Substrate Hydrolysis by Matrix Metalloproteinase-9*

Received for publication, January 30, 2001, and in revised form, March 12, 2001
Published, JBC Papers in Press, March 14, 2001, DOI 10.1074/jbc.M100900200

Steven J. Kridel‡, Emily Chen‡, Lakshmi P. Kotra¶, Eric W. Howard‡, Shahriar Mobashery§, and Jeffrey W. Smith**

From the ‡Program On Cell Adhesion and the Cancer Research Center, Burnham Institute, La Jolla, California 92037, the ¶Institute for Drug Design and the Department of Chemistry, Wayne State University, Detroit, Michigan 48202, and the §Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104

The catalytic clefts of all matrix metalloproteinases (MMPs) have a similar architecture, raising questions about the redundancy in substrate recognition across the protein family. In the present study, an unbiased phage display strategy was applied to define the substrate recognition profile of MMP-9. Three groups of substrates were identified, each occupying a distinct set of subsites within the catalytic pocket. The most prevalent motif contains the sequence Pro-X-X-Hy-(Ser/Thr) at P_2 through P_2. This sequence is similar to the MMP cleavage sites within the collagens and is homologous to substrates that have been selected for other MMPs. Despite this similarity, most of the substrates identified here are selective for MMP-9 over MMP-7 and MMP-13. This observation indicates that substrate selectivity is conferred by key subsite interactions at positions other than P_3 and P_1. This study shows that MMP-9 has a unique preference for Arg at both P_2 and P_1, and a preference for Ser/Thr at P_2. Substrates containing the consensus MMP-9 recognition motif were used to query the protein data bases. A surprisingly limited list of putative physiologic substrates was identified. The functional implications of these proteins lead to testable hypotheses regarding physiologic substrates for MMP-9.

Matrix metalloproteinase-9 (MMP-9) is a member of the matrixin family of metallo-endopeptidases (1–3). MMP-9 is historically referred to as gelatinase B because of its ability to cleave gelatin, a denatured form of collagen, in vitro. Along with MMP-2, MMP-9 differs from other MMPs because it contains three fibronectin type II repeats that have high binding affinity for collagen. These repeats are thought to mediate the action brings the catalytic pocket of the MMP in proximity to collagen, thereby enhancing its rate of hydrolysis. Despite these well characterized biochemical interactions, it is now clear that MMP-9 is also able to cleave a number of other proteins and may have a rather wide range of physiologic substrates (4–8).

Much of our understanding of the biological function of MMP-9 comes from the study of mice lacking this gene. For example, MMP-9-deficient mice have impaired ossification of the skeletal growth plate, a defect that has been partially attributed to poor vascularization of developing bone (9). Studies on these mice also show that MMP-9 is essential for the recruitment of osteoclasts into developing bones (10). Other work indicates that MMP-9-deficient mice are resistant to dermal blistering in a bullous pemphigoid model, an effect that has been attributed to the inability of these mice to cleave the SERPIN α-proteinase inhibitor (5). Finally, recent studies in the RIP1-Tag2 transgenic mouse model of multistage carcinogenesis indicate that MMP-9 is part of the angiogenic “switch” that is essential for tumor growth (11, 12). Other reports suggest that MMP-9 may play a role in inflammation in the nervous system. MMP-9 is elevated in encephalomyelitis (7, 8), in the cerebrospinal fluid of patients with multiple sclerosis (13), and in patients with AIDS-related dementia (14).

A first approximation of the substrate recognition specificity of MMP-9 has been gleaned from alignments of its cleavage site within a number of different proteins (4, 8, 9, 15–17). Nevertheless, a detailed understanding of subsite preferences for MMP-9 is lacking. Such an analysis is particularly important because the catalytic cleft of MMP-9 is closely related to other MMPs, raising questions about the distinction among substrates for these proteases. These issues are particularly important because many of the current pharmacologic antagonists of the MMPs have overlapping inhibition profiles. Furthermore, in the post-genomic era, where the sequences of all proteins will soon be available, information on substrate recognition could help identify important physiologic substrates.

With these ideas in mind, we applied an unbiased approach to define the substrate recognition preference of MMP-9. MMP-9 was used to cleave substrates within a vastly complex phage display library of random hexamers. Substrates within this library can be cleaved at any position within the hexamer, allowing information on the substrate specificity on both sides of the scissile bond to be obtained. Three families of substrates were identified. The largest group contains a Pro-X-X-Hy-(Ser/Thr) motif (X is any residue, and Hy is a hydrophobic residue) that occupies positions P_3 through P_2. This general motif is cleaved by a number of MMPs and is presumed to represent a collagen-like substrate (18–23). Nevertheless, substrates within this family that were selected show considerable selec-
Substrate Recognition by Matrix Metalloproteinase-9

20573

activity for MMP-9. The second group of substrates are defined by a Gly-Leu-(Lys/Arg) motif at positions P1 through P2. Members of the third family of substrates are unique in that they contain Arg residues at both P1 and P2. Altogether, these findings reveal multiple modes of substrate recognition by MMP-9 and provide important insights into the hydrolysis of physiologic substrates that may be important in biology and pathology.

EXPERIMENTAL PROCEDURES

Purified forms of full-length MMP-7, MMP-13, and TIMP-2 were purchased from Chemicon (Temecula, CA). MMP-9 was expressed and purified using the AMS Scientific (Concord, CA). Restriction enzymes were from Roche Biosciences or New England Biolabs. Oligonucleotides were synthesized by Integrated DNA Technologies, Inc. Tissue culture media and reagents were from Irvine Scientific (Irvine, CA). All other reagents, chemicals, and plasticware were from Sigma or Fisher.

Construction of Substrate Phage Display Library—Substrate phage libraries were generated using a modified version of the IUSE5 phage-emed (24-26). A FLAG epitope was engineered at the NH2 terminus of the geneIII protein by annealing oligonucleotides 5'-CGCGTTTGGTCGTGCGTCTTTTGACTGTCATG-3' and 5'-CGACTCAGAAAAACGAGCACGACAAAAC-3' and ligation them into IUSE5 at the KpnI and XbaI restriction sites. The random hexamers were generated by PCR extension of oligonucleotide in Incubation Buffer (50 mM Tris, pH 9.0, 100 mM NaCl, 5 mM CaCl2, 100 mM KCl). Elongation of the template oligonucleotide was performed using Sequenase (United States Biochemical Corp.) (27). The final cDNA product was precipitated with ethanol, re-suspended in water, and digested with SfiI. The DNA insert and IUSE5 were mixed and ligated at a 5:1 molar ratio and electroporated into Escherichia coli MC1061(F-). Several phage were selected for sequencing to confirm randomness in the insert sequences and the correct reading frame.

Expression and Purification of MMP-9—The cDNA encoding the catalytic domain of MMP-9 (28, 29) was generated by PCR and cloned into pCDNA3 (Invitrogen). The expression vector was used to transfect HEK 293 cells by electroporation, and individual antibiotic-resistant colonies were isolated and cultured. Colonies were screened by reverse transcription-PCR for MMP-9 mRNA. Expression of the proteinaceous catalytic domain was determined by zymography using conditioned medium from the transfected cells. Cells expressing the MMP-9 catalytic domain were grown in 150-mm dishes and conditioned in Dulbecco's modified Eagle's medium with 10% fetal bovine serum supplemented with G418 (200 μg/ml) for 5 days. The catalytic domain of MMP-9 from the conditioned medium was isolated from gelatin-Sepha-

zymography. Pro-MMP-13 was activated with APMA for 3 hr at room temperature (29). Activation was monitored by altered migration of the protein on SDS-polyacrylamide gel electrophoresis and zymography. Pro-MMP-13 was activated with APMA for 3 hr at room temperature, and MMP-7 required no activation.

The level of active protease was always quantified by active site molar ratio and electroporated into E. coli (28, 29) was generated by PCR and cloned into pCDNA3 (Invitrogen). The expression vector was used to transfect HEK 293 cells by electroporation, and individual antibiotic-resistant colonies were isolated and cultured. Colonies were screened by reverse transcription-PCR for MMP-9 mRNA. Expression of the proteinaceous catalytic domain was determined by zymography using conditioned medium from the transfected cells. Cells expressing the MMP-9 catalytic domain were grown in 150-mm dishes and conditioned in Dulbecco's modified Eagle's medium with 10% fetal bovine serum supplemented with G418 (200 μg/ml) for 5 days. The catalytic domain of MMP-9 from the conditioned medium was isolated from gelatin-Sepha-

zymography. Pro-MMP-13 was activated with APMA for 3 hr at room temperature, and MMP-7 required no activation.

The level of active protease was always quantified by active site molar ratio and electroporated into E. coli (28, 29) was generated by PCR and cloned into pCDNA3 (Invitrogen). The expression vector was used to transfect HEK 293 cells by electroporation, and individual antibiotic-resistant colonies were isolated and cultured. Colonies were screened by reverse transcription-PCR for MMP-9 mRNA. Expression of the proteinaceous catalytic domain was determined by zymography using conditioned medium from the transfected cells. Cells expressing the MMP-9 catalytic domain were grown in 150-mm dishes and conditioned in Dulbecco's modified Eagle's medium with 10% fetal bovine serum supplemented with G418 (200 μg/ml) for 5 days. The catalytic domain of MMP-9 from the conditioned medium was isolated from gelatin-Sepha-

zymography. Pro-MMP-13 was activated with APMA for 3 hr at room temperature, and MMP-7 required no activation.

The level of active protease was always quantified by active site molar ratio and electroporated into E. coli (28, 29) was generated by PCR and cloned into pCDNA3 (Invitrogen). The expression vector was used to transfect HEK 293 cells by electroporation, and individual antibiotic-resistant colonies were isolated and cultured. Colonies were screened by reverse transcription-PCR for MMP-9 mRNA. Expression of the proteinaceous catalytic domain was determined by zymography using conditioned medium from the transfected cells. Cells expressing the MMP-9 catalytic domain were grown in 150-mm dishes and conditioned in Dulbecco's modified Eagle's medium with 10% fetal bovine serum supplemented with G418 (200 μg/ml) for 5 days. The catalytic domain of MMP-9 from the conditioned medium was isolated from gelatin-Sepha-

zymography. Pro-MMP-13 was activated with APMA for 3 hr at room temperature, and MMP-7 required no activation.

The level of active protease was always quantified by active site molar ratio and electroporated into E. coli (28, 29) was generated by PCR and cloned into pCDNA3 (Invitrogen). The expression vector was used to transfect HEK 293 cells by electroporation, and individual antibiotic-resistant colonies were isolated and cultured. Colonies were screened by reverse transcription-PCR for MMP-9 mRNA. Expression of the proteinaceous catalytic domain was determined by zymography using conditioned medium from the transfected cells. Cells expressing the MMP-9 catalytic domain were grown in 150-mm dishes and conditioned in Dulbecco's modified Eagle's medium with 10% fetal bovine serum supplemented with G418 (200 μg/ml) for 5 days. The catalytic domain of MMP-9 from the conditioned medium was isolated from gelatin-Sepha-

zymography. Pro-MMP-13 was activated with APMA for 3 hr at room temperature, and MMP-7 required no activation.

The level of active protease was always quantified by active site molar ratio and electroporated into E. coli (28, 29) was generated by PCR and cloned into pCDNA3 (Invitrogen). The expression vector was used to transfect HEK 293 cells by electroporation, and individual antibiotic-resistant colonies were isolated and cultured. Colonies were screened by reverse transcription-PCR for MMP-9 mRNA. Expression of the proteinaceous catalytic domain was determined by zymography using conditioned medium from the transfected cells. Cells expressing the MMP-9 catalytic domain were grown in 150-mm dishes and conditioned in Dulbecco's modified Eagle's medium with 10% fetal bovine serum supplemented with G418 (200 μg/ml) for 5 days. The catalytic domain of MMP-9 from the conditioned medium was isolated from gelatin-Sepha-

zymography. Pro-MMP-13 was activated with APMA for 3 hr at room temperature, and MMP-7 required no activation.

The level of active protease was always quantified by active site molar ratio and electroporated into E. coli (28, 29) was generated by PCR and cloned into pCDNA3 (Invitrogen). The expression vector was used to transfect HEK 293 cells by electroporation, and individual antibiotic-resistant colonies were isolated and cultured. Colonies were screened by reverse transcription-PCR for MMP-9 mRNA. Expression of the proteinaceous catalytic domain was determined by zymography using conditioned medium from the transfected cells. Cells expressing the MMP-9 catalytic domain were grown in 150-mm dishes and conditioned in Dulbecco's modified Eagle's medium with 10% fetal bovine serum supplemented with G418 (200 μg/ml) for 5 days. The catalytic domain of MMP-9 from the conditioned medium was isolated from gelatin-Sepha-

zymography. Pro-MMP-13 was activated with APMA for 3 hr at room temperature, and MMP-7 required no activation.
to draw meaningful conclusions. Therefore, the modeling work was concentrated on groups I and III. The molecular models of the sequences flanking the sites of MMP-9 cleavage in clones A10 and C11 (Ac-SGPLFYSVTANH2 and Ac-SGRRLHHTANH2, respectively) were constructed using the SYBYL molecular modeling program and were docked into the active site of MMP-9 (42). Each enzyme-substrate complex contained the hydrolytic water molecule positioned in the active site near the Glu-402 side chain, the general base that promotes it for the hydrolytic reaction. The enzyme-substrate complexes were energy-minimized using the AMBER 5.0 software package (44, 45). The protocol for the energy minimization was described previously (42).

RESULTS

Construction of the Phage Library—We created a system for displaying random hexamer substrates on the surface of phage. The fUSE5 polyvalent phage display vector (24–26) was modified to express random hexamers at the amino terminus of the geneIII protein, and an octapeptide FLAG epitope at the amino-terminal end of the random hexamers. The library comprises 2.4 x 10^10 independent transformants, giving a 75% confidence that each of the 6.4 x 10^10 possible random hexamer sequences are represented in the library. Sequencing of phage confirmed the randomness of the hexamer insert. Under the selection conditions, greater than 95% of phage could be immunodepleted using an anti-FLAG antibody (data not shown).

Selection of Peptide Substrates for MMP-9—Optimal substrates were selected by exposing the phage library to a recombinant form of the catalytic domain of MMP-9 expressed in HEK 293 cells (29). The recombinant catalytic domain of MMP-9 was purified on gelatin-Sepharose, followed by ion exchange chromatography (29). The protease was activated with 2 mM APMA for 16 h at room temperature, and the active site was titrated with the hydroxamate inhibitor Ibonastat (29, 30). Phage selections were performed with 2.5 µg/ml (56 nM) active MMP-9. Following three rounds of exposure to MMP-9, individual phage clones were selected for sequencing. An alignment of the motifs revealed three groups of structurally distinct substrates (Table I). Substrates in group I contain the motif with sequence Pro-X-X-Hy-(Ser/Thr). Further analysis shows that, within this larger motif, Arg is favored at P2 and Ser/Thr is favored at P3 (consensus motif noted in Table I). Substrates in group II contain a Gly-Leu-(Lys/Arg) motif. Interestingly, both groups I and II are related to substrates that have been described previously for other MMPs and are somewhat related to sites of cleavage in collagen (21–23). Group III substrates, however, appear to represent a novel recognition sequence as it contains the Arg-Arg-Hy-Leu (group IIIA) and Arg-X-Leu (group IIIB) motifs. The two subclasses that comprise group III have not been described previously as a substrate motif for MMP-9.

The ability of MMP-9 to hydrolyze each of the phage clones was assayed in a semiquantitative manner using a modified ELISA. Individual phage clones were captured into microtiter plates using anti-M13 antibody. Captured clones were exposed to MMP-9 (2.5 µg/ml). A polyclonal anti-FLAG antibody was used to measure the liberation of the FLAG epitope by hydrolysis of the substrate with MMP-9. Results are expressed as the extent of hydrolysis relative to untreated phage clones (Table I). In general, the group I substrates were hydrolyzed most efficiently. However, a number of the substrates within groups II and III were cleaved to the same degree as group I substrates. Identification of Scissile Bonds within MMP-9 Substrates—To identify the position of the scissile bonds, 10 peptide products by MALDI-TOF analysis (Table II). This analysis revealed that each substrate contains a hydrophobic residue at the P1 position. The substrates in group I contain Pro at the P2 residue, and a Lys at P3. The motifs of groups I and II, and locations of their scissile bonds, are similar to MMP cleavage sites in collagen and gelatin (21–23). Interestingly, however, the peptides from group III were all cleaved after an Arg residue, and in several cases an Arg-Arg occupied the P1 and P2 positions.

Kinetic Characterization of Substrate Hydrolysis by MMP-9—The Michaelis constant (Km) and first-order rate constant of substrate peptide turnover (kcat) were measured by incubating a range of each peptide with MMP-9. Peptide hydrolysis was measured by incorporation of fluorescamine onto newly formed amino termini as previously described (31, 33, 34). From these measurements, kcat and Km were derived for each peptide using double-reciprocal plots of 1/[S] versus 1/v. Results from this

| Group | Peptide sequence | Clone | % Hydrolysis |
|-------|-----------------|-------|-------------|
| I     | P L F Y S V     | A10   | 100 100     |
|       | K I P R T L T   | A11   | 100 100     |
|       | P L R L S W     | A12   | 100 100     |
|       | P R A V S T     | A6    | 100 99      |
|       | K G P R Q I T   | C15   | 100 98      |
|       | P R P L S G     | D12   | 82 94       |
|       | W P L G L A     | D35   | 89 91       |
|       | F R P R S I T   | D36   | 87 91       |
|       | R L P V G L T   | D20   | 89 90       |
|       | R S P K S L T   | D21   | 88 90       |
|       | P V W L A A     | D22   | 85 90       |
|       | I H P S S L T A  | D4    | 87 89       |
|       | G Q P H L S T   | D29   | 87 86       |
|       | M K P A S W T    | B2    | 77 75       |
|       | T H P Y T M T    | B10   | 74 82       |
|       | T P A Y M L T    | A13   | 78 79       |
|       | P L Y L T       | D2    | 72 85       |
|       | P G L I G T     | D1    | 75 85       |
|       | S P A N S N      | D24   | 82 82       |
|       | R L P A S Y T    | D30   | 82 60       |
|       | N P P R Y L T    | D5    | 70 67       |
|       | P P K T Q I S    | D8    | 73 65       |
|       | S V K P R F T    | D17   | 70 52       |
|       | L L L P A W L T  | D5    | 56 48       |
|       | T H P Y T M T    | D14   | 50 47       |
|       | L R P A K S T    | A18   | 28.9 30     |
|       | S G P S T S T    | A9    | 13.1 5.8    |

Consensus: P R S/T Hy S/T

II

| Group | Peptide sequence | Clone | % Hydrolysis |
|-------|-----------------|-------|-------------|
|       | G S G L K A     | A7    | 79.8 78     |
|       | A M G L K S     | B9    | 73.8 86     |
|       | K V G L R T     | B5    | 13.3 12     |

IIIA

| Group | Peptide sequence | Clone | % Hydrolysis |
|-------|-----------------|-------|-------------|
|       | G R R L I H     | C11   | 94.9 74     |
|       | H P R R S I T    | C1    | 13.9 29     |
|       | A L P L L S R A  | R16   | 20.2 11     |
|       | A L R R L E T    | C13   | 8.3 24      |
|       | F Y K R L V T    | B6    | 11.7 11.1   |
|       | F R R C V        | C7    | 17.9 0      |
|       | V F F R R Q T A  | C8    | 8.9 0       |

IIIB

| Group | Peptide sequence | Clone | % Hydrolysis |
|-------|-----------------|-------|-------------|
|       | G L A R N I T A  | B3    | 66 70       |
|       | F G S R Y L T A  | A19   | 45 50       |
|       | Q D R Y L N T    | C6    | 40 27       |

Table I

Sequence alignment and grouping of phage selected by MMP-9

Phage from the third round MMP-9 selection were selected for their ability to be cleaved by MMP-9 and sequenced. The substrates are separated into three groups based on distinguishable motifs. Residues that define each group are in bold. The ability of MMP-9 to cleave the substrates on phage was measured in a modified ELISA format. A rank order of preference within each group is demonstrated as a function of phage cleaved by MMP-9 relative to a non-treated control. This is expressed as percentage of hydrolysis. The results of two independent experiments are shown. Hy, hydrophobic amino acids.
Each peptide was incubated with protease and hydrolysis was detected as described under “Experimental Procedures.” The peptides are named according to their phage clone designation. All peptides were synthesized with acetylated NH₂ termini and amidated COOH-termini. A11m1 and A11m2 are analogues of peptide A11. Changes are indicated in bold. Scissile bonds are designated with **. Amino acids to the right of the ** are primed residues (\(\ldots P_1 \ldots P_n \ldots\)), and amino acids to the left are non-primed residues (\(\ldots P_n \ldots P_1 \ldots\)). ** Measurements were not made; * not detected. Standard deviations of triplicate experiments ranged from 11% to 22%.

### TABLE II
Kinetic analysis of MMP-9 peptide substrates based on phage clones

| Peptide | Sequence | MMP-9 | MMP-13 | MMP-7 |
|---------|----------|-------|--------|-------|
| | | \(k_{cat}/K_m\) | \(k_{cat}\) | \(K_m\) | \(k_{cat}/K_m\) | \(k_{cat}/K_m\) |
| | \(\mu M^{-1} s^{-1}\) | s⁻¹ | M⁻¹ | s⁻¹ | \(\mu M^{-1} s^{-1}\) | \(\mu M^{-1} s^{-1}\) |
| C15 | SGGKPQ | ITA | 188,000 | 15,400 | 13,800 |
| A11 | SGKIPRT | LTA | 67,500 | 23,500 | 7,000 |
| A6 | SGPR | VSTTA | 61,000 | 12,400 | 1,300 |
| A11m1 | SGKIPRTR | AATA | 160,000 | 15,200 | 24,000 |
| A11m2 | SGKIPRT | AATA | 2000 | 4.7 | ** |
| A10 | SGLFL | YVTTA | 25,600 | 0.5 | ** |
| B1 | SQPHY | LTTA | 8800 | 9 | ** |
| A18 | SGLRP | STA | 5800 | 13.5 | ** |
| A7 | SG | LKALMITA | 12,400 | 12,300 | 41,000 |
| A19 | SGFSRY | LTA | 13,200 | 2.4 | ** |
| A16 | SGRR | LLSRTA | 1200 | 1.4 | 1000 |
| C11 | SGRR | LIHHTA | * | * | * |

Analysis are shown in Table II. Among the peptides, the \(k_{cat}\) values ranged from 9 s⁻¹ to 703 s⁻¹, and \(K_m\) values were generally in the high micromolar to low millimolar range. Overall, peptides from group I appeared to represent the best substrates, with peptide C15 having a \(k_{cat}/K_m\) ratio of 180,000 M⁻¹ s⁻¹, the highest of any of the substrates. Peptide A11, which closely matches the consensus recognition motif of group I, also exhibits a high \(k_{cat}/K_m\) ratio (67,500 M⁻¹ s⁻¹) and considerable selectivity for MMP-9.

Interestingly, the three peptides cleaved most efficiently by MMP-9 were from group I and contained Arg at the \(P_1\) position. This led us to hypothesize that group I and III substrates might represent a larger set of substrates whose relationship to one another is not entirely evident from the sequences of only 100 clones. To test this idea, we synthesized a mutant peptide (A11m1) that contains a Thr → Arg substitution at \(P_1\) within the context of the sequence of the A11 peptide. Hence, the mutant peptide contained elements of both group I (Pro at \(P_3\) and Leu at \(P_1\)) and group III (Arg at \(P_2\) and \(P_1\)). Peptide A11m1 had a \(k_{cat}/K_m\) twice that of the parent peptide, an effect resulting primarily from an increase in \(k_{cat}\). This finding suggests that the substitution of Arg at \(P_1\) lowers the transition state energy of the protease-substrate interaction. This observation also indicates that, within the Pro-X-X-Hy (Ser/Thr) motif, Arg residues at \(P_2\) and \(P_1\) are favored for MMP-9, and that the substrates from groups I and III may have a related binding mechanism.

In almost every case, the MMP-9 substrates selected from the phage library contain a hydrophobic residue at \(P_1\), a finding that is entirely consistent with the fact that MMPs are known to have a deep hydrophobic S₁ pocket (1, 2, 46). To assess the contribution of this hydrophobic residues to substrate binding and to substrate turnover, we synthesized another mutant peptide (A11m2) based on the sequence of peptide A11, but containing Ala, rather than Leu, at \(P_1\) (Ac-SGKIPRT \(\rightarrow\) ATA-NH₂). This substitution had deleterious effects on both \(k_{cat}\) and \(K_m\), and reduced the \(k_{cat}/K_m\) ratio nearly 30-fold. These results are consistent with the idea that the S₁ subsite of MMP-9 coordinates substrate binding and also influences the rate of hydrolysis. Importantly, however, even this mutant peptide had a measurable \(k_{cat}/K_m\) ratio (2000 M⁻¹ s⁻¹), indicating that efficient hydrolysis can be enacted by MMP-9 if the rest of the substrate sequence is optimal.

**Kinetic Characterization of Substrate Hydrolysis by MMP-7 and MMP-13**—Since many of the selected substrates contained a motif similar to that described for other MMPs (P-X-X-Hy), we measured the degree to which MMP-7 and MMP-13 could cleave these MMP-9 substrates. These two MMPs were used for comparison because substrates for both proteases have been selected using a similar phage display approach. Interestingly, most of the substrates we tested were cleaved more efficiently by MMP-9. The \(k_{cat}/K_m\) ratios ranged from 2.6- to 47-fold higher for MMP-9 than for either MMP-7 or MMP-13 (Table II). Only peptide A7 deviated from this trend. It is also worth noting that, although the substrates in group IIIA (Arg-Arg) were not cleaved rapidly by MMP-9, we were unable to detect any hydrolysis of these peptides by either MMP-7 or -13.

**Modeling Substrate Interactions with MMP-9**—Molecular modeling studies were conducted to help visualize how substrates might dock into the enzymatic cleft of MMP-9. Energy-minimized models of MMP-9 were constructed according to procedures described under “Experimental Procedures,” and then docked with peptides A10 (Ac-SGPLFYSVTA-NH₂) representing group I and C11 (Ac-SGRRLIHHTA-NH₂) representing group III.

A primary feature of the group I substrates is a Pro residue at the \(P_1\) position. The unique conformational features of Pro introduce the appropriate “bend” needed for the optimal substrate binding at this site. This Pro residue, occupying the S₃ subsite of MMP-9, is illustrated in Fig. 1A (white arrow). The group I substrates also contain an invariant hydrophobic residue at the \(P_1\) position. This residue protrudes into the deep S₁ pocket of the protease, as depicted by the orange arrows.

Amino acids with long basic side chains at \(P_2\) and \(P_1\), such as Arg, are the defining features of the substrates in group III. Although the presence of these residues at \(P_3\) and \(P_1\) is somewhat surprising, the energy-minimized models support the observation that these residues bind favorably. An Arg at \(P_1\) is likely to interact with the backbone carbonyl moieties of His-405, Gly-408, and the side chain of Asp-410, all of which contribute to the S₂ subsite within MMP-9 (white arrow in Fig. 1B). These electrostatic interactions are predicted to contribute to favorable binding of this class of substrates. Many of the group III substrates also contain an Arg residue at \(P_4\). The favorable interaction of Arg into the S₃ subsite can be explained by the somewhat unusual nature of this subsite. It is essentially a hydrophobic binding surface that would be predicted to accommodate the hydrophobic side chains of amino acids such as Ala, Phe, and Tyr. However, the docking studies of peptides with Arg at \(P_1\) show that the hydrophobic surface of the S₂ subsite could also bind to the extended methylene group in the
side chain of Arg. In addition, the hydrophobic channel of the S₂ subsite contains the backbone carbonyl of Pro-180, which is likely to engage in electrostatic interaction with the basic side chain of Arg, stabilizing the interaction.

**DISCUSSION**

Because of their association with a number of diseases, the MMPs have received considerable attention as drug targets (15–17). Much of the effort in this area has focused on the design of small molecule antagonists with two primary features; 1) a hydroxamate moiety that binds to the proteases catalytic zinc, and 2) a rather large hydrophobic moiety that fits into the deep S₁ pocket present in all MMPs (30). This synthetic strategy has focused structure-activity studies to essentially two positions within the catalytic pocket. An understanding of the interactions between the substrate and other key subsites within the catalytic pocket is lacking.

We have identified three families of peptide substrates for MMP-9 that each appear to interact differently with the catalytic cleft of the protease. Substrates in group I are cleaved most efficiently by MMP-9. These peptides all contain a Pro at P₃ and a large hydrophobic residue at P₁. In this respect, the group I substrates are similar to collagen-like sequences that are known to be cleaved by the MMPs (19, 20). A prior analysis of a small series of synthetic peptides based on collagen, and containing a similar Pro-X-X-Hy core, showed k_{cat}/K_m ratios ranging from 340 to 1000 mM⁻¹ h⁻¹. These values are ~100 fold lower than those exhibited by the best peptides reported here. Consequently, the added diversity afforded by phage libraries allowed us to identify better substrates and a wider range of substrates. Interestingly, many different MMPs recognize the Pro-X-X-Hy core sequence. Can subsite interactions at positions other than P₃ and P₁, generate selectivity within this family of sequences?

FIG. 1. Three-dimensional models of distinct binding modes between substrate and MMP-9. To ascertain whether peptides from different substrate families exhibit different modes of binding in the active site of MMP-9, models with two representative peptides binding to MMP-9 were constructed. The stereo view represents the energy-minimized complexes of peptide A10 (SGPLFYSVTA, panel A) and peptide C11 (SRRRLRSRTA, panel B) in the active site of the MMP-9 catalytic domain. The active site of MMP-9 is shown as a green Conolly surface, and the peptide atoms are colored according to type (carbon, white; oxygen, red; nitrogen, blue). The catalytic zinc ion is represented as a yellow sphere, and the backbone of MMP-9 outside the active site is shown in magenta. The orange arrows in both panels show the deep hydrophobic pocket at the S₁ subsite. In panel A, the white arrow shows the Pro of substrate A10 penetrating well into the S₃ pocket. In contrast, a Gly from substrate C11 occupies this subsite (panel B).

Apparently, individual MMPs do exhibit a great deal of selectivity for peptides containing the P-X-X-Hy core sequence. For example, most of the MMP-9 substrates within group I are selective for MMP-9 over MMP-7 and MMP-13. Some of the k_{cat}/K_m ratios for these peptides are up to 47-fold higher for MMP-9 than for the other MMPs tested. These findings suggest that substrate selectivity can be conferred by subsite interactions outside of the dominant P₃ and P₁ subsites that are common among MMP substrates.

Since phage substrate selections have been performed for MMP-3, -7, -9, and -13, an analysis of the frequency by which individual residues occupy distinct subsites can be used as a first test of this idea. This comparison reveals considerable distinction in the residues that occupy the P₂, P₁, and P₃ subsites. Nearly one-third of all group I substrates for MMP-9 contain Arg at P₂. Although Arg can also be found at P₃ in peptide substrates for MMP-13, its frequency is much lower than in the MMP-9 substrates we selected (47). Furthermore, Arg is rarely, if ever, found at P₂ in MMP-3 or -7 peptide substrates (32). Ser or Thr most frequently occupies the P₁ position of the MMP-9 substrates. However, a Gly residue is preferred at this position by MMP-13, and Asp or Glu are preferred by MMP-3 and -7 (32, 47). Significant differences are also observed at P₂. In the group I MMP-9 substrates, 23 of 28 substrates have Ser or Thr at P₂, but neither residue is prevalent at this subsite in the substrates selected for the other MMPs (32, 47). These observations support the contention that subsite interactions other than P₂ and P₁ have a significant impact on substrate selectivity among the MMP family.

Here we observed two additional families of substrates for MMP-9 that are distinct from the P-X-X-Hy family. Group II substrates were selected least frequently and contain only three members, making it difficult to identify subsite prefer-
Substrate Recognition by Matrix Metalloproteinase-9

The group I MMP-9 consensus motif is compared to the cleavage sites in proteins known to be cleaved by MMP-9. The protein substrates are divided into two groups: collagen and non-collagen. The residues in the protein substrates that match the MMP-9 consensus are shown in bold. Scissile bonds are identified by ↓.

### Table III

| MMP-9 consensus collagen substrates | P R S/T | ↓ | Hy S/T |
|-----------------------------------|---------|---|--------|
| α1(I) collagen                    | P Q G ↓ V R |   |        |
| CB3 α1(IV) collagen               | P Q G ↑ V Q |   |        |
| α1(V) collagen                    | P Q G ↓ V V |   |        |
| α1(XI) collagen                   | P Q G ↓ L R |   |        |
| Non-collagen substrates           |          |   |        |
| Aggrecan                          | P E N ↓ F F |   |        |
| Tissue factor pathway inhibitor   | P L K ↓ L M |   |        |
| Galectin 3                        | P G A ↓ Y H |   |        |
| Cartilage link protein            | A I H ↓ I Q |   |        |
| Myelin basic protein              | H F F ↓ K N |   |        |
| Glial S ↓ L S                    | G L S ↓ L S |   |        |
| A S D ↓ Y K                      | A S D ↓ Y K |   |        |

This comparison is used to query the SWISSPROT data base to identify potential physiological substrates of MMP-9. The search was limited to proteins that contained these potential MMP-9 cleavage sites. The putative protein substrates and potential cleavage sites are shown.

### Table IV

| MMP-9 consensus: | P R S/T | Hy S/T | Potential protein substrate (residues) |
|------------------|---------|--------|--------------------------------------|
| α1(I) collagen   | P Q G ↓ V R |   | Kallikrein 14 (180–184) |
| CB3 α1(IV) collagen | P Q G ↑ V Q |   | Ladinin 1 (371–375) |
| α1(V) collagen   | P Q G ↓ V V |   | Endoglin 231–235 |
| α1(XI) collagen  | P Q G ↓ L R |   | Endothelin receptor 82–86 |
| Non-collagen substrates |          |   | Laminin α3 chain 1568–1572 |
| Aggrecan         | P Q G ↓ F F |   | Phosphate regulating neutral endopeptidase 566–570 |
| Tissue factor pathway inhibitor | P Q G ↓ L M |   | ADAM 2 299–302 |
| Galectin 3       | P Q G ↓ Y H |   | Desmoglein 3 579–583 |
| Cartilage link protein | A Q I H ↓ I Q |   | Integrin β3 57–61 |

Given this new understanding of substrate recognition by MMP-9, one can arrive at an optimal substrate consensus (Pro-Arg-(Ser/Thr)-↓-Leu/Ile)-(Ser/Thr)). A comparison of this motif to previously reported substrates for MMP-9 reveals some interesting similarities and distinctions (Table III). For example, all of the collagen substrates and even aggrecan, tissue factor pathway inhibitor and galectin 3 contain a Pro at the P3 position and a hydrophobic amino acid at the P1 position (6, 48, 50–52). However, none of the substrates contain residues found to be optimal at other positions. The remaining substrates have significantly lower homology to the MMP-9 consensus recognition motif of group I (8, 53).

These comparisons raise questions as to whether all of the physiologically relevant substrates for MMP-9 have been identified. We reasoned that a query of the protein data bases with the optimal recognition motif might reveal a short list of putative substrates. Indeed, this search revealed only nine human proteins that contain this sequence within their extracellular domain (Table IV). Although at this juncture these proteins can only be considered hypothetical substrates for MMP-9, many of them are functionally linked to MMP-9-related pathologies. For example, MMP-9 influences skin blistering in the autoimmune disorder bullous pemphigoid (5, 55). Interestingly, two of the hypothetical substrates for MMP-9, ladinin 1 and desmoglein 3, are also associated with autoimmune skin blistering disorders (56, 57). Similarly, MMP-9 has been suggested as a regulator of the angiogenic switch in tumor development (11). Two of the hypothetical substrates for MMP-9, integrin β1 and endoglin 1, are also involved in angiogenesis (49, 54, 58). Along with the presence of a potential MMP-9 cleavage site, these functional associations lead to important and testable hypotheses about physiologic substrates for MMP-9.
