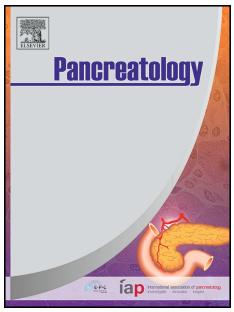

Genetic and functional analysis of chymotrypsin-like protease (CTRL) in chronic pancreatitis

Katharina Eiseler, Lea Neppl, Andreas W. Schmidt, Beate Rauscher, Maren Ewers, Emmanuelle Masson, Jian-Min Chen, Claude Férec, Vinciane Rebours, Tassos Grammatikopoulos, Pierre Foskett, William Greenhalf, Christopher Halloran, John Neoptolemos, Tobias B. Haack, Stephan Ossowski, Marc Sturm, Jonas Rosendahl, Helmut Laumen, Heiko Witt

PII: S1424-3903(23)01839-2

DOI: https://doi.org/10.1016/j.pan.2023.11.002

Reference: PAN 1893

To appear in: *Pancreatology*

Received Date: 22 September 2023

Revised Date: 30 October 2023

Accepted Date: 1 November 2023

Please cite this article as: Eiseler K, Neppl L, Schmidt AW, Rauscher B, Ewers M, Masson E, Chen J-M, Férec C, Rebours V, Grammatikopoulos T, Foskett P, Greenhalf W, Halloran C, Neoptolemos J, Haack TB, Ossowski S, Sturm M, Rosendahl J, Laumen H, Witt H, Genetic and functional analysis of chymotrypsin-like protease (CTRL) in chronic pancreatitis, *Pancreatology* (2023), doi: https://doi.org/10.1016/j.pan.2023.11.002.

Genetic and functional analysis of chymotrypsin-like protease (CTRL) in chronic pancreatitis

Katharina Eiseler[1], Lea Neppl[1], Andreas W. Schmidt[1,2,3], Beate Rauscher[1], Maren Ewers[1], Emmanuelle Masson[4,5], Jian-Min Chen[4], Claude Férec[4], Vinciane Rebours[6], Tassos Grammatikopoulos[7,8], Pierre Foskett[8], William Greenhalf[9], Christopher Halloran[9], John Neoptolemos[10], Tobias B. Haack[3], Stephan Ossowski[3] , Marc Sturm[3], Jonas Rosendahl[2], Helmut Laumen[1,2], Heiko Witt[1]*

* corresponding author

[1] Paediatric Nutritional Medicine, Else Kröner Fresenius Center for Nutritional Medicine (EKFZ), Technical University of Munich (TUM), Freising, Germany

[2] Department of Internal Medicine I, Martin Luther University (MLU), Halle (Saale), Germany

[3] Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany

[4] Univ Brest, Inserm, EFS, UMR 1078, GGB, F-29200 Brest, France

[5] Service de Génétique Médicale et de Biologie de la Reproduction, CHRU Brest, F-29200 Brest, France

[6] Pancreatology and Digestive Oncology Department, Beaujon Hospital, APHP - Clichy, Université Paris Cité, Paris, France

[7] Paediatric Liver, GI & Nutrition Centre and MowatLabs, King's College Hospital NHS Foundation Trust, London, UK

[8] Institute of Liver Studies, King's College London, London, UK

[9] Department of Molecular and Clinical Cancer Medicine, University of Liverpool, Liverpool, UK

[10] Department of Surgery, University of Heidelberg, Heidelberg, Germany

***Corresponding author:**
Heiko Witt, M.D.

Paediatric Nutritional Medicine & Else Kröner-Fresenius-Centre for Nutritional Medicine (EKFZ) Technical University of Munich (TUM)
Gregor-Mendel-Str. 2
85354 Freising
Germany
Phone: +49 8161 71 2466, Fax: +49 8161 71 2465
E-Mail: heiko.witt@tum.de
ORCID ID: 0000-0001-7791-7341

Running title: *CTRL* and chronic pancreatitis

Key words: CTRL, chymotrypsin-like protease, chronic pancreatitis, genetics

Disclosure statement: The authors report no conflicts of interest.

Financial support: The authors thank all study participants for providing clinical data and blood samples. This work was supported by the Else Kröner-Fresenius-Foundation (EKFS) 2017_A108 – EKFZ-Witt (to HW), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) WI 2036/3-1 (to HW), RO 3929/5-1 and RO 3929/6-1 (to JR) and 433158657 (to TBH and to JR), the DFG Research Training Group InCuPanc GRK 2751/1 - 2022 Project number: 449501615 (to JR), the Association Gaétan-Saleün, the Association des Pancréatites Chroniques Héréditaires, and the Institut National de la Santé et de la Recherche Médicale (INSERM), France.

Author contributions:
H.W. conceived and directed the study.
K.E., L.N. and B.R. performed experimental analyses.
A.W.S. and M.S. performed the bioinformatic analysis.
K.E., M.E. and H.W. drafted and revised the manuscript.
All other co-authors recruited study subjects, collected clinical data and/or provided genomic DNA samples or data. All authors approved the final manuscript and contributed critical revisions to its intellectual content.

List of abbreviations:
CP, chronic pancreatitis;
CTRL, chymotrypsin-like protease
REVEL, Rare Exome Variant Ensemble Learner

Abstract

Background. Genetic predisposition is crucial in the pathogenesis of early-onset chronic pancreatitis (CP). So far, several genetic alterations have been identified as risk factors, predominantly in genes encoding digestive enzymes. However, many early-onset CP cases have no identified underlying cause. Chymotrypsins are a family of serine proteases that can cleave trypsinogen and lead to its degradation. Because genetic alterations in the chymotrypsins *CTRC*, *CTRB1*, and *CTRB2* are associated with CP, we genetically and functionally investigated chymotrypsin-like protease (CTRL) as a potential risk factor.

Methods. We screened 1,005 non-alcoholic CP patients and 1,594 controls for *CTRL* variants by exome sequencing. We performed Western blots and activity assays to analyse secretion and proteolytic activity. We measured *BiP* mRNA expression to investigate the potential impact of identified alterations on endoplasmic reticulum (ER) stress.

Results. We identified 13 heterozygous non-synonymous *CTRL* variants: five exclusively in patients and three only in controls. Functionality was unchanged in 6/13 variants. Four alterations showed normal secretion but reduced (p.G20S, p.G56S, p.G61S) or abolished (p.S208F) activity. Another three variants (p.C201Y, p.G215R and p.C220G) were not secreted and already showed reduced or no activity intracellularly. However, intracellular retention did not lead to ER stress.

Conclusion. We identified several *CTRL* variants, some showing potent effects on protease function and secretion. We observed these effects in variants found in patients and controls, and *CTRL* loss-of-function variants were not significantly more common in patients than controls. Therefore, *CTRL* is unlikely to play a relevant role in the development of CP.

Introduction

Chronic pancreatitis (CP) is an inflammatory condition of the pancreas that manifests as recurrent or continuous disease. It often results in significant impairment of both endocrine and exocrine function, leading to maldigestion, steatorrhea, and diabetes mellitus. While alcohol abuse is the most common cause, non-alcoholic forms also exist. These entities, termed hereditary and idiopathic CP, are often genetic, particularly in the case of early-onset disease [1].

Several disease-predisposing genes have been identified over the past few decades. Except for *CFTR*, *TRPV6,* and *CLDN2* [2-4], which show mainly ductal expression, most risk genes encode pancreatic digestive enzymes. Mutations that increase cationic trypsinogen (PRSS1) function or decrease the function of either the trypsin inhibitor SPINK1, or the trypsin-degrading enzyme chymotrypsinogen C (CTRC), are associated with CP. These findings led to the trypsin-centred pathogenic concept of an imbalance between proteases and their inhibitors as a central disease mechanism [5-7].

However, in recent years, variants in other digestive enzymes such as carboxypeptidase A1 (*CPA1*), carboxyl ester lipase (*CEL*), chymotrypsinogen B1 and B2 (*CTRB1-CTRB2*), pancreatic lipase (*PNLIP*), and chymotrypsin-like elastase 3B (*CELA3B*) have also been described in non-alcoholic CP cases. Thus, additional mechanisms are involved in disease pathogenesis [8-12].

In addition to the chymotrypsinogens B1, B2, and C, Larsen and colleagues identified in 1993 a fourth member of the chymotrypsin family on chromosome 16q22.1, termed chymotrypsin-like protease (CTRL) [13]. Because of its sequence similarity to chymotrypsin and its hydrolytic activity, it has been suggested to be a novel digestive enzyme [14]. As chymotrypsin, CTRL also hydrolyses polypeptides at tyrosine (Tyr) and phenylalanine (Phe) residues [15]. CTRL represents a low abundance isoform whose precise physiological function is still unclear. *Ctrl*-deficient mice were generated to study the role of CTRL in caerulein-induced pancreatitis [16]. In an acute pancreatitis model, $Ctrl^{-/-}$ mice showed insignificantly higher intrapancreatic trypsin

activity, but the severity of pancreatitis remained unchanged compared to littermates. However, these mice were not tested in an experimental model of chronic pancreatitis. To clarify the role of CTRL in pancreatitis, we analysed CTRL variants found in CP patients for their secretory and proteolytic function and compared the data with a control collective.

Methods

Study subjects. We investigated 1,005 unrelated European patients with paediatric non-alcoholic CP. Patients were recruited in Germany, France, the UK and Sweden. Additionally, 1,594 individuals served as controls. We diagnosed CP based on two or more findings: a typical history of recurrent pancreatitis, radiological findings such as pancreatic calcifications or pancreatic ductal irregularities revealed by endoscopic retrograde pancreatography, magnetic resonance imaging of the pancreas, or pathological sonographic findings. We excluded patients with alcohol abuse from the study. The institutional review boards of all participating study centres approved this study. All participants gave written informed consent. Control genotypes from unrelated European samples were received from the Institute for Medical Genetics and Applied Genomics Tübingen.

Exome sequencing (ES). ES of CP patients was performed by BGI Genomics in Hong Kong. Genomic DNA from blood leukocytes was analysed. Exon capture was performed using SureSelect Human All Exon V6 (Agilent) target regions with an average mean coverage of 120x and exon-enriched DNA library was sequenced using the Illumina HiSeq 1500 platform. We aligned samples to hg19, followed by indel realignment with ABRA2 [17] and filtering for contamination (SNV allele frequency deviation >5%) and coverage outliers (less than 90% of target regions covered with at least 20x) using the megSAP pipeline (https://github.com/imgag/megSAP/tree/pancreatitis_project). European samples (iAdmix score ≥0.95, based on HapMap V3 TSI and CEU populations [18]) were considered for further analyses. Next, we applied base recalibration, haplotype calling, joint genotyping and allele-specific filtering (AS-VQSR) implemented in GATK 4.2.6.1. Based on a high confidence set of autosomal SNPs (AS-VQSR variant filter status = PASS, call rate >99%, coverage ≥50 in a subset of randomly sampled 100 samples and not located in any low confidence region obtained from github.com/imgag/megSAP/blob/GRCh37/data/misc/low_conf_regions.bed), up to 2nd degree relatives were excluded with the KING algorithm [19] implemented in PLINK 2.0 [20]. Variants were annotated with basic annotations, REVEL [21] scores and allele frequencies from gnomAD 2.1.1 [22] by VEP v.105 [23].

Sanger sequencing of *CTRL* variants. We verified all variants found in CP patients by ES by Sanger sequencing. Primers complementary to intronic sequences flanking *CTRL* exons based on the published nucleotide sequence (NM_001907.2) (**Table 1**) were synthesized by TIB MOLBIOL (Berlin, Germany). We performed polymerase chain reaction (PCR) using 0.75 U OneTaq Hot Start DNA Polymerase (New England Biolabs), 1x OneTaq GC reaction buffer, 400 µmol/L deoxynucleoside triphosphates, and 0.1 µM primers in a total volume of 20 µL. Cycle conditions were as follows: initial denaturation for 5 min at 95 °C; 50 cycles of 20 s at 95 °C, 40 s at 64 °C and 2 min at 68 °C; and a final extension for 5 min at 68 °C. We digested PCR products with antarctic phosphatase and exonuclease I (both New England Biolabs) for 40 min at 37 °C followed by inactivation for 20 min at 85 °C. Sequencing using internal primers (**Table 1**) was performed by Eurofins Genomics (Ebersberg, Germany).

Vector plasmids. We ordered the coding sequence of human *CTRL* in a basic pcDNA3.1+/C-(K)-DYK from GenScript Biotech (Leiden, Netherlands): we used wild-type *CTRL* (OHu00893D; NM_001907.2) and 13 mutations (p.L2S, p.S15Y, p.G20S, p.G56S, p.G61S, p.R106W, p.S112N, p.T150I, p.H173R, p.C201Y, p.S208F, p.G215R, and p.C220G) for functional analyses. The basic vector resulted in an in-frame C-terminal Flag (DYK) tag.

Cell culture and transfection. We cultured human embryonic kidney (HEK293T; RRID:CVCL_0063) cells in high-glucose DMEM (Sigma-Aldrich) fortified with 10% foetal bovine serum (Sigma-Aldrich) and 1% penicillin/streptomycin (Sigma-Aldrich) at 37 °C in a humidified atmosphere containing 5% CO_2. The day after seeding 500,000 cells, we performed transfection in 6-well tissue culture plates using 10 µl Lipofectamine 2000 (Invitrogen) and 4 µg expression plasmid in a total of 2 ml. After 18 h incubation, we aspirated the medium transfection mix and washed the cells once with warm phosphate-buffered saline (pH 7.4; Sigma-Aldrich). We then incubated the cells with 2 ml Opti-MEM I reduced serum medium (Gibco) and harvested conditioned media and cells 72 h after transfection.

Western blotting. We precipitated 360 µl conditioned medium with 40 µl 10% trichloroacetic acid, dissolved the resulting protein pellet in 15 µl 1x Laemmli buffer-DTT solution (pH 6.8; 50 mM Tris, 10% glycerol, 2% SDS, 12.5 mM EDTA, 0.02% Bromphenol Blue, 1% β-ME, 100 mM

DTT), and denatured it for 5 min at 90 °C. Simultaneously, we prepared total lysate from harvested cells. We resuspended cell pellets in 50 µl RIPA buffer solution (pH 7.4; 10 mM Tris, 150 mM NaCl, 5 mM EDTA, 1% Triton X 100, and 10 mM phenylmethylsulphonyl fluoride). After two exposures to sonification (5 beats, cycle 0.5, amplitude 40; UP200S, Hielscher Ultrasonics) and a 5 min centrifugation at 1,000 g and 4 °C, we transferred the supernatant containing the total lysate into a new reaction tube. We mixed 20 µg total lysate with ddH_2O and Laemmli buffer. We loaded the precipitated medium and the lysate on a 10% sodium dodecyl sulphate-polyacrylamide gel for electrophoresis (SDS-PAGE). We then transferred the proteins onto a nitrocellulose membrane (Sigma-Aldrich) and detected CTRL with a monoclonal anti-Flag antibody (mouse, Sigma-Aldrich, cat# F1804). We first blocked the membrane in a 5% BSA solution and added the primary antibody at a dilution of 1:1,000 overnight at 4 °C. For detection of β-actin we used a polyclonal antibody (goat, abcam, cat# Ab8229) at a dilution of 1:2,500. After incubation with secondary antibodies (donkey anti-mouse, LI-COR Biosciences, cat# 926-68072; donkey anti-goat, LI-COR Biosciences, cat# 926-32214) for 2 h, we detected the proteins using the Azure Sapphire Biomolecular Imager (Azure Biosystems).

Enzyme activity measurements. We performed the assay in a 48-well plate in technical triplicates. We activated 48.75 µl conditioned medium with 1.25 µl porcine trypsin (Sigma-Aldrich; stock: 0.1 mg/ml) in a total volume of 50 µl for 2 h at 37 °C. We then measured CTRL activity 8 min after adding 150 µl of 0.2 mM N-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (AAPF; Sigma-Aldrich) as substrate in the VarioScan Flash Multimode Reader at 405 nm (Thermo Fisher Scientific).

Endoplasmic reticulum (ER) stress analysis. We conducted ER stress analyses with cDNA from HEK293T cells after overexpression of CTRL wild-type or mutants. As positive control served 0.5 µM thapsigargin (TG; stock: 0.1 mM) 6 h before RNA isolation. cDNA of untransfected and of pcDNA3.1+/C-(K)-DYK transfected cells served as negative controls. We harvested transfected cells for RNA isolation using the RNeasy Mini Kit (Qiagen) and transcribed 1 µg RNA into cDNA with the QuantiTect Reverse Transcription Kit (Qiagen). We

conducted TaqMan assays on *BiP* mRNA expression by real-time qPCR on the LightCycler 480 (Roche Diagnostics) in technical triplicates on four biological replicates. We used as primers 5-AAATTTGAAGAGCTCAACATGGATCT-3 and 5-GGTTTATGCCACGGGATGGT-3 and 6FAM-CCGGTCTACTATGAAGCCCGTCCAGAAAGTG-BBQ as probe. The housekeeper genes β-actin (5-AGCCTCGCCTTTGCCGA-3, 5-CTGGTGCCTGGGGCG-3, 6FAM-CCGCCGCCCGTCCACACCCGCC-BBQ) and GAPDH (5-GAAGGTGAAGGTCGGAGTC-3, 5- GAAGATGGTGATGGGATTTC-3, 6FAM-CAAGCTTCCCGTTCTCAGCC-BBQ) served as internal controls.

Statistical association analysis. To examine the association of potential disease-causing variants in the canonical *CTRL* transcript (ENST00000574481) between our patients and the general population, we compared the number of individuals carrying a REVEL (Rare Exome Variant Ensemble Learner) [21] positive (score ≥0.5) or nonsense mutation to the gnomAD 2.1.1 Non-Finnish European exome dataset (https://gnomad.broadinstitute.org) by applying a one-tailed Fisher's exact test [24]. Only variants with a maximum allele frequency across gnomAD populations of 0.5% (excluding Ashkenazi Jews (asj), European-Finnish (fin) and "Other" (oth)) not filtered in either the case or control cohort, at sites with at least 10x coverage in at least 90% of cases and gnomAD controls, were considered.

In addition, we analysed the genotype results of the patient and control cohort from Tübingen using the Fisher's exact test with a 2x3 contingency table for genotypes and a 2x2 contingency table for heterozygous carriers using the Interactive Statistical Calculation Pages (https://statpages.info). A two-tailed *p*-value less than 0.05 was considered significant. We tested the distributions of the genotypes for the Hardy-Weinberg equilibrium.

Results

Genetic analysis. We analysed ES data from 1,005 CP patients for rare *CTRL* variants (allelic frequency <0.1%) and compared data for predicted deleterious variants (REVEL score ≥0.5 or nonsense annotation) with gnomAD data. We identified three missense variants in four patients with a score REVEL ≥0.5 in our patient cohort (p.C220G twice) (0.4%), all variants were heterozygous. In comparison, 89 REVEL-predicted deleterious or nonsense variants were detected in the gnomAD dataset (*P*=0.079, OR=2.55) (**Table 2**). We also compared the patient cohort to a control cohort consisting of 1,594 unrelated Caucasians and found two heterozygotes with a predicted deleterious variant, p.G61S and p.G215R (0.13%) (*P*=0.21, OR=3.18) (**Table 3**). Although the results in both comparisons were statistically insignificant, we decided to further functionally examine the variants found in cases and controls, because the odds ratio was slightly elevated in both comparisons, and all prediction tools are error-prone.

Western blot reveals abolished secretion of some CTRL variants. Since secretion levels were too low for detection by Coomassie staining, we conducted Western blot analyses to investigate whether *CTRL* variants alter secretion. We overexpressed wild-type and mutated CTRL in HEK293T cells and collected supernatant and total lysate (**Figure 1**). Since CTRL is a secreted enzyme, we expected that the protein would not only be in the lysate but would be found mainly in the supernatant.

We detected all variants and the CTRL wild-type in the lysate, indicating successful intracellular protein synthesis. Examination of the supernatant gave different results, reflecting the secretory properties of the enzyme. Several changes resulted in secretory dysfunction with non-detectable bands in the Western blots: p.C201Y and p.C220G (found in patients), and p.G215R (found in controls). Interestingly, all variants with abolished secretion are at the C-terminal end.

Some CTRL variants show abrogated or reduced activity. Since CTRL cleaves polypeptides after tyrosine and phenylalanine residues, we used the synthetic substrate

N-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (AAPF) to test the enzymatic activity of the mutants in the supernatant (**Figure 2**). Six variants showed unaltered activity compared to the wild-type: p.L2S (88.2%), p.S15Y (112.8%), p.R106W (90.2%), p.S112N (91.0%), p.T150I (96.1%), and p.H173R (117.1%). Three variants showed an almost complete loss of activity: p.C201Y (2.2%), p.S208F (1.1%), and p.G215R (0.7%). Notably, p.S208F showed lost activity despite normal secretion. We also identified four variants with a reduced activity between 20-70%: p.G20S (65.3%), p.G56S (51.7%), p.G61S (35.0%), and p.C220G (24.8%).

ER stress is not a consequence of CTRL variants. Since some CTRL variants showed no secretion, we hypothesised that these might result in intracellular accumulation of misfolded proteins with subsequent endoplasmic reticulum (ER) stress. To investigate this possible effect, we analysed ER chaperone protein immunoglobulin-binding protein (*BiP*) mRNA expression [25]. As a positive control, we treated HEK293T cells overexpressing CTRL wild-type with the known ER stress inducer thapsigargin (TG) 6 h before RNA isolation. qPCR data showed that none of the variants increased the expression of *BiP* mRNA compared to the wild-type (**Table 3**). In contrast, CTRL wild-type transfected cells treated with TG showed 3.9-fold increased expression, indicating successful induction of ER stress.

Discussion

Chronic pancreatitis has a strong genetic background, particularly in early-onset CP. Several genes have been associated, but the underlying cause remains elusive in many early-onset cases. Chymotrypsins are the second most abundant proteases in pancreatic fluid. With a capability to cleave trypsinogen, they might participate in the protective mechanism of intrapancreatic trypsinogen degradation. CTRL degrades human anionic trypsinogen, however, much slower than CTRB2 [26]. In this context, and with the additional observation that *CTRC* and *CTRB1/CTRB2* have already been associated with CP [7, 10], we examined the fourth chymotrypsin, *CTRL*, as a potential risk gene.

We found several rare variants in the patient cohort that were predicted as deleterious by REVEL (score ≥0.5). Although the differences to gnomAD and further controls were insignificant (P=0.079 and P=0.21, respectively), both comparisons showed an increased odds ratio of 2.55 and 3.18, respectively. Since the validity of prediction tools is limited, we decided to examine the variants found in the patient and control collective functionally to compare the two cohorts based on these results. We identified several *CTRL* alterations that impair secretion or proteolytic function.

Three variants impaired the secretion of CTRL. The enzyme possesses five disulphide bonds that are crucial to its tertiary structure. The variants p.C201Y and p.C220G disrupt the C187-C201 and C155-C220 disulphide bonds, respectively, most likely leading to a severe conformational change with a complete loss-of-function and intracellular retention [27]. Accordingly, both variants and additionally p.G215R were not secreted. Nevertheless, a clear visible band in the cell lysate on the Western blot proved successful synthesis. Therefore, intracellular retention due to misfolding in the ER lumen might cause impaired secretion. We hypothesised that *CTRL* variants might lead to ER stress. However, we could not detect ER stress in any CTRL variant by measuring *BiP* mRNA expression levels. Due to its low abundance, intracellular retention of CTRL might not have a profound impact on provoking ER stress. Similar results have been described for misfolded SPINK1 variants in which the

mutations p.D50E, p.Y54H, and p.R67C led to intracellular retention [28], without eliciting ER stress [29, 30].

Seven CTRL variants caused reduced or lost activity. Interestingly, changes that diminish activity are predominantly located at the C-terminal end of CTRL. We detected an activity loss (<10%) in variants p.C201Y and p.S208F, and p.G215R. The variants p.G56S and p.G61S are close to the C60-C76 disulphide bond. Our functional analyses revealed decreased activity (51.7% and 35.0%) in these mutants, suggesting an impairment of the enzyme's conformation for proper catalytic activity. The same applied to p.S208F, which may interfere with the disulphide bond C210-C239. Considering the conventional chymotrypsin numbering, which allows an easy comparison of serine proteases in their primary structural features, disulphide bonds and tertiary structures, it reveals that p.S208F (p.S189 in conventional chymotrypsin numbering) lays at the bottom of the S1 primary specificity pocket which facilitates binding of the substrate and determines its specificity [31].

CTRL contains active site residues typical of the serine protease family (VTAA**H**C and GD**S**GG) [13]. The catalytic triad consists of a histidine residue p.H76 (basic; p.H57 in chymotrypsin numbering), an aspartic acid residue p.D121 (acidic; p.D102 in chymotrypsin numbering) and a serine residue p.S214 (nucleophilic; p.S195 in chymotrypsin numbering), representing the nucleophile-base-acid pattern conserved for serine peptidases in general [32]. The sequence surrounding the active site residues may be required to form the catalytic triad successfully. p.G215R (p.G196 in chymotrypsin numbering) is relatively close to the serine component (p.S214). An amino acid change could affect the triad formation and substrate binding, impairing hydrolysis. Similar effects have also been described for genetic alterations affecting the corresponding glycines in CELA3B (p.G218) and CTRC (p.G217) (both p.G196 in the chymotrypsin numbering). These variants showed significantly reduced catalytic activity (p.G218A in CELA3B and p.G217S in CTRC) or secretion (p.G217R in CTRC) [12, 33]. At the N-terminal end of CTRL, we observed slightly reduced proteolytic activity in p.G20S, which is close to the signal peptide cleavage site.

A recent study investigated the role of CTRL in caerulein-induced acute pancreatitis in *Ctrl*$^{-/-}$ mice generated from C57BL/6N mice using CRISPR-Cas9 genome engineering. Even though no Ctrl was detectable in the pancreas of homozygous *Ctrl*$^{-/-}$ mice, the overall chymotrypsinogen level was reduced by only 10% [16]. This finding indicates that CTRL is a rather low-abundance isoform of chymotrypsin, consistent with public domain data from GTEx or Human Protein Atlas. Pancreatic histology and morphology were unchanged in *Ctrl*$^{-/-}$ mice compared to controls, and they did not differ in phenotype or behaviour [16]. Caerulein overstimulation in *Ctrl*$^{-/-}$ and control mice did not result in significant metabolic, morphological, or proteolytic differences. However, the authors studied the effects only in an acute pancreatitis model. Phenotype and experimental outcome may differ whether an acute or chronic caerulein-induced pancreatitis model is applied. For example, *Trp6*$^{mut/mut}$ mice, in which aspartic acid was replaced by alanine in codon 541 (p.D541A), showed no effect in a caerulein-induced acute pancreatitis model (unpublished data). However, these mice develop severe CP after repeated caerulein administration with increased levels of serum amylase and lipase after 24 hours, histological alterations, and fibrosis [3]. Although the differences between the patients and the control cohorts were statistically insignificant, we observed increased odds ratios (2.55 and 3.18, respectively). Therefore, we cannot completely rule out that our study was still underpowered despite the large number of patients.

In summary, some *CTRL* variants substantially affect protease activity and secretion. However, these variants were found with similar frequency in patients as in controls and were therefore not associated with CP. Loss of CTRL might have only minor effects, as it accounts for only about 10% of total chymotrypsin content [16]. Therefore, we assume the presence of backup mechanism by other chymotrypsin isoforms. Due to the lack of understanding of the actual function of CTRL, we did not identify possible compensating mechanisms for a loss-of-function.

References

[1] Witt H, Apte MV, Keim V, Wilson JS. Chronic pancreatitis: challenges and advances in pathogenesis, genetics, diagnosis, and therapy. Gastroenterology 2007; 132: 1557-1573.

[2] Sharer N, Schwarz M, Malone G, Howarth A, Painter J, Super M, et al. Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. N Engl J Med 1998; 339: 645-652.

[3] Masamune A, Kotani H, Sörgel FL, Chen JM, Hamada S, Sakaguchi R, et al. Variants that affect function of calcium channel TRPV6 are associated with early-onset chronic pancreatitis. Gastroenterology 2020; 158: 1626-1641.

[4] Whitcomb DC, LaRusch J, Krasinskas AM, Klei L, Smith JP, Brand RE, et al. Common genetic variants in the *CLDN2* and *PRSS1-PRSS2* loci alter risk for alcohol-related and sporadic pancreatitis. Nat Genet 2012; 44: 1349-1354.

[5] Whitcomb DC, Gorry MC, Preston RA, Furey W, Sossenheimer MJ, Ulrich CD, et al. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. Nat Genet. 1996; 14: 141-145.

[6] Witt H, Luck W, Hennies HC, Claßen M, Kage A, Laß U, et al. Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis. Nat Genet. 2000; 25: 213-216.

[7] Rosendahl J, Witt H, Szmola R, Bhatia E, Ózsvári B, Landt O, et al. Chymotrypsin C (CTRC) alterations that diminish activity or secretion are associated with chronic pancreatitis. Nat Genet 2008;40(1):78–82.

[8] Witt H, Beer S, Rosendahl J, Chen JM, Chandak GR, Masamune A, et al. Variants in CPA1 are strongly associated with early onset chronic pancreatitis. Nat Genet 2013; 45: 1216-1220.

[9] Fjeld K, Weiss FU, Lasher D, Rosendahl J, Chen JM, Johansson BB, et al. A recombined allele of the lipase gene CEL and its pseudogene CELP confers susceptibility to chronic pancreatitis. Nat Genet 2015; 47: 518-522.

[10] Rosendahl J, Kirsten H, Hegyi E, Kovacs P, Weiss FU, Laumen H, et al. Genome-wide association study identifies inversion in the CTRB1-CTRB2 locus to modify risk for alcoholic and non-alcoholic chronic pancreatitis. Gut 2018; 67:1855-1863.

[11] Lasher D, Szabó A, Masamune A, Chen JM, Xiao X, Whitcomb DC, et al. Protease-Sensitive Pancreatic Lipase Variants Are Associated With Early Onset Chronic Pancreatitis. Am J Gastroenterol 2019; 114: 974-983.

[12] Tóth A, Demcsák A, Zankl F, Oracz G, Unger LS, Bugert P, et al. Loss-of-function variant in chymotrypsin like elastase 3B (CELA3B) is associated with non-alcoholic chronic pancreatitis. Pancreatology 2022; 22: 713-718.

[13] Larsen F, Solheim J, Kristensen T, Kolstø A-B, Prydz H. A tight cluster of five unrelated human genes on chromosome 16q22.1. Hum Mol Genet 1993; 2: 1589-1595.

[14] Reseland JE, Larsen F, Solheim J, Eriksen JA, Hanssen LE, Prydz H. A novel human chymotrypsin-like digestive enzyme. J Biol Chem 1997; 272: 8099-8104.

[15] Szabó A, Sahin-Tóth M. Determinants of chymotrypsin C cleavage specificity in the calcium-binding loop of human cationic trypsinogen. FEBS J 2012; 279: 4283-4292.

[16] Mosztbacher D, Jancsó Z, Sahin-Tóth M. Loss of chymotrypsin-like protease (CTRL) alters intrapancreatic protease activation but not pancreatitis severity in mice. Sci Rep 2020; 10: 11731.

[17] Mose LE, Perou CM, Parker JS. Improved indel detection in DNA and RNA via realignment with ABRA2. Bioinformatics 2019; 35: 2966-2973.

[18] Bansal V, Libiger O. Fast individual ancestry inference from DNA sequence data leveraging allele frequencies for multiple populations. BMC Bioinformatics 2015;16: 4.

[19] Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010; 26: 2867-2873.

[20] Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015; 4: 7.

[21] Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet 2016; 99: 877-885.

[22] Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020; 581: 434-443.

[23] McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biol 2016; 17: 122.

[24] Guo MH, Plummer L, Chan YM, Hirschhorn JN, Lippincott MF. Burden testing of rare variants identified through exome sequencing via publicly available control data. Am J Hum Genet 2018; 103: 522-534.

[25] Oslowski CM, Urano F. Measuring ER stress and the unfolded protein response using mammalian tissue culture system. Meth Enzymol 2011; 490: 71-92.

[26] Németh BZ, Nagy ZA, Kiss B, Gellén G, Schlosser G, Demcsák A, et al. Substrate specificity of human chymotrypsin-like protease (CTRL) characterized by phage display-selected small-protein inhibitors. Pancreatology 2023; S1424-3903(23)01608-3. doi: 10.1016/j.pan.2023.08.004.

[27] Braakman I, Helenius J, Helenius A. Manipulating disulfide bond formation and protein folding in the endoplasmic reticulum. EMBO J 1992; 11: 1717-1722.

[28] Király O, Wartmann T, Sahin-Tóth M. Missense mutations in pancreatic secretory trypsin inhibitor (SPINK1) cause intracellular retention and degradation. Gut 2007; 56: 1433-1438.

[29] Boulling A, Le Maréchal C, Trouvé P, Raguénès O, Chen J-M, Férec C. Functional analysis of pancreatitis-associated missense mutations in the pancreatic secretory trypsin inhibitor (SPINK1) gene. Eur J Hum Genet 2007; 15: 936-942.

[30] Boulling A, Keiles S, Masson E, Chen J-M, Férec C. Functional analysis of eight missense mutations in the SPINK1 gene. Pancreas 2012; 41: 329-330.

[31] Perona JJ, Craik CS. Structural basis of substrate specificity in the serine proteases. Protein Sci 1995; 4: 337-360.

[32] Polgár L. The catalytic triad of serine peptidases. Cell Mol Life Sci 2005; 62: 2161-2172.

[33] Beer S, Zhou J, Szabó A, Keiles S, Chandak GR, Witt H, et al. Comprehensive functional analysis of chymotrypsin C (CTRC) variants reveals distinct loss-of-function mechanisms associated with pancreatitis risk. Gut 2013; 62: 1616-1624.

Table 1. Oligonucleotides for PCR and sequencing of CTRL.

Exon	Direction	Sequence 5'->3'
PCR primers		
Exon 1-3	forward	CAAAGTTGACCGCAGTGAGC
	reverse	CTCCGTGTCTGCAGCCCAGG
Exon 4-7	forward	CTTCCTGCCTCCTCCCCTGG
	reverse	GTCTCTGGGGACCTGAAGCC
Sequencing primers		
Exon 1	forward	CTGTGATCCCAGCCACCAGG
Exon 2 and 3	forward	GTGCTGTACCCTCTCGACAG
Exon 4 and 5	forward	TGCCTCCTCCCCTGGCTGTC
Exon 6 and 7	forward	TCAGGCCAAAGCTCAGGGTG

Table 2. Nonsense and predicted deleterious *CTRL* variants

Gene	case count			control count (gnomAD)		
	het	hom	total	het	hom	total
CTRL	4	0	4	89	0	89
	p-value	OR	Position (hg19)	Nucleotide change		AA change
	0.079	2.55	16:67963974	A>C		p.C220G
			16:67964098	G>A		p.S208F
			16:67964119	C>T		p.C201Y

REVEL = Rare Exome Variant Ensemble Learner; het = heterozygous; hom = homozygous, OR = odds ratio; AA = amino acid; case count = 1,005; control count = 56,885. Note, the variant p.C220G was found twice in cases.

Revel positive (score ≥0.5) and nonsense variants were considered if maximum allele frequency among gnomAD populations was below 0.1% and if coverage in cases and gnomAD individuals was ≥10 in at least 90% of samples. For coverage calculations we used subsets of 10% of the two cohorts.

Table 3. *CTRL* variants found in CP patients and controls

Exon	Position (hg19)	Nucleotide change	Amino acid change	rs number	REVEL	CP (n=1,005) (%)	Controls (n=1,594) (%)	Secretion	% of wt activity (S)	% of wt *BiP* expression (SD)	*p*-value
1	16:67965751	c.5T>C	p.L2S	rs144681992	0.086	1 (0.1)	1 (0.06)	Yes	88.2	86.6 (±25.1)	1.0
1	16:67965712	c.44C>A	p.S15Y	rs144605661	0.290	3 (0.3)	1 (0.06)	Yes	112.8	115.0 (±47.0)	0.31
2	16:67965099	c.58G>A	p.G20S	rs140026167	0.266	2 (0.2)	3 (0.19)	Yes	65.3	160.5 (±72.1)	1.0
2	16:67964893	c.166G>A	p.G56S	rs761948651	0.273	1 (0.1)	-	Yes	51.7	101.0 (±13.5)	0.34
3	16:67964878	c.181G>A	p.G61S	rs141577010	**0.673**	-	1 (0.06)	Yes	35.0	126.0 (±80.6)	1.0
5	16:67964631	c.316C>T	p.R106W	rs1188402196	0.220	1 (0.1)	-	Yes	90.2	98.8 (±32.6)	0.34
5	16:67964475	c.335G>A	p.S112N	rs137951739	0.028	-	1 (0.06)	Yes	91.0	102.0 (±11.1)	1.0
5	16: 67964361	c.449C>T	p.T150I	rs11552953	0.080	1 (0.1)	5 (0.3)	Yes	96.1	88.3 (±19.4)	1.0
6	16: 67964203	c.518A>G	p.H173R (het)	rs1134760	0.020	265 (26.4)	474 (29.7)	Yes	117.1	81.8 (±19.4)	0.12
			p.H173R (homo)			34 (3.4)	62 (3.9)				
6	16:67964119	c.602G>A	p.C201Y	rs201073485	**0.923**	1 (0.1)	-	No	2.2	102.0 (±16.8)	0.34
6	16:67964098	c.623C>T	p.S208F	rs762981045	**0.506**	1 (0.1)	-	Yes	1.1	84.5 (±8.7)	0.34
7	16:67963989	c.643G>A	p.G215R	rs780563969	**0.714**	-	1 (0.06)	No	0.7	119.5 (±37.5)	1.0
7	16:67963974	c.658T>G	p.C220G	rs147657411	**0.599**	2 (0.2)	-	Reduced	24.8	87.3 (±19.7)	0.11
Predicted deleterious variants (REVEL score ≥0.5)						4 (0.4)	2 (0.13)		0.7-35.0		0.21

REVEL = Rare Exome Variant Ensemble Learner (http://www.mulinlab.org/vportal/index.html); CP = chronic pancreatitis; S = supernatant; SD = standard deviation; wt = wild-type. P-values were calculated by two-tailed Fisher's exact test with a 2x3 contingency table comparing genotype frequencies and with a 2x2 contingency table comparing carrier frequencies.

Figure Legends

Figure 1. Western blot of HEK293T supernatant and total lysate after overexpression of CTRL

We prepared supernatant and total lysate from HEK293T cells and used a monoclonal anti-Flag antibody to detect overexpressed CTRL (28 kDa). We used the housekeeper gene β-actin (42 kDa) as an internal loading control in the total lysate samples. We loaded 400 µl precipitated supernatant and 40 µg total lysate. In some replicates, p.C220G showed no secretion. The figures are representative of a total of four biological replicates.

CTRL = chymotrypsin-like protease; empty = pcDNA3.1+/C-(K)-DYK; wt = wild-type; (S) = supernatant; (L) = lysate

Figure 2. CTRL activity in the supernatant of transfected HEK293T cells

Supernatant was prepared from HEK293T cells after overexpression of CTRL wild-type (wt) or mutants. We activated 48.75 µl conditioned medium of overexpressed CTRL with 1.25 µl porcine trypsin (stock: 0.1 mg/ml) for 2 h and performed activity assays in technical triplicates on four biological replicates. The figure shows enzyme activity 8 min after the substrate AAPF was added to the activated CTRL. All measurements were set in relation to CTRL wild-type activity (= 100%; dark grey bar). Error bars depict the standard deviation. Dotted bars depict mutations found in controls.

AAPF = N-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide; wt = wild-type.

Figure 1

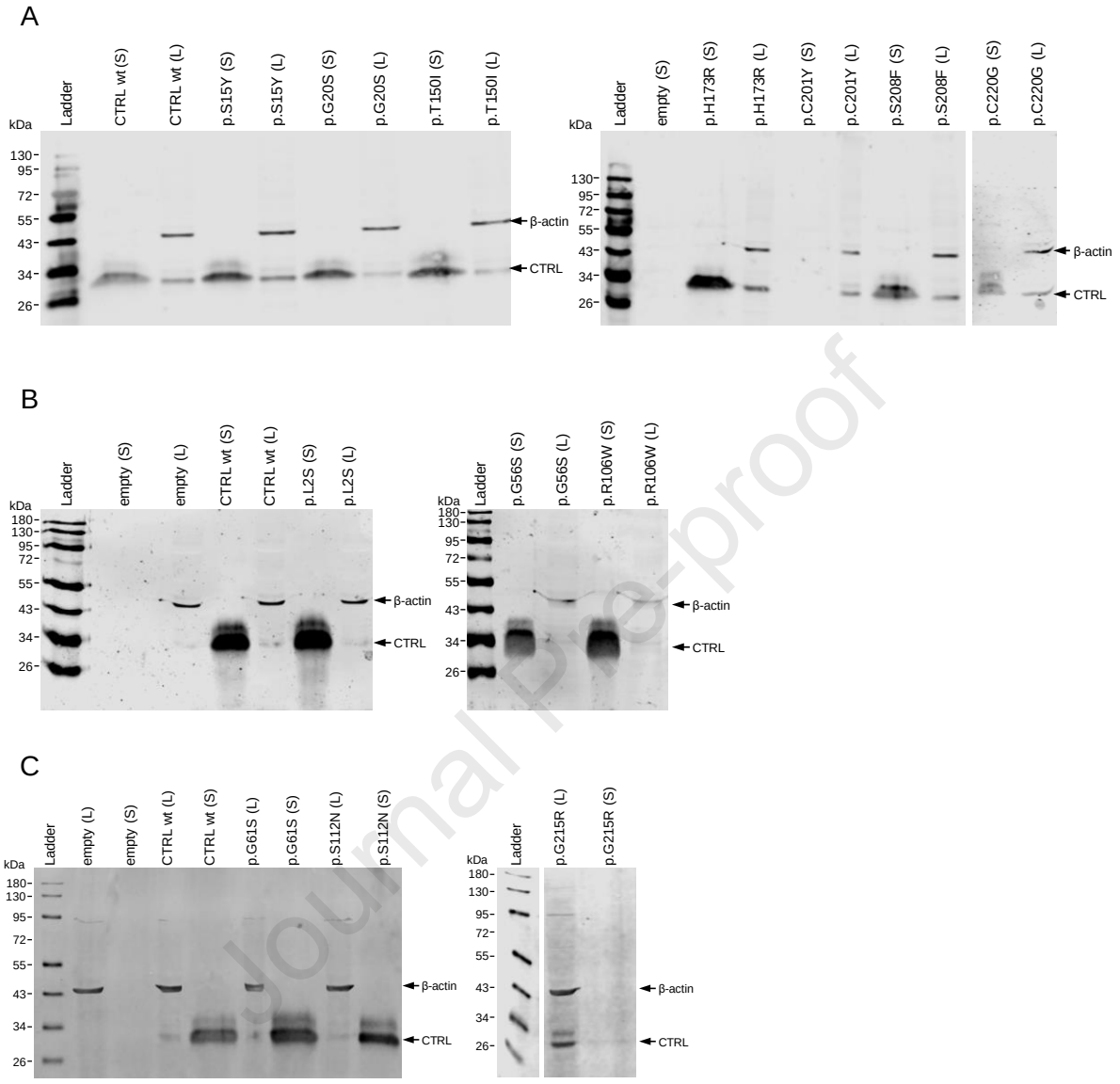

Figure 2