Conservation of the Extended Substrate Specificity Profiles Among Homologous Granzymes Across Species*[S]

Kim Plasman‡§§§, Sebastian Maurer-Stroh¶‖, Jamshaid Ahmad, Han Hao¶, Dion Kaiserman**, Fernanda L. Sirota¶, Veronique Jonckheere‡§, Phillip I. Bird**, Kris Gevaert‡§, and Petra Van Damme‡§§§‡‡**

Granzymes are structurally related serine proteases involved in cell death and immunity. To date four out of five human granzymes have assigned orthologs in mice; however for granzyme H, no murine ortholog has been suggested and its role in cytotoxicity remains controversial. Here, we demonstrate that, as is the case for granzyme C, human granzyme H is an inefficient cytotoxin that together with their similar pattern of GrB divergence and functional similarity strongly hint to their orthologous relationship. Besides analyzing the substrate specificity profile of granzyme H by substrate phage display, substrate cleavage susceptibility of human granzyme H and mouse granzyme C was assessed on a proteome-wide level. The extended specificity profiles of granzymes C and H (*i.e.* beyond cleavage positions P4-P4′) match those previously observed for granzyme B. We demonstrate conservation of these extended specificity profiles among various granzymes as granzyme B cleavage susceptibility of an otherwise granzyme H/C specific cleavage site can simply be conferred by altering the P1-residue to aspartate, the preferred P1-residue of granzyme B. Our results thus indicate a conserved, but hitherto underappreciated specificity-determining role of extended protease-substrate contacts in steering cleavage susceptibility. ***Molecular & Cellular Proteomics 12: 10.1074/mcp.M113.028670, 2921–2934, 2013.***

Several molecular mechanisms are in place to combat transformed malignant cells and virally infected cells. Granzymes (Gr)[1], a family of structurally related serine proteases found in the granules of many immune cells, play crucial roles in such cellular defense mechanisms. The granzyme family consists of five human proteases (granzymes A, B, H, K, and M) and 10 murine members (granzymes A to G, K, M, and N). To date for four human granzymes (A, B, K, and M) clear murine orthologs have been assigned, while the most probable murine ortholog of human granzyme H (GrH) is granzyme C (GrC) based on their 70% sequence similarity, 61% sequence identity, and identical chromosomal location relative to granzyme B (GrB). Furthermore, both granzymes are expressed by NK and CD4+ T-cells (1, 2): GrH is constitutively expressed at high levels in NK cells and less in CD4+ T cells, whereas GrC expression can be observed after stimulation. Thus, their overlapping expression profiles further support possible functional similarities between mouse GrC and human GrH.

The physiological role of granzymes was presumed to be the induction of death in target cells. Granzyme B is a highly efficient cytotoxin (3) and other granzymes such as GrA, GrC, GrF, and GrK can cause cell death at high concentrations (4–8). Recently, granzymes A, K, and M were shown to steer inflammatory processes when used at physiological levels (9–11). Two previous studies identified GrH as an alternative cytotoxic effector protease. Although both studies showed typical hallmarks of apoptosis, including mitochondrial depolarization, reactive oxygen species (ROS) production, DNA degradation as well as chromatin condensation, Fellows *et al.* (12) found that, in contrast to GrB mediated cell death, GrH induced cell death did not result in caspase activation, cytochrome *c* release, or cleavage of Bid and/or ICAD. In sharp contrast however, Hou *et al.* (13) demonstrated that GrH induced apoptosis depended on caspase activation and that GrH cleaved ICAD and Bid, the latter ultimately resulting in mitochondrial cytochrome *c* release. Such discrepancies

From the ‡Department of Medical Protein Research, VIB, B-9000 Ghent, Belgium; §Department of Biochemistry, Ghent University, B-9000 Ghent, Belgium; ¶Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), Singapore 138671; ‖School of Biological Sciences (SBS), Nanyang Technological University (NTU), Singapore 637551; **Department of Biochemistry and Molecular Biology, Monash University, Victoria 3800, Australia

Received February 13, 2013, and in revised form, June 13, 2013
Published, MCP Papers in Press, June 20, 2013, DOI 10.1074/mcp.M113.028670

[1] The abbreviations used are: Gr, granzyme; Abz, aminobenzoyl; ACC, 7-amino-4-carbamoylmethylcoumarin; ACT, alpha-1-anti-Chymotrypsin; CD, cluster of Differentiation; COFRADIC, combined fractional diagonal chromatography; EC, effective concentration; FEN-1, Flap endonuclease 1; IAA, iodoacetamide; MTT, methylthiazol tetrazolium; NK cell, natural killer cell; PS-SCL, positional scanning synthetic combinatorial libraries; RCL, reactive center loop; ROS, reactive oxygen species; SCX, strong cation exchange; SILAC, stable isotope labeling by amino acids in cell culture; SLO, streptolysin-O.

have been documented in other granzyme studies and are usually linked to the use of different granzyme delivery systems, sources of recombinantly produced granzymes, differences in the granzyme concentration and species-specific differences in substrate specificities (14).

GrC induces cell death reminiscent of GrH induced cell death as observed by Fellows *et al.* (12), as both exert their cytotoxic functions independent of caspase activation, Bid or ICAD cleavage, or by mitochondrial release of cytochrome *c* (5). GrC induced apoptosis was characterized by the rapid externalization of phosphatidylserine, nuclear condensation and collapse, and single-stranded DNA nicking. Supporting evidence implying a role for GrC in lymphocyte induced cytotoxicity was inferred from the fact that GrB cluster-deficient mice (mice that do not express GrB and show a five- to sixfold reduced expression of GrC and GrF respectively) display a more pronounced defect in the clearance of allogeneic tumor cells when compared with GrB-only knockout mice (4). This suggests that GrC and/or GrF may be important for correct functioning of cytotoxic lymphocytes. Besides, in GrB-only knockout mice, a likely compensatory mechanism occurs, given that during cytotoxic lymphocyte activation, peak expression of GrC occurs earlier, giving rise to overall higher GrC levels as compared with wild-type mice (1, 4, 15). Of note, despite its implication in cytotoxicity, GrC was shown to be an inefficient cytotoxin, with a 2900-fold greater EC_{50} value as compared with hGrB when delivered into P815 cells with recombinant mouse perforin (16).

Positional Scanning Synthetic Combinatorial Libraries (PS-SCL) revealed the chymotrypsin-like activity of GrH, which it shares with granzyme M (17–19), with an optimal P4-P1 peptide substrate sequence Pro-Thr-Ser-Tyr. Less stringent specificities were observed at positions P4, P3, and P2 (19) where GrH seems to tolerate multiple amino acids with different chemical characteristics (especially neutral and aliphatic amino acids). Although GrH and GrM (optimal peptide identified as Lys-Val-Pro-Leu) share a P1 chymotryptic activity, GrH prefers bulkier, aromatic amino acids (Tyr and Phe) at P1 whereas GrM prefers Leu. Both further recognize Leu and Met at P1, implying that some substrates could be cleaved by both granzymes. GrM also shows broader specificities at P3, but at P4 and P2 it prefers basic residues and Pro respectively. Besides, GrC chymase activity could be inferred from N-terminal COFRADIC and substrate phage display screens, which defined the P4-P3′ substrate specificity of GrC as [Ile/Val]-X-[Phe-Tyr]-[Phe-Leu-Tyr-Met] ↓ X-[Gly-Ser]-[Asp-Glu] (16).

Recently, the crystal structure of the D_{102}>N GrH variant in complex with a decapeptide substrate (PTSYAGDDSG) or inhibitor (Ac-PTSY-chloromethylketone) was resolved (20). The electron density maps clearly showed the full length protease adopting a canonical structure of 2 α-helices and 13 β-strands assembled into two juxtaposed β-barrel domains bridged by the catalytic triad (composed of His_{57}, Asp_{102} and Ser_{195}). The S1 specificity pocket is built up from residues of 2 loops; loop 189 (from residue 183–196) and loop 220 (from residue 215–226) with the determinants being Thr_{189}, Gly_{216}, and Gly_{226}. From these, Gly_{226} was assigned as the most important determinant for the preference of bulky aromatic amino acids (such as Tyr and Phe) at P1. Where GrH contains a Gly at position 226, GrB, GrC, and GrM harbor an Arg, Gln, and Pro residue respectively (supplemental Fig. S1). Mutation of Gly_{226} in GrH to Arg_{226} compromised binding of bulky aromatic residues and enabled interaction with negatively charged amino acids. Hydrogen bond formation between a P1 Tyr and Asn_{217} of GrH could furthermore strengthen the observed preference of Tyr over Phe at P1 (19). The presence of Pro at position 226 in GrM instead of Gly narrows the S1 pocket and might be indicative for the preference of Leu instead of Phe and Tyr at P1. In addition, the structure revealed that the S4′ pocket formed by the backbones of the Arg_{39}-Lys_{40}-Arg_{41} motif resulted in a preference for acidic residues at P4′, in addition to influencing the P3′ preference for acidic residues via salt bridge formation with this Lys_{40}. Although the Arg_{39}-Lys_{40}-Arg_{41} motif is a unique GrH feature, GrB possesses a partially degenerate Leu-Lys-Arg motif in which Lys_{40} also enables interaction with acidic residues in P3′. Next to P3′, proteome-wide screening for GrB substrates led to the identification of a clear preference for acidic residues at P4′ caused by salt bridge formation with Arg_{41} (21), a characteristic that was previously assigned to Lys_{40} of GrB (22). To validate these structural observations, similar to the P4′ Asp mutation in the GrB substrate PI-9 (22), both Asp residues found at P3′ and P4′ in the nuclear phosphoprotein La, previously identified as a macromolecular GrH substrate (23), were mutated to Ala. These mutations completely abolished GrH mediated proteolysis, a first indication that the P4-P1 specificity profile is not a sole determinant for substrate recognition by GrH. Next to the La phosphoprotein, cleavage of the viral adenovirus DNA binding protein (DBP) and the 100K assembly protein (L4–100K), the latter resulting in the relieve of GrB inhibition by L4–100K, has been observed, implicating GrH in host antiviral defense and indicative for a functional synergism between GrH and GrB (24).

Elucidation of the crystal structure, with an electron density map showing residues Ile_{16}-His_{244} (*i.e.* 94% of full length GrC and 99.6% of active GrC), furthermore showed that wild-type GrC can be restrained in its proteolytic function despite the presence of the catalytic triad residues His_{57}-Asp_{102}-Ser_{195} (16). Comparing the crystal structures of GrC and GrA showed an unusual conformation of the active site, which could explain the inactivity of GrC. Apparently, the 190-strand, preceding Ser_{195} of the catalytic triad, has undergone a register shift on the structural level leading to Phe_{191} filling the S1 pocket. This pocket is furthermore covered by Glu_{192}, which forms a salt bridge with Arg_{99} and a hydrogen bond with the backbone amide of Ser_{195}. Because of this unusual conformation, the Glu_{192}-Glu_{193} peptide bond points away from the substrate, as such leading to an improperly formed oxyanion

hole, which normally stabilizes the negatively charged substrate oxygen atom during catalysis. Mutation of Glu_{192}-Glu_{193} to the corresponding amino acids in its closest related homolog GrB (Arg_{192}-Gly_{193}) disrupted the Glu_{192}-Ser_{195} hydrogen bond and caused a shift of the 190-strand, thereby clearing the S1 pocket and giving rise to an active GrC mutant. These results indicate allosteric control of wild-type GrC in which binding of a substrate or cofactor might stabilize the 190-strand, which becomes extremely mobile due to breaking the Glu_{192}-Ser_{195} hydrogen bond and turning the region surrounding the active site more rigid.

To elucidate a possible functional homology between GrH and GrC, we performed differential degradome analyses using N-terminal COFRADIC in the species-matching proteome backgrounds. For these analyses, we made use of the active $E_{192}E_{193}$>RG GrC mutant as described in (16) (further referred to as mut GrC). These analyses, further complemented with phage display data on granzyme H, clearly show that both granzymes display a highly similar substrate specificity profile, analogous to other orthologous granzymes. Besides, and in contrast to GrA, a general conservation of the extended substrate specificity profiles among the homologous granzymes B, C, H, and M could be observed across species, highlighting the importance of the extended substrate specificity in steering substrate cleavage susceptibility.

EXPERIMENTAL PROCEDURES

Production of Recombinant Granzymes—Recombinant human granzyme B was expressed and purified from *Pichia pastoris* and its activity was determined as described previously (25, 26). Recombinant, enterokinase activated wild-type GrC and the mutated form of GrC (E192R/E193G) were prepared as described previously (16). Recombinant human granzyme H was expressed in *Pichia pastoris* and purified to homogeneity. Correct folding and activity were assessed by serpin-mediated GrH complex formation.

Serpin-mediated Granzyme H Complex Formation—GrH was incubated with or without a 50-fold molar excess of alpha-1-anti-chymotrypsin (ACT) in Tris buffered saline (20 mM Tris, pH 7.4, 150 mM NaCl) for 30 min at 37 °C. Samples were then mixed with SDS loading buffer (62.5 mM Tris, pH 6.8, 2% SDS, 20% glycerol, 0.1 M dithiothreitol), boiled and resolved by 12.5% SDS-PAGE. In-gel protein staining was performed using Rapid-Ag-Stain™ (MP Biomedicals, Solon, OH) to assay complex formation of GrH and ACT.

Cell Culture—Human K-562 cells (ATCC; American Type Culture Collection, Manassas, VA, CCL-243) and mouse YAC-1 cells (ATCC; TIB-160) were grown in RPMI 1640 medium (Invitrogen, Carlsbad, CA, #61870–010) containing either natural, $^{13}C_6$ or $^{13}C_6{}^{15}N_4$ L-arginine (Cambridge Isotope Labs, Andover, MA) at a concentration of 57.5 μM for K-562 cells (*i.e.* 5% of the suggested concentration present in RPMI at which L-arginine to proline conversion was not detectable) and 575 μM for YAC-1 cells (50% of the suggested concentration). Media were supplemented with 10% dialyzed fetal bovine serum (Invitrogen, 26400–044), 100 units/ml penicillin (Invitrogen), and 100 μg/ml streptomycin (Invitrogen). For YAC-1 cells, β-mercaptoethanol (M7522, Sigma) was added freshly on subcultivating cells at a f.c. (final concentration) of 55 μM. Cell populations were cultured at 37 °C and 5% CO_2 for at least six population doublings for complete incorporation of labeled arginine.

Granzyme H Substrate Phage Display—Phage display was performed as described previously (16) using the random 9-mer library for 4 rounds of selection.

SLO-mediated Granzyme Cytotoxicity—Cell death mediated by granzymes and recombinant streptolysin-O (SLO) was performed according to (25) and (27).

Preparation of Cell Lysates for N-Terminal COFRADIC Analysis—Stable isotope labeling by amino acids in cell culture (SILAC)-labeled K-562 and YAC-1 cells were washed in D-PBS and resuspended at 7×10^6 cells per ml in 50 mM Tris-HCl (pH 8.0) and 100 mM NaCl. Cells were subjected to three rounds of freeze-thaw lysis and supernatants, cleared by centrifugation for 10 min at 16,000 $\times$ *g*, were treated with 200 nM of recombinant human granzyme H (GrH), recombinant mutant mouse granzyme C (mut GrC) or left untreated (control). Solid guanidinium hydrochloride was added to a final concentration of 4 M to inactivate granzymes and denature all proteins. Before mixing the samples in equal amounts, proteins were reduced and alkylated using TCEP.HCl (1 mM f.c.) and IAA (2 mM f.c.) respectively, for 1 h at 30 °C. Subsequent steps of the N-terminal COFRADIC protocol were performed as described previously (28). The proteome was digested overnight at 37 °C with sequencing-grade, modified trypsin (Promega, Madison, WI) (enzyme/substrate of 1/100 w/w).

LC-MS/MS Analysis and Data Processing—Electrospray ionization liquid chromatography-tandem MS (ESI LC-MS/MS) analysis was performed on a Q-TOF (Waters Corporation) mass spectrometer as described before (29). ESI-Q-TOF MS/MS peptide fragmentation spectra were converted to pkl files using the Masslynx® software (version 4.1, Waters Corporation). N-terminal peptides were identified using a locally installed version of the MASCOT database search engine version 2.2.1 and 2.2.04 (www.matrixscience.com) and the Swiss-Prot database (version 2012_09 of UniProtKB/Swiss-Prot protein database, containing in total 538,010 sequence entries from which 20,234 human and 16,563 mouse entries) was searched with restriction to human or mouse proteins. Truncated peptide databases made by DBToolkit (30), were searched in parallel to pick up protein processing events more efficiently (*e.g.* (31, 32)). The following search parameters were used. Peptide mass tolerance was set at 0.2 Da and peptide fragment mass tolerance at 0.1 Da with the ESI-QUAD-TOF as selected instrument for peptide fragmentation rules for the Q-TOF Premier data. The protease setting semi-ArgC/P (*i.e.* no restriction toward arginine-proline cleavage) was used allowing one missed cleavage. Variable modifications were pyroglutamate formation of N-terminal glutamine, pyrocarbamidomethyl formation of N-terminal alkylated cysteine, deamidation of asparagine, acetylation or tri-deuteroacetylation of the alpha-N terminus. Fixed modifications were methionine oxidation (sulfoxide), carbamidomethyl for cysteine, tri-deuteroacetylation of lysine and for identifying heavy labeled peptides, [$^{13}C_6$] or [$^{13}C_6{}^{15}N_4$]-arginine were additionally set as fixed modifications. Only MS/MS-spectra and corresponding identifications that exceeded the corresponding MASCOT's threshold score of identity (at 95% confidence level) and that were ranked one, were withheld. In addition, spectra that received a low MASCOT ion score (5 or less points above threshold for identity) were further interrogated and only spectra that contained b- and y- fragment ions covering a stretch of three consecutive amino acids were considered identified. The estimated false discovery rate by searching decoy databases was typically found to lie between 2 and 4% on the spectrum level (28). Whenever a peptide matched to multiple members of a protein family (redundancy), all protein entries were reported (column "Description," supplemental Tables S1 and S2) and ranked alphabetically. Corresponding orthologous mouse or human cleavage sites, as determined in (21), are reported in supplemental Tables S1 and S2. The peptide intensity ratios were manually calculated from the MS data in the Masslynx® software version 4.1 environment. To correct for varia-

tions in initial protein concentrations, the $^{12}C_6/^{13}C_6$ and $^{12}C_6/^{13}C_6{}^{15}N_4$ ratios of the 2670 human and 1866 mouse database-annotated protein N-terminal peptides (peptides starting at position 1 or 2) were first subjected to robust statistics on the base-2 logarithms of their ratios (33). All the mass spectrometry is deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) with the data set identifier PXD000152 and DOI 10.6019/PXD000152. Under the ProteomeXchange accession code; PXD000152 and PRIDE accession codes; PRIDE accessions: 28635–28636.

Cloning and Expression of (Truncated) FEN-1—The pOTB7-FEN-1 (RZPD Imagenes, cat # IRAUp969C102D, Germany) vector served as a template to generate PCR products of full-length and truncated FEN-1 suitable for use as substrates in a Gateway® BP recombination reaction with a donor vector (pDONR221, Invitrogen, cat # 12536–017) thereby creating an entry clone. The following PCR primers were used to generate *att*B flanked PCR products (forward primer FEN-1: 5′-ggggacaagtttgtacaaaaaagcaggcttcaccatgggaattcaaggcctggcc-3′ and truncated FEN-1: 5′-ggggacaagtttgtacaaaaaagcaggcttcaccatgtggtgaaaagcagttctctgagg-3′; reverse primer 5′-ggggaccactttgtacaagaaagctgggtcttttccccttttaaacttccctgc-3′). The reverse primer was designed to fuse the desired PCR products in frame with a C-terminal GFP tag.

Subsequently the FEN-1 inserts were shuttled into the pcDNA 6.2/C-EmGFP-DEST destination vector (Invitrogen, cat # V355–20) using the LR recombination reaction.

K-562 cells were transiently transfected for 24 h with 0.4 μg/ml (final concentration) of the eukaryotic expression vectors using Lipofectamine™ LTX and PLUS™ Reagents (Invitrogen) according to the manufacturer's instructions. Following transfection, cells were resuspended in 50 μl 50 mM Tris-HCl (pH 8.0) and 100 mM NaCl and subjected to three rounds of freeze-thaw lysis and centrifuged for 10 min at 16,000 $\times$ *g*. Whenever indicated, cleared cell lysates were incubated for 15 min with 5 mM iodoacetamide at 37 °C in the dark before incubation with 200 nM mut GrC for 1 h at 37 °C. Following mut GrC incubation, NuPAGE® LDS Sample Buffer (Invitrogen) was added and samples were heated for 10 min at 70 °C. Proteins were then separated in MOPS buffer using 4–12% polyacrylamide Criterion XT-gels (1.0 mm thick gels) in MOPS buffer (Bio-Rad) at 150 V. Subsequently, proteins were transferred onto a PVDF membrane. Membranes were blocked for 30 min in a 50:50 Tris-buffered saline and 0.1% Tween-20 (TBS-T) Odyssey Blocking solution (LI-COR, cat n° 927–40003) and probed for 1 h with primary anti-GFP antibody (Invitrogen, cat # A11122) in TBS-T/Odyssey blocking buffer. Following three 10 min washes in TBS-T, membranes were incubated with secondary anti-rabbit antibody (IRDye 800 CW goat anti-Rabbit antibody IgG, LI-COR, cat n° 926–32211) for 1 h in TBS-T/Odyssey blocking buffer followed by three washes in TBS-T or TBS (last wash step), bands were visualized using an Odyssey infrared imaging system (LI-COR).

Granzyme Cleavage of In Vitro Transcribed and Translated FEN-1—pOTB7-FEN-1 (RZPD Imagenes, cat # IRAUp969C102D, Germany) was used as template for *in vitro* coupled transcription/translation in a rabbit reticulocyte lysate system according to the manufacturer's instructions (IVTT; Promega) to generate [^{35}S]methionine-labeled FEN-1. The SP6 TNT Quick Master Mix (40 μl) - containing TNT rabbit reticulocyte lysate, TNT reaction buffer, TNT SP6 RNA polymerase, an amino acid mixture without methionine and RNasin Ribonuclease Inhibitor - was mixed with 2 μl [^{35}S]-methionine (20 μCi/ml), 1 μl of plasmid DNA template (1 μg/μl) and 7 μl nuclease-free water. Following 90 min incubation at 30 °C, translates were incubated with iodoacetamide (f.c. 5 mM) for 20 min at 37 °C as to avoid downstream caspase activity upon granzyme treatment (34). The translates (5 μl) were subsequently incubated for 1.5 h at 37 °C in granzyme assay buffer with the indicated concentrations of recombinant granzymes in a total volume of 30 μl. Cleavage reactions were stopped by the addition of NuPAGE® LDS Sample Buffer (Invitrogen) and heating the samples for 10 min at 70 °C. Proteins were separated on 4–12% NuPAGE® Bis-Tris gradient gels and subsequently transferred onto a PVDF membrane as indicated above, these membranes were then air-dried and exposed to a film suitable for radiographic detection (ECL Hyperfilms, Amersham Biosciences, Buckinghamshire, UK). Radiolabeled proteins and their cleavage fragments were visualized by autoradiography.

Site-directed PCR-Mutagenesis—The cDNA encoding the FM_{310}>FD, FM_{310}>FY, FM_{310}>FF, FM_{310}>AA, FM_{310}>AD and FM_{310}>DA mutants of FEN-1 were generated by site-directed PCR-mutagenesis (QuickChange, Stratagene) according to the manufacturer's instructions using the following primer pairs: FEN-1-FM_{310}>FD-forward, 5′-tgaagaagagctgatcaagttcgactgtggtgaaaagcagttctctg-3′ and FEN-1-FM_{310}>FD-reverse, 5′-acttcttctcgactagttcaagctgacaccacttttcgtcaagagac-3′; FEN-1-FM_{310}>FY-forward, 5′-tgaagaagagctgatcaagttctactgtggtgaaaagcagttctctg-3′ and FEN-1-FM_{310}>FY-reverse, 5′-acttcttctcgactagttcaagatgacaccacttttcgtcaagagac-3′; FEN-1-FM_{310}>FF-forward, 5′-aagaagagctgatcaagttcttctgtggtgaaaagcagttctc-3′ and FEN-1-FM_{310}>FF-reverse, 5′-ttcttctcgactagttcaagaagacaccacttttcgtcaagag-3′; FEN-1-FM_{310}>AA-forward, 5′-ccaaatgaagagctgatcaaggccgcgtgtggtgaaaagcagttctctga-3′ and FEN-1-FM_{310}>AA-reverse, 5′-ggtttacttcttctcgactagttccggcgcacaccacttttcgtcaagagact-3′; FEN-1-FM_{310}>AD-forward, 5′-gaagagctgatcaaggacgcgtgtggtgaaaag-3′ and FEN-1-FM_{310}>AD-reverse, 5′-cttctcgactagttcctgcgccacccacttttc-3′; FEN-1-FM_{310}DA-forward, 5′-aagagctgatcaaggccgactgtggtgaaaagcagttc-3′ and FEN-1-FM_{310}DA-reverse, 5′-ttctcgactagttccggctgacaccacttttcgtcaag-3′. The correctness of all (mutant) cDNA sequences generated was confirmed by DNA-sequencing.

Physical Property Motif Analysis—The positional amino acid frequency vectors of the ungapped substrate motif alignment calculated over sliding windows of 1 to 12 positions was correlated with an extensive collection of physical property scales and selected those that were significant with $R^2 > 0.6$. Properties that themselves closely correlated ($R^2 > 0.9$) with average database composition scales were excluded. If multiple, related properties (cross-correlation of $R^2 > 0.4$) were found for the same region, only those with the highest correlation were kept to eliminate redundant motif descriptions. To favor stretches of continuous significant correlations over high correlations on single positions, the values for the selection were normalized by the square root of the number of positions involved (21, 35).

Structural Classification of Cleavage Sites—Structural classification of cleavage sites was conducted as follows. First, we compared the cleavage locations with the annotation of available structures for sequence entries in UniProtKB (13). For each overlap, we manually confirmed if the site is part of the resolved 3D structure or unresolved loops or termini without atomic coordinates of the PDB entries using YASARA (14). In the former case, the site would be classified as "known structure" and in the latter case as "known disorder." Furthermore, we extended the search for structure locations by homology searches of the respective region with BLASTP (15) against PDB (16), which ensures only highly similar hits and hence reliable structure prediction. As before, all structural locations of cleavage sites predicted by similarity were manually confirmed, giving rise to the classification of "predicted structure." Finally, cleavage sites in regions outside of known or predicted structure are divided into those with "unknown structure status" and those that are "predicted disorder" as determined by IUPred (17) using a threshold setting of 0.54, which was recently benchmarked (18) to give reliable disorder assignments with less or equal to 5% false positive predictions.

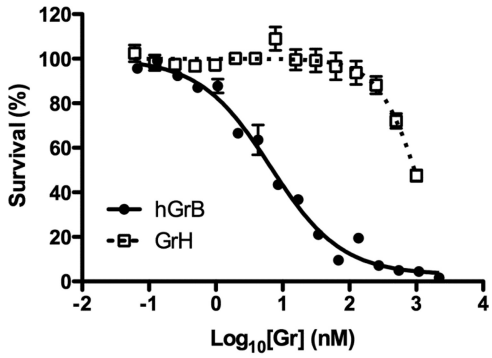


FIG. 1. **Granzyme H is a poor cytotoxin.** Jurkat cells were treated for 1 h with a sublytic concentration of streptolysin-O (SLO) and varying concentrations of GrH (squares, dotted line) or GrB (circles, solid line). SLO was inactivated by addition of RPMI containing 10% fetal calf serum and cells were allowed to recover overnight, survival was measured by the methylthiazol tetrazolium (MTT) assay. EC_{50} values were determined by fitting a sigmoidal dose-response curve with variable slope and are indicated below the curves together with the standard deviation (S.D.) values derived from two independent experiments.

RESULTS

Granzyme H Displays a Low Cytotoxicity Potential Compared With Granzyme B—Human granzyme H (GrH) was expressed in the yeast *Pichia pastoris* and purified to homogeneity. To confirm that GrH was correctly folded, an active site titration was performed by incubating recombinant granzyme H (GrH) with the serpin inhibitor alpha-1-antichymotrypsin (ACT). Serpin-mediated protease inhibition requires substrate recognition of an exposed reactive center loop (RCL), followed by cleavage at the reactive bond leading to a conformational change that distorts the protease active site, thereby trapping the protease in a 1:1 covalent, inhibitory complex with the serpin. ACT was selected as it contains a double Leu motif at the reactive site (supplemental Fig. S2), which PS-SCL predicts would be preferred as a substrate by GrH (19). Making use of excess ACT, gel-based analyses revealed formation of the SDS-stable complex, indicative of serpin-protease interaction, demonstrating that essentially all the GrH moves into complex (supplemental Fig. S2).

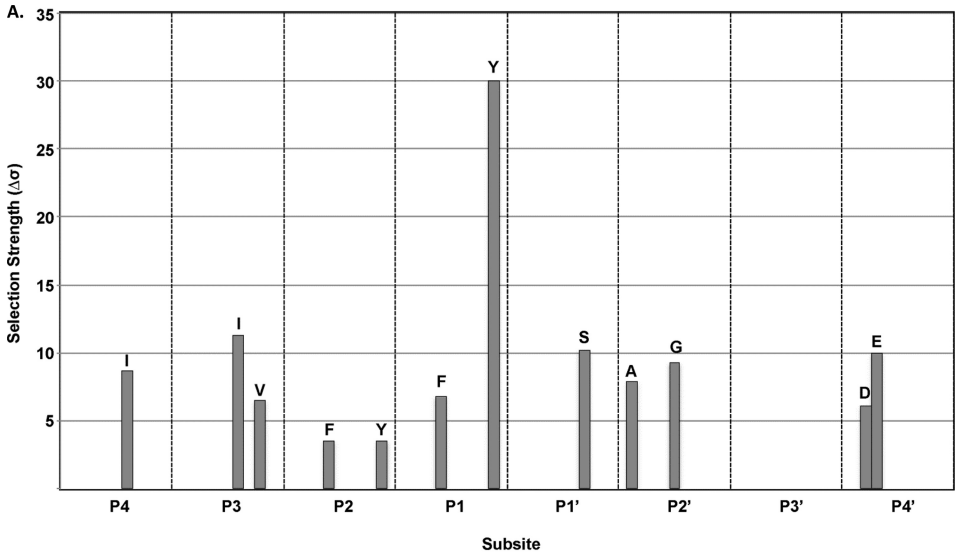
Having demonstrated the full catalytic potential of GrH, GrH cytotoxicity was studied by monitoring cell survival on streptolysin-O (SLO) mediated granzyme delivery in cell culture (Fig. 1). Quantification of granzyme cytotoxicity was performed using a dose-response analysis, in which the amount of granzyme required to kill 50% of the target cells (EC_{50}) was measured (25). The over 170-fold greater EC_{50} value for human GrH (1.658 μM $\pm$ 0.457) as compared with human GrB (9.9 nM $\pm$ 4.7) (Fig. 1), suggests that GrH is an inefficient cytotoxin similar to (unlocked) GrC (EC_{50} values of 39.6 μM and 11.2 μM for locked and unlocked GrC respectively) (16), an observation supporting its role in viral clearance rather than cell death (24).

Substrate Phage Display—By employing a substrate phage display system, we examined the extended substrate specificity of GrH. Here, random 9-mer peptides linked to a His-tag were fused to the C terminus of the T7 phage coat protein (36). Phages were captured on a nickel-Sepharose support and eluted by GrH incubation. After several rounds of capture and release, sequencing of 96 phages yielded 77 unique sequences and a consensus recognition sequence Ile-[Ile/Val]-[Phe/Tyr]-[Tyr/Phe] $\downarrow$ Ser-[Gly/Ala]-X-[Glu/Asp] (Fig. 2). Incubation with the peptide substrate aminobenzoyl (Abz)-Arg-Ile-Met-Phe-Phe-Ser-Gly-Ala-Thr-Lys(dnp(dinitrophenyl))-Asp indicated cleavage after the double Phe (data not shown). Despite the identical P1 preference observed, a less restricted substrate specificity profile was identified by combinatorial peptide analysis using a completely diversified PS-SCL of ACC (7-amino-4-carbamoylmethylcoumarin) substrates yielding the consensus sequence Pro-Thr-Ser-Tyr (19). Especially, the broad preferences at P2 in contrast with the more restricted preference for Tyr and Phe observed in this study, potentially signifies the importance of a more extended substrate GrH specificity profile beyond P4-P1.

Differential N-Terminomics Analysis—N-terminomics in combination with SILAC (37) was performed to unravel the (differential) macromolecular substrate repertoires of GrH and mut GrC. Here the proteolytic activities of GrH and mut GrC were investigated simultaneously in either a human (K-562) or murine (YAC-1) proteome background. Overall, these analyses resulted in the identification of 6196 MS/MS spectra from 2310 peptides derived from 1029 human and 1244 mouse proteins (isoforms not taken into account). Sixty-nine cleavage sites in 67 human proteins were identified, of which 46 (67%) were cleaved by both granzymes, whereas 14 (20%) and 9 (13%) sites were unique to GrH or mut GrC respectively (supplemental Table S1). Of note, 12 cleavage sites (17%) - including 9 P1 Leu specific cleavages were previously also generated on GrM cleavage (38). Besides, 35 of the 67 GrH and/or GrC substrates (52%) reported here were previously identified to be cleaved by GrM, indicative for a shared substrate repertoire among these homologous granzymes.

In the mouse proteome, 83 cleavage sites in 75 proteins were identified. GrH generated 23 unique sites (28%), whereas 31 sites (37%) were cleaved by mut GrC, leaving 29 sites (35%) cleaved by both proteases (supplemental Table S2). Overall, 18 orthologous human and mouse substrate-pairs were identified, of which 17 were found to be cleaved at the same position(s) in both species.

Proteins playing a role in transcriptional regulation (*e.g.* PHD finger protein 6), RNA binding/processing (*e.g.* insulin-like growth factor 2 mRNA binding protein 1), microtubule dynamics (*e.g.* centrosomal protein of 170 kDa) and ribosome biogenesis (*e.g.* ribosome biogenesis regulatory protein homolog) were identified. Although cleavage of the phosphopro-



B.

	P4	P3	P2	P1	P1′	P2′	P3′	P4′
GrH	I	I/V	F/Y	Y/F	S	G/A	-	E/D

FIG. 2. **Extended substrate specificity of granzyme H by substrate phage display.** *A*, A random 9-mer phage display library was subjected to four rounds of selection using 200 nM GrH. Variable regions of 96 random phage from two independent experiments were sequenced yielding 77 unique sequences. These were manually aligned and the occurrence of each amino acid at each subsite determined (43). Residues occurring with a frequency of at least four standard deviations ($\Delta\sigma$) from random are shown. *B*, Consensus cleavage sequence for GrH derived from phage display.

tein La (previously identified as a physiological GrH substrate) cannot be identified via this approach (*i.e.* because of the size and the presence of two SCX nontolerated histidines at the GrH-introduced neo-N terminus of La ($KTKF_{364}\downarrow$ASDDEHDEHDENGATGPVKR)), the La-related protein-1 was identified in the human proteome as a sole mut GrC substrate, whereas La-related protein-7 was found in the YAC-1 background as a single GrH substrate.

Proteome-wide Substrate Specificity Profiles of GrH and Mutant GrC—The iceLogos (39) created for both proteases revealed an extended substrate specificity profile in line with previous data deduced from PS-SCL, phage display and N-terminal COFRADIC (16, 19, 20), with 86% of all human and mouse cleavage events after hydrophobic and/or aromatic amino acids (Fig. 3). For both granzymes we observed a pronounced preference for acidic residues from P3′ up to P8′, an extended specificity profile shared with granzyme B (16, 21) and which was recently shown for GrH, at least for P3′ and P4′ (20). Preliminary data obtained by proteome-wide screening of GrC and GrM substrate specificity also hinted to such an extended specificity profile (16, 38). As is the case for many proteases, small amino acids like Ala and Ser were over-represented at P1′ and also at P2′, an observation in line with phage display data (16). The preference for small P2′ residues was also apparent when comparing human and mouse GrB cleavage susceptibility; here efficiency of cleavage by mGrB increased significantly when Gly was present at P2′ (7), whereas for hGrB, other amino acids at P2′ were tolerated.

The major difference in cleavage efficiency between GrH and mut GrC was observed when a Phe was present at P1 (Fig. 4). In general mut GrC cleaved P1 Phe-containing substrates 4 to 5-fold more efficiently than GrH. Furthermore, 48% of all unique mut GrC cleavages held a Phe at P1 whereas this only occurred in 5% of the unique GrH cleavage sites identified.

Physical Property Analyses—To complement substrate sequence conservation based on amino acid identity we next analyzed aligned GrH and mut GrC cleavage motifs according to physical property constraints as described previously (21, 40), thereby emphasizing conservation by merging amino acid type preferences sharing similar physical properties. Regions with significant physical property correlations as well as deviations from the database averages of selected properties are listed in supplemental Tables S3 and S4 and summarized in Fig. 5. As also apparent from the IceLogo visualizations (Fig. 3), both enzymes highly favor small and/or polar residues at positions P1′ and P2′, whereas next to P1; P2 clearly favors the occupancy of bulky hydrophobic residues. Overall, P1 surrounding regions correlate well with hydrophilic and flexible properties, showing significant correlations encompassing stretches over 10 amino acids (Fig. 5). Classification of cleavage sites based on their structural location indicates that the

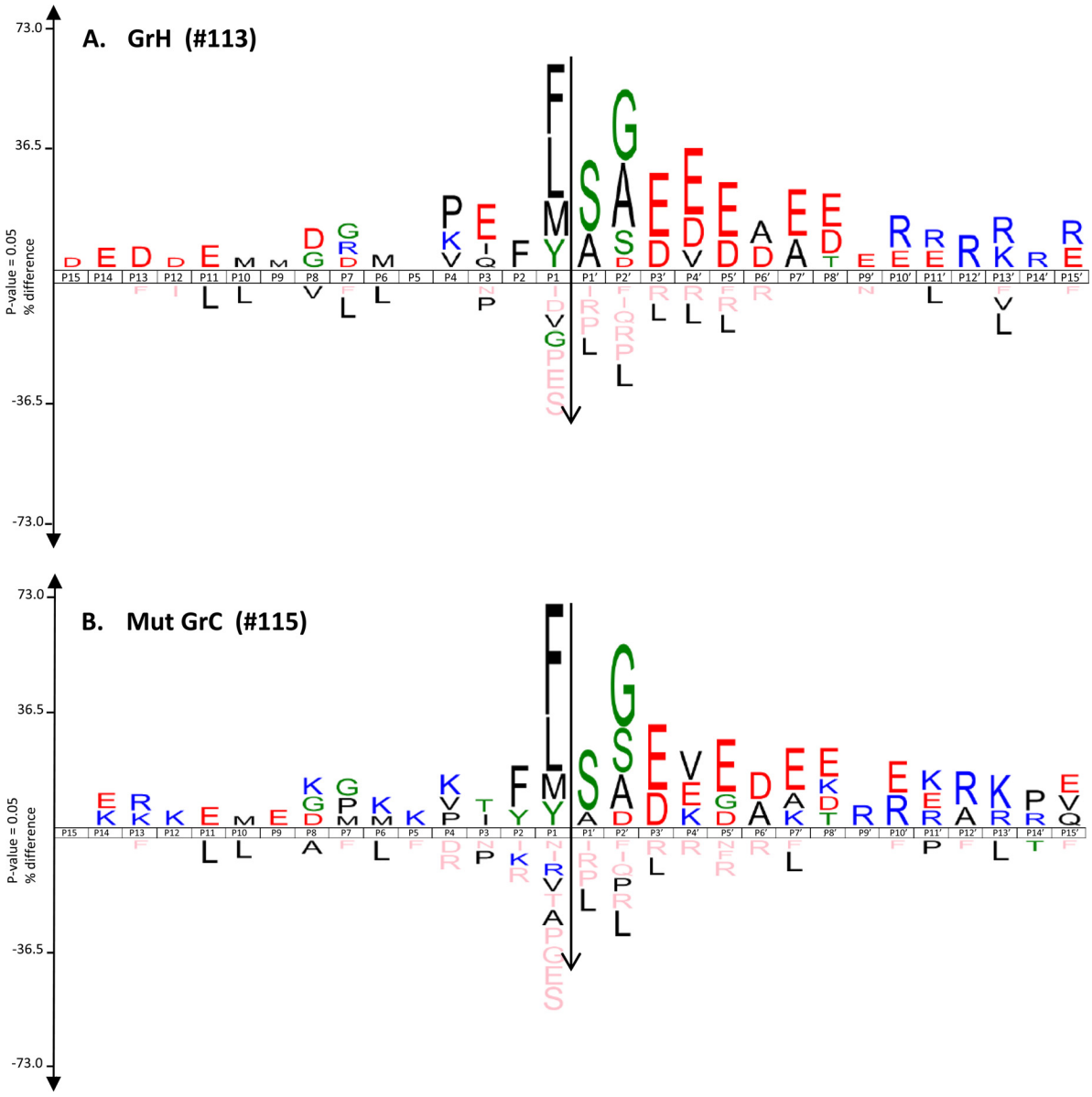


FIG. 3. **IceLogo visualization of 113 and 115 nonredundant P15-P15′ human granzyme H (*A*) and mouse mutant granzyme C (*B*) cleavage site motifs identified in the proteomes of a human (K-562) or murine (YAC-1) cell line.** Multiple sequence alignments (39) of peptide substrate motifs from P15 to P15′ are given with cleavage of the substrate occurring between P1 and P1′ (black arrow). Statistically significant residues with a *p* value threshold of 0.05 and less are plotted. The amino acid heights are indicative for their degree of conservation at the indicated position. The frequency of the amino acid occurrence at each position in sequence set was compared with the human (hGrH) or mouse (mGrC) Swissprot 2012_09 database (release date October 3, 2012). Residues that are significantly underrepresented as compared with SwissProt are plotted at the negative site of the *x*-axis (pink coloring representing the complete absence of the amino acid indicated).

majority of the cleavage sites identified is located in disordered or unstructured regions, which is compatible with extended flexible conformations that would be accessible to the protease (supplemental Tables S1 and S2). More specifically, only 9 (or 11%) of the human and 13 (or 15%) of the mouse cleavage sites are located in known or predicted 3D structures. Furthermore, analyzing compliance of individual sites with the average characteristic physical properties described in Fig. 5, shows that cleavage events found in structured elements have a much lower score as compared with the majority of cleavage sites found in disordered or unstructured regions, which implies that they do not match the flexible, unfolded extended physical property motif in agreement with their observed structural constraints (supplemental Fig. S5).

Structural Analysis of the Substrate Specificity Profiles of GrH and Mutant GrC—To further elucidate the structural protease features underlying the observed substrate specificity profiles of GrH and mut GrC, next to the available decapep-

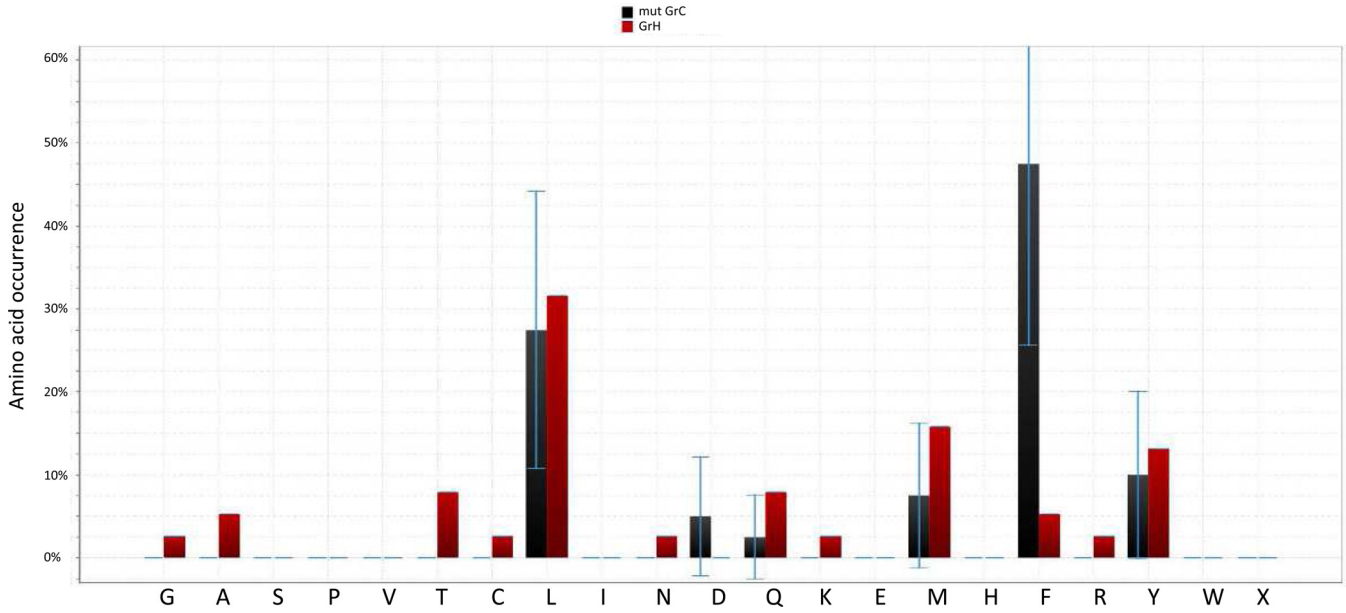


FIG. 4. **Bar chart representation of the P1 amino acid occurrence in mut GrC substrates (black bars) *versus* GrH substrates (red bars).** The error bars indicate the variation that can be expected from random sampling as compared with the reference of amino acid occurrence in the SwissProt 2012_09 database.

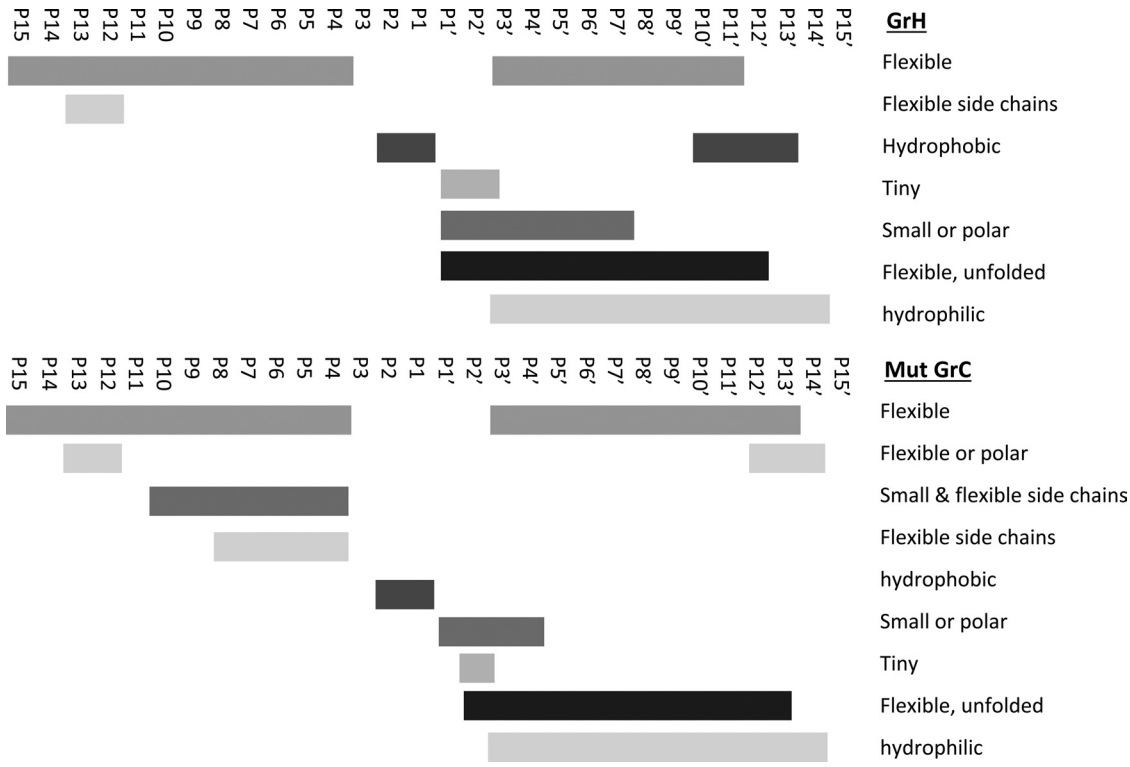


FIG. 5. **Overview of physical properties preferred by GrH and mut GrC from P15-P15′.**

tide (PTSYAGDDS) bound structure of hGrH (Protein Data Bank (PDB) code 3tjv) (20), a conservative model for ligand-bound unlocked GrC was created. Viewing the fact that several residues forming the substrate binding pocket are not resolved in the available crystal structure of unlocked GrC (PDB:3g01) (16), we created a model based on the crystal structure of the locked conformation (PDB:3fzz) (16) after mutating to the unlocked form (EE to RG double mutant). Relating the conformation of the template ligand of GrH by structural superimposition onto the unlocked model of GrC,

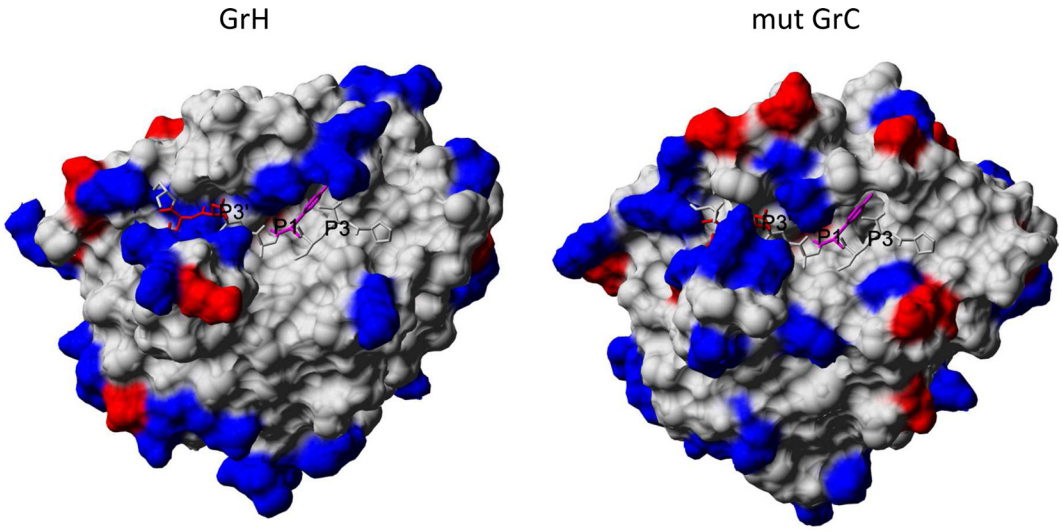


FIG. 6. **Crystal structure of GrH (*left*) and model of mut GrC (*right*).** *Left* side: crystal structure of GrH (PDB:3TJV) with bound substrate peptide PTSYAGDDS. *Right* side: model of "unlocked" mut GrC substrate recognition based on crystal structure of "locked" wt GrC (PDB:3FFZ) with double mutation of the EE motif to an "unlocked" RG configuration and superimposition of substrate ligand with same conformation as in GrH, followed by short simulated annealing molecular dynamics simulations of enzyme residues within 5 Ångström from the ligand for energy minimization using the AMBER03 force field in YASARA (41). Coloring: Asp, Glu… red; Arg, Lys = blue, P1 residue… magenta).

followed by short simulated annealing molecular dynamics simulations of residues within 5 Ångström from the peptidic ligand for energy minimization using the AMBER03 force field in YASARA (41), resulted in a model for peptide ligand-bound unlocked version of mut GrC. As a consequence of energy minimization in the presence of ligand, the model shows opening of the initial loop causing the lock, thereby returning the specificity region back into register with the GrH conformation (supplemental Fig. S3). In both cases, the bound substrate confirms the observed size restrictions inferred by the binding pocket constraints at P1′ and P2′. Besides, the observed preferences of acidic amino acids at P3′ and P4′ and the presence of larger residues at P2 are well supported by respectively the surrounding positive charges (blue residues in Fig. 6) and the fact that P2 is positioned on top of the open part of the binding pocket, not restricted in space.

Further, the crystal and modeled structure predicts a number of differences in the extended substrate binding cleft, which may explain the increased tolerance of positively charged lysines at P4-P6 by mut GrC as compared with GrH, viewing the exclusive availability of a negative surface charge in this region in mut GrC (red patch in lower right corner of mut GrC; Fig. 6) and the increased preference for acidic residues at P4′ and P3 because of the complementary charge differences at the enzyme surface lining the binding pocket (Fig. 3 and 6), all indicative for differences in the extended substrate binding cleft that could be causative for the (subtle) differences in substrate interactions observed.

The Extended Specificity Profile is a Major Determinant for Granzyme Substrate Recognition—Western blot analysis was performed to validate substrate cleavage of overexpressed Flap endonuclease 1 (FEN-1). Human and mouse FEN-1 were found to be cleaved by GrH and mut GrC in both differential degradomics analyses, generating an identical neo-N terminus ($_{311}$CGEKQFSEER$_{320}$) though more efficiently cleaved by mut GrC (K-562 GrH/mut GrC ratio of 0.38, YAC-1 GrH/mut GrC ratio of 0.22), and moreover harboring acidic residues at positions P3′, P8′ and P9′ (supplemental Table S1). Fig. 7*A* shows the MS-spectrum of $_{2}$GIQGLAKLIADVAPSAIR$_{19}$ identified as the protein N terminus of human FEN-1 with overall about equal intensities in the three samples (1:0.9:0.8), whereas the neo-N terminus was exclusively present in the granzyme-treated samples with a GrH/mut GrC ratio of 0.38 and 0.22 in the human (Fig. 7*B*) and mouse proteome background respectively. C-terminally GFP tagged FEN-1 and its presumed GrH/GrC truncated form were overexpressed in K-562 cells. Cell lysates overexpressing full-length GFP-FEN1 and incubated with 200 nM mut GrC (Fig. 8) revealed a fragment of equal molecular weight as the truncated GFP-FEN-1; indicative for the fact that mut GrC cleavage occurred at IKFM$_{310}$.

Next, we investigated the influence of P1 cleavage susceptibility on the efficiency of GrH/mut GrC cleavage by mutation of the FEN-1 P1 residue from Met to Phe or Tyr. Using site-directed mutagenesis PCR, the FEN-1 P1 methionine (M$_{310}$) was mutated to Tyr (M$_{310}$>Y) and Phe (M$_{310}$>F) and the wild-type and mutant protein variants were *in vitro* transcribed and translated in the presence of ^{35}S-Met to investigate GrH/mut GrC cleavage profiles by autoradiography. Cleavage was assessed by incubating the radiolabeled protein with the different granzymes at various concentrations. Similar to the Western blot data (Fig. 8), cleavage of wt FEN-1

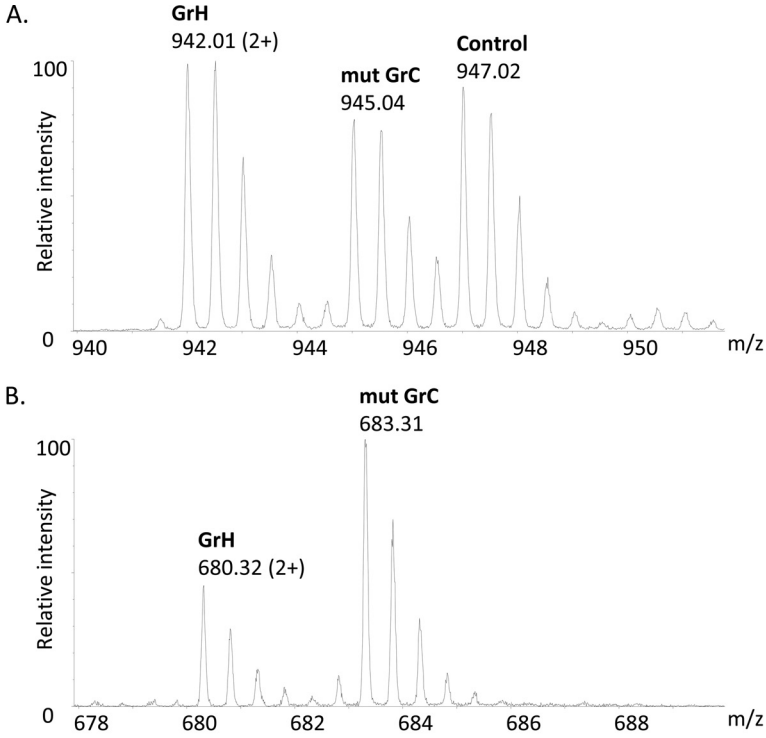


FIG. 7. **Representative MS-spectra of FEN-1.** The amino acid sequence of the Flap endonuclease 1 protein is shown at the bottom. *A*, The peptide corresponding to the protein N terminus of FEN-1 (underlined) was identified as a peptide ion triplet, implying its recovery from all setups analyzed with about equal intensities (corrected ratios of 1:0.9:0.8). *B*, A peptide ion couple separated by 6 Da, representative for their presence in the setups incubated with GrH ($^{12}C_6$) or mut GrC ($^{13}C_6$) and not in the control setup, was identified as $_{311}$CGEKQFSEER$_{320}$ (bold in protein sequence) and assigned to FEN-1 cleavage after IKFM$_{310}$ (italic)with a GrH/mut GrC ratio of 0.38.

MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQGGDVLQNEEGET
TSHLMGMFYRTIRMMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQAQAAGAE
QEVEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEASCAALVKAGKVYAAATEDM
DCLTFGSPVLMRHLTASEAKKLPIQEFHLSRILQELGLNQEQFVDLCILLGSDYCESIRG
IGPKRAVDLIQKHKSIEEIVRRLDPNKYPVPENWLHKEAHQLFLEPEVLDPESVELKWSE
PNEEEL*IKFM***CGEKQFSEER**IRSGVKRLSKSRQGSTQGRLDDFFKVTGSLSSAKRKEPEP
KGSTKKKAKTGAAGKFKRGK

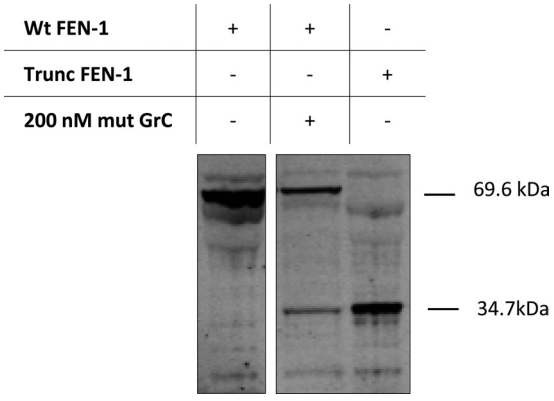


FIG. 8. **Assessment of *in vitro* FEN-1 cleavage by mut GrC.** Western blot analyses on K-562 cell lysates overexpressing full length or truncated FEN-1. Lysates expressing full length FEN-1 served either as a control or were incubated with iodoacetamide and mut GrC.

(~43 kDa) by GrH and mut GrC resulted in generation of a 35 kDa fragment (black arrow), again indicative for cleavage at IKFM$_{310}$ (Fig. 9*A*). The wild-type FEN-1 precursor was cleaved more efficiently by mut GrC as compared with GrH in line with the COFRADIC data. Mutating M$_{310}$>Y did not drastically affect cleavage susceptibility by GrH and mut GrC. However, for mut GrC a small increase in cleavage efficiency could be observed when mutating M$_{310}$>F, and a decrease in GrH cleavage susceptibility was observed, overall correlating with the substrate specificity data obtained using N-terminal COFRADIC (Fig. 4 and supplemental Table S1).

We further observed a shared extended specificity profile between GrB, GrM, GrH and mut GrC comprising a prevalence of acidic residues from P3′-P8′, which was not observed for GrA (Fig. 3 and supplemental Fig. S4) (42). We therefore mutagenized FEN-1 to investigate whether the extended specificity profiles are sufficient to steer cleavage by specific granzymes based on an optimal P1 residue. To investigate whether GrB, processing wt FEN1 at distinct sites (Fig. 9*B*, dashed and gray arrows), can cleave in an otherwise GrH/mut GrC cleavage context, a P1 M$_{310}$>D mutant was assayed for granzyme B cleavage susceptibility. Remarkably, GrB generated an additional fragment with an equal molecular weight as compared with the GrC/GrH M$_{310}$ cleavage product, accounting for about 33% of the precursor cleavage observed. This indicates that the presence of a P1 Asp residue in the extended substrate specificity profile of an otherwise GrH/mut GrC-specific cleavage event suffices to confer

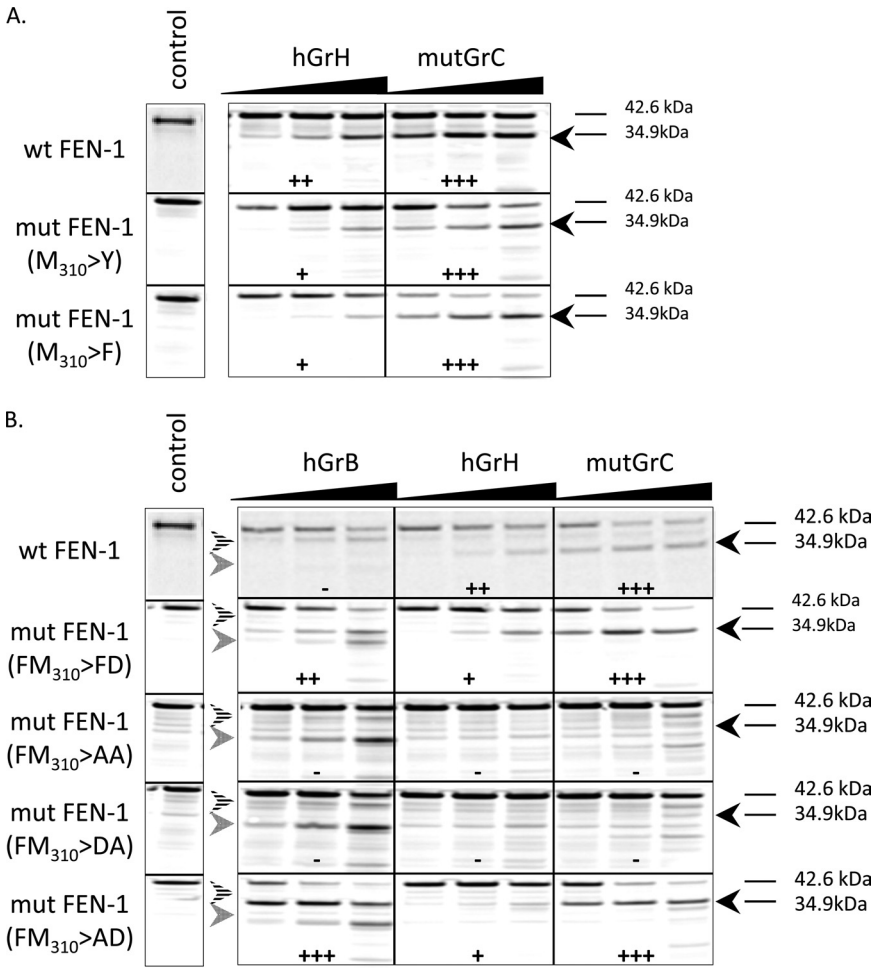


FIG. 9. **In vitro cleavage of FEN-1 and FEN-1 mutants by GrB, GrH or mut GrC.** *A*, Autoradiographs of wild type FEN-1 and P1 mutagenized FEN-1 variants (M_{310}>Y and M_{310}>F), incubated with various concentrations of human GrH or mouse mut GrC. *B*, Autoradiographs showing cleavage of P1 mutagenized M_{310}>D and in addition P2 and P1 mutagenized (FM_{310}>AA, FM_{310}>DA and FM_{310}>AD) FEN-1 variants by human GrB, GrH and mouse mut GrC. The degree of cleavage at the GrH/mut GrC cleavage site is indicated by "-" and "+," with "+++" being cleaved most efficiently. The cleavage fragment generated upon GrH/mut GrC cleavage is indicated with a black arrow, whereas GrB cleavage fragments are indicated with dashed and gray arrows.

GrB cleavage susceptibility. Unexpectedly, the P1 M_{310}>D mutation did not abolish cleavage by GrH or mut GrC and still resulted in respectively 42 and 68% of the FEN-1 precursor being cleaved.

To investigate if cleavage at P2 (Phe_{309}) could account for this, an $IKA_{309}A_{310}$ double mutant was created and shown to resist cleavage by either GrH or mut GrC. When mutating the P2-P1 residues FM_{310} to DA to investigate if GrB cleavage can occur at position 309 similar to cleavage after Phe_{309} by GrH and mut GrC, an $IKAA_{310}$-like pattern could be observed (*i.e.* no cleavage at position 309 or 310), whereas analysis of the $IKAD_{310}$ mutant resulted in 41 and 71% of cleavage by GrB and mut GrC respectively. These results suggest that besides the P1 identity, the extended specificity profile is a key determinant for recognition by GrB, GrH, and mut GrC, and indicates that cleavages are predominantly confined to restricted protein areas.

DISCUSSION

In this study we investigated the postulated orthologous relationship between human granzyme H (GrH) and murine granzyme C (GrC) and whether they may exert similar functions. Therefore we complemented their high sequence similarities, sequence identities, highly similar chromosomal locations and overlapping expression profiles with information on their cytotoxicity and their substrate repertoires and specificities. By means of dose-response analyses, GrH (this study) and mutant unlocked GrC (7) were found to be very inefficient cytotoxins as compared with GrB because the amount of GrH and mut GrC needed to kill 50% of granzyme targeted cells is ~170-fold and up to 1000-fold higher as compared with GrB respectively. Next, substrate phage display with GrH was performed as described previously for assessing GrC subsite requirements (16). Whereas P4 Ile and P2′ Gly preferences in GrH and mut GrC cleavage sites are similar, the biggest differences were observed at the P2 and P1 positions. Preferences for Tyr at P2 are quite similar whereas cleavage sites harboring P2 Phe residues are ~5-fold more efficiently cleaved by GrC as compared with GrH. At P1 preferences for Tyr and Phe could be observed for both orthologous granzymes, with a Phe/Tyr ratio of 1/4 in GrH cleavage sites, whereas the inverse ratio for these amino acids was observed in GrC cleaved sites. N-terminal COFRADIC, allowing direct comparison of both granzymes, was applied as to examine whether the by phage display

determined subsite features also apply on the macromolecular level. N-terminal COFRADIC results obtained by comparing wild type GrC and an "unlocked" mut GrC previously resulted in the identification of 18 mut GrC substrates, with six of these found to be cleaved by both GrC variants although mut GrC cleaved these 10-fold more efficient (16). Among the substrates identified in this study, all 6 wt GrC substrates were identified. In accordance to phage display data, mut GrC cleaved P1 Phe-containing substrates four- to fivefold more efficiently than GrH. We furthermore observe that cleavage sites harboring either Phe or Tyr at P2 are cleaved more efficiently by mut GrC. Although these differences in cleavage efficiency are only subtle, substrate specificity profiles derived by phage display or positional proteomics gave highly similar, but none-identical patterns, most likely due to the process of phage display selection, specifically enriching for higher affinity substrate peptides through multiple rounds of selection and amplification. By contrast, the positional proteomics strategy used identifies cleavage events in their macromolecular context to provide a more all-encompassing pattern of potential substrates, though not restricted to the most efficiently cleaved substrates (5). Since the substrate specificity profiles and thus biochemical protease characterization were obtained using extensive *in vitro* degradome analyses, extrapolation to the functional implication and physiological importance of substrate cleavage requires caution. However, using the DAVID Functional Annotation Clustering Tool (7, 8), two annotation clusters encompassing ribosomal proteins and helicases showed group enrichment scores of $p \leq 0.01$ when compared with the experimentally derived mouse and human proteome backgrounds respectively (*i.e.* respectively 826 and 654 unique mouse and human proteins based on the identification of their database annotated protein N termini). Interestingly, representatives of these categories of substrates have previously been found to be targeted by other granzymes under physiological conditions (9, 10), which might hint to the fact that similar to hnRNPK for instance, other hitherto unidentified substrates might represent pan-granzyme substrates.

In addition to these subtle subsite differences surrounding the P1 residue, a shared extended specificity profile for acidic residues was found, similar to former analyses on human GrB and GrM substrate motifs (21, 38). We here showed that this extended protease-substrate interaction can steer protein cleavage because cleavage by GrB can be conferred by a sole mutation at P1 into GrB's preferred P1 aspartic acid residue. These observations suggest granzyme cleavage sites to reside in restricted substrate regions for the aforementioned granzymes, whereas for hGrA (and other proteases) such an extended specificity-profile has not been found (42). As such, the preferential targeting of acidic substrate regions may have exerted a selective force to conserve substrate recognition among specific granzymes and thus to target specific key substrates. On the other hand, investigation of cleavage site locations in macromolecular substrates reveals that cleavage events found in structures appear accessible at the protein surface but can sometimes be found within or at the border of secondary structure elements, as is the case for the FEN-1 $IKFM_{310}$ cleavage site. Although rare in our case, in line with the previously reported proteolytic cleavage sites residing in structural elements (*i.e.* α-helices) of proteins with resolved structures (11), we suggest an explanation through the dynamic nature of proteins, in which presumably stable structures exist in a thermodynamic equilibrium with a fraction of (partially) unfolded states, or by the structural rearrangement of the substrate and/or protease active site, as disorder/order transitions have been demonstrated on protein complex formation (12).

Furthermore, the significant overlap of cleavage sites identified (50%) as compared with another pair of granzyme orthologs; *i.e.* hGrB and mGrB (21)(47%), are further indicative for a potential functional homology of GrH and GrC.

Of note is the fact that all these analyses were conducted making use of an unlocked, mutant GrC variant because GrC activation was suggested to require an allosteric control mechanism (16). Unlocking GrC by mutating those residues to their GrB counterparts, activates GrC and potentiates the killing potential of mut GrC fourfold.

As the proteomics data suggest similar substrate preferences between GrH and mut GrC, the original crystal conformation of the GrH substrate PTSYAGDDS superimposed into mut GrC as example ligand could be used to interpret relative positions of the surface charges to the substrate positions as well as binding pocket space constraints. Because based on the structural alignment no register shift for GrH can be observed (supplemental Fig. S3), it is difficult to predict whether because of evolution an allosteric control mechanism for only one of these proteases was preserved/has evolved (thus making GrC a strictly controlled protease for specialized functions), more so viewing the fact that granzyme paralogs could have diversified or taken over (parts) of GrH/C function.

In conclusion, we provide further evidence to assign GrH and mut GrC as orthologous granzymes exerting similar functions because they both were found to be poor cytotoxins and they target their substrates making use of similar specificity profiles. In addition, we showed that the extended specificity profile is a critical determinant among several granzymes for cleavage to occur because mutation of the P1 methionine into the GrB preferable aspartic acid infers GrB cleavage susceptibility and furthermore does not abolish GrH and mut GrC cleavage because a neighboring Phe residue together with the overlapping extended specificity profile of the substrate can still be recognized by mut GrC and GrH as such.

* K.P. is supported by a PhD grant from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). P.V.D. is a Postdoctoral Fellow of the Research Foundation - Flanders (FWO-Vlaanderen). J.A., D.K. and P.I.B. are supported by Program Grant #490900 from the National Health and Medical Research Council (Australia).

S This article contains supplemental Figs. S1 to S5 and Tables S1 to S4.

§§ These authors contributed equally to this work.

‡‡ To whom correspondence should be addressed: Department of Medical Protein Research, Flanders Interuniversity Institute for Biotechnology, Ghent University, A. Baertsoenkaai 3, B9000 Ghent, Belgium. Tel.: +32-92649279; Fax: +32-92649496; E-mail: petra.vandamme@vib-ugent.be.

REFERENCES

1. Cai, S. F., Fehniger, T. A., Cao, X., Mayer, J. C., Brune, J. D., French, A. R., and Ley, T. J. (2009) Differential expression of granzyme B and C in murine cytotoxic lymphocytes. *J. Immunol.* **182,** 6287–6297
2. Sedelies, K. A., Sayers, T. J., Edwards, K. M., Chen, W., Pellicci, D. G., Godfrey, D. I., and Trapani, J. A. (2004) Discordant regulation of granzyme H and granzyme B expression in human lymphocytes. *J. Biol. Chem.* **279,** 26581–26587
3. Afonina, I. S., Cullen, S. P., and Martin, S. J. (2010) Cytotoxic and non-cytotoxic roles of the CTL/NK protease granzyme B. *Immunol. Rev.* **235,** 105–116
4. Revell, P. A., Grossman, W. J., Thomas, D. A., Cao, X., Behl, R., Ratner, J. A., Lu, Z. H., and Ley, T. J. (2005) Granzyme B and the downstream granzymes C and/or F are important for cytotoxic lymphocyte functions. *J. Immunol.* **174,** 2124–2131
5. Johnson, H., Scorrano, L., Korsmeyer, S. J., and Ley, T. J. (2003) Cell death induced by granzyme C. *Blood* **101,** 3093–3101
6. Shi, L., Wu, L., Wang, S., and Fan, Z. (2009) Granzyme F induces a novel death pathway characterized by Bid-independent cytochrome c release without caspase activation. *Cell Death Differ.* **16,** 1694–1706
7. Kaiserman, D., Bird, C. H., Sun, J., Matthews, A., Ung, K., Whisstock, J. C., Thompson, P. E., Trapani, J. A., and Bird, P. I. (2006) The major human and mouse granzymes are structurally and functionally divergent. *J. Cell Biol.* **175,** 619–630
8. Jenkins, M. R., Trapani, J. A., Doherty, P. C., and Turner, S. J. (2008) Granzyme K expressing cytotoxic T lymphocytes protects against influenza virus in granzyme AB−/− mice. *Viral Immunol.* **21,** 341–346
9. Joeckel, L. T., Wallich, R., Martin, P., Sanchez-Martinez, D., Weber, F. C., Martin, S. F., Borner, C., Pardo, J., Froelich, C., and Simon, M. M. (2011) Mouse granzyme K has pro-inflammatory potential. *Cell Death Differ.* **18,** 1112–1119
10. Metkar, S. S., Menaa, C., Pardo, J., Wang, B., Wallich, R., Freudenberg, M., Kim, S., Raja, S. M., Shi, L., Simon, M. M., and Froelich, C. J. (2008) Human and mouse granzyme A induce a proinflammatory cytokine response. *Immunity* **29,** 720–733
11. Anthony, D. A., Andrews, D. M., Chow, M., Watt, S. V., House, C., Akira, S., Bird, P. I., Trapani, J. A., and Smyth, M. J. (2010) A role for granzyme M in TLR4-driven inflammation and endotoxicosis. *J. Immunol.* **185,** 1794–1803
12. Fellows, E., Gil-Parrado, S., Jenne, D. E., and Kurschus, F. C. (2007) Natural killer cell-derived human granzyme H induces an alternative, caspase-independent cell-death program. *Blood* **110,** 544–552
13. Hou, Q., Zhao, T., Zhang, H., Lu, H., Zhang, Q., Sun, L., and Fan, Z. (2008) Granzyme H induces apoptosis of target tumor cells characterized by DNA fragmentation and Bid-dependent mitochondrial damage. *Mol. Immunol.* **45,** 1044–1055
14. Susanto, O., Trapani, J. A., and Brasacchio, D. (2012) Controversies in granzyme biology. *Tissue Antigens* **80,** 477–487
15. Getachew, Y., Stout-Delgado, H., Miller, B. C., and Thiele, D. L. (2008) Granzyme C supports efficient CTL-mediated killing late in primary alloimmune responses. *J. Immunol.* **181,** 7810–7817
16. Kaiserman, D., Buckle, A. M., Van Damme, P., Irving, J. A., Law, R. H., Matthews, A. Y., Bashtannyk-Puhalovich, T., Langendorf, C., Thompson, P., Vandekerckhove, J., Gevaert, K., Whisstock, J. C., and Bird, P. I. (2009) Structure of granzyme C reveals an unusual mechanism of protease autoinhibition. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 5587–5592
17. Wu, L., Wang, L., Hua, G., Liu, K., Yang, X., Zhai, Y., Bartlam, M., Sun, F., and Fan, Z. (2009) Structural basis for proteolytic specificity of the human apoptosis-inducing granzyme M. *J. Immunol.* **183,** 421–429
18. Edwards, K. M., Kam, C. M., Powers, J. C., and Trapani, J. A. (1999) The human cytotoxic T cell granule serine protease granzyme H has chymotrypsin-like (chymase) activity and is taken up into cytoplasmic vesicles reminiscent of granzyme B-containing endosomes. *J. Biol. Chem.* **274,** 30468–30473
19. Mahrus, S., and Craik, C. S. (2005) Selective chemical functional probes of granzymes A and B reveal granzyme B is a major effector of natural killer cell-mediated lysis of target cells. *Chem. Biol.* **12,** 567–577
20. Wang, L., Zhang, K., Wu, L., Liu, S., Zhang, H., Zhou, Q., Tong, L., Sun, F., and Fan, Z. (2012) Structural insights into the substrate specificity of human granzyme H: the functional roles of a novel RKR motif. *J. Immunol.* **188,** 765–773
21. Van Damme, P., Maurer-Stroh, S., Plasman, K., Van Durme, J., Colaert, N., Timmerman, E., De Bock, P. J., Goethals, M., Rousseau, F., Schymkowitz, J., Vandekerckhove, J., and Gevaert, K. (2009) Analysis of protein processing by N-terminal proteomics reveals novel species-specific substrate determinants of granzyme B orthologs. *Mol. Cell Proteomics* **8,** 258–272
22. Sun, J., Whisstock, J. C., Harriott, P., Walker, B., Novak, A., Thompson, P. E., Smith, A. I., and Bird, P. I. (2001) Importance of the P4′ residue in human granzyme B inhibitors and substrates revealed by scanning mutagenesis of the proteinase inhibitor 9 reactive center loop. *J. Biol. Chem.* **276,** 15177–15184
23. Romero, V., Fellows, E., Jenne, D. E., and Andrade, F. (2009) Cleavage of La protein by granzyme H induces cytoplasmic translocation and interferes with La-mediated HCV-IRES translational activity. *Cell Death Differ.* **16,** 340–348
24. Andrade, F., Fellows, E., Jenne, D. E., Rosen, A., and Young, C. S. (2007) Granzyme H destroys the function of critical adenoviral proteins required for viral DNA replication and granzyme B inhibition. *EMBO J.* **26,** 2148–2157
25. Sun, J., Bird, C. H., Thia, K. Y., Matthews, A. Y., Trapani, J. A., and Bird, P. I. (2004) Granzyme B encoded by the commonly occurring human RAH allele retains pro-apoptotic activity. *J. Biol. Chem.* **279,** 16907–16911
26. Sun, J., Bird, C. H., Buzza, M. S., McKee, K. E., Whisstock, J. C., and Bird, P. I. (1999) Expression and purification of recombinant human granzyme B from Pichia pastoris. *Biochem. Biophys. Res. Commun.* **261,** 251–255
27. Bird, C. H., Sun, J., Ung, K., Karambalis, D., Whisstock, J. C., Trapani, J. A., and Bird, P. I. (2005) Cationic sites on granzyme B contribute to cytotoxicity by promoting its uptake into target cells. *Mol. Cell Biol.* **25,** 7854–7867
28. Staes, A., Van Damme, P., Helsens, K., Demol, H., Vandekerckhove, J., and Gevaert, K. (2008) Improved recovery of proteome-informative, protein N-terminal peptides by combined fractional diagonal chromatography (COFRADIC). *Proteomics* **8,** 1362–1370
29. Ghesquière, B., Van Damme, J., Martens, L., Vandekerckhove, J., and Gevaert, K. (2006) Proteome-wide characterization of N-glycosylation events by diagonal chromatography. *J. Proteome Res.* **5,** 2438–2447
30. Martens, L., Vandekerckhove, J., and Gevaert, K. (2005) DBToolkit: processing protein databases for peptide-centric proteomics. *Bioinformatics* **21,** 3584–3585
31. Van Damme, P., Martens, L., Van Damme, J., Hugelier, K., Staes, A., Vandekerckhove, J., and Gevaert, K. (2005) Caspase-specific and non-specific in vivo protein processing during Fas-induced apoptosis. *Nat. Methods* **2,** 771–777
32. Vande Walle, L., Van Damme, P., Lamkanfi, M., Saelens, X., Vandekerckhove, J., Gevaert, K., and Vandenabeele, P. (2007) Proteome-wide Identification of HtrA2/Omi Substrates. *J. Proteome Res.* **6,** 1006–1015
33. Huber, P. J., Ronchetti, E. M. (2009) *Robust Statistics*, Second Ed., Wiley, John & Sons, Incorporated.
34. Demon, D., Van Damme, P., Vanden Berghe, T., Deceuninck, A., Van Durme, J., Verspurten, J., Helsens, K., Impens, F., Wejda, M., Schymkowitz, J., Rousseau, F., Madder, A., Vandekerckhove, J., Declercq, W., Gevaert, K., and Vandenabeele, P. (2009) Proteome-wide substrate analysis indicates substrate exclusion as a mechanism to generate caspase-7 versus caspase-3 specificity. *Mol. Cell Proteomics* **8,** 2700–2714
35. Maurer-Stroh, S., Debulpaep, M., Kuemmerer, N., Lopez de la Paz, M., Martins, I. C., Reumers, J., Morris, K. L., Copland, A., Serpell, L., Serrano,

L., Schymkowitz, J. W., and Rousseau, F. (2010) Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nat. Methods* **7,** 237–242
36. Van Damme, P., Vandekerckhove, J., and Gevaert, K. (2008) Disentanglement of protease substrate repertoires. *Biol. Chem.* **389,** 371–381
37. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell Proteomics* **1,** 376–386
38. de Poot, S. A., Westgeest, M., Hostetter, D. R., Van Damme, P., Plasman, K., Demeyer, K., Broekhuizen, R., Gevaert, K., Craik, C. S., and Bovenschen, N. (2011) Human and mouse granzyme M display divergent and species-specific substrate specificities. *Biochem. J.* **437,** 431–442
39. Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009) Improved visualization of protein consensus sequences by iceLogo. *Nat. Methods* **6,** 786–787
40. Maurer-Stroh, S., and Eisenhaber, F. (2005) Refinement and prediction of protein prenylation motifs. *Genome Biol.* **6,** R55
41. Krieger, E., Joo, K., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., and Karplus, K. (2009) Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* **77,** 114–122
42. Van Damme, P., Maurer-Stroh, S., Hao, H., Colaert, N., Timmerman, E., Eisenhaber, F., Vandekerckhove, J., and Gevaert, K. (2010) The substrate specificity profile of human granzyme A. *Biol. Chem.* **391,** 983–997
43. Matthews, D. J., Goodman, L. J., Gorman, C. M., and Wells, J. A. (1994) A survey of furin substrate specificity using substrate phage display. *Protein Sci.* **3,** 1197–1205