SUPPORTING INFORMATION FOR: Reprogramming Caspase-7 Specificity by Regio-Specific Mutations and Selection Provides Alternate Solutions for Substrate Recognition

Maureen E. Hill,[1¶] Derek J. MacPherson,[1¶] Peng Wu,[2] Olivier Julien,[3] James A. Wells,[3,4] Jeanne A. Hardy*[1]

[¶]These authors contributed equally to this work.

Affiliations

[1]Department of Chemistry, 104 LGRT, 710 N. Pleasant St., University of Massachusetts, Amherst, MA 01003, United States

[2]Present address: Sunnybrook Research Institute, Program in Biological Sciences, University of Toronto, 2075 Bayview Ave, Toronto, ON M4N 3M5, CANADA

[3]Department of Pharmaceutical Chemistry and [4]Department of Cellular & Molecular Pharmacology, 504 Byers Hall,1700 Fourth Street, University of California, San Francisco, CA 94158, United States

*corresponding author: phone (413) 545-3486; fax (413) 545-4490; hardy@chem.umass.edu

Keywords: apoptosis, conformational change, cysteine protease, caspase reporter, GFP, selection, flow cytometry, exosite, substrate binding, specificity, directed evolution, protease/protein engineering

Abbreviations: CHO – aldehyde; Ac - acetyl, AMC – 7-amino-4-methylcoumarin; caspase – casp; esCasp-7 – evolved specificity caspase-7

Methods

Caspase-6 and -7 GFP Reporters

The construction of the CA-GFP family of genetically encoded fluorescent reporters has been described previously.[1,2] The casp-7 reporter C7A-GFP^{e} contains a casp-7 recognizable linker DEVDF. The casp-6 reporter C6A-GFP^{e} contains a casp-6 recognizable linker VEIDG. Both reporters contain the D27E mutation in GFP. This mutation eliminates the intrinsic casp-6 cleavage site in GFP (D27E GFP is named GFP^{e} here).

Caspase-7 Variants

A wild-type human constitutively-two-chain (activated) casp-7 (C7 CTC) construct in pET23b,[3] was used as the DNA template for the library of esCaspases. The derivative casp-7 variants, either from library screening or from Quikchange site-directed mutagenesis (Agilent/Stratagene), were all expressed from this active C7 CTC construct.

Caspase-7 Library (Pre-screening)

Prior to running a full-scale screen, a pre-screening step with FACS gated only by the FITC signal was performed in order to improve efficiency of the sorting routine. For pre-screening, WT C7 CTC served as the DNA template. Saturation mutagenesis at residues Trp232, Ser234 and Gln276 was performed using the QuikChange lightning multi site-directed mutagenesis (Agilent/Stratagene) with degenerate primers: 5'-[phos]GACAG AGTTG CCAGG CACTTT GAGTC TNNST CTGAT GACCC ACA-3' and 5'-[phos]GTTCC AGGCT ATTAC TCGNN SAGGN NSCCA GGAAG AGGCT CCTG-3', where n represents all four bases and s represents a 50:50 mixture of g and c bases. The prescreening library contained greater than 95% coverage of all possible mutational combinations.

Caspase-7 S2 & S4 Library (Screening)

For the casp-7 S2 & S4 library, two degenerate primers, 5'-[phos]ccagg ctatn nstcg nnsag gnnsc cagga agagg ctcc-3' and 5'-[phos]aaccg tggaa taggc gaaga ggaag tc-3' were used to perform saturation mutagenesis on residues Tyr230, Trp232 and Ser234 by the Phusion site-directed mutagenesis (Thermo Scientific). Four DNA libraries were separately constructed with different DNA templates: wild-type C7 CTC (library size >1x10^{5}, >99.7% coverage) and three variants selected from the pre-screening (W232Y/S234C/Q276A, S234C/Q276C and Q276D, size 4x10^{4} each, ~80% coverage each). The four acquired DNA libraries were then combined into the casp-7 S2 & S4 library-A (total size >2x10^{5}) with randomized residues at Tyr230, Trp232 and Ser234, and four residue possibilities - Gln, Ala, Cys or Asp at position 276.

C6A$^{LMNA'}$-GFP^{e} Reporter

To create the C6A$^{LMNA'}$-GFP^{e} reporter two primers, 5'- C AAG GTG GAG ATT GAT AAC GGG AAA CAG TGT AAT GAT TCG TC -3' as the forward primer and 5'- TAC AGC TCG TCC ATG CCG TGA GTG ATC CC -3' as the reverse primer were used to perform Phusion site-directed mutagenesis (Thermo Scientific) at the linker region of C6A-GFP^{e} to change the sequence from VEIDGQGP to VEIDNGKQ as found in lamin A/C.

CA-GFP-based High-Throughput Screening by Fluorescence Activated Cell Sorting (FACS)

For high-throughput screening, 100 ng of plasmid DNA from the casp-7 variant library in pET23b, ampicillin (Amp) resistant, was co-transformed with 150 ng of GFP reporter DNA in pBB75, kanamycin (Kan) resistant. The transformed cells were spread onto about 60 LB plates containing 100 μg/mL Amp and 40 μg/mL Kan. More than 3x10^{5} colonies from these plates were collected and pooled into 50 mL of LB medium with Amp and Kan. Of these pooled cells,

200 µL (>10^8 cells) were inoculated into a 250 mL baffled flask containing 50 mL auto-induction medium[4] with 100 µg/mL Amp and 40 µg/mL Kan. The cells were cultured at 37°C until they reached OD_{600} of 0.9 then incubated overnight at 16°C to produce both the esCasp-7 variants and the C6A-GFP^e reporter. Cells were then harvested, rinsed, re-suspended in PBS buffer and sorted on a FACS Aria II 16-colored cell sorter. In pre-screening, cells were gated by FITC signal only, and sorted twice to improve the purity. In the screening of casp-7 S2 and S4 library-A, sorting was based on both the FSC and FITC channels, two gates (gate p6 and p10) were applied (Supporting Information Figure S7) because we had observed more cell debris and lower cell viability in the high FSC-A fraction. About 1 million events were collected for each gating condition.

On-plate Assay and Selection of the esCasp Variants

Positive cells from cell sorting were collected and diluted with LB medium. Four aliquots, each with about 2500 events, were inoculated onto nitrocellulose membranes. The membranes were first incubated on LB agar plates, containing both Amp and Kan, overnight at 37°C. The nitrocellulose membrane was then transferred onto a LB agar induction plate with 100 µg/mL Amp, 40 µg/mL Kan and 500 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated for 4-6 hrs at room temperature, then for 18-24 hrs at 4°C. The result was recorded using an InGenius (SynGene) gel dock under long wavelength UV and a white illumination picture was also taken as a control. The brightest colonies were selected. Plasmids were extracted by DNA miniprep kit (Qiagen). Those plasmid DNAs, containing both the casp-7 variant and GFP-reporter, were transformed into low competency TAM-1 *E. coli* cells and selected for Amp resistant, but Kan sensitive colonies. The plasmid DNA from such colonies was then extracted and sequenced.

From the prescreening, we obtained three individual variants that showed improved activity toward C6A-GFP^e and reduced activity toward C7A-GFP^e: sort1 (W232Y/S234C/Q276A), sort9 (S234C/Q276C) and sort14 (Q276D). Western blotting using an anti-GFP antibody was used to assess which clones showed increased cleavage of C6A-GFP^e reporters. The correspondence of the increased activities of these variants confirmed the improved proteolytic activity toward C6A-GFP^e in those variants. These variants were taken forward as templates for library construction and additional screening at S2 and S4 subsites. Following sorting in the screening step, we plated the media collected from 10,000 FACS events. We obtained ~2000 colonies from 10,000 events and selected the 14 brightest colonies from the p6 gated population and 15 colonies from p10 gated population. Of these 29 colonies, four individual variant sequences are identified: 26 are of the same variant V1 (Y230A/W232M/ S234N), the other three colonies harbored V2 (Y230V/W232M/Q276C), V3 (Y230V/W232Y/S234V/Q276C) and F1 (Y230D/W232S/ S234V/Q276A). V4 (Y230V/W232Y/S234V/Q276D) was obtained by screening the S2 & S4 library against the LEHDG-GFP reporter, which is also recognized by caspase-6.

To confirm the positivity of the screening, fresh protease reporters (C6A-GFP^e or C7A-GFP) in the pBB75 vector (Kan resistant) were co-transformed with selected esCasp-7 variants or with wild-type active constitutively two-chain versions of casp-7 and casp-6 as controls. Cells were incubated in 2 mL LB culture overnight and 0.5 µL of this culture was inoculated onto a nitrocellulose membrane. The nitrocellulose membrane was incubated and induced as described above. The result was recorded using a G:BOX (SynGene) under Alexa Fluor 488 channel and a white illumination picture was taken as a control (Supporting Information Figure S8).

Fluorescent Protease Activity Assays in *E. coli.*

Selected esCasp-7 variants in the pET23b vector (Amp) and expression constructs for C6A-GFP^e or C7A-GFP in the pBB75 vector (Kan) were transformed simultaneously into the

BL21(DE3) strain of *E .coli.* The wild-type active constitutively two-chain versions of casp-7 and casp-6 were applied as controls. Two mL auto-induction media (AIM) cultures were inoculated from a dense 100 µL overnight LB (Amp/Kan) culture with 100 µg/mL Amp and 40 µg/mL Kan and incubated until reaching an OD_{600} of 0.6-0.9. The cultures were then incubated for 18 hrs at 16°C. Cells from each sample were harvested by centrifugation, rinsed with PBS, resuspended and diluted in PBS to a final OD_{600} of 2.0. The GFP-based fluorescence of 200 µL of this diluted suspension was measured (Ex 475/Em 512) nm in a co-star 96-well black plate on a Molecular Devices Spectramax M5 spectrophotometer.

Western Blotting

After the whole cell fluorescence analysis, samples of *E. coli* cell suspension were analyzed by SDS-PAGE, transferred onto a PVDF membrane (Millipore) and blotted with an anti-GFP monoclonal mouse primary antibody (Millipore). The western blots were then treated with anti-mouse IgG alkaline phosphatase produced in goat (Sigma) and visualized using 1-Step™ NBT/BCIP (Thermo Scientific).

Protein Expression and Purification

esCasp-7 variants in the pET23b vector were transformed into BL21(DE3) *E. coli* cells. The cells were inoculated in 50 mL small-scale cultures of 2xYT with 100 µg/mL Amp (Sigma) at 37°C overnight and transferred to one liter 2xYT cultures with 100 µg/mL of Amp at 37°C until an absorbance of 0.6 at 600 nm was reached. The temperature was lowered to increase expression conditions and protein expression was induced with 1 mM IPTG. WT casp-7, esCasp-7 V1, V2 and V3 expressed soluble protein after 18 hours at 18°C whereas esCasp-7 V4 expression was optimal after 3 hours at 25°C. The cells were harvested by centrifugation at 5,000 rpm and the pellets were stored at -80°C. The pellets were thawed at 37°C before being lysed in a microfluidizer (Microfluidics, Inc.) in a buffer consisting of 50 mM sodium phosphate, pH 8.0, 300 mM NaCl and 2 mM imidazole. Cellular components were removed from the lysate by centrifugation at 27,000 rcf for 1 hour. A 5-mL HiTrap nickel-affinity column (GE Healthcare) was loaded with the supernatant and washed with a buffer of 50 mM sodium phosphate, pH 8.0, 300 mM NaCl and 50 mM imidazole, then eluted with a buffer containing 50 mM sodium phosphate, pH 8.0, 300 mM NaCl and 300 mM imidazole. The protein collected from the elution was diluted 6-fold into a buffer of 20 mM Tris, pH 8.0 and 2 mM DTT to reduce the salt concentration and loaded onto a 5 mL HiTrap Q column (GE Healthcare). A NaCl gradient was used to elute casp-7 and variants in a buffer containing 20 mM Tris, pH 8.0, 200 mM NaCl and 2 mM DTT and was stored at -80°C in the above conditions. Fractions containing protein were analyzed by SDS-PAGE to ensure casp-7 identity and assess purity. Procaspase-3 and WT casp-6 were purified as previously described.[5] C7A-GFP was purified as previously described.[6]

The lamin C gene in the pDNR-Dual vector was transformed into BL21(DE3) *E. coli* cells. 50 mL cultures were grown at 37°C overnight in 2xYT with 100 µg/mL Amp (Sigma), transferred to a one-liter 2xYT culture with 100 µg/mL Amp at 37°C until an absorbance of 0.6 at 600 nm was reached. The temperature was lowered to 18°C and protein expression was induced with 0.1 mM IPTG. Soluble lamin C was produced after 18 hours at 18°C. The cells were harvested by centrifugation at 5,000 rcf. The collected pellet was stored at -80°C until use. The pellet was thawed at 37°C before being lysed in a microfluidizer (Microfluidics, Inc.) in a buffer consisting of 20 mM sodium phosphate, pH 8.0, 100 mM NaCl, 2mM imidazole. Cellular components were removed from the lysate by centrifugation at 27,000 rcf for 1 hour. A 5-mL HiTrap nickel-affinity column (GE Healthcare) was loaded with the supernatant and eluted using an imidazole gradient in a buffer containing 20 mM sodium phosphate, 100 mM NaCl, and 1 M imidazole. The eluted protein was collected and diluted 6-fold to reduce the salt concentration into a buffer containing 20 mM Tris, pH 8.0, 2 mM EDTA and 2 mM DTT and loaded onto a 5 mL HiTrap Q column (GE Healthcare). A linear NaCl gradient was using 20 mM Tris, pH 8.0, 2 mM EDTA, 1

M NaCl, 2 mM DTT resulted in separation of the protein. Fractions containing lamin C were analyzed by SDS-PAGE to confirm identity and assess purity. Fractions containing lamin C were collected and concentrated to 5 mg/mL in an Amicon Ultra-15 10,000 NMWL regenerated cellulose spin column (EMD Millipore). The concentrated protein was loaded on to a HiLoad 26/600 Superdex 200 pg column (GE Healthcare, Inc.) and eluted at 2 mL/min in a buffer containing 20 mM Tris, pH 8.0, 100 mM NaCl, 2 mM EDTA and 2 mM DTT. SDS-PAGE identified fractions containing lamin C and assessed purity. Protein was eluted after 90 minutes (180 mL) and stored at -80°C in the above conditions.

Caspase Activity Assays

To measure caspase activity, WT casp-7 and esCasp variants were assayed over 7 minutes at 37°C in casp-7 activity assay buffer consisting of 100 mM HEPES, pH 7.5, 5 mM $CaCl_2$, 10% PEG 400, 0.1% CHAPS and 5 mM DTT. Casp-6 was tested in casp-6 activity assay buffer consisting of 100 mM HEPES, pH 7.5, 100 mM NaCl, 10% sucrose, 0.1% CHAPS and 5 mM DTT. Substrate titrations were initiated by the addition of 0-400 µM fluorogenic substrates *N*-acetyl-Asp-Glu-Val-Asp-7-amino-4-methylcoumarin (Ac-DEVD-AMC) or *N*-acetyl-Val-Glu-Ile-Asp-7-amino-4-methylcoumarin (Ac-VEID-AMC) (Enzo Life Sciences) to 100 nM of caspase. Non-optimal substrates including *N*-acetyl-Tyr-Val-Ala-Asp-7-amino-4-methylcoumarin (Ac-YVAD-AMC), *N*-acetyl-Trp-Glu-His-Asp-7-amino-4-trifluoromethylcoumarin (Ac-WEHD-AFC), *N*-acetyl-Leu-Glu-Thr-Asp-7-amino-4-trifluoromethylcoumarin (Ac-LETD-AFC) and *N*-acetyl-Leu-Glu-His-Asp-7-amino-4-trifluoromethylcoumarin (Ac-LEHD-AFC) (Enzo Life Sciences) were assayed from 0 to 1 mM. Fluorescence of the cleaved coumarin was measured in a Spectramax M5 spectrometer (Molecular Devices) at Ex 365/Em 495 nm for AMC and Ex 380/Em 505 for AFC. Kinetic assays were run in duplicate using 100 µL reaction volumes in a 96-well microtiter. Prism (GraphPad Software) was used to fit initial velocities versus substrate concentration to determine kinetic values K_M and k_{cat}.

Caspase concentrations were determined using an active site titration to accurately determine the active caspase concentration. Active site titrations of wild-type casp-7, esCasp V1 and esCasp V2 were performed with the covalent inhibitor Ac-DEVD-CHO (*N*-acetyl-Asp-Glu-Val-Asp-aldehyde, Enzo Life Sciences) because these enzymes prefer DEVD over VEID. Enzyme concentrations of WT casp-6, esCasp V3 and esCasp V4 were determined using VEID-CHO (*N*-acetyl-Val-Glu-Ile-Asp-aldehyde, Enzo Life Sciences) because the enzymes prefer VEID. Protein was added to inhibitor solvated in DMSO in 96-well V-bottom plates at room temperature for 1 hr in casp-7 activity assay buffer. After labeling, 90 µL was transferred to a black well plate in duplicate and assayed with 50-fold molar excess of DEVD-AMC or VEID-AMC substrate. The concentration at which full inhibition was observed was considered to be equivalent to the total enzyme concentration, and was thus used for k_{cat} calculations. Estimation of enzyme concentration by the active site titration method sometimes produced estimates of protein concentration that differed from the concentrations calculated based on absorbance at 280 nm. These differences were never observed to be greater than 2-fold; thus any errors in k_{cat} reported should be less than 2-fold.

Inhibition Rate Constant k_2 Assessment

Progress curve method as described by Tian and Tsou[7,8] monitors the rate of substrate hydrolysis in the presence of fast, irreversible covalent inhibitors. We have found Ac-DEVD-CHO and Ac-VEID-CHO to behave irreversibly at the DTT concentrations used. Caspases were diluted to 125 nM in their respective activity buffers (see Caspase Activity Assays). 80 µL of 125 nM caspase solution was mixed with 10 µL of Ac-DEVD-CHO, (final inhibitor concentration, 200 nM) and 10 µL of Ac-DEVD-AMC (final substrate concentration of 100 µM), in a 96-well black bottom plate yielding a final protein concentration of 100 nM. The same process was repeated utilizing Ac-VEID-CHO and Ac-VEID-AMC. Casp-7 and esCasp-7 variants were assayed in

casp-7 activity assay buffer; casp-6 was assayed in casp-6 activity assay buffer. Reactions were run in duplicate and monitored at Ex. 365/Em. 495 nm once a minute for 1 hour at 37°C. Data from the first 20 minutes was utilized for the determination of the rate constant k_2.

Progress of inhibition is described by the equation: $[P]=[P_\infty]\ (1-e^{-A[I]t})$, where [P] and $[P_\infty]$ are the product concentrations measured at times t and t=∞ and A is the apparent rate of formation of the inhibited enzyme complex. The value for the apparent rate constant A is determined by the slope of the plot of $\ln([P_\infty]-[P])$ versus time (t) where: A=slope/[I], where [I] is the concentration of inhibitor . By assuming irreversible covalent modification, the rate constant k_2 can be derived as a second order equation as follows: $k_2 = A(1+[S]/K_M)$, where [S] is the concentration of substrate.

Structures of Caspase-7 Variants

Purified esCasp variants in a buffer of 200 mM NaCl, 20 mM Tris and 2 mM DTT were incubated on ice with three molar excess of either Ac-DEVD-CHO or Ac-VEID-CHO for 3 hours. Inhibited protein complexes were concentrated to 4-8 mg/mL using Amicon Ultrafree 3K NMWL Membrane Concentrators (Millipore). Crystals grew in 1-3 µL hanging drops by varying the ratio of protein to mother liquor from 0.5-4:1 respectively. Crystals were obtained in buffers containing 300 mM diammonium citrate, pH 5.0-6.0, 14% PEG 3350, 10 mM guanidine hydrochloride and 10 mM DTT within 5 days at 4°C by vapor diffusion. Crystals of WT casp-7, esCasp-7 V1 and esCasp-7 V4 with VEID were harvested and cryoprotected for 30-60 seconds in an identical buffer containing 20% glycerol and quickly frozen in liquid nitrogen. Crystals of esCasp-7 V2 and esCasp-7 V4 with VEID were grown in mother liquor composed of 300 mM diammonium citrate, pH 5.0-6.0, 14% PEG 3350, 10 mM guanidine hydrochloride and 20% glycerol and quickly frozen in liquid nitrogen. Data sets for WT casp-7:DEVD, esCasp-7 V1:DEVD or VEID, and esCasp-7 V4:VEID were collected at Brookhaven National Laboratories National Synchrotron X6A beamline (Upton, NY). Data sets for esCasp-7 V2:DEVD or VEID, and esCasp-7 V4:DEVD were collected at Argonne National Laboratory 24-ID-E beamline (Chicago, Il).

Crystallographic phases were generated by molecular replacement using 3EDR as a search model for PhaserMR,[9] part of the CCP4 software suite.[10] Substrates were modeled into density attached to the catalytic cysteine. Structures were refined using iterative rounds of rebuilding in Coot[11] and refinement in Refmac5[12] using individual B-factors. Water molecules were modeled into the structures, verified with refinement and checked by stereo-chemical viability. A final round of refinement with NCS restraints imposed was performed in Phenix.[13] Structural analysis of the complexes was accomplished by mapping both hydrophobic and polar interactions of each atom in the DEVD or VEID substrates within 4Å by manual inspection and in conjunction with LigPlot analysis.[14]

Protein Cleavage Assays

Protein cleavage assays were used to monitor the ability of caspases to cleave folded protein substrates. Procaspase-3 C163A is inactive due to replacement of the catalytic cysteine and was chosen as a casp-6 substrate to demonstrate casp-6-like specificity. Cleavage of 3 µM procaspase-3 by 150 nM WT casp-7, casp-6 or esCasp-7 variants was assessed in their respective activity assay buffers at 37°C. Samples were taken at various time-points and analyzed by 16% SDS-PAGE. C7A-GFP^e, a Caspase Activatable-GFP reporter with a DEVD linker and D27E mutation, was used to assess casp-7-like activity. C7A-GFP^e was transferred to a buffer containing into 100 mM HEPES, pH 7.5, 5 mM $CaCl_2$, 10% PEG, 0.1% CHAPS and 5 mM DTT using a NAP-5 column (GE Healthcare) and incubated with 300 nM caspase at 37°C. Samples were taken at various time points and assessed by 12% SDS-PAGE. Proteins were transferred to a PVDF membrane (Millipore) and blotted with a monoclonal mouse primary antibody against GFP (Millipore). Western blots were incubated with peroxidase-conjugated

Affinipure goat anti-mouse IgG (Jackson ImmunoResearch) and visualized using TMB stabilized substrate for horseradish peroxidase (Promega).

Protein cleavage assays were used to monitor the ability of caspases to cleave the casp-6 substrate lamin C. An enzyme titration was performed to calculate the hydrolysis rate of lamin C as a substrate of caspases -6, -7 and esCasp-7 V4. A 2-fold dilution series (50 μM-2 nM) of the various caspases were made in a buffer containing 50 mM HEPES, pH 7.5, 100 mM NaCl, 0.5% CHAPS, 1 mM EDTA and 10 mM DTT. The lamin C added to a final concentration of 2.5 μM, vortexed and incubated at 37°C for 4 hours to allow the reaction to come to completion. Samples were electrophoresed in a 16% SDS-PAGE. Coomassie-stained gels were quantified using a Bio-Rad ChemiDoc MP imaging system (Bio-Rad Laboratories, Inc.) to assess the cleavage of the substrate lamin C as a function of time. Quantification was fit to a single phase decay using GraphPad Prism (GraphPad, Inc.) to find the CF_{50}, which is the concentration of enzyme (casp-6, casp-7, esCasp-7 V4) where 1/2 of the protein substrate (lamin C) was cleaved. The enzyme concentration at CF_{50} was used to calculate the hydrolysis rate from the equation $k = (-\ln(P))/(E*t)$ where k is the hydrolysis rate, P is the portion of the fraction cleaved (CF_{50}), E is the enzyme concentration (M) at which CF_{50} is achieved and t is time (s).

C6A-GFP^{e} and C6A$^{LMNA'}$-GFP^{e} cleavage assays were used to monitor the ability of caspases to cleave the P4-P4' sequence of lamin C. Cleavage of 3 μM C6A-GFP^{e} or C6A$^{LMNA'}$-GFP^{e} by 500 nM casp-6, casp-7 or esCasp-7 V4 was assessed in a buffer containing 50 mM HEPES, pH 7.5, 100 mM NaCl, 0.5% CHAPS, 1 mM EDTA and 10 mM DTT at 37°C. Samples were taken at various time-points and analyzed by 16% SDS-PAGE. Coomassie-stained gels were quantified using a Bio-Rad ChemiDoc MP imaging system (Bio-Rad Laboratories, Inc.) to assess the cleavage of the substrates as a function of time.

Circular Dichroism Analysis

WT Caspase-7 and esCasp variants were prepared by a 1-hour incubation on ice with or without 3-molar excess of either DEVD-CHO or VEID-CHO at 10-20 μM in 20 mM Tris, pH 8.0, 200 mM NaCl and 2 mM DTT. To verify active site occupancy, the proteins were assayed with 50-molar excess DEVD-AMC. Inhibition greater than 95% indicated that the active site was fully occupied with the covalent active site inhibitor. Liganded and unbound proteins were then buffer exchanged using a NAP-5 column (GE Healthcare) into 10 mM phosphate buffer, pH 7.5 with 120 mM NaCl and protein concentrations were measured using absorbance at 280 nm (Nanodrop 2000C spectrophotometer).

Thermal denaturation of 5-12 μM WT casp-7 and esCasp variants was observed by a loss of circular dichroism signal at 222 nm over the range from 20-90°C on a J-720 CD spectrometer (Jasco). Spectra from 200-250 nm were collected before and after thermal denaturation. Samples were run in duplicate on separate days. The melting curves were plotted using GraphPad Prism (GraphPad Software) using a Boltzman-Sigmoidal fit to determine the T_m.

Assessment of esCasp V4 Activity by *N*-terminomics

For each discovery replicate, $1x10^9$ Jurkat cells were harvested by centrifugation and lysed (0.1% Triton x-100) in the presence of protease inhibitors (5 mM EDTA, 1 mM 4-(2-Aminoethyl)benzenesulfonyl fluoride (AEBSF, Sigma), 1 mM phenylmethylsulfonyl fluoride (PMSF, Sigma), 10 mM 2-Iodoacetamide (IAM, Sigma)). Excess cysteine-protease inhibitors were neutralized with 20 mM DTT prior to addition of 500 nM of purified caspase and incubated for 4 hours (see Supporting Information Figure S5 for *N*-terminomics workflow). Exogenous caspase activity was then neutralized with 100 uM zVAD-fmk (VWR). *N*-termini labeling was performed for two hours with 1 μM subtiligase and 1 mM TEVTest4B probe.[15] Tagged protein fragments were precipitated using acetonitrile, then denatured (8 M Gdn-HCl), reduced (2 mM TCEP) and alkylated (4 mM IAM), before ethanol precipitation. Biotinylated N-termini peptides

were then captured with High Capacity Neutravidin Agarose Resin (Life Technologies) for 48 hours. The beads were washed using 4 M Gdn-HCl, trypsinized, and released from the beads using TEV protease. The tryptic peptides were then desalted and concentrated with a C_{18} Ziptip (Millipore), then fractionated into 11 fractions/sample using high pH reverse phase chromatography.

Discovery Mass Spectrometry

LC-MS/MS was carried out by reverse phase LC interfaced with a LTQ Orbitrap Velos (ThermoFisher Scientific) mass spectrometer. A nanoflow HPLC (NanoAcquity UPLC system, Waters Corporation) was equipped with an Easy-Spray C18 column ES800 (Thermo). Peptides were eluted over a linear gradient over 60 min from 2% to 30% acetonitrile in 0.1% formic acid. MS and MS/MS spectra were acquired in a data-dependent mode with up to 6 HCD MS/MS spectra acquired for the most intense parent ions per MS. For data analysis, peptide sequences were assigned using the ProteinProspector (v5.13.2) database search engine (http://prospector.ucsf.edu/prospector/mshome.htm) against the Swiss-Prot human protein database (2015.3.5). Search parameters included a precursor mass tolerance of 20 ppm, fragment ion mass tolerance of 20 ppm, up to 2 missed trypsin cleavages, constant carbamidomethylation of Cys, variable modifications of *N*-terminal addition of non-standard aminobutyric acid residue (Abu), acetylation of protein N-terminus, and oxidation of methionine. The identified peptides were searched against a random decoy protein database for evaluating the false positive rates. The false discovery rate is below 1.0% at the maximum expectation value of peptide set at 0.05.

Accession Codes

The following complexes are assigned the PDB ID listed: Wild Type caspase-7 bound to VEID is 4ZVU; esCasp-7 V1 bound to DEVD is 4ZVS; esCasp-7 V1 bound to VEID is 4ZVT, esCasp-7 V2 bound to DEVD is 4ZVP; esCasp-7 V2 bound to VEID 4ZVQ; esCasp-7 V4 bound to DEVD 4ZVR; esCasp-7 V4 bound to VEID is 4ZVO.

Supporting Information Files

Supporting Information File 1. This PyMol Session File can be downloaded from http://people.chem.umass.edu/jhardy/Figures/esCasp_Superpositions.pse . This file contains the three dimensional crystal structures of casp-7 and the esCasp-7 Variants bound to the canonical substrate DEVD (green) or VEID (yellow). Structures are superimposed base substrate coordinates to highlight differences adopted by the variants upon binding to substrate. Selections are labeled by PDB ID, enzyme name and substrate. Scene 1 is a superposition of all DEVD bound structures; scene 2, all VEID bound structures; Scene 3, WT casp-7 bound to DEVD or VEID; Scene 4, esCasp-7 V1 bound to DEVD or VEID; Scene 5; esCasp-7 V2 bound to DEVD or VEID; Scene 6, esCasp-7 V4 bound to DEVD or VEID. Lastly, scene 7 returns to the original centered position.

Supporting Information Files S2 and S3. These files contain a full list of all of all of the peptides, cleavage sites and substrates identified in our studies on WT casp-7 (cb5b00971_si_002.xlsx) and esCasp-7 V4 (cb5b00971_si_003.xlsx). The peptides for WT casp-6 are being published separately.[16]

Supporting Information File S4. This file contains a list of all of the common peptides compared between the variant caspases (cb5b00971_si_004.xlsx).

***N*-terminomics Data Repository.** The mass spectrometry data obtained during the *N*-

terminomics analysis is hosted at: http://prospector2.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer . For each of the caspases two independent biological replicates were analyzed and are noted as A or B.

wt-Casp7A: http://prospector2.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key= zmg7d0g9k1&search_name=msviewer

wt-Casp7B: http://prospector2.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=i7c8zdqfpx&search_name=msviewer

esCasp7-V4B: http://prospector2.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=x2nfjcdwdd&search_name=msviewer

esCasp7-V4B: http://prospector2.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=83rw5ksbsf&search_name=msviewer

Table S1. Crystallographic data collection and refinement statistics for esCasp-7 variants with VEID or DEVD substrates.

	Wild Type Caspase-7: VEID	esCasp-7 V1: DEVD	esCasp-7 V1: VEID	esCasp-7 V2: DEVD	esCasp-7 V2: VEID	esCasp-7 V4: DEVD	esCasp-7 V4: VEID
PDB ID	**4ZVU**	**4ZVS**	**4ZVT**	**4ZVP**	**4ZVQ**	**4ZVR**	**4ZVO**
Data collection statistics							
Wavelength (Å)	0.954	0.954	1.54	0.979	0.979	0.979	0.954
Diffraction Resolution (Å)	36.0 – 2.60 (2.64–2.60)	48.4-2.50 (2.54-2.50)	48.3 – 2.85 (3.00-2.85)	93.8-2.50 (2.60-2.50)	93.9-2.50 (2.60-2.50)	76.2-2.3 (2.38-2.30)	50.0-2.90 (2.95-2.90)
Measured Reflections (n)	200486	226296	103000	217941	196222	282011	91952
Unique Reflections	25303	28395	20141	29117	29209	38059	20424
Completeness (%)	99.8 (98.0)	99.7 (97.9)	98.8 (98.0)	100.0 (100.0%)	100.0 (100.0%)	99.9 (99.7%)	99.8 (100%)
Redundancy	7.5	4.0	5.1	7.5	6.7	7.4	4.5
I/σ(I)	24.0 (2.14)	14.3 (2.8)	8.9 (1.9)	8.4 (3.3)	12.4 (3.4)	12.3 (2.8)	15.7 (2.2)
R_{sym} (%)	12.6 (<100)	11.3 (84.7%)	12.2 (69.6%)	14.2 (51.6)	9.7 (53.0)	10.4 (79.4)	11.1 (59.5)
R_{pim}	N/A	0.065 (.458)	0.076 (0.430)	0.082 (0.298)	0.058 (0.320)	0.061 (0.459)	N/A
Space Group	P 3_2 2 1	P 3_2 2 1	P 3_2 2 1	P 3_2 2 1	P 3_2 2 1	P 3_2 2 1	P 3_2 2 1
a (Å)	88.3	88.1	88.4	86.8	86.9	87.9	88.2
b (Å)	88.3	88.1	88.4	86.8	86.9	87.9	88.2
c (Å)	187.0	187.2	186.5	187.6	187.7	187.3	187.2
α = β (°)	90	90	90	90	90	90	90
γ (°)	120	120	120	120	120	120	120
Refinement statistics							
No. of atoms	3834	3900	3785	3851	3874	3921	3796
No. waters	31	111	4	76	99	130	7
R_{work}/R_{free} (%)	20.5 / 25.2	20.73 / 23.91	19.98 / 23.44	16.66/22.39	17.53 / 20.79	18.29 / 22.39	20.92 / 25.13
RMSD bond length (Å)	0.003	0.002	0.004	0.007	0.003	0.009	0.005
RMSD bond angle (°)	0.84	0.70	0.90	1.13	0.87	1.32	1.18
Average B-factor (Å^2)	61.3	48.7	52.2	45.5	41.5	43.2	65.5
B-factor (Protein)	61.1	48.7	52.0	45.5	41.5	43.0	65.4
B-factor (Inhibitor)	70.7	48.4	62.2	50.5	48.4	49.6	70.2
B-factor (Water)	60.3	39.4	37.4	41.6	39.4	41.1	59.5
Ramachandran plot							
Core (%)	98.1	98.5	96.3	98.3	98.5	97.6	94.8
Allowed (%)	1.9	1.5	3.7	1.5	1.5	2.4	4.3
Disallowed (%)	0	0	0	0.2	0	0	0.9

Values in parentheses are for the highest resolution bin

DEVD and VEID represent the Ac-DEVD-CHO or Ac-VEID-CHO tetrapeptide inhibitors

A

Enzyme	Substrate	K_M (μM)	k_{cat} (s^{-1})	10^6 x k_{cat}/K_M (s^{-1}/M)
Casp-6[a]	VEID	38 ± 4.3	0.7 ± 0.02	0.02
Casp-7[b]	DEVD	42 ± 11	2.4 ± 0.07	0.06
dsCasp-7	DEVD	>200	ND	ND
	VEID	>200	ND	ND

B

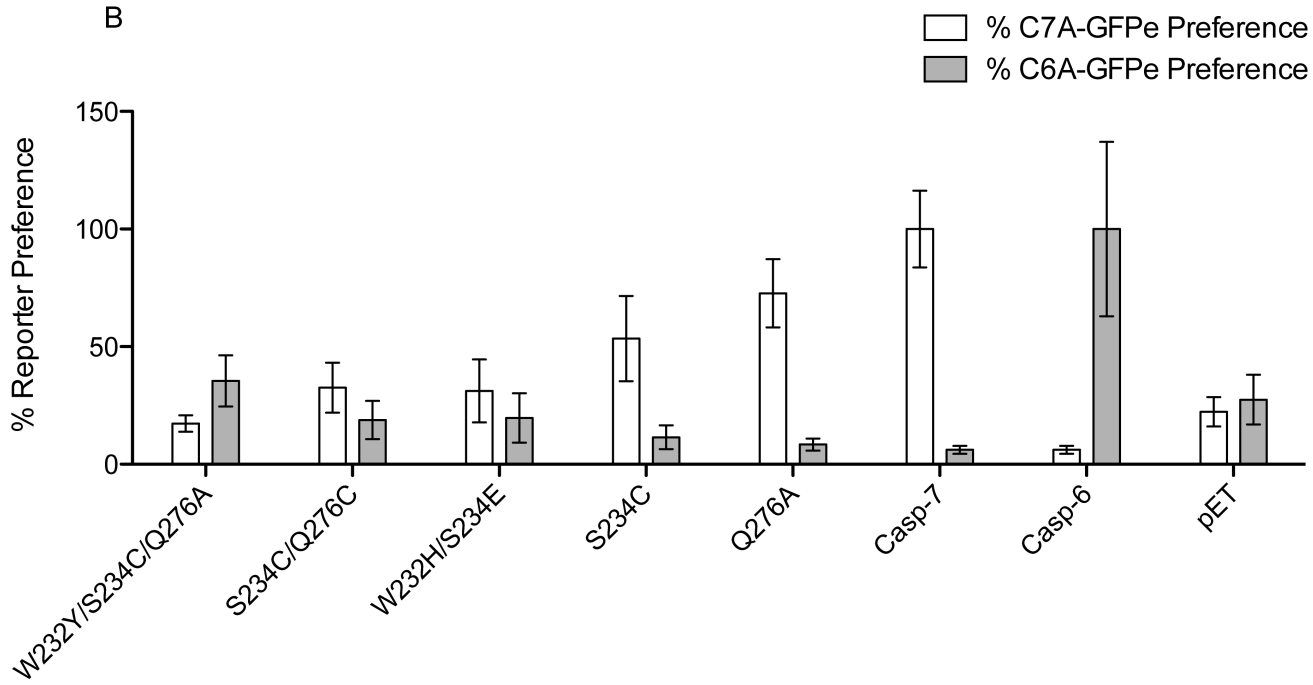
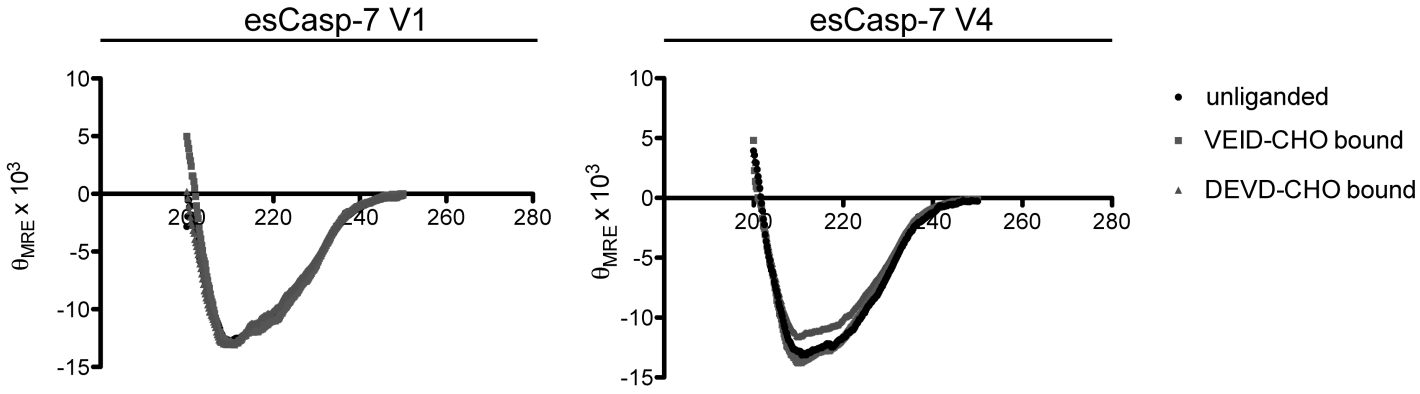


Figure S1. (A) Direct substitution (ds) of residues in S2 and S4 subsites of casp-7 by the homologous residues in casp-6 (W232H/S234E/Q276F) yielded protein with no measurable catalytic rate. ND: Not Determined due to ambiguous or incomplete saturation. [a]Previously Reported Values.[5] [b]Previously Reported Values.[3] (B) The % casp-7 preferences (white) were calculated by dividing the in cell fluorescence response to the C7A-GFP[e] reporter by their fluorescence response to C6A-GFP[e] and normalizing it to the fluorescence response of WT casp-7 to C7A-GFP[e]. Likewise, % casp-6 preferences (gray) were calculated by dividing the fluorescence response to C6A-GFP[e] by the C7A-GFP[e] response and normalizing it to the fluorescence response of WT casp-6 to C6A-GFP[e].

A



B

		unliganded (°C)	Ac-VEID-CHO-bound (°C)	Ac-DEVD-CHO-bound (°C)
WT	Casp-6	77.0 ± 0.5[a]	80.3 ± 0.4[a]	ND
	Casp-7	59.8 ± 0.0	NM[b]	77.7 ± 3.8
esCasp	V1	59.3 ± 0.5	78.8 ± 1.1	68.9 ± 2.4
	V4	61.3 ± 0.7	75.0 ± 2.0	76.6 ± 0.5

Figure S2. esCasp-7 V1 and V4 maintain the same fold and similar stability to WT casp-7. (A) CD spectra of esCasp-7 V1 and esCasp-7 V4 unliganded or bound to VEID-CHO or DEVD-CHO demonstrate that the overall fold is largely unchanged between bound and unbound states. (B) The melting temperatures (T_m) for esCasp-7 V1 and V4 were assessed by monitoring loss of CD signal at 222 nm as a function of temperature. Binding of DEVD or VEID to esCasp-7 variants induces dramatic increase in stability, as is typical of WT casp-7 when bound to DEVD. [a]as previously reported.[17] [b]Not Measured (NM) due to an inability of the indicated inhibitor to completely occupy the active site.

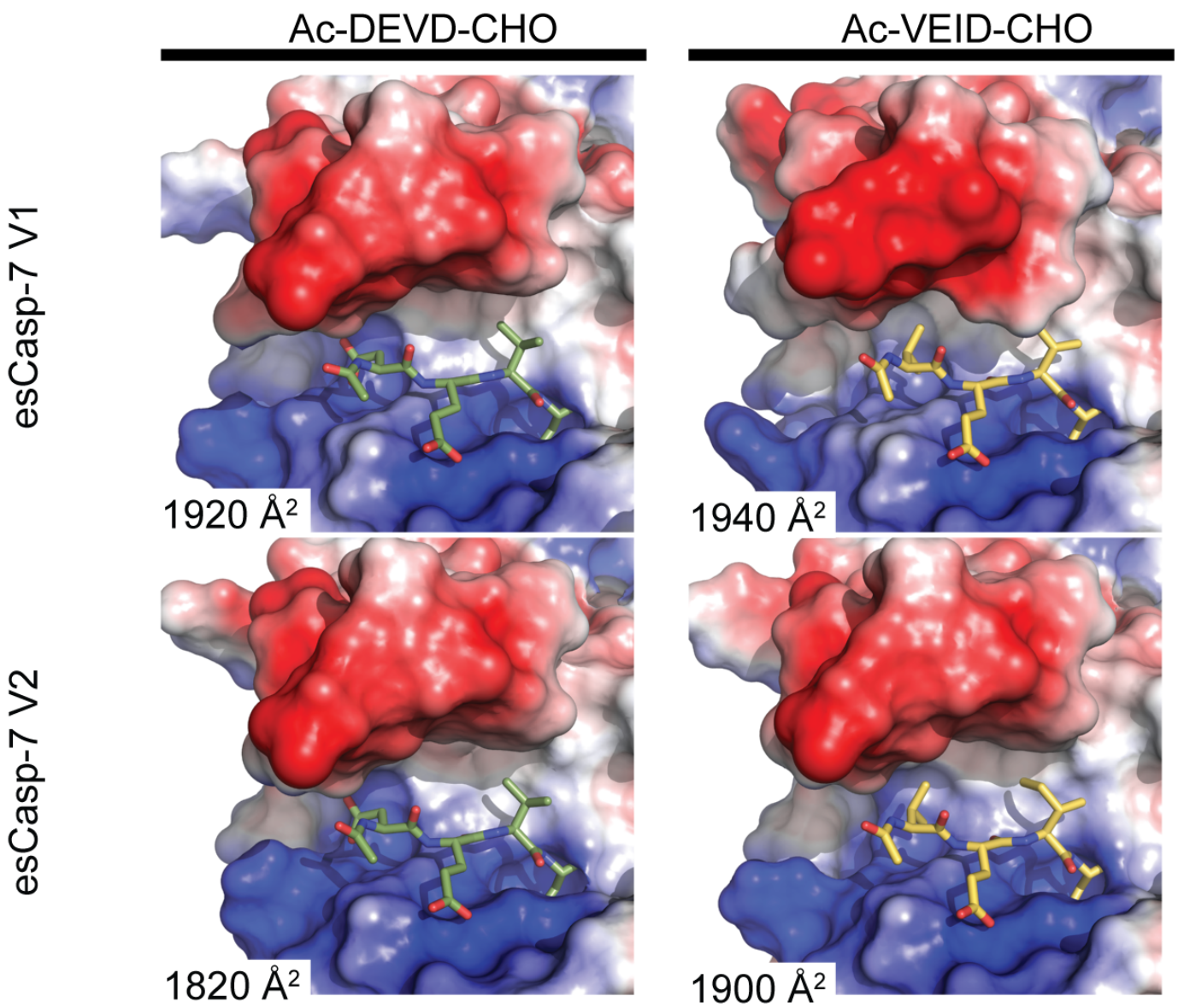


Figure S3. Plasticity of the active site fosters recognition of substrates. Consistent orientation of the L4 loop preserves the shape of the surface patches of positive (blue) or negative (red) charge or hydrophobic (white) character with DEVD (green sticks) or VEID (yellow) substrates. The solvent assessable surface areas of the interacting residues for each structure are listed.

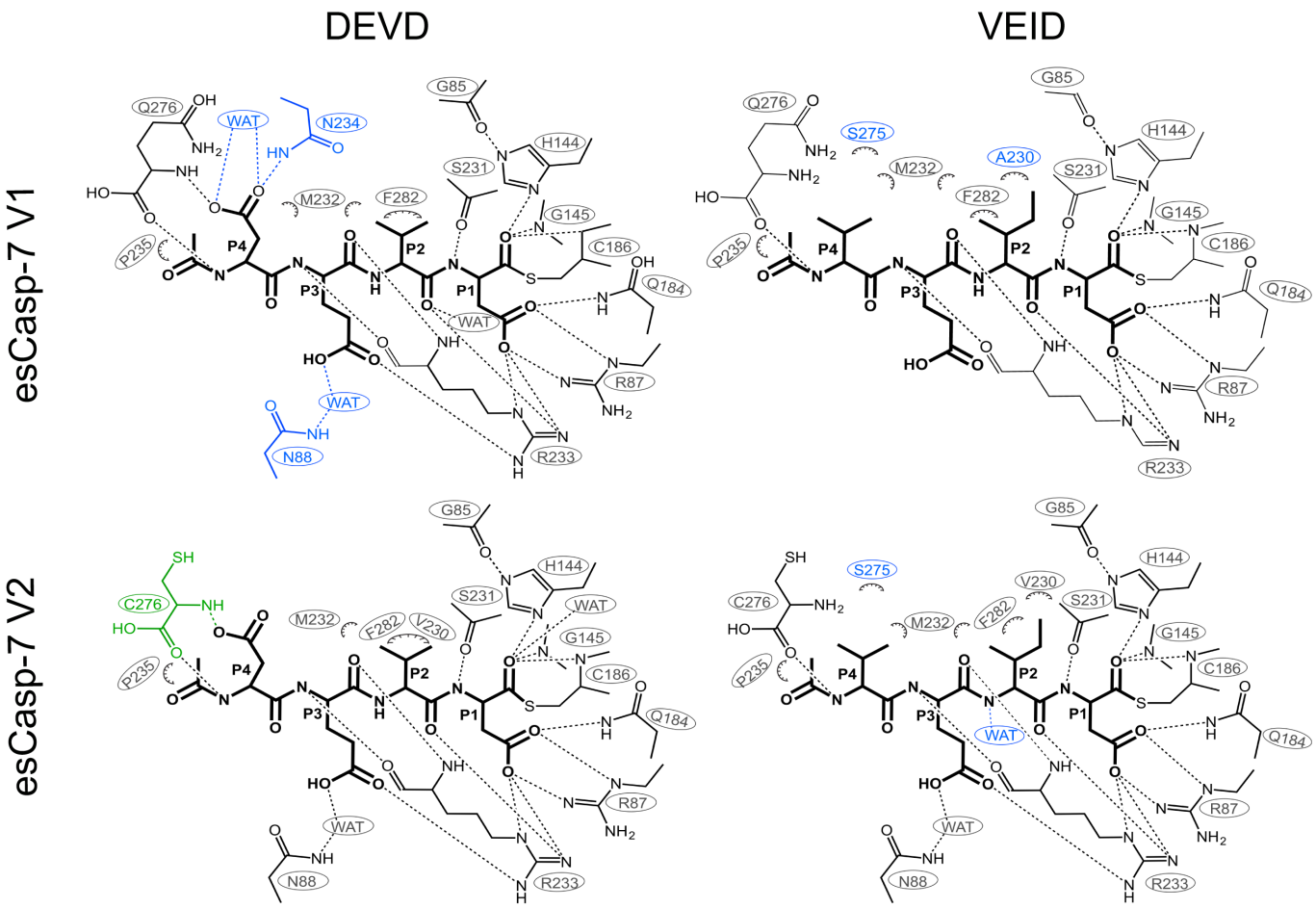


Figure S4. Interactions upon binding of esCasp-7 V1 or esCasp-7 V2 upon binding to DEVD or VEID. VEID results in three hydrophobic interactions around P2 Ile. Projections of substrate were drawn in ChemDraw (CambridgeSoft, PerkinElmer). Hydrophobic interactions of heavy atoms within 4Å are shown as hashed semicircles. Polar interactions within 3.5Å are shown as dashed lines. Interactions that differ between DEVD- and VEID-bound structures are shown in green while those lost are shown in blue.

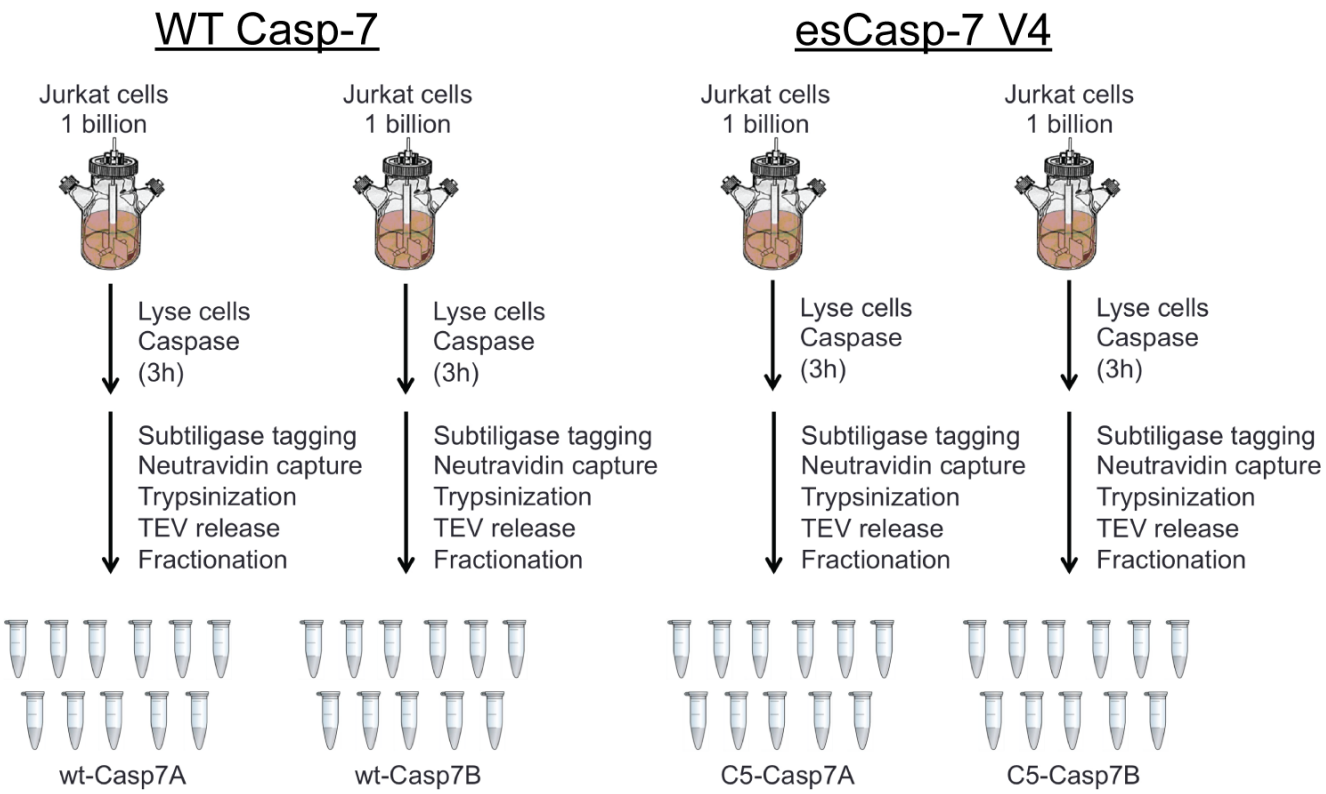


Figure S5. Experimental flow for *N*-terminomics experiments used to determine substrates of WT casp-7, WT casp-6 and esCasp-7 V4.

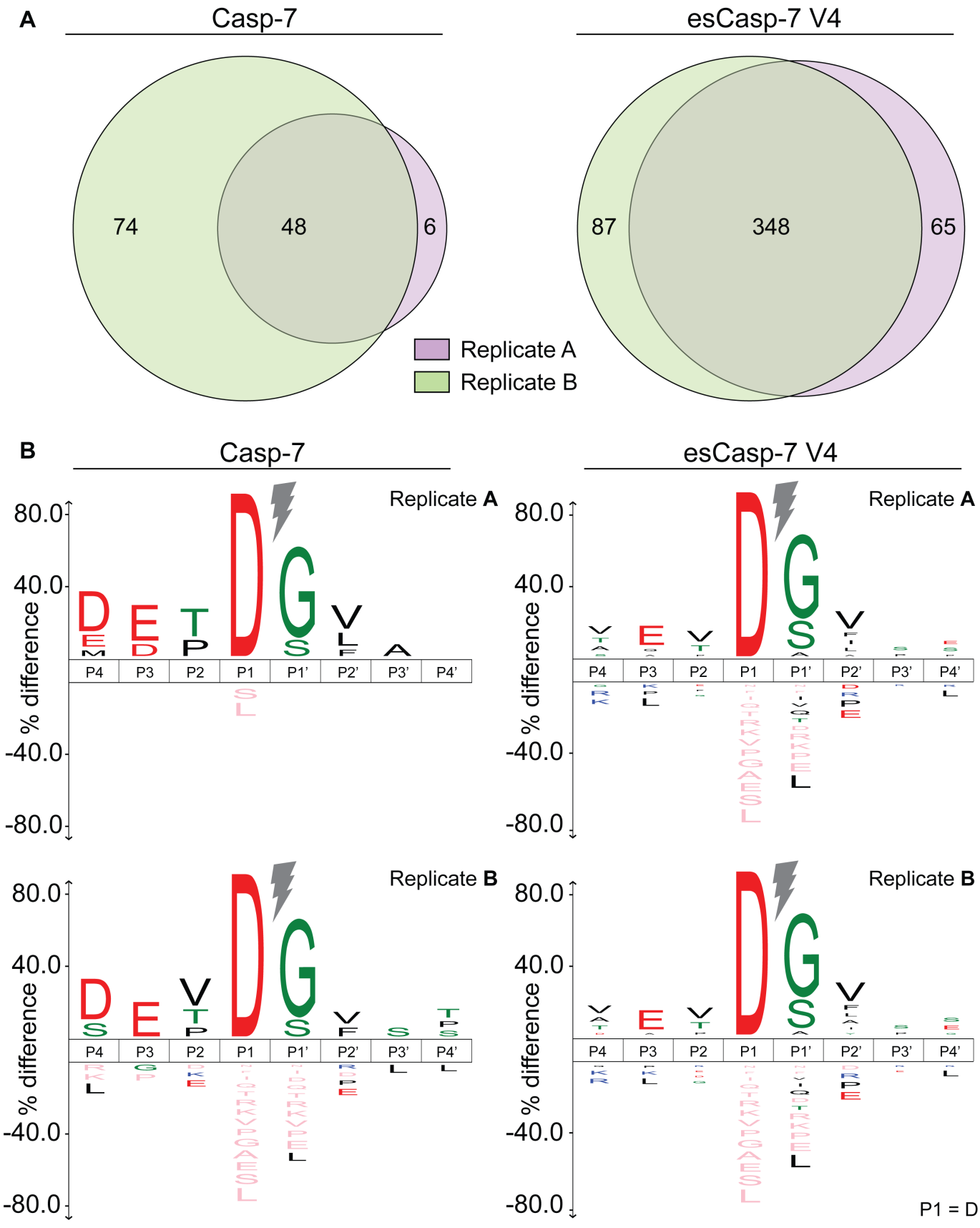
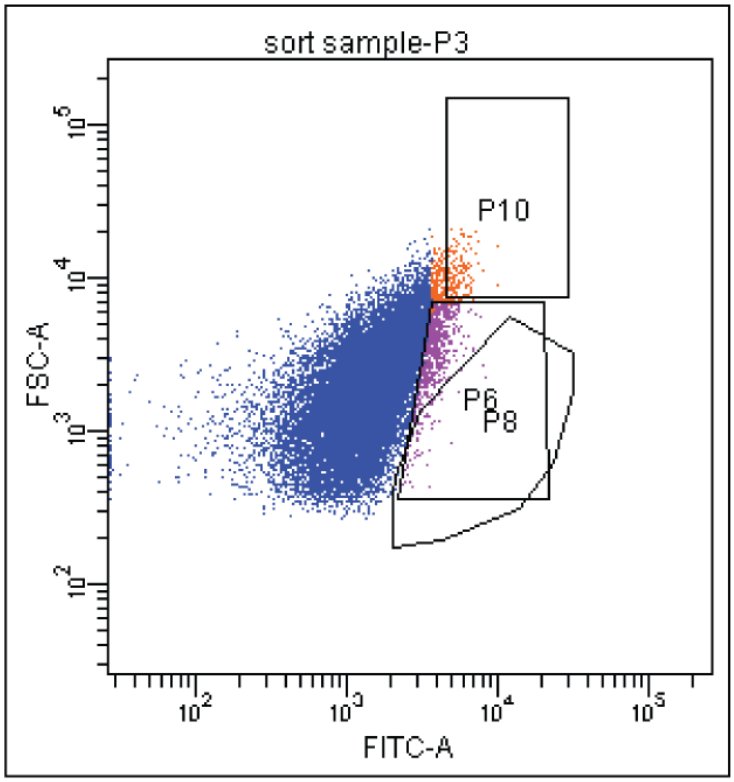
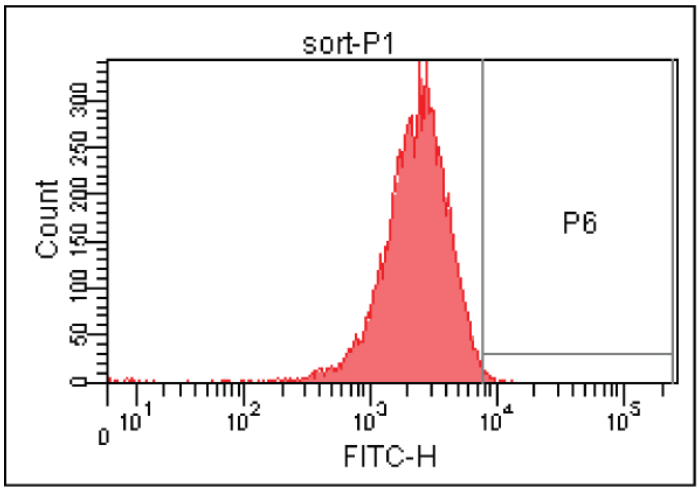


Figure S6. Replicates of substrate proteins identified by *N*-terminomics of WT casp-7 and esCasp-7 V4.

A



B



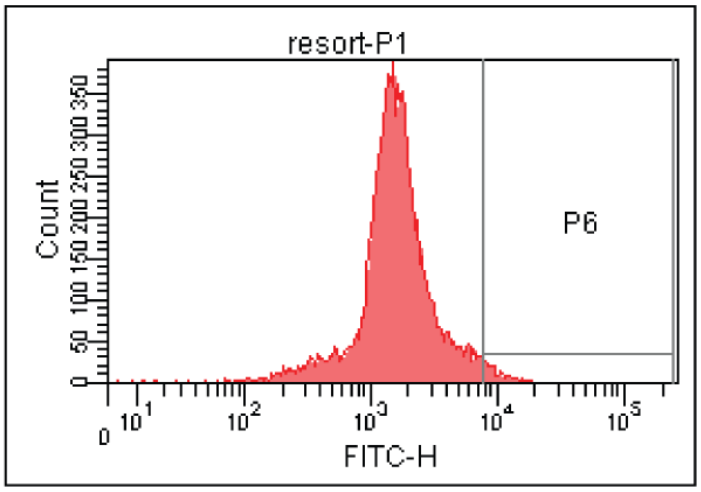


Figure S7. (A) Gating parameters for screening of casp-7 S2 and S4 library-A based on both FSC and FITC channels. Two gating parameters were applied using p6 and p10. (B) Events were sorted twice using p6, selecting for the top 0.5% fluorescent cells to enrich the selected population.

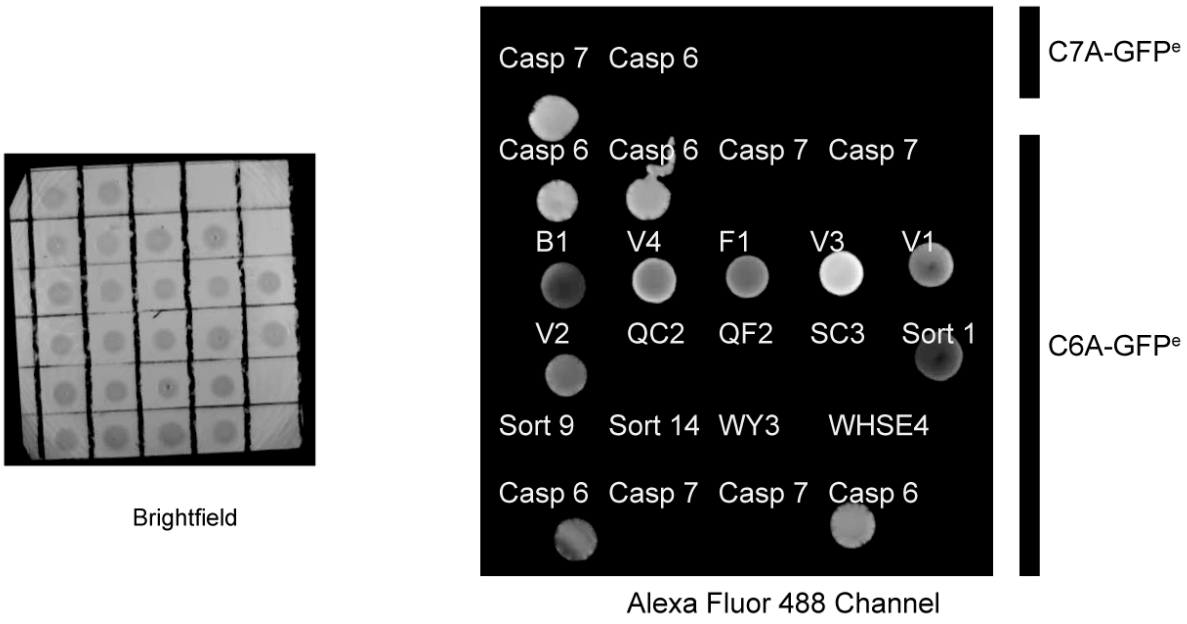


Figure S8. On-plate assay of WT casp-7 and WT casp-6 with selected variants co-transformed with C7A-GFP^{e} or C6A-GFP^{e} reporters. The bright field image shows that all colonies were of approximately similar diameter. The 488 nm illumination shows in cell fluorescence resulting cleavage of the respective reporter.

References:

(1) Nicholls, S. B., Chu, J., Abbruzzese, G., Tremblay, K. D., and Hardy, J. A. (2011) Mechanism of a genetically encoded dark-to-bright reporter for caspase activity. *J. Biol. Chem. 286*, 24977–86.
(2) Wu, P., Nicholls, S. B., and Hardy, J. A. (2013) A tunable, modular approach to fluorescent protease-activated reporters. *Biophys. J. 104*, 1605–1614.
(3) Witkowski, W. A., and Hardy, J. A. (2011) A designed redox-controlled caspase. *Protein Sci. 20*, 1421–31.
(4) Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif. 41*, 207–234.
(5) Velázquez-Delgado, E. M., and Hardy, J. A. (2012) Phosphorylation regulates assembly of the caspase-6 substrate-binding groove. *Structure 20*, 742–51.
(6) Nicholls, S. B., and Hardy, J. A. (2013) Structural basis of fluorescence quenching in caspase activatable-GFP. *Protein Sci. 22*, 247–257.
(7) Tian, W. X., and Tsou, C. L. (1982) Determination of the rate constant of enzyme modification by measuring the substrate reaction in the presence of the modifier. *Biochemistry 21*, 1028–1032.
(8) Ekici, Ö. D., Li, Z. Z., Campbell, A. J., James, K. E., Asgian, J. L., Mikolajczyk, J., Salvesen, G. S., Ganesan, R., Jelakovic, S., Grütter, M. G., and Powers, J. C. (2006) Design, synthesis, and evaluation of aza-peptide Michael acceptors as selective and potent inhibitors of caspases-2, -3, -6, -7, -8, -9, and -10. *J. Med. Chem. 49*, 5728–5749.
(9) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658–674.
(10) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. a., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr. 67*, 235–242.
(11) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr. 66*, 486–501.
(12) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr. 53*, 240–55.
(13) Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr. 66*, 213–221.
(14) Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng. Des. Sel. 8*, 127–134.
(15) Agard, N. J., Mahrus, S., Trinidad, J. C., Lynn, A., Burlingame, A. L., and Wells, J. A. (2012) Global kinetic analysis of proteolysis via quantitative targeted proteomics. *Proc. Natl. Acad. Sci. 109*, 1913–1918.
(16) Julien, O., Zhuang, M., Wiita, A. P., O'Donoghue, A. J., Knudsen, G. M., Craik, C. S., and Wells, J. A. (2016) Quantitative MS-based enzymology of caspases reveals distinct protein substrate specificities, hierarchies, and cellular roles. *Proc. Natl. Acad. Sci. U.S.A.* (in press)
(17) Vaidya, S., and Hardy, J. A. (2011) Caspase-6 latent state stability relies on helical propensity. *Biochemistry 50*, 3282–3287.