SUPPORTING INFORMATION

Crystal structure and mechanism of human carboxypeptidase O, insights for the specific activity toward acidic residues

Maria C. Garcia-Guerrero[a,1], Javier Garcia-Pardo[a,b,1], Esther Berenguer[a], Roberto Fernandez-Alvarez[a], Gifty B. Barfi[c], Peter J. Lyons[c], Francesc X. Aviles[a], Robert Huber[d,e,f,2], Julia Lorenzo[a,2] and David Reverter[a,2]

[a]Institute for Biotechnology and Biomedicine, and Department of Biochemistry and Molecular Biology, Universitat Autonoma de Barcelona, 08193 Bellaterra, Barcelona, Spain.

[b] Catalan Institute of Nanoscience and Nanotechnology (ICN2), CSIC and The Barcelona Institute of Science and Technology, Campus UAB, Bellaterra 08193, Barcelona, Spain.

[c]Department of Biology, Andrews University, Price Hall 216, Berrien Springs, MI, 49104 USA.

[d]Max Planck Institut für Biochemie, Am Klopferspitz 18, D-82152 Martinsried, Germany.

[e]Zentrum für Medizinische Biotechnologie, Universität Duisburg-Essen, D-45117 Essen, Germany.

[f]Fakultät für Chemie, Technische Universität München, Lichtenbergstraße 4. D-85747 Garching Germany

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed: Robert Huber, Max Planck Institut für Biochemie, 82152 Martinsried, Germany. Ph: +49 (0)89 8578 2677. Fax: +49 (0)89 8578 3516. E-mail: huber@biochem.mpg.de .

-To whom correspondence should be addressed: Julia Lorenzo, Universitat Autonoma de Barcelona, 08193 Bellaterra, Barcelona, Spain, Ph: +34 93 5868936, Fax: +34 (93)5812011, E-mail: Julia.lorenzo@uab.cat

-To whom correspondence should be addressed: David Reverter, Universitat Autonoma de Barcelona, 08193 Bellaterra, Barcelona, Spain, Ph: +34 (93)5868955, Fax: (93)5812011, E-mail: david.reverter@uab.cat

SI Materials and Methods

Protein expression and purification. A clone encoding hCPO (clone ID 8327546; GenBankTM accession number BC112078) was used as a template to generate by PCR a construct encoding soluble hCPOΔC, without the endogenous signal peptide and lacking the C-terminal 25 amino acids (encompassing amino acids Tyr1 to Trp329 of the full-length mature hCPO). This construct was cloned into the expression vector pTriEx-7 (Novagen) between BamHI and NotI restriction sites to incorporate an IgM signal sequence plus an N-t Strep-tag® II fusion protein. hCPO Arg275Asp and Arg275Ala mutants (referred to the mature hCPO numbering and equivalent to Ile255 in bovine CPA1) termed here as hCPOΔC R275D and hCPOΔC R275A, respectively, were generated by PCR-driven overlap extension (1, 2).

hCPOΔC, hCPOΔC R275D and hCPOΔC R275A were produced extracellularly by transient transfection of HEK293 F mammalian cells grown in suspension, similar to the method described before for human carboxypeptidase D (hCPD) (1). Briefly, 500 μg of DNA of each construct were mixed with polyethyleneimine (PEI) (linear 25 kDa, Polysciences Inc.) at a ratio of 1:3 DNA:PEI (w:w) in 50 ml of fresh FreeStyle 293 Expression Medium (Gibco, Thermo Scientific) and incubated for 20 minutes at room temperature (RT). DNA-PEI complexes were then added to the culture at a cell density of $1 \cdot 10^6$ cells/ml and incubated in 500 ml flasks on a rotary shaker (120 rpm) at 37°C, 8% CO_2 and 70% humidity for 9 days. Each culture supernatant containing recombinant hCPOΔC, hCPOΔC R275D or hCPOΔC R275A was recovered by centrifugation at 300 xg for 15 min and filtered through 0.8 μm nitrocellulose membrane filters (PALL). For protein purification, the culture supernatants were equilibrated with 30% ammonium sulfate (Sigma-Aldrich), bound to a hydrophobic chromatography column (TOYOPEARL Butyl-650 M, Tosoh Bioscience) connected to an ÄKTA™ FPLC System (GE Healthcare) and eluted with a decreasing linear gradient (from 30 to 0%) of ammonium sulfate in 50 mM Tris-HCl, pH 7.5. The eluted fractions containing hCPOΔC were pooled and loaded into a 5 ml Strep-Tactin® MacroPrep® column (IBA GmbH) previously equilibrated in 100 mM Tris-HCl, pH 7.5, 150 mM NaCl (buffer W). After washing with 5 column volumes (cv) of

buffer W, the protein was eluted from the resin in 6 cv of the same buffer supplemented with 2.5 mM D-desthibiotin (IBA GmbH). The resultant hCPOΔC, hCPOΔC R275D or hCPOΔC R275A from affinity purification were further purified by size-exclusion chromatography using a Superdex 75 HR 10/30 column (GE Healthcare) equilibrated with 25 mM Tris-HCl buffer, pH 7.5, 150 mM NaCl. Eluted fractions were analyzed by SDS-PAGE and the fractions containing the enzyme were pooled, flash-frozen at approximately 0.2 $mg \cdot ml^{-1}$ and stored at -80 ºC. To perform crystallization experiments, recombinant hCPOΔC was buffer-exchanged to 5 mM Tris-HCl, pH 7.3, 100 mM NaCl and 1 mM β-mercaptoethanol, and concentrated to ~5 $mg.ml^{-1}$ using a centrifugal filter (Amicon Ultra-15 Centricon filter device 10-kDa, Millipore).

NvCI was produced by heterologous expression in *Pichia pastoris* and purified as described in (3), with slight modifications. In brief, purification of NvCI was performed using a combination of two chromatographic methods: an initial cation exchange chromatography (Streamline Direct HST, GE Healthcare) in 100 mM citrate buffer, pH 3.2, and a linear gradient up to 100 mM phosphate buffer, pH 8.0, followed by a second step of size-exclusion chromatography (Superdex 30 High Load, GE Healthcare) in 50 mM Tris-HCl, pH 8.0, 1 M NaCl. The purity of NvCI was determined by MALDI-TOF mass spectrometry and by Tris/Tricine/SDS-PAGE.

Crystallization and data collection. Monoclinic crystals of recombinant hCPOΔC were obtained at 18 °C using the hanging-drop vapor-diffusion method. The reservoir solution contained 100 mM HEPES, pH 7.0, 200 mM ammonium chloride and 20% (w/v) PEG 6000. Single crystals appeared after 20 days with drops mixed from equal volumes of protein solution (about 5 mg/ml of hCPOΔC in 5 mM Tris-HCl, pH 7.3, 100 mM NaCl and 1 mM β-mercaptoethanol, and NvCI in molar ratio 1:0.5 of hCPOΔC:NvCI) and reservoir solution. All crystals were cryoprotected in reservoir buffer containing 15% glycerol and flash-frozen in liquid nitrogen prior to diffraction analysis. Diffraction data were recorded from cryo-cooled crystals (100 K) at the ALBA synchrotron in Barcelona (BL13-XALOC beamline (4). Data

were integrated and merged using XDS software (5), scaled, reduced, and further analyzed using CCP4 (6).

Structure determination and refinement. The structure of hCPOΔC was determined from the X-ray data by molecular replacement using a previously solved structure of hCPA4 in complex with a cleaved hexapeptide (PDB accession code 2PCU) (7) as a reference model. The initial electron density maps produced from molecular replacement programs were manually improved to build up complete models for hCPOΔC using the program COOT (8). Model refinement was performed with Refmac (6) and Phenix (9). The asymmetric unit of the crystal showed two molecules of hCPOΔC, however only one of them is bound to NvCI. Ramachandran analysis showed that 97.6% of residues are in preferred regions, 2.4% of residues are in allowed regions, and 0% of residues are in outlier regions. **Table S2** provides a summary of the refined model parameters. Structural representations were prepared with PyMOL (10).

Kinetic measurements. Carboxypeptidase activity was assayed with the colorimetric substrates 3-(2-furyl)acryloyl-Glu-Glu-OH (FA-EE) (Bachem), 3-(2-furyl)acryloyl-Ala-Lys-OH (FA-AK) (Bachem) and 3-(2-furyl)acryloyl-Phe-Phe-OH (FA-FF) (Bachem), as described (11). In brief, the enzymatic cleavage of the substrate was measured by a decrease in absorbance at 340 nm at 25 ºC in 100 mM Tris-HCl buffer, pH 7.5, 150 mM NaCl. For kinetic analysis, purified hCPOΔC, hCPOΔC R275D and hCPOΔC R275A were assayed with FA-EE, FA-AK and FA-FF at different substrate concentrations ranging from 0 to 1 M. As control, the kinetics of bovine CPB against the substrate FA-AK were determined under the same experimental conditions. For all the enzymes the concentration assayed was 30 nM. Reactions were prepared in triplicate in a final volume of 200 µl using 96-well flat-bottom plates. For all the substrates, the decrease in absorbance was continuously monitored for 10 min using a Victor3 spectrophotometer (Perkin Elmer). Kinetic parameters were determined by fitting the obtained data to the *Michaelis-Menten* equation (Eq. 1) and analyzed by non-linear regression using GraphPad Prism 5.0 (12).

(Eq. 1)

$$y = \frac{V_{max} \text{ x } [S]}{K_M + [S]}$$

Where [S] is the substrate concentration, V_{max} is the maximum enzyme velocity and K_M is the *Michaelis-Menten* constant.

Inhibitory assays and K_i determination. The method for tight-binding inhibitors (13) was used for the K_i determination of PCI, LCI, TCI and ACI. These protein inhibitors were produced by recombinant expression and purified in our group as published elsewhere (14, 15, 16, 17). Inhibitory assays were performed at 25 ºC in equilibrium conditions ($[E_0]/ K_i \leq 10$) with a fixed concentration of enzyme (30 nM) and substrate (500 μM FA-EE) in 100 mM Tris-HCl buffer, pH 7.5, 150 mM NaCl. Samples were pre-incubated for 15 min at RT before substrate addition. The decrease in absorbance was continuously measured at 340 nm using a Victor3 spectrophotometer (Perkin Elmer). K_i values were obtained by direct fitting of fractional activities (v_i/v_0) to the Morrison equation for tight-binding inhibitors (Eq. 2) with a competitive inhibition modality (Eq. 3) using GraphPad Prism 5.0 (12).

(Eq. 2)

$$\frac{V_i}{V_0} = 1 - \frac{([E]_T + [I]_T + K_i^{app}) - \sqrt{([E]_T + [I]_T + K_i^{app})^2 - 4[E]_T[I]_T}}{2[E]_T}$$

(Eq. 3)

$$K_i = \frac{K_i^{app}}{([S_0]/K_M) + 1}$$

Where v_i/v_o is the fractional activity, $[E]_T$ is the total active enzyme concentration, $[I]_T$ is the total inhibitor concentration and K_i^{app} and K_i are the apparent and true inhibition constants, respectively.

The K_i value of NvCI against hCPOΔC was calculated by steady-state kinetics using FA-EE. The assays were performed in 100 mM Tris-HCl buffer, pH 7.5, 150 mM NaCl, with a fixed concentration of enzyme (30 nM) at different substrate concentrations (50, 100, 200, 400, 600, 800 and 1000 μM) in the absence or in the presence of 0.5, 1, 5 and 10 μM of NvCI. Reactions were pre-incubated for 15 min at RT and, after substrate addition, the decrease in absorbance was continuously measured as described above. The *Ki* of NvCI was determined by fitting the initial velocities (v_0) to the following equation (Eq. 4) describing a simple competitive inhibition mechanism.

(Eq. 4)

$$v_0 = \frac{V_{max} \text{ x } [S]}{\left(K_M \left(1 + \frac{[I]}{K_i} \right) + [S] \right)}$$

Where v_0 is the initial velocity, V_{max} is the maximum enzyme velocity, [S] is the substrate concentration, K_M is the *Michaelis-Menten* constant, [I] is the total inhibitor concentration and K_i is the inhibition constant.

Peptide and EGF mass spectrometry experiments. To study the substrate specificity of hCPOΔC, two peptides differing only in the C-t residue (ARLSQKFPKAE and ARLSQKFPKAD) were custom-synthesized (GenScript). Met-enkephalin peptides with C-t Arg, Lys or Phe were purchased from Phoenix Peptides (Phoenix Pharmaceuticals, Inc). Approximately 1.6 μM of each peptide was incubated with 80 nM hCPOΔC in 100 mM Tris-

HCl buffer, pH 7.5, containing 150 mM NaCl, at 37 ºC for 1, 5, 15 and 30 min. After incubation, reactions were stopped by addition of four volumes of 0.1% TFA.

In order to study the C-t cleavage of hEGF by hCPOΔC, 1.5 µg of recombinant human EGF_{1-53} (R&D Systems) was incubated with both 30 nM bCPA (Sigma-Aldrich) and 50 nM pCPB (Sigma-Aldrich) in the absence or in the presence of 30 nM hCPOΔC for 5, 15 and 30 min at 37 °C in 20 mM Tris-HCl buffer, pH 7.5, 250 mM NaCl. After incubation, reactions were stopped by addition of an equal volume of 0.1% TFA.

For MALDI-TOF mass spectrometry analysis, samples were mixed with an equal volume of α-cyano-4-hydroxycinnamic acid (hcca) matrix solution, spotted onto a MTP 384 target plate polished steel T F (Bruker Daltonics) and evaporated to dryness at RT. Mass spectra were acquired on an Ultraflextreme MALDI-TOF mass spectrometer (Bruker Daltonics) operating in reflectron or in linear positive ion mode (for synthetic/Met-enkephalin peptides and for hEGF, respectively) at 25 kV. An external calibration was performed using standard peptide or protein calibration mixtures (Bruker Daltonics).

Miscellaneous

Data deposition: The atomic coordinates and structure factors for the crystal structure of human carboxypeptidase O in complex with NvCI reported in this work have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 5MRV).

References

1. Garcia-Pardo J, et al. (2017) Substrate specificity of human metallocarboxypeptidase D: Comparison of the two active carboxypeptidase domains. PloS one 12(11):e0187778.
2. Heckman KL & Pease LR (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. Nature protocols 2(4):924-932.
3. Covaleda G, del Rivero MA, Chavez MA, Aviles FX, & Reverter D (2012) Crystal Structure of Novel Metallocarboxypeptidase Inhibitor from Marine Mollusk Nerita

versicolor in Complex with Human Carboxypeptidase A4. J Biol Chem 287(12):9250-9258.

4. Juanhuix J, et al. (2014) Developments in optics and performance at BL13-XALOC, the macromolecular crystallography beamline at the ALBA synchrotron. Journal of synchrotron radiation 21(Pt 4):679-689.

5. Kabsch W (2010) Xds. Acta crystallographica. Section D, Biological crystallography 66(Pt 2):125-132.

6. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. Acta crystallographica. Section D, Biological crystallography 67(Pt 4):235-242.

7. Bayes A, et al. (2007) Caught after the act: A human a-type metallocarboxypeptidase in a product complex with a cleaved hexapeptide. Biochemistry 46(23):6921-6930.

8. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta crystallographica. Section D, Biological crystallography 60(Pt 12 Pt 1):2126-2132.

9. Adams PD, et al. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta crystallographica. Section D, Biological crystallography 58(Pt 11):1948-1954.

10. DeLano WL (2002) The PyMol Molecular Graphics System (DeLano Scientific, San Carlos, CA).

11. Lyons PJ & Fricker LD (2011) Carboxypeptidase O Is a Glycosylphosphatidylinositol-anchored Intestinal Peptidase with Acidic Amino Acid Specificity. J Biol Chem 286(45):39023-39032.

12. Motulsky HJ (GraphPad Prism (GraphPad Sofware Inc., San Diego).

13. Morrison JF (1982) The slow-binding and slow, tight-binding inhibition of enzyme-catalysed reactions. Trends in Biochemical Sciences 7(3).

14. Lufrano D, et al. (2015) Biochemical characterization of a novel carboxypeptidase inhibitor from a variety of Andean potatoes. Phytochemistry 120:36-45.

15. Sanglas L, Aviles FX, Huber R, Gomis-Ruth FX, & Arolas JL (2009) Mammalian metallopeptidase inhibition at the defense barrier of Ascaris parasite. Proceedings of the National Academy of Sciences of the United States of America 106(6):1743-1747.

16. Reverter D, et al. (2000) Structure of a novel leech carboxypeptidase inhibitor determined free in solution and in complex with human carboxypeptidase A2. Nature Structural Biology 7(4):322-328.

17. Arolas JL, et al. (2005) The three-dimensional structures of tick carboxypeptidase inhibitor in complex with A/B carboxypeptidases reveal a novel double-headed binding mode. Journal of Molecular Biology 350(3):489-498.

SI Figure Legends

Figure S1. N- and C-terminal hCPO truncations retain full activity. (A) Plasmids expressing full-length hCPO or hCPOΔN (ΔSer24-Ser41, retaining endogenous signal peptide) or hCPOΔC (Met1-Trp349, with endogenous signal peptide) were transiently expressed in HEK293 T cells. Lysates (extracted in 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% (v/v) Triton X-100 and protease inhibitor cocktail lacking EDTA) and media were collected 2 days following transfection with polyethylenimine (PEI) and analyzed by western blotting with the RP3-CPO antibody (Triple Point Biologics). (B) Protein loading was confirmed by staining the membrane with Ponceau S, (C) Protein expression, as determined by western blotting, was quantified using ImageJ. The enzymatic activity of hCPO was determined incubating 1/40 and 1/400 of total lysates and media, respectively, with 500 μM of the synthetic substrate FA-EE in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl. (D) Initial rate of reaction was measured and (E) normalized to hCPO protein quantity, showing that truncations do not result in any significant change in enzymatic activity. A small decrease in hCPOΔC quantity in lysates is balanced by an increase in hCPOΔC in media, as judged by enzymatic activity. Error bars indicate standard deviation(n=3).

Figure S2. Inhibition of hCPOΔC by NvCI. (A) Representative *Michaelis-Menten* and (B) *Lineweaver-Burk* plots for the inhibition of hCPOΔC by NvCI. Increasing concentrations of the substrate FA-EE (ranging from 0 to 1000 μM) were incubated with hCPOΔC in the absence or in the presence of NvCI at indicated concentrations. Assays were performed in 100 mM Tris-HCl buffer, pH 7.5, containing 150 mM NaCl, with an enzyme concentration of 30 nM.

Figure S3. Structural comparison of hCPOΔC with hCPA1, hCPA2 and hCPB. (A) Structural alignment of the catalytic carboxipeptidase domain of hCPO with pancreatic MCPs hCPA1, hCPA2 and hCPB, in ribbon representation. hCPOΔC is colored in blue. hCPA1 and hCPA2 are colored in green and red, respectively. hCPB is shown in grey. The zinc atom is

shown as a yellow sphere. Atomic coordinates of hCPA1 (PDB accession code: 3FJU) (15), hCPA2 (PDB accession code: 1DTD) (16) and hCPB (PDB accession code: 1ZLI) (17) were obtained from the Protein Data Bank (PDB) (www.rcsb.org). Structural models were prepared with PyMOL. (B) Structure-based sequence alignment corresponding to the carboxypeptidase domains of hCPO (UniProt accession code: Q8IVL8), hCPA1 (UniProt accession code: P15085), hCPA2 (UniProt accession code: P48052) and hCPB (UniProt accession code: P15086). Sequences were retrieved from UniprotKB/SWISS-PROT database (www.uniprot.org). Numbering above the sequences corresponds to the amino acid position in mature hCPO, without the signal peptide. Amino acids numbered below the sequences and colored in blue correspond to the catalytic and substrate-binding residues conserved in all M14 MCPs. These residues were numbered according to mature bCPA, which is the standard numbering system for pancreatic MCPs. In this system, number 1 is assigned to the first residue of the mature form after the cleavage of the signal peptide and pro-domain. Residues that shape the S1' specificity pocket are colored in blue. The residue equivalent to I255 of bCPA, that determines the substrate preference of M14-A family of MCPs, is shown in orange, highlighted with an orange arrow below the sequence alignment and numbered. N-glycosylated Asn residues (Asn154, Asn167 and Asn231) in hCPO are colored in green and marked with a green square above the sequence of hCPO. The regular secondary structural elements are depicted above the alignment and labelled, with α-helices shown with empty grey rectangles and β-strands with empty red arrows. Gaps are denoted by dashes (-).

Figure S4. Effect of mutations of Arg275 on the kinetic constants of hCPO. *Michaelis-Menten* plots for (A) hCPOΔC and for hCPO Arg275 single-point mutants (B) hCPOΔC R275D and (C) hCPOΔC R275A. Increasing concentrations of the substrate FA-EE, FA-AK or FA-FF (ranging from 0 to 1000 μM) were incubated with hCPOΔC, hCPOΔC R275D and hCPOΔC R275. Assays were performed in 100 mM Tris-HCl buffer, pH 7.5, containing 150 mM NaCl. The enzyme concentration was 30 nM in all the cases.

Figure S5. Interaction of NvCI and hCPOΔC. (A) Ribbon representation of the NvCI-hCPOΔC complex (upper panel) and NvCI sequence (lower panel) showing the location of the primary (PS) and secondary (SS) binding sites. Residues of NvCI involved in hCPOΔC binding are depicted in red in the ribbon structure and highlighted in the NvCI sequence (in grey and blue for PS and SS residues, respectively). Residues of the C-t tail of NvCI, Ala53, Tyr52 and Cys51, that enter into the active site and contribute to the primary interaction with hCPOΔC are indicated. (B) Close-up view of the C-t residues of NvCI bound to the active site of hCPOΔC. Interactions of the C-t residues of NvCI (Ala53, Tyr52 and Cys51, labelled and shown in orange) with residues in the active site of hCPOΔC (labelled and shown in blue) are indicated. Red dashed lines indicate the interactions (both electrostatic as well as hydrophobic) between residues of the inhibitor and the enzyme. (C) Close-up view of residues of the secondary binding region of NvCI (orange model) to hCPOΔC (blue model). (D) Same as in (C) but with a rotation of the NvCI-hCPOΔC complex of 180º. In all panels, the three disulfide bridges of NvCI and the catalytic zinc atom are shown as yellow sticks or as a yellow sphere, respectively. In panels (C) and (D) the NAG molecule bound to Asn154 is shown as a green stick model with blue nitrogen and red oxygen atoms. (E) Sequence alignment of the C-terminal tail of NvCI with the C-terminal tails of other exogenous proteinaceous carboxypeptidase inhibitors, depicting the P1', P1, P2 and P3 positions of the inhibitor that interact with the S1', S1, S2 and S3 subsites of the enzyme.

Figure S6. Comparison of the interactions of NvCI with hCPA4 and hCPOΔC. (A) Structural superimposition of the NvCI-hCPA4 (PDB accession code: 4A94) (3) and NvCI-hCPOΔC complexes. (B) NvCI sequence indicating those residues of the inhibitor involved in hCPA4 (grey asterisks) or hCPOΔC (blue asterisks) binding. Residues of NvCI within the primary binding site are highlighted in grey. In red are shown those residues from NvCI which interact exclusively with hCPA4. In bold are shown those residues from NvCI which interact with hCPOΔC, but not with hCPA4. (C) Comparison of the primary binding regions for NvCI in hCPA4 and in hCPOΔC. hCPA4 and hCPOΔC, both in complex with NvCI, were overlapped

and presented in grey and blue, respectively. Residues from the C-t tail of the NvCI complexed to hCPA4 (yellow) or to hCPOΔC (orange) are numbered and shown as sticks. Side chains of residues Glu163 and Gln182 from hCPA4 and hCPOΔC, respectively, are labelled. (D) Comparison of the secondary binding regions for NvCI in hCPA4 and in hCPOΔC. Side chains of some selected residues for both enzymes and inhibitor are shown as sticks and labelled with different colors as described in (B). (E) Same as in (C) but with a rotation of the NvCI-hCPOΔC complex of 180º to display the interactions of the opposite site of the structure. See **Table S3** for more details about the molecular interactions presented. In (C), (D) and (E), the catalytic zinc atom is shown as a yellow sphere.

Figure S7. Substrate specificity of hCPOΔC. (A-B) MALDI-TOF MS spectra of the synthetic peptides (A) ARLSQKFPKAE and (B) ARLSQKFPKAD incubated for different times at 37 ºC with 80 nM purified hCPOΔC. Numbers above the major peaks indicate the monoisotopic masses of the MH^+ ion (m/z). The peaks with a mass of 1274.70 Da and 1260.69 Da correspond to the peptides ARLSQKFPKAE (theoretical monoisotopic mass=1273.70 Da) and ARLSQKFPKAD (theoretical monoisotopic mass=1259.68 Da), respectively. The peak with a mass of 1145.67 Da formed in presence of hCPOΔC corresponds to the peptide ARLSQKFPKA (theoretical monoisotopic mass=1144.66 Da) produced by the cleavage of the C-t Glu or Asp amino acids from the peptides ARLSQKFPKAE and ARLSQKFPKAD, respectively. (C-E) MALDI-TOF MS spectra of Met-enkephalin peptides with C-t (C) Phe, (D) Arg or (E) Lys (YGGFMRF, YGGFMKR and YGGFMKK, respectively) after 60 min of incubation in the absence (peptide) or in the presence of 80 nM hCPOΔC (peptide + hCPOΔC). Numbers above the major peaks indicate the monoisotopic masses of the MH+ ion (m/z) for the different Met-enkephalin peptides or for its predicted products generated by hCPOΔC cleavage; YGGFMRF (mass = 877.38 Da), YGGFMKR (mass = 858.41 Da), YGGFMKK (mass = 830.40 Da), YGGFMR (mass = 730.32 Da) and YGGFMK (mass = 702.31 Da).

Figure S8. Docking of model peptides with C-t Glu or Asp amino acids into the catalytic site of hCPO. (A) Close-up view of a C-t Glu (from the peptide PFKAE, in red) docked in the S1' region of the catalytic groove of hCPOΔC. For clarity, only the last two residues from the peptide are shown. Side chains of the most important substrate-binding and catalytic residues of hCPOΔC are labelled and shown as sticks in blue. Interactions between the side chain of Glu in the P1' position and the side chain of Arg275 within the S1' specificity pocket region of hCPOΔC are shown. (B) Same as in (A) but with an Asp from the peptide PFKAD (in green) docked in the S1' specificity pocket of hCPOΔC. Numbers close to dashed lines correspond to the distance measured in Å. The docking of both peptides was performed using the standard-precision docking protocol for peptides from Glide. Graphics were prepared with PyMOL.

SI Tables

Table S1. Kinetic constants determined for FA-EE hydrolysis by hCPOΔC.

	K_M μM	V_{max} μM/s	K_{cat} s^{-1}	K_{cat}/K_M $\mu M^{-1} \cdot s^{-1}$
hCPOΔC	374.1 ± 48.36	0.23 ± 0.01	7.85 ± 0.41	0.021
hCPO-WT	352 ± 49	ND	8.6 ± 0.5	0.026

Enzymatic activity measurements were performed in triplicate. Data are shown as mean ± SE. See **SI Appendix**, **SI Materials** and **Methods** for more details about the equations used for kinetic constant determinations. Data for the kinetic parameters of hCPO-WT were reported in (18). ND, Parameter not determined.

Table S2. Data collection and refinement statistics

Data collection	
Wavelength	0.9792
Space group	C2 (C 1 2 1)
Cell dimensions	
A, b, c (Å)	150.11, 72.14, 90.19
α, β, γ, (°)	90.00, 94.66, 90.00
Resolution (Å)	1.85
R_{merge} [a]	0.069
R_{sym} [b]	0.084
I/σ(I)	13.9
Completeness (%)	98
Redundancy	3.5
Refinement	
Resolution (Å)	1.85
Nº reflections	79.833
R_{work} [c]/R_{free} [d]	18.6 / 20.8
Total number of atoms	5995
Nº aa protein	
hCPOΔC	307
NvCI	53
Water	518
rmsd[e]	
Bond lengths (Å)	0.011
Bond angles (°)	0.96
Ramachandran statistics (%)	
Favored	97.6
Allowed	2.4
Outliers	0

[a] $R_{merge} = \sum|I_i - \langle I\rangle| / \sum I_i$, where I_i is the i^{th} measurement of the intensity of an individual reflection or its symmetry-equivalent reflections and $\langle I\rangle$ is the average intensity of that reflection and its symmetry-equivalent reflections.

[b] $R_{sym} = \sum hkl \sum i\ |Ii - \langle I\rangle| / \sum hkl \sum i\ I_i$

[c] $R_{work} = \sum\big||F_o| - |F_C|\big| / \sum|F_o|$ for all reflections and [d] $R_{free} = \sum\big||F_o| - |F_C|\big| / \sum|F_o|$, calculated based on the 5% of data excluded from refinement.

[e] rmsd, root-mean-square deviation.

Table S3. Interactions between the primary and secondary regions in hCPOΔC and NvCI.

NvCI	hCPO	Distance (Å)
Primary interaction region		
Ala-53 O	Zn^{2+}	2.28
Ala-53 Oτ	Zn^{2+}	2.55
His-88 Nδ1	Zn^{2+}	2.11
Glu-91 Oε1	Zn^{2+}	2.42
Glu-91 Oε2	Zn^{2+}	2.08
His-216 Nδ1	Zn^{2+}	2.16
Ala-53 O	Glu-290 Oε2	2.69
Ala-53 Oτ	Arg-146 Nη2	2.81
Ala-53 Cβ	Glu-290 Oε1	3.57
Ala-53 N	Tyr-268 Oη	2.83
Tyr-52 O	Arg-90Nη2	3.00
Tyr-52 O	Arg-146Nη2	3.36
Tyr-52 O	Phe-299 Cζ	3.24
Tyr-52 Cδ1	Tyr-268 Cζ	3.53
Tyr-52 Cδ2	Thr-183 Cγ1	3.65
Cys-51 N	Arg-90Nη2	3.72
Secondary binding site		
Pro-16 Cβ	Tyr-268 Cδ2	3.58
Leu-17 Cδ1	Tyr-218 Cε1	3.85
Leu-17 Cδ1	Phe-299 Cζ	3.70
Asn-32 Nδ2	Gly-155 O	3.20
Asn-32 Cβ	Asn-154 O	3.74
Asp-33 Oδ1	Asn-154-NAG O7	3.69
Asp-33 Oδ2	Asn-154 Nδ2	3.00
Asp-33 Oδ2	Asn-154-NAG O7	3.13
Phe-34 Cε2	Asn-154 Oδ1	3.90
Phe-34 Cζ	Asn-154-NAG C8	3.62
Phe-34 Cε2	Asn-154 Cβ	3.62
Phe-34 Cδ2	Asn-154 Cβ	3.82
Phe-34 Cδ2	Thr-156 Cγ2	3.89
Tyr-36 O	Thr-156 Cγ2	3.61
Tyr-36 O	Thr-156 Cα	3.54
Glu-37 Cα	Thr-156 O	3.40
Glu-37 C	Thr-156 O	3.57
Cys-38 N	Thr-156 O	2.81
Cys-38 O	Thr-156 O	3.69
Cys-38 O	Cys-157 Cβ	3.80
His-40 N	Gln-180 O	3.52
His-40 O	Gln-182 Nε2	3.20
His-40 Nδ1	Gln-180 O	3.94
His-40 Nε2	Gln-180 Oε1	3.34
Arg-48 O	Leu-144 Cδ1	3.46
Arg-48 Nη2	Asp-142 Oδ1	2.56
Arg-48 Nη2	Asp-142 Oδ2	3.60
Gly-50 Cα	Arg-90 Nη1	3.51
Gly-50 Cα	Phe-299 Cε2	3.85

Figure S1.

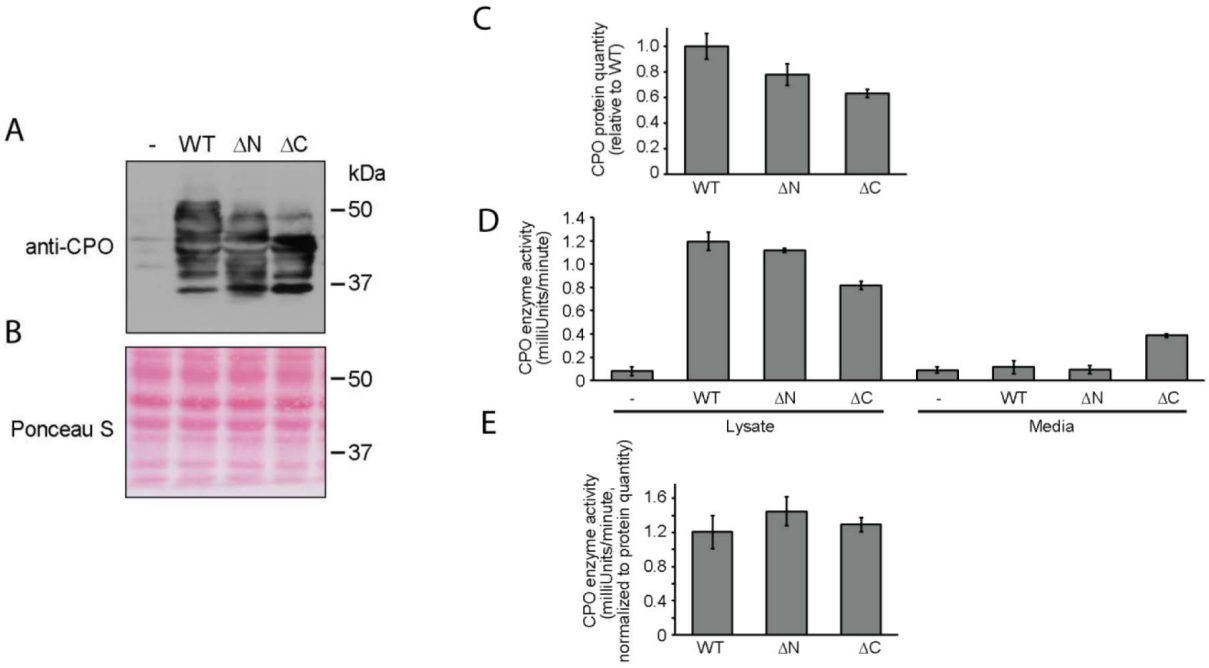

A
- WT ΔN ΔC
kDa
anti-CPO
50
37
B
Ponceau S
50
37
C
CPO protein quantity (relative to WT)
1.0
0.8
0.6
0.4
0.2
0
WT
ΔN
ΔC
D
CPO enzyme activity (milliUnits/minute)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
- WT ΔN ΔC - WT ΔN ΔC
Lysate
Media
E
CPO enzyme activity (milliUnits/minute, normalized to protein quantity)
1.6
1.2
0.8
0.4
0
WT
ΔN
ΔC

Figure S2.

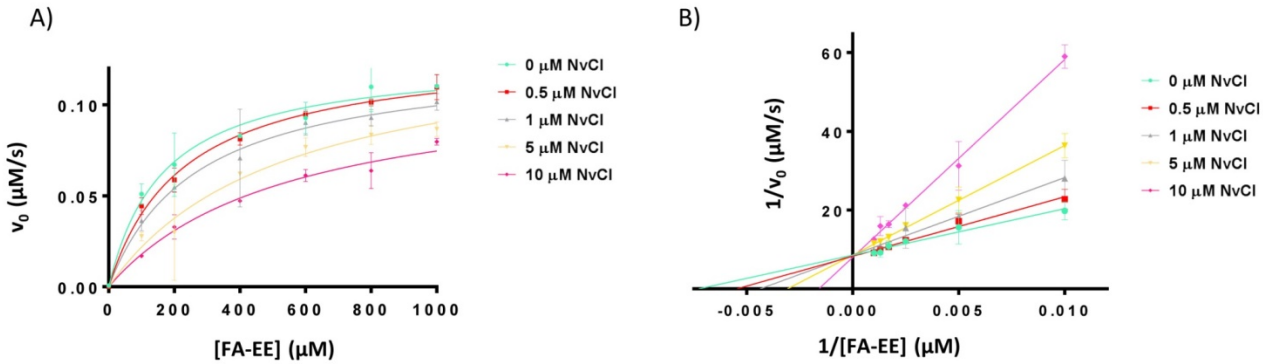

A)
0.10
0.05
0.00
v_0 (μM/s)
0
200
400
600
800
1000
[FA-EE] (μM)
0 μM NvCl
0.5 μM NvCl
1 μM NvCl
5 μM NvCl
10 μM NvCl
B)
60
40
20
$1/v_0$ (μM/s)
-0.005
0.000
0.005
0.010
1/[FA-EE] (μM)
0 μM NvCl
0.5 μM NvCl
1 μM NvCl
5 μM NvCl
10 μM NvCl

Figure S3.

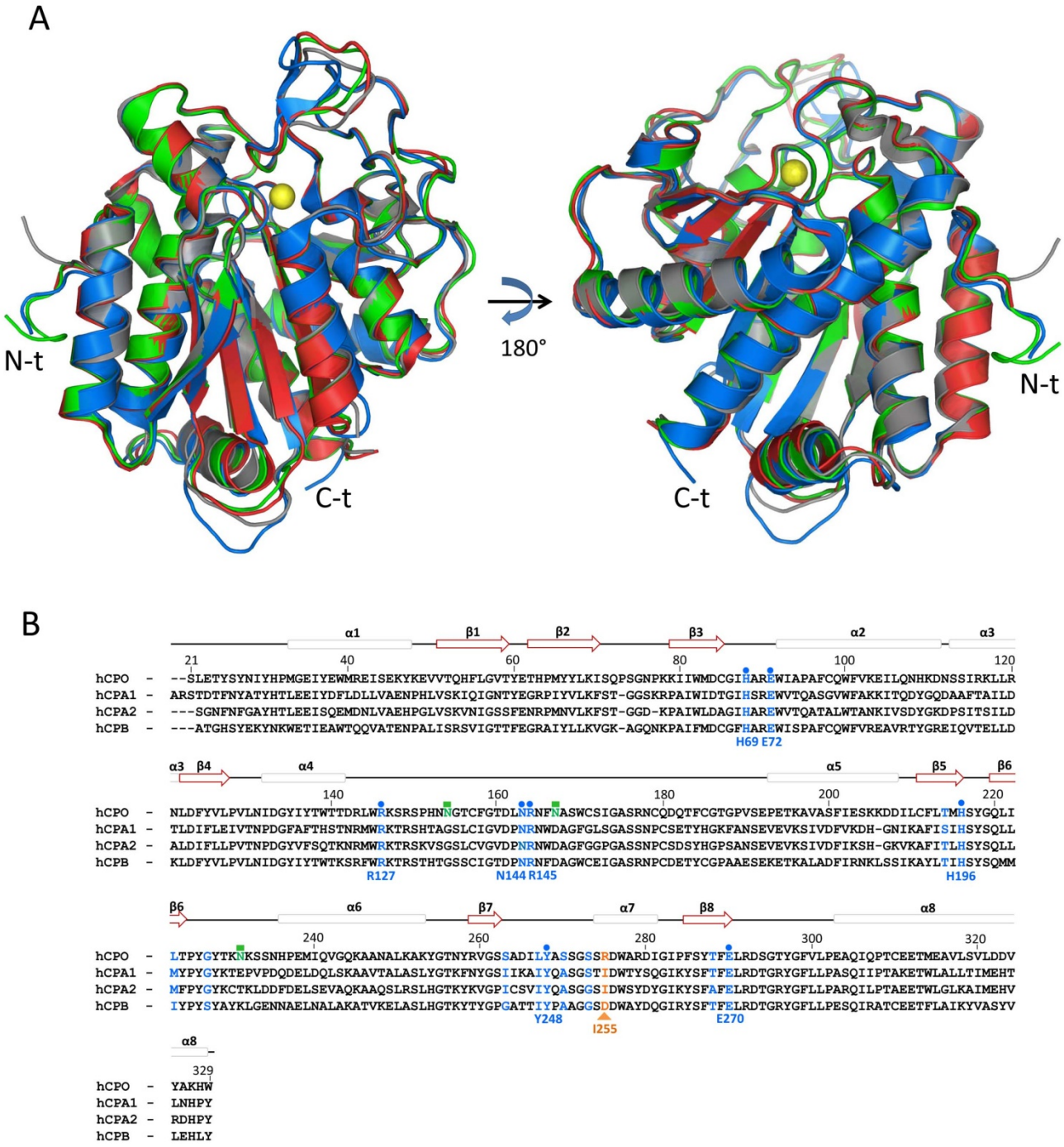

A
N-t
C-t
180°
N-t
C-t
B
α1 β1 β2 β3 α2 α3
21 40 60 80 100 120
hCPO - --SLETYSYNIYHPMGEIYEWMREISEKYKEVVTQHFLGVTYETHPMYYLKISQPSGNPKKIIWMDCGIHAREWIAPAFCQWFVKEILQNHKDNSSIRKLLR
hCPA1 - ARSTDTFNYATYHTLEEIYDFLDLLVAENPHLVSKIQIGNTYEGRPIYVLKFST-GGSKRPAIWIDTGIHSREWVTQASGVWFAKKITQDYGQDAAFTAILD
hCPA2 - ---SGNFNFGAYHTLEEISQEMDNLVAEHPGLVSKVNIGSSFENRPMNVLKFST-GGD-KPAIWLDAGIHAREWVTQATALWTANKIVSDYGKDPSITSILD
hCPB - ---ATGHSYEKYNKWETIEAWTQQVATENPALISRSVIGTTFEGRAIYLLKVGK-AGQNKPAIFMDCGFHAREWISPAFCQWFVREAVRTYGREIQVTELLD
H69 E72
α3 β4 α4 α5 β5 β6
140 160 180 200 220
hCPO - NLDFYVLPVLNIDGYIYTWTTDRLWRKSRSPHNNGTCFGTDLNRNFNASWCSIGASRNCQDQTFCGTGPVSEPETKAVASFIESKKDDILCFLTMHSYGQLI
hCPA1 - TLDIFLEIVTNPDGFAFTHSTNRMWRKTRSHTAGSLCIGVDPNRNWDAGFGLSGASSNPCSETYHGKFANSEVEVKSIVDFVKDH-GNIKAFISIHSYSQLL
hCPA2 - ALDIFLLPVTNPDGYVFSQTKNRMWRKTRSKVSGSLCVGVDPNRNWDAGFGGPGASSNPCSDSYHGPSANSEVEVKSIVDFIKSH-GKVKAFITLHSYSQLL
hCPB - KLDFYVLPVLNIDGYIYTWTKSRFWRKTRSTHTGSSCIGTDPNRNFDAGWCEIGASRNPCDETYCGPAAESEKETKALADFIRNKLSSIKAYLTIHSYSQMM
R127 N144 R145 H196
β6 α6 β7 α7 β8 α8
240 260 280 300 320
hCPO - LTPYGYTKNKSSNHPEMIQVGQKAANALKAKYGTNYRVGSSADILYASSGSSRDWARDIGIPFSYTFELRDSGTYGFVLPEAQIQPTCEETMEAVLSVLDDV
hCPA1 - MYPYGYKTEPVPDQDELDQLSKAAVTALASLYGTKFNYGSIIKAIYQASGSTIDWTYSQGIKYSFTFELRDTGRYGFLLPASQIIPTAKETWLALLTIMEHT
hCPA2 - MFPYGYKCTKLDDFDELSEVAQKAAQSLRSLHGTKYKVGPICSVIYQASGGSIDWSYDYGIKYSFAFELRDTGRYGFLLPARQILPTAEETWLGLKAIMEHV
hCPB - IYPYSYAYKLGENNAELNALAKATVKELASLHGTKYTYGPGATTIYPAAGGSDDWAYDQGIRYSFTFELRDTGRYGFLLPESQIRATCEETFLAIKYVASYV
Y248 I255 E270
α8
329
hCPO - YAKHW
hCPA1 - LNHPY
hCPA2 - RDHPY
hCPB - LEHLY

Figure S4.

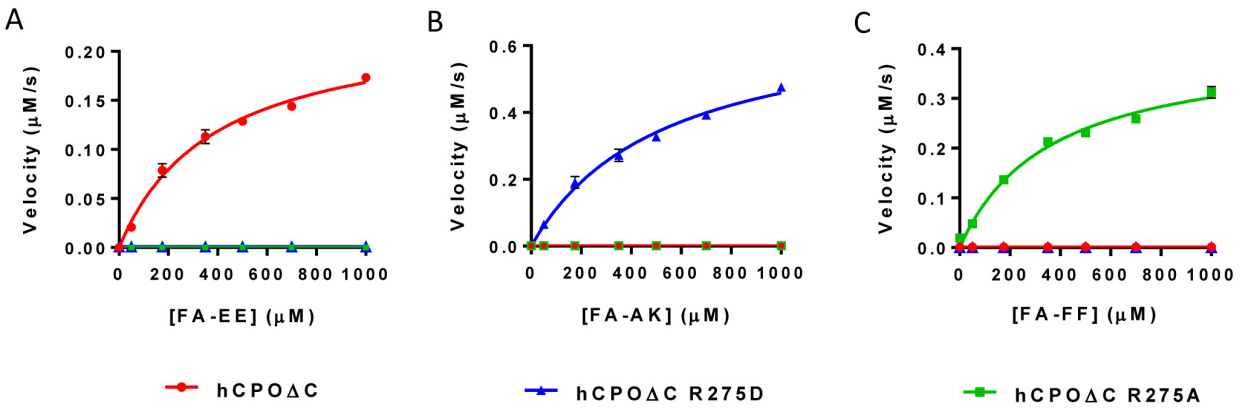
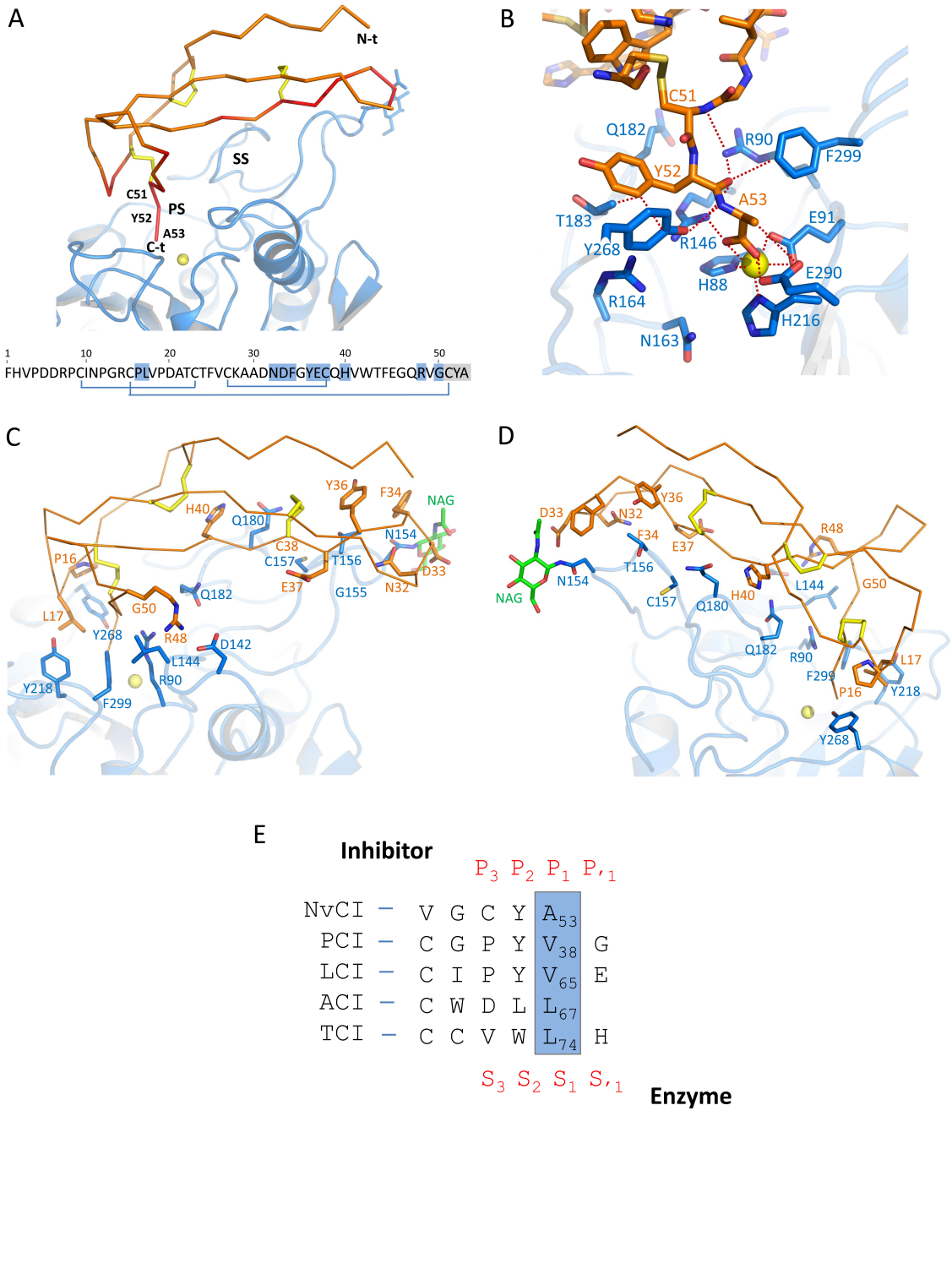

A
Velocity (μM/s)
0.20
0.15
0.10
0.05
0.00
0
200
400
600
800
1000
[FA-EE] (μM)
hCPOΔC
B
Velocity (μM/s)
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
[FA-AK] (μM)
hCPOΔC R275D
C
Velocity (μM/s)
0.4
0.3
0.2
0.1
0.0
0
200
400
600
800
1000
[FA-FF] (μM)
hCPOΔC R275A

Figure S5.



E

Inhibitor

	P_3	P_2	P_1	P'_1	
NvCI	V	G	C	Y	A_{53}
PCI	C	G	P	Y	V_{38} G
LCI	C	I	P	Y	V_{65} E
ACI	C	W	D	L	L_{67}
TCI	C	C	V	W	L_{74} H

S_3 S_2 S_1 S'_1 **Enzyme**

Figure S6.

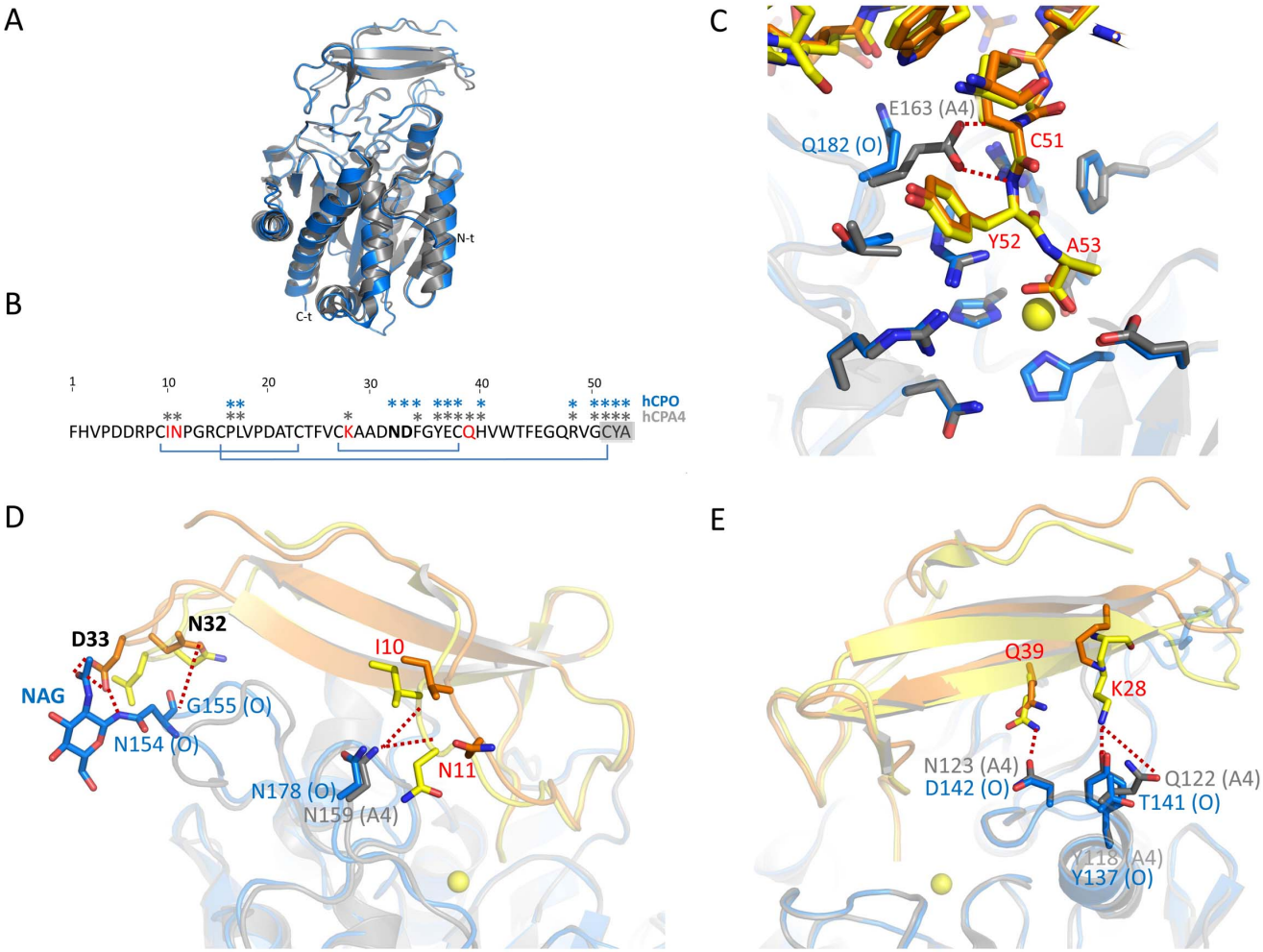

A
N-t
C-t
B
1
10
20
30
40
50
hCPO
hCPA4
FHVPDDRPCINPGRCPLVPDATCTFVCKAADNDFGYECQHVWTFEGQRVGCYA
C
E163 (A4)
Q182 (O)
C51
Y52
A53
D
D33
N32
NAG
G155 (O)
N154 (O)
I10
N11
N178 (O)
N159 (A4)
E
Q39
K28
N123 (A4)
D142 (O)
Q122 (A4)
T141 (O)
Y118 (A4)
Y137 (O)

Figure S7

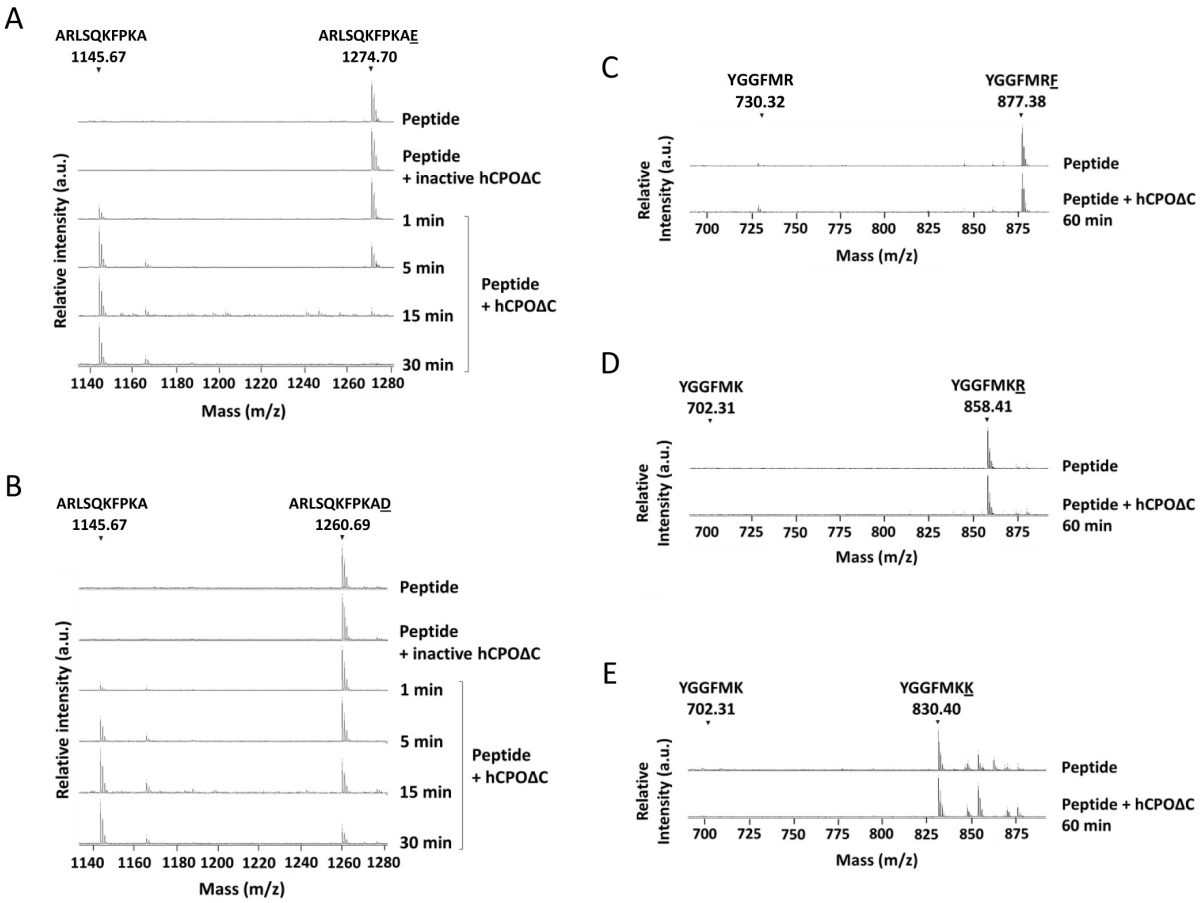
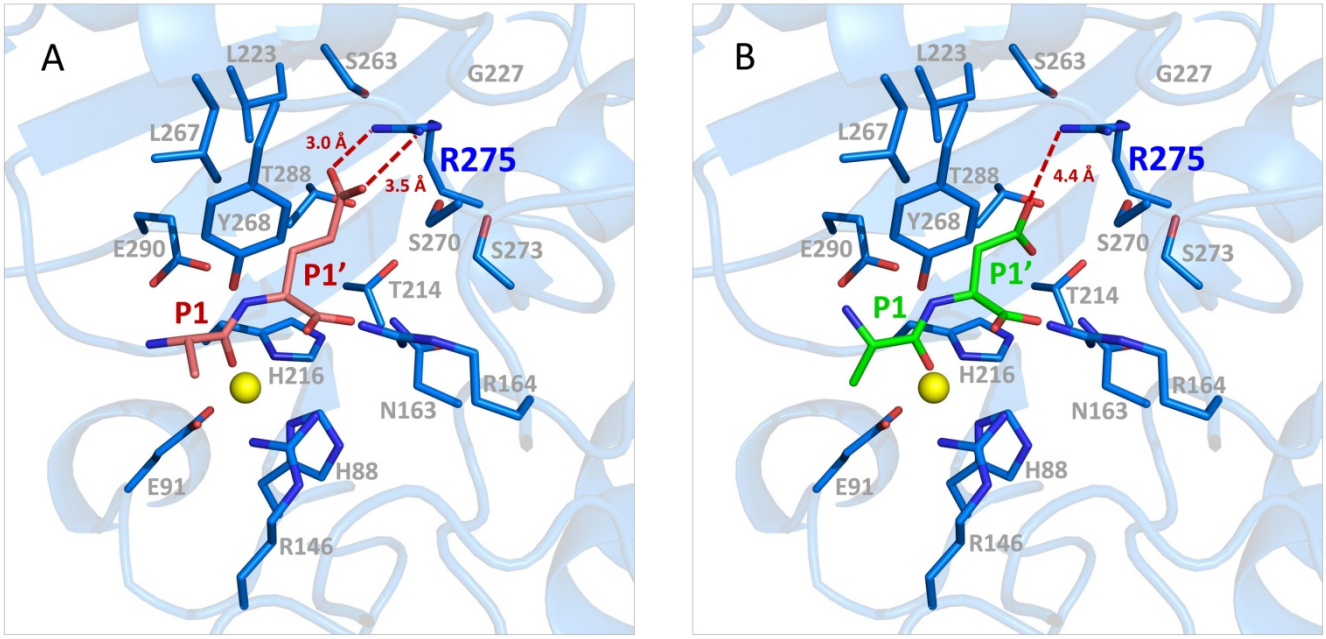

A
ARLSQKFPKA
1145.67
ARLSQKFPKAE
1274.70
Peptide
Peptide
+ inactive hCPOΔC
1 min
5 min
15 min
30 min
Peptide
+ hCPOΔC
Relative intensity (a.u.)
1140 1160 1180 1200 1220 1240 1260 1280
Mass (m/z)
B
ARLSQKFPKA
1145.67
ARLSQKFPKAD
1260.69
Peptide
Peptide
+ inactive hCPOΔC
1 min
5 min
15 min
30 min
Peptide
+ hCPOΔC
Relative intensity (a.u.)
1140 1160 1180 1200 1220 1240 1260 1280
Mass (m/z)
C
YGGFMR
730.32
YGGFMRF
877.38
Peptide
Peptide + hCPOΔC
60 min
Relative Intensity (a.u.)
700 725 750 775 800 825 850 875
Mass (m/z)
D
YGGFMK
702.31
YGGFMKR
858.41
Peptide
Peptide + hCPOΔC
60 min
Relative Intensity (a.u.)
700 725 750 775 800 825 850 875
Mass (m/z)
E
YGGFMK
702.31
YGGFMKK
830.40
Peptide
Peptide + hCPOΔC
60 min
Relative Intensity (a.u.)
700 725 750 775 800 825 850 875
Mass (m/z)

Figure S8.


A
L223
S263
G227
L267
3.0 Å
3.5 Å
R275
T288
Y268
E290
S270
S273
P1'
T214
P1
H216
R164
N163
E91
H88
R146
B
L223
S263
G227
L267
4.4 Å
R275
T288
Y268
E290
S270
S273
P1'
T214
P1
H216
R164
N163
E91
H88
R146