Inorg. Chem. **2002**, *41*, 5659–5666

Catalytic Mechanism of Matrix Metalloproteinases: Two-Layered ONIOM Study

Vladimir Pelmenschikov and Per E. M. Siegbahn*

Department of Physics, Stockholm Center for Physics, Astronomy and Biotechnology (SCFAB), Stockholm University, S-106 91 Stockholm, Sweden

Received February 27, 2002

The two-layered ONIOM(B3LYP:MNDO) method has been used to investigate the hydrolytical mechanism of matrix metalloproteinases (MMPs), a large family of zinc-dependent endopeptidases capable of degrading a wide range of macromolecules of the extracellular matrix. Human stromelysin-1 (MMP-3) was chosen as a physiologically important member of the MMP family. As a structural reference, X-ray data on the stromelysin-1 catalytic domain (SCD) complexed to the transition state analogue diphenyl piperidine sulfonamide inhibitor was used. The backbone spacer of 11 residues (201−211) was included in the final model, spanning the catalytic Glu202 residue and the three structural His201,205,211 zinc ligands. The polypeptide framework incorporated, partly accounting for the protein rigidity, reduces the activation free energy slightly by 1.6 kcal/mol. Essentially a single-step catalytic mechanism was obtained, generally following a classical proposal for MMPs. Glu202 here acts as a base, abstracting a proton from the metal-bound reactant water and delivering this proton to the peptide nitrogen. An auxiliary water molecule is suggested to be of crucial importance acting as an electrophilic agent to the carbonyl oxygen of the substrate. The direct inclusion of the auxiliary water molecule decreases the activation free energy by about 5 kcal/mol via donation of a strong hydrogen bond. The calculated activation barrier of 13.1 kcal/mol agrees well with experimental rates.

I. Introduction

Matrix metalloproteinases (MMPs, or matrixins) along with astacins, adamalysins, and serralysins form a large metzincins[1] superfamily of Zn^{2+}-dependent proteases, sharing a common metal-binding HExxHxxGxxH amino acid consensus motif at the active site. Endopeptidases with this zinc binding signature are involved in a wide range of biologically relevant events of tissue degrading and remodeling. Depending on their substrate specificity, they may participate in growth, embryonic development, and wound healing. Disruption of MMP regulation in cells results in pathological conditions and leads to degenerative diseases, such as arthritis, tumor growth, and metastasis. Modulation of the proteolytic activity of the MMPs is currently an attractive trend in structure-based drug design. Indeed, extensive research in the field of synthetic MMP inhibitors[2,3] during the past decade has led pharmaceutical companies such as Roche, Glycomed, and British Biotech to clinical implementations.

Human stromelysin-1 (MMP-3, classification number EC 3.4.24.17) is one of the most attractive targets in drug discovery today[4] because of its broad physiological specificity. This extracellular endopeptidase of vertebrate tissues degrades various proteoglycan components of the extracellular matrix as well as fibronectin and laminin.[5] Stromelysin also plays a unique role among the MMPs because of its involvement in activation of other MMP proenzymes.[6] The 19 kDa stromelysin catalytic domain (SCD) alone is able to hydrolyze peptides, with kinetic parameters similar to those of the full-length protein.[7,8] The small size, purity, and stability have made SCD highly suitable for biophysical

* To whom correspondence should be addressed. E-mail: ps@physto.se.

(1) Bode, W.; Gomis-Rüth, F.-X.; Stöker, W. *FEBS Lett.* **1993**, *331*, 134.
(2) Borkakoti, N. *Prog. Biophys. Mol. Biol.* **1998**, *70*, 73−94.
(3) Blundell, T. L. *Nat. Struct. Biol.* **1994**, *1*, 73−75.
(4) Browner, M. F. *Perspect. Drug Discovery Des.* **1995**, *2*, 343−351.
(5) Welgus, H. G. Stromelysin: Structure and Function. In *Progress in Inflammation Research and Therapy*; Verlag: Basel, Switzerland, 1991; pp 61−67.
(6) Flannery, C.; Lark, M. V.; Sandt, J. D. *J. Biol. Chem.* **1992**, *267*, 1008−1014.

10.1021/ic0255656 CCC: $22.00
Published on Web 09/05/2002

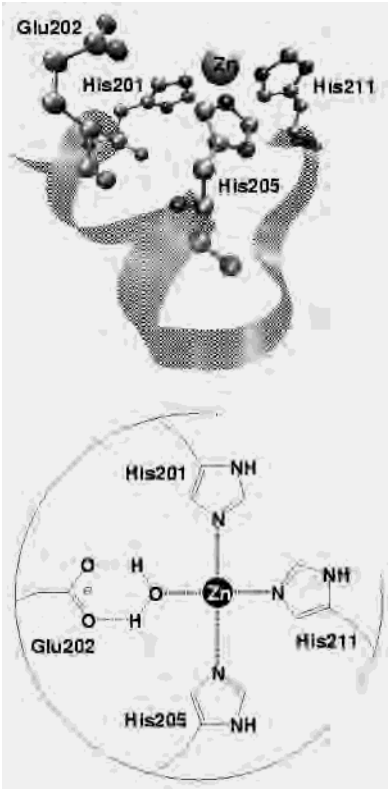

Figure 1. Stromelysin-1 active site structure, generated from X-ray data[12] of the catalytic domain, inhibited with the diphenyl piperidine sulfonamide. The topology of the active site is clarified in the described pattern.

studies. Therefore, analysis of the three-dimensional structure of the enzyme has been devoted mainly to SCD−inhibitor complexes. Both X-ray crystallography and NMR spectroscopy data[9−12] are available.

The X-ray structures reveal the smallest Zn binding site known, with the three histidines of the HExxHxxGxxH "short spacer" cluster chelating the metal cation, Figure 1. Two of them, His201 and His205, project from the catalytic helix, while a sharp turn of the backbone at the conserved glycine, Gly208, allows the third one (His211) to coordinate to the metal. The conserved glutamate, Glu202, is presumed to play an essential role in catalysis, supported by mutagenesis results.[13] Similarly to the classic proposal for the thermolysin-like proteases (TLPs),[14−18] the mechanistic studies assign a general base role to the Glu202 equivalent in MMPs. The water molecule, coordinated to the metal in the reactant complex, is being extremely polarized between the glutamate base and zinc Lewis acid in the Glu−H_2O−Zn bridge, Figure 2. A pK_a value of 5.6 was attributed to this bridging complex.[19] Upon substrate delivery, the water oxygen performs a nucleophilic attack on the peptide carbon. Concertedly, the glutamate abstracts the proton from the water and shuttles it toward the nitrogen of the scissile amide. The tetrahedral *gem*-diolate intermediate is formed, bidentately coordinated to the metal. According to the classic proposals for zinc endopeptidases,[16] the final breakdown of the C−N bond occurs only after the second proton transfer from the water oxygen, mediated again by the conserved glutamate. However, it was recently shown[20] for thermolysin that the immediate amine product can be released already in its neutral form, omitting the second proton transfer by Glu202. The second proton should be accepted by Glu202, however. The protonated form fitting the standard pH near 7.0 can be rapidly formed via contact with solution.

An interesting aspect of the catalysis by MMPs is an apparently simplified arrangement of the active site, as compared to the related endopeptidase families. Unlike the astacins and serralysins, no zinc-coordinated tyrosine is present in the MMPs. On the other hand, the zinc structural ligands are homogeneous (all three are histidines), as compared to TLPs (two histidines and one glutamate), and substituents for His231 or Tyr157 of thermolysin are absent. It has been pointed out[2] that histidine instead of glutamate as a third ligand to the catalytic zinc results in a more positive net charge of the metal center, allowing enough stabilization to the negatively charged *gem*-diolate intermediate even without His231 or Tyr157 analogues. However, a careful examination of the X-ray structures of the inhibited matrixins (as adamalysin II[21,22]) displays a crystal water molecule at the location equivalent to His231. Only a few studies[23,24] stress the importance of this auxiliary water molecule which could act as a potent electrophilic agent in the reaction mechanism of MMPs; see Figure 2.

The most essential part of the mechanism of the MMPs has been investigated in the present study, starting with the formation of the Michaelis complex and ending before the

(7) Marcy, A. I.; Eiberger, L. L.; Harrison, R.; Chan, H. K.; Hutchinson, N. I.; Hagmann, W. K.; Cameron, P. M.; Boulton, D. A.; Hermes, J. D. *Biochemistry* **1991**, *30*, 6476−6483.
(8) Ye, Q.-Z.; Johnson, L. L.; Hupe, D. J.; Baragi, V. *Biochemistry* **1992**, *31*, 11231−11235.
(9) Gooley, P. R.; O'Connell, J. F.; Marcy, A. I.; Cuca, G. C.; Salowe, S. P.; Bush, B. L.; Hermes, J. D.; Esser, C. K.; Hagmann, W. K.; Springer, J. P.; Johnson, B. A. *Nat. Struct. Biol.* **1994**, *1*, 111−118.
(10) Becker, J. W.; Marcy, A. I.; Rokosz, L. L.; Axel, M. G.; Burbaum, J. J.; Fitzgerald, P. M. D.; Cameron, P. M.; Esser, C. K.; Hagmann, W. K.; Hermes, J. D.; Springer, J. P. *Protein Sci.* **1995**, *4*, 1966−1976.
(11) Dhanaraj, V.; Ye, Q.-Z.; Johnson, L. L.; Hupe, D. J.; Ortwine, D. F.; Dunbar, J. B.; Rubin, R. J.; Pavlovsky, A.; Humblet, C.; Blundell, T. L. *Structure* **1996**, *4*, 375−386.
(12) Pavlovsky, A.; Ye, Q.-Z.; Ortwine, D. F.; Humblet, C.; Purchase, C.; White, A.; Roth, B.; Johnson, L.; Hupe, D.; Williams, M.; Dhanaraj, V.; Blundell, T. *Protein Sci.* **1999**, *8*, 1455−1462.
(13) Arza, B.; De Maeyer, M.; Felez, J.; Collen, D.; Lijnen, H. R. *Eur. J. Biochem.* **2001**, *268*, 826−831.
(14) Kester, W. R.; Matthews, B. M. *Biochemistry* **1977**, *16*, 2506−2516.
(15) Holmes, M. A.; Matthews, B. W. *Biochemistry* **1981**, *20*, 6912−6920.
(16) Hangauer, D. G.; Monzingo, A. F.; Matthews, B. W. *Biochemistry* **1984**, *23*, 5730−5741.
(17) Holden, H. M.; Tronrud, D. E.; Monzingo, A. F.; Weaver, L. H.; Matthews, B. W. *Biochemistry* **1987**, *26*, 8542−8553.
(18) Matthews, B. W. *Acc. Chem. Res.* **1988**, *21*, 333−340.
(19) Johnson, L. L.; Pavlovsky, A. G.; Johnson, A. R.; Janowicz, J. A.; Man, C.-F.; Ortwine, D. F.; Purchase, C. F.; White, A. D.; Hupe, D. J. *J. Biol. Chem.* **2000**, *275*, 11026−11033.
(20) Pelmenschikov, V.; Blomberg, M. R. A.; Siegbahn P. E. M. *J. Biol. Inorg. Chem.* **2002**, *7*, 284−298.
(21) Cirilli, M.; Gallina, C.; Gavuzzo, E.; Giordano, C.; Gomis-Rüth, F.-X.; Gorini, B.; Kress, L. F.; Mazza, F.; Paradisi, M. P.; Pochetti, G.; Politi, V. *FEBS Lett.* **1997**, *418*, 319−322.
(22) D'Alessio, S.; Gallina, C.; Gavuzzo, E.; Giordano, C.; Gorini, B.; Mazza, F.; Paradisi, M. P.; Panini, G.; Pochetti, G.; Sella, A. *Bioorg. Med. Chem.* **1999**, *7*, 389−394.
(23) Borkakoti, N.; Winkler, F. W.; Williams, D. H.; D'Arcy, A.; Broadhurst, M. J.; Brown, P. A.; Johnson, W. H.; Murray, E. J. *Nat. Struct. Biol.* **1994**, *1*, 106−110.
(24) Browner, M. F.; Smith, W. W.; Castelhano, A. L. *Biochemistry* **1995**, *34*, 6602−6610.

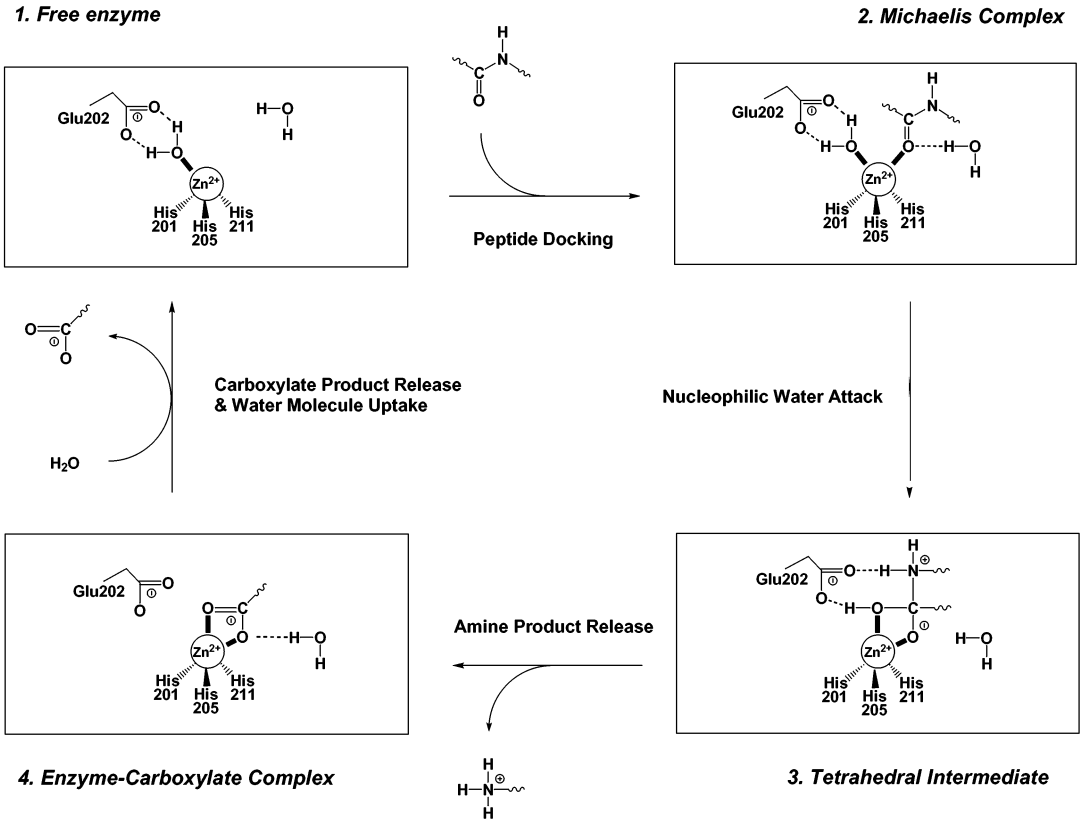

Figure 2. Catalytic cycle for the proteolytical mechanism of stromelysin-1.

actual release of the products from the active site region. Detailed information on the local minima and transition state geometries is given, as well as the corresponding free energy profile. The importance of the auxiliary water molecule for the mechanism is demonstrated. Finally, the influence of including the HExxHxxGxxH backbone motif, coordinating the amino acids of the model, is discussed. The recently developed ONIOM method has been applied at this stage, allowing high accuracy theoretical results to be obtained for this large-scale quantum chemical molecular system.

II. Computational Details

The two-layered hybrid ONIOM[25−27] method as implemented in the GAUSSIAN-98 program package[28] was used in the present study. The inner model and the total real parts for the ONIOM system remain the same throughout the calculations. Only carbon−carbon single covalent bonds were cut (and automatically saturated with hydrogen link atoms) during the generation of the model subsystem to avoid possible chemical artifacts. The model part was selected such that bond formations and cleavages occur relatively far from the boundary region, see description later in the text and figures. Using the terms of the ONIOM concept, the two-layered ONIOM extrapolated energy is defined as[25]

$$E(\text{ONIOM2}) = E(\text{high, model}) + E(\text{low, real}) - E(\text{low, model}) = E(\text{high, model}) + \Delta E(\text{low, real} \leftarrow \text{model})$$

Therefore, E(high, model) should be regarded as an energy contribution from the inner model part treated within a high level method, and ΔE(low, real ← model) should be regarded as a low level energy perturbation from the rest of the real system. Both contributions are discussed further later, showing ΔE(low, real ← model) to be insignificantly small for the activation energy (less than 1 kcal/mol), which is the key parameter of the mechanism. For the overall potential energy surface, the relative ΔE(low, real ← model) contribution has its absolute maximum of 4.7 kcal/mol.

The B3LYP hybrid density functional method[29−32] was used for all the present high level calculations. For the geometry optimization, the LANL2DZ basis set was used. For the zinc atom, this means the Los Alamos nonrelativistic ECP by Hay and Wadt.[33] The valence shell, including the 3d orbitals, is described by a basis set of essentially double-ζ quality including a diffuse 3d function. For the rest of the atoms, LANL2DZ implies the Dunning−

(25) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. J. *J. Phys. Chem.* **1996**, *100*, 19357.
(26) Humbel, S.; Sieber, S.; Morokuma, K. J. *J. Chem. Phys.* **1996**, *105*, 1959.
(27) Froese, R. D. J.; Morokuma, K. Hybrid Method. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; Wiley: New York, 1998.
(28) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.9; Gaussian, Inc.: Pittsburgh, PA, 1998.
(29) Becke, A. D. *Phys. Rev.* **1988**, *A38*, 3098.
(30) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
(31) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(32) Lee, C.; Yang, W.; Paar, R. G. *Phys. Rev.* **1988**, *B37*, 785.
(33) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1988**, *82*, 299.

Huzinaga full double-ζ basis set.[34] Within the low level of theory, the semiempirical MNDO[35] Hamiltonian was used for all the calculations, on the basis of the NDDO[36] approximation. For the optimized stationary points along the reaction pathway, accurate energy evaluations were carried out using an extended 6-311+G-(1d,1p) triple-ζ basis set for the high level, which includes a set of polarization functions on each atom.

The geometry optimization of the ONIOM model was followed by a Hessian calculation using the same assignment of the high (H) and the low (L) levels of the theory, verifying that it contains only one negative eigenvalue, with vibrational vectors leading to the desired neighboring minima. Zero-point thermal-corrected vibrational effects and entropy contributions[37] to the energy were obtained at 298.15 K, on the basis of the computed Hessians. The zero-point corrections are within 2.7 kcal/mol for the different stationary points. The entropy effects gave a substantial relative contribution of −5.9 kcal/mol for the very last minimum, which is discussed in the following section. The self-consistent isodensity polarized continuum model (SCI-PCM)[38,39] was used to estimate the effects of the protein environment, when mentioned. The cavity containing the molecular system of interest is defined self-consistently by the surface with the default isodensity value of $0.0004e/B^3$, which has been shown to yield volumes close to the measured molar ones. The dielectric constant was chosen equal to 4, as commonly used for proteins.[40] The effects of the dielectric environment were found to be within 2 kcal/mol.

For the purely B3LYP hybrid density functional calculations performed in this study, the level of theory was the same as for the high level of the ONIOM model. The results of the different models should thus be comparable.

III. Results and Discussion

The experimentally suggested mechanism, as shown in Figure 2, has been followed in general in the present study. Originating from the crystallographic studies by Matthews et al.[14−18] for the hydrolysis by thermolysin, it has later been generalized[2,24] to the entire MMP class. Recently, a detailed reaction pathway for the catalysis by thermolysin[20] was obtained at the B3LYP level of theory. Because it agreed well with experimental data, it has been used here as a starting point; see the scheme in Figure 3.

Three models of the active site are discussed in this section in sequence of increased complexity, as briefly described in Table 1. At the B3LYP level, the zinc-coordinated structural histidines are represented by imidazoles, while the conserved catalytic glutamate Glu202 is modeled by a carboxylate, retaining a varying number of carbon atoms of the side chain. The extended butyrate model for the glutamate in the A and B cases incorporates the entire side chain of the residue plus the α-carbon of the backbone. This was done to constrain the possible extra mobility of the second shell Glu202 residue, which is artificially loose in the A and B models; see Figure 4. An alternative solution would be to lock the heavy edge atoms of the side chains in their original positions in the X-ray structure to approximately model the rigid protein. However, in contrast to our recent work on thermolysin, no artificial edge fixation has been applied here, because only small differences between the released and fixed edge approximations were obtained in the previous study. Instead, the backbone of the consensus HExx-HxxGxxH sequence has been included explicitly in the B-ONIOM system, covalently holding one catalytic and three structural residues. This allows a contraction of the chemical model for Glu202, with formate as the simplest one which should still be accurate enough.[41] This considerably decreases

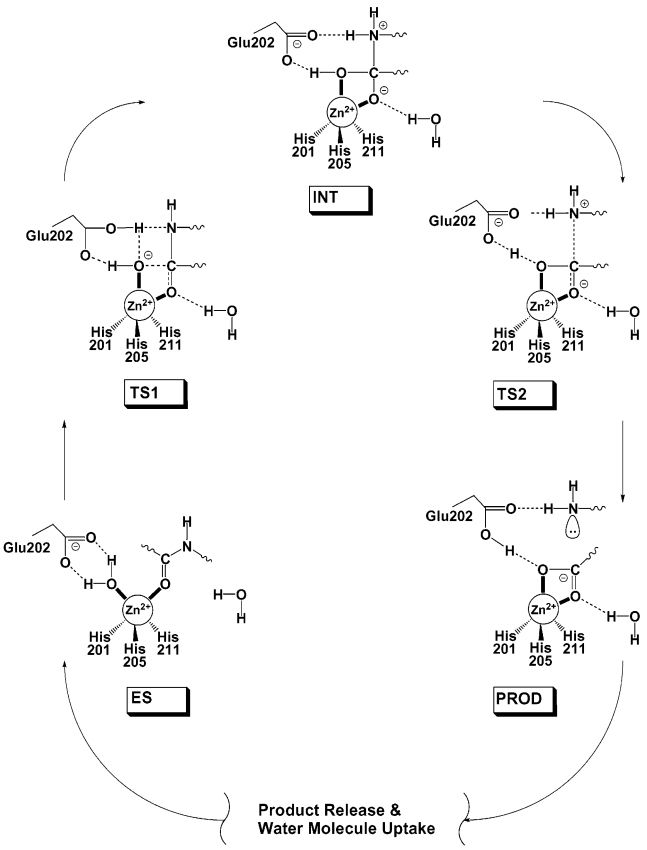

Figure 3. Reaction pathway investigated for the peptide bond cleavage by MMPs. Strong hydrogen bonds and bond rearrangements in transition states are given by dashed lines. Charges given are the formal ones.

Table 1. Model Systems Used for the Substrate and Active Site Amino Acid Residues

model system	size, atoms	auxiliary H_2O	substrate	Glu202	His201,205,211
A	56	no	*N*-methyl-acetamide	butyrate	imidazole
B	59	yes	*N*-methyl-acetamide	butyrate	imidazole
B-ONIOM	50[a]	yes	*N*-methyl-acetamide	formate	imidazole

[a] The model part of the ONIOM system is implied here; see the Computational Details section. For the real ONIOM system, the total number of atoms is 132.

(34) Dunning, J.; Hay, P. J. In *Modern Theoretical Chemistry*; Schaefer, H. F., III, Ed.; Plenum: New York, 1976; Vol. 3, p 1.
(35) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.
(36) Pople, J. A.; Santry, D. P. *J. Chem. Phys.* **1965**, *43*, S129.
(37) McQuarrie, D. A. *Statistical Thermodynamics*; Harper and Row: New York, 1973.
(38) Wiberg, K. B.; Rablen, P. R.; Rush, D. J.; Keith, T. A. *J. Am. Chem. Soc.* **1995**, *117*, 4261.
(39) Wiberg, K. B.; Keith, T. A.; Frisch, M. J.; Murcko, M. *J. Phys. Chem.* **1995**, *99*, 9072.
(40) Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421−437.
(41) Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Phys. Chem.* **2001**, *105*, 9375−9386.

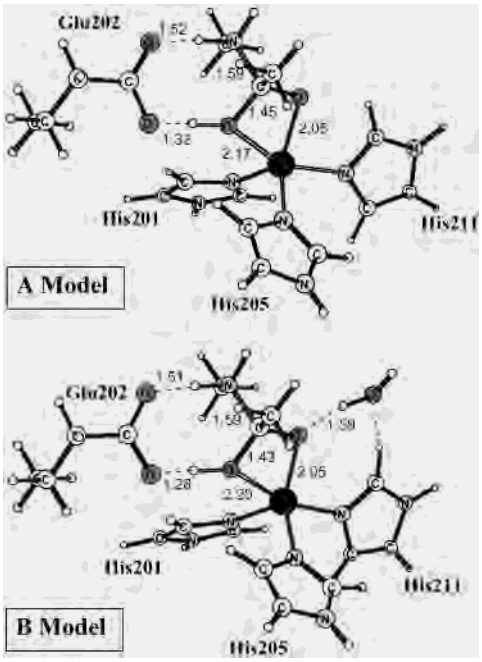

Figure 4. Optimized *gem*-diolate intermediate INT structures for the A and B models.

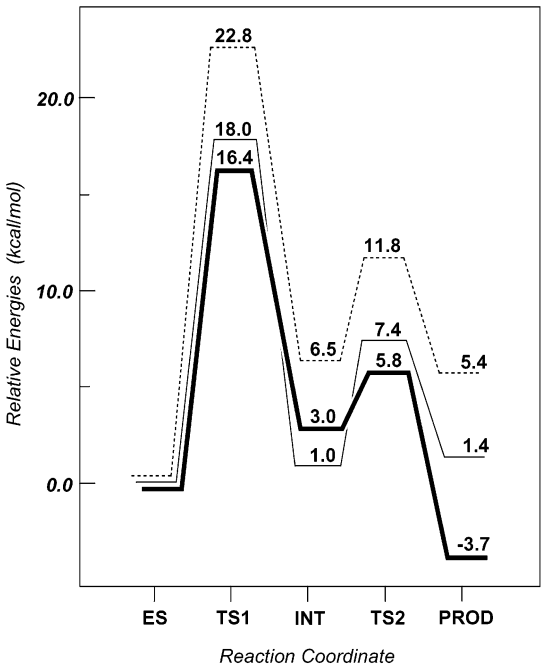

Figure 5. Potential energy surfaces obtained for the A, B, and B-ONIOM model systems at the Lanl2DZ level (dashed, plain, and bold lines, respectively).

the computational effort, because the rest of the side chains can be left out from the high level model and treated by MNDO, as well as the backbone fragment. The "x" residues were replaced by hydrogen atoms (glycines) in the B-ONIOM model. Finally, the peptide substrate was modeled by *N*-methylacetamide in all the A, B, and B-ONIOM models; see figures for detailed structures.

The starting coordinates for the optimization procedure were extracted from the 1CAQ[12] PDB file containing the SCD complexed to the diphenyl piperidine sulfonamide transition state analogue inhibitor. The active site inhibited with a transition state analogue should be the best structural match to the midpoint of the reaction path, which is the INT structure in Figure 3. The metal-bound inhibitor part was therefore eliminated and substituted with the *gem*-diolate tetrahedral intermediate, formed from the reactants. The initial conformation of the tetrahedral intermediate was taken from the INT-like structure obtained in the recent study for thermolysin. Geometry optimization of the resulting system led directly to the INT minimum. A continuous scan forward and backward along the reaction coordinate led to the rest of the extrema on the potential energy surface.

a. Preliminary Reaction Model. At the first stage of our study, the simplest A model was investigated at the B3LYP/Lanl2DZ level. The *gem*-diolate intermediate INT structure was obtained as described previously; see Figure 4. The shape of the potential energy profile for the MMP hydrolysis was found to be qualitatively very similar to the catalytic pathway for thermolysin, with the three equilibrium structures and the two connecting transition states as shown in Figure 3. No solvent water was included in the A model, to estimate the strength of the Zn cation alone as an electrophile in the ligand field of the three histidines. Having two histidines and one glutamate as zinc ligands was shown to be insufficient for the metal center to stabilize the negatively charged intermediate in thermolysin. In that case the second ligand sphere histidine His231 effectively stabilizes the "oxyanion hole",[16] reducing the activation barrier by about 10 kcal/mol.[20] Also for MMPs, the energetics given in Figure 5 suggest that some electrophilic agent is still missing at the level of the A model, since the TS1 rate-limiting step is as high as 22.8 kcal/mol, which means a kinetically too slow mechanism. The large basis correction reduces this value by only 1.1 kcal/mol. The entropy contribution and zero-point vibrational effects are of the same scale, increasing the barrier by 1.2 kcal/mol. Surprisingly, the dielectric effects (obtained for the A model) actually increase the barrier height by 1.7 kcal/mol.

b. Direct Inclusion of an Assisting Water Molecule. Only a few studies[23,24] stress the importance of the additional water molecule for the hydrolytic mechanism of MMPs; see Figure 2 or 3. However, a water molecule, or even two of them, are the only available hydrogen bond donors in the immediate vicinity to the carbonyl oxygen, which has a quite negative charge at the transition state. The Mulliken population analysis for the A model predicts a charge of $-0.65e$ on the carbonyl oxygen in the TS1 transition state. The auxiliary water molecule has been directly included in the B model. An internuclear separation of 2.57 Å between the oxygens of the substrate and the water molecule in the optimized *gem*-diolate INT intermediate (Figure 4) indicates a quite strong interaction in this hydrogen bond. The charge of the substrate oxygen decreases to $-0.85e$, which is the maximum negative atomic charge obtained along the reaction pathway. A resulting stabilization of about 5 kcal/mol by adding the water molecule is obtained at the transition state (4.8 kcal/mol for TS1 and 5.5 kcal/mol for INT), as can be seen from the profiles for the A and B systems in Figure 5. Transformed to rate constants according to transition state

theory, this value gives a factor of 3000 in substrate conversion efficiency (number of peptides degraded per unit of time) between the A and B models. An effect of adding an organic cosolvent on the kinetic parameters of two MMPs, namely gelatinase A and stromelysin-1, confirmed[42] the importance of the *weakly bound* water for the MMP activity. The interpretation of the experiments is that the weakly bound water molecules are essential for mediating the interaction between the substrate and the enzyme. The present study provides additional support for this effect.

c. Large-Scale ONIOM(B3LYP:MNDO) Model. Accurate Results and Role of the Backbone. The Zn^{2+} ion is the second most abundant in nature following iron. With its fully occupied 3d shell, it is one of the easiest metals to handle by quantum chemical methods because there are no ligand field effects, and the binding conformation is therefore dictated solely by ligand features. This explains the availability of relatively reliable semiempirical parameters for Zn. The MNDO method augmented with the Zn parameters[43] makes it possible to apply the two-layered ONIOM(B3LYP:MNDO) method for the extended B-ONIOM model of the stromelysin catalytic domain. It can be added that because the zinc complex, with all its ligands, is included in the high accuracy part of the model treated by the B3LYP method, the results should furthermore be very insensitive to the zinc parameters. The backbone connecting the three structural histidines His201,205,211 and the catalytic glutamate Glu202 was included in the ONIOM model as described in the first paragraph of this section. It is worth noting that stromelysin is an unusually simple example for backbone incorporation, because only 11 links of the polypeptide chain need to be included. For thermolysin, which belongs to the HExxH-(>20)H "long spacer" zinc proteases, it would be necessary to include 90 links. That would considerably slow the computations, particularly for the Hessian evaluation.

Optimized geometries for the B-ONIOM model are given in Figures 6−10, corresponding to the reaction sequence in Figure 3. The energy profile is similar to the one obtained for the B model; see Figure 5. The rate-limiting TS1 barrier becomes only 1.6 kcal/mol lower, with a value of 16.4 kcal/mol at the B3LYP/Lanl2DZ level of theory. The maximum deviation of 2.3 kcal/mol between the B and B-ONIOM profiles is reached for the final PROD products. The interpretation of this result is that the MMP catalysis is not dependent on the specific strained *entatic*[44,45] state of the enzyme in the active site region. This conclusion was drawn also for thermolysin, where the approximation with fixed edge atoms was made; see comments at the beginning of this section. The structures in Figures 6−10 suggest that the conformation of the backbone is quite stable during catalysis, because the polypeptide bends insignificantly during the substrate cleavage. Inspecting the position of the solvent water molecule, which is treated inside the high level model part, again confirms its stabilizing function. The closest approach of the water molecule to the substrate was obtained exactly at the rate-limiting TS1 step.

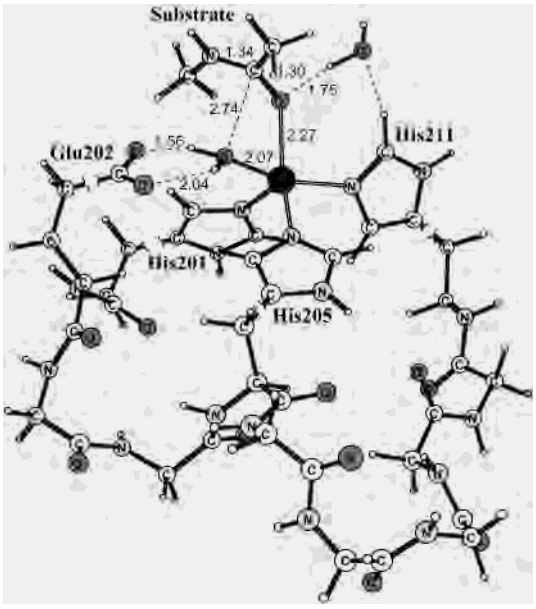

Figure 6. Optimized enzyme−substrate ES structure. Boundaries of the model part are shown as broken bonds.

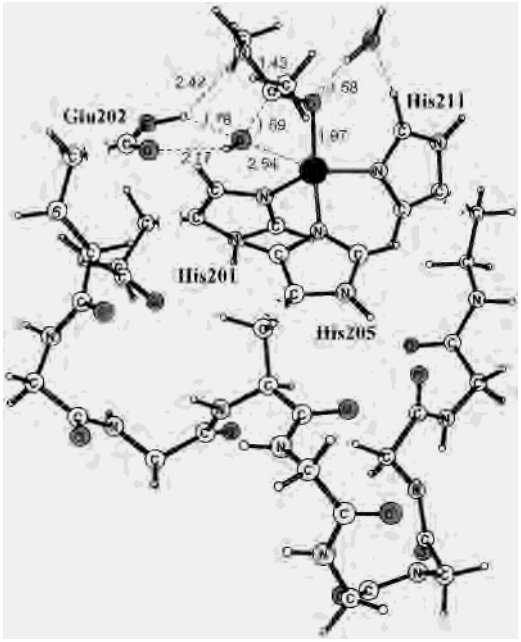

Figure 7. Optimized transition state TS1 structure for the nucleophilic attack by water. Boundaries of the model part are shown as broken bonds.

Single point calculations were performed for the optimized geometries using the large 6-311G+(1d,1p) basis set for the high level, as described in section II. For the TS1 transition state, only a small correction of −0.5 kcal/mol was obtained. The largest effect was found for the INT intermediate, which becomes considerably higher in energy by 6.3 kcal/mol. Similar to what was found for thermolysin, the most accurate values place the energy of the TS2 transition state for the direct C−N bond cleavage below the energy of the INT structure. This means that there is essentially a *single step*

(42) Willenbrock, F.; Knight, C. G.; Murphy, G.; Phillips, I. R.; Brocklehurst, K. *Biochemistry* **1995**, *34*, 12012−12018.
(43) Dewar, M. J. S.; Merz, K. M. *Organometallics* **1986**, *5*, 1494.
(44) Galdes, A.; Vallee, B. L. *Metal Ions in Biological Systems*; Siegel, H., Ed.; Marcel Dekker, Inc.: New York, 1983; Vol. 15, pp 6−7.
(45) Williams, R. J. P. *Eur. J. Biochem.* **1995**, *234*, 363−381.

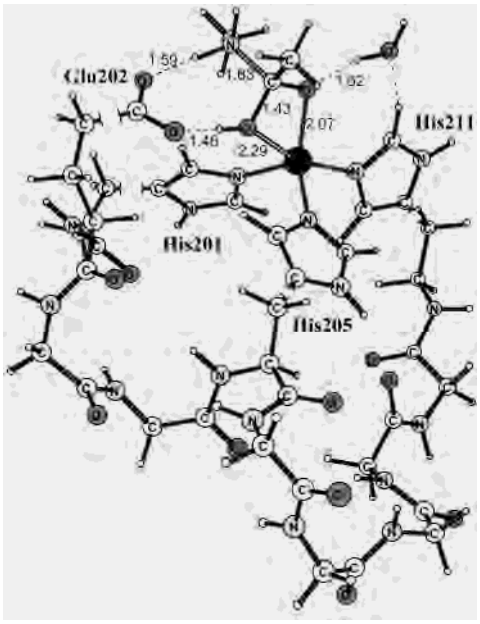

Figure 8. Optimized *gem*-diolate intermediate INT structure. Boundaries of the model part are shown as broken bonds.

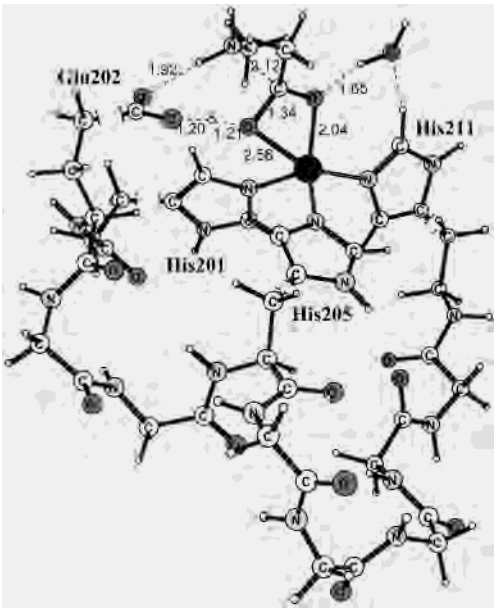

Figure 9. Optimized transition state TS2 structure for the direct C−N bond cleavage. Boundaries of the model part are shown as broken bonds.

mechanism for peptide hydrolysis by MMPs. The final free energy profile in Figure 11 includes also zero-point vibrational effects and entropy contributions. The activation barrier of 13.1 kcal/mol is kinetically feasible and agrees well with standard-state Gibbs activation energies of 15−16 kcal/mol, derived from k_{cat} data[46] for thioester substrates. The high exothermicity of 15.2 kcal/mol obtained for the overall reaction is due to the large basis correction (−5.4 kcal/mol) and entropy contribution (−5.9 kcal/mol). This is a problematic point in the present model (not the mechanism), because the reverse reaction of peptide synthesis becomes too slow. In the terminal PROD complex of the reaction,

(46) Stein, R. L.; Izquierdo-Martin, M. *Arch. Biochem. Biophys.* **1994**, *308*, 274−277.

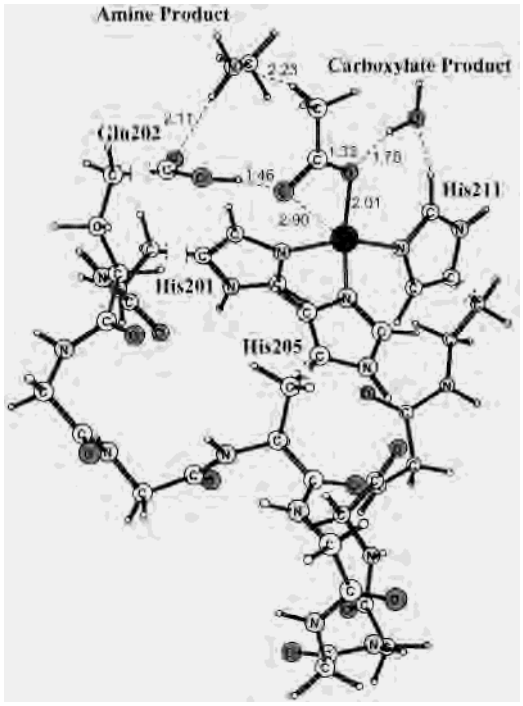

Figure 10. Optimized PROD structure for the hydrolysis products. Boundaries of the model part are shown as broken bonds.

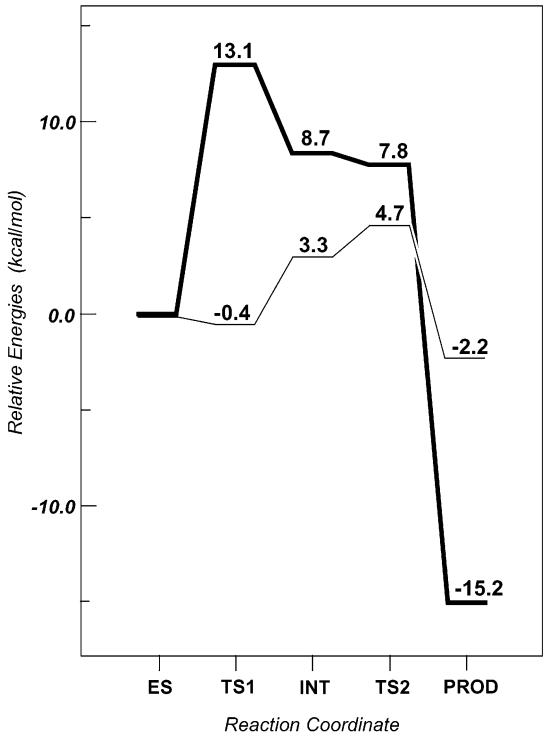

Figure 11. Free energy surface obtained for the advanced B-ONIOM model system. The low level MNDO ΔE(low, real ← model) contributions to the energetics are given by the thin line.

the methylamine part of the products is only weakly bound to the active site via hydrogen bonding, giving rise to increased entropy. Clearly, after or even during the second transition state TS2, the model is probably less valid, because motion of the amine part of the cleaved peptide would be restricted in the actual enzyme and should cost at least a few kilocalories.

The low level MNDO ΔE(low, real ← model) contribution to the energetics (Figure 11) is −0.4 kcal/mol for the TS1 transition state. This can be regarded as another confirmation of the small role of the backbone during the catalytic step. The largest value over the profile is +4.7 kcal/mol, affecting the second transition state, TS2. However, it does not modify any of the main conclusions here. An indication of a very small protein environment effect of about 1 kcal/mol was suggested from the recent study on thermolysin,[20] which has a high degree of similarity with MMPs. As mentioned at the end of section IIIa, the dielectric effects are rather small for the A model, where they might have been expected to contribute significantly because of the extra negative charge at the carbonyl oxygen. Even smaller effects are obtained for the B model system, where this oxygen is screened by the solvent water included in the model. On the basis of these results, the calculation of the dielectric effects was omitted for the B-ONIOM model.

IV. Conclusions

In the present study, the mechanism for peptide hydrolysis by human stromelysin-1 has been investigated using quantum chemical methods. The structural reference was taken from crystallographic X-ray data[12] for the enzyme, inhibited with the transition state analogue piperidine sulfonamide inhibitor. The model was set up to investigate the catalytic action. Only the characteristic HExxHxxGxxH zinc-binding motif was included, with the structural zinc-coordinating histidines and the catalytic glutamate considered explicitly and the rest of the motif as a backbone. Because the HExxHxxGxxH spacer is a hallmark of metzincins proteases, our results are generally applicable to this large family of zinc enzymes, which includes also matrix metalloproteases.

Three models (Table 1) were investigated, basically following the single mechanism described in Figure 3. The difference between models A and B (Figure 4) reveals the effect of including the solvent water, which acts as a potent electrophile to the zinc-coordinated substrate oxygen and reduces the activation barrier by about 5 kcal/mol; see Figure 5. Therefore, a support is provided for the experimentally suggested[42] importance of weakly bound water. The recently implemented two-layered ONIOM method has been used for the third B-ONIOM model, allowing the consideration of a more realistic approximation for the active site, containing 132 atoms and including the backbone between the residues. However, no significant effect of protein strain was observed for the catalysis, and the barrier for the rate-limiting step TS1 (Figure 7) was decreased by only 1.6 kcal/mol. Essentially a single-step reaction mechanism has been obtained; see Figure 11. The intermediate INT (Figure 8) at the midpoint of the profile thus becomes metastable. The Glu202 residue is confirmed to play a key role, acting as a base during the reactant water deprotonation. The Zn^{2+} center remains pentacoordinated during the reaction, with a distorted trigonal bipyramidal coordination. However, for the PROD (Figure 10) final structure, a tetrahedral coordination sphere can be assigned, because one of the carboxylate oxygens is quite far away (2.9 Å) from the metal. The key parameter obtained is the activation energy, which is 13.1 kcal/mol and agrees well with the experimental results of 15−16 kcal/mol for thioester substrates.

The final comment made here is that the outcome of the current study for the catalysis by MMPs closely matches the previous results for the catalysis by thermolysin-like proteases (TLPs). This relates to the similarity of the energy profile and structure obtained, and to the insignificance of protein strain in both types of zinc proteases.

IC0255656