Supporting Information

Sanglas *et al.* 10.1073/pnas.0812623106

SI Materials and Methods

Materials. Adult *Ascaris suum* worms were collected from pigs at a local slaughterhouse and stored in either ice or liquid nitrogen. All chemical reagents were purchased from Sigma unless otherwise stated. Antibiotics and restriction enzymes were from Roche Applied Science. T4 DNA ligase was from New England Biolabs, and DNA polymerases were from Ecogen and Biotools. Roche Applied Science synthesized all of the oligonucleotides. Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from Apollo Scientific. Human thrombin was from Sigma, and rabbit lung thrombomodulin was from American Diagnostica. Recombinant His-tagged tobacco-etch virus (TEV) protease was produced in *Escherichia coli* and purified as described elsewhere (1). Bovine carboxypeptidase A1 (bCPA1) was purchased from Sigma, and human CPN was from Elastin Product Co. Recombinant human CPA1, CPA2, CPA4, and CPB1, and *Drosophila* CPD domain I were produced in yeast and purified as described elsewhere (2, 3). Briefly, the proteins were expressed in *Pichia pastoris* strain KM71, in accordance with the instructions of the EasySelect *Pichia* expression kit (Invitrogen). The secreted proteins were purified from the supernatant by hydrophobic chromatography on a Toyopearl Butyl 650M column (Tosohaas) connected to an ÄKTA-purifier system (GE Healthcare) by applying a decreasing gradient of 30% $(NH_4)_2SO_4$. The eluted samples were desalted by dialysis against 20 mM Tris·HCl (pH 8.0) and then loaded onto an anion-exchange TSK-DEAE 5PW column (Tosohaas) by using a linear gradient from 0 to 10% of 0.4 M ammonium acetate. Human CPA1, CPA2, CPA4, and CPB1 zymogens were activated by a controlled digestion process with bovine trypsin (Sigma), by using a 1:100 (wt/wt) enzyme:substrate ratio at 4 °C for 24 h. Purification and activation were monitored by SDS/PAGE. Human CPA3 was kindly provided by G. Pejler (Swedish University of Agricultural Sciences, Uppsala, Sweden). Human plasma CPB2, also known as thrombin-activatable fibrinolysis inhibitor (TAFI), was a generous gift from J. J. Enghild (University of Århus, Århus, Denmark). The chromogenic metallocarboxypeptidase (MCP) substrates *N*-(4-methoxyphenylazoformyl)-Phe-OH and *N*-(4-methoxyphenylazoformyl)-Arg-OH were purchased from Bachem, and *N*-(3-[2-Furyl]acryloyl)-Ala-Lys was from Sigma.

Synthesis and Cloning of the ACI Gene According to the Reported Sequence. A gene encoding ACI according to the reported sequence (4) was synthesized by recursive PCR using overlapping oligonucleotides and optimized codons for overexpression in *E. coli*. The 220-bp PCR product was cloned into the pGEM-T Easy vector (Promega), transformed into competent *E. coli* XL1-Blue cells (Novagen), and the clones were analyzed by DNA sequencing. The gene was subsequently cloned into the following expression vectors: pBAT-4 (5) and pET-32a (Novagen) for intracellular expression in *E. coli*, pBAT-4-OmpA for periplasmic expression in *E. coli*, and pPICZα (Invitrogen) for extracellular expression in *P. pastoris*. Specific oligonucleotides were designed with the appropriate restriction sites to enable cloning into the different vectors. The pGEM-T Easy vector with the inserted construct was used as a template for all PCRs.

Heterologous Expression and Purification of ACI with the Reported Sequence. ACI as cloned into the pBAT-4 vector was expressed as described for tick carboxypeptidase inhibitor (TCI) (6). For expression with the pET-32a vector, which attaches an N-terminal thioredoxin-His_6 fusion and an enterokinase cleavage site, *E. coli* Origami2 (DE3) cells (Novagen) were transformed with the plasmid and grown overnight at 30 °C in Luria–Bertani (LB) medium with 100 μg/mL carbenicillin and 13 μg/mL tetracycline. The medium was then diluted 1:10 (vol/vol) in fresh LB containing the same antibiotics. Cells were grown at 37 °C and induced to express the protein at an OD_{550} of 0.8 with 0.5 mM IPTG (final concentration). The cells were incubated overnight at 18 °C and then harvested by centrifugation. The pellet was washed three times in 20 mM Tris·HCl, 150 mM NaCl (pH 7.5) and resuspended in 50 mM Tris·HCl, 50 mM NaCl, 10 mM imidazole (pH 8.0). Cells were lysed by sonication in an ice bath, and the cell debris was subsequently removed by centrifugation at 35,000 $\times$ *g* for 30 min at 4 °C. The supernatant was loaded onto a 5-mL HiTrap HP column (GE Healthcare) equilibrated with 50 mM Tris·HCl, 50 mM NaCl, 10 mM imidazole (pH 8.0). The column was washed in the same buffer further containing 100 mM imidazole at a flow rate of 1 mL/min, and the bound proteins were eluted with 450 mM imidazole. The eluted fractions were pooled, checked by 15% SDS/PAGE stained with Coomassie blue, desalted with a HiPrep 26/10 desalting column (GE Healthcare), and digested with 0.25 unit of enterokinase (Invitrogen) per mg of recombinant protein. The digestion products were analyzed by MALDI-TOF mass spectrometry (MS), and the recombinant ACI protein with the reported sequence was cleaved from the thioredoxin fusion protein and purified by reversed-phase high performance liquid chromatography (RP-HPLC) in a Waters Alliance apparatus as follows. The protein was loaded onto a 4.6-mm protein C4 column (Vydac Grace) by using a linear gradient from 10 to 50% acetonitrile with 0.1% trifluoroacetic acid at a flow rate of 0.75 mL/min for 50 min. ACI, with the reported sequence as cloned into the pBAT-4-OmpA vector, was expressed and purified as described in detail for potato carboxypeptidase inhibitor (PCI) and TCI (7, 8). ACI protein with the reported sequence cloned into the pPICZα vector was expressed in the *P. pastoris* strain KM71H as described above for different CPs. In all cases, protein expression was monitored by MALDI-TOF MS (see below for details).

Intensity-Fading MALDI-TOF MS Assay. *Ascaris* worms (10 specimens) maintained on ice were washed three times in PBS and homogenized by using a Polytron and 50 mM Tris·HCl, 150 mM NaCl (pH 7.5) as buffer. The insoluble fraction of the homogenate was removed by centrifugation at 35,000 $\times$ *g* for 30 min at 4 °C. The soluble fraction was loaded onto a Sep-Pak light C18 Cartridge (Waters) and washed in water, and the bound proteins were eluted with 60% acetonitrile and subsequently lyophilized. The protein extract was resuspended in 0.5 mL of 20 mM Tris·HCl (pH 7.5), analyzed by MALDI-TOF MS, and then treated as detailed below. bCPA1 was coupled to cyanide bromide-activated Sepharose 4B following the manufacturer's guidelines (GE Healthcare) to give CPA–Sepharose. For the intensity-fading MS assays (9), 2–3 μL of the *Ascaris* extract was mixed with 2–3 μL of CPA–Sepharose on a small strip of parafilm and incubated for 5 min at room temperature. Then, 0.5 μL from the top of the drop was recovered and analyzed by MALDI-TOF MS. The CPA–Sepharose was washed twice with 5 μL of 10 mM Tris·HCl (pH 7.5) and with 5 μL of 50 mM sodium phosphate (pH 11.0). Two to three microliters of 100% formic acid was mixed with the CPA–Sepharose and incubated for 5 min to release the bound carboxypeptidase inhibitor; 0.5 μL

from the top of the drop was removed for the final MALDI-TOF MS analysis. All samples were mixed with a matrix solution (1:1, vol/vol) of 10 mg/mL 2,6-dihydroxyacetophenone (Sigma) dissolved in 30% acetonitrile containing 20 mM dibasic ammonium citrate (pH 5.5); 0.5 μL of the mixture was spotted onto the MALDI-TOF plate by using the dried-droplet method. Mass spectra were acquired in an Ultraflex mass spectrometer (Bruker) equipped with a 337-nm laser in linear-mode geometry under 20 kV and ≈300 laser shots. A mixture of proteins from Bruker (protein calibration standard I; mass range 3,000–25,000 Da) was used as a standard.

Purification of Natural ACI. Approximately 1 kg of *Ascaris* worms was homogenized in 50 mM Tris·HCl, 150 mM NaCl (pH 7.5), by using a Polytron and centrifuged at 35,000 × *g* for 30 min at 4 °C. The soluble fraction of the extract was loaded onto a 20-mL CPA–Sepharose affinity column prepared as described above. The column was washed with 50 mM Tris·HCl (pH 9.0) followed by 50 mM Tris·HCl (pH 10.0) and eluted with 20 mM sodium phosphate (pH 12.0). The proteins eluted at this pH were lyophilized, and ACI was finally purified by RP-HPLC by using the conditions detailed above. All of the purification steps were monitored by MALDI-TOF MS.

Carboxymethylation and Amino Acid Sequencing. To determine the number of cysteine residues, the natural ACI inhibitor was reduced with 50 mM DTT in 100 mM Tris·HCl (pH 8.5) for 3 h and subsequently incubated with 200 mM 4-vinylpyridine for 45 min at room temperature in the dark. The solution was diluted 10 times with deionized water and analyzed by MALDI-TOF MS by using 2,6-dihydroxyacetophenone as matrix solution. Each free cysteine residue incorporates one *S*-pyridylethyl group with the consequent increase in mass of 105 Da. For the sequence analyses, ACI was reduced and carboxymethylated as described, freed from reagents by RP-HPLC, and either directly analyzed by automated Edman degradation in an Applied Biosystems model Procise 492 protein sequencer or digested with endoproteinase Lys-C (Roche Applied Science) in 20 mM Tris·HCl (pH 8.0) for 3 h at 37 °C by using an enzyme:substrate ratio of 1:100 (wt/wt). The resulting peptides were isolated by RP-HPLC on a Vydac C18 column with a linear gradient of 10–60% acetonitrile with 0.1% trifluoroacetic acid at 0.75 mL/min for 50 min and lyophilized. Peptide sequences were obtained by automated Edman degradation.

Molecular Cloning of Full-Length ACI cDNA. Total RNA was extracted from frozen *Ascaris* worms by using the nucleospin RNA L kit (Macherey Nagel) following the manufacturer's instructions. The first strand of the *Ascaris* cDNA was synthesized by using the adaptor oligonucleotide R_0R_1-dT (R_0, 5′-CCGGAATTCACTGCAG-3′; R_1, 5′-GGTACCCAATACGACTCACTATAGGGC-3′) and avian myeloblastosis virus reverse transcriptase (Roche Applied Science). To clone ACI cDNA, two degenerate oligonucleotides were designed from the amino acid sequence published for ACI (4): N1, GAYACIGAYTGYACIAAYGGIGARAARTG (corresponding to residues Asp-9–Cys-18) and N2, GGIGARAARTGYGTICARAARAAYAARAT (residues Gly-15–Ile-24), with Y = C or T, R = A or G, and I = inosine. The 3′-end fragment of ACI cDNA was obtained by using the 3′-RACE protocol. For the first round of PCR, the *Ascaris* cDNA was amplified by using oligonucleotide N1 and adaptor oligonucleotide R_0. PCR was performed in 25 cycles comprising each denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min, and extension at 72 °C for 2 min. PCR mixtures were diluted 20-fold and reamplified by using the N2 and R_1 nested oligonucleotides. The prominent PCR products were separated by electrophoresis on 2% agarose gels, recovered by using the QIAEX II gel extraction kit (Qiagen), and cloned into the pGEM-T Easy vector. Fifteen clones containing the amplified 3′ cDNA fragment were sequenced. The 5′ end fragment of ACI cDNA was obtained by 5′-RACE by using the Smart RACE cDNA amplification kit (BD Biosciences Clontech). Two nested oligonucleotides were designed on the basis of the sequence of the 3′ cDNA end: gene-specific primer 1 (5′-CGATTAGAGTAAATCCCAGCAGCAACC-3′, corresponding to nucleotides 290–316 of the full-length ACI cDNA) and gene-specific primer 2 (5′-GCAAATCTTTTCATGTGCCC-3′, nucleotides 264–283). After synthesis of the second cDNA strand and ligation of the adaptor oligonucleotide, "semi-nested" PCR was performed with the adaptor primer mix and gene-specific primers 1 and 2 in the first and second rounds of PCR, respectively. The PCR products were analyzed and cloned as described above. Twenty clones containing the amplified 5′ cDNA fragment were sequenced to generate a consensus sequence.

Heterologous Expression and Purification of Recombinant ACI. The C terminus of the reported ACI sequence (-LPWGL[65]) was substituted for the C-terminal residues as determined in this study (-GCCWDLL[67]) by using oligonucleotides designed considering optimal *E. coli* codon usage. After PCR amplification, the new *aci* gene was cloned into the pBAT-4-OmpA and pET-32a vectors for expression in *E. coli*. The enterokinase recognition site of the pET-32a vector was replaced by a TEV protease site to improve the cleavage efficiency of the fusion protein. Recombinant ACI was expressed and purified by using the pBAT-4-OmpA and pET-32a systems as described above. After the purification of the thioredoxin-His_6-ACI fusion protein, eluted from the Ni^{2+}-affinity column with 450 mM imidazole, the purified fractions were pooled and dialyzed in cellulose membranes (cutoff 1,000 Da; Membrane Filtration Products) against 50 mM Tris·HCl, 0.5 mM EDTA (pH 7.5), further containing 0.5 mM oxidized glutathione and 3 mM reduced glutathione. TEV protease was added during dialysis at an enzyme:substrate ratio of 1:150 (wt/wt) for 24 h at room temperature. Recombinant ACI was finally purified to homogeneity by RP-HPLC as detailed above for the inhibitor with the incorrect sequence. Purity and molecular mass of the inhibitor were assessed, respectively, by SDS/PAGE and MALDI-TOF MS.

Circular Dichroism and NMR Spectroscopy. Samples for far-UV circular dichroism (CD) spectroscopy were prepared by dissolving natural and recombinant ACI at a final protein concentration of 0.5 mg/mL in 20 mM sodium phosphate (pH 8.0). CD analyses were carried out in a Jasco J-715 spectrometer at 25 °C by using a cell with 2-mm path length. Denaturation and reduction studies were accomplished by incubating recombinant ACI for 24 h in the same phosphate buffer, which was supplemented with one of the following reagents: 8 M urea, 6–8 M guanidine hydrochloride (GdnHCl), and 10 mM DTT, or in 50 mM Tris·HCl (pH 8.5), containing both 8 M GdnHCl and 0.25 mM β-mercaptoethanol. For one-dimensional NMR spectroscopy experiments, lyophilized protein powders of natural and recombinant ACI were resuspended to a final protein concentration of 1 mg/mL in 20 mM sodium phosphate (pH 7.0), prepared at a H_2O:D_2O volume ratio of 9:1 (99.9% D_2O). NMR spectra were acquired at different temperatures (25–90 °C) in a Bruker AVANCE 600-MHz spectrometer by using solvent-suppression WATERGATE techniques. A two-dimensional NOESY spectrum of recombinant ACI (8 mg/mL) was recorded at 25 °C with the mixing times set to 200 ms. The NMR spectra were processed and analyzed by using the TopSpinv2.0 software packages from Bruker Biospin.

MCP Inhibitory Activity. The inhibitory activity of natural ACI and of the recombinant forms comprising the reported sequence and the one determined in the present study was tested by measuring

inhibition of the hydrolysis of the chromogenic substrate *N*-(4-methoxyphenylazoformyl)-Phe-OH by A-type MCPs and *N*-(4-methoxyphenylazoformyl)-Arg-OH by B-type MCPs at 350 nm in a Cary 400 UV-visible spectrophotometer (Varian). All inhibitory assays were performed in 50 mM Tris·HCl, 100 mM NaCl (pH 7.5) containing 100 μM substrate. Inhibition constants (K_i) for the complexes of ACI with the MCPs assayed were determined by using presteady-state kinetics as described for tight binding inhibitors (10). Titration experiments revealed that recombinant ACI is fully functional (specific activity >90% of the theoretical values). The concentration of ACI used in the assays was determined based on A_{280} and a calculated extinction coefficient of $E_{0.1\%} = 1.5$. TAFI was activated to generate TAFIa immediately before use by incubation of the zymogen (2.2 μM) with thrombin (160 nM) and thrombomodulin (200 nM) in 20 mM Tris·HCl, 150 mM NaCl (pH 7.5), for 40 min at 23 °C. The inhibitory activity of ACI toward CPN was tested by measuring its effect on the hydrolysis of the chromogenic substrate *N*-(3-[2-Furyl]-acryloyl)-Ala-Lys. The assay was performed in 50 mM Tris·HCl (pH 7.5), containing 100 mM NaCl and 0.35 mM substrate at 37 °C and 340 nm.

Western Blotting of Natural ACI. Mouse polyclonal antibodies anti-ACI were raised at the Servei de Producció d'Anticossos (Universitat Autònoma de Barcelona, Bellaterra, Spain). Three 8-week-old BALB/c mice were immunized by i.p. injection of 40 μg of recombinant ACI as immunogen. Freund's complete adjuvant was used only for the first immunization, and incomplete adjuvant was used in other immunizations. Blood samples were collected from animals before the first injection and after four immunizations. The immune response was monitored by ELISA, and antibody titers were determined as the inverse of the dilution that gave OD_{450} values corresponding to 50% of maximal signal. Total serum was collected and stored at −20 °C until use. IgGs were purified from serum by using a 1-mL HiTrap protein A HP column (GE Healthcare) following the manufacturer's guidelines. Body wall, intestine, testis and vas deferens (male), seminal vesicle (male), ovary and oviduct (female), and uterus (female) were excised from adult *Ascaris* worms. Tissue samples were washed three times in PBS and homogenized in 50 mM Tris·HCl, 150 mM NaCl (pH 7.5), by using a Polytron. Homogenates were centrifuged at 35,000 × *g* for 30 min at 4 °C, and the soluble fraction was analyzed by 15% SDS/PAGE. Gels were stained with Coomassie blue or transferred electrophoretically to polyvinylidene fluoride membranes (Millipore). The membranes were blocked with 5% nonfat milk in Tris-buffered saline (TBS) containing 0.1% Tween 20 for 1 h at room temperature and subsequently incubated overnight at 4 °C with polyclonal anti-ACI antibodies at 1:500 dilution. After washing in 0.1% Tween/TBS, the membrane was incubated for 2 h at room temperature with a 1:3,000 dilution of goat anti-mouse IgG peroxidase-conjugated secondary antibody (Pierce). Immunolabeling was detected by using Immobilon Western chemiluminescent horseradish peroxidase substrate (Millipore) and visualized on a Versadoc imaging system (Bio-Rad).

Immunohistochemistry of Natural ACI. Adult male and female *Ascaris* worms were fixed in 4% paraformaldehyde for 10 min at room temperature, injected with the same fixative, cut in 2-cm pieces, and then fixed for another 24 h. Segments from various regions were embedded in paraffin and cut in transverse sections of 5-μm thickness. The sections were deparaffinized with xylene and hydrated through descending graded series of ethanol. The slides were then treated with 2% hydrogen peroxide in 70% methanol and blocked for 1 h in PBS containing 5% FBS and 0.2% Triton X-100, followed by treatment with an avidin/biotin blocking kit (Vector Laboratories). Sections were subsequently incubated overnight at 4 °C with anti-ACI antibody diluted at 1:50 in the same blocking solution. Mouse serum obtained before immunization was used as negative control. After the immunological reaction, sections were washed in PBS containing 0.2% Triton X-100 and incubated with biotinylated horse anti-mouse IgG (Vector Laboratories) diluted at 1:500 for 1 h. The washed slides were incubated in avidin–biotin–peroxidase complex reagent of the Vectastain Elite ABC kit (Vector Laboratories) for 30 min, washed, and incubated for 10 min in 3.3′-diaminobenzidine tetrahydrochloride (DAB+; Dako). Sections were then washed in distilled water, counterstained with Harris hematoxylin (Sigma), dehydrated in ethanol, and cleared in xylene. Coverslips were applied over DPX mountant (Sigma). Each section was visualized in a Nikon Eclipse 90i microscope, and images were taken with a Nikon digital camera DXM 1200F coupled to computerized image processing software (ACT-1 version 2.63).

Crystallization of the Complex. The ACI–hCPA1 complex was prepared by incubating equimolar quantities of inhibitor and protease for 1 h at room temperature. The complete inhibition of the carboxypeptidase was checked by a continuous spectrophotometric assay with the chromogenic substrate *N*-(4-methoxyphenylazoformyl)-Phe-OH. The complex was purified by a HiLoad Superdex 75 26/60 column (GE Healthcare) equilibrated with 50 mM Tris·HCl, 250 mM NaCl (pH 7.5). The complex was buffer-exchanged to 10 mM Tris·HCl, 50 mM NaCl (pH 7.5) and concentrated to 16.5 mg/mL by using an Amicon Centricon (10-kDa cutoff, Millipore). The purity and integrity of the complex were assessed by SDS/PAGE and MALDI-TOF MS (data not shown). Crystallization assays were performed following the sitting-drop vapor diffusion method. Reservoir solutions were prepared by a Tecan robot, and 200-nL crystallization drops were dispensed on 96- × 3-well CrystalQuick plates (Greiner) by a Cartesian nanodrop robot (Genomic Solutions) at the joint IBMB-CSIC/IRB/Barcelona Science Park High-Throughput Crystallography Platform (PAC). Best crystals appeared after 2–3 weeks in a Bruker steady-temperature crystal farm at 20 °C with 0.2 M zinc acetate dihydrate, 0.1 M sodium cacodylate, 9% PEG 8000 (pH 6.5) as reservoir solution. These conditions were successfully scaled up to the microliter range with Cryschem crystallization dishes (Hampton Research) by using drops containing 3 μL of protein solution (16.5 mg/mL) and 1 μL of reservoir solution. Crystals were cryoprotected through stepwise replacement of the mother liquor with 0.2 M zinc acetate, 0.1 M sodium cacodylate, 35% PEG 8000 (pH 6.5). A complete diffraction dataset was collected at 100 K from a single N_2 flash-cryocooled (Oxford Cryosystems) crystal on an ADSC Q315R CCD detector at beam line ID23–1 of the European Synchrotron Radiation Facility (ESRF; Grenoble, France) within the Block Allocation Group "BAG Barcelona." Crystals were orthorhombic and harbored one complex per asymmetric unit. Diffraction data were integrated, scaled, merged, and reduced with programs XDS (11) and SCALA within the CCP4 suite (12) (see Table S2).

Structure Solution and Refinement. The structure was solved by Patterson search methods with program AMoRe (13) by using all diffraction data between 15- and 4-Å resolution. The coordinates of a hCPA1 monomer [Protein Data Bank (PDB) code 2v77] were used as a searching model. A single solution was found at 76.0, 64.4, 76.2 (α, β, γ in Eulerian angles) and 0.163, 0.420, 0.304 (x, y, z, as fractional unit cell coordinates) after rigid-body refinement. This solution gave a correlation coefficient in structure factor amplitudes of 56.3% and a crystallographic R factor of 39.4% (for definitions, see Table S2 and ref. 13). An electron density map based on calculations with phases derived from the correctly positioned and oriented model revealed extra density for the inhibitor. Subsequently, manual

model building on a Silicon Graphics workstation with program TURBO-Frodo alternated with crystallographic refinement with REFMAC5 within the CCP4 suite until completion of the model (see Table S2). This model contained the residues of the protease moiety (molecule A) from Ser-2 to Pro-308 and the catalytic zinc ion and the inhibitor residues from Val-3 to Leu-67 (molecule B). All residues lie in allowed regions of a Ramachandran plot according to ref. 14 except Ser-199 ($\Phi = 145°$; $\Psi = -12°$), as usual in funnelin structures. In addition, a further nine zinc ions, six acetate molecules, and four cacodylate molecules (attributed to the crystallization conditions) were identified in the electron density maps, as well as 564 solvent molecules (see Table S2).

Miscellaneous. Nucleotide and amino acid similarity searches were carried out with BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Fig. 3 was prepared with programs SETOR (15) and TURBO-Frodo. Structural similarity searches were performed with DALI (http://ekhidna.biocenter.helsinki.fi/dali_server), SSM within SCOP (www.ebi.ac.uk/msd-srv/ssm), VAST (www.ncbi.nlm.nih.gov/structure/vast/vastsearch.html), and the CATHEDRAL server within CATH (www.cathdb.info/cgi-bin/cath). None of these searches revealed substantial structural similarity beyond generic strand–turn–strand–helix-containing motifs to any protein of known structure. Surfaces and intramolecular close contacts (<4 Å) were calculated by using a sphere radius of 1.4 Å with program CNS (16). The value of the complex interface was estimated taking the half of the difference between the sum of the individual molecular surfaces and the total complex surface.

Structure of Human Carboxypeptidase A1. The structure of hCPA1 consists of a central eight-stranded β-sheet (strands β1–β8), which vertically spans the molecule top to bottom, accumulating a vertical twist. Three helices (α5, α7, α8) and the active-site cleft nestle into the concave face of the sheet, and six helices (α1–α4, α6, and α9) plus the N and C termini of the molecule lie on the convex side of the sheet (Fig. 3*A*). The access to the active site is like a funnel, whose rim is shaped by a series of irregular loop segments required for interactions of such A/B-type protease moieties with pro-domains and cognate protein inhibitors (17, 18). These segments comprise the loop connecting strand β3 with helix α2 (Lβ3α2), Lβ5β6, Lα7α8, and Lβ8α9, and the initial and central parts of the 52-residue segment connecting α4 and α5 (Lα4α5). This long segment features a compact subdomain, cross-linked by a disulfide bond (Cys-138–Cys-161), and it closes the front and the bottom of the active site and the specificity pocket, thus contributing to the characteristic cul-de-sac of funnelin exopeptidases.

The catalytic zinc ion of hCPA1 resides at the bottom of the funnel-like cleft and is coordinated by His-69 Nδ1 (2.08 Å apart) and, asymmetrically, by Glu-72 Oε2 (1.97 Å) and Oε2 (2.85 Å) (Fig. 7*B*). These 2 residues are embedded in a consensus motif sequence, HXXE (amino acid one-letter code; X for any residue), characteristic of A/B- and N/E-type MCPs (17, 19). The third zinc ligand is His-196 Nδ1 (2.10 Å). As found in related enzymes (17, 20, 21), the protein residues engaged in substrate binding are Tyr-198, Ser-199, Tyr-248, and Phe-279, which form substrate-binding subsite S_1 (for subsite nomenclature, see ref. 22); Arg-71, Glu-163, Thr-164, and Arg-127 contributing to S_2. In addition, Arg-71, Tyr-198, and Phe-279 also contribute to S_3. The commonly accepted mechanism foresees that a zinc-bound solvent molecule, which is further polarized by the general base/acid Glu-270, attacks the scissile bond carbonyl group (for a detailed review, see ref. 17 and references therein). Typical A-type MCP specificity toward aliphatic side chains in substrates is caused by a rather small hydrophobic $S_{1'}$ pocket. In hCPA1, it is shaped by the side chains of Asn-144, Arg-145, Ser-194, Met-203, Ile-243, Ile-247, Ala-250, Ser-253, Thr-254, Thr-268 (present in alternate conformation for its side chain), Tyr-248, and Ile-255. The terminal carboxylate group of a substrate, when it is trapped for scission, is fixed by Asn-144, Arg-145, and Tyr-248, whereas the scissile carbonyl group is near Glu-270, Arg-127, and the catalytic zinc (Fig. 7*B*).

1. Fang L, *et al.* (2007) An improved strategy for high-level production of TEV protease in *Escherichia coli* and its purification and characterization. *Protein Expr Purif* 51:102–109.
2. Reverter D, Ventura S, Villegas V, Vendrell J, Aviles FX (1998) Overexpression of human procarboxypeptidase A2 in *Pichia pastoris* and detailed characterization of its activation pathway. *J Biol Chem* 273:3535–3541.
3. Ventura S, *et al.* (1999) Mapping the pro-region of carboxypeptidase B by protein engineering: Cloning, overexpression, and mutagenesis of the porcine proenzyme. *J Biol Chem* 274:19925–19933.
4. Homandberg GA, Litwiller RD, Peanasky RJ (1989) Carboxypeptidase inhibitors from *Ascaris suum*: The primary structure. *Arch Biochem Biophys* 270:153–161.
5. Peranen J, Rikkonen M, Hyvonen M, Kaariainen L (1996) T7 vectors with modified T7lac promoter for expression of proteins in *Escherichia coli*. *Anal Biochem* 236:371–373.
6. Arolas JL, *et al.* (2005) A carboxypeptidase inhibitor from the tick *Rhipicephalus bursa*: Isolation, cDNA cloning, recombinant expression, and characterization. *J Biol Chem* 280:3441–3448.
7. Arolas JL, Bronsoms S, Ventura S, Aviles FX, Calvete JJ (2006) Characterizing the tick carboxypeptidase inhibitor: Molecular basis for its two-domain nature. *J Biol Chem* 281:22906–22916.
8. Arolas JL, *et al.* (2004) Secondary binding site of the potato carboxypeptidase inhibitor: Contribution to its structure, folding, and biological properties. *Biochemistry* 43:7973–7982.
9. Yanes O, Villanueva J, Querol E, Aviles FX (2007) Detection of noncovalent protein interactions by ''intensity fading'' MALDI-TOF mass spectrometry: Applications to proteases and protease inhibitors. *Nat Protoc* 2:119–130.
10. Bieth JG (1995) Theoretical and practical aspects of proteinase inhibition kinetics. *Methods Enzymol* 248:59–84.
11. Kabsch W (2001) XDS. *International Tables for Crystallography. Volume F. Crystallography of Biological Macromolecules*, eds Rossmann MG, Arnold E (Kluwer, for the International Union of Crystallography), Dordrecht, The Netherlands), 1st Ed, pp 730–734.
12. CCP4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
13. Navaza J (1994) AMoRe: An automated package for molecular replacement. *Acta Crystallogr A* 50:157–163.
14. Davis IW, *et al.* (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35:W375–W383.
15. Evans SV (1993) SETOR: Hardware lighted three-dimensional solid model representations of macromolecules. *J Mol Graphics* 11:134–138.
16. Brünger AT, *et al.* (1998) Crystallography & NMR System: A new software suite for macromolecular structure determination. *Acta Crystallogr D* 54:905–921.
17. Gomis-Ruth FX (2008) Structure and mechanism of metallocarboxypeptidases. *Crit Rev Biochem Mol Biol* 43:319–345.
18. Pallares I, *et al.* (2005) Structure of human carboxypeptidase A4 with its endogenous protein inhibitor, latexin. *Proc Natl Acad Sci USA* 102:3978–3983.
19. Hooper NM (1994) Families of zinc metalloproteases. *FEBS Lett* 354:1–6.
20. Auld DS (2004) Carboxypeptidase A. *Handbook of Proteolytic Enzymes*, eds Barrett AJ, Rawlings ND, Woessner JF, Jr (Elsevier, London), 2nd Ed, Vol 1, pp 812–821.
21. Vendrell J, Querol E, Aviles FX (2000) Metallocarboxypeptidases and their protein inhibitors: Structure, function and biomedical properties. *Biochim Biophys Acta* 1477:284–298.
22. Abramowitz N, Schechter I, Berger A (1967) On the size of the active site in proteases. II. Carboxypeptidase A. *Biochem Biophys Res Commun* 29:862–867.

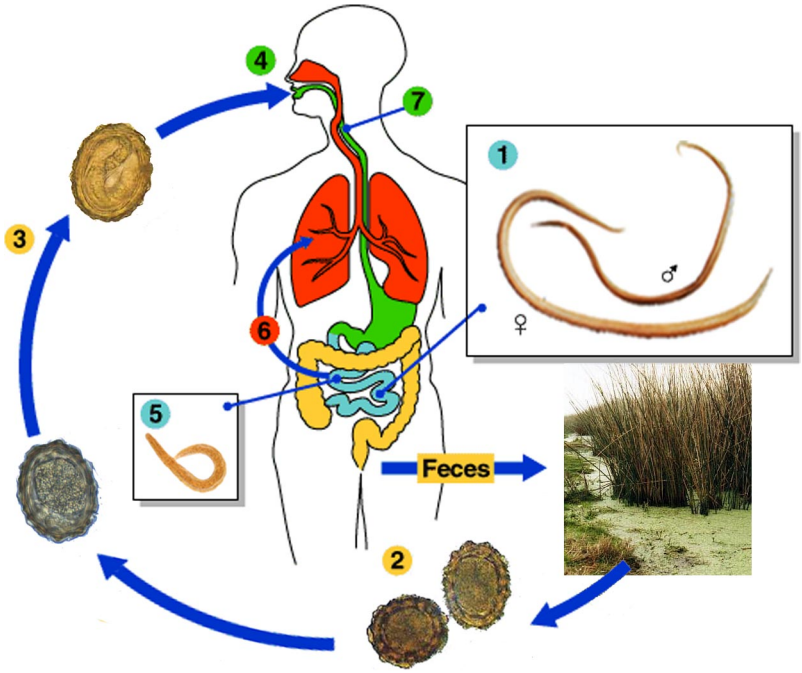

Fig. S1. Life cycle of the helminth *Ascaris*. Adult worms (point 1) live in the lumen of the small intestine. A mature female may produce daily up to 200,000 eggs, which are secreted with the feces (point 2). Unfertilized eggs may be ingested but are not infective. Fertile eggs embryonate and become infective after 18 days to several weeks (point 3), depending on the environmental conditions (optimum: moist, warm, shaded soil). After infective eggs are swallowed (point 4), larvae hatch (point 5), invade the intestinal mucosa, and are carried via the blood system to the lungs (point 6). The larvae mature further in the lungs (10–14 days), penetrate the alveolar walls, ascend the bronchial tree to the throat, and are then (re)swallowed (point 7). Upon reaching the small intestine, they develop into adult worms (point 1) that reach a maximal length of 35 cm. Between 2 and 3 months are required from ingestion of the infective eggs to oviposition by the adult female. Adult worms can live 1–2 years. Figure adapted from DPDx, the CDC parasitology website (www.dpd.cdc.gov/dpdx/default.htm).

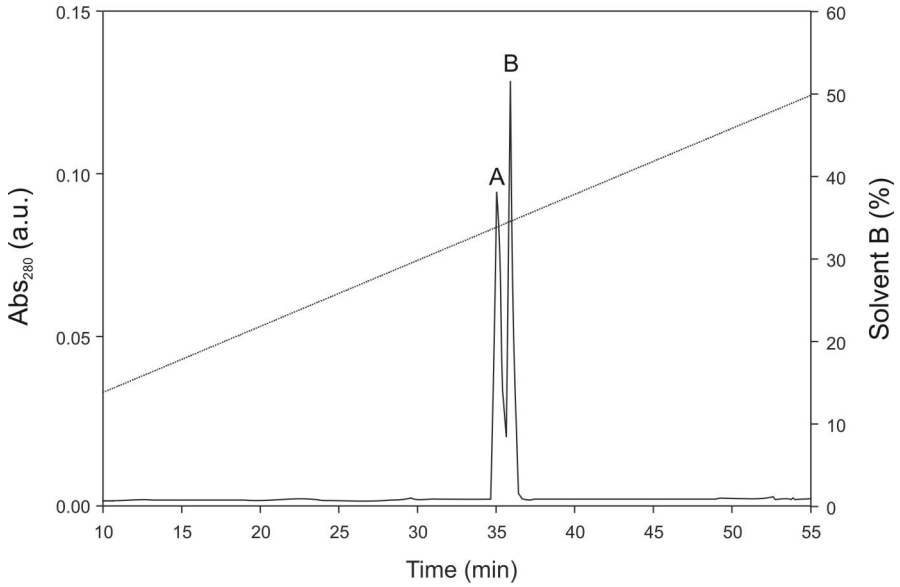

Fig. S2. RP-HPLC analysis of recombinant ACI comprising the reported sequence (4). ACI expressed in *E. coli* by using either the pBAT-4-OmpA or the pET-32a vectors was analyzed by RP-HPLC on a Vydac C4 column by using a linear gradient from 10 to 50% acetonitrile with 0.1% trifluoroacetic acid for 50 min. Two forms of the inhibitor, named A and B, were detected eluting, respectively, after 35 and 36 min of chromatography.

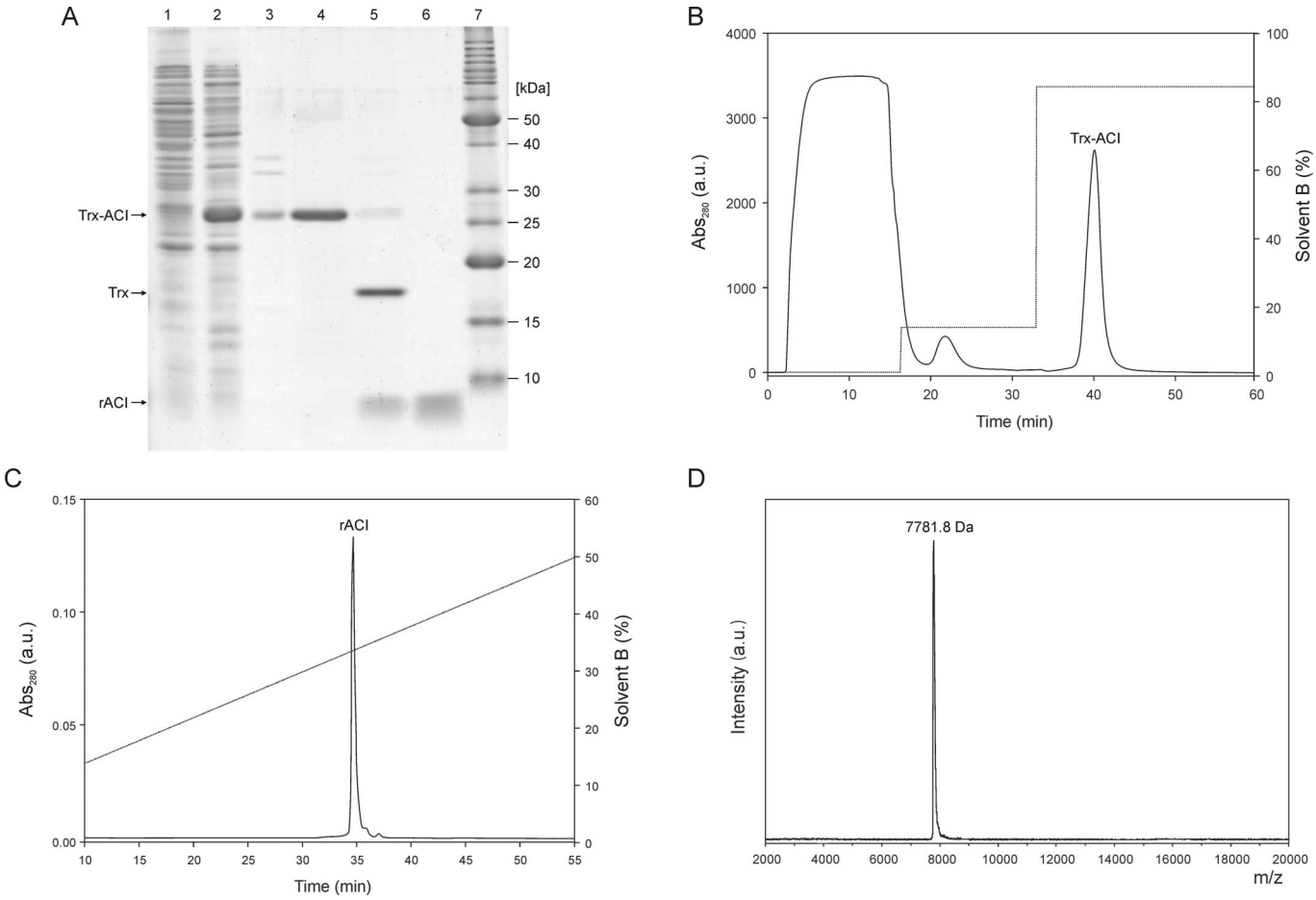

Fig. S3. Heterologous expression and purification of recombinant ACI. (*A*) Coomassie-stained SDS/PAGE of recombinant ACI produced as a fusion protein with thioredoxin A and purified by Ni^{2+}-affinity chromatography and RP-HPLC. Lane 1, cell lysate of *E. coli* Origami2(DE3) cells transformed with the mock pET-32a vector. Lanes 2 and 3, total proteins of the soluble and insoluble fractions, respectively, of cell lysates from Origami cells expressing the fusion protein with induction at 18 °C. Lane 4, fusion protein purified by Ni^{2+}-affinity chromatography. Lane 5, protein mixture generated by digestion of the fusion protein with TEV protease. Lane 6, recombinant ACI purified by RP-HPLC. Lane 7, molecular mass standard (BenchMark Protein Ladder; Invitrogen) ranging from 10 to 220 kDa. (*B*) Purification of the fusion protein by Ni^{2+}-affinity chromatography, washing the column with 100 mM imidazole and eluting with 450 mM imidazole. (*C*) RP-HPLC analysis of the cleaved recombinant ACI on a Vydac C4 column by using a linear gradient from 10 to 50% acetonitrile with 0.1% trifluoroacetic acid for 50 min. (*D*) MALDI-TOF mass spectrum of pure recombinant ACI with an extra N-terminal glycine residue resulting from the TEV digestion.

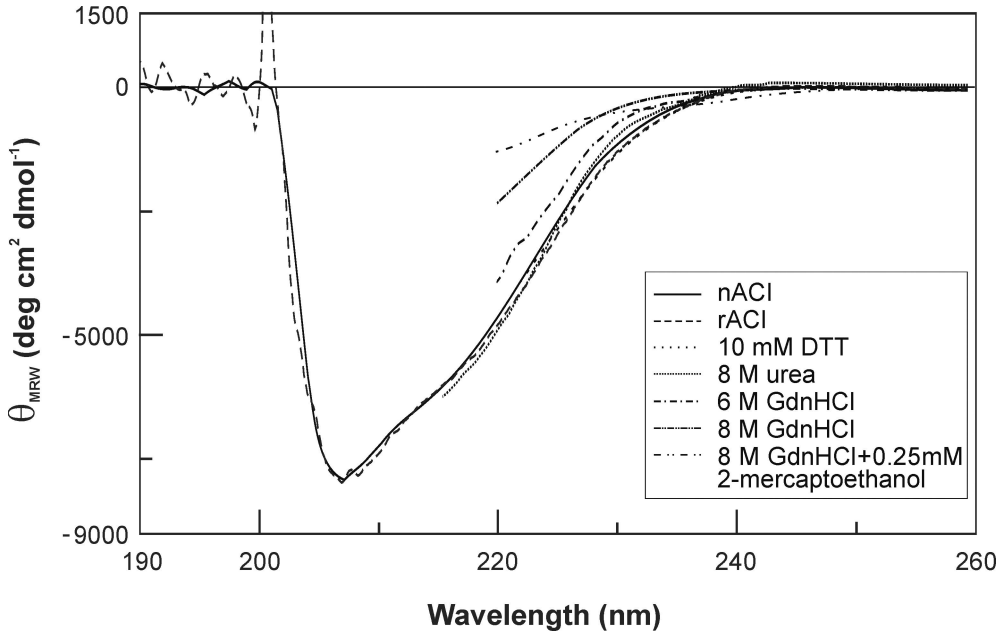

Fig. S4. Spectrophotometric analysis of ACI. CD analysis of natural and recombinant ACI was carried out at 0.5 mg/mL in 20 mM sodium phosphate (pH 8.0) and 25 °C. CD spectra for recombinant ACI were also recorded after incubation of the protein for 24 h in the same buffer containing the indicated concentrations of reducing agent and/or denaturant.

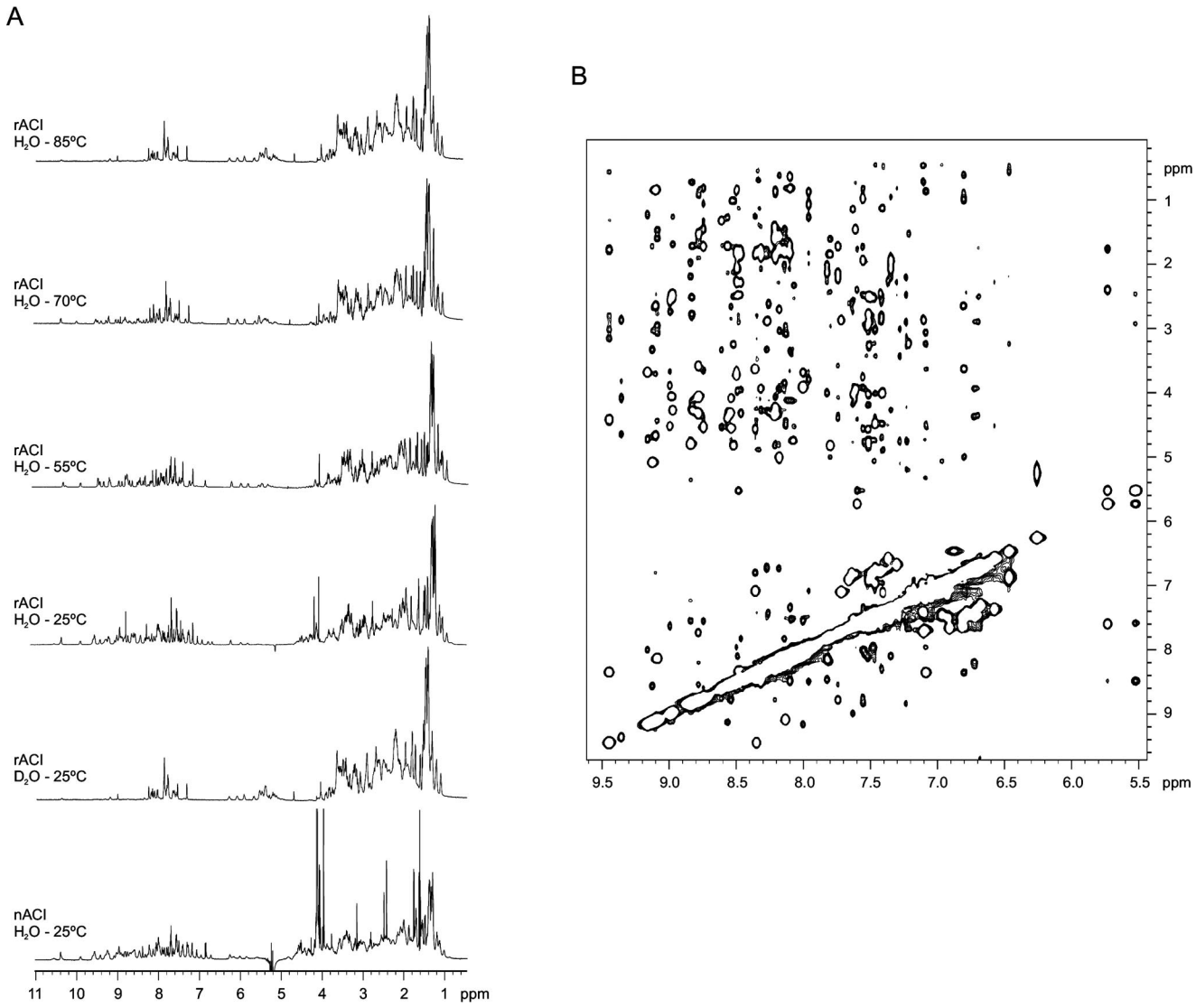

Fig. S5. Western blot of natural ACI. (*A*) Coomassie-stained SDS/PAGE of total extracts of male and female *Ascaris* and of dissected regions of the worm. Lane 1, molecular mass standard (SeeBlue prestained standard; Invitrogen) ranging from 4 to 250 kDa. Lanes 2 and 3, total proteins of male and female worms, respectively. Proteins are from the dissected body wall (lane 4), intestine (lane 5), testis and vas deferens (lane 6), seminal vesicle (lane 7), ovary and oviduct (lane 8), and uteri (lane 9). (*B*) Western blot analysis of the samples loaded onto the SDS/PAGE. The blot was probed with an anti-ACI polyclonal antibody.

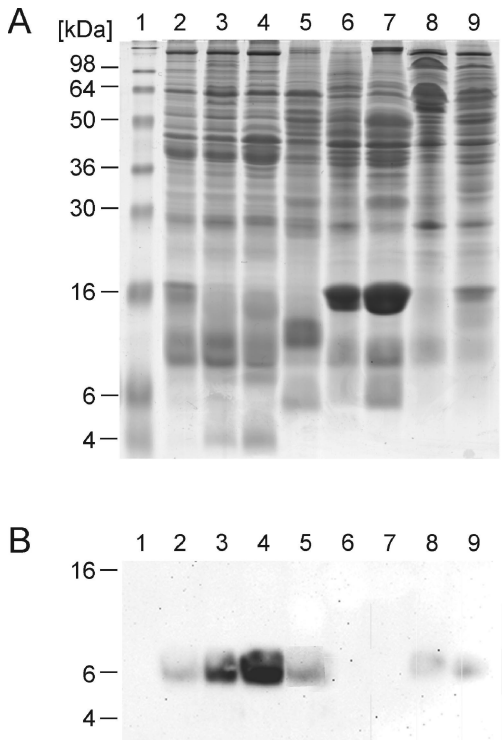

Fig. S6. One- and two-dimensional NMR analysis of ACI. (*A*) Monodimensional spectra of natural and recombinant ACI (1 mg/mL) were acquired at 25 °C and 600 MHz in 20 mM sodium phosphate (pH 7.0) containing 10% D_2O. For recombinant ACI, 1D NMR spectra were also recorded at increasing temperatures (25–85 °C) in 10% D_2O and at 25 °C in 99.9% D_2O. (*B*) Selected region of the NOESY spectrum of recombinant ACI (8 mg/mL) obtained at 25 °C and 600 MHz in 20 mM sodium phosphate (pH 7.0) containing 10% D_2O.

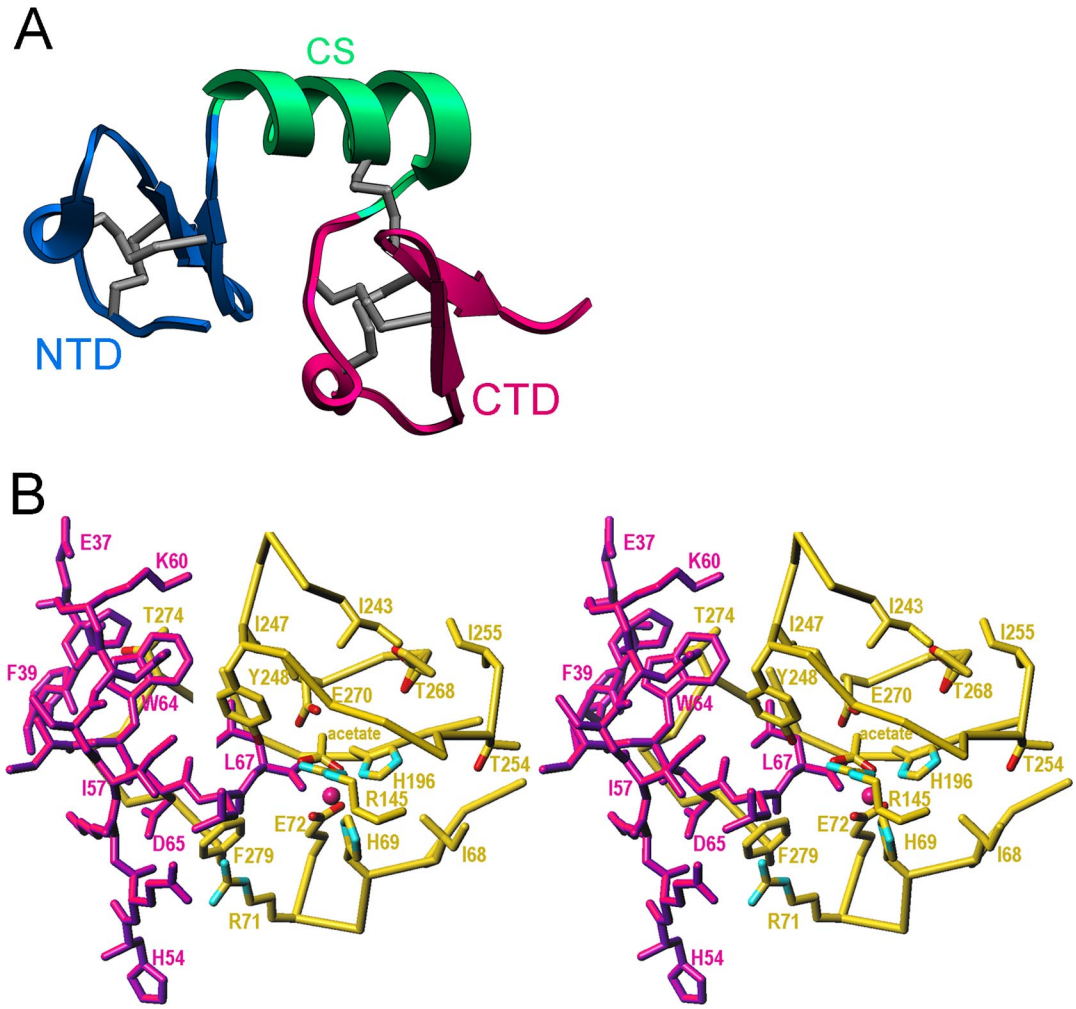

Fig. S7. Structure of ACI and its binding with hCPA1. (*A*) Richardson plot of ACI showing its domain architecture, regular secondary structure elements (ribbons for the concatenated 1,4-turns and the α-helix; arrows for β-strands), and disulfide bonds (in gray). The N-terminal domain is shown in blue, the connecting segment in green, and the C-terminal domain in magenta. (*B*) Close-up view in stereo of the interface between ACI (magenta stick model; selected residues are labeled) and hCPA1 (yellow Cα model). Selected side chains of the protease are displayed and labeled. The active-site zinc is shown as a magenta sphere.

Table S1. Interactions between ACI and hCPA1

ACI	hCPA1	Distance, Å
Hydrogen bonds and polar interactions		
Leu-67I O	His69 Nδ1	3.03
Leu-66I O	Arg-71 Nη1	3.02
Leu-66I O	Arg-127 Nη2	3.32
Leu-67I O	Arg-127 Nη2	2.87
Glu-55I Oϵ1	Glu-163 N	2.94
Leu-67I OT	His-196 Nδ1	3.35
Asp-65I Oδ1	Tyr-198 Oη	2.62
Lys-60I Nζ	Lys-245 O	3.14
Lys-60I Nζ	Ala-246 O	3.36
Asn-59I Nδ2	Ile-247 O	2.75
Lys-60I Nζ	Ile-247 O	3.35
Leu-67I N	Tyr-248 Oη	2.81
Residues making hydrophobic C–C interactions		
His-54I	Met-125	3.45
Trp-64I	Ile-247	3.78
Leu-67I	Ile-247	3.88
Ile-57I	Tyr-248	3.81
Phe-39I	Gly-275	3.54
Asp-65I	Phe-279	3.53

Table S2. Crystallographic data

Space group/cell constants (*a*, *b*, and *c*), Å	$P2_12_12_1$/55.47, 78.70, 84.50
Wavelength, Å	1.0000
No. of measurements/unique reflections	352,964/49,284
Resolution range, Å (outermost shell)*	46.4–1.60 (1.69–1.60)
Completeness, %	99.6 (97.7)
$R_{r.i.m.}$ (= R_{meas})[†]/$R_{p.i.m.}$[†]	0.086 (0.499)/0.032 (0.184)
Average intensity $\langle[\langle I\rangle/\sigma(\langle I\rangle)]\rangle$	18.8 (4.6)
B factor (Wilson), Å^2/average multiplicity	11.6/7.2 (7.1)
Resolution range used for refinement, Å	46.4–1.60
No. of reflections used (test set)	48,532 (752)
Crystallographic *R* factor (free *R* factor)[‡]	0.157 (0.200)
No. of protein atoms[§]/solvent molecules/ions and ligands	2,967/564/9 (zinc), 6 (acetate), 4 (cacodylate)
Rmsd from target values	
Bonds, Å/angles, °	0.011/1.30
Bonded *B* factors (main chain/side chain), Å^2	0.89/2.18
Average *B* factors for protein atoms, Å^2	10.4
Main-chain conformational angle analysis[¶]	
Residues in favored regions/outliers/all residues	360/1/368

*Values in parentheses refer to the outermost resolution shell.

[†]$R_{r.i.m.} = \Sigma_{hkl}(n_{hkl}/[n_{hkl} - 1]^{1/2})\Sigma_i \mid I_i(hkl) - \langle I(hkl)\rangle \mid /\Sigma_{hkl}\Sigma_i I_i(hkl)$ and $R_{p.i.m.} = \Sigma_{hkl}(1/[n_{hkl} - 1]^{1/2})\Sigma_i \mid I_i(hkl) - \langle I(hkl)\rangle \mid /\Sigma_{hkl}\Sigma_i I_i(hkl)$.

[‡]Crystallographic *R* factor $= \Sigma_{hkl} \| F_{obs} \mid - k \mid F_{calc} \| / \Sigma_{hkl} \mid F_{obs} \mid$; free *R* factor, same for a test set of reflections not used during refinement.

[§]Including atoms in alternate conformation.

[¶]According to ref. 14.