteine switch from the catalytic zinc ion, enabling enzymatic activity (Stricklin, Jeffrey, Roswit, & Eisen, 1983). The catalytic domain (CAT) (Figure 1(A)) is approximately 170 amino acids in length and contains the conserved zinc-binding domain amino acid motif HEXXHXXGXXH (Van Wart & Birkedal-Hansen, 1990) (Figure 1(B)). Three histidine residues and a water molecule make up the complete coordination sphere of the catalytic zinc ion. A second zinc ion is found buried close to the rear of the active site cavity (Figure 1(C)), with a coordination sphere consisting of three histidine residues and an aspartic acid residue. A C-terminal hemopexin-like (HPX) domain (Figure 1(E)) of approximately 200 residues in length,

*Corresponding author. Email: a.nash@ucl.ac.uk

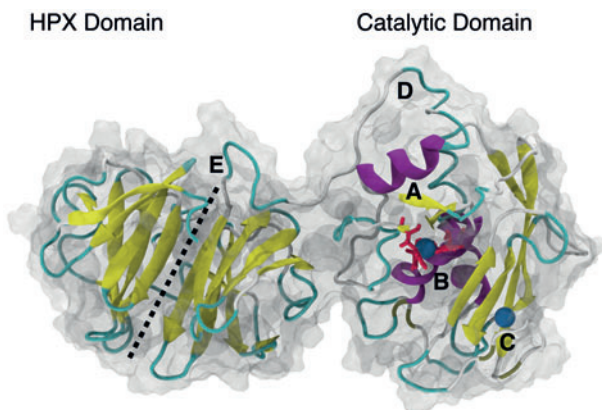


Figure 1. A secondary structure representation of the MMP-1 4AUO molecular structure. (A) The structure begins with the N-terminus at the catalytic domain, followed by the catalytic active site zinc-binding domain (B) and the structural zinc domain (C). Both zincs are shown as spheres and the three coordinating histidine residues of the catalytic zinc as red sticks. The coiled hinge linker region begins at (D) and ends at the beginning of the HPX domain (E). A dashed line indicates the orientation of the β -propeller funnel in the HPX domain.

arranged as a fourfold symmetrical beta-propeller structure (bladed regions 1–4) around a funnel-like opening, is connected to the CAT domain by a short hinge linker region (Figure 1(D)). Finally, the HPX domain has been implicated as a binding mediator for the degradation of ECM substrates (Bode, 1995; Gomis-Rüth et al., 1996).

The proteolytic mechanism of MMP is thought to depend on the electrostatic attraction between the backbone carbonyl group of an ECM protein with the Zn ion in the active site cavity. Of particular interest to this study is the binding of the collagenases, MMP1, MMP8 and MMP13 to a collagen molecule: a trimeric coiled-coil arrangement of helical polypeptide collagen alpha chains. The zinc ion becomes penta-coordinated during a cascade of nucleophilic rearrangements between a conserved glutamic acid, a carbonyl group from a collagen polypeptide chain backbone and a water molecule (Verma & Hansch, 2007). Quantum mechanical calculations reveal that the presence of water reduces the electrophilic activation barrier of the carbonyl oxygen by approximately 5 kcal/mol, thus increasing the likelihood of a reaction (Pelmenchikov & Siegbahn, 2002).

Even with the current understanding of the proteolytic mechanism of zinc-dependent metzincins, the inceptive period of protein insertion into the active site of MMP, prior to any proteolytic activity, remains unknown. From crystallographic observations of the catalytic domains of MMP-1 (Borkakoti et al., 1994; Manka et al., 2012) and MMP-8 (Bode et al., 1994) and the full length of MMP-1 (Li et al., 1995), the entrance to the active site appears too narrow for the width of a triple-helical collagen molecule. An early report suggests that

the HPX domain might bind to collagen through a folding mechanism of the hinge linker region. Not only would this expose the scissile bond of collagen to the active site for a prolonged period of time but favourable interactions between the HPX domain and the collagen residues could also facilitate collagen unwinding (Bode, 1995; Bode et al., 1994; Gomis-Rüth et al., 1996; Ottl et al., 2000). Interestingly, a study by Zhao et al., yielded atomic force microscopy images of a potential fold mechanism in MT1-MMP (Zhao et al., 2015). However, this mechanism is at odds with a recent crystal structure of collagen-bound MMP-1, which shows no indication of a folding type mechanism (Manka et al., 2012). Furthermore, a study based on the unwinding and enzymatic cleavage of collagen confirmed that although inactive, mutated enzyme MMP-1 (E200A) induces a change to the collagen structure (Chung et al., 2004). In particular, collagen digestion was detected after an active secondary cutter had been introduced. It was remarked that the discovery of fragmented collagen from a secondary cutter, typically ineffective until an inactive MMP-1 has been introduced, could rule out the possibility that MMPs themselves undergo the conformational change required to accommodate the substrate. Finally, it would be reasonable to suggest that a conformational folding of the MMP-1 could sterically hinder the secondary cutter from accessing the scissile collagen bond.

An alternative mechanism proposed by de Souza et al., suggests that the collagenase hinge linker, rich in proline residues, adopts a transient poly-Pro II-like helical structure. Upon interacting with collagen, by forming a proline-zipper, the hinge domain competes for a single collagen polypeptide by disrupting the native collagen fold (de Souza, Pereira, Jacchieri, & Brentani, 1996). In the case of MMP-1, the flexibility in the hinge region is attributed to G272 (Tsukada & Pourmotabbed, 2002).

How MMP-1 facilitates collagen unwinding and collagen insertion remains unclear. Whether these events are dependent on conformational dynamics of the enzyme itself, or whether the enzyme retains a consistent structure is also open to debate. Here, we use a Molecular Dynamics (MD) modelling approach to provide further insight into the structure and dynamics of the active site of MMP-1. Firstly, we report on the structural stability of the MMP-1 in complex with a model collagen protein from crystallographic coordinates (4AUO). This also helps to identify any issues with the force-field parameterisation of the protein and the zinc ions. Secondly, the MD model from the first stage is used to elucidate the free energy of insertion and non-bonded association with the active site of a single collagen peptide, the initial event prior to later proteolytic activity, using a combined molecular dynamics and umbrella sampling (MDUS) approach. Given that this manuscript focuses on non-bonded interactions and without the added complexity of

covalent bond rearrangement, free energy calculations derived from MD principles are appropriate for a protein of this size.

Computational methods

Simulation details

Atomistic MD simulations were performed using the GROMACS simulation package, version 5.0.5 (Pronk et al., 2013). The time step used in the simulations was 2 fs, and was integrated using the leapfrog algorithm. All bonded interactions were holonomically constrained using the LINCS constraint algorithm (Hess, Bekker, Berendsen, & Fraaije, 1997). The short-range neighbour interaction list cut-off was fixed to 0.8 nm and updated every 10 fs. Short-range interactions were modelled using a 12-6 Lennard–Jones potential and were truncated at 0.8 nm. Short-range electrostatic interactions were truncated at 0.8 nm and long-range electrostatics were calculated using the Particle Mesh Ewald scheme (Darden, York, & Pedersen, 1993). The pressure was maintained at one atmosphere by applying an isotropic coupling Parrinello–Rahman scheme (NPT ensemble) (Parrinello & Rahman, 1981). The temperature was maintained at 294.5 K using the Nose–Hoover thermostat (NVT and NPT ensemble) (Hoover, 1985; Nosé, 1984). Loose coupling was applied between the protein complex and the water environment. Periodic boundary conditions were applied in three dimensions. Coordinates, velocities and energies were saved every 100 ps.

Meta-dynamics were performed using the PLUMED 2.0.2 plugin for Gromacs (Tribello, Bonomi, Branduardi, Camilloni, & Bussi, 2014). The height and width of the bias potentials were maintained at 0.25 kJ/mol and 1.5 Å, respectively. Gaussian potentials were deployed every 250 MD step calculations (500 fs). Umbrella sampling was performed along an absolute distance reaction coordinate defined by the meta-dynamics trajectory (see results and discussion). Fourteen individual umbrella windows were set up to cover the distance interval from 0.38 nm to 1.5 nm and each was equilibrated for 1 ns by applying the above pressure and temperature simulation details whilst subjecting the protein atoms and zinc ions to a 1000 kJ/mol position restraint. An umbrella potential of 1000 kJ/mol in three dimensions was applied between the scissile hydroxyl group oxygen on the collagen peptide backbone and the catalytic zinc ion. Individual umbrella simulations were performed for 20 ns. A total of 380 ns of simulation was performed to adequately sample the reaction coordinate. Force constants were recorded every 2 ps, totalling 10,000 records per simulation. A potential of mean force (PMF) profile was constructed from the individual umbrella windows using the weighted histogram analysis method (WHAM) (Kumar, Rosenberg, Bouzida, Swendsen, & Kollman, 1992). A

detailed description of the meta-dynamics and umbrella sampling methodology, including convergence of free energy, can be found in the Supplementary Information.

Derivation of partial charges

In order to accurately model the two zinc-binding domains, zinc and zinc-ligand partial charges were calculated. Two zinc-dependent coordination models were constructed from the crystal structure geometry. The first comprised residues ^{199}H , ^{203}H and ^{209}H in addition to the catalytic zinc and the second comprised residues ^{149}H , ^{164}H , ^{177}H and ^{151}D and the structural zinc. A single water molecule was added to the extracted zinc catalytic coordination model to make up a full coordination shell as would be present *in vivo*. Hydrogen atoms were added to the extracted crystal structure residues and partial backbones were terminated according to the R.E.D (RESP ESP charge Derived) Server instructions for restrained partial charge fitting. Partial charges were fitted to the Amber99SB-FF using PyR.E.D (Vanqualef et al., 2011). Geometries were optimised at the Hartree-Fock level of theory with a 6-31G* basis-set to keep in line with the philosophy of the Amber force field.

Model construction

Having defined the simulation parameters and derived partial charges for both zinc-binding sites, a MMP-1 collagen-bound protein complex was generated from crystal structure coordinates using the pdb2gmx program. Protonation of nitrogen epsilon or nitrogen delta of histidine sidechains was achieved at random to represent the almost equal propensity for protonation in experimental systems. The *ab initio* optimised electronic structure of both structural and catalytic zinc-binding domains was used to replace the respective residues in the newly generated atomistic structure. Bond distances and angles were constrained between zinc and their respective nitrogen and oxygen coordination donors. Force constants were taken from Tuccinardi et al. (2006). Partial charges of each backbone fragment and accompanying sidechain along with the zinc ion were replaced with the R.E.D. derived partial charges. The partial charge on heavy atoms of the main functional group per side chain ligand, in addition to the intermolecular distance between coordinating side chain atoms with each zinc ion, can be seen in Figure 2 (see Supplementary Information for a complete list of partial charges, including charges for hydrogen atoms and the amino acid backbone). The new partial charges over the coordinating zinc ligands and each zinc ion ensure the compatibility with computational techniques used to compute free energy values. The unit cell, adjusted to prevent periodic image artefacts, was fully solvated using TIP4P-ew water

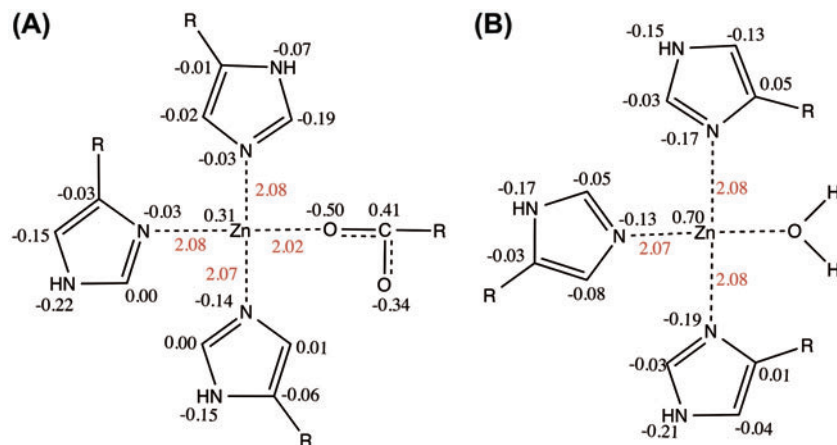


Figure 2. Derived partial charges on the zinc ion and accompanying ligand side chains of (A) the structural zinc coordination sphere and (B) the catalytic zinc coordination sphere.

molecules enhanced for PME calculations (Horn et al., 2004). Calcium ions were added to yield a net neutral charge.

The protein complex and water environment were subjected to a steepest descent energy minimisation to correct for overlapping atoms. The maximum force of 1000 kJ/mol on any one atom was used as an energy minimisation stopping criteria. To ensure that the temperature of the system had converged to 294.5 K before applying the pressure-coupling scheme, a short 200 ps simulation using the constant number of particles, volume and temperature (NVT) ensemble was performed. During this short simulation, harmonic position restraints of 1000 kJ/mol were applied to all protein heavy atoms. The position restraints were maintained and the system was subjected to a 200 ps simulation at constant pressure using the Parrinello–Rahman pressure-coupling scheme (NPT ensemble). Position restraints were reduced to 100 kJ/mol for an additional 200 ps. A final 200 ps were added with the position restraints reduced to 10 kJ/mol. All position restraints were finally lifted and the system ran for 50 ns. During the unrestrained 50 ns, the pressure, temperature and total energy had all converged at a very early stage.

Results and discussion

The MMP-1 collagen complex

To prepare the MMP-1 model, new parameters were defined for the catalytic and structural zinc ions along with their respective coordination donor residues. The *ab initio* calculations were performed with a water molecule present to act as the fourth coordination donor. It is understood that water coordinates with ²⁰⁰E in the active site of the wild type, acting as a nucleophilic agent during the proteolytic reaction (Pelmenschikov & Siegbahn,

2002). It must be noted, that the protein sequence, 4AUO, used in this study contains the experimental substitution E200A.

After inserting the optimised zinc-coordination geometry into the structural and catalytic domains, the single catalytic water molecule was removed from the *ab initio* output. To do otherwise would have resulted in either a bonded water model that prevented an exchange of water ligands during the simulation, or a non-bonded model that could result in a unique water molecule escaping into the bulk. The small difference between the derived partial charge and the forcefield partial charge of water suggests that water native to the forcefield would suffice as a zinc ligand. The structural zinc ion along with ligands ¹⁶⁴H, ¹⁴⁹H, ¹⁷⁷H and ¹⁵¹D showed very little deviation from the crystal structure.

Properties of the intermolecular association interface

An unrestrained production simulation was performed for 50 ns. There was little in the way of structural deviation of the MMP-1 production run from the original crystallographic structure (Figure 3(A)), nor did alteration to the secondary structure occur (see Supplementary Figure 1), indicating that the system had equilibrated. We therefore inferred that this duration was sufficient to explore the intermolecular associations between enzyme and substrate.

After equilibration, the polarity between interfacial residues within a 5 Å cut-off was evaluated. There was a consistent match in polarity at the interface between collagen and the HPX domain: ²⁸⁵R, ²⁸⁹F, ²⁹⁰Y, ²⁹²E, ²⁹⁴E, ²⁹⁶N, ²⁹⁷F, ³⁰⁰V, ³⁰¹F, ³¹⁶A, ³¹⁷D, ³¹⁹D, ³³⁴G and ³³⁵Q. Interestingly, ²⁹⁰Y and ²⁸⁹F protruded from the polar HPX domain in the shape of a small clasp, and gradually shifted across the interfacial region of the collagen.

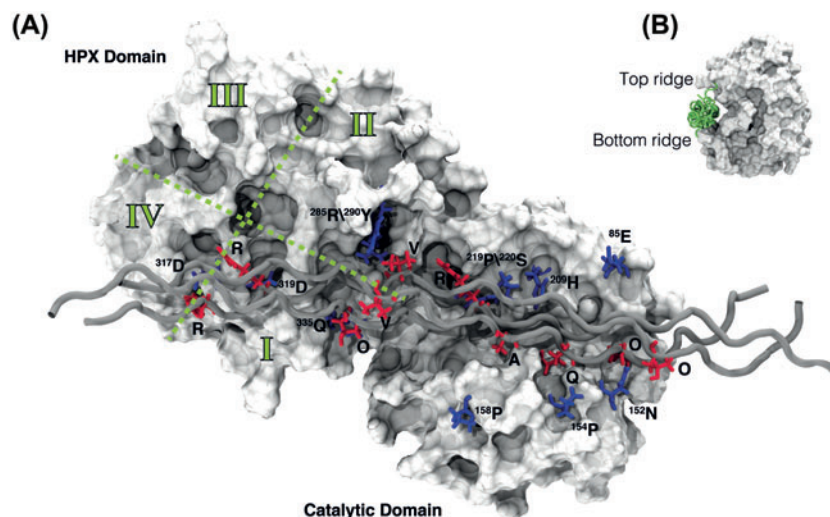


Figure 3. (A) The model collagen triple helical protein bound to MMP-1. Residues identified to fulfil the criteria for potential intermolecular hydrogen bonding. MMP-1 residues have been illustrated in purple and collagen residues in red. The HPX domain can be identified by the globular region on the left, and the CAT domain as the globular region on the right. The narrowing protein density between both domains indicates the hinge linker. The collagen protein is represented by a ribbon, and MMP-1 as a surface. (B) Side view of the catalytic domain bound to collagen. The top and bottom ridges are clearly visible. The first frame from the unrestrained 50 ns simulation was used for illustrative purposes only.

Besides this subtle rearrangement, the paired protein complex was remarkably stable. The polar interface of HPX was followed by a small number of non-polar HPX interfacial residues adjacent to the MMP-1 hinge linker region. These corresponded with the non-polar nature of the accompanying stretch of collagen turns.

The interfacial residues of the CAT domain found within 5 Å of collagen were: ⁸⁵E, ¹⁴⁹H, ¹⁵¹D, ¹⁵²N, ¹⁵³S, ¹⁵⁴P, ¹⁵⁵F, ¹⁵⁶D, ¹⁵⁸P, ¹⁶⁰G, ¹⁶¹N, ¹⁶⁴H, ¹⁶⁶F, ¹⁶⁷Q, ¹⁷⁰P, ¹⁷²I, ¹⁷⁷H, ¹⁹¹Y, ¹⁹⁹H, ²⁰³H, ²⁰⁷L, ²⁰⁹H, ²¹⁷M, ²¹⁸Y, ²¹⁹P, ²²⁰S, ²²¹Y, ²⁷⁰T, ²⁷¹I, ²⁷²R, ²⁷³G and ²⁷⁴E. The cavity *ridges* (Figure 3(B)) at the top and bottom of the active site were composed predominantly of polar residues glutamine, glutamic acid, histidine, serine and asparagine. The collagen motif G-L-A, revealed to be the collagen cut site (Chung et al., 2004), remained within close proximity of the active site for the entirety of the simulation.

A thorough and verbose description of a network of hydrogen bonds between the MMP-1 and collagen interface is documented in the Supplementary Information. In summary, the simulated dynamics do not deviate far from the initial crystal structure, providing a stable configuration from which a set of free energy calculations can be derived with confidence.

The free energy of local collagen insertion into the active site

A 200 ns meta-dynamic MD simulation was performed using the final time step from the simulation described above. The absolute distance between the scissile bond

and the catalytic zinc ion made up the reaction coordinate. It has been noted that at physiological temperature, the numerous non-polar residues surrounding the collagenase cleavage site of the polypeptide chain destabilise and become vulnerable to proteolysis (Fields, 1991; Leikina, Merts, Kuznetsova, & Leikin, 2002). This made the proposed reaction coordinate an apt starting point for investigating collagen insertion and association with the MMP-1 active site prior to proteolytic activity. During the simulation, the unwinding of the scissile polypeptide chain from the remaining collagen protein was almost an immediate effect. The relaxation of the N-terminal collagen structure could have influenced the free energy surface (FES) landscape during the separation of a single polypeptide collagen chain. However, free energy calculations of collagen-I have yielded metastable states of unwound collagen local to the scissile bond (Nerenberg & Stultz, 2008) An increase in Δ FES by approximately 10.5 kJ/mol (2.509 kcal/mol) resulted in a peak from the initial non-inserted structure (Figure 4(A)) before a significant free energy descent by Δ FES of -44.7 kJ/mol (-10.686 kcal/mol) upon entry of a collagen polypeptide chain into the active site cavity (Figure 5). This minimum resulted in coordination between the carbonyl group from glycine and the zinc at an approximate distance of 0.3 nm.

The inserted peptide chain was also shown to form a tightly packed association within the active site, notably through a match in polarity with a beta-sheet, a helix, and the accompanying random coils, which make up the rear of the active cavity.

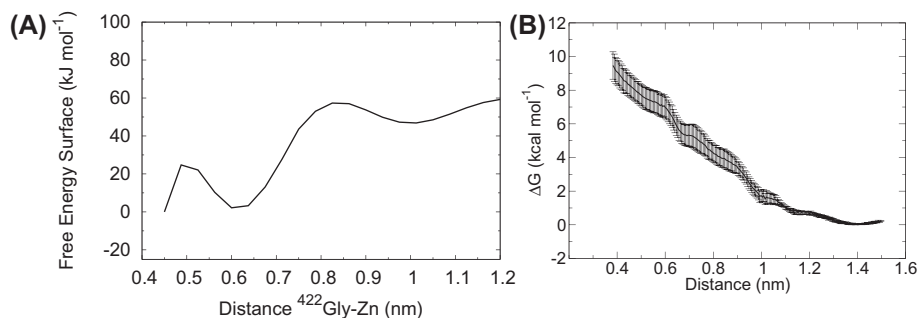


Figure 4. (A) The absolute distance-based FES taken between the carbonyl oxygen atom of glycine from the scissile sequence of a collagen peptide and the catalytic zinc ion from a 200 ns meta-dynamics simulation. (B) The PMF profile based on the recovered trajectory from the earlier meta-dynamics study of the collagen scissile bond approaching the catalytic zinc. Error estimates were generated using bootstrap analysis.

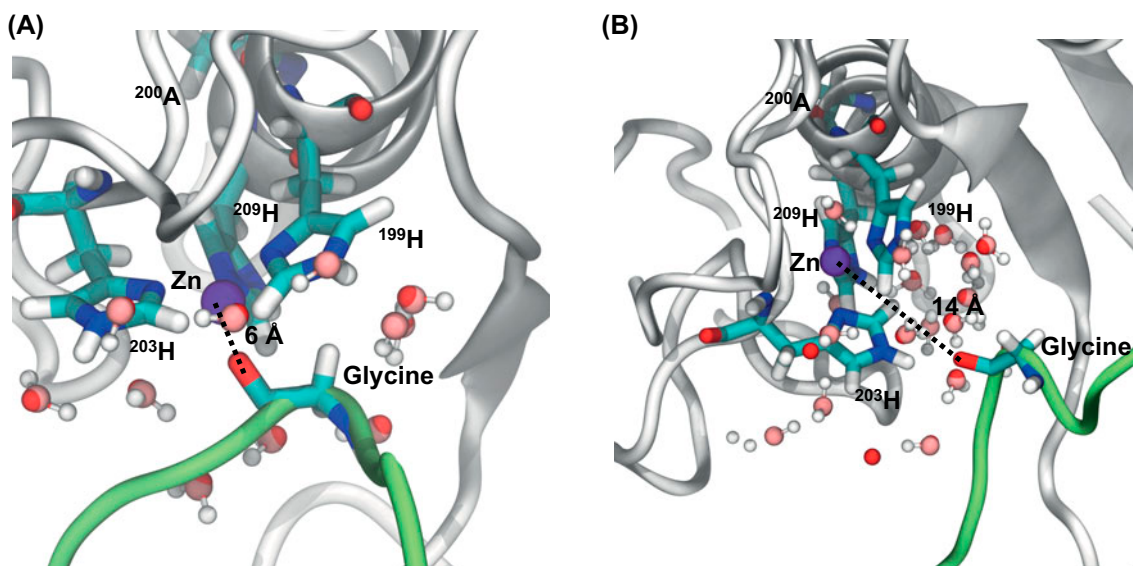


Figure 5. A representative structure extracted from the inflection point, approximately 0.65 nm (A), and from the minimum (B), approximately 1.4 nm, along the reaction coordinate, respectively. The collagen peptide is in green ribbon; the surrounding structure of the CAT domain is represented in structural illustrations (sheets, helices and turns). E200A substitution has been identified, situated on the α -helix directly behind the coordinating histidine residues. Water within 8 Å of the zinc ion has been identified explicitly.

The meta-dynamics study was augmented by performing umbrella sampling using recovered frames that sampled the absolute distance between 0.38 and 1.5 nm from the scissile bond to the catalytic zinc ion. This range does not include a representation of the FES minimum 0.3 nm along the reaction coordinate, and it became clear early on that umbrella windows at that distance resulted in an exponential growth in the PMF profile as approaching atoms began to repel one another. Fourteen individual umbrella simulations were sufficient to sample the absolute distance-based reaction coordinate. Individual umbrella windows were simulated for 20 ns and the first nano-second was used as equilibration of the system and therefore disregarded from WHAM calculations. It is immediately obvious that the FES from the meta-dynam-

ics analysis is very different from the final PMF profile (Figure 4(B)). It is reasonable to assume that the meta-dynamics simulation sampled degenerate energy states of unique structures at identical points along the single reaction coordinate. The MDUS combined approach though, would take into account all of these different conformations within the confines of the reaction coordinate during the construction of the PMF profile. The greatest level of ΔG deviation was approximately 5.65 kJ/mol (1.35 kcal/mol) as the collagen peptide was within 0.4 nm of the zinc ion. Further from the zinc ion, within the region of 1.2–1.45 nm, there is almost no deviation from the predicted ΔG . The increase in ΔG from the coiled collagen conformation, 1.4 nm along the reaction coordinate, towards the inserted collagen peptide configuration is

indicative of a system approaching an energy barrier. Given that ^{200}E was substituted for ^{200}A during the experimental preparation of this crystal structure, any potential energy minima expected prior to the cleavage of the collagen chain may be missing. Furthermore, an increase in ΔG after the loss of favourable energetic contributions from the entwined collagen triple helix to a loosened collagen molecule model (Figure 6) is expected. Such distortion to the collagen triple-helical structure will result in a clash in polarity as hydrophobic side chains associate with water molecules. In the PMF profile, there is a small inflection 0.61–0.71 nm along the reaction coordinate after a $\Delta\Delta\text{G}$ of approximately +21 kJ/mol (+5 kcal/mol) from the lowest minimum. Within this particular umbrella sampling configuration, the collagen peptide is inserted in the active site (Figure 5(A)). Such a small, yet significant, difference in free energy may be indicative of a free collagen molecule undergoing an insertion event due to a fluctuation in the dynamics of the microenvironment, such as temperature (Fields, 1991; Leikina et al., 2002).

A radial distribution function (RDF) was performed over the solvent from the scissile bond (see Figure 7(A)) of the non-inserted trajectory, 1.4 nm along the reaction coordinate. Two pronounced peaks, each indicative of a solvation sphere were observed. The first peak can be seen 0.8 nm from the scissile bond followed by the second 1.8 nm from the bond. Observations clearly show a

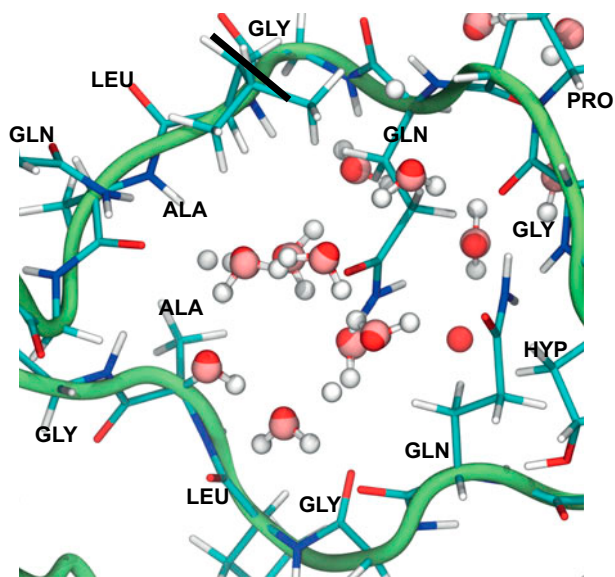


Figure 6. A representative structure extracted from the inflection point, approximately 0.65 nm along the reaction coordinate. The collagen peptide around the scissile bond (as indicated) was found to have disassociated from the neighbouring collagen peptide. Water atoms present between the two chains are clearly marked (within 8 Å of the central glutamine side chain).

separation of the active site from the collagen peptide by water. Given the occupation of the active site by the collagen peptide, it comes as little surprise that the trajectory pertaining to a position 0.65 nm along the reaction coordinate demonstrates little or no order from the scissile bond to bulk water. The RDF calculations of both trajectories yield an expectation of bulk water 3 nm from the bond, with absolute certainty of complete bulk water a further 1.5 nm away.

This study was concluded with a brief investigation into the stability of an unrestrained simulation of the inserted collagen peptide taken from the saddle points observed in the PMF profile. A set of 1000 kJ/mol restraints was applied to the heavy atoms of the complete protein complex and the system progressed for 10 ns. The restraints were then lifted from all but the α -carbon of the zinc-coordinating glycine and the system was extended by a further 10 ns using stored velocities from the previous time step. The two position restraints were reduced to 100 kJ/mol followed by 10 kJ/mol, per additional 10 ns of simulation time. Finally, position restraints were lifted altogether and the system was allowed to progress for 10 ns.

The average distance during the first 8 ns between the carbonyl oxygen of glycine and the zinc ion was approximately 6 Å (Figure 7(B)). The water molecules in the active site cavity were not shown to permeate between the carbonyl group and zinc during this time. During the final 2 ns, the distance was shown to rapidly increase by approximately 2 nm, rendering the interaction ineffective due to the short-range cut-off. Simultaneously, density in water molecules between glycine and the zinc ion also increased. Given the inactive enzymatic substitution of glutamic acid for alanine, there is a loss of the water bridge between the carbonyl groups on the backbone of collagen and the carboxylate functional group on the absent glutamic acid. However, it was still possible to capture a prolonged presence of the polypeptide chain within the active site cavity. There are understood limitations of this particular simulation. Firstly, there would be a discrepancy in partial charges during an attempt to donate the carbonyl group of glycine native to the forcefield to a zinc ion, which had been parameterised without the glycine present. Secondly, it is thought that the presence of a glutamic acid would also have an effect on the polarity of a neighbouring water molecule during the parameterising of the forcefield. We are confident, though, that the short-lived presence of the scissile bond carbonyl within the active site represents the coordination, if only for a brief moment in time, of the collagen backbone during the start of the collagenase reaction mechanism before adjusting to a favourable conformation, as suggested by the ΔG of collagen insertion.

Our results have indicated that there exists only a small positive free energy gradient from a non-inserted

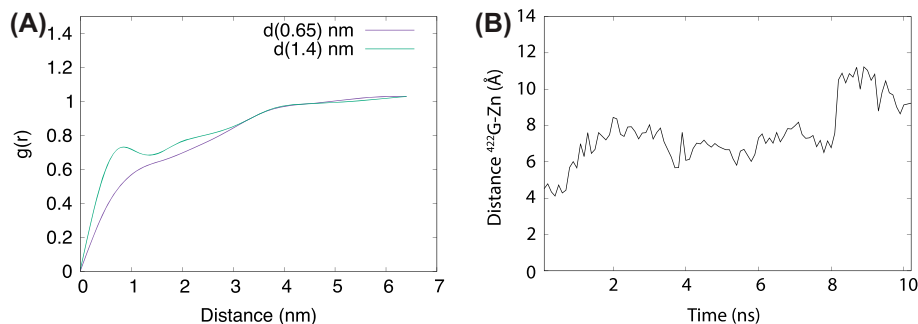


Figure 7. (A) Plot of the radial pair distribution function, $g(r)$, vs. the distance between the carbonyl oxygen of ^{422}G of the collagen peptide of interest, and the catalytic zinc. Distributions were generated using an acspline best fitting curves from umbrella trajectories covering the respective points 0.65 and 1.4 nm along the reaction coordinate of the free energy profile. (B) The distance between the scissile bond of ^{422}G of the inserted collagen peptide and the zinc ion in the catalytic domain of MMP-1 during a 10 ns unrestrained simulation.

collagen to the inserted collagen state. Reflecting on to the three proposed mechanisms of insertion and cleavage, it is entirely possible that the collagen molecule local to the active site may unwind and insert, easily overcoming without assistance the free energy gradient, as proposed by Chung et al. (2004).

Conclusions

The force-field parameterisation and MD simulations of the MMP-1-collagen complex, 4AUO, yielded a stable conformation within the confines of the simulation time. The meta-dynamics calculations suggest that the collagen sequence adjacent to the active site is amenable to unwinding, thus allowing insertion of a single polypeptide chain into the active site. However, free energy characterisation from umbrella sampling of the transition from non-inserted collagen to inserted collagen was energetically unfavourable. The energy barrier may be surmounted via a rare thermal fluctuation to a triple-coiled collagen molecule unwinding and inserting into the active site. Our final simulation of an unrestrained inserted collagen peptide, demonstrated a short-lived interaction between the catalytic active site and the collagen peptide carbonyl group before returning to a favourable conformation.

A quantitative measure of the interaction strength between enzyme and substrate is beyond the scope of this study, although it will be the subject of future work. Our detailed documentation of the intermolecular associations between substrate and enzyme will aid the mapping of human type I collagen onto MMP-1 until a crystal structure has been calculated. Finally, our presented energetic calculations are fundamental in the elucidation of the interaction between MMPs and their substrates and will aid the rational design of *de novo* MMP-1 inhibitors.

Supplementary material

The supplementary material for this paper is available online at <http://dx.doi.org/10.1080/07391102.2016.1153521>.

Acknowledgements

The authors would like to thank the BBSRC (Grant no: BB/K007785) for funding and the Research Computing Platform, UCL, for the use of Legion. Also with thanks for input, Thomas Collier and Marion Cubitt.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Biotechnology and Biological Sciences Research Council [grant number BB/K007785].

ORCID

Helen L. Birch  <http://orcid.org/0000-0002-7966-9967>

Nora H. de Leeuw  <http://orcid.org/0000-0002-8271-0545>

References

- Bernhard, E. J., Gruber, S. B., & Muschel, R. J. (1994). Direct evidence linking expression of matrix metalloproteinase 9 (92-kDa gelatinase/collagenase) to the metastatic phenotype in transformed rat embryo cells. *Proceedings of the National Academy of Sciences USA*, 91, 4293–4297. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8183903>
- Bode, W. (1995). A helping hand for collagenases: The haemopexin-like domain. *Structure*, 3, 527–530. doi:10.1016/S0969-2126(01)00185-X
- Bode, W., Reinemer, P., Huber, R., Kleine, T., Schnierer, S., & Tschesche, H. (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. *The EMBO Journal*, *13*, 1263–1269. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC394940/>
- Borkakoti, N., Winkler, F. K., Williams, D. H., D'Arcy, A., Broadhurst, M. J., Brown, R. A., ... Murray, E. J. (1994). Structure of the catalytic domain of human fibroblast collagenase complexed with an inhibitor. *Nature Structural & Molecular Biology*, *1*, 106–110. doi:10.1038/nsb0294-106
- Brew, K., & Nagase, H. (2010). The tissue inhibitors of metalloproteinases (TIMPs): An ancient family with structural and functional diversity. *Biochimica et Biophysica Acta*, *1803*, 55–71. doi:10.1016/j.bbamcr.2010.01.003
- Chaudhary, A. K., Singh, M., Bharti, A. C., Asotra, K., Sundaram, S., & Mehrotra, R. (2010). Genetic polymorphisms of matrix metalloproteinases and their inhibitors in potentially malignant and malignant lesions of the head and neck. *Journal of Biomedical Science*, *17*, 10–23. doi:10.1186/1423-0127-17-10
- Chung, L., Dinakarandian, D., Yoshida, N., Lauer-Fields, J. L., Fields, G. B., Visse, R., & Nagase, H. (2004). Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. *The EMBO Journal*, *23*, 3020–3030. doi:10.1038/sj.emboj.7600318
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, *98*, 10089–10092. doi:10.1063/1.464397
- Deryugina, E. I., & Quigley, J. P. (2006). Matrix metalloproteinases and tumor metastasis. *Cancer and Metastasis Reviews*, *25*, 9–34. Retrieved from <http://link.springer.com/article/10.1007%2Fs10555-006-7886-9>
- Duffield, J. S., Lupher, M., Thannickal, V. J., & Wynn, T. A. (2013). Host responses in tissue repair and fibrosis. *Annual Review of Pathology: Mechanisms of Disease*, *8*, 241–276. doi:10.1146/annurev-pathol-020712-163930
- Fields, G. B. (1991). A model for interstitial collagen catabolism by mammalian collagenases. *Journal of Theoretical Biology*, *153*, 585–602. doi:10.1016/S0022-5193(05)80157-2
- Gomis-Rüth, F. X., Gohlke, U., Betz, M., Knäuper, V., Murphy, G., López-Otín, C., & Bode, W. (1996). The helping hand of collagenase-3 (MMP-13): 2.7 Å crystal structure of its C-terminal haemopexin-like domain. *Journal of Molecular Biology*, *264*, 556–566. doi:10.1006/jmbi.1996.0661
- Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, *18*, 1463–1472. doi:10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H
- Hoover, W. G. (1985). Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, *31*, 1695–1697. doi:10.1103/PhysRevA.31.1695
- Horn, H. W., Swope, W. C., Pitara, J. W., Madura, J. D., Dick, T. J., Hura, G. L., & Head-Gordon, T. (2004). Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of Chemical Physics*, *120*, 9665–9678. doi:10.1063/1.1683075
- Jemal, A., Tiwari, R. C., Murray, T., Ghafour, A., Samuels, A., Ward, E., ... Thun, M. J. (2004). Cancer statistics, 2004. *CA: A Cancer Journal for Clinicians*, *54*, 8–29. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14974761>
- Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., & Kollman, P. A. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, *13*, 1011–1021. doi:10.1002/jcc.540130812
- Leikina, E., Merts, M. V., Kuznetsova, N., & Leikin, S. (2002). Type I collagen is thermally unstable at body temperature. *Proceedings of the National Academy of Sciences USA*, *99*, 1314–1318. doi:10.1073/pnas.032307099
- Li, J., Brick, P., O'Hare, M. O., Skarzynski, T., Lloyd, L. F., Curry, V. A., ... Blow, D. M. (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. doi:10.1016/S0969-2126(01)00188-5
- Liotta, L. A., Abe, S., Robey, P. G., & Martin, G. R. (1979). Preferential digestion of basement membrane collagen by an enzyme derived from a metastatic murine tumor. *Proceedings of the National Academy of Sciences USA*, *76*, 2268–2272. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC383580/>
- Liotta, L. A., Tryggvason, K., Garbisa, S., Hart, I., Foltz, C. M., & Shafie, S. (1980). Metastatic potential correlates with enzymatic degradation of basement membrane collagen. *Nature*, *284*, 67–68. doi:10.1038/284067a0
- Manka, S. W., Carafoli, F., Visse, R., Bihan, D., Raynal, N., Farndale, R. W., ... Nagase, H. (2012). Structural insights into triple-helical collagen cleavage by matrix metalloproteinase 1. *Proceedings of the National Academy of Sciences USA*, *109*, 12461–12466. doi:10.1016/j.matbio.2015.01.005
- Mitchell, P. G., Magna, H. A., Reeves, L. M., Lopresti-Morrow, L. L., Yocum, S. A., Rosner, P. J., ... Hambor, J. E. (1996). Cloning, expression, and type II collagenolytic activity of matrix metalloproteinase-13 from human osteoarthritic cartilage. *The Journal of Clinical Investigation*, *97*, 761–768. doi:10.1172/JCI118475
- Nerenberg, P. S., & Stultz, C. M. (2008). Differential unfolding of alpha1 and alpha2 chains in type I collagen and collagenolysis. *Journal of Molecular Biology*, *382*, 246–256. doi:10.1016/j.jmb.2008.07.009
- Nosé, S. (1984). A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, *81*, 511–519. doi:10.1063/1.447334
- Ottl, J., Gabriel, D., Murphy, G., Knäuper, V., Tominaga, Y., Nagase, H., ... Moroder, L. (2000). Recognition and catabolism of synthetic heterotrimeric collagen peptides by matrix metalloproteinases. *Chemistry & Biology*, *7*, 119–132.
- Overall, C. M., & López-Otín, C. (2002). Strategies for mmp inhibition in cancer: Innovations for the post-trial era. *Nature Reviews Cancer*, *2*, 657–672. doi:10.1038/nrc884
- Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, *52*, 7182–7190. doi:10.1063/1.328693
- Pelmenschikov, V., & Siegbahn, P. E. M. (2002). Catalytic mechanism of matrix metalloproteinases: Two-layered ONIOM study. *Inorganic Chemistry*, *41*, 5659–5666. doi:10.1021/ic0255656
- Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., ... Lindahl, E. (2013). GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, *29*, 845–854. doi:10.1093/bioinformatics/btt055
- Reboul, P., Pelletier, J. P., Tardif, G., Cloutier, J. M., & Martel-Pelletier, J. (1996). The new collagenase, collagenase-3, is expressed and synthesized by human chondrocytes but not by synoviocytes. A role in osteoarthritis. *The Journal of*

- Clinical Investigation*, 97, 2011–2019. doi:10.1172/JCI118636
- Sekhon, B. S. (2010). Matrix metalloproteinases – an overview. *Research and Reports in Biology*, 1, 1–20. doi:10.2147/RRB.S12043
- de Souza, S. J., Pereira, H. M., Jacchieri, S., & Brentani, R. R. (1996). Collagen/collagenase interaction: Does the enzyme mimic the conformation of its own substrate? *The Journal of the Federation of American Societies for Experimental Biology*, 10, 927–930.
- Stricklin, G. P., Jeffrey, J. J., Roswit, W. T., & Eisen, A. Z. (1983). Human skin fibroblast procollagenase: Mechanisms of activation by organomercurials and trypsin. *Biochemistry*, 22, 61–68. doi:10.1021/bi00270a009
- Thiennu, H. V., & Werb, Z. (2000). Matrix metalloproteinases: Effectors of development and normal physiology. *Genes & Development*, 14, 2123–2133. doi:10.1101/gad.815400
- Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C., & Bussi, G. (2014). PLUMED 2: New feathers for an old bird. *Computer Physics Communications*, 185, 604–613. doi:10.1016/j.cpc.2013.09.018
- Tsukada, H., & Pourmotabbed, T. (2002). Unexpected crucial role of residue 272 in substrate specificity of fibroblast collagenase. *The Journal of Biological Chemistry*, 277, 27378–27384. doi:10.1074/jbc.M201367200
- Tuccinardi, T., Martinelli, A., Nuti, E., Carelli, P., Balzano, F., Uccello-Barretta, G., ... Rossell, A. (2006). Amber force field implementation, molecular modelling study, synthesis and MMP-1/MMP-2 inhibition profile of (R)- and (S)-N-hydroxy-2-(N-isopropoxybiphenyl-4-ylsulfonamido)-3-methylbutanamides. *Bioorganic & Medicinal Chemistry*, 14, 4260–4276. doi:10.1016/j.bmc.2006.01.056
- Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J. C., ... Dupradeau, F.-Y. (2011). R.E.D. Server: A web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Research*, 39, W511–W517. doi:10.1093/nar/gkr288
- Verma, R. P., & Hansch, C. (2007). Matrix metalloproteinases (MMPs): Chemical-biological functions and (Q)SARs. *Bioorganic & Medicinal Chemistry*, 15, 2223–2268. doi:10.1016/j.bmc.2007.01.011
- Vihinen, P., & Kähäri, V. M. (2002). Matrix metalloproteinases in cancer: Prognostic markers and therapeutic targets. *International Journal of Cancer*, 99, 157–166. doi:10.1002/ijc.10329
- Van Wart, H. E., & Birkedal-Hansen, H. (1990). The cysteine switch: A principle of regulation of metalloproteinase activity with potential applicability to the entire matrix metalloproteinase gene family. *Proceedings of the National Academy of Sciences USA*, 87, 5578–5582. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC54368/>
- Wernicke, D., Seyfert, C., Hinzmann, B., & Gromnica-Ihle, E. (1996). Cloning of collagenase 3 from the synovial membrane and its expression in rheumatoid arthritis and osteoarthritis. *The Journal of Rheumatology*, 23, 590–595. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8730110>
- Zhao, Y., Marcink, T. C., Sanganna Gari, R. R., Marsh, B. P., King, G. M., Stawikowska, R., ... Van Doren, S. R. (2015). Transient collagen triple helix binding to a key metalloproteinase in invasion and development. *Structure*, 23, 257–269. doi:10.1016/j.str.2014.11.021