Substrate cleavage profiling suggests a distinct function of *Bacteroides fragilis* metalloproteinases (fragilysin and metalloproteinase II) at the microbiome – inflammation - cancer interface§

Sergey A. Shiryaev1, Albert G. Remacle1, Andrei V. Chernov1, Vladislav S. Golubkov1, Khatereh Motamedchaboki1, Norihiro Muranaka2, Corey M. Dambacher2, Petr Capek2, Muskan Kukreja2, Igor A. Kozlov2, Manuel Perucho1, Piotr Cieplak1 and Alex Y. Strongin1*

1Sanford-Burnham Medical Research Institute, La Jolla, CA 92037, and 2Prognosys Biosciences Inc., San Diego, California, CA 92121

*To whom correspondence should be addressed: (strongin@sanfordburnham.org)

§The work reported here was supported by R01CA83017, R01CA157328 and R01DE022757 (AYS), R44GM085884 (IAK), and R01GM98835 (PCieplak) grants from the NIH.

Running title: *Bacteroides fragilis* toxigenic metalloproteinases

**Background:** Two distinct metalloproteinase types (fragilysin and metalloproteinase II/MPII) are encoded by *Bacteroides fragilis* pathogenicity island.

**Results:** Our assays determined substrate cleavage characteristics of fragilysin and MPII.

**Conclusion:** MPII is the first zinc metalloproteinase with the dibasic cleavage preferences.

**Significance:** Our results are important for understanding *Bacteroides fragilis* virulence and fundamental roles of the microbiome in human health and disease.

**Summary**

Enterotoxigenic anaerobic *Bacteroides fragilis* is a significant source of inflammatory diarrheal disease and risk factor for colorectal cancer. Two distinct metalloproteinase types [the homologous 1, 2 and 3 isoforms of fragilysin (FRA1, FRA2 and FRA3, respectively)] and metalloproteinase II (MPII) are encoded by *B. fragilis* pathogenicity island. FRA was demonstrated to be important to pathogenesis, while MPII, also a potential virulence protein, remained completely uncharacterized. Here, we, for the first time, extensively characterized MPII in comparison with FRA3, a representative of the FRA isoforms. We employed a series of multiplexed peptide cleavage assays to determine substrate specificity and proteolytic characteristics of MPII and FRA. These results enabled implementation of an efficient assay of MPII activity using a fluorescent-quenched peptide, and contributed to structural evidence for the distinct substrate cleavage preferences of MPII and FRA. Our data imply that MPII specificity mimics the dibasic ArgArg cleavage motif of furin-like proprotein convertases, while the cleavage motif of FRA [Pro-X-X-Leu-(Arg/Ala/Leu)] resembles that of human MMPs. To the best of our knowledge, MPII is the first zinc metalloproteinase with the dibasic cleavage preferences, suggesting a high level of versatility of metalloproteinase proteolysis. Based on these data, we now suggest that the combined (rather than individual) activity of MPII and FRA is required for the overall *B. fragilis* virulence in vivo.

A variety of microbial communities (the microbiome) exists in the human body, playing fundamental roles in human health and disease (1,2). Normally, bacteria outnumber human cells within an individual by at least an order of magnitude. The gastrointestinal tract is part of the direct interface between the human organism and the microbiome. In certain circumstances, the beneficial relationship between the microbiome and the gastrointestinal tract becomes disrupted, causing disturbances leading to disease.

Chronic inflammation affects all phases of carcinogenesis, from favoring the initial genetic alterations that drive cancer formation, to acting as a tumor promoter by establishing conditions in the
surrounding tissues that allow the tumor to progress and metastasize (3-6). For example, chronic hepatitis B or C virus infections frequently leads to liver cancer (7), and chronic Helicobacter pylori infection leads to gastric cancer in some patients (8-10). Increased cancer incidence is likewise found in experimental mouse models of both infection-induced and noninfectious inflammation (11,12).

The role of infectious and inflammatory processes in colon carcinogenesis is of great interest. Enterotoxigenic Bacteroides fragilis (B. fragilis) is both a significant source of chronic inflammation (e.g., inflammatory diarrhea and ulcerative colitis) as well as risk factor for colorectal cancer (CRC) (4,13-19). B. fragilis comprises typically only 0.5-2% of the cultured fecal flora (20-23) but causes over 80% of anaerobic infections (24). It is likely that the pro-inflammatory, pro-tumorigenic role of B. fragilis in CRC and Helicobacter pylori in stomach/gastric cancer is similar (4,19,21,23,25).

There is a consensus among researchers that metalloproteinase activity is essential for B. fragilis virulence and that this activity is encoded by the 6-kb pathogenicity island in enterotoxigenic B. fragilis strains (14,21,26,27). The island contains, at least, two metalloproteinase genes. These genes encode fragilysin (FRA; also termed as B. fragilis toxin or BFT), demonstrated to be important to pathogenesis, and metalloproteinase II (MPII), also a potential virulence protein. FRA exists in three homologous isoforms (FRA1, 2 and 3) with the sequence identity over 90%. In turn, sequence identity between FRAs and MPII is only 25% (Fig. 1).

FRAs and MPII are secretory zinc-metalloproteinases with a zinc-binding HEXXXHXXGXXH motif and a characteristic Met-turn. These structural features, especially when combined, indicate that both FRAs and MPII exhibit the Matrix Metalloproteinases (MMP)/A Disintegrin and Metalloproteinase (ADAM) fold (28-31). The overall level of homology between the catalytic domain of bacterial FRAs and MPII, and mammalian MMPs/ADAMs, however, is low.

While there is limited information about the structural-functional features of FRAs (17,32), the biochemical characteristics of MPII remain completely unknown. Because MPII has not been characterized as yet, it is unclear if MPII can facilitate the toxigenic effect of FRAs in causing diarrhea, inflammatory bowel disease and CRC (14,26,30,33,34). As a result, we cannot decipher, at the molecular level, how the proteolytic activity of B. fragilis tailors the normal luminal epithelium for inflammation and disease onset. Understanding the substrate cleavage specificity of MPII relative to FRAs may help to determine how infection-associated inflammation enhances carcinogenesis in the affected organs and how we find a means to fight the disease.

Here, we performed a comparative characterization of MPII and FRA3, a representative of the FRA isoforms. Our data imply that, in contrast with the FRA family members, the unconventional MPII cleavage preferences mimic those of furin-like protoprotein convertases. To the best of our knowledge, MPII is the first zinc metalloproteinase with the dibasic cleavage preferences, suggesting a high level of versatility of metalloproteinase proteolysis. Based on our results, we suggest the combined (rather than individual) activity of MPII and FRAs is required for B. fragilis virulence.

Materials and Methods

Reagents. The reagents were purchased from Sigma-Aldrich, unless indicated otherwise. 5-FAM-SLGRKIQIQK(QXL520)-NH2 fluorescent-quenched peptide substrate was acquired from AnaSpec. GM6001/Iломastat, BB94/Batimastat and AG3340/Prinomastat were obtained from EMD Millipore, Tocris Biosciences and Allergan, respectively. Anthrax protective antigen-83 (PA83) was purchased from List Biological Laboratories. Recombinant human TIMP-2 was expressed in Madin-Darby canine kidney (MDCK) cells and then purified from conditioned medium as reported earlier (35). Human TIMP-1 and TIMP-3 were purchased from Invitrogen.

Detection of B. fragilis DNA in the patient samples. The frozen tumor and matching normal tissue de-identified biopsies were obtained from our pre-existing collection of proximal CRC cancer specimens. Genomic DNA was extracted from the tissue samples using the DNeasy Blood
and Tissue DNA purification system (Qiagen). The 501-bp fragment of the *B. fragilis* 16S rRNA gene was amplified in the 100 µl PCR reactions containing genomic DNA (100 ng), the forward and reverse primers (5′-ATAGCCCTTTCGAAAGRAAGAT-3′ and 5′-CCGATATCAACTGCAATTTTA-3′ respectively; 0.3 µM each), Crimson Taq DNA polymerase (1 unit) and 12.5 mM Tricine buffer, pH 8.5, supplemented with 42.5 mM KCl, 1.5 mM MgCl<sub>2</sub>, 6% Dextran and 0.2 mM dNTP mix. DNA amplification were performed using denaturing of the samples at 95°C for 5 min followed by 35 PCR amplification (95°C for 30 s, 52°C for 30 s, 72°C for 1 min). The products were separated by 2% agarose gel-electrophoresis. Amplified 501-bp products were purified and sequenced to confirm their authenticity and identity. Two-sided Fisher’s exact test was used to evaluate the statistical significance of the association of the bacteria with colorectal cancer.

Cloning, expression and purification of MPII and FRA3. The full-length cDNA coding for the wild type MPII proenzyme (gi 3046922) and the FRA3 proenzyme (PDB accession 3P24; gi:315583580) were synthesized by Genewiz. PCR with the 5′–CACCATGCACCATCACCATCACCATGGAC
CCTGTGCGATGACCTG–3′ and 5′–TCAATGGTGGTGATGGTGGTGCTTGTCATC
GTCATCTTTGTAGTCCTTTTGGATGCACTC
CAG–3′ oligonucleotides as the forward and reverse primers, respectively, was then used to insert the 6xHis tags (both N- and C-terminally) and the Flag tag (C-terminally) into the MPII template. Similarly, the 5′–CACCATGCACCATCACCATCACCATGGAG
CCTGCAGCAATGAGGCC–3′ and 5′–GATGGCCCACGCACTGGCCACATC–3′ as the forward and reverse primers, respectively, and the wild type FRA3 template were used to generate the FRA3 E349A inactive mutant (mutant nucleotides are underlined). The authenticity of the constructs was confirmed by DNA sequencing.

The wild type and mutant tagged constructs were re-cloned into the pET101 expression vector (Invitrogen). The recombinant pET101 plasmids were used to transform competent *E. coli* BL21 (DE3) Codon Plus cells (Stratagene). Transformed cells were grown at 30°C in LB broth containing ampicillin (0.1 mg/ml). Cultures were induced with 0.6 mM IPTG for 16 h at 18°C. Cells were collected by centrifugation, re-suspended in Tris-HCl buffer, pH 8.0, containing 1 M NaCl, and disrupted by sonication. The pellet was removed by centrifugation (40,000g; 30 min). The constructs were then purified from the supernatant fraction on a Co<sup>2+</sup>-chelating Sepharose Fast Flow column (GE Healthcare), equilibrated with 20 mM Tris-HCl buffer, pH 8.0, supplemented with 1 M NaCl. After washing out the impurities using the same buffer supplemented with 35 mM imidazole, the bound material was eluted using a 35-500 mM gradient of imidazole. The FRA fractions were combined, dialyzed against 20 mM Tris-HCl buffer, pH 8.0, containing 150 mM NaCl. The purified material was kept at -80°C until use. The purity of the material was tested by SDS gel-electrophoresis (12% NuPAGE-MOPS, Invitrogen) followed by Coomassie staining and by Western blotting with an anti-Flag tag antibody.

Activation of MPII and FRA3. The MPII E352A and FRA3 constructs (3 µg; 3 µM each) were co-incubated in PBS for 1 h at 37°C with increasing concentrations of trypsin. Similarly, FRA3 (3 µg; 3 µM) was co-incubated for 1 h at 37°C with increasing concentrations of active MPII in 50 mM HEPES buffer, pH 8.0, containing 1 mM CaCl<sub>2</sub> and 50 µM ZnCl<sub>2</sub>. The cleavage was stopped by adding a 5xSDS sample buffer to the reactions. The digested samples were analyzed by
SDS-gel electrophoresis in 12% NuPAGE-MOPS gels.

**Multiplex peptide cleavage assay.** Peptide pool preparation and the following *in vitro* cleavage assays were performed using methodology similar to those described in our recent publications (36-38). Template DNA encoding 10 amino acid residue long peptide substrates was transcribed *in vitro* to produce the corresponding mRNAs. *In vitro* translation was then used to generate peptides covalently attached to their mRNA templates (39). To increase their stability, the peptide-mRNA fusions were converted to the corresponding covalent peptide-cDNA fusions (40). The peptide-cDNA fusions were biotinylated at the N-terminus and immobilized on streptavidin-coated magnetic beads. We also spiked six biotinylated oligonucleotides into the peptide-cDNA pool to use as internal standards for normalization after PCR reactions. The magnetic beads with the immobilized peptide-cDNA pool (2 pmol) were co-incubated with MPII and FRA3 (1.2 or 12 pmol) at 37°C for 30 min, in 3 μl reactions containing 50 mM HEPES, pH 6.8, 10 mM CaCl₂ and 10 μM ZnCl₂. Reactions without the proteinases were used as negative controls. To identify cleaved peptide substrates, the cDNA molecules released by the proteinase treatment were collected, the DNA adapters required for sequencing were installed by PCR (41), and the obtained DNA constructs were sequenced using MiSeq sequencing instrument (Illumina). In this study, we produced and tested 316 peptide sequences.

Peptide abundance in solution was quantified by counts of DNA reads corresponding to each peptide sequence. The cleavage efficiencies were determined by comparing counts in the proteinase-treated samples versus proteinase-less, buffer-only controls (in triplicate each). Next, we normalized the data for each of six different biotinylated substrates (internal standard) across the whole data set. In order to calculate an enrichment ratio for each substrate, we considered counts (in triplicate) in a positive sample over average of counts of a control sample (buffer-only, no proteinase) +3X standard deviation. An enrichment Ratio (ER) = 1 indicates that counts in a positive sample are 3 standard deviations above the “buffer-only, no proteinase” negative sample.

**Mass-spectrometry analysis of the cleavage peptides.** Eleven 10-residue long peptides (GHSRRSRRSG, AGLRRAAALGG, AGLRRASLGG, GRHRQQIDRG, GNKRRGGTAG, SGHMHAALTA, SGPVSMMRFTA, SGPRSLKSTA, SGPMSLRMTA and PTKIYDNIYD) were synthesized by Spyder Institute (Prague, Czech Republic). The peptides (1 μg; 50 μM each) were incubated for 1 h at 37°C with the MPII or FRA3 (0.4 μg; 0.5 μM) in 20 μl 50 mM HEPES buffer, pH 8.0, containing 1 mM CaCl₂, 0.5 mM MgCl₂ and 10 μM ZnCl₂. The molecular mass of the intact peptides and the digest products was determined by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI–TOF MS) analysis using an Autoflex II mass-spectrometer (Brucker Daltonics).

**Sequence logos.** The sequence logos were obtained by calculating cleavage efficiency over the entire set of substrates and then selecting substrates with ER value >3. These substrates were considered susceptible to proteolysis. Substrates with ER value <3 were considered resistant to proteolysis. The resulting logos were created by the web-based IceLogo program (42).

**Molecular modeling of MPII, FRA1, FRA2 and FRA3.** Molecular modeling of the substrate complexes of MPII and FRA3 was initiated from the crystal structure of FRA3 (PDB 3P24) (30). In this structure, the β11 strand of the FRA3 prodomain directly interacts with and tightly fits into the active site of the catalytic domain. The strand, however, runs in the opposite orientation relative to the substrate. To model the PRPLRAWGAA substrate, the Leu186-Val201 segment of the β11 strand was deleted *in silico*. The substrate was then modeled by placing the substrate's side chain heavy atoms at the positions of the heavy atoms in the β11 strand. The further optimization of the main and side chain atoms of the substrate was performed using molecular mechanical minimization and short molecular dynamics simulations *via* the Amber12 modeling package and FF99SB force field, and generalized-Born method for representing solution environment implicitly (43-46). In the course of our molecular mechanical optimization procedure, the proper orientation of the carbonyl group of the
scissile bond relative to the active site Zn ion was maintained, and acetyl and N-methyl groups for capping the N- and C-termini of the modeled substrate were used. The structure of MPII was prepared by homology modeling using the PDB 3P24 as a template, and FFAS03 (47,48) and Modeller (49) software. The fold of the PGRLR↓RSGAA substrate mimics the fold of the PRPLR↓AWGAA substrate in the FRA3 structure. The structures of FRA1 and FRA2 were modeled using FRA3 (PDB 3P24) as a template and Modeller software. The final models were displayed using PyMol (www.pymol.org).

Activity assays with the fluorescent-quenched peptide. The assay for MPII cleavage activity was performed in 50 mM HEPES buffer, pH 8.0, containing 1 mM CaCl₂, 0.5 mM MgCl₂ and 10 μM ZnCl₂. The 5-FAM-SLGRKIQIQK(QXL520)-NH₂ substrate and enzyme concentrations were 10 μM and 50 nM, respectively. The total assay volume was 0.2 ml. Initial reaction velocities were monitored continuously at λ_ex=488 nm and λ_em=520 nm using a Spectramax Gemini EM fluorescence spectrophotometer (Molecular Devices). All assays were performed in triplicate in wells of a 96 well plate.

Where indicated, MPII (50 nM) was co-incubated for 30 min at 20°C with increasing concentrations of natural tissue inhibitors of MMPs (TIMP-1, TIMP-2 and TIMP-3) and small-molecule hydroxamate inhibitors of MMPs (GM6001/Ilomatstat, BB94/Batimastat and AG3340/Prinomastat). The residual activity of the samples was then measured as above. EC₅₀ values were calculated by determining the concentrations of TIMPs and hydroxamate inhibitors needed to inhibit 50% of the MPII cleavage activity. GraphPad Prism was used as fitting software.

MPII proteolysis of proteins in vitro. Anthrax PA83 (2 μg; ~1 μM) was co-incubated for 1 h at 37°C with increasing concentrations of MPII in 50 mM HEPES, pH 8.0, containing 1 mM CaCl₂, 0.5 mM MgCl₂ and 10 μM ZnCl₂. The total volume of the reactions was 20 μl. Where indicated, GM6001 (1 μM) was added to the reactions to inhibit MPII. The cleavage reaction was stopped by adding a 5xSDS sample buffer. The digest samples were analyzed by SDS-PAGE using 4-20% polyacrylamide gel.

Results

B. fragilis in human CRC. To test if B. fragilis is more prevalent in the CRC tumor biopsies as compared with the matching normal tissue, we isolated the total genomic DNA from 30 tumor biopsies and 30 matching normal tissue specimens. We then amplified by PCR the 501-bp fragment of the B. fragilis 16S RNA gene using the isolated DNA samples. Selected amplified bands were sequenced to confirm their B. fragilis identity. Our data clearly demonstrated the presence of the B. fragilis 16S RNA in 22 from 30 tumor biopsies (~73%) but only in 13 from 30 matching normal tissue samples (~43%), thus suggesting a statistically significant association of the bacteria with CRC (Fig. 2). In addition, the bacterium was found only in those matching normal tissues, which were derived from the patients with the B. fragilis-positive tumor samples.

Recombinant constructs of MPII and FRA3. To determine the potential pro-inflammatory, cancer-promoting role of B. fragilis toxin, we successfully cloned the full-length genes coding for B. fragilis MPII and FRA3. To facilitate their purification, the constructs were flanked by an N-terminal and C-terminal 6xHis tag and, additionally, by a C-terminal Flag tag. We also constructed the catalytically inactive MPII and FRA3 (E352A and E349A mutants, respectively) in which the catalytically essential Glu residue of the active site was mutated into Ala. The constructs were expressed in and purified from lysates of E. coli (Fig. 3). MPII was readily self-activated while the FRA3 proform was noticeably more stable and only a minor fraction of the active enzyme accumulated in the purified FRA3 samples during storage. As judged by the mobility of the digest products in SDS-gels and in agreement with the observations by others (30), trypsin proteolysis converted the FRA3 proenzyme into the mature proteinase (data not shown). In turn, MPII was unable to convert the FRA3 proenzyme into the active proteinase (data not shown), suggesting that either human trypsin or an unidentified bacterial proteinase activates FRA3 in the gut in vivo.

Multiplex cleavage assay of MPII and FRA3. In order to quickly characterize the substrate
preferences of the FRA cleavage, we used recently published technology for screening of customizable peptide pools. The utility of peptide-cDNA fusion pools in multiplexed assays were demonstrated in our previous reports using multiple proteinases, including human furin, hepatitis C virus NS3/4A proteinase, and thrombin (36-38).

The design of the set of 316 10-mer peptides used in the assay was based on the known cleavage site of FRA3 in E-cadherin (17,21) and on the consensus cleavage sites (Pro-X-X↓Leu) of human MMPs. The original cleavage data are presented in Table S1 and summarized in Table 1. To determine the position of scissile bonds, eleven representative individual peptides were re-synthesized using standard peptide synthesis methodology. The individual peptides were then digested by MPII and FRA3. The digest reactions were analyzed by MALDI-TOF mass spectrometry to identify resulting digest fragments and, consequently, the scissile bonds. The MS data directly correlated with the results of the multiplex assay. The representative mass spectrometry analysis data are shown in Fig. 4A.

Out of 316 tested peptide sequences >80 and >230 were cleaved by MPII and FRA3, respectively (ER>3). In order to identify the most probable cleavage sequences, we included only peptides with ER>3 in our analysis. While the peptide set used in the experiments was limited in size, it was sufficient to show a clear difference in substrate preferences of MPII versus FRA3 (Fig. 4B). Our data imply that MPII prefers Arg at both the P1 and P1’ positions. The presence of Leu at the P2 position, Arg or Gly at both the P3 and P3’ positions, and Gly at the P4 position are all characteristic of substrates efficiently cleaved by MPII. In turn, FRA3 prefers the presence of Leu and Pro at the P2 and P5 positions, respectively. Arg/Ala, Pro/Gly/Ser and Arg/Ala/Leu are well tolerated at the P4, P3 and P1 positions, respectively. Overall, the resulting cleavage signature of FRA3 is most similar to the Pro-X-X-Leu consensus cleavage motif of MMPs.

Inhibitors and fluorescence-quenched peptide substrates of MPII. In cells/tissues, the activity of MMPs/ADAMs is regulated by TIMPs (51,52). In addition, small-molecule inhibitors of a hydroxamate class are also readily available for the MMP/ADAM studies in vitro and in cell-based tests (53-56). These hydroxamates chelate the active site zinc atom and inactivate MMPs/ADAMs. Using the activity assay conditions we established, we identified compounds that inhibit MPII (Fig. 5C). The GM6001 hydroxamate inhibitor (N-[(2R)-2-(hydroxamidocarbonylmethyl)-4-methylpentanoyl]-L-tryptophan methylamide; potent wide-range inhibitor of multiple individual specificities) appeared to be cleaved at the Arg193-Lys-Lys-Arg196 furin cleavage motif. As a result, MPII transformed PA83, albeit less efficiently than furin, into the C-terminal PA63 and the N-terminal PA20 fragments (Fig. 4C).

In vitro cleavage assay of MPII activity. Our cleavage studies allowed us to elucidate a cleavage signature of MPII and FRA3 and thus to identify the commercially-available fluorescent quenched peptide 5-FAM-SLGRK↓IQIQK(QXL520)-NH₂ as a substrate for MPII. This peptide employs a 5FAM/QXL520 FRET pair for quantification of enzyme activity, providing a convenient direct assay for MPII activity. As we expected based on the multiplex assay results, this peptide is cleaved by MPII but not by FRA3. The availability of this peptide substrate allowed us to readily establish the fundamental enzymological features of the MPII activity, including its pH and temperature optimum dependence, effects of salts and glycerol (Fig. 5). Our results imply that MPII is a thermolabile protease with pH optimum at pH 8.0, that MPII is fairly resistant to high salt concentrations and that the MPII activity is not enhanced by glycerol, which is normally used for enzyme stabilization and activity enhancement. Because we were not able to locate any commercially available fluorescent quenched peptide substrates for FRA3, we could not perform the similar enzymological studies with FRA3. Overall, we conclude that the cleavage preferences we determined for FRA3 roughly represent those of the entire FRA sub-family and that the cleavage preferences of MPII and FRAs are distinct.
MMPs/ADAMs; known also as Galardin and Ilomastat) (57-59), was capable of inhibiting MPII (EC₅₀=31 nM). However, GM6001 is approximately 100-fold more potent in inhibiting MMPs [e.g., metalloproteinase-2 (MMP-2) with the EC₅₀=0.4 nM] (Table 2). Additional potent, broad-spectrum hydroxamate inhibitors, including AG3340/Prinomastat \{3(S)-2,2-dimethyl-4-[4-pyridin-4-yloxy]-benzenesulfonyl\}-thimorpholine-3-carboxylic acid hydroxyamide\} and BB-94/Batimastat \{(2R,3S)-N⁴-hydroxy-N¹-[(1S)-2-(methylamino)-2-oxo-1-(phenylmethyl)ethyl]-2-(2-methylpropyl)-3-[(2-thienylthio)methyl]butanediamide\} were poorly efficient against MPII. In principle, medicinal chemistry structure optimization and/or screening of hydroxamate class of inhibitors could identify more potent FRA inhibitors. In turn, TIMP-1, TIMP-2 and TIMP-3 (51), which are potent, sub-nanomolar MMP inhibitors are significantly less potent inhibitors of MPII (EC₅₀= >700 nM, 540 and 275 nM, respectively).

**Structural evidence for specificity of MPII and FRA3.** To elucidate structural elements that guide the cleavage preferences of MPII and FRA3, especially at the P2-P1' positions, we modeled the PRPLR↓AWGAA substrate bound to FRA3 using the recently reported structure of the FRA3 proenzyme (PDB 3P24) as a template (30). The sequence of the peptide substrate (PRPLR↓AWGAA) was based on the FRA3 IceLogo plot (Fig. 4B). The structure of MPII was prepared by homology modeling using the PDB 3P24 as a template. The sequence of the peptide substrate (PGRLR↓RSGAA) was based on the MPIII IceLogo plot (Fig. 4B). The fold of the PGRLR↓RSGAA substrate in MPII mimics the fold of the PRPLR↓AWGAA substrate in the FRA3 structure (Fig. 6A).

It becomes evident from this modeling that a large size S1' pocket can accommodate both small and large, hydrophobic and hydrophilic, and positive charged residues (but not the negatively charged Asp/Glu residues). This structural feature of the catalytic groove could explain a broad-range specificity of FRA3 at the P1' position. The S1' pocket of FRA3 is organized by the side chains of Ser319, Tyr347, Tyr374 and Ser374 and by the backbone of Ile321 and Leu368. Arg of the peptide substrate (and, probably, Lys as well) matches a highly negatively charged S1’ pocket of MPII. The presence of Glu276 in S1 of MPII (versus Ala276 in corresponding position of FRA3) explains the preference of MPII (but not FRA3) for the P1 Arg. This preference, however, does exclude hydrolysis of the substrates with P1’ residues distinct from Arg, such as Gly and Ala, both of which are accepted, albeit less efficiently, by MPII (Table S1). Self-activation of MPII at the LSSR↓A site provides an additional support for the above suggestion. The extended S2 site readily accommodates the hydrophobic Leu substrate residue in both MPII and FRA3. The 2.8 A resolution X-ray crystal structure of MPII we recently solved correlates very closely with our modeled MPIII structure, thus corroborating our modeling studies (manuscript in preparation).

Because of the high sequence homology among FRA1, FRA2 and FRA3, the modeled structures of FRA1 and FRA2 are highly similar to that of FRA3 (PDB 3P24). There are a very few substitutions in FRA1 and, especially in FRA2, which may affect their substrate binding mode and cleavage preferences as compared with those of FRA3 (Fig. 6B). Thus, the presence of Asp320 and Arg357 in FRA2 (versus Asn320 and Asn357 in FRA3) may have an insignificant effect on the P2 and P4 sub-site specificity. In a similar way, the presence of Lys312, Met316, Glu357, Lys277 and Phe319 in FRA1 (versus Asn312, Ile316, Asn357, Asp277 and Leu319 in FRA3) may have a limited impact on the promiscuous P1-P4 and P2’ sub-sites that accept multiple residue type in FRA3 (Fig. 4).

**Discussion**

The determination of the substrate recognition specificity of a protease is a necessary first step in developing drug-like antagonists. Knowledge of an optimal peptide substrate greatly facilitates many of the subsequent steps in drug development, and in some instances, can provide a lead that is usually a few steps from a drug.
Enterotoxigenic *B. fragilis* is the most frequent disease-causing anaerobe (13,21,23). The production of the secretory metalloproteinases is essential for virulence of *B. fragilis*. There are two distinct secretory metalloproteinase types (MPII and FRA) encoded by the pathogenicity island in enterotoxigenic *B. fragilis* strains (21,60-62). FRA exists in the three highly homologous isoforms (FRA1, FRA2 and FRA3) which exhibit over 90% sequence identity. In turn, there is only a low, 25%, identity of the peptide sequence of MPII with that of FRAs (21,26,30). While FRAs were the focus of multiple studies by others (13,14,19,21,30,32,63-69), the characteristics of MPII are completely unknown.

Here, we, for the first time, extensively characterized the purified recombinant MPII and FRA3, a representative isoform of the FRA subfamily. Based on the results of the high throughput multiplexed cleavage assay, we recently developed (36-38), followed by the digest of the individual selected peptides combined with the mass-spectrometry analysis of the cleavage reactions, we revealed that that MPII strongly prefers Arg residues at both the P1 and P1’ positions. The presence of Leu at the P2 position, Arg or Gly at both the P3 and P3’ positions, and Gly at the P4 position is also characteristic for the efficient substrates of MPII. From these perspectives, MPII mimics the dibasic Arg↓Arg cleavage motif of the furin-like proprotein convertases. We suspect that MPII interferes with the processing of those membrane and soluble precursor proteins, which are incompletely processed by proprotein convertases. As a result, MPII proteolysis may result in an abnormal precursor-mature protein ratio for the multiple targets of proprotein convertase processing, leading to pathological consequences (e.g., via aberrations of the furin-regulated functions, including cell-to-cell signaling, cell movement, Rho and Notch signaling and many others) (37). Intriguingly, MPII appears to be the first zinc metalloproteinase with dibasic cleavage preferences, suggesting a high level of versatility of metalloproteinase proteolysis. Unfortunately, the readily available pyroglutamic acid-Arg-Thr-Lys-Arg-methylcoumaryl-7-amide and other similar furin substrates cannot be used with MPII because the Arg-methyl-coumaryl-7-amide cleavage product is as quenched as the intact substrate.

In turn, FRA3 prefers the presence of Leu and Pro at the P2 and P5 positions of the substrate peptide, respectively. Arg/Ala, Pro/Gly/Ser and Arg/Ala/Leu are well tolerated at the P4, P3 and P1, respectively. Overall, the cleavage signature of FRA3 is related to the Pro-X-X-Leu consensus cleavage motif of MMPs. Based on the recently published crystal structure of FRA3 (PDB 3P24) and the follow-on our in silico modeling of the MPII structure, we provided structural evidence for specificity of MPII and FRA3. Our efforts identified the commercially-available fluorescent quenched peptide 5-FAM-SLGRK↓IQIQK(QXL520)-NH₂ as a substrate for the routine cleavage studies of MPII. This substrate is efficiently cleaved by MPII but not by FRA3, for which a commercially-available peptide substrate is currently unavailable. Our results, however, should allow us to develop FRA3 substrates to facilitate the laboratory studies of the toxin isoforms.

We, however, appreciate that many proteases harbor exosites that can be important elements for substrate selectivity and specificity. Likewise, structural properties of the substrate can be relevant for proper exposure of cleavage sites. Naturally, the use of short peptides in protease screenings is hampered by the lack of this structural information. Nevertheless, our results represent the first step in the right direction and cell-based assays in the search for natural substrates of MPII in colon epithelial cells will follow shortly.

Our data suggest that because of their inefficient binding capacity, TIMPs, natural inhibitors of MMPs, cannot neutralize the FRA activity in the course of *B. fragilis* infection (Table 2). More efficient inhibitors might limit FRA activity in a clinically beneficial manner. Therefore, peptide substrates that mimic the cleavage preference of the FRA isoforms are urgently required for inhibitor design and the follow-on inhibitory studies.

In sum, our results may help to determine, at the molecular level, generalized mechanisms by which the microbiota contributes to inflammation and increases cancer risk in infected patients (70).
Our results may be useful in the development of tools to monitor the presence of bacterial toxigenic proteinases and the bacterium itself in infected patients. Such tools may help to determine where, how and when B. fragilis fragilysins contribute to infection, to validate the value of novel inhibitors of these proteinases, and, lastly, to control these proteinases in a clinically beneficial manner for the infected patients.

**Abbreviations.** CRC, colorectal cancer; E352A and E349A, the catalytically inactive mutants of MPII and FRA3, respectively; FRA, fragilysin; MMP, matrix metalloproteinase; MPII, metalloproteinase II; PA, anthrax protective antigen; TIMP, tissue inhibitor of MMP.

**Acknowledgements.** We would like to thank all employees of Prognosys Biosciences for their help and support. In particular, we are thankful to T. Kitaoka and Erina He for assistance with DNA sequencing, Mark Chee for useful discussions and reading the manuscript, Wayne Delport for help with sequencing data analysis, and Brittan Van Beuge for administrative support.

**References**

1. Cho, I., and Blaser, M. J. (2012) The human microbiome: at the interface of health and disease. *Nat Rev Genet* 13, 260-270
2. Lee, S. M., Donaldson, G. P., Mikulski, Z., Boyajian, S., Ley, K., and Mazmanian, S. K. (2013) Bacterial colonization factors control specificity and stability of the gut microbiota. *Nature* 501, 426-429
3. Karin, M., and Greten, F. R. (2005) NF-kappaB: linking inflammation and immunity to cancer development and progression. *Nat Rev Immunol* 5, 749-759
4. Wu, S., Rhee, K. J., Albesiano, E., Rabizadeh, S., Wu, X., Yen, H. R., Huso, D. L., Brancati, F. L., Wick, E., McAllister, F., Housseau, F., Pardoll, D. M., and Sears, C. L. (2009) A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat Med* 15, 1016-1022
5. Chan, A. T., Arber, N., Burn, J., Chia, W. K., Elwood, P., Hull, M. A., Logan, R. F., Rothwell, P. M., Schror, K., and Baron, J. A. (2012) Aspirin in the chemoprevention of colorectal neoplasia: an overview. *Cancer Prev Res (Phila)* 5, 164-178
6. Rothwell, P. M., Fowkes, F. G., Belch, J. F., Ogawa, H., Warlow, C. P., and Meade, T. W. (2011) Effect of daily aspirin on long-term risk of death due to cancer: analysis of individual patient data from randomised trials. *Lancet* 377, 31-41
7. Arzumanyan, A., Reis, H. M., and Feitelson, M. A. (2013) Pathogenic mechanisms in HBV- and HCV-associated hepatocellular carcinoma. *Nat Rev Cancer* 13, 123-135
8. Salama, N. R., Hartung, M. L., and Muller, A. (2013) Life in the human stomach: persistence strategies of the bacterial pathogen Helicobacter pylori. *Nat Rev Microbiol* 11, 385-399
9. Noto, J. M., and Peek, R. M., Jr. (2012) Helicobacter pylori: an overview. *Methods Mol Biol* 921, 7-10
10. Polk, D. B., and Peek, R. M., Jr. (2010) Helicobacter pylori: gastric cancer and beyond. *Nat Rev Cancer* 10, 403-414
11. Trinchieri, G. (2011) Innate inflammation and cancer: Is it time for cancer prevention? *F1000 Med Rep* 3, 11
12. Saleh, M., and Trinchieri, G. (2011) Innate immune mechanisms of colitis and colitis-associated colorectal cancer. *Nat Rev Immunol* 11, 9-20
13. Holton, J. (2008) Enterotoxigenic Bacteroides fragilis. *Curr Infect Dis Rep* 10, 99-104
14. Riegler, M., Lotz, M., Sears, C., Pothoulakis, C., Castagliuolo, I., Wang, C. C., Sedivy, R., Sogukoglu, T., Cosentini, E., Bischof, G., Feil, W., Teleky, B., Hamilton, G., LaMont, J. T., and Wenzl, E. (1999) Bacteroides fragilis toxin 2 damages human colonic mucosa in vitro. *Gut* 44, 504-510
15. Wells, C. L., van de Westerlo, E. M., Jechorek, R. P., Feltis, B. A., Wilkins, T. D., and Erlandsen, S. L. (1996) Bacteroides fragilis enterotoxin modulates epithelial permeability and bacterial internalization by HT-29 enterocytes. *Gastroenterology* 110, 1429-1437
16. Wexler, H. M. (2007) Bacteroides: the good, the bad, and the nitty-gritty. *Clin Microbiol Rev* 20, 593-621
17. Wu, S., Lim, K. C., Huang, J., Saidi, R. F., and Sears, C. L. (1998) Bacteroides fragilis enterotoxin cleaves the zona adherens protein, E-cadherin. *Proc Natl Acad Sci U S A* 95, 14979-14984
18. Yoon, Y. M., Lee, J. Y., Yoo, D., Sim, Y. S., Kim, Y. J., Oh, Y. K., Kang, J. S., Kim, S., Kim, J. S., and Kim, J. M. (2010) Bacteroides fragilis enterotoxin induces human beta-defensin-2 expression in intestinal epithelial cells via a mitogen-activated protein kinase/I kappaB kinase/NF-kappaB-dependent pathway. *Infect Immun* 78, 2024-2033
19. Housseau, F., and Sears, C. L. (2010) Enterotoxigenic Bacteroides fragilis (ETBF)-mediated colitis in Min (Apc+/-) mice: a human commensal-based murine model of colon carcinogenesis. *Cell Cycle* 9, 3-5
20. Rhee, K. J., Wu, S., Wu, X., Huso, D. L., Karim, B., Franco, A. A., Rabizadeh, S., Golub, J. E., Mathews, L. E., Shin, J., Sartor, R. B., Golenbock, D., Hamad, A. R., Gan, C. M., Housseau, F., and Sears, C. L. (2009) Induction of persistent colitis by a human commensal, enterotoxigenic Bacteroides fragilis, in wild-type C57BL/6 mice. *Infect Immun* 77, 1708-1718
21. Sears, C. L. (2009) Enterotoxigenic Bacteroides fragilis: a rogue among symbiotes. *Clin Microbiol Rev* 22, 349-369
22. Toprak, N. U., Yagei, A., Gulluoglu, B. M., Akin, M. L., Demirkalem, P., Celenk, T., and Soyletir, G. (2006) A possible role of Bacteroides fragilis enterotoxin in the etiology of colorectal cancer. *Clin Microbiol Infect* 12, 782-786
23. Wick, E. C., and Sears, C. L. (2010) Bacteroides spp. and diarrhea. *Curr Opin Infect Dis* 23, 470-474
24. Dejea, C., Wick, E., and Sears, C. L. (2013) Bacterial oncogenesis in the colon. *Future Microbiol* 8, 445-460
25. Sears, C. L., Islam, S., Saha, A., Arjumand, M., Alam, N. H., Faruque, A. S., Salam, M. A., Shin, J., Hecht, D., Weintraub, A., Sack, R. B., and Qadri, F. (2008) Association of enterotoxigenic Bacteroides fragilis infection with inflammatory diarrhea. *Clin Infect Dis* 47, 797-803
26. Chung, G. T., Franco, A. A., Wu, S., Rhie, G. E., Cheng, R., Oh, H. B., and Sears, C. L. (1999) Identification of a third metalloprotease toxin gene in extraintestinal isolates of Bacteroides fragilis. *Infect Immun* 67, 4945-4949
27. Franco, A. A. (2004) The Bacteroides fragilis pathogenicity island is contained in a putative novel conjugative transposon. *J Bacteriol* 186, 6077-6092
28. Bode, W., Fernandez-Catalan, C., Tschesche, H., Grams, F., Nagase, H., and Maskos, K. (1999) Structural properties of matrix metalloproteinases. *Cell Mol Life Sci* 55, 639-652
29. Gomis-Ruth, F. X. (2003) Structural aspects of the metzincin clan of metalloendopeptidases. *Mol Biotechnol* 24, 157-202
30. Goulaas, T., Arolas, J. L., and Gomis-Ruth, F. X. (2011) Structure, function and latency regulation of a bacterial enterotoxin potentially derived from a mammalian adamalysin/ADAM xenolog. *Proc Natl Acad Sci U S A* 108, 1856-1861
31. Van Tassell, R. L., Lyerly, D. M., and Wilkins, T. D. (1992) Purification and characterization of an enterotoxin from Bacteroides fragilis. *Infect Immun* 60, 1343-1350
32. Wu, S., Rhee, K. J., Zhang, M., Franco, A., and Sears, C. L. (2007) Bacteroides fragilis toxin stimulates intestinal epithelial cell shedding and gamma-secretase-dependent E-cadherin cleavage. *J Cell Sci* 120, 1944-1952
33. Kling, J. J., Wright, R. L., Moncrief, J. S., and Wilkins, T. D. (1997) Cloning and characterization of the gene for the metalloprotease enterotoxin of Bacteroides fragilis. *FEMS Microbiol Lett* 146, 279-284
34. Moncrief, J. S., Obiso, R., Jr., Barroso, L. A., Kling, J. J., Wright, R. L., Van Tassell, R. L., Lyerly, D. M., and Wilkins, T. D. (1995) The enterotoxin of Bacteroides fragilis is a metalloprotease. *Infect Immun* 63, 175-181
35. Remacle, A. G., Shiryaev, S. A., Golubkov, V. S., Freskos, J. N., Brown, M. A., Karwa, A. S., Naik, A. D., Howard, C. P., Sympon, C. J., and Strongin, A. Y. (2013) Non-destructive and selective imaging of the functionally active, pro-invasive membrane type-1 matrix metalloproteinase (MT1-MMP) enzyme in cancer cells. J Biol Chem 288, 20568-20580

36. Kozlov, I. A., Thomsen, E. R., Munchel, S. E., Villegas, P., Capek, P., Gower, A. J., Pond, S. J., Chudin, E., and Chee, M. S. (2012) A highly scalable peptide-based assay system for proteomics. PLoS One 7, e37441

37. Shiryaev, S. A., Chernov, A. V., Golubkov, V. S., Thomsen, E. R., Chudin, E., Chee, M. S., Kozlov, I. A., Strongin, A. Y., and Cieplak, P. (2013) High-resolution analysis and functional mapping of cleavage sites and substrate proteins of furin in the human proteome. PLoS One 8, e54290

38. Shiryaev, S. A., Thomsen, E. R., Cieplak, P., Chudin, E., Cheltsov, A. V., Chee, M. S., Kozlov, I. A., and Strongin, A. Y. (2012) New details of HCV NS3/4A proteinase functionality revealed by a high-throughput cleavage assay. PLoS One 7, e35759

39. Roberts, R. W., and Szostak, J. W. (1997) RNA-peptide fusions for the in vitro selection of peptides and proteins. Proc Natl Acad Sci U S A 94, 12297-12302

40. Kurz, M., Gu, K., Al-Gawari, A., and Lohse, P. A. (2001) cDNA - protein fusions: covalent protein - gene conjugates for the in vitro selection of peptides and proteins. ChemBioChem 2, 666-672
42. Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009) Improved visualization of protein consensus sequences by iceLogo. *Nat Methods* 6, 786-787

43. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26, 1668-1688

44. Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Goetz, A. W., Kolossvary, I., Wong, F., Paesani, F., Vanicek, J., Wolf, R. M., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Salomon-Ferrer, R., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2012) Amber 12. *University of California, San Francisco*

45. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65, 712-725

46. Onufriev, A., Bashford, D., and Case, D. A. (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* 55, 383-394

47. Jaroszewski, L., Li, Z., Cai, X. H., Weber, C., and Godzik, A. (2011) FFAS server: novel features and applications. *Nucleic Acids Res* 39, W38-44

48. Jaroszewski, L., Rychlewski, L., Li, Z., Li, W., and Godzik, A. (2005) FFAS03: a server for profile–profile sequence alignments. *Nucleic Acids Res* 33, W284-288

49. Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U., and Sali, A. (2006) Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* Chapter 5, Unit 5 6

50. Seidah, N. G., and Prat, A. (2012) The biology and therapeutic targeting of the proprotein convertases. *Nat Rev Drug Discov* 11, 367-383

51. Brew, K., and Nagase, H. (2010) The tissue inhibitors of metalloproteinases (TIMPs): an ancient family with structural and functional diversity. *Biochim Biophys Acta* 1803, 55-71

52. Egeblad, M., and Werb, Z. (2002) New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer* 2, 161-174

53. Devel, L., Czarny, B., Beau, F., Georgiadis, D., Stura, E., and Dive, V. (2010) Third generation of matrix metalloprotease inhibitors: Gain in selectivity by targeting the depth of the S1' cavity. *Biochimie* 92, 1501-1508

54. Yiotakis, A., and Dive, V. (2008) Synthetic active site-directed inhibitors of metzincins: achievement and perspectives. *Mol Aspects Med* 29, 329-338

55. Hidalgo, M., and Eckhardt, S. G. (2001) Development of matrix metalloproteinase inhibitors in cancer therapy. *J Natl Cancer Inst* 93, 178-193

56. Woessner, J. F., Jr. (1999) Matrix metalloproteinase inhibition. From the Jurassic to the third millennium. *Ann NY Acad Sci* 878, 388-403

57. Levy, D. E., Lapierre, F., Liang, W., Ye, W., Lange, C. W., Li, X., Grobelny, D., Casabonne, M., Tyrrell, D., Holme, K., Nadzan, A., and Galardy, R. E. (1998) Matrix metalloproteinase inhibitors: a structure-activity study. *J Med Chem* 41, 199-223

58. Galardy, R. E., Cassabonne, M. E., Giese, C., Gilbert, J. H., Lapierre, F., Lopez, H., Schaefer, M. E., Stack, R., Sullivan, M., Summers, B., and et al. (1994) Low molecular weight inhibitors in corneal ulceration. *Ann NY Acad Sci* 732, 315-323

59. Grobelny, D., Ponec, L., and Galardy, R. E. (1992) Inhibition of human skin fibroblast collagenase, thermolysin, and Pseudomonas aeruginosa elastase by peptide hydroxamic acids. *Biochemistry* 31, 7152-7154

60. Claros, M. C., Claros, Z. C., Hecht, D. W., Citron, D. M., Goldstein, E. J., Silva, J., Jr., Tanglefield, Y., and Rodloff, A. C. (2006) Characterization of the Bacteroides fragilis pathogenicity island in human blood culture isolates. *Anaerobe* 12, 17-22
61. Franco, A. A., Cheng, R. K., Chung, G. T., Wu, S., Oh, H. B., and Sears, C. L. (1999) Molecular evolution of the pathogenicity island of enterotoxigenic Bacteroides fragilis strains. *J Bacteriol* 181, 6623-6633

62. Sanchez, E., Laparra, J. M., and Sanz, Y. (2012) Discerning the role of Bacteroides fragilis in celiac disease pathogenesis. *Appl Environ Microbiol* 78, 6507-6515

63. Foulon, I., Pierard, D., Muyldermans, G., Vandoorslaer, K., Soetens, O., Rosseel, P., and Lauwers, S. (2003) Prevalence of fragilysin gene in Bacteroides fragilis isolates from blood and other extraintestinal samples. *J Clin Microbiol* 41, 4428-4430

64. Obiso, R. J., Azghani, A. O., and Wilkins, T. D. (1997) The Bacteroides fragilis toxin fragilysin disrupts the paracellular barrier of epithelial cells. *Infect Immun* 65, 1431-1439

65. Sears, C. L., Buckwold, S. L., Shin, J. W., and Franco, A. A. (2006) The C-terminal region of Bacteroides fragilis toxin is essential to its biological activity. *Infect Immun* 74, 5595-5601

66. Vines, R. R., Perdue, S. S., Moncrief, J. S., Sentz, D. R., Barroso, L. A., Wright, R. L., and Wilkins, T. D. (2000) Fragilysin, the enterotoxin from Bacteroides fragilis, enhances the serum antibody response to antigen co-administered by the intranasal route. *Vaccine* 19, 655-660

67. Saidi, R. F., Jaeger, K., Montrose, M. H., Wu, S., and Sears, C. L. (1997) Bacteroides fragilis toxin rearranges the actin cytoskeleton of HT29/C1 cells without direct proteolysis of actin or decrease in F-actin content. *Cell Motil Cytoskeleton* 37, 159-165

68. Wu, S., Powell, J., Mathioudakis, N., Kane, S., Fernandez, E., and Sears, C. L. (2004) Bacteroides fragilis enterotoxin induces intestinal epithelial cell secretion of interleukin-8 through mitogen-activated protein kinases and a tyrosine kinase-regulated nuclear factor-kappaB pathway. *Infect Immun* 72, 5832-5839

69. Wu, S., Shin, J., Zhang, G., Cohen, M., Franco, A., and Sears, C. L. (2006) The Bacteroides fragilis toxin binds to a specific intestinal epithelial cell receptor. *Infect Immun* 74, 5382-5390

70. Chow, J., Tang, H., and Mazmanian, S. K. (2011) Pathobionts of the gastrointestinal microbiota and inflammatory disease. *Curr Opin Immunol* 23, 473-480

71. Kashwagi, M., Tortorella, M., Nagase, H., and Brew, K. (2001) TIMP-3 is a potent inhibitor of aggrecanase 1 (ADAM-TS4) and aggrecanase 2 (ADAM-TS5). *J Biol Chem* 276, 12501-12504

72. Rasmussen, H. S., and McCann, P. P. (1997) Matrix metalloproteinase inhibition as a novel anticancer strategy: a review with special focus on batimastat and marimastat. *Pharmacol Ther* 75, 69-75

73. Shalinsky, D. R., Brekken, J., Zou, H., McDermott, C. D., Forsyth, P., Edwards, D., Margosiak, S., Bender, S., Truitt, G., Wood, A., Varki, N. M., and Appelt, K. (1999) Broad antitumor and antiangiogenic activities of AG3340, a potent and selective MMP inhibitor undergoing advanced oncology clinical trials. *Ann N Y Acad Sci* 878, 236-270
**Tables**

**Table 1. Peptide sequences efficiently cleaved by MPII and FRA3.** High numbers indicate a high cleavage efficiency of the peptides in the multiplexed cleavage assay. Mass spectrometry (MS) of the digest reactions was used to identify the mass of the cleavage fragments and, consequently, the scissile bonds (arrow). Enrichment Ratio (ER) is a measure of cleavage activity.

| Peptide sequences with the scissile bonds | Peptides identified by MS in the cleavage reactions | Fragment mol. wt. (Da) | Cleavage assay data, ER |
|-----------------------------------------|---------------------------------------------------|------------------------|-------------------------|
|                                         |                                                   |                        | MPII | FRA3    |
| GHSRRSRRSG                             | GHSRRSRRSG                                       | 1155                   | 81.7 | 3.4     |
|                                         | GHSRRS                                           | 699                    |      |         |
|                                         | GHSRR                                             | 612                    |      |         |
|                                         | RSRR                                               | 574                    |      |         |
|                                         | RRSRG                                             | 474                    |      |         |
| AGLRRARALGG                            | AGLRRARALGG                                       | 941                    | 69.5 | 78.1    |
|                                         | AGLR                                               | 415                    |      |         |
| AGLRRASLGG                             | AGLRRASLGG                                       | 957                    | 61.1 | 49.7    |
|                                         | RRASLGG                                           | 715                    |      |         |
|                                         | RASLGG                                             | 559                    |      |         |
|                                         | AGLR                                               | 416                    |      |         |
| LRRKRLRRGSYS                          | LRRKRLRRGSYS                                      | 1265                   | 49.1 | 40.9    |
|                                         | RLSYS                                               | 711                    |      |         |
|                                         | LRRK                                              | 572                    |      |         |
| GRHRQIDRG                              | GRHRQIDRG                                         | 1250                   | 49.0 | 8.4     |
|                                         | RRQIDRG                                            | 900                    |      |         |
|                                         | RQIDRG                                             | 744                    |      |         |
| GNKRGGGTAG                             | GNKRGGGTAG                                        | 973                    | 18.4 | 1.9     |
|                                         | RRGGGTAG                                           | 673                    |      |         |
|                                         | GNKRR                                               | 631                    |      |         |
|                                         | RGGTAG                                              | 517                    |      |         |
| SGPVSMRYTA                             | SGPVSMRYTA                                        | 1068                   | 3.3  | 130.9   |
|                                         | SGPVSMR                                           | 733                    |      |         |
| SGHMHAALTSA                           | SGHMHAALTSA                                       | 995                    | 14.8 | 141.1   |
|                                         | SGHMHAALTSA                                       | 995                    | 14.8 | 141.1   |
| SGPVSMRYTA                             | SGPVSMRYTA                                        | 1068                   | 3.3  | 130.9   |
|                                         | SGPVSMR                                           | 733                    |      |         |
| SGPRSLKSTA                             | SGPRSLKSTA                                        | 1003                   | 22.4 | 110.5   |
|                                         | SGPRSLK                                           | 744                    |      |         |
| SGPMISRMTA                             | SGPMISRMTA                                        | 1050                   | 16.0 | 110.4   |
|                                         | SGPMISR                                           | 747                    |      |         |
| AGLRRAALGG                             | AGLRRAALGG                                        | 941                    | 69.5 | 78.1    |
|                                         | AGLRRAALGG                                        | 941                    |      |         |
|                                         | AGLRRRAA                                          | 714                    |      |         |
| GHSRRSRRSRG                           | GHSRRSRRSRG                                       | No cleavage            | 81.7 | 3.4     |
|                                         | GHSRRSRRSRG                                       | No cleavage            |      |         |
| PTKIYYNIDY                             | PTKIYYNIDY                                        | 1241                   | n/a  | n/a     |
|                                         | PTKIY                                              | 621                    |      |         |
Table 2. MPII inhibitors. The EC$_{50}$ values of hydroxamate inhibitors and TIMPs against MPII were measured using 5-FAM-SLGRKIQIK(QXL520)-NH$_2$ as a fluorescent-quenched peptide substrate. *The data indicate the k$_{on}$ constant for the individual inhibitory N-terminal domain of TIMP-3 and the k$_i$ values for both AG3340 and BB-94/Batimastat (71-73).

| Inhibitor | EC$_{50}$, nM |
|-----------|---------------|
|           | MPII MMP-2    |
| TIMP-1    | >700 <0.1     |
| TIMP-2    | 540 <0.05     |
| TIMP-3    | 275 ~3*       |
| GM6001    | 31 0.4        |
| AG3340    | 230 0.083*    |
| BB-94     | >1000 4*      |
Figure legends

Fig. 1. Sequence alignment of the fragilysin isoforms (FRA1, FRA2 and FRA3) and MPII encoded by *B. fragilis* pathogenicity island. Dots indicate identical residue positions. Asterisks indicate the conserved active site histidine residues.

Fig. 2. *B. fragilis* is predominantly associated with CRC tumors rather than normal tissue. The 501-bp fragment of the *B. fragilis* 16S RNA gene was PCR amplified in the CRC1-8 and XT1-22 tumor and matching normal tissue specimens. Numbers of *B. fragilis*-positive and negative samples are shown in the bottom table. Arrows show the samples in which the amplified DNA bands were sequenced to determine their nucleotide sequence and to confirm their identity. The data are statistically significant (P=0.035) as judged by two-sided Fisher’s exact test.

Fig. 3. MPII and FRA3 constructs. (A) Recombinant MPII and FRA3. Left, *B. fragilis* wild type MPII and FRA3. Arrows indicate the cleavages at the prodomain excision sites in the proform. These cleavages release the prodomain and the mature protease. SP, PRO and CAT, signal peptide, prodomain and catalytic domain, respectively. Right, recombinant MPII and FRA3 tagged with the 6xHis and Flag tags. E352A and E349A indicate the catalytically inactive mutants of MPII and FRA3, respectively, in which the catalytically essential Glu residue of the active site was mutated into Ala. (B) SDS-gel electrophoresis of the purified constructs. Note that the wild-type MPII is readily self-activated while the wild-type FRA3 is significantly more stable in its proenzyme form during storage. CS, Coomassie staining; WB, Western blotting with an anti-Flag tag antibody.

Fig. 4. Cleavage preferences of MPII and FRA3. (A) Representative MALDI-TOF MS spectra of the individual cleavage peptides. Top, GNKRRGGTAG was co-incubated with MPII. Middle and bottom, AGLRRAALGG and SGHMHAALTA, respectively, were co-incubated with FRA3. The spectra of the digest, the intact peptide, the enzyme alone and the buffer alone are shown by red, black, green and magenta lines, respectively. AU, arbitrary units. (B) Frequency plot of the cleavage sequences of MPII and FRA3 in an IceLogo format. The height of a character is proportional to the frequency of the amino acid residue at the individual position of the cleaved peptide and is normalized for the amino acid encoded in the entire human genome according to RefSeq. (C) MPII cleaves anthrax PA83. Left, MPII and furin were co-incubated with PA83 at the indicated enzyme-substrate molar ratio. The resulting PA63 and PA20 are shown by arrows. The digests were separated by SDS-gel electrophoresis. Where indicated, GM6001 was added to the reactions.

Fig. 5. Biochemical characteristics and inhibitors of MPII. The cleavage activity of MPII was measured using 5-FAM-SLGRKIQIQK(QXL520)-NH2 as a fluorescent-quenched peptide substrate. (A, B, C, and D) The effect of buffer pH, glycerol concentration, salt concentration and temperature on the MPII cleavage activity, respectively. (E) Selected inhibitors of MPII: representative dose-response curves. The EC50 values of TIMP-2 and the hydroxamate inhibitors (GM6001 and AG3340) were measured using 5-FAM-SLGRKIQIQK(QXL520)-NH2 as a fluorescent-quenched peptide substrate.

Fig. 6. Structural parameters of FRA1-3 and MPII. (A) Structural modeling of MPII and FRA3 with peptide substrates. Left, the structure of MPII with the PGRLR|RSGAA modeled substrate. Right, the structure of FRA3 (PDB 3P24) with the PRPLR|AWGAA modeled substrate. Substrates are shown as sticks. The molecular surface of MPII and FRA3 with the labeled substrate sub-sites is colored according to the electrostatic potential (red, blue and white are negative, positive and neutral electrostatic potential values, respectively). The active site zinc ion is shown as a grey sphere. (B) Modeled structures of FRA1 and FRA2 versus FRA3 (PDB 3P24). The modeled PRPLR|AWGAA substrate is shown as green sticks. The structures of the mature FRA1, FRA2 and FRA3 are shown as grey cartoons in the background and as transparent electrostatic potential surfaces (red, blue and white are negative, positive and neutral electrostatic potential values, respectively). Arrows point to the residue positions, which may affect substrate binding and which are distinct in FRA1 and FRA2 versus FRA3.
Fig. 1
**Fig. 2**

| B. fragilis | + | − |
|-------------|---|---|
| TUMOR       | 22| 8 |
| NORMAL      | 13| 17|
Fig. 3
Fig. 4
Fig. 5
Substrate cleavage profiling suggests a distinct function of Bacteroides fragilis metalloproteinases (fragilysin and metalloproteinase II) at the microbiome-inflammation-cancer interface

Sergey A. Shiryaev, Albert G. Remacle, Andrei V. Chernov, Vladislav S. Golubkov, Khatereh Motamedchaboki, Norihito Muranaka, Corey M. Dambacher, Petr Capek, Muskan Kukreja, Igor A. Kozlov, Manuel Perucho, Piotr Cieplak and Alex Y. Strongin

J. Biol. Chem. published online October 21, 2013

Access the most updated version of this article at doi: 10.1074/jbc.M113.516153

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2013/10/21/M113.516153.DC1