SUPPORTING INFORMATION

SI Materials and Methods

Strains and cultivation conditions. *Escherichia coli* strain DH5α and GeneHogs® *E. coli* (Invitrogen) were used for cloning and amplification of plasmids; *E. coli* strain BL21(DE3) was used for bacterial expression of proteins and biotoxicity assays. *E. coli* was cultivated on LB medium as described (1). *Caenorhabditis elegans* strains were cultivated on nematode growth medium (NGM) seeded with *E. coli* strain OP50 as described (2). The Bristol isolate N2 was used as wild type strain. Strains *pmk-1(km25), bre-1(ye4), bre-2(ye31), bre-3(ye26), bre-4(ye27), bre-5(ye17), fut-1(ok892), fut-2(gk360), fut-2(ok509), fut-3(gk103), fut-4(gk111), fut-5(ok242), fut-6(ok475), fut-8(ok2558), gly-2(qa703), gly-12(is47), dpy-6(e14);gly-13(ok712), gly-14(id48), gly-20(ok826),* were obtained from the *Caenorhabditis* Genetics Center (CGC) at the University of Minnesota (USA). Strains resulting from *Mos1*-mediated mutagenesis (see below) were *pmk-1(km25);bre-3(op504), pmk-1(km25);bre-3(op505), pmk-1(km25);bre-3(op506)* and *pmk-1(km25);bre-3(op508).* Primers used for genotyping are listed in Supplementary Table S1.

Cloning and site-directed mutagenesis. The full-length coding sequence (Fig. S9) of wild type MOA was subcloned from plasmid MOApT7LO (3) into the pET22 vector (Novagen) using *Nde*I and *Eco*RI restriction sites. The coding sequence for MOA(C215A) was constructed by PCR using MOA_F, C215A_R, C215A_F and MOA_R as primers and MOApET22 as template. MOA(Q46A,W138A) was generated by PCR using MOA_F, I222M_R, I222M_F and MOA_R as primers and plasmid MOA(Q46A,W138A)pT7LOH (4) as template. MOA(ΔC) was constructed by PCR using MOA_F and ΔC_R as primers and MOApET22 as template. PCR products were ligated into pET22 using *Nde*I and *Eco*RI restriction sites. Primer sequences are listed in Table S1.

Protein expression and purification. MOA and MOA(C215A) were expressed in *E. coli* BL21(DE3) in LB medium by cultivation of the respective transformants to OD600nm ~ 1 at 37 °C, addition of 1 mM IPTG and further incubation for 5 h at 28 °C. Bacterial cell pellets were resuspended in ice-cold phosphate-buffered saline, pH 7.4 (PBS) and lysed using a French press. Cell debris were removed in two consecutive steps of centrifugation at 12 000 g for 15 min and at 27 000 g for 30 min. The supernatant was incubated with melibiose-sepharose at 4°C for 1 h and the protein was finally eluted at room temperature in PBS containing 200 mM melibiose. Purified MOA and MOA(C215A) were desalted on a PD-10 column (Amersham Biosciences) and concentrated using an Amicon Ultra-4 centrifugal filter device (Millipore) with a molecular weight cutoff of 10 kDa. Protein concentration was calculated by measuring the absorbance at 280 nm, assuming a calculated extinction coefficient of 87 480 M^{-1} cm^{-1}.

***In vitro* protease assay.** RNase A was denatured by heating at 100°C for 10 min with glycoprotein denaturing buffer (New England BioLabs). 0.1 ng to 1 µg of purified recombinant MOA or MOA(C215A) was incubated with 20 µg of denatured RNase A in assay buffer (40 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 1% NP-40) in a final reaction volume of 30 µl for 1 h at 37 °C. The reaction was stopped by addition of SDS sample buffer and heating at 95 °C for 10 min. Samples were analyzed by SDS-PAGE and Coomassie blue

staining. For inhibition assays, MOA was preincubated with 50 µM E-64, 2 mM PMSF or 2 mM $AgNO_3$ for 20 min at RT before the substrate was added. To test the effect of divalent cations the reaction was performed in assay buffer without $CaCl_2$ as well as in buffer containing 1 mM $CaCl_2$, $MgCl_2$, $MnCl_2$, and $ZnCl_2$ respectively. The influence of the pH was assayed by using the following buffers supplemented with 1mM $CaCl_2$ and 1% NP-4: 40 mM sodium acetate buffer pH 4.5, 50 mM sodium phosphate buffer pH 6.0, 50 mM sodium phosphate buffer pH 7.5, 40 mM Tris-HCl pH 8.0, 50 mM Tris-HCl pH 8.5, 40 mM Tris-HCl pH 9.0 and 10 mM CAPS pH 9.7.

Cleavage site specificity. Samples were analyzed on an LTQ-Orbitrap XL mass spectrometer (Thermo Fischer Scientific, Bremen, Germany) coupled to an Eksigent-Nano-HPLC system (Eksigent Technologies, Dublin (CA), USA). Solvent composition at the two channels was 0.2% formic acid, 1% acetonitrile for channel A and 0.2% formic acid, 80% acetonitrile for channel B. Peptides were loaded on a self-made tip column (75 µm × 80 mm) packed with reverse phase C18 material (AQ, 3 µm 200 Å, Bischoff GmbH, Leonberg, Germany) and eluted with a flow rate of 200 nl per min by a gradient from 0 to 10% of B in 5 min, 47% B in 55 min, 97% B in 58 min. Full-scan MS spectra (300−2000 *m*/*z*) were acquired with a resolution of 60000 at 400 *m*/*z* after accumulation to a target value of 500000. Collision induced dissociation (CID) MS/MS spectra were recorded in data dependent manner in the ion trap from the five most intense signals above a threshold of 500, using a normalized collision energy of 35% and an activation time of 30 ms. Charge state screening was enabled and singly charge states were rejected. Precursor masses already selected for MS/MS were excluded for further selection for 90 s and the exclusion window was set to 20 ppm. The size of the exclusion list was set to a maximum of 500 entries. Peptides were assigned to MS2 spectra using Mascot v2.3 and the UniProt Swiss-Prot release 15.13 database with the following parameters: parent tolerance: 5 ppm; fragment tolerance: 0.6 Da; enzyme specificity: none; missed cleavages: 0; variable Modifications: -17 on n (Pyro-cmC), +16 on M (Oxidation), +42 on n (Acetyl), +80 on S (Phospho). Scaffold 3.0 (Proteome Software, Inc.) was used for sample merging and secondary validation of Mascot results. Only peptides with an identification probability of >95% matching to one of the test substrate proteins were included in further analyses (Table S2). Consensus cleavage sites were derived as described in Schilling *et al.* (5) using a custom database only comprising the test substrates, and sequence logos were generated using iceLogo (http://iomics.ugent.be/icelogoserver/main.html) with Swiss-Prot composition as reference set.

Figure S1. MOA-mediated toxicity towards *Acanthamoeba castellanii*.

E. coli BL21(DE3) cells expressing MOA, the catalytic site mutant MOA(C215A), the C-terminal deletion mutant MOA(ΔC), the carbohydrate-binding defective mutant MOA(Q46AW138A) or containing empty vector (VC) were fed to *A. castellanii*. Growth of amoebae was evaluated as clearing zone of a lawn of respective *E. coli* BL21 cells after 6 days (6). Columns represent the average of 5 biological replicates. Error bars indicate standard deviations. Toxicity data were evaluated by one-way ANOVA and post-hoc comparisons were performed using LSD.

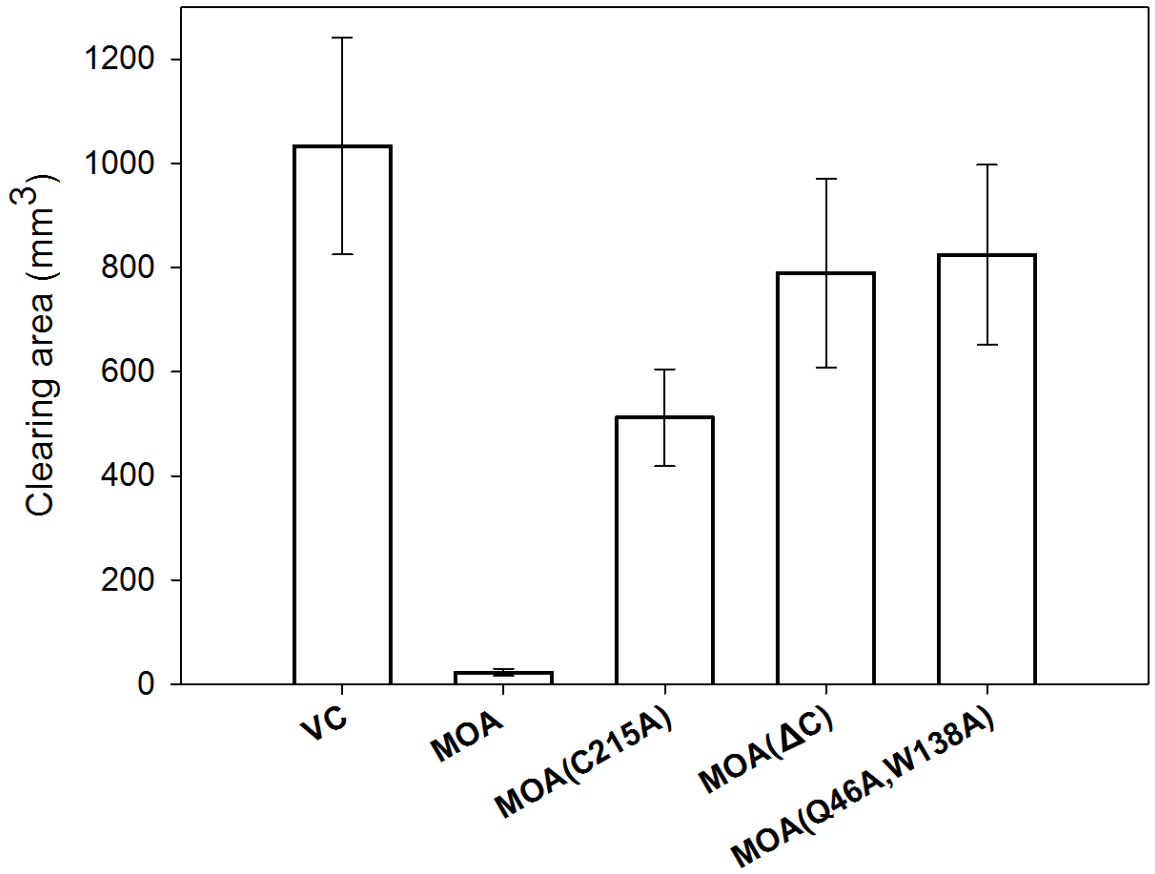
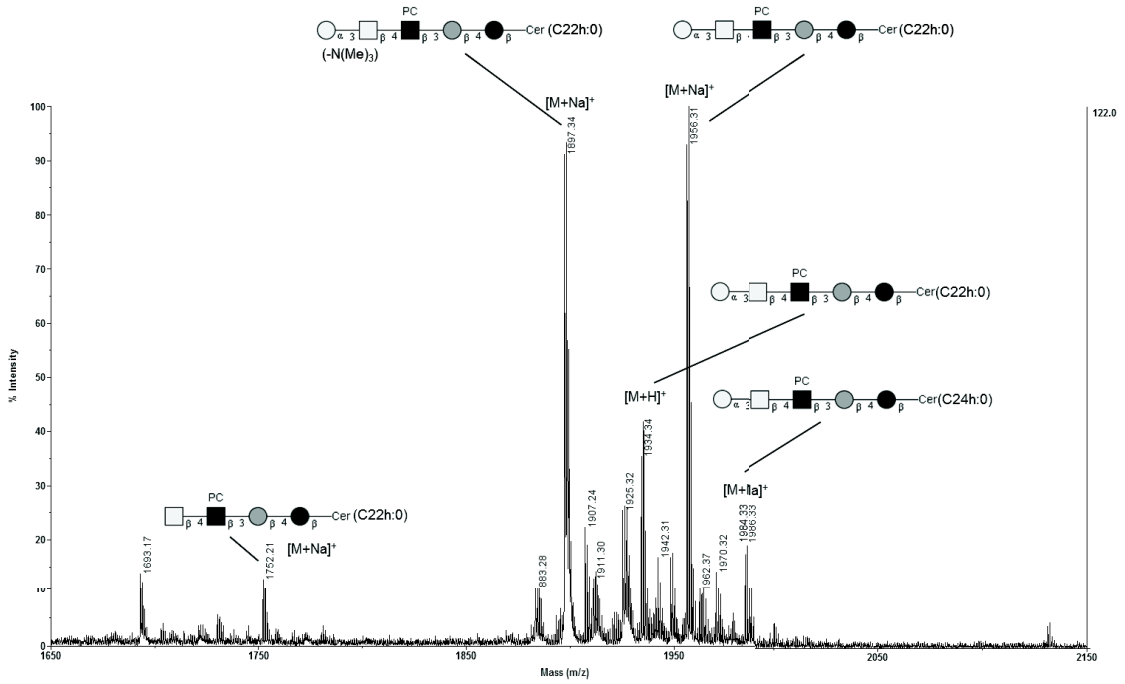


Figure S2. MS analysis of permethylated glycolipid extracts from MOA-positive TLC-band.

Purified upper phase glycolipids from *C. elegans* N2 were separated by TLC as described in Materials and Methods. The TLC plate was cut in two pieces, one of which was overlaid with biotinylated MOA as described. Glycolipids at the position corresponding to band D in the overlay (Fig. 2*C*) were extracted from the non-treated part of the TLC plate using MeOH and sonication. Permethylation of glycolipid extracts was performed as described (7). Derivatized glycolipids were dissolved in 20 µl of methanol, and 1 µl of dissolved sample was premixed with 1 µl of matrix (1% 6-aza-2-thiothymine/1 mM ammonium citrate/80% acetonitrile) before loading onto a metal plate. MALDI data were acquired using a 4800 Plus MALDI-TOF/TOF Analyzer (AB Sciex, Foster City, CA). For the calculation of the molecular weight a permethylation of all amide and of all hydroxyl groups, apart from the hydroxyl group of the phosphorylcholine (PC), was assumed. **A.** MALDI-MS spectrum acquired in the reflector mode. The peak at *m/z* 1934.34 corresponds to the protonated molecular ion of permethylated component D including the ceramide. The ceramide moiety composed of a saturated hydroxy fatty acid with 22 carbon atoms (C22h:0) has previously been described (8,9). Elimination of trimethylamine ($N(Me)_3$) from this structure leads to the signal at *m/z* 1897.34. The peaks at *m/z* 1956.31 and 1984.33 can be assigned to the $[M+Na]^+$ ions of component D including a ceramide moiety composed of (C22h:0) and (C24h:0) fatty acids, respectively. Besides these structures signals corresponding to glycosphingolipids with different fatty acid chain length are present (*m/z* 1942.31 and 1970.32). **B.** MALDI-MS/MS spectrum of the $[M+H]^+$ ion at *m/z* 1934.34. The fragmentation pattern of *m/z* 1934.34 strongly supports the identity of component D.

A


PC
Cer (C22h:0)
(-N(Me)3)
[M+Na]+
1897.34
1956.31
[M+H]+
1934.34
Cer(C24h:0)
[M+Ia]+
1752.21
1693.17
883.28
1907.24
1911.30
1925.32
1942.31
1962.37
1970.32
1984.33
1986.33
% Intensity
Mass (m/z)

B

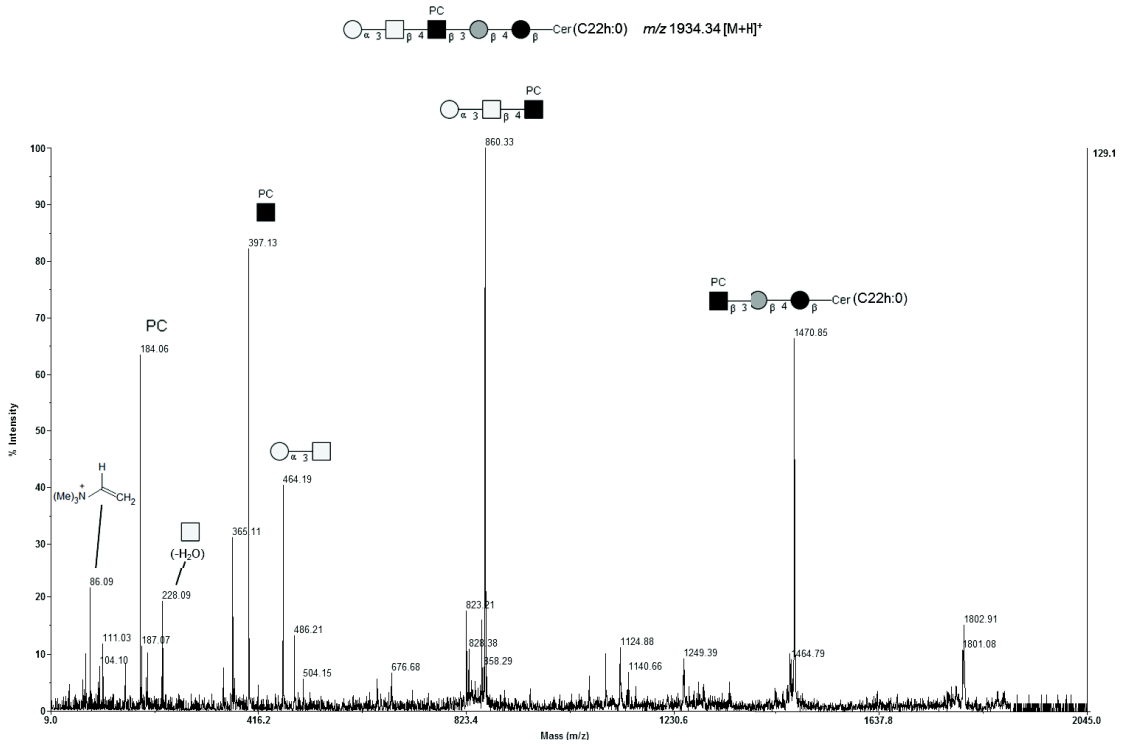

PC
Cer (C22h:0)
m/z 1934.34 [M+H]+
860.33
397.13
184.06
1470.85
464.19
365.11
(-H2O)
228.09
86.09
111.03
104.10
187.07
486.21
504.15
676.68
823.21
828.38
858.29
1124.88
1140.66
1249.39
1464.79
1802.91
1801.08
% Intensity
Mass (m/z)

Figure S3. Influence of Ca^{2+} concentration on proteolytic activity of MOA.

20 µg of denatured RNase A were digested with 10 ng of purified recombinant MOA in assay buffer (40 mM Tris-HCl pH 8.0, 1% NP-40) at $CaCl_2$ concentrations of 10 µM to 5 mM or 10 mM EDTA. The assay was incubated at 37 °C for 1 h followed by addition of SDS sample buffer and heating at 95 °C for 10 min. Samples were analyzed by SDS-PAGE and Coomassie blue staining. m, molecular weight standard.

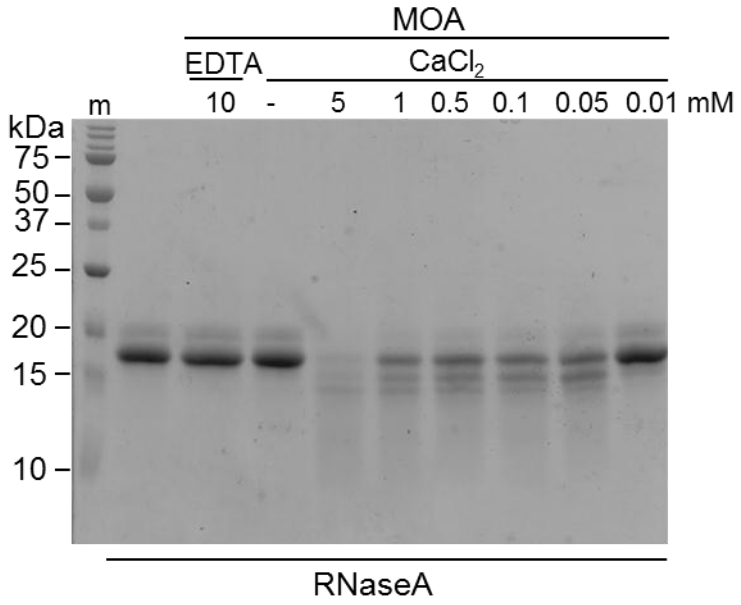


Figure S4. Influence of carbohydrate ligand on proteolytic activity of MOA.

0.5 µg of purified recombinant MOA was incubated with 0.16 nmol of Galα1,3Galβ1,4GlcNAc in assay buffer (40 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 1% NP-40) for 30 min at RT before the addition 20 µg of denatured RNase A. After 1 h incubation at 37 °C the reaction was stopped by addition of SDS sample buffer and heating at 95 °C for 10 min. Samples were analyzed by SDS-PAGE and Coomassie blue staining. m, molecular weight standard.

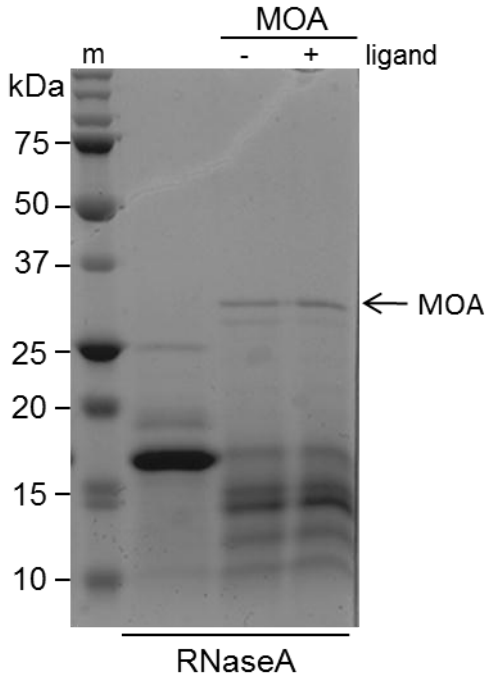


Figure S5. MOA protease activity on denatured and native protein substrates.

1 µg of purified recombinant MOA was added to 20 µg of protein substrate in assay buffer (40 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 1% NP-40) and incubated for 1 h at 37 °C. Samples were analyzed as described above. m, molecular weight standard. **A.** Protein substrates were denatured by heating at 100 °C for 10 minutes in denaturing buffer (0.5% SDS, 40 mM DTT) before the start of the assay. **B.** Native protein was used as substrate in the protease assay.

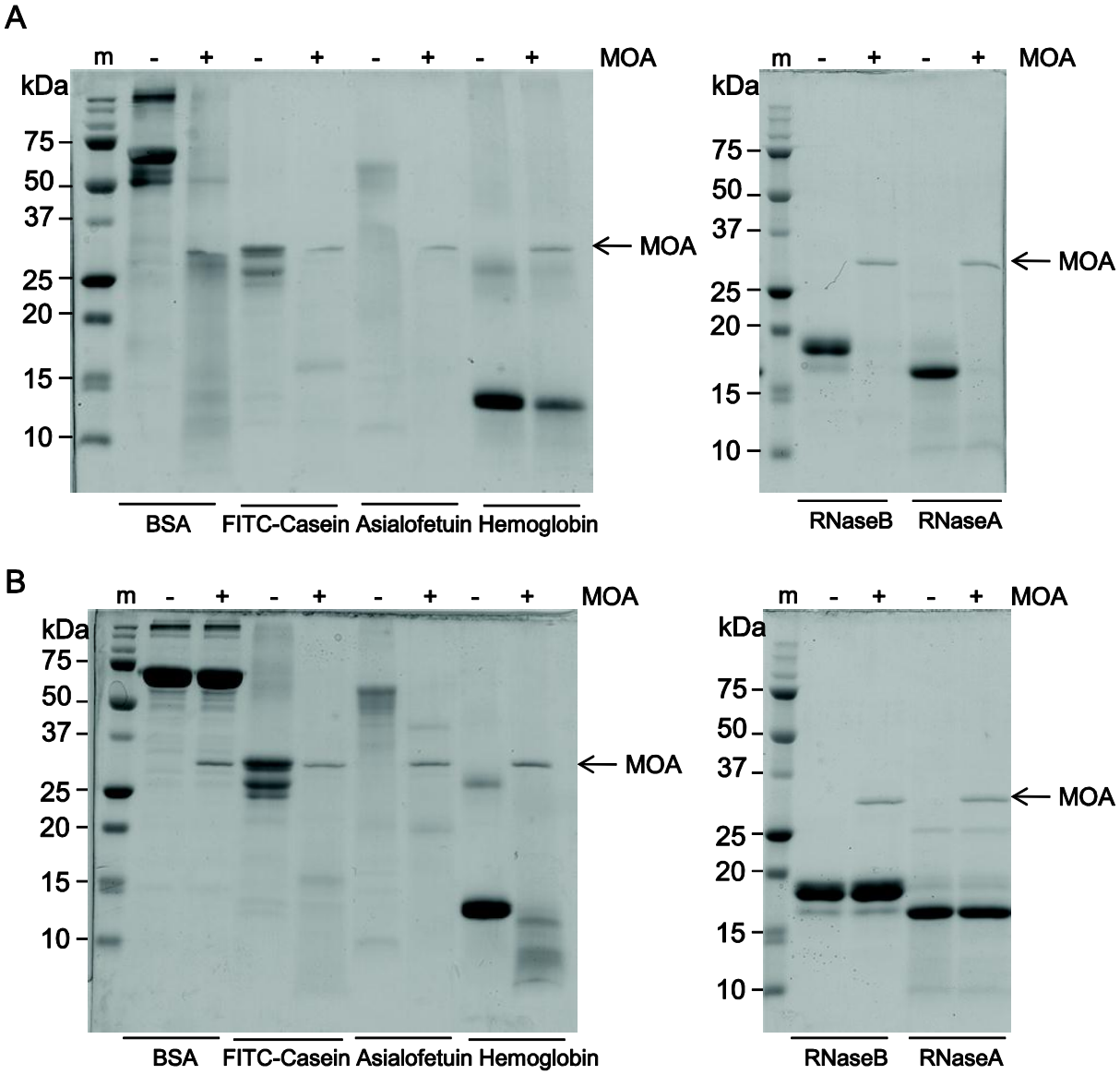


Figure S6. Activity of MOA in FITC-casein protease assay.

100 ng (3 pmol) of MOA or MOA(C215A) were added to 40 µg of fluorescein isothiocyanate (FITC) labeled casein substrate in assay buffer (40 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 1% NP-40) and incubated for 1h at 37 °C. The reaction was stopped by the addition of 0.6 M TCA and samples were measured using a fluorimeter at an excitation wavelength of 485 nm and an emission wavelength of 535 nm. A reaction with 100 ng papain (4 pmol) in 200 mM sodium phosphate buffer pH 7.5 was included for comparison. The experiment was performed in duplicates. Error bars indicate standard deviations.

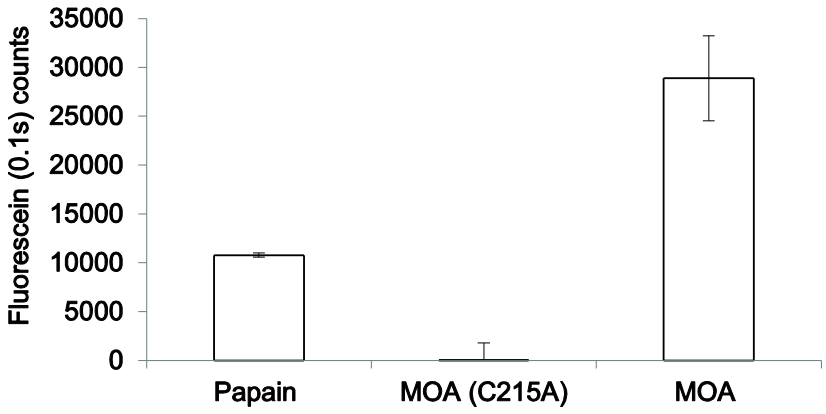


Figure S7. Oligomerization of MOA and MOA(C215A).

MOA(C215A) (solid line) and MOA (dashed line) both elute at 8.8 ml from a Superdex 75 column in FPLC. Based on runs with standard proteins (dotted lines, numbers indicate size in kDa), this elution volume corresponds to an apparent molecular weight of 48.2 kDa. BSA (67 kDa), ovalbumin (45 kDa) and chymotrypsinogen (25 kDa) were used as standard proteins. The void volume (6.8 ml) was determined by elution of Blue Dextran. Gel filtration was performed at a flow rate of 0.5 ml/min in PBS containing 0.2 mM $CaCl_2$ and 0.1 M lactose. Samples of 0.1-0.7 mg of protein in 0.1 ml were injected, and the eluate was monitored at 280 nm.

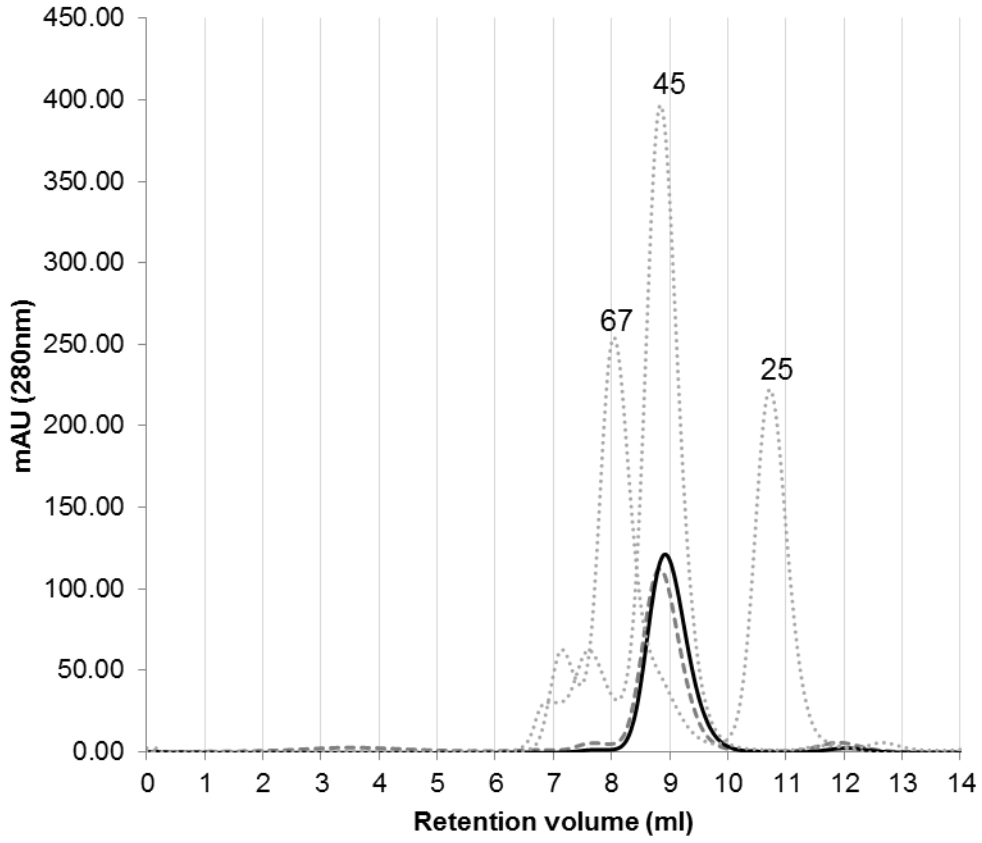


Figure S8. Hypothetical toxicity mechanism of MOA.

Upon ingestion by the nematode, the lectin domain of MOA (blue square) binds to terminal Galα1,3 of glycosphingolipids on the surface of intestinal epithelial cells triggering internalization to the early endosomal (EE) compartment. MOA is then trafficking *via* the Golgi to the endoplasmic reticulum (ER), where the protease domain (green sphere) is activated at a Ca^{2+} concentration of >100 µM. MOA-mediated toxicity is finally induced by degradation of proteins in the ER.

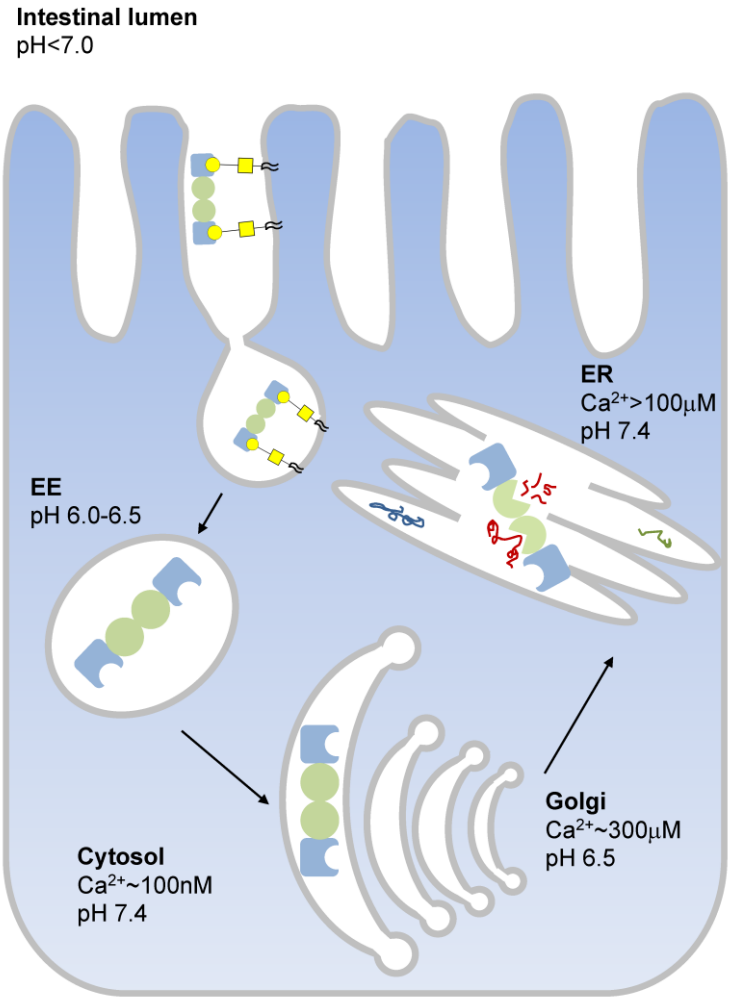


Figure S9. Nucleotide (A) and amino acid sequence (B) of MOA used in this study.

The coding sequence of MOA used in this study was derived from MOApT7LO (3) and contains two point mutations, Q196H and D200N, and several silent mutations compared to Genbank sequence AY066013.

A

ATGTCTCTGCGACGCGGAATTTACCACATCGAGAATGCTGGGGTTCCCAGTGCCA
TTGATCTCAAAGACGGCAGCTCCAGTGACGGCACACCTATCGTTGGCTGGCAGTT
TACACCAGACACGATCAACTGGCATCAGCTCTGGCTTGCTGAACCAATCCCCAAC
GTTGCTGATACCTTTACCCTTTGCAACCTGTTCAGCGGTACCTACATGGATCTCTA
CAACGGTTCTTCCGAAGCGGGCACCGCAGTCAATGGTTGGCAAGGAACTGCCTTT
ACGACCAATCCCCACCAGCTCTGGACCATCAAGAAGTCGAGCGACGGTACGAGC
TACAAGATCCAGAATTATGGAAGTAAAACCTTCGTCGATCTTGTCAATGGCGACA
GCTCTGATGGGGCCAAAATTGCTGGATGGACCGGCACTTGGGATGAAGGTAACC
CTCACCAGAAATGGTACTTCAATAGGATGAGCGTCTCCAGCGCGGAGGCCCAAG
CGGCTATCGCGCGAAACCCTCATATTCACGGGACTTACAGAGGATACATCCTCGA
TGGAGAATATCTTGTCCTCCCTAACGCTACTTTCACTCACATTTGGAAAAACTCC
GGTCTTCCTGGTAGCAAATGGCGTGAGCAAATCTATGATTGCGATGACTTCGCTA
TAGCCATGAAGGCCGCCGTTGGGAAGTGGGGCGCCGACTCCTGGAAGGCTAATG
GCTTCGCCATCTTTTGTGGAGTTATGCTTGGTGTCAACAAGGCTGGAGATGCGGC
CCATGCTTACAACTTCACCCTCACCAAGGACCATGCTGACATTGTCTTCTTTGAGC
CTCAGAACGGTGGATACCTGAACGACATTGGCTATGACAGCTACATGGCCTTCTA
CTGA

B

MSLRRGIYHIENAGVPSAIDLKDGSSSDGTPIVGWQFTPDTINWHQLWLAEPIPNVAD
TFTLCNLFSGTYMDLYNGSSEAGTAVNGWQGTAFTTNPHQLWTIKKSSDGTSYKIQ
NYGSKTFVDLVNGDSSDGAKIAGWTGTWDEGNPHQKWYFNRMSVSSAEAQAAIAR
NPHIHGTYRGYILDGEYLVLPNATFTHIWKNSGLPGSKWREQIYDCDDFAIAMKAAV
GKWGADSWKANGFAIFCGVMLGVNKAGDAAHAYNFTLTKDHADIVFFEPQNGGY
LNDIGYDSYMAFY

Table S1. Oligonucleotide primers.

Primer Name	Sequence 5'-3'	Purpose
MOA_F	TTTCATATGTCTCTGCGACGCGGAATTTACC	SDM[1)]
MOA_R	TTTGAATTCTCAGTAGAAGGCCATGTAGCTGTCATAGC	SDM
C215A_F	GCAAATCTATGATGCGGATGACTTTGCTATAGCC	SDM
C215A_R	GGCTATAGCAAAGTCATCCGCATCATAGATTTGC	SDM
I222M_F	GACTTTGCTATAGCCATGAAGGCCGCCGTTGGG	SDM
I222M_R	CCCAACGGCGGCCTTCATGGCTATAGCAAAGTC	SDM
ΔC_R	TTTGAATTCTCAGTAAGTCCCATGAATATGAGGG	SDM
bre-3 (op504)_F	GCTCTGTCATTGTATGCCTGGGCC	Genotyping
bre-3 (op504)_R	CGATATAGAATTGATCCTTCCGTCC	Genotyping
oJL102	CAACCTTGACTGTCGAACCACCATAG	Sequencing/PCR
oJL115	GCTCAATTCGCGCCAAACTATG	Sequencing/PCR
oJL103	TCTGCGAGTTGTTTTTGCGTTTGAG	PCR
oJL104	ACAAAGAGCGAACGCAGACAGT	PCR
oJL114	AAAGATTCAGAAGGTCGGTAGATGGG	PCR
oJL116	GAACGAGAGGCAGATGGAGAGG	PCR
iPCR1a	GACCTTGTGAAGTGTCAACCTTGACTG	PCR
iPCR1b	GACAATCGATAAATATTTACGTTTGCGAGAC	PCR
iPCR2b	CATCTATATGTTCGAACCGACATTCCC	PCR

[1)]SDM: Site-Directed Mutagenesis

References in Supporting Information

1. Sambrook, J., and Russell, D. W. (2001) *Molecular cloning: A laboratory manual.*, Cold Spring Harbor Laboratory, New York
2. Brenner, S. (1974) *Genetics* **77**, 71-94
3. Kruger, R. P., Winter, H. C., Simonson-Leff, N., Stuckey, J. A., Goldstein, I. J., and Dixon, J. E. (2002) *J Biol Chem* **277**, 15002-15005
4. Tateno, H., and Goldstein, I. J. (2004) *Arch Biochem Biophys* **427**, 101-109
5. Schilling, O., and Overall, C. M. (2008) *Nat Biotechnol* **26**, 685-694
6. Kunzler, M., Bleuler-Martinez, S., Butschi, A., Garbani, M., Luthy, P., Hengartner, M. O., and Aebi, M. (2010) *Methods Enzymol* **480**, 141-150
7. Jang-Lee, J., North, S. J., Sutton-Smith, M., Goldberg, D., Panico, M., Morris, H., Haslam, S., and Dell, A. (2006) *Methods Enzymol* **415**, 59-86
8. Gerdt, S., Lochnit, G., Dennis, R. D., and Geyer, R. (1997) *Glycobiology* **7**, 265-275
9. Gerdt, S., Dennis, R. D., Borgonie, G., Schnabel, R., and Geyer, R. (1999) *Eur J Biochem* **266**, 952-963