THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 292, NO. 7, pp. 2690–2702, February 17, 2017

Overlapping Specificity of Duplicated Human Pancreatic Elastase 3 Isoforms and Archetypal Porcine Elastase 1 Provides Clues to Evolution of Digestive Enzymes*

Received for publication, November 30, 2016, and in revised form, January 1, 2017 Published, JBC Papers in Press, January 6, 2017, DOI 10.1074/jbc.M116.770560

Eszter Boros[‡§], **András Szabó**[§], **Katalin Zboray**[‡], **Dávid Héja**[‡], **Gábor Pál**[‡1,2], **and Miklós Sahin-Tóth**[§¶1,3]

From the [‡]*Department of Biochemistry, Eötvös Loránd University, Budapest 1117, Hungary and the* [¶]*Center for Exocrine Disorders,* [§]*Department of Molecular and Cell Biology, Boston University Henry M. Goldman School of Dental Medicine, Boston, Massachusetts 02118*

Edited by George N. DeMartino

Chymotrypsin-like elastases (CELAs) are pancreatic serine proteinases that digest dietary proteins. CELAs are typically expressed in multiple isoforms that can vary among different species. The human pancreas does not express CELA1 but secretes two CELA3 isoforms, CELA3A and CELA3B. The reasons for the CELA3 duplication and the substrate preferences of the duplicated isoforms are unclear. Here, we tested whether CELA3A and CELA3B evolved unique substrate specificities to compensate for the loss of CELA1. We constructed a phage library displaying variants of the substrate-like *Schistocerca gregaria* proteinase inhibitor 2 (SGPI-2) to select reversible high affinity inhibitors of human CELA3A, CELA3B, and porcine CELA1. Based on the reactive loop sequences of the phage display-selected inhibitors, we recombinantly expressed and purified 12 SGPI-2 variants and determined their binding affinities. We found that the primary specificity of CELA3A, CELA3B, and CELA1 was similar; all preferred aliphatic side chains at the so-called P1 position, the amino acid residue located directly N-terminal to the scissile peptide bond. P1 Met was an interesting exception that was preferred by CELA1 but weakly recognized by the CELA3 isoforms. The extended substrate specificity of CELA3A and CELA3B was comparable, whereas CELA1 exhibited unique interactions at several subsites. These observations indicated that the CELA1 and CELA3 paralogs have some different but also overlapping specificities and that the duplicated CELA3A and CELA3B isoforms did not evolve distinct substrate preferences. Thus, increased gene dosage rather than specificity divergence of the CELA3 isoforms may compensate for the loss of CELA1 digestive activity in the human pancreas.

Digestive serine proteinases are secreted by the pancreas and include trypsins, chymotrypsins, and elastases, which cleave dietary protein substrates with unique substrate specificities primarily determined by the P1 amino acid forming the scissile peptide bond (Schechter and Berger nomenclature of interactions between proteinases and their substrates (1)). Pancreatic elastases were recently renamed chymotrypsin-like elastases to avoid confusion with the homologous neutrophil elastase. Digestive enzymes are typically expressed in multiple isoforms, and the number and properties of these isoenzymes can vary significantly among the different vertebrate species making clear assignment of orthology challenging at best. The advent of the genomic era facilitated the comprehensive identification and annotation of digestive enzyme isoforms; however, functional studies have been slow to keep up with the omics revolution, and the physiological significance of digestive proteinase diversity has remained relatively poorly understood even in humans.

Elastases were first defined by their ability to solubilize insoluble elastin by Hungarian scientists Baló and Banga in 1949 (2), who demonstrated the presence of such elastolytic activity in pancreatic homogenates of the pig (3). Subsequent purification of porcine elastase 1 (CELA1) allowed unambiguous assignment of the elastolytic activity to this serine proteinase (4–6). CELA1 was shown to adsorb to the surface of elastin because of its basic character and then to cleave elastin at multiple Ala-Ala and Ala-Gly peptide bonds (7–9). The P1 Ala specificity was confirmed using various peptide and protein substrates; furthermore, cleavages after P1 Val, Leu, Ile, Gly, and Ser were also documented indicating that CELA1 prefers P1 amino acids with smaller aliphatic side chains (5, 6, 8–15). Finally, X-ray crystallography demonstrated that the specificity pocket of porcine CELA1 was restricted at the mouth by Val-216 (chymotrypsin numbering) and partially occluded at the bottom by Thr-226, which explained the characteristic P1 preference and the exclusion of aromatic side chains (16). Curiously, the ortholog of porcine CELA1 is not expressed in the human pancreas due to evolutionary mutations in the 5′ promoter and enhancer regions (17, 18).

A second proteinase with elastolytic activity, named elastase 2 (CELA2), was isolated from the pig (8, 19–21) and human pancreas (23, 24) and was subsequently cloned from the human (25–27), rat (28), and bovine pancreas (29). Surprisingly, when

* This work was supported by National Institutes of Health Grants R01DK095753, R01DK082412, and R01DK058088 (to M. S. T.), National Research, Development and Innovation Office (NKFIH) Grants NK100769 and K119386, National Development Agency Grant KMOP-4.2.1/B-10-2011, and the MedInProt Program of the Hungarian Academy of Sciences (to G. P.). The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

[1] Both authors contributed equally to this work.
[2] To whom correspondence may be addressed: Pázmány Péter Sétány 1/C, H-1117 Budapest, Hungary. E-mail: gabor.pal@ttk.elte.hu.
[3] To whom correspondence may be addressed: 72 East Concord St., Evans-433, Boston, MA 02118. Tel.: 617-414-1070; E-mail: miklos@bu.edu.

tested on elastin (8, 9), oxidized insulin (30), or *p*-nitroanilide peptide substrates (31, 32), CELA2 exhibited chymotrypsin-like primary substrate specificity and cleaved these substrates after Leu, Met, Phe, and Tyr rather than Ala or Val residues. Relative to chymotrypsins, however, the catalytic efficiency was about 100-fold lower when measured on *p*-nitroanilide substrates (32). The chymotrypsin-like substrate specificity was consistent with the presence of Gly-216 and Ser-226 in the specificity pocket of CELA2. Cloning of the human cDNA resulted in the identification of two highly similar genes, named *CELA2A* and *CELA2B*, indicating recent gene duplication (27). Although expressed at the mRNA level, the CELA2B protein has never been isolated or detected. When expressed recombinantly, human CELA2B did not function as an active proteinase as a result of natural evolution-driven mutations absent in the functional CELA2A enzyme (33).

A third major human pancreatic elastase was isolated by multiple groups independently and named proteinase E (or protease E) (34), elastase 1 (23), and cholesterol-binding protein (36, 37). N-terminal protein sequencing and cDNA cloning confirmed the identity of these proteinases of acidic character as elastase 3 (CELA3) and also revealed that the human gene underwent duplication giving rise to *CELA3A* and *CELA3B* (38–40). Both isoforms are expressed at the mRNA and protein levels to a comparable extent (39, 40). Human CELA3B was found to have poor elastolytic activity but readily degraded casein and hemoglobin and hydrolyzed small ester and amide substrates with a P1 Ala (34). Further characterization of the human and porcine enzymes on peptide and protein substrates indicated cleavages after not only a P1 Ala but also Val, Leu, Ser, and Thr residues indicating that the primary substrate specificity is similar to that of porcine CELA1 (23, 37, 41–43). Interestingly, the proenzyme of human CELA3B, but not that of human CELA3A, forms binary complexes with procarboxypeptidase A1 and A2, and a similar binary complex was isolated from the pig pancreas (see Ref. 44 and references therein). Ruminants such as cattle, goat, and sheep secrete ternary complexes that also contain chymotrypsinogen C (CTRC)[4] (45). During purification of the bovine ternary complex, proproteinase E (CELA3) becomes autolyzed at the Val-30–Asn-31 peptide bond giving rise to subunit III, an N-terminally truncated inactive proteinase E, the true identity of which has been shrouded in mystery for decades (Refs. 44, 46, 47 and references therein).

Silencing of the archetypal *CELA1* in the human pancreas and duplication of the *CELA3* genes raise the intriguing possibility that these two evolutionary events are linked, and CELA3A or CELA3B might serve as a surrogate for CELA1 with respect to substrate specificity. Therefore, in this study we characterized the substrate specificity of human CELA3A and CELA3B and compared it with pig CELA1 using phage display-selected proteinase inhibitor variants that bind in a substrate-like manner.

[4] The abbreviations used are: CTRC, chymotrypsin C; Boc, *t*-butoxycarbonyl; TEV, tobacco etch virus; Suc, *N*-succinyl.

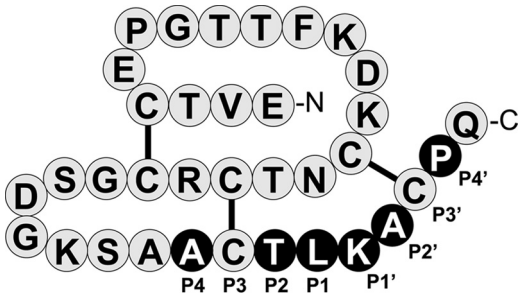

FIGURE 1. **Primary structure of wild-type SGPI-2.** The amino acid positions within the binding loop that were subjected to randomization in the phage display library are highlighted in *white*. Figure taken from Ref. 49. This figure was originally published in Ref. 49. © the American Society for Biochemistry and Molecular Biology.

Results

Selection of SGPI-2 Inhibitor Variants against Human CELA3A and CELA3B and Porcine CELA1 by Phage Display—To analyze the substrate specificity of CELA3A and CELA3B, we performed directed evolution of the *Schistocerca gregaria* proteinase inhibitor 2 (SGPI-2) that binds in a substrate-like manner to its cognate proteinases. To compare the characteristics of the CELA3 substrate-binding sites to those of the CELA1 paralog, we also evolved variants against porcine CELA1. The pig enzyme served as a surrogate model for the evolutionarily silenced human ortholog (17, 18). We previously successfully selected tight-binding inhibitors from SGPI-2 libraries against several serine proteinases, including bovine and crayfish trypsins, human CTRC, and human mannan-binding lectin-associated serine proteinases 1 and 2 (48–50). Furthermore, we demonstrated that wild-type SGPI-2 bound to CELA3A and CELA3B with low but measurable affinity (K_D ~2 and ~0.3 μM, respectively) (49). For this study, we constructed an M13 bacteriophage library displaying SGPI-2 variants on the p8 coat protein where we fully randomized the P4, P2, P1, P1′, P2′, and P4′ positions in the reactive loop while keeping the structurally important P3 and P3′ Cys residues intact (Fig. 1). Selection of tight-binding phage clones from the library containing 3×10^8 variants was performed against CELA3A and CELA3B, and in a separate experiment a library of 7.5×10^8 variants was selected against CELA1. Three rounds of panning were performed. Clones from the third selection cycle were tested for target binding by phage ELISA, and binders were sequenced. The enrichment after the third panning was 3500-fold on CELA3A, 1100-fold on CELA3B, and 3500-fold on CELA1.

CELA3A and CELA3B Selected Pools of SGPI-2 Variants with Similar Binding Loop Sequences—DNA sequencing of selected phages revealed 34 and 49 unique phage clones against CELA3A and CELA3B, respectively (Table 1). To visualize the sequence patterns in the reactive loop of the inhibitors, we generated codon-normalized sequence logos (Fig. 2). The logos indicated similar binding loop sequences selected against the two elastases, with some small but notable differences. Somewhat surprisingly, Ala was not the preferred amino acid at the P1 position where longer side-chain aliphatic residues (Ile, Leu, Met, Val) were primarily selected. CELA3A preferred a P1 Ile over Leu, but Met was excluded, whereas a P1 Leu over Ile was

TABLE 1

Amino acid sequence of the reactive loop of SGPI-2 variants selected by phage display against human CELA3A (34 clones), CELA3B (49 clones), and pig CELA1 (59 clones)

The primary structure of variants unique at the DNA level is shown for positions P4-P3-P2-**P1**-P1′-P2′-P3′-P4′ with P1 residues in bold type.

CELA3A	CELA3B		CELA1	
ACTA**M**LCH	ACTA**Q**LCY	SCT**L**MACW	ACTA**M**YCA	ACT**M**LYCG
ACT**L**MACH	ACT**L**MACA	SCT**L**MLCH	ACTA**R**WCA	ACT**M**LYCG
ACT**L**MLCA	ACT**L**MLCE	WCT**L**MACD	ACTA**R**WCG	ACT**M**LYCQ
ACT**L**MLCF	ACT**L**MLCG	WCT**L**MLCD	ACT**I**MWCN	ACT**M**LYCS
ACT**L**MLCH	ACT**L**MLCN	YCTA**M**YCF	ACT**I**RYCQ	ACT**M**MYCA
ACT**L**MLCH	ACT**M**ELCW	YCTA**R**ACA	ACT**L**IWCG	ACT**M**MYCG
ACT**L**MLCH	ACT**M**LLCA	YCT**I**KLCH	ACT**L**MACH	ACT**M**MYCW
ACT**V**MLCH	ACT**M**LPCG	YCT**I**RLCY	ACT**L**MFCH	ACT**M**RFCG
HCT**I**MACH	ACT**V**MLCV	YCT**L**MACV	ACT**L**MWCG	ACT**M**RLCV
HCT**L**MLCH	ACT**V**MYCP	YCT**L**MACY	ACT**L**MYCG	ACT**M**RLCY
HCT**L**MLCY	ACS**I**RYCE	YCT**L**RECF	ACT**L**MYCL	ACT**M**RWCA
HCT**L**MLCY	FCTA**M**ECG	YCT**L**RFCA	ACT**L**RFCP	ACT**M**RWCD
LCT**I**MLCH	FCTA**R**FCH	YCT**L**RYCQ	ACT**L**RWCA	ACT**M**RWCD
LCT**I**MLCH	FCT**L**MWCW	YCT**V**MLCS	ACT**L**RWCG	ACT**M**RWCD
LCT**I**MLCH	HCT**I**RLCH	YCT**V**RWCW	ACT**L**LYCG	ACT**M**RWCG
LCT**I**RLCW	HCT**L**RWCY		ACT**L**LYCG	ACT**M**RWCG
LCT**L**MACH	LCT**I**MLCH		ACT**M**AYCA	ACT**M**RWCS
LCT**L**MACH	LCT**I**RHCE		ACT**M**IWCG	ACT**M**RYCD
LCT**L**MLCY	LCT**L**MACA		ACT**M**IWCG	ACT**M**RYCG
LCT**V**MLCW	LCT**L**MACY		ACT**M**LFCQ	ACT**M**RYCG
WCT**I**AECH	LCT**L**MLCA		ACT**M**LFCY	ACT**M**RYCG
WCT**L**MLCD	LCT**L**MLCH		ACT**M**LLCV	ACT**M**RYCV
WCT**V**MLCH	LCT**L**MLCM		ACT**M**LPCA	ACT**M**RYCS
WCT**V**MYCP	LCT**L**MLCS		ACT**M**LPCF	GCT**M**LYCG
YCTA**M**LCW	LCT**L**MSCH		ACT**M**LPCH	YCT**L**RWCD
YCT**I**MACH	LCT**L**RHCH		ACT**M**LPCI	
YCT**I**MLCH	LCT**L**RYCP		ACT**M**LPCS	
YCT**I**RECH	LCT**L**RYCW		ACT**M**LWCG	
YCT**I**RLCH	LCT**M**RYCH		ACT**M**LWCG	
YCT**L**MLCH	LCT**V**MLCM		ACT**M**LWCG	
YCT**V**MACH	LCT**V**MLCN		ACT**M**LWCN	
YCT**V**MACY	LCT**V**RACY		ACT**M**LWCV	
YCT**V**MLCH	MCT**L**MLCD		ACT**M**LYCA	
YCT**V**MLCH	MCT**L**MYCS		ACT**M**LYCG	

preferred by CELA3B, and Met was also allowed. The striking dominance of Thr at the P2 position in all logos does not represent elastase-specific selection, but rather it is the result of hydrogen bonding interactions of the Thr side chain with the core of the protein that are important for stabilizing the canonical conformation of the proteinase binding loop. At the P1′ position Met was selected overwhelmingly against both elastases. Arg also occurred in this position with a higher frequency in the CELA3B logo. Compared with P1′, a somewhat weaker conservation emerged at the P2′ position, where CELA3A and CELA3B showed preference for Leu and Tyr, respectively. Furthermore, a P2′ Glu was also selected more frequently against CELA3A than CELA3B. Finally, at position P4′ CELA3A displayed preference for His, but no clear amino acid conservation was apparent at this position for CELA3B. Similarly, only a low level of selection was observed at position P4 with some preference for Tyr and Ala by both elastases.

Porcine CELA1 Selected an SGPI-2 Binding Loop Sequence Pattern Characteristically Different from Those Obtained for the Human CELA3 Enzymes—The sequence logo (Fig. 2*C*) representing 59 unique binding clones (Table 1) revealed a narrower specificity of CELA1 at several binding loop positions. Most notably, in sharp contrast with the more permissive S4 subsite of human CELA3 proteinases, Ala was almost exclusively selected by pig CELA1 at P4. Moreover, the S1 binding pocket of the CELA1 enzyme strongly preferred a P1 Met, although this residue was rare or even absent among CELA3 binders. On the primed side of the scissile bond amino acids Met, Arg, and Leu were selected with similar frequency at the P1′ position, although both CELA3 enzymes preferred Met. CELA1 showed an overwhelming preference for aromatic side chains at P2′, but the human CELA3 enzymes either preferred or at least readily tolerated the presence of aliphatic residues at this position. Finally, the P4′ side chain seemed to have the least impact on the interaction of the inhibitor and the CELA1 enzyme.

Inhibitors Derived from the Phage-selected CELA3 Sequence Logos Were Produced as Recombinant Proteins—On the basis of the selected sequence patterns, we produced recombinant SGPI-2 variants. To generate a meaningful set of inhibitors while keeping their number manageable, we focused on the CELA3 selected patterns. We designed 12 variants (Table 2) to test the roles of the P4, P1, P1′, P2′, and P4′ positions in determining inhibitory affinity and selectivity. The variants were named SGPI-2 E1 through E12. Variants E1–E6 had the common P4–P4′ reactive loop sequence, Tyr-Cys-Thr-(*Xaa*)-Met-Leu-Cys-His, where the highlighted Xaa (P1) residue was varied by incorporating Leu, Ile, Val, Met, Ala, and Ser. Although not selected from our library, Ser was included among the tested P1 residues because cleavage after Ser by CELA3B was reported previously (43), and we also observed autolytic cleavages in the CELA3A and CELA3B activation peptides after Ser residues (44). In variants E7–E12, we tested the effects of a P4 Ala (*versus* Tyr), a P1′ Arg (*versus* Met), a P2′ Glu and Tyr (*versus* Leu), and a P4′ Ala (*versus* His). Note that variant E7 contained a P1 Ile, whereas variants E8–E12 carried a P1 Leu residue. The SGPI-2 variants were expressed in *Escherichia coli* and purified to homogeneity. The binding affinities of the inhibitors against

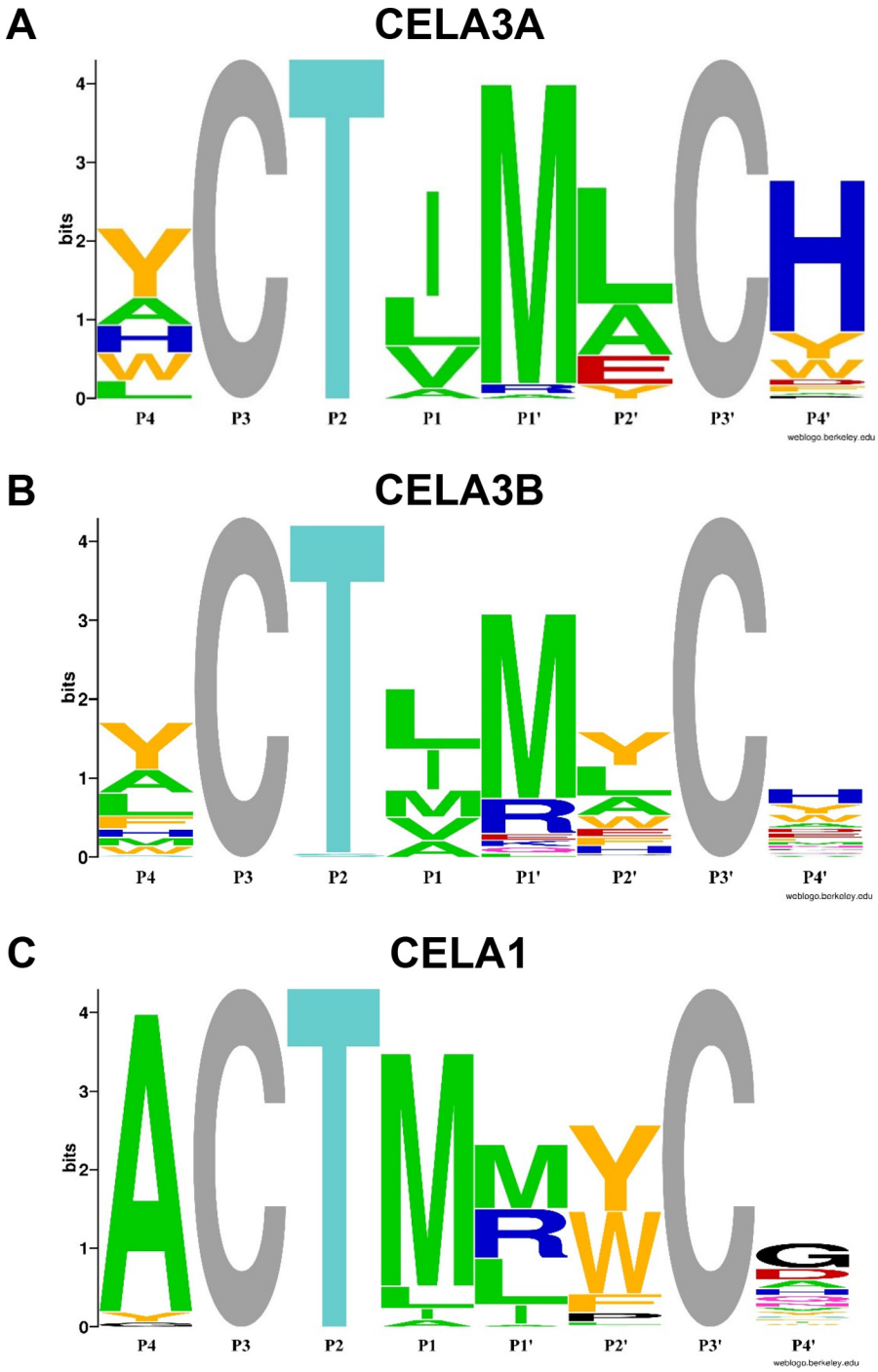

FIGURE 2. **SGPI-2 variants selected by phage display against human elastases CELA3A and CELA3B and pig CELA1.** Sequence logo representation of the binding loop of phage display-selected inhibitors against human CELA3A (34 clones), CELA3B (49 clones), and porcine CELA1 (59 clones). The overall *height of the stack of symbols* indicates the sequence conservation at that position, and the *height of symbols within the stack* indicates the relative frequency of each amino acid. The colors indicate the chemical properties of the amino acid side chains as follows: aliphatic is *green*; aromatic is *orange*; acidic is *red*; basic is *dark blue*; and polar with no charge is *light blue*. Note that the Cys residues shown in *gray* at the P3 and P3′ positions were not randomized. Their *symbol height* represents the maximal height corresponding to a completely conserved residue. The logo was generated with the WebLogo program (55).

human CELA3A, CELA3B, and porcine CELA1 were determined in equilibrium binding assays. Furthermore, to characterize the selectivity of the variants, their affinity toward the human chymotrypsins CTRC, CTRB1, and CTRB2 was also measured. Table 2 lists the equilibrium dissociation constants (K_D) in nanomolar units for each inhibitor variant against the six pancreatic serine proteinases.

Inhibitors Carrying the Isoform-selected Consensus Strongly Improve CELA3 Binding but Reveal Negligible Isoform Specificity—The logos indicate only a loose consensus for CELA3A and CELA3B corresponding to variants E2 and E9, respectively. The parent molecule SGPI-2 inhibited CELA3A and CELA3B with moderate affinities (K_D values of 1952 nM and 337 nM, respectively (49)). In contrast, E2 inhibited

TABLE 2

Inhibition of pancreatic elastases and chymotrypsins by SGPI-2 variants

The table lists the equilibrium dissociation constant (K_D) values in nanomolar units of concentration. Averages of two independent K_D determinations are shown. Binding loop sequences are indicated with P1 residues in bold type. Differences at positions P4, P1′, P2′, and P4′ are underlined and italicized. ND means no inhibition was detected.

Variant	P4 P3 P2 P1 P1′ P2′ P3′ P4′	CELA3A	CELA3B	CELA1	CTRC	CTRB1	CTRB2
E1	Tyr-Cys-Thr-**Leu**-Met-Leu-Cys-His	11	12	30	0.001	1579	128
E2	Tyr-Cys-Thr-**Ile**-Met-Leu-Cys-His	4.5	16	38	0.22	ND	ND
E3	Tyr-Cys-Thr-**Val**-Met-Leu-Cys-His	13	35	113	5.9	ND	ND
E4	Tyr-Cys-Thr-**Met**-Met-Leu-Cys-His	142	246	41	0.088	ND	799
E5	Tyr-Cys-Thr-**Ala**-Met-Leu-Cys-His	27	20	64	1992	ND	78565
E6	Tyr-Cys-Thr-**Ser**-Met-Leu-Cys-His	1170	478	1492	4240	ND	ND
E7	Tyr-Cys-Thr-**Ile**-Met-*Glu*-Cys-His	3.6	17	990	1.6	13928	ND
E8	Tyr-Cys-Thr-**Leu**-*Arg*-Leu-Cys-His	42	102	89	0.21	7049	2281
E9	Tyr-Cys-Thr-**Leu**-Met-*Tyr*-Cys-His	43	10	1.3	0.001	1177	226
E10	*Ala*-Cys-Thr-**Leu**-Met-Leu-Cys-His	10	6.1	9.2	0.002	32	0.47
E11	*Ala*-Cys-Thr-**Leu**-Met-*Tyr*-Cys-His	78	18	0.70	0.009	16	1.1
E12	Tyr-Cys-Thr-**Leu**-Met-Leu-Cys-*Ala*	40	38	59	0.014	1953	174

CELA3A with K_D of 4.5 nM corresponding to a 434-fold improvement, and E9 inhibited CELA3B with a K_D of 10 nM representing a 34-fold improved binding. Additional small increases in affinity were achieved by changing the P2′ to Glu (*versus* Tyr) in E2 (variant E7, K_D 3.6 nM against CELA3A) and the P4 position to Ala (*versus* Tyr) in E9 (variant E10, K_D 6.1 nM against CELA3B). These strongest variants showed 542- and 55-fold increased binding to CELA3A and CELA3B, respectively, relative to wild-type SGPI-2. Overall, all tested variants exhibited comparable binding toward the two elastases with no significant isoform selectivity, indicating that the substrate specificity of CELA3A and CELA3B is highly similar. The most selective CELA3A and CELA3B binders (variants E7 and E11) showed only a 4.7- and 4.3-fold selectivity, respectively, over the other isoform.

Canonical serine proteinase inhibitors become slowly cleaved after the P1 residue by their cognate enzymes, and if no other cleavage occurs, the process leads to equilibrium as the cleaved inhibitor remains functional. However, the cleaved inhibitor may be more susceptible to additional degradation and inactivation by the cognate proteinase. To test for this possibility, we characterized the stability of the six P1 variants by measuring K_D values after 1 h and overnight (16 h) incubations with CELA3A and CELA3B. We found that extended incubation caused either no change or only a small loss in affinity, which was less than 2-fold in most cases (range 0.9–3.0-fold), indicating that inhibitor degradation is not a significant confounding factor in our experiments (Table 3).

P1 Preference of Elastases and Chymotrypsins—Inhibitors E1–E6 differed only at their P1 position and allowed evaluation of the P1 preference of the proteinases studied. We observed a striking similarity with respect to the P1 specificity of human CELA3A and CELA3B with relatively minor differences in side-chain preference. Branched aliphatic side chains (Leu, Ile, and Val) and Ala conferred the highest affinity (K_D range 4.5–27 nM for CELA3A and 12–35 nM for CELA3B), whereas Met was less favored (K_D 142 and 246 nM, respectively), and Ser was inferior (K_D 1170 and 478 nM, respectively). Porcine CELA1 bound inhibitors with P1 branched side chains and Ala weaker than the human elastases (K_D range 30–113 nM) but did not discriminate against Met (K_D 41 nM), which was the most frequently selected P1 amino acid by CELA1. A P1 Ser was poorly recognized by porcine CELA1 as well (K_D 1492 nM).

The K_D values of purified P1 variants E1–E6 against CELA3A reflected the amino acid preference order at the same position of the sequence logo (Table 2 and Fig. 3). Thus, Ile at the P1 position provided the highest affinity toward the enzyme and was followed by Leu, Val, and Ala with 2.4-, 2.9-, and 6-fold higher K_D values, respectively. No SGPI-2 clones with P1 Met or Ser were selected in the phage display experiments, and accordingly, variants with Met(E4) or Ser(E6) at the P1 position were 32- and 260-fold weaker inhibitors of CELA3A, respectively, relative to the consensus E2 variant with P1 Ile. CELA3B exhibited a small 1.3-fold preference toward a P1 Leu over Ile, whereas variants with P1 Val and Ala bound 2.9- and 1.7-fold weaker, respectively (Table 2 and Fig. 3). As seen with CELA3A, inhibitor variants with P1 Met and Ser exhibited a 21- and 40-fold decrease in binding affinity toward CELA3B, respectively. The poor binding of variant E4 with a P1 Met is somewhat surprising considering that Met was more frequently selected in the logo than Val or Ala. Moreover, the binding affinity of variant E6 with a P1 Ser was only 2-fold weaker than that of E4 with P1 Met, even though Ser was completely absent from the P1 position of the sequence logo.

These discrepancies suggest that in terms of binding constraints the E4 variant is not an ideal representative of the P1 Met-containing pool of phage display-selected variants. This can be due to a synergistic contribution of individual binding loop positions to the overall binding energy. Such context dependence of the P1 Met function is supported by the sequences of the selected clones (Table 1). The CELA3A-selected pool contained no P1 Met variant at all, but we tested the P1 contribution in the context of the CELA3A-selected consensus, Tyr-Cys-Thr-Xaa-Met-Leu-Cys-His. If even a single residue of the CELA3A consensus is incompatible with a P1 Met, it could readily explain the observed effect. In fact, out of the 49 unique

TABLE 3

Effect of extended incubation on the inhibition of human CELA3A and CELA3B by SGPI-2 variants

The table lists the equilibrium dissociation constant (K_D) values in nanomolar units of concentration determined after 1- and 16-h incubation times. Binding loop sequences are indicated with P1 residues in bold type. Averages of two independent K_D determinations are shown.

Variant	P4 P3 P2 P1 P1′ P2′ P3′ P4′	CELA3A		CELA3B	
		1 h	16 h	1 h	16 h
E1	Tyr-Cys-Thr-**Leu**-Met-Leu-Cys-His	11	15	12	15
E2	Tyr-Cys-Thr-**Ile**-Met-Leu-Cys-His	4.5	8.8	16	37
E3	Tyr-Cys-Thr-**Val**-Met-Leu-Cys-His	13	39	35	56
E4	Tyr-Cys-Thr-**Met**-Met-Leu-Cys-His	142	268	246	222
E5	Tyr-Cys-Thr-**Ala**-Met-Leu-Cys-His	27	64	20	26
E6	Tyr-Cys-Thr- **Ser**-Met-Leu-Cys-His	1170	1295	478	808

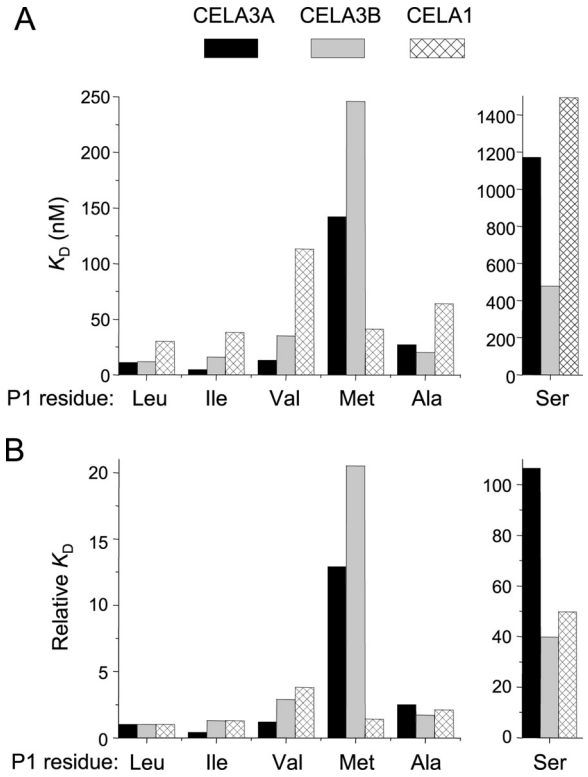

FIGURE 3. **Effect of the P1 amino acid in SGPI-2 on the inhibition of human CELA3A, CELA3B and porcine CELA1.** *A,* equilibrium dissociation constant (K_D) values determined for six inhibitors (E1–E6, see Table 2) containing the indicated P1 residues. *B,* relative K_D values normalized to the E1 inhibitor with a P1 Leu. Measurements were carried out as described under "Experimental Procedures."

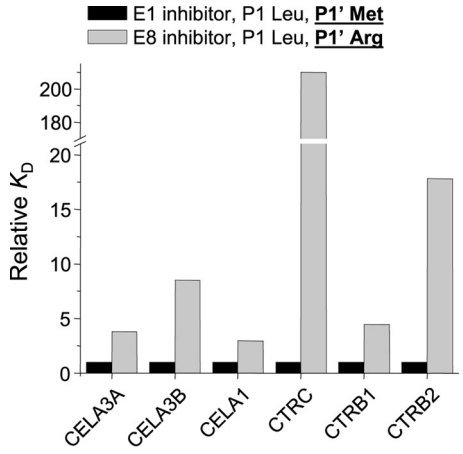

FIGURE 4. **Effect of the P1′ residue (Met *versus* Arg) on the inhibition of elastases and chymotrypsins by SGPI-2 variants E1 and E8 containing a P1 Leu.** The *graph* shows pairwise comparisons of relative K_D values of inhibitor binding to the indicated proteinases, normalized to the E1 value.

CELA3B-selected clones, four have a P1 Met, three accompanied with a P4 Ala and one with a P4 Leu, but none with a P4 Tyr. Moreover, 28 of the 49 CELA3B-selected clones (57%) have a P1′ Met, but none of these combines with a P1 Met. Taken together, this suggests that a P1 Met functions poorly in the dual context of P4 Tyr and P1′ Met, although based on the individual sequences, all other tested P1 residues are compatible with these side chains.

Human CTRC bound inhibitors with a P1 Leu with remarkable (picomolar) affinity confirming its previously described primary substrate specificity (49). Binding of SGPI-2 variants with P1 Met, Ile, and Val were 88-, 220-, and 5900-fold weaker, respectively, whereas variants with P1 Ala and Ser were poor CTRC inhibitors with micromolar affinity.

Binding of all P1 variants to CTRB1 and CTRB2 was very weak and in most cases undetectable. Measurable binding with both enzymes was observed only with variant E1 carrying a P1 Leu. Interestingly, CTRB2 bound this variant more than 10-fold tighter than CTRB1, and a similar relationship was observed for all other inhibitor variants (E8–E12) where binding to both chymotrypsins could be measured (Table 2).

Met at the P1′ Position Is Preferred Over Arg by Elastases—The CELA3A- and CELA3B-selected sequence logos indicated a strong preference for Met at the P1′ position, which was even more strictly conserved than the P1 position. Nevertheless, CELA3B appeared to be permissive toward Arg in the phage selections. The CELA1 enzyme selected Met followed by Arg and Leu in comparable proportions. In direct binding experiments, however, variant E8 carrying a P1′ Arg (*versus* a P1′ Met in E1) exhibited 8.5-fold decreased affinity toward CELA3B, while binding to CELA3A and porcine CELA1 was reduced by 3.8- and 3-fold, respectively (Table 2 and Fig. 4). Note that sequences of E1 and E8 differ from the CELA1-selected consensus (Ala-Cys-Thr-Met-Arg-Trp-Cys-Gly) at 3–4 positions, which might explain why the Arg-containing variant was not preferred by CELA1.

Remarkably, the P1′ Arg inhibited binding to CTRC by 210-fold, and affinity toward CTRB1 and CTRB2 was diminished 4.5- and 18-fold, respectively.

P2′ Position Controls Binding to CELA1 but Not to CELA3 Isoforms—The phage-selected sequence patterns suggested that CELA1 strongly prefers Tyr or Trp at the P2′ position, whereas the two CELA3 isoforms have slightly different requirements. CELA3A showed a clear preference for Leu, whereas Tyr was selected against CELA3B more often. Furthermore, CELA3A seemed more tolerant toward an acidic P2′ Glu side chain than CELA3B. To test whether these differences can be exploited for the design of an isoform-selective elastase inhibitor, we compared binding of inhibitors with a P1 Ile and P2′ Leu *versus* Glu (E2 and E7) and inhibitors with a P1 Leu and P2′ Leu *versus* Tyr (inhibitor pairs E1–E9 and E10-E11). A P2′ Glu (*versus* Leu) caused no significant changes in affinity toward CELA3A or CELA3B (Table 2 and Fig. 5*A*) but reduced binding 26- and 7.3-fold to porcine CELA1 and human CTRC, respectively. In contrast, a P2′ Tyr (*versus* Leu) decreased binding almost 4-fold to CELA3A but had no impact on inhibitor affinity toward CELA3B, which is in good accordance with the corresponding logos of CELA3A and CELA3B. Remarkably, the P2′ Tyr (*versus* Leu) increased binding affinity to pig CELA1 by 13–23-fold in perfect accordance with the CELA1 logo, but it had no appreciable effect on binding to CTRC, CTRB1, or CTRB2 (Fig. 5, *B* and *C*).

P4′ Side Chain Does Not Confer Selectivity among Elastase Isoforms—The P4′ position of the sequence logos exhibits an interesting difference. Although this is the least conserved position for CELA3B and CELA1, a distinct preference for His is

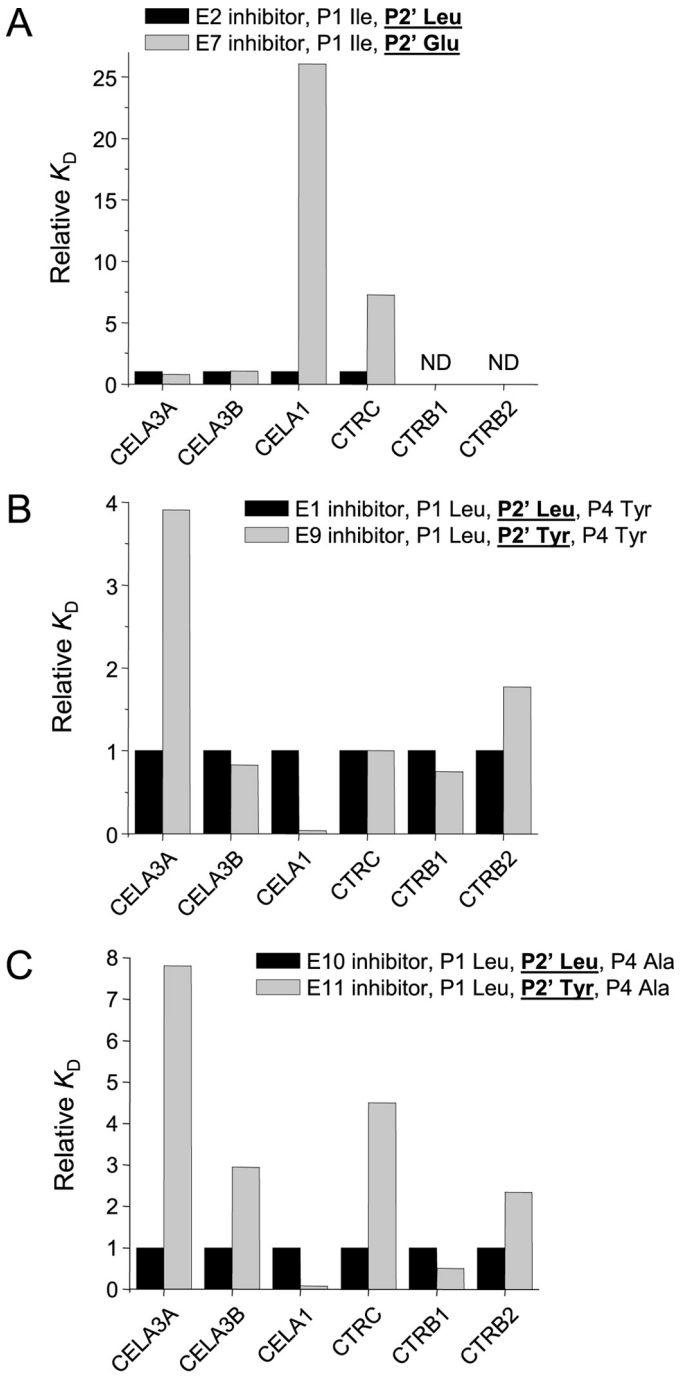

FIGURE 5. **Effect of the P2′ residue on the inhibition of elastases and chymotrypsins by SGPI-2 variants.** *A,* inhibition by variants E2 and E7 containing a P1 Ile and a P2′ Leu *versus* Glu. Pairwise comparisons of relative K_D values of inhibitor binding to the indicated proteinases are shown, normalized to the E2 value. *ND*, no inhibition was detected. *B,* inhibition by variants E1 and E9 containing a P1 Leu and a P2′ Leu *versus* Tyr. Pairwise comparisons of relative K_D values normalized to the E1 value are shown. *C,* inhibition by variants E10 and E11 containing a P1 Leu, a P4 Ala, and a P2′ Leu *versus* Tyr. Pairwise comparisons of relative K_D values normalized to the E10 value are shown.

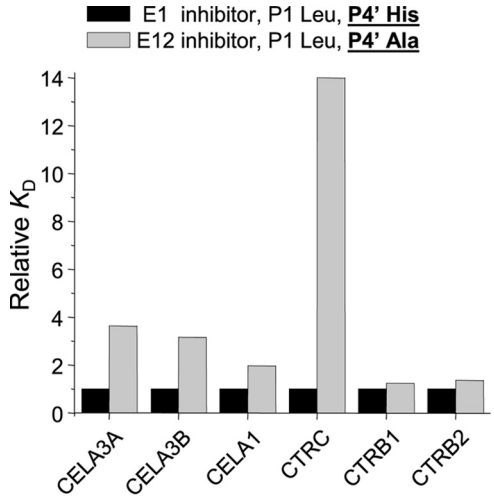

FIGURE 6. **Effect of the P4′ residue (His *versus* Ala) on the inhibition of elastases and chymotrypsins by SGPI-2 variants E1 and E12 containing a P1 Leu.** The *graph* shows pairwise comparisons of relative K_D values of inhibitor binding to the indicated proteinases, normalized to the E1 value.

seen with CELA3A suggesting that the P4′ position might be a selectivity determinant. Unexpectedly, however, when binding of inhibitor variants with P4′ His(E1) *versus* Ala(E12) to CELA3A and CELA3B was tested, a similar 3.6- and 3.2-fold reduction in affinity was observed, respectively; whereas binding to porcine CELA1 was decreased by 2-fold (Fig. 6), indicating that the P4′ position plays a minor role in determining binding affinity and does not confer selectivity among elastases. This conclusion is consistent with the published view that porcine CELA1 contains only three important subsites (S1′–S3′) interacting with the primed side of the substrate (12).

With respect to chymotrypsins, the presence of a P4′ Ala (*versus* His) reduced affinity toward CTRC by 14-fold, while binding to CTRB1 and CTRB2 remained unchanged. The exceptionally strong binding of inhibitor variants with a P4′ His to CTRC was somewhat surprising as we previously found that an acidic P4′ residue (Glu or Asp) was required for high affinity (49). The significantly better performance of His over Ala at this position suggests that His is capable of forming stabilizing, most probably H-bonding, or other polar interactions with the S4′ site.

P4 Position Is an Important Determinant of CELA1, CTRB1, and CTRB2 Inhibition—The CELA1-selected sequence logo showed dramatic conservation of an Ala at position P4, whereas the CELA3A- and CELA3B-selected logos indicated low conservation with a slight preference for Tyr and Ala. We evaluated the effect of P4 Ala *versus* Tyr combined either with a P2′ Leu (E1 *versus* E10) or a P2′ Tyr (E9 *versus* E11) in direct binding experiments. Consistent with the logos, the presence of either Ala or Tyr at the P4 position had small effects (range 1.1–2-fold) on binding of the inhibitors to CELA3 isoforms. Surprisingly, similarly small (~2–3-fold) effects were observed with CELA1 despite the strong selection for an Ala at the P4 position. This suggests that the preference for a P4 Ala is context-dependent and requires the presence of optimal residues at the P1 and P2′ positions.

A P4 Ala (*versus* Tyr) decreased inhibition of CTRC by 2-fold (with P2′ Leu) and 9-fold (with P2′ Tyr). Remarkably, however, a P4 Ala (*versus* Tyr) improved binding to CTRB1 and CTRB2 by 74- and 205-fold (with a P2′ Tyr) and 49- and 272-fold (with a P2′ Leu), respectively (Fig. 7).

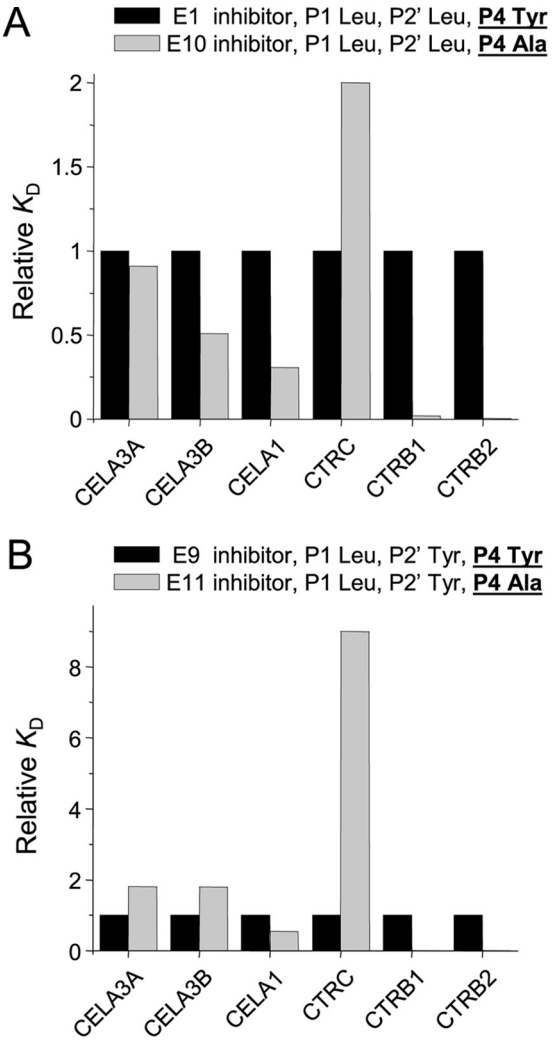

FIGURE 7. **Effect of the P4 residue on the inhibition of elastases and chymotrypsins by SGPI-2 variants.** *A,* inhibition by variants E1 and E10 containing a P1 Leu, a P2′ Leu, and a P4 Tyr *versus* Ala. Pairwise comparisons of relative K_D values of inhibitor binding to the indicated proteinases are shown, normalized to the E1 value. *B,* inhibition by variants E9 and E11 containing a P1 Leu, a P2′ Tyr, and a P4 Tyr *versus* Ala. Pairwise comparisons of relative K_D values normalized to the E9 value are shown.

P1 Preference of Elastases on p-Nitroanilide Peptide Substrates Is Skewed toward Ala—The widely accepted notion that elastases favor a P1 Ala over other amino acid side chains has been supported by a biased literature where most studies used various synthetic ester and amide substrates with a P1 Ala (11, 13, 23, 32, 34, 36). To re-investigate this phenomenon in a more systematic manner, we measured kinetic parameters for human CELA3A, CELA3B, and pig CELA1 on a series of small peptide substrates with the general sequence of Suc-Ala-Ala-Pro-Xaa-*p*-nitroanilide, where Xaa was Ala, Val, Leu, Ile, Met, Ser, or Gly (Table 4). Considering catalytic efficiency (k_{cat}/K_m), we found that CELA3B was on average 2.8-fold (range 1.6–4-fold) more active on these substrates than CELA3A. Relative to the human elastases, porcine CELA1 was superior on all substrates to a varying extent (range 1.1–47-fold), with the largest differences observed on the alanyl, leucyl, and seryl substrates. This higher activity might be partially due to the overwhelming Ala preference of the S4 site of CELA1. There was no quantitative correlation between K_D values measured for the inhibitor variants with different P1 residues and catalytic efficiencies of the analogous *p*-nitroanilide substrates. Remarkably, when relative catalytic efficiencies were compared, the substrate with P1 Ala was cleaved by all elastases 8–10-fold better compared with the second best substrate with a P1 Val or Leu (Fig. 8). The results confirm the apparent Ala specificity of elastases on small peptide substrates, which stands in contrast with published cleavage patterns on larger protein substrates and our binding data using phage display-selected SGPI-2 variants.

Discussion

In this study we used phage display-assisted directed evolution of a small protein proteinase inhibitor to characterize the substrate specificity of human pancreatic elastases CELA3A and CELA3B. These experiments were undertaken to clarify whether the primate-specific duplication of the *CELA3* gene resulted in the evolution of unique proteinase specificities. This question seemed particularly interesting in light of the previously reported observation that the human ortholog of the archetypal pig pancreatic elastase CELA1 is not expressed in the human pancreas (17, 18). It seemed reasonable to speculate that gene duplication might have created a functional substitute for the missing CELA1. To address this problem, we used phage-displayed libraries of the SGPI-2 proteinase inhibitor to select tight-binding variants against human CELA3A and CELA3B as well as porcine CELA1. In this library, six amino acid positions in the reactive binding loop of the inhibitor were fully randomized (see Fig. 1). On the basis of the CELA3-selected sequence patterns, we recombinantly expressed and purified 12 SGPI-2 inhibitor variants and determined their binding to CELA3A, CELA3B, and porcine CELA1. Furthermore, to characterize selectivity, affinity toward human chymotrypsins CTRC, CTRB1, and CTRB2 was measured. The success of the selection process was confirmed by the markedly improved binding of inhibitor variants carrying the loose consensus sequence compared with the parent SGPI-2 molecule. Interestingly, affinity of the CELA3 consensus variants was comparable and showed little isoform selectivity.

With respect to the primary substrate specificity of the elastases studied, we made the following observations. (i) The P1 specificity of human CELA3A, CELA3B, and porcine CELA1 was broad and overall very similar, directed toward aliphatic amino acid side chains Leu, Ile, Val, and Ala. (ii) A P1 Met was preferred by porcine CELA1, but it was less favored by human elastases. (iii) Inhibitors with a P1 Ser were poorly recognized by all elastases tested. The observations confirm and extend previous studies where cleavage specificity of pig CELA1, pig proteinase E, and human CELA3B was examined on macromolecular substrates. Thus, porcine CELA1 was shown to hydrolyze elastin at Ala-Ala and Ala-Gly peptide bonds and to cleave oxidized insulin A and B chains after Ala, Val, Ser, and Gly (6, 8, 9). A 1977 "meta-analysis" of all published CELA1 cleavage sites in proteins found that P1 Ile, Val, and Ala were preferred, but cleavages after Leu, Thr, and Ser were also reported (14). Porcine proteinase E cleaved oxidized ribonuclease and insulin chains mostly after Ala, Val, Ser, and Thr residues; one Ile-Thr peptide bond in ribonuclease was also extensively hydrolyzed

TABLE 4

Enzyme kinetic parameters of human CELA3A, CELA3B, and porcine CELA1 on Suc-Ala-Ala-Pro-Xaa-*p*-nitroanilide substrates, where Xaa stands for the amino acids indicated in the table heading.

Measurements were performed at 22 °C as described under "Experimental Procedures." Although not shown, substrates containing Lys, Arg, Trp, Tyr, Phe, Asn, or Gln at the Xaa position were not cleaved by the elastases. The units of measure were as follows: k_{cat} (s^{-1}), K_m (μM), and k_{cat}/K_m ($M^{-1}\,s^{-1}$). Average values ± S.E. ($n = 3$) are shown. ND means no cleavage was detected.

	Ala	Val	Leu	Ile	Met	Ser	Gly
CELA3A							
k_{cat}	22.9 ± 0.7	2.44 ± 0.08	1.16 ± 0.03	0.99 ± 0.01	0.24 ± 0.01	0.14 ± 0.01	ND
K_m	1089 ± 70	1178 ± 82	722 ± 53	1060 ± 27	452 ± 52	1447 ± 156	ND
k_{cat}/K_m	2.1×10^4	2.1×10^3	1.6×10^3	9.3×10^2	5.3×10^2	9.9×10^1	ND
CELA3B							
k_{cat}	24.8 ± 0.5	3.66 ± 0.10	1.01 ± 0.01	1.25 ± 0.02	0.24 ± 0.01	0.26 ± 0.01	ND
K_m	374 ± 22	430 ± 32	259 ± 8	484 ± 23	288 ± 19	984 ± 51	ND
k_{cat}/K_m	6.6×10^4	8.5×10^3	3.9×10^3	2.6×10^3	8.3×10^2	2.6×10^2	ND
CELA1							
k_{cat}	115 ± 1	16.9 ± 0.7	57.8 ± 1.3	2.87 ± 0.06	0.70 ± 0.01	5.62 ± 0.16	1.86 ± 0.05
K_m	166 ± 4	1029 ± 93	785 ± 30	995 ± 43	540 ± 30	2311 ± 108	1555 ± 74
k_{cat}/K_m	6.9×10^5	1.7×10^4	7.4×10^4	2.9×10^3	1.3×10^3	2.4×10^3	1.2×10^3

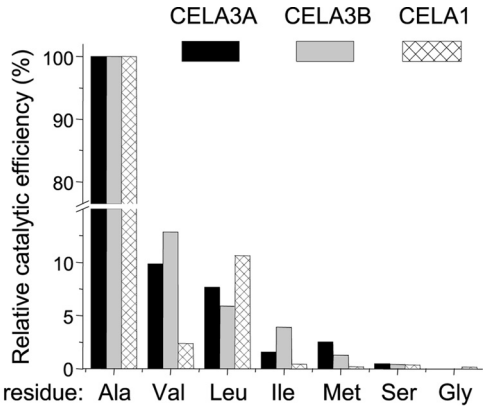

FIGURE 8. **Catalytic efficiency of human elastases CELA3A and CELA3B and porcine CELA1 on small peptide substrates.** Enzyme kinetic parameters were determined on Suc-Ala-Ala-Pro-Xaa-*p*-nitroanilide substrates, where Xaa stands for the indicated P1 amino acid residues, as described under "Experimental Procedures" and Table 4. Specificity constants (k_{cat}/K_m) were expressed as percent of the value obtained with the Ala substrate.

(30, 42). Finally, human CELA3B digested oxidized insulin chains and glucagon preferably after Ala, Val, and Leu, although cleavages after Ser, Thr, and His were also observed (43). Interestingly, in protein substrates elastases often cleaved after a P1 Ser and Thr, but our substrate-like inhibitor with a P1 Ser exhibited poor binding. Although the reason for this discrepancy is not readily apparent, cleavage after the relatively polar Ser may be more context-dependent and may require different subsite interactions than those selected here with more hydrophobic P1 side chains. Alternatively, the more rigid binding loop in the inhibitor may pose unique constraints that prevent accommodation of a P1 Ser in the elastase primary binding pocket.

While the published literature on elastase cleavage sites in protein substrates seems to agree with our results, quantitative data on cleavage rates are scarce, and a direct comparison is difficult. K_D values for substrate-like protein inhibitors tend to correlate with k_{cat}/K_m values for hydrolysis of the analogous protein substrates, but this may not be generally applicable (52). Additional studies are required to see how closely the substrate specificity profiles determined with substrate-like inhibitors actually mimic the cleavage preferences of elastases on protein substrates.

The commonly cited Ala specificity of elastases is partly explained by a biased literature that preferentially used small amide and ester substrates with a P1 Ala to study elastase activity. Curiously, amide substrates with a P1 Ala are markedly better hydrolyzed by elastases than those with any other P1 residue (11, 13, 32). We confirmed this phenomenon by measuring kinetic parameters on Suc-Ala-Ala-Pro-Xaa-*p*-nitroanilide peptides, where Xaa was Ala, Val, Ile, Leu, Met, Ser, or Gly. All elastases studied cleaved the P1 Ala substrate with 8–10-fold better catalytic efficiency than the next best substrate. Although an explanation for the superiority of a P1 Ala in the context of small amide substrates is not obvious, our binding data with substrate-like proteinase inhibitors and published cleavage site preferences in proteins do not support the contention that a P1 Ala would be favored over other small side chains by elastases. Interestingly, when a series of tripeptide thiobenzyl ester substrates (Boc-Ala-Ala-Xaa-thiobenzyl ester) was tested, the order of cleavage preference by pig CELA1 was Leu > Met > Ala ~Val ~ Ile ~ Ser with only about 2-fold differences between substrates (15).

Our results also offered insight into the extended subsite specificities for human CELA3A, CELA3B, and pig CELA1. In this regard, prior studies on pig CELA1 found that it contained eight subsites extending from P5 through P3′ (12). The S1′ site preferred Lys (*versus* Ala, Tyr, or Glu); the S2′ site preferred Phe (*versus* Ala), and the S3′ site preferred aromatic side chains (10, 12, 53). Another study found that P1′ Ala, Leu, Met, Val, and even Arg were well tolerated, whereas a P1′ Gly was not preferred (54). Based on the cleavage pattern of oxidized insulin chains by porcine proteinase E, the authors speculated on subsite preference and suggested that at P2 Gln or Ile are not favored but Glu and Leu are allowed and at P1′ Gln and at P2′ Lys also have negative effects (42). In our phage display and inhibitor binding experiments, we did not find a marked difference in the extended subsite preference of human CELA3A and CELA3B; both isoforms preferred a P1′ Met, whereas interactions with positions P4, P2′, and P4′ were less selective. In contrast, the porcine CELA1-selected logo indicated significant subsite preference at position P4 (Ala) and P2′ (Tyr/Trp). The importance of a P2′ Tyr was verified in direct binding experiments, whereas that of a P4 Ala appeared to be context-dependent.

Selectivity of the newly evolved SGPI-2 variants was tested against three human chymotrypsins, CTRB1, CTRB2, and CTRC. In general, variants inhibited CTRB1 and CTRB2 poorly; measurable inhibition was only observed with a P1 Leu. A strong selectivity determinant against these two chymotrypsins was identified at the P4 position where the presence of a Tyr (*versus* Ala) markedly reduced affinity. Remarkably, inhibitor variants with a P1 Leu turned out to be superior CTRC inhibitors with picomolar affinity. Indeed, the reactive loop sequences of the strongly binding inhibitors ((Tyr/Ala)-Cys-Thr-**Leu**-Met-(Leu/Tyr)-Cys-His) were very similar to the best CTRC inhibitor variant selected in our preceding study (Ala-Cys-Thr-**Leu**-Met-Leu-Cys-Glu) (49). Although previously we assigned significance to the presence of the acidic residue at the P4′ position in determining high affinity binding, the current data indicate that a His residue is also compatible at this site and is significantly better than an Ala. As we observed before, CTRC is relatively tolerant to changes at the P2′ position; however, we found now that a P1′ Arg inhibited binding to CTRC by 20-fold, although it had only a small effect on other human chymotrypsins. In fact, the absence of a P1′ Arg in our earlier phage display study already predicted that this residue would be inhibitory for CTRC binding (49). As discussed above, the somewhat unexpected efficacy of our selected inhibitor variants against CTRC indicates that within the human chymotrypsin-like enzymes the SGPI-2 scaffold is uniquely compatible with this proteinase.

In summary, we compared the substrate specificity of human CELA3A, CELA3B, and porcine CELA1 elastases using phage display selection of an SGPI-2 proteinase inhibitor library, direct binding experiments with purified SGPI-2 variants, and cleavage experiments with *p*-nitroanilide substrates. The results demonstrate that the substrate specificity of CELA3A, CELA3B, and CELA1 are similar with a P1 specificity broadly directed toward aliphatic amino acid side chains. In accordance with their more distant evolutionary relationship, we detected specificity differences between the CELA3 and CELA1 paralogs at the P4 and P2′ positions. However, the extended subsite specificities of the two CELA3 isoforms were comparable. Thus, gene duplication of the *CELA3* isoforms did not give rise to a CELA1 surrogate, but it may have compensated for the loss of CELA1 in the human pancreas by increasing the digestive enzyme pool.

Experimental Procedures

Purification of Digestive Serine Proteinases—Human CELA3A, CELA3B, CTRC, CTRB1, and CTRB2 proenzymes carrying C-terminal His tags were expressed in HEK 293T cells and purified from the conditioned medium by nickel-affinity chromatography as described previously (44). The proenzymes were activated with immobilized trypsin (catalog no. 20230, Thermo Fisher Scientific, Waltham, MA), and the trypsin beads were removed by centrifugation. Porcine CELA1 (catalog no. E0127) used for phage selection was purchased from Sigma, and the enzyme used for kinetic and inhibition assays (code ESFF, catalog no. LS006363) was from Worthington.

Construction of the SGPI-2 Library—A protein-phage library displaying SGPI-2 on the surface of M13 bacteriophage was created as described previously (48–50). Briefly, the coding sequence of SGPI-2 was fused to the p8 coat protein of M13, and six positions in the binding loop of the inhibitor (Fig. 1) were fully randomized in two consecutive Kunkel mutagenesis steps. First, stop codons were introduced to eliminate any wild-type background from the selection. Next, stop codons were replaced by NNK codons (where N = A, C, T, G and K = G, T), which encode all 20 amino acids. The resulting phagemid library was electroporated into *E. coli* SS320. Cells were infected with M13KO7 helper phage (New England Biolabs, Ipswich, MA) and grown overnight to produce the protein-phage library.

Selection of the SGPI-2 Library on Human CELA3A and CELA3B—As simpler surface immobilization methods proved to be inefficient with these isoforms, solution phase selection was performed followed by a His tag/immobilized metal affinity chromatography capture. In the first round of selection, 10^{12} phage particles displaying SGPI-2 variants were incubated in the absence or presence of 50 nM His-tagged CELA3A or CELA3B in 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 0.1% Tween 20, and 20 mM imidazole (wash buffer) for 1 h. Phage particles incubated in wash buffer containing no elastase served as control for enrichment determination. Aliquots of Dynabeads His Tag Isolation and Pulldown (Thermo Fisher Scientific) (5 μl) were blocked with 500 μl of 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 0.1% Tween 20, 5 mg/ml BSA, and 50 mM imidazole (block buffer) to prevent nonspecific binding of the phage particles to the beads and washed with 500 μl of wash buffer. Elastase-bound phages were then captured through the His tag by incubating with the pre-blocked Dynabeads for 5 min. To separate bound phages, the beads were collected using a magnet, and unbound phages were removed by rinsing the beads twice with 1.5 ml of wash buffer. Bound phages were eluted with 1.4 ml of 250 mM imidazole in 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, and 0.1% Tween 20 (elution buffer). Then, 4.5 ml of midlog-phase *E. coli* XL1 Blue cells were infected by 0.5 ml of phages. After 30 min of incubation at 37 °C with shaking, 20 μl of cells were removed, and 10-fold serial dilutions were prepared. Cell dilutions (10 μl) were transferred onto pre-dried LB/Amp plates and incubated at 37 °C overnight. The number of infected cells generated by elastase-selected *versus* control phages defined the enrichment of target-binding phage clones over nonspecific binders. The rest of the 5 ml of infected *E. coli* cultures were superinfected with M13KO7 helper phage and grown overnight to propagate the phages. The second and third rounds of selection were carried out similarly as described above with the exception that the propagated and isolated phages were pre-cleared on blocked Dynabeads to eliminate background-binding clones.

Phage ELISA of Human CELA3A and CELA3B Selected SGPI-2 Clones—Individual clones from the third round of selection were tested for target binding in phage ELISA. Clones were grown in single loose tubes (National Scientific), and the supernatant containing the inhibitor-displaying phages was used in the assays. Aliquots (1 μl) of Dynabeads were transferred to 96-well microtiter plates and blocked with 100 μl of block buffer for 30 min. After removal of the block buffer, the beads were washed with 100 μl of wash buffer. On a separate microtiter plate, 50 μl of phage supernatant was incubated in

the absence (control) or presence of 7 nM CELA3A or CELA3B (final concentration) in a 200-μl final volume for 1 h. The elastase-phage incubates were transferred to the blocked beads and incubated for 5 min. Non-bound phages were removed, and the beads were washed four times with 200 μl of wash buffer. Then, 50 μl of HRP-conjugated anti-M13 antibody (GE Healthcare) diluted 5000-fold in wash buffer was added to each well for 30 min. The unbound antibody was discarded, and the beads were first washed with 200 μl of wash buffer and then with 200 μl of 50 mM Tris-HCl (pH 8.0) and 300 mM NaCl. For color development, 50 μl of 3,3′,5,5′-tetramethylbenzidine substrate was added to the wells, and the reaction was stopped with 50 μl of 1 M HCl. The absorbance at 450 nm was measured after the removal of the beads from the solutions.

Selection of the SGPI-2 Library on Porcine CELA1—Three selection cycles were performed as described previously (48–51). Briefly, 5 μg/well porcine CELA1 in phosphate-buffered saline (pH 7.2) (PBS) was immobilized in 12 wells of Nunc Immobilizer Amino Plates for 3 h. Wells without CELA1 served as control. The wells were blocked with 200 μl of 5 mg/ml BSA in PBS for 1 h. Then, 100 μl of phages in PBS, 5 mg/ml BSA, and 0.05% Tween 20 was added to the blocked wells and incubated for 2.5 h. After rinsing the wells with PBS, bound clones were eluted with 100 μl of 0.1 M HCl, and the eluate was brought to neutral pH by adding 1 M Tris base solution. Mid-log XL1-Blue cells were infected with the eluted phages, titrated for determination of the enrichment ratio, superinfected with M13 KO7 helper phage, and grown overnight for target-binding phage amplification.

Phage ELISA of Porcine CELA1-selected SGPI-2 Clones—Individual clones from the third selection cycle were tested for binding to CELA1 in phage ELISAs as described previously (48–51). Porcine CELA1 (1.25 μg/well in PBS) was immobilized in Nunc Immobilizer Amino Plates.

DNA Sequence Analysis of Selected Phages—Clones with a signal-to-noise ratio higher than 2 were considered positive and were Sanger sequenced (Table 1). In each of the randomized positions, the frequency of the amino acids was normalized to the expected frequency of their corresponding codons in the NNK set. A dataset of 1000 sequences was generated based on these normalized frequencies that served as input for sequence logo generation by the WebLogo program (55).

Expression and Purification of SGPI-2 Variants—Based on the sequence pattern of the inhibitor binding loops selected against the human elastases, 12 different inhibitor variants were designed and created using PCR-based mutagenesis (Table 2). The recombinant SGPI-2 variants were expressed into the cytoplasm of *E. coli* SHuffle T7 cells (C3026, New England Biolabs) fused to the C terminus of the *E. coli* protein-disulfide isomerase DsbC following the strategy of Nozach *et al.* (56). The construct also incorporated an N-terminal His tag and a TEV proteinase cleavage site between the inhibitor and DsbC. The fusion protein was purified by nickel-affinity chromatography and processed by His-tagged TEV proteinase. His-tagged DsbC and TEV proteinase as well as any unprocessed SGPI-2 were removed in a second nickel-affinity chromatography step. Finally, inhibitor variants were purified to homogeneity on a Phenomenex Jupiter 10u C4 300A column using reverse phase-HPLC. The correct molecular weight of the proteins was verified by ESI-MS as described previously (49).

Determination of Concentrations—Active site titrated bovine trypsin was used to determine the concentration of the pan-specific serine proteinase inhibitor ecotin, which was then used to measure the concentrations of the CTRC, CTRB1, CTRB2, and pig CELA1 preparations. Because ecotin binds CELA3A and CELA3B with lower affinity, eglin C was used to measure active concentrations of human elastases. Eglin C was purified as reported previously, and its concentration was determined by titration against human CTRC (22). SGPI-2 variants E1, E2, E4, E8, E9, E10, and E12 were titrated against CTRC and E11 against porcine CELA1. All titrations were performed at enzyme concentrations at least 2 orders of magnitude higher than the equilibrium dissociation constant of the interaction. Concentration of SGPI-2 variants E3, E5, E6, and E7 was measured by a protein assay (Micro BCA Protein Assay Kit, Thermo Fisher Scientific) using E8 as standard.

Equilibrium Binding Assays—The K_D values of the binding reaction between SGPI-2 variants and pancreatic proteinases were determined according to Empie and Laskowski (35) as we described previously (48–50). Incremental amounts of the inhibitor were incubated with a fixed enzyme concentration at 22 °C for 1 h in 0.1 M Tris-HCl (pH 8.0), 1 mM $CaCl_2$, and 0.05% Tween 20. The amount of free enzyme was then measured using elastase substrates Suc-Ala-Ala-Pro-Ala-*p*-nitroanilide or Suc-Ala-Pro-Ala-amido-methylcoumarin or chymotrypsin substrates Suc-Ala-Ala-Pro-Phe-*p*-nitroanilide or Suc-Ala-Ala-Pro-Phe-amido-methylcoumarin. The experimental data were fitted with the following equation: $y = E - (E + x + K - \sqrt{((E + x + K)\wedge 2 - 4Ex)})/2$, where the independent variable x represents the total inhibitor concentration; the dependent variable y is the free proteinase concentration in equilibrium; K is K_D, and E designates the total proteinase concentration. Each K_D value was determined in two independent assays, and the average values are presented.

Enzyme Kinetic Analysis—The activity of human CELA3A, CELA3B, and porcine CELA1 was determined on a series of chromogenic peptide substrates with the general structure of Suc-Ala-Ala-Pro-Xaa-*p*-nitroanilide, where Xaa stands for Ala, Arg, Asn, Gln, Gly, Ile, Leu, Met, Lys, Phe, Ser, Trp, Tyr, or Val. Measurements were performed at 22 °C in 0.1 M Tris-HCl (pH 8.0), 1 mM $CaCl_2$, and 0.05% Tween 20. Rates of substrate hydrolysis were plotted as a function of the substrate concentration, and K_m and k_{cat} values were determined from hyperbolic fits to the data points.

Author Contributions—E. B., A. S., K. Z., and D. H. designed and performed the experiments, analyzed the data, prepared the figures for publication, and approved the final version of the manuscript. M. S. T. and G. P. designed the project and analyzed the data. M. S. T., G. P., and E. B. wrote the manuscript.

Note Added in Proof—András Szabó was inadvertently omitted as an author in the version of this manuscript that was published as a Paper in Press on January 6, 2017. This error has now been corrected.

References

1. Schechter, I., and Berger, A. (1967) On the size of the active site in proteinases. I. Papain. *Biochem. Biophys. Res. Commun.* **27,** 157–162
2. Baló, J., and Banga, I. (1949) Elastase and elastase-inhibitor. *Nature* **164,** 491
3. Baló, J., and Banga, I. (1950) The elastolytic activity of pancreatic extracts. *Biochem. J.* **46,** 384–387
4. Lamy, F., Craig, C. P., and Tauber, S. (1961) Studies on elastase and elastin. I. Assay and properties of elastase. *J. Biol. Chem.* **236,** 86–91
5. Naughton, M. A., and Sanger, F. (1961) Purification and specificity of pancreatic elastase. *Biochem. J.* **78,** 156–163
6. Narayanan, A. S., and Anwar, R. A. (1969) The specificity of purified porcine pancreatic elastase. *Biochem. J.* **114,** 11–17
7. Gertler, A. (1971) The non-specific electrostatic nature of the adsorption of elastase and other basic proteins on elastin. *Eur. J. Biochem.* **20,** 541–546
8. Gertler, A., Weiss, Y., and Burstein, Y. (1977) Purification and characterization of porcine elastase II and investigation of its elastolytic specificity. *Biochemistry* **16,** 2709–2715
9. Vered, M., Burstein, Y., and Gertler, A. (1985) Digestion of elastin by porcine pancreatic elastase I and elastase II. *Int. J. Pept Protein Res.* **25,** 76–84
10. Atlas, D., Levit, S., Schechter, I., and Berger, A. (1970) On the active site of elastase: partial mapping by means of specific peptide substrates. *FEBS Lett.* **11,** 281–283
11. Bieth, J., and Wermuth, C. G. (1973) The action of elastase on *p*-nitroanilide substrates. *Biochem. Biophys. Res. Commun.* **53,** 383–390
12. Atlas, D. (1975) The active site of porcine elastase. *J. Mol. Biol.* **93,** 39–53
13. Kasafírek, E., Fric, P., Slabý, J., and Malis, F. (1976) *p*-Nitroanilides of 3-carboxypropionyl-peptides. Their cleavage by elastase, trypsin, and chymotrypsin. *Eur. J. Biochem.* **69,** 1–13
14. Powers, J. C., Gupton, B. F., Harley, A. D., Nishino, N., and Whitley, R. J. (1977) Specificity of porcine pancreatic elastase, human leukocyte elastase and cathepsin G. Inhibition with peptide chloromethyl ketones. *Biochim. Biophys. Acta* **485,** 156–166
15. Harper, J. W., Cook, R. R., Roberts, C. J., McLaughlin, B. J., and Powers, J. C. (1984) Active site mapping of the serine proteinases human leukocyte elastase, cathepsin G, porcine pancreatic elastase, rat mast cell proteinases I and II. Bovine chymotrypsin Aα, and *Staphylococcus aureus* proteinase V-8 using tripeptide thiobenzyl ester substrates. *Biochemistry* **23,** 2995–3002
16. Shotton, D. M., and Watson, H. C. (1970) Three-dimensional structure of tosyl-elastase. *Nature* **225,** 811–816
17. Tani, T., Kawashima, I., Furukawa, H., Ohmine, T., and Takiguchi, Y. (1987) Characterization of a silent gene for human pancreatic elastase I: structure of the 5′-flanking region. *J. Biochem.* **101,** 591–599
18. Rose, S. D., and MacDonald, R. J. (1997) Evolutionary silencing of the human elastase I gene (ELA1). *Hum. Mol. Genet.* **6,** 897–903
19. Uram, M., and Lamy, F. (1969) Purification of two proelastases from porcine pancreas. *Biochim. Biophys. Acta* **194,** 102–111
20. Ardelt, W. (1974) Partial purification and properties of porcine pancreatic elastase II. *Biochim. Biophys. Acta* **341,** 318–326
21. Ledoux, M., and Lamy, F. (1975) Electrophoretic characterization of porcine pancreatic (pro)elastases A and B. *Can. J. Biochem.* **53,** 421–432
22. Komiyama, T., and Fuller, R. S. (2000) Engineered eglin c variants inhibit yeast and human proprotein processing proteinases, *Kex2* and furin. *Biochemistry* **39,** 15156–15165
23. Largman, C., Brodrick, J. W., and Geokas, M. C. (1976) Purification and characterization of two human pancreatic elastases. *Biochemistry* **15,** 2491–2500
24. Ohlsson, K., and Olsson, A. S. (1976) Purification and partial characterization of human pancreatic elastase. *Hoppe Seylers Z Physiol. Chem.* **357,** 1153–1161
25. Shirasu, Y., Yoshida, H., Matsuki, S., Takemura, K., Ikeda, N., Shimada, Y., Ozawa, T., Mikayama, T., Iijima, H., and Ishida, A. (1987) Molecular cloning and expression in *Escherichia coli* of a cDNA encoding human pancreatic elastase 2. *J. Biochem.* **102,** 1555–1563
26. Fletcher, T. S., Shen, W. F., and Largman, C. (1987) Primary structure of human pancreatic elastase 2 determined by sequence analysis of the cloned mRNA. *Biochemistry* **26,** 7256–7261
27. Kawashima, I., Tani, T., Shimoda, K., and Takiguchi, Y. (1987) Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed in human pancreas. *DNA* **6,** 163–172
28. MacDonald, R. J., Swift, G. H., Quinto, C., Swain, W., Pictet, R. L., Nikovits, W., and Rutter, W. J. (1982) Primary structure of two distinct rat pancreatic preproelastases determined by sequence analysis of the complete cloned messenger ribonucleic acid sequences. *Biochemistry* **21,** 1453–1463
29. Gestin, M., Le Huërou-Luron, I., Wicker-Planquart, C., Le Dréan, G., Chaix, J. C., Puigserver, A., and Guilloteau, P. (1997) Bovine pancreatic preproelastases I and II: comparison of nucleotide and amino acid sequences and tissue specific expression. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* **118,** 181–187
30. Ardelt, W., Tomczak, Z., and Dudek-Wojciechowska, G. (1979) Specificity of elastases: degradation of the oxidized β-chain of insulin by porcine pancreatic elastase II and dog leucocyte elastase. *Acta Biochim. Pol.* **26,** 267–273
31. Del Mar, E. G., Largman, C., Brodrick, J. W., Fassett, M., and Geokas, M. C. (1980) Substrate specificity of human pancreatic elastase 2. *Biochemistry* **19,** 468–472
32. Szabó, A., and Sahin-Tóth, M. (2012) Determinants of chymotrypsin C cleavage specificity in the calcium-binding loop of human cationic trypsinogen. *FEBS J.* **279,** 4283–4292
33. Szepessy, E., and Sahin-Tóth, M. (2006) Inactivity of recombinant ELA2B provides a new example of evolutionary elastase silencing in humans. *Pancreatology* **6,** 117–122
34. Mallory, P. A., and Travis, J. (1975) Human pancreatic enzymes: purification and characterization of a nonelastolytic enzyme, proteinase E. Resembling elastase. *Biochemistry* **14,** 722–730
35. Empie, M. W., and Laskowski, M., Jr. (1982) Thermodynamics and kinetics of single residue replacements in avian ovomucoid third domains: effect on inhibitor interactions with serine proteinases. *Biochemistry* **21,** 2274–2284
36. Sziegoleit, A. (1982) Purification and characterization of a cholesterol-binding protein from human pancreas. *Biochem. J.* **207,** 573–582
37. Sziegoleit, A. (1984) A novel proteinase from human pancreas. *Biochem. J.* **219,** 735–742
38. Shen, W. F., Fletcher, T. S., and Largman, C. (1987) Primary structure of human pancreatic proteinase E determined by sequence analysis of the cloned mRNA. *Biochemistry* **26,** 3447–3452
39. Tani, T., Ohsumi, J., Mita, K., and Takiguchi, Y. (1988) Identification of a novel class of elastase isozyme, human pancreatic elastase III, by cDNA and genomic gene cloning. *J. Biol. Chem.* **263,** 1231–1239
40. Shirasu, Y., Takemura, K., Yoshida, H., Sato, Y., Iijima, H., Shimada, Y., Mikayama, T., Ozawa, T., Ikeda, N., and Ishida, A. (1988) Molecular cloning of complementary DNA encoding one of the human pancreatic proteinase E isozymes. *J. Biochem.* **104,** 259–264
41. Kobayashi, R., Kobayashi, Y., and Hirs, C. H. (1978) Identification of a binary complex of procarboxypeptidase A and a precursor of proteinase E in porcine pancreatic secretion. *J. Biol. Chem.* **253,** 5526–5530
42. Kobayashi, R., Kobayashi, Y., and Hirs, C. H. (1981) The specificity of porcine pancreatic proteinase E. *J. Biol. Chem.* **256,** 2460–2465
43. Sziegoleit, A., Linder, D., Schlüter, M., Ogawa, M., Nishibe, S., and Fujimoto, K. (1985) Studies on the specificity of the cholesterol-binding pancreatic proteinase and identification as human pancreatic elastase 1. *Eur. J. Biochem.* **151,** 595–599
44. Szabó, A., Pilsak, C., Bence, M., Witt, H., and Sahin-Tóth, M. (2016) Complex formation of human proelastases with procarboxypeptidases A1 and A2. *J. Biol. Chem.* **291,** 17706–17716
45. Kerfelec, B., Chapus, C., and Puigserver, A. (1985) Existence of ternary complexes of procarboxypeptidase A in the pancreas of some ruminant species. *Eur. J. Biochem.* **151,** 515–519
46. Figarella, C. (1992) What is human pancreatic proelastase 1? *Int. J. Pancreatol.* **11,** 213–215
47. Gomis-Rüth, F. X., Gómez-Ortiz, M., Vendrell, J., Ventura, S., Bode, W., Huber, R., and Avilés, F. X. (1998) Cutting at the right place–the importance of selective limited proteolysis in the activation of proproteinase E. *Eur. J. Biochem.* **251,** 839–844

48. Szenthe, B., Patthy, A., Gáspári, Z., Kékesi, A. K., Gráf, L., and Pál, G. (2007) When the surface tells what lies beneath: combinatorial phage-display mutagenesis reveals complex networks of surface-core interactions in the pacifastin proteinase inhibitor family. *J. Mol. Biol.* **370,** 63–79
49. Szabó, A., Héja, D., Szakács, D., Zboray, K., Kékesi, K. A., Radisky, E. S., Sahin-Tóth, M., and Pál, G. (2011) High affinity small protein inhibitors of human chymotrypsin C (CTRC) selected by phage display reveal unusual preference for P4′ acidic residues. *J. Biol. Chem.* **286,** 22535–22545
50. Héja, D., Harmat, V., Fodor, K., Wilmanns, M., Dobó, J., Kékesi, K. A., Závodszky, P., Gál, P., and Pál, G. (2012) Monospecific inhibitors show that both mannan-binding lectin-associated serine proteinase-1 (MASP-1) and -2 are essential for lectin pathway activation and reveal structural plasticity of MASP-2. *J. Biol. Chem.* **287,** 20290–20300
51. Sidhu, S. S., Lowman, H. B., Cunningham, B. C., and Wells, J. A. (2000) Phage display for selection of novel binding peptides. *Methods Enzymol.* **328,** 333–363
52. Laskowski, M., and Qasim, M. A. (2000) What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? *Biochim. Biophys. Acta* **1477,** 324–337
53. Renaud, A., Lestienne, P., Hughes, D. L., Bieth, J. G., and Dimicoli, J. L. (1983) Mapping of the S′ subsites of porcine pancreatic and human leucocyte elastases. *J. Biol. Chem.* **258,** 8312–8316
54. Schellenberger, V., Schellenberger, U., Mitin, Y. V., and Jakubke, H. D. (1989) Characterization of the ′-subsite specificity of porcine pancreatic elastase. *Eur. J. Biochem.* **179,** 161–163
55. Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14,** 1188–1190
56. Nozach, H., Fruchart-Gaillard, C., Fenaille, F., Beau, F., Ramos, O. H., Douzi, B., Saez, N. J., Moutiez, M., Servent, D., Gondry, M., Thaï, R., Cuniasse, P., Vincentelli, R., and Dive, V. (2013) High throughput screening identifies disulfide isomerase DsbC as a very efficient partner for recombinant expression of small disulfide-rich proteins in *E. coli*. *Microb. Cell Fact.* **12,** 37

Overlapping Specificity of Duplicated Human Pancreatic Elastase 3 Isoforms and Archetypal Porcine Elastase 1 Provides Clues to Evolution of Digestive Enzymes

Eszter Boros, András Szabó, Katalin Zboray, Dávid Héja, Gábor Pál and Miklós Sahin-Tóth

J. Biol. Chem. 2017, 292:2690-2702.
doi: 10.1074/jbc.M116.770560 originally published online January 6, 2017