

The role of elastases in pancreatic diseases

Ph.D. Thesis

Anna Zsófia Tóth M.D.

Supervisors:

Prof. Péter Hegyi, M.D., Ph D., D.Sc.

Prof. Miklós Sahin-Tóth, M.D., Ph.D., D.Sc



Doctoral School of Theoretical Medicine

Szeged

2018.

Content

| | |
|---|----|
| Content | 1 |
| 1 List of abbreviations | 3 |
| 2 Introduction | 5 |
| 2.1 Exocrin pancreatic insufficiency | 5 |
| 2.2 Pancreatic elastases | 5 |
| 2.3 Pancreatic function tests based on detection of elastases | 6 |
| 2.4 Possible role of elastases in CP | 7 |
| 3 Aims | 8 |
| 3.1 Elastase detection study | 8 |
| 3.2 Genetic analysis study | 8 |
| 4 Matherials and methods..... | 8 |
| 4.1 Experiments for both study | 8 |
| 4.1.1 Plasmid construction and mutagenesis | 8 |
| 4.1.2 Cell culture and transfection..... | 8 |
| 4.1.3 Enzyme activity measurements | 9 |
| 4.1.4 SDS-page..... | 9 |
| 4.1.5 Purification of pancreatic proteinases..... | 9 |
| 4.2 Experiments with the ScheBo Pancreatic Elastase 1 test | 10 |
| 4.2.1 Materials | 10 |
| 4.2.2 ELISA assays | 10 |
| 4.3 Performing genetic study..... | 10 |
| 4.3.1 Study Subjects | 10 |
| 4.3.2 DNA Sequencing..... | 10 |
| 5 Results | 11 |
| 5.1 ScheBo Pancreatic Elastase 1 Test results..... | 11 |
| 5.1.1 The ScheBo Pancreatic Elastase 1 Stool Test detects elastase 3 isoforms CELA3A and CELA3B 11 | |
| 5.1.2 Proelastases, active elastases and autolyzed elastase forms are measured with equal efficacy 11 | |
| 5.1.3 Effect of CELA3A and CELA3B genetic variants on the performance of the ScheBo test 11 | |
| 5.1.4 Glu154 in CELA3B is a critical determinant of recognition by the ScheBo test | 12 |
| 5.1.5 Mutations of Lys154 and nearby Arg179 improve detection of CELA3A | 13 |
| 5.1.6 DNA Sequence Analysis of Exon 7 of Human CELA3A and CELA3B | 13 |
| 5.1.7 Functional Analysis of Missense Variants and the Gene Conversion Event..... | 14 |
| 6 Discussion | 15 |

| | | |
|-----|--------------------------------|----|
| 6.1 | Elastase detection study..... | 15 |
| 6.2 | Genetic analysis study | 16 |
| 7 | Summary and new findings | 18 |
| 8 | Acknowledgements | 18 |
| 9 | Financial support | 19 |

1 List of abbreviations

A: adenin
ACP: alcoholic chronic pancreatitis
C: citosine
CELA: chymotripsin-like elastase
CFTR: cystic fibrosis transmembrane conductance regulator
CI: confidential interval
CO₂: carbon dioxide
CP: chronic pancreatitis
CPA1: carboxypeptidase A1
CTRB: chymotrypsin B
CT: computer tomography
CTRC: chymotrypsin
CTRL1: chymotrypsin-like protease
DNA: deoxyribonucleic acid
DMEM: Dulbecco's Modified Eagle's Medium
ELISA: enzyme-linked immunosorbent assay
ERCP: endoscopic retrograde cholangiopancreatography
G: guanin
HCl: hydrochloric acid
ICP: idiopathic chronic pancreatitis
MRCP: magnetic resonance cholangiopancreatography
min: minute
NHLBI: National Heart, Lung, and Blood Institute (USA)
OR: odd ratio
PCR: polymerase chain reactions
PNGase f: Peptide:N-glycosidase F
PRSS1: protease serine 1, human cationic trypsinogen
SPINK: serine protease inhibitor Kazal type 1
SDS: sodium dodecyl sulfate
T: timin
UV: ultraviolet
9 His: 9 histidine

Publications related to the subject of the thesis:

Detection of human elastase isoforms by the ScheBo pancreatic elastase 1 test

Tóth Anna Zsófia, Szabó András, Hegyi Eszter, Hegyi Péter, Sahin-Tóth Miklós

American Journal of Physiology-Gastrointestinal and Liver Physiology 2017. March **IF:3,4**

Genetic Analysis of Human Chymotrypsin-Like Elastases 3A and 3B (CELA3A and CELA3B) to Assess the Role of Complex Formation between Proelastases and Procarboxypeptidases in Chronic Pancreatitis.

Párniczky Andrea, Hegyi Eszter, **Tóth Anna Zsófia**, Szücs Ákos, Szentesi Andrea, Vincze Áron, Izbéki Ferenc, Németh Balázs Csaba, Hegyi Péter, Sahin-Tóth Miklós

International Journal of Molecular Sciensis. 2016. December **IF: 3,2**

Publications closely related to the subject of the thesis:

Pain in the Early Phase of Pediatric Pancreatitis (PINEAPPLE Trial): Pre-Study Protocol of a Multinational Prospective Clinical Trial.

Zsoldos Fanni, Párniczky Andrea, Mosztbacher Dóra, **Tóth Anna Zsófia**, Lásztity Natália, Hegyi Péter; Hungarian Pancreatic Study Group and the International Association of Pancreatology.

Digestion. 2016. December **IF: 2,088**

Analysis of Pediatric Pancreatitis (APPLE Trial): Pre-Study Protocol of a Multinational Prospective Clinical Trial.

Párniczky Andrea, Mosztbacher Dóra, Zsoldos Fanni, **Tóth Anna Zsófia**, Lásztity Natália, Hegyi Péter; Hungarian Pancreatic Study Group and the International Association of Pancreatology.

Digestion. 2016 November **IF: 2,088**

| |
|---|
| Number of full publications: 4 (first author: 1) |
|---|

| |
|--|
| Cumulative impact factor: 10,76 |
|--|

2 Introduction

2.1 Exocrin pancreatic insufficiency

Diseases of the pancreas that result in loss of functional acinar cells can compromise digestive enzyme production and eventually lead to maldigestion. The main etiologic factors of exocrine pancreatic insufficiency in adults are chronic pancreatitis, pancreatic cancer and resection of pancreas. It can also occur in progression of cystic fibrosis especially in infancy and childhood; in conditions with severe protein-energy malnutrition or after a total/subtotal gastric resection. There are other rare diseases which can lead to exocrine pancreatic insufficiency (Shwachmann-Diamond syndrome, Johanson-Blizzard syndrome, or congenital enzyme insufficiencies such as congenital insufficiency of trypsinogen-, enteropeptidase-, alpha-1 antitrypsin-, amylase-, and lipase). The principal symptoms of pancreatic insufficiency are abdominal pain, bloating, cramping, increased flatulence, diarrhea, steatorrhea and malnutrition. The functional reserve capacity of the exocrine human pancreas is high. According to this observation the main symptoms of exocrine pancreatic insufficiency generally appear at the late stage of the pancreatic diseases, for example steatorrhea does not occur until approximately 90 percent of glandular function has been lost. Impaired enzyme secretion may be complicated by latent fat-soluble vitamin deficiencies.

Clinical laboratory tests that quantify decreased digestive enzyme output can aid in the diagnosis of pancreatic insufficiency. The most widely used tests measure levels of the pancreatic elastase enzyme in the stool.

2.2 Pancreatic elastases

The chymotrypsin-like elastases (CELAs) are digestive serine proteinases secreted by the pancreas in multiple isoforms. Due to its cationic character, CELA1 can adsorb to the surface of the negatively charged elastin fibers and cleave multiple Ala-Ala and Ala-Gly peptide bonds. Despite its name, CELA1 is not a specific elastin-degrading enzyme and it readily digests a variety of dietary protein substrates. The primary specificity pocket of CELA1 accommodates small (Ala, Ser) and aliphatic (Ile, Leu, Met, Val) amino-acid side-chains at the so-called P1. Curiously, while the human CELA1 gene appears to be potentially functional, it is not expressed in the pancreas due to evolutionary mutations in its promoter and enhancer regions. A second pancreatic elastase (CELA2) was identified based on its

ability to solubilize elastin. Unlike CELA1, this elastase exhibits chymotrypsin-like P1 specificity and prefers to cleave after aromatic (Tyr, Phe) and aliphatic (Leu, Met) P1 amino-acids. In humans evolutionary duplication of CELA2 gave rise to the CELA2A and CELA2B genes. Even though both genes are expressed at the mRNA level, only the CELA2A enzyme is functional as CELA2B seems to have accumulated inactivating evolutionary mutations. The CELA2A content of pancreatic juice corresponds to about 10% of total protein. Arguably, CELA3 has the most interesting history and characteristics among the human elastases. This elastase gene is also duplicated in humans and the two closely related isoforms were designated as CELA3A and CELA3B. Both are expressed in the pancreas at comparable mRNA and protein levels. Substrate specificity of human CELA3A and CELA3B appear to be similar to that of porcine CELA1, broadly directed toward aliphatic P1 side chains. It was also determined that the CELA3B content of human pancreatic juice accounts for 4-6% of total protein. Thus, the combined levels of CELA3A and CELA3B are similar to those of CELA2A. Spurred by the observation that CELA3B suffers no proteolytic degradation during intestinal transit and appears in the stool in high concentrations. ELISA tests have been developed for the detection of stool elastase and their clinical utility in the diagnosis of pancreatic insufficiency has been demonstrated.

2.3 Pancreatic function tests based on detection of elastases

Among non-invasive pancreatic function tests the determination of fecal elastases by ELISA method is a non-invasive, simple, practical and reliable clinical test. The first sandwich-type ELISA for detecting human pancreatic elastases was developed by Sziegoleit et al. in 1989. Nowadays there are two commercially available pancreatic elastase tests: Schebo Pancreatic Elastase-1 test - based on monoclonal ELISA technique and the BioServ Diagnostics Elastase-1 test – based on polyclonal ELISA technique. According to previous studies both assays are fundamentally equivalent in their ability to classify patients as normal, or as moderately or severely pancreatic insufficient.

One of the most widely used assays is the Pancreatic Elastase 1 Stool Test by ScheBo Biotech AG (Giessen, Germany), which utilizes two monoclonal antibodies raised against CELA3B to measure enzyme levels. However, it remains unclear whether the test also detects other elastases, CELA3A in particular, and to what extent the homologous pancreatic trypsins and chymotrypsins might interfere with the assay. More importantly, the potential confounding effect of natural CELA3 variants on test performance has never been evaluated.

2.4 Possible role of elastases in CP

Chronic pancreatitis (CP) is a progressive, relapsing inflammatory disorder of the pancreas which often develops in the background of genetic susceptibility. Among mutations in the best characterized risk genes PRSS1 (cationic trypsinogen), SPINK1 (pancreatic secretory trypsin inhibitor), and CTSC (chymotrypsin C) which stimulate the activation of trypsinogen resulting elevated trypsin activity in the pancreas, the loss-of-function variants in the CPA1 gene encoding carboxypeptidase A1 were shown to increase risk for early onset CP. The majority of impaired CPA1 variants exhibited a secretion defect due to intracellular misfolding and retention. The mechanism of action of CPA1 variants was unrelated to trypsinogen activation or trypsin activity and seemed to involve endoplasmic reticulum stress as a consequence of misfolding. It is also possible that reduced CPA1 secretion might cause predisposition to CP by other mechanisms. In this respect, we noted that procarboxypeptidase A1 (proCPA) often forms complexes with proelastases in the mammalian pancreas. Consequently, changes in CPA1 levels in the secretory pathway and pancreatic juice might have a significant impact on proelastases; possibly facilitating ectopic elastase activation and thereby contributing to pancreatic injury. CELA2A and CELA3B bind to proCPA1 and CELA3B also binds to proCPA2. Even though CELA3A is 92% identical with CELA3B in its primary structure, it does not form tight complexes with proCPA1 or proCPA2. It is known that a major determinant of binding affinity was amino acid 241, which is Gly in CELA3A and Ala in CELA3B. Mutation p.G241A in CELA3A increases whereas mutation p.A241G in CELA3B reduces binding to proCPA1. Notably, position 241 is polymorphic in both elastases with minor allele frequencies of about 2% in subjects of European origin. This genetic variation should translate to individual differences in complex formation between proelastases and proCPA1. More importantly, the polymorphic nature of amino-acid 241 offers the unique opportunity to perform a genetic association study to investigate whether changes in complex formation between proelastases and procarboxypeptidases affect CP risk.

3 Aims

3.1 Elastase detection study

This study was undertaken to clarify:

- I. which human pancreatic elastase isoforms are detected by the ScheBo Pancreatic Elastase 1 Stool Test and
- II. whether naturally-occurring genetic variants influence the performance of this test.

3.2 Genetic analysis study

In this study we performed a genetic association study to investigate

- III. whether changes in complex formation between proelastases and procarboxypeptidases affect on the risk of CP.
- IV. In addition to the position 241 variants we wanted to discover other elastase mutations which might have been associated with pancreatic diseases.

4 Matherials and methods

4.1 Experiments for both study

4.1.1 Plasmid construction and mutagenesis

Expression plasmids for human elastases CELA2A, CELA3A, CELA3B and chymotrypsins CTB1, CTB2, CTB3 and CTRL1 constructed in the pcDNA3.1(-) vector and for human trypsins PRSS1, PRSS2 and PRSS3 in the pTrapT7 vector were described previously. The plasmids contain the coding DNA for the proenzyme (zymogen) form of the indicated pancreatic proteinases. Mutations in CELA3A and CELA3B were introduced by overlap extension PCR mutagenesis.

4.1.2 Cell culture and transfection

For small scale expression studies, human embryonic kidney (HEK) 293T cells were grown in six-well tissue culture plates (1.5×10^6 cells per well) in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum, 4 mM glutamine and 1%

penicillin/streptomycin, at 37 °C in a humidified atmosphere containing 5% CO₂. Transfections were performed with 4 µg expression plasmid with 10 µL Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) in 2 mL DMEM. After overnight incubation cells were rinsed and covered with 2 mL Opti-MEM Reduced Serum Medium. The conditioned media were harvested after 48 h.

4.1.3 Enzyme activity measurements

Enzymatic activity of human CELA3A and CELA3B in the conditioned medium of transfected cells was determined using the Suc-Ala-Ala-Pro-Ala-p-nitroanilide substrate. To activate proelastases, aliquots of conditioned media (100 µL) were supplemented with 10 µL of 1 M Tris-HCl (pH 8.0) and 1 µL 0.1 M CaCl₂ and incubated with 100 nM human cationic trypsin at 37 °C for 30 min (final concentrations). Activated elastases (20 µL) were then mixed with 175 µL assay buffer (0.1 M Tris-HCl (pH 8.0), 1 mM CaCl₂, 0.05% Tween 20) and elastase activity was measured by adding 5 µL of 6 mM substrate. The increase in absorbance at 405 nm was followed for 5 min in a microplate reader at 22 °C. Rates of substrate cleavage were calculated from the linear portion of the curves and expressed in mOD/min units.

4.1.4 SDS-page

Conditioned media (180 µL) were precipitated with 10% trichloroacetic acid (final concentration); the precipitate was collected by centrifugation and resuspended in 20 µL Laemmli sample buffer containing 100 mM dithiothreitol, heat-denatured at 95 °C for 5 min and run on 15% SDS-polyacrylamide gels. The gels were stained with Coomassie Blue R-250. Densitometric quantitation of bands was carried out with the Gel Doc XR+ gel documentation system and Image Lab software (Bio-Rad, Hercules, CA, USA).

4.1.5 Purification of pancreatic proteinases

Human proelastases and chymotrypsinogens were expressed in transiently transfected HEK 293T cells and purified from the conditioned medium through their C-terminal His tags by nickel-affinity chromatography. Human trypsinogens were expressed in *E. coli* as insoluble inclusion bodies. Refolding and purification on immobilized ecotin were performed according to our earlier published protocol.

4.2 Experiments with the ScheBo Pancreatic Elastase 1 test

4.2.1 Materials

The ScheBo Pancreatic Elastase 1 Stool Test was purchased from the manufacturer. For some of the experiments we used the ScheBo Pancreatic Elastase 1 Serum Test which contains essentially the same ELISA components.

4.2.2 ELISA assays

The ELISA detection of human pancreatic proteinases by the ScheBo test was performed according to the manufacturer's instructions with the ready-to-use reagents supplied. All assays were performed in duplicates. Data points plotted represent absorbance readings corrected for the average of two blank values.

4.3 Performing genetic study

4.3.1 Study Subjects

De-identified genomic DNA samples were obtained from the registry of the Hungarian Pancreatic Study Group (ethical approval: TUKEB 22254-1/2012/EKU, biobanking approval: IF702-19/2012). Individuals in the registry were recruited from 11 Hungarian centers between 2012 and 2016 and all gave informed consent according to the ethical guidelines of the Declaration of Helsinki. The study was approved by the Institutional Review Board at Boston University ("Analysis of susceptibility genes in patients with chronic pancreatitis"; IRB number H-35382). A total of 225 unrelated patients with CP, including 120 with alcoholic CP (ACP) and 105 with idiopathic CP (ICP) and 300 control subjects with no pancreatic disease were studied. Sequence analysis of CELA3A and CELA3B was successfully completed for 295 and 293 of the 300 control samples, respectively.

4.3.2 DNA Sequencing

Primers were designed against intronic sequences flanking exon 7 in CELA3A and CELA3B. Polymerase chain reactions (PCR) were performed using 0.75 U HotStar Taq DNA polymerase (Qiagen, Valencia, CA, USA), 0.2 mM dNTP, 5 μ L 5 \times Q-solution and 2.5 μ L 10 \times PCR buffer (Qiagen), 0.5 μ M primers, and 10–50 ng genomic DNA template in a volume of 25 μ L. PCR reactions were started by a 15-min initial heat activation at 95 $^{\circ}$ C followed by 35 cycles of 30 s denaturation at 94 $^{\circ}$ C, 30-s annealing at 55 $^{\circ}$ C, and 40 s extension at 72 $^{\circ}$ C; and finished by a final extension for 5 min at 72 $^{\circ}$ C. Products of all PCR reactions were verified by 2% agarose gel electrophoresis. The PCR amplicons (5 μ L) were treated with 1

μL FastAP Thermosensitive Alkaline Phosphatase and 0.5 μL Exonuclease I (Thermo Fisher Scientific, Waltham, MA, USA) for 15 min at 37 °C and the reaction was stopped by heating the samples to 85 °C for 15 min. Sanger sequencing was performed using the reverse PCR primer as sequencing primer. Amplicons containing heterozygous deletion variants in intron 6 of CELA3A were also sequenced with the forward PCR primer. New variants were confirmed by a second independent PCR amplification and also by sequencing the other DNA strand.

5 Results

5.1 ScheBo Pancreatic Elastase 1 Test results

5.1.1 The ScheBo Pancreatic Elastase 1 Stool Test detects elastase 3 isoforms CELA3A and CELA3B

To determine the specificity of the ScheBo test, we recombinantly expressed and purified human pancreatic serine proteinases and performed ELISA assays according to the manufacturer's instructions. Elastases CELA3A and CELA3B were tested at 100 pM final concentration while all other proteinases were tested at the 10-fold higher 1 nM concentration. We obtained strong signals for CELA3A and CELA3B whereas none of the other proteinases was detected to a significant extent.

5.1.2 Proelastases, active elastases and autolyzed elastase forms are measured with equal efficacy

To characterize the detection of CELA3A and CELA3B by the ScheBo test in a more quantitative manner, we performed the ELISA assay using elastases over a concentration range from 20 to 200 pM. CELA3B was detected on average 4.3-fold better (range 3.4 – 5.1-fold) than CELA3A over the concentration range tested. Identical signals were obtained when proelastases were compared to active elastases indicating that the test measures the zymogen and active forms with equal efficacy. Autolyzed forms of CELA3B missing either 9 amino acids (9del) or 13 amino acids (13del) from the N terminus also produced ELISA signals identical to that of the intact CELA3B proelastase..

5.1.3 Effect of CELA3A and CELA3B genetic variants on the performance of the ScheBo test

To identify genetic variants of CELA3A and CELA3B in the population, we interrogated the Exome Variant Server database of the NHLBI Exome Sequencing Project. When considering missense variants only with an allele frequency above 1%, we found one CELA3A variant (G241A) and five CELA3B variants (W79R, Q134L, I209V, R210H and

A241G). To evaluate whether common genetic variants of CELA3A and CELA3B might alter the performance of the ScheBo test, we purified these six variants and tested their detection in the ELISA assay. With the exception of the CELA3B W79R variant, none of the variants had an appreciable effect on signal development. Variant W79R was detected on average 1.4-fold (range 1.2 – 1.5-fold) better than wild-type CELA3B over the concentration range tested. However, considering that in most carriers the variant is heterozygous, this difference should have no meaningful impact on the clinical interpretation of the ScheBo test results.

5.1.4 Glu154 in CELA3B is a critical determinant of recognition by the ScheBo test

CELA3A and CELA3B share 92% identity at the amino-acid level yet CELA3B is detected circa 4-fold better than CELA3A by the ELISA assay. To identify the reason for this difference, we aligned the two isoforms and then individually mutated all divergent amino-acid positions in CELA3B to the corresponding CELA3A amino acid. Positions where differences occurred in neighboring amino acids were mutated en bloc. Overall 11 new CELA3B mutants were constructed and tested. For these qualitative screening experiments we used conditioned media of HEK 293T cells transfected with the mutant constructs. Remarkably, mutant E154K gave no ELISA signal whatsoever while all other mutants were robustly detected with some variations in signal yield. Since the mutations may alter secretion and enzymatic activity of CELA3B, we also verified the expression of all mutants by SDS-PAGE and Coomassie staining and by direct activity measurements after activation with trypsin. All mutants, including E154K, exhibited measurable elastase activity and for the majority of mutants activity was comparable or even higher than that of the wild-type CELA3B. The higher activity of mutants S77R,S78D,W79L and D89N,R90L was due to higher elastase amounts secreted to the conditioned medium (not shown) and this was also consistent with the stronger ELISA signal. Similarly, mutant A241G was secreted to higher levels and produced a higher ELISA signal but this change was not obvious in the activity measurement as this mutation decreases catalytic activity of CELA3B. Taken together, this initial screen conclusively identified Glu154 in CELA3B as a major determinant of recognition by the ScheBo test.

To confirm the importance of Glu154, we purified the E154K mutant and compared detection to wild-type CELA3B over a 20-200 pM concentration range. No signal was obtained with the mutant. The ScheBo test uses a sandwich assay format with separate capturing and detection antibodies directed at different regions of the elastase molecules. To ascertain whether the defect with the E154K mutant is at the level of capturing or detection, we

eliminated the capturing step by immobilizing wild-type and mutant CELA3B to nickel plates (Ni-NTA HisSorb plate, Qiagen, Valencia, CA) via their His-tag. Under these assay conditions both elastase forms were detected comparably, indicating that mutation E154K interferes with the capturing step in the ELISA protocol. Finally, structural modeling indicated that Glu154 is located on the surface of CELA3B far removed from the active site.

5.1.5 Mutations of Lys154 and nearby Arg179 improve detection of CELA3A

The experiments presented above strongly indicated that detection of CELA3A should be improved by changing the Lys154 residue to Glu. Surprisingly, however, when the CELA3A mutant K154E was purified and tested, no improvement in detection was observed over wild-type CELA3A. To explain these puzzling observations, we speculated that the presence of another divergent amino-acid that lies in proximity interferes with the recognition of Glu154 in the CELA3A K154E mutant. Inspection of the CELA3B structural model indicated that amino acid 179, which is Arg in CELA3A and Leu in CELA3B, might be important in this. Indeed, when mutations K154E and R179L were introduced simultaneously in CELA3A, the ELISA signal was increased by about 2.5-fold and approximated that of CELA3B.

5.1.6 DNA Sequence Analysis of Exon 7 of Human CELA3A and CELA3B

To investigate whether changes in complex formation between human procarboxypeptidases and proelastases alter risk for CP, we investigated the frequency of variants c.722G>C (p.G241A) in CELA3A and c.722C>G (p.A241G) in CELA3B in subjects with CP and controls without pancreatic disease. We sequenced exon 7 and flanking intronic regions of CELA3A and CELA3B in 225 patients and 300 controls from the registry of the Hungarian Pancreatic Study Group. This CP cohort consisted of 120 alcoholic chronic pancreatitis (ACP) and 105 idiopathic chronic pancreatitis (ICP) patients. Sequence analysis of CELA3A and CELA3B was successfully completed for all patient samples and for 295 and 293 of the 300 control samples, respectively. In CELA3A we found 8 variants which included 4 intronic variants, 3 synonymous variants and 1 non-synonymous (missense) variant. Synonymous variants c.750C>T (p.P250=) and c.753G>A (p.T251=) were found in complete linkage disequilibrium. In CELA3B we detected 13 variants which included 6 intronic variants, 3 synonymous variants, 3 non-synonymous (missense) variants and a gene-conversion event resulting in five nucleotide alterations that changed three amino-acids at the protein level. Synonymous variants c.699T>C (p.H233=) and c.702C>T (p.G234=) were found in the same ACP patient..

When allele frequency was considered, distribution of the variants between patients and controls showed a significant difference only for the CELA3B variant c.643-7G>T in intron 6, which occurred with an allele frequency of 16% in patients and 21.3% in controls (OR = 0.7; 95% CI 0.51–0.97; $p = 0.03$). Subgroup analysis revealed that the association was driven by the ACP cohort (OR = 0.59, 95% CI = 0.39–0.89, $p = 0.01$) while it was not significant in ICP patients ($p = 0.4$). Importantly, neither p.G241A in CELA3A nor p.A241G in CELA3B was significantly enriched in patients or controls. Novel missense variants identified in CELA3B were rare; the heterozygous c.694G>C (p.V232L) variant was found in one ACP patient and one control, whereas the heterozygous c.740G>C (p.R247P) variant was found in one control subject only. Both subjects with the p.V232L variant also carried a heterozygous c.643-7G>T CELA3B variant, while the subject with the p.R247P variant carried a heterozygous c.643-26C>T CELA3B variant. An unexpected observation in CELA3B was a heterozygous gene conversion event detected in a patient with ICP. A minimum of seven nucleotides in CELA3B between c.736 and c.742 was converted with the corresponding CELA3A sequence. This resulted in five nucleotide changes (c.736A>T, c.737C>T, c.739C>A, c.740G>T, c.742A>T) and three amino-acid changes (p.T246F, p.R247I, p.R248W) in CELA3B.

Genotypes were also analyzed for common variants using dominant and recessive models but no significant differences were found between CP patients and controls. In subgroup analysis of variant c.643-7G>T genotypes using a dominant model (GT plus TT in patients versus controls) we confirmed the underrepresentation of this variant in ACP patients (OR = 0.6, 95% CI 0.37–0.96, $p = 0.03$), whereas the same association could not be verified in the ICP group ($p = 0.52$). Using a recessive model (TT genotypes in patients versus controls) a more obvious difference was observed in the ACP group (1.7%) than in ICP patients (3.8%) versus controls (5.8%), however, due to the limited sample size it did not reach statistical significance.

5.1.7 Functional Analysis of Missense Variants and the Gene Conversion Event

To characterize how missense mutations in CELA3A and CELA3B and the gene conversion in CELA3B alter elastase secretion and activity, we transfected human embryonic kidney (HEK) 293T cells with expression plasmids and measured proelastase levels in the conditioned medium by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and elastase activity by enzymatic assays. When elastase secretion into the medium was determined at 48 h after transfection, all variants were secreted normally, as judged by

Coomassie Blue staining and densitometry of stained gels. CELA3B variant p.V232L exhibited a slight (~20%) but reproducible increase in secretion. Similarly, elastase activity of the conditioned medium after activation with trypsin showed no major differences among the variants and wild-type enzymes when measured with the Suc-Ala-Ala-Pro-Ala-p-nitroanilide chromogenic peptide substrate. Consistent with its increased levels of secretion, variant p.V232L exhibited higher activity while the activity of the gene conversion variant (p.T246F, p.R247I, p.R248W) was slightly (1.4-fold) reduced relative to wild-type CELA3B. Interestingly, variants p.G241A in CELA3A and p.A241G in CELA3B had opposite effects on elastase activity. Thus, activity of the CELA3A p.G241A variant was increased by 1.8-fold, whereas activity of the CELA3B p.A241G variant was decreased by 2.2-fold relative to their respective wild-type controls. Even though these changes in activity controlled by the amino acid at position 241 (Gly versus Ala) are intriguing mechanistically, they have no impact on CP risk as demonstrated by the lack of genetic association described above.

6 Discussion

6.1 Elastase detection study

In this study, we evaluated the isoform specificity of the ScheBo Pancreatic Elastase 1 Stool Test. Previously the exact molecular targets of this ELISA assay have not been characterized; it has been unclear which elastase isoform(s) the test detects, and whether other homologous pancreatic proteinases interfere with the assay. Thus, we expected the test would detect CELA3B; yet the vexing questions remained whether the 92% identical CELA3A isoform was equally well detected and whether natural genetic variants of CELA3A and CELA3B affected test performance. Our findings confirmed that the primary target of the test is CELA3B, however, CELA3A is also detected, approximately 4-fold weaker than CELA3B. Importantly, CELA2A, chymotrypsins CTRB1, CTRB2, CTRC, CTRL1 and trypsins PRSS1, PRSS2, PRSS3 produced minimal or no signal in this ELISA assay. We compared test performance on all molecular forms of CELA3B and found no difference in detection efficiency, indicating that test results are unaffected by the activation state or the degree of autolysis of secreted elastases.

We characterized the effect on test performance of all genetic variants of CELA3A and CELA3B, which occur at or above 1% frequency in the population and found no clinically

meaningful changes. Only one variant, CELA3B W79R, exhibited increased detection by approximately 1.4-fold, which should be inconsequential in heterozygous carriers. Our studies do not rule out the possibility that rare or private genetic variants in certain patients may interfere with the test, however, for the large majority of the population the ScheBo test should not be affected by common genetic variants in the CELA3 isoforms.

We performed limited epitope mapping to identify amino-acids that are responsible for the preferential detection of CELA3B over CELA3A by the ScheBo test. We identified Glu154 in CELA3B, which is Lys in CELA3A, as a key determinant of recognition by the capturing monoclonal antibody. Mutation E154K in CELA3B abolished detection by the ELISA assay. Surprisingly, however, the opposite mutation K154E in CELA3A did not improve detection by the test and simultaneous mutation of the nearby Arg179 to Leu (R179L) was required to achieve signal levels that were comparable to those of CELA3B.

6.2 Genetic analysis study

In the present study we tested the hypothesis that changes in complex formation between human proelastases and procarboxypeptidases might alter risk for CP. We speculated that individual variations in the levels of these zymogen complexes may influence digestive enzyme activation as the free, uncomplexed enzyme should be more susceptible to activation than its complexed form. Indeed, we recently demonstrated that activation of human procarboxypeptidases A1 and A2 were delayed when bound to human proCELA3B. Although activation of proCELA3B did not seem to be directly affected by complex formation, changes in folding, solubility, packaging and secretion might be altered which can ultimately promote ectopic elastase activation inside the pancreas. To identify whether such mechanisms play a role in CP risk, we took advantage of the observation that the primary amino-acid determinants of complex formation in CELA3A and CELA3B are polymorphic in the population. Thus, normally proCELA3A binds poorly to procarboxypeptidases due to the presence of Gly241, whereas proCELA3B which contains Ala241 forms tight complexes. Individuals carrying a heterozygous p.G241A mutation in CELA3A should have increased complex levels in their pancreatic secretions. Conversely, subjects with the p.A241G variant in CELA3B should exhibit decreased complex formation and higher levels of free zymogens. Should CP risk be elevated by decreased complex formation, one would expect to observe underrepresentation of the CELA3A p.G241A variant and/or enrichment of the CELA3B p.A241G variant in a cohort of CP patients versus population controls. To determine allele

frequencies of these variants, we sequenced exon 7 and flanking intronic regions of CELA3A and CELA3B in CP patients and controls of Hungarian origin. We found that allele frequencies of CELA3A p.G241A and CELA3B p.A241G were not significantly different between the two groups, indicating that these variants do not alter CP risk. Despite the lack of statistical significance, an opposing trend for enrichment of the two variants was apparent. Thus, CELA3A variant p.G241A was slightly overrepresented in controls while CELA3B variant p.A241G was overrepresented in patients. Assuming the OR values represent the true effect sizes, we would need a much larger cohort to confirm these findings with statistical significance. These effects, on the other hand, would be still clinically insignificant, which does not justify further extension of the studies.

In addition to the position 241 variants discussed above, we identified six synonymous variants, two missense variants, a gene conversion event and ten variants in the flanking intronic regions. Intriguingly, CELA3B variant c.643-7G>T in intron 6 was significantly underrepresented in ACP patients, indicating an approximately 1.7-fold protective effect against CP. This variant is located relatively close to the pre-mRNA splice site and it is conceivable that it might disrupt splicing and decrease CELA3B expression. This finding suggests that elastase activity contributes to pancreatitis risk and warrants extension of sequence analysis of the CELA3A and CELA3B genes and replication of the observed association(s) in an independent cohort. None of the other variants showed a significant difference in frequency between CP patients and controls. Finally, we performed functional analysis on all missense variants detected in this study and found no major defects in proelastase secretion or elastase activity.

We discovered a rare gene conversion event between exons 7 of CELA3A and CELA3B in an individual with ICP. Gene conversion is a non-reciprocal exchange of genetic information between homologous DNA sequences, most likely as a result of mismatch repair following a heteroduplex formation between the donor and acceptor genes. DNA exchange by gene conversion is of paramount importance for the evolution of gene families but it can also cause human diseases including CP and other pancreatic pathologies. Gene conversion events between PRSS1 and PRSS2 or between PRSS1 and the pseudogene PRSS3P2 were shown to generate pathogenic alleles that cause hereditary pancreatitis. More recently, a recombination allele of the carboxyl ester lipase gene (CEL) and its pseudogene CELP was described as a novel genetic risk factor for CP. The newly found gene conversion event between CELA3A and CELA3B was detected in a CP patient, however, this is likely an accidental finding as we

found no functional defect with respect to secretion or activity of the converted CELA3B that might suggest a pathogenic role in CP.

7 Summary and new findings

In this work we investigated to clarify the role of pancreatic elastases in the exocrine pancreatic diseases from molecular mechanism to the bedside.

- I. We characterized the molecular targets of the ScheBo Pancreatic Elastase 1 Stool Test and demonstrated that it predominantly measures CELA3B but also detects CELA3A with lower efficacy.
- II. Other pancreatic proteinases or genetic variants of the CELA3 isoforms have no appreciable impact on test performance.
- III. Our study demonstrated that variants affecting amino-acid position 241 in human CELA3A and CELA3B are not associated with CP, indicating that changes in complex formation between human proelastases and procarboxypeptidases do not influence the risk for CP.
- IV. The observation that intronic variant c.643-7G>T in CELA3B was significantly underrepresented in ACP patients suggests this might be a protective variant.

8 Acknowledgements

First, I would like to express my deepest and sincere gratitude to my mentors **Professor Péter Hegyi** at the First Department of Internal Medicine, University of Szeged; Institute for Translational Medicine, University of Pécs and **Professor Miklós Sahin-Tóth** at the Department of Molecular and Cell Biology, Boston University Medical Center for all of their guidance, patience, inspiration and support during my time as a PhD student. Without their outstanding supervision, help and encouragement this work and this PhD thesis would not have been possible. I would like to express my thanks for thrusting me and giving me the possibility to have an international experience. They both are real inspiring experts and I really appreciate their knowledge.

I gratefully acknowledges the support and personal help of **Dr. Csaba Bereczki** (Department of Pediatrics, University of Szeged, Szeged, Hungary) who let me interrupt my clinical work and made me possible to working on these project in Boston. This work would not have been possible to accomplish without the help and support of **Dr. Andrea Párniczky, Dr. Eszter Hegyi and Dr. András Szabó**. I would like to express my sincere gratitude for all of their advices, knowledge and patience for designing and performing the experiments. I am grateful to **Dr. Andrea Geisz, Dr. Balázs Németh, Dr. Dóra Mosztbacher, Dr. Fanni Zsoldos, Dr. Ákos Szűcs, Dr. Áron Vince, Dr. Ferenc Izbéki, Dr. Dorottya Berki, Dr. Alexandra Demcsák, Dr. Anita Balázs, Andrea Szentesi, all the members of the Hungarian Pancreatic Study Group** and all the contributors to the Registry for Pancreatic Patients for their help and contribution to this work.

I would like to express my deep and sincere gratitude for **Professor Ferenc Bari** and **Dr. Erzsébet Forczek** for their support and guidance during the medical university, their opinion was really important for me.

I am grateful to **Dr. Endre Neparáczki** for helping me editing this PhD thesis.

Finally i would like to say special thanks to **My Parents** and **Grandparents** for their unconditional help and support in my life.

9 Financial support

This work was supported by NIH grants R01DK095753, R01DK082412 and R01DK058088 (to MST) and the Momentum Grant of the Hungarian Academy of Sciences LP2014-10/2014.