JBC RESEARCH ARTICLE

The position of single-base deletions in the VNTR sequence of the carboxyl ester lipase (*CEL*) gene determines proteotoxicity

Received for publication, December 22, 2020, and in revised form, April 5, 2021 Published, Papers in Press, April 14, 2021,
https://doi.org/10.1016/j.jbc.2021.100661

Anny Gravdal[1,2,3], Xunjun Xiao[4], Miriam Cnop[5,6], Khadija El Jellas[1,2], Stefan Johansson[2,3], Pål R. Njølstad[2,7], Mark E. Lowe[4], Bente B. Johansson[2,7], Anders Molven[1,2,8,*], and Karianne Fjeld[1,2,3]

From the [1]The Gade Laboratory for Pathology, Department of Clinical Medicine and [2]Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway; [3]Department of Medical Genetics, Haukeland University Hospital, Bergen, Norway; [4]Department of Pediatrics, Division of Gastroenterology, Hepatology and Nutrition, Washington University School of Medicine, St Louis, Missouri, USA; [5]ULB Center for Diabetes Research and [6]Division of Endocrinology, ULB Erasmus Hospital, Université Libre de Bruxelles, Brussels, Belgium; [7]Department of Pediatrics and Adolescent Medicine, Haukeland University Hospital, Bergen, Norway; and [8]Department of Pathology, Haukeland University Hospital, Bergen, Norway

Edited by Qi-Qun Tang

Variable number of tandem repeat (VNTR) sequences in the genome can have functional consequences that contribute to human disease. This is the case for the *CEL* gene, which is specifically expressed in pancreatic acinar cells and encodes the digestive enzyme carboxyl ester lipase. Rare single-base deletions (DELs) within the first (DEL1) or fourth (DEL4) VNTR segment of *CEL* cause maturity-onset diabetes of the young, type 8 (MODY8), an inherited disorder characterized by exocrine pancreatic dysfunction and diabetes. Studies on the DEL1 variant have suggested that MODY8 is initiated by CEL protein misfolding and aggregation. However, it is unclear how the position of single-base deletions within the *CEL* VNTR affects pathogenic properties of the protein. Here, we investigated four naturally occurring CEL variants, arising from single-base deletions in different VNTR segments (DEL1, DEL4, DEL9, and DEL13). When the four variants were expressed in human embryonic kidney 293 cells, only DEL1 and DEL4 led to significantly reduced secretion, increased intracellular aggregation, and increased endoplasmic reticulum stress compared with normal CEL protein. The level of O-glycosylation was affected in all DEL variants. Moreover, all variants had enzymatic activity comparable with that of normal CEL. We conclude that the longest aberrant protein tails, resulting from single-base deletions in the proximal VNTR segments, have highest pathogenic potential, explaining why DEL1 and DEL4 but not DEL9 and DEL13 have been observed in patients with MODY8. These findings further support the view that *CEL* mutations cause pancreatic disease through protein misfolding and proteotoxicity, leading to endoplasmic reticulum stress and activation of the unfolded protein response.

Variable number of tandem repeat (VNTR) is a term used to describe short DNA sequence motifs that are consecutively repeated several times in the genome (1). As these motifs are very polymorphic and inherited in a Mendelian pattern, they have had a tremendous impact on our ability to identify individuals based on biological samples (2). The VNTRs are most often located in noncoding regions of the genome, but some are found in promoter regions (3) or within coding DNA sequences (4). In such cases, a VNTR may influence the expression levels or functional properties of the corresponding protein. Moreover, as repeated DNA motifs are prone to undergo mutations by DNA slippage and unequal crossing-over, VNTRs affecting the protein product are intriguing candidates for human disease associations (1).

One such example is found in the *CEL* gene, encoding the digestive enzyme carboxyl ester lipase, which is mainly expressed in the acinar cells of the pancreas (5). *CEL* contains 11 exons, where the last exon has a VNTR that consists of nearly identical 33–base pair segments, each encoding 11 amino acids (5). The number of repeated segments within the VNTR may vary between three and 23. However, in all populations examined so far, the most common *CEL* allele contains 16 repeats, encoding a protein with a predicted molecular mass of ∼79 kDa (6–11) (Fig. 1). As the CEL protein is both N- and O-glycosylated, the observed mass is considerably larger (12, 13). In particular, threonine residues within the VNTR-encoded tail region undergo O-glycosylation (13). This modification may serve to prevent rapid degradation of the CEL protein but might also facilitate secretion or increase solubility (14). Intriguingly, the pattern of O-glycosylation in CEL reflects the ABO blood group of the individual (15).

A single-base deletion (DEL) in the first repeat of the *CEL* VNTR leads to maturity-onset diabetes of the young, type 8 (MODY8), a dominantly inherited syndrome of exocrine pancreatic dysfunction and diabetes (16). The corresponding pathogenic protein has been referred to as CEL-MODY or CEL-MUT in previous publications (17, 18) but will hereafter be denoted as CEL-DEL1 (abbreviated: DEL1) to emphasize that the deletion is positioned in the first repeat. Another

* For correspondence: Anders Molven, anders.molven@uib.no.

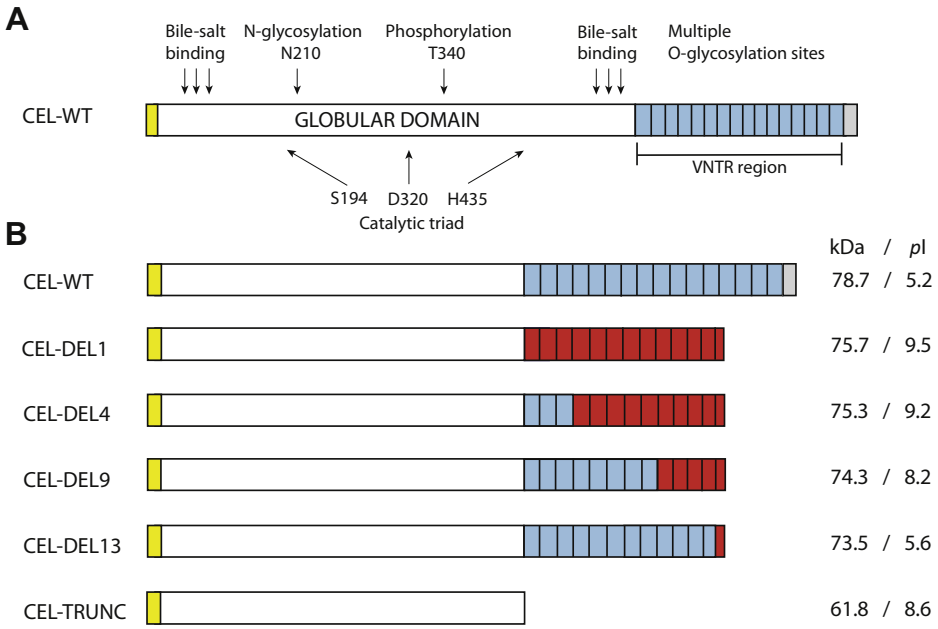

Figure 1. Schematic overview of the different CEL protein variants encoded by constructs used in the present study. *A*, the normal CEL protein (CEL-WT) with 16 repeated segments, each consisting of 11 amino acids and encoded by the VNTR. This is the most common CEL variant in the general population. The *yellow box* symbolizes the N-terminal signal peptide, whereas *light blue boxes* represent the normal repeats. The unique C-terminal sequence of CEL-WT is symbolized by a *gray box*. Also indicated are residues required for bile salt binding, N- and O-glycosylation, phosphorylation, and catalysis. *B*, overview of the different CEL protein variants investigated. Predicted molecular mass (kDa) and the isoelectric point (*pI*) of each variant are shown on the *right* side. All DEL constructs were based on a 16-repeat backbone. *Light blue boxes* represent normal repeat sequences, whereas *red boxes* indicate aberrant repeats resulting from the frameshifts introduced by single–base deletions in the *CEL* VNTR. CEL-TRUNC is an artificial variant lacking the VNTR region and was included as control construct. Elements of A and B are not drawn to scale. DEL, deletion; VNTR, variable number of tandem repeat.

single-base deletion in the fourth repeat of the *CEL* VNTR was found in a family with a similar phenotype as the original MODY8 family (16), and this protein will be denoted as DEL4. Owing to frameshifts and premature stop codons, DEL1 and DEL4 are both truncated CEL variants with aberrant protein C termini (Fig. 1). Very rare single-base deletions in the more distal repeats of the *CEL* VNTR have been observed when screening population-based materials (own unpublished data). It is unknown whether such variants also associate with the disease.

We have previously shown that DEL1 has a high propensity to form both intracellular and extracellular aggregates (18). We have also found that after secretion, there is cellular re-uptake followed by lysosomal degradation of the DEL1 protein, resulting in reduced viability of pancreatic cell lines (17, 19). Furthermore, we have reported that DEL1 causes endoplasmic reticulum (ER) stress, induction of the unfolded protein response, and subsequent apoptosis (20). Taken together, these results indicate that the CEL protein in MODY8 has acquired a negative gain-of-function effect and that the disease follows the protein misfolding–dependent pathway of pancreatitis (21).

The aim of the present study was to examine the pathogenicity of single-base deletions naturally occurring in the *CEL* VNTR. We tested how the position of a deletion within this region affects the functional properties of the encoded CEL protein variant. We therefore compared the DEL1 and DEL4 variants causing MODY8 with two deletions (DEL9 and DEL13) that we have observed in the general population. Our data suggest a correlation between the position of a *CEL* VNTR single-base deletion and its functional impact on the protein, where increasingly distal mutations have less pathogenic potential.

Results

Cellular properties of CEL proteins are influenced by the V5/His tag

In previous cellular studies of the pathogenic DEL1 variant, we have expressed the protein both with and without a V5/His epitope tag fused to the C terminus (15–20). Although epitope tags are very useful tools for protein purification and detection, the charged and hydrophilic residues of such tags may sometimes have a significant impact on the biochemical properties of the protein (22, 23). Before initiating the functional study of our set of CEL DEL variants, we therefore investigated how the V5/His tag might influence the CEL protein.

Human embryonic kidney 293 (HEK293) cells were transiently transfected with plasmids expressing DEL1 with and without tag, followed by either cellular fractionation and Western blotting or immunocytochemistry. As controls, we included a construct encoding the normal CEL protein with 16 repeats (WT) and a construct with a stop codon introduced at the beginning of VNTR segment 1 (TRUNC, Fig. 1). The latter construct encodes a truncated CEL protein that has no repeated segments and is unlikely to undergo O-glycosylation. It is an artificial variant not being reported in human materials.

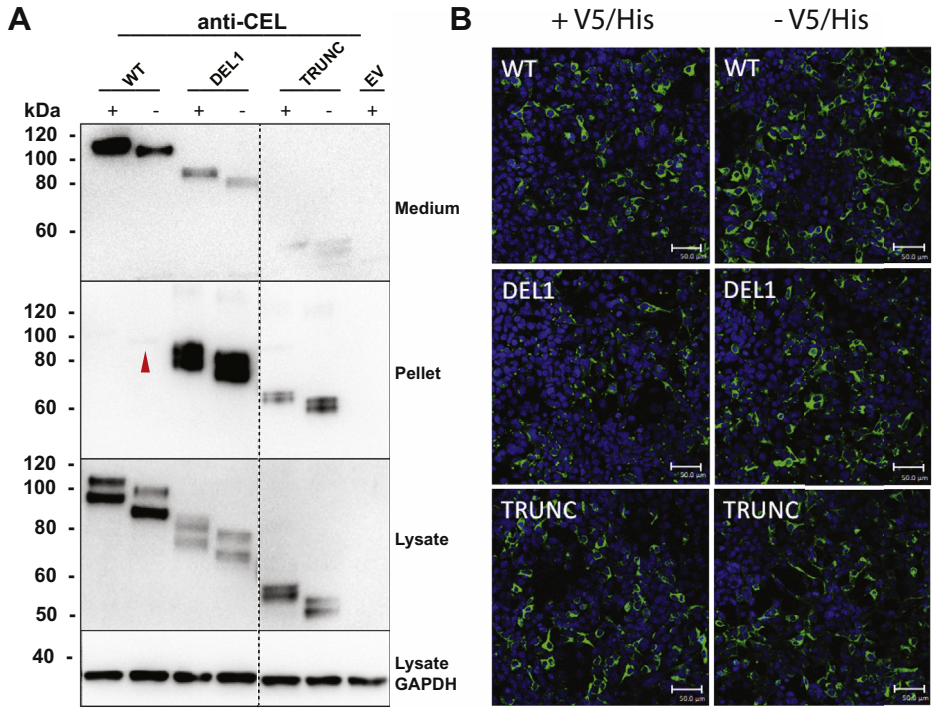

Figure 2. Secretion and intracellular distribution of tagged *versus* untagged CEL variants in HEK293 cells. HEK293 cells were transiently transfected with plasmids encoding CEL-WT, CEL-DEL1, or CEL-TRUNC with or without the V5/His tag (+/−). The expressed CEL proteins were examined by Western blotting (*A*) and immunocytochemistry (*B*). Representative images from n = 3 independent experiments are shown. *A*, secretion was evaluated as CEL levels detected in the conditioned medium. Intracellular distribution was assessed by analyzing the soluble (lysate) and insoluble (pellet) cellular fractions. Cells transfected with the empty vector (EV) were used as the negative control, whereas GAPDH expression was monitored for control of loading. The *red arrow* points at a very weak band for untagged CEL-WT protein in the pellet fraction. The *stippled line* indicates an unrelated lane removed from the images. *B*, transfected cells were subjected to immunostaining and confocal microscopy. CEL protein is stained *green* and cell nuclei are stained *blue*. Scale bars represent 50 μm. DEL, deletion; HEK293, human embryonic kidney 293.

We found the WT and DEL1 proteins expressed without the V5/His tag to be less secreted than the tagged variants (Fig. 2*A*). In addition, the untagged variants showed slightly stronger bands in the detergent-insoluble pellet fraction, indicating that they are more prone to aggregate. Consistent with these findings, confocal imaging demonstrated a stronger intracellular signal for the untagged WT and DEL1 proteins than for their tagged counterparts (Fig. 2*B*). The TRUNC protein did not display clear differences with or without the tag in our experiments.

Based on these results, we concluded that the V5/His tag may increase the solubility of CEL protein variants, possibly *via* interactions involving the VNTR-encoded region. In previous studies with tagged variants (17–19), we may therefore have underestimated the detrimental cellular effects of DEL1. Consequently, we decided to use only untagged protein variants in the present study.

VNTR composition affects CEL protein secretion and intracellular distribution

We have previously shown that the disease-causing DEL1 variant exhibits intracellular retention and reduced secretion (20). In the first article describing MODY8, we also identified a family with a DEL4 mutation of *CEL* (16), but this variant has not been studied functionally. When screening 1287 subjects of the general Norwegian population, we observed three subjects having a DEL9 mutation, that is, a single–base deletion in segment number 9 of the *CEL* VNTR. This corresponded to a carrier frequency of 0.0023. Similarly, we identified one subject with a single–base deletion in VNTR segment 13 (DEL13, carrier frequency 0.0008).

We therefore decided to test how the position of the deletion within the VNTR might affect pathogenicity by investigating the identified set of *CEL* variants with mutations in different repeat segments (Table 1; Fig. 1). Constructs encoding the CEL-WT and CEL-TRUNC proteins were included as reference variants in all experiments. Initially, we studied the effect of the four protein variants on secretion and intracellular distribution. HEK293 cells were transiently transfected with the six constructs, and the conditioned medium, soluble lysate, and insoluble pellet were analyzed by Western blotting.

When compared with CEL-WT, secretion and intracellular distribution varied substantially between the variants

Table 1
Analyzed single-base deletions in the VNTR region of the *CEL* gene

Variant designation	Nucleotide change[a]	Amino acid change
DEL1	c.1686delT	p.Val563Cysfs*132
DEL4	c.1785delC	p. Val596Cysfs*99
DEL9	c.1950delC	p.Val651Cysfs*44
DEL13	c.2082delC	p.Val695*

DEL, deletion; VNTR, variable number of tandem repeat.
[a] Position according to NCBI reference mRNA sequence NM_001807.4 (www.ncbi.nlm.nih.gov/nuccore/NM_001807.4).

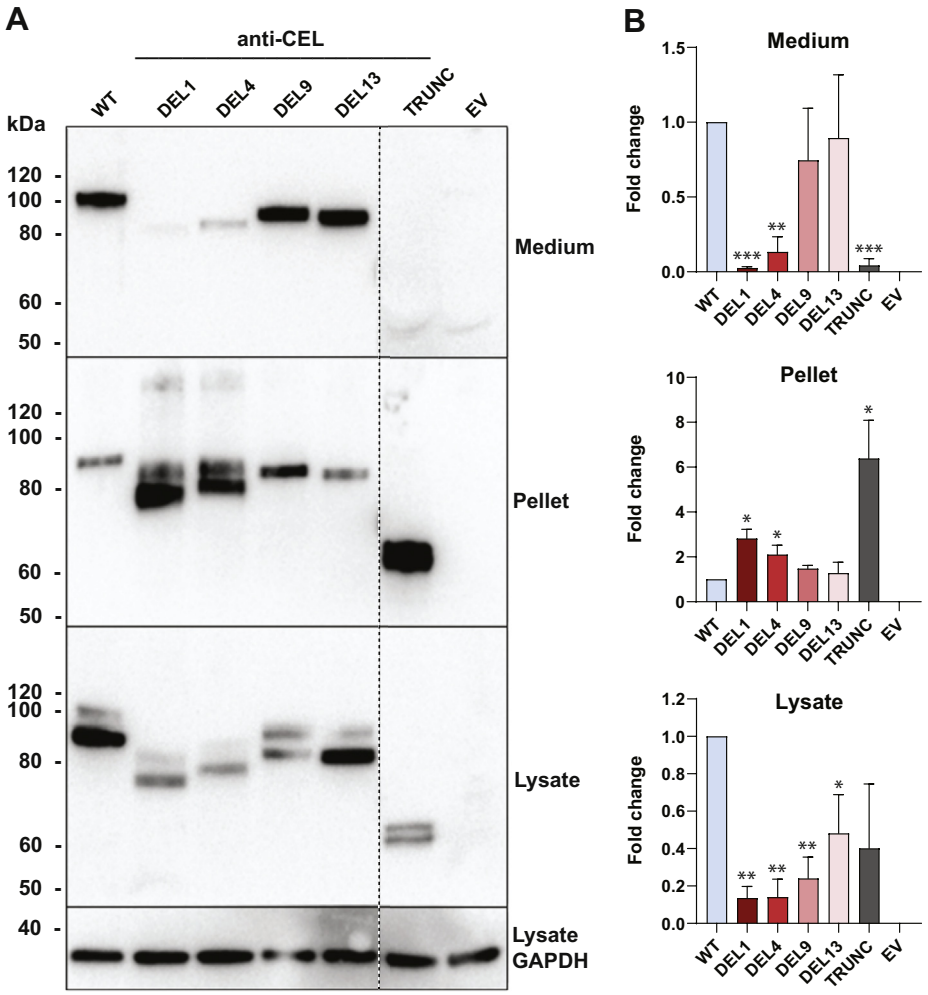

Figure 3. Secretion and cellular distribution of different CEL protein variants. *A*, HEK293 cells were transiently transfected with plasmids encoding the CEL variants shown in Figure 1 and subjected to cellular fractionation and Western blotting. Cells transfected with the empty vector (EV) were used as negative control, whereas GAPDH expression was monitored for control of loading. Secretion was assessed as CEL levels detected in the conditioned medium. Intracellular distribution was evaluated by analyzing the soluble (lysate) and insoluble (pellet) fractions after lysis of the cells. The *stippled line* indicates an unrelated lane removed from the images. Representative images from n = 3 independent experiments are shown. *B*, quantification of Western blot band intensities in the three experiments, adjusted to the GAPDH levels and normalized to the CEL-WT lane. Error bars are SD. Statistical significance is indicated as follows for band intensities different from that of CEL-WT: *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$. HEK293, human embryonic kidney 293.

(Fig. 3). We found significantly reduced secretion for DEL1 (38-fold), DEL4 (7.5-fold), and TRUNC (23-fold). In contrast, variants DEL9 and DEL13 displayed a secretion level similar to that of the WT protein. In the pellet fraction, there was a 2.8-fold increase for DEL1, 2.1-fold increase for DEL4, and 6.4-fold increase for TRUNC. Band intensities of DEL9 and DEL13 appeared slightly stronger than for the WT protein, but differences were not statistically significant. Moreover, significantly lower levels of all DEL variants (DEL1: 7.4-fold, DEL4: 7.1-fold, DEL9: 4.2-fold, DEL13: 2.1-fold) were detected in the soluble lysate fraction (Fig. 3). TRUNC exhibited a 2.5-fold reduction that did not reach statistical significance because of large experimental variation. The general decrease in abundance of all variants in the lysate fraction when compared with the CEL-WT level is likely to reflect their increased abundance in the pellet fraction. The intracellular levels of the variants should be considered as a total of the lysate and pellet fraction and appear as such to be relatively constant.

The extent of O-glycosylation varies between CEL DEL variants

Interestingly, the four DEL variants migrated with clearly different band sizes in the SDS-PAGE gels despite having very similar predicted molecular masses (compare Figs. 1 and 3). This suggested that the variants differ at the level of post-translational modification. Because CEL is heavily O-glycosylated (13, 15, 24), a likely explanation for the observed size differences could be variations in O-glycosylation. Prediction of theoretical O-glycosylation sites in the VNTR region revealed that CEL-WT has the highest number of potential sites (n = 36), whereas DEL1, DEL4, DEL9, and DEL13 contain 13, 18, 23, and 26 sites, respectively (Fig. S1). Thus, the proximal DEL variants DEL1 and DEL4 have fewer predicted O-glycosylation sites than the distal variants DEL9 and DEL13. This observation might at

least to some degree explain the migration differences observed in Figure 3.

To experimentally investigate the effect of O-glycosylation on the various DEL variants, an isogenic cell model with (HEK293 SimpleCells) and without (regular HEK293 cells) KO of the gene encoding the Cosmc chaperone was used (25, 26). As a result of this KO, the enzymatic activity that provides O-glycan elongation beyond the initial GalNAc-α1 residue on O-linked glycoproteins (T-synthase) is absent. We transfected SimpleCells and their nonaltered HEK293 counterparts with the different CEL variants and analyzed the protein mass by Western blotting (Fig. 4). We found that most variants showed a ∼20-kDa reduction in band migration when expressed in SimpleCells, reflecting the presence of truncated O-glycans in the C terminus of CEL. Notably, DEL1 was the only variant whose migration in the lysate fraction of the SimpleCells corresponded to its predicted theoretical mass (∼75 kDa). It was also the DEL variant with least efficient secretion in both cell lines. While SimpleCells secreted CEL proteins with two different sizes into the medium, HEK293 cells secreted the variants as only one prominent band, representing the mature and heavily glycosylated form of the protein (compare the right panels of Fig. 4). Proteins can undergo the initial O-glycosylation step

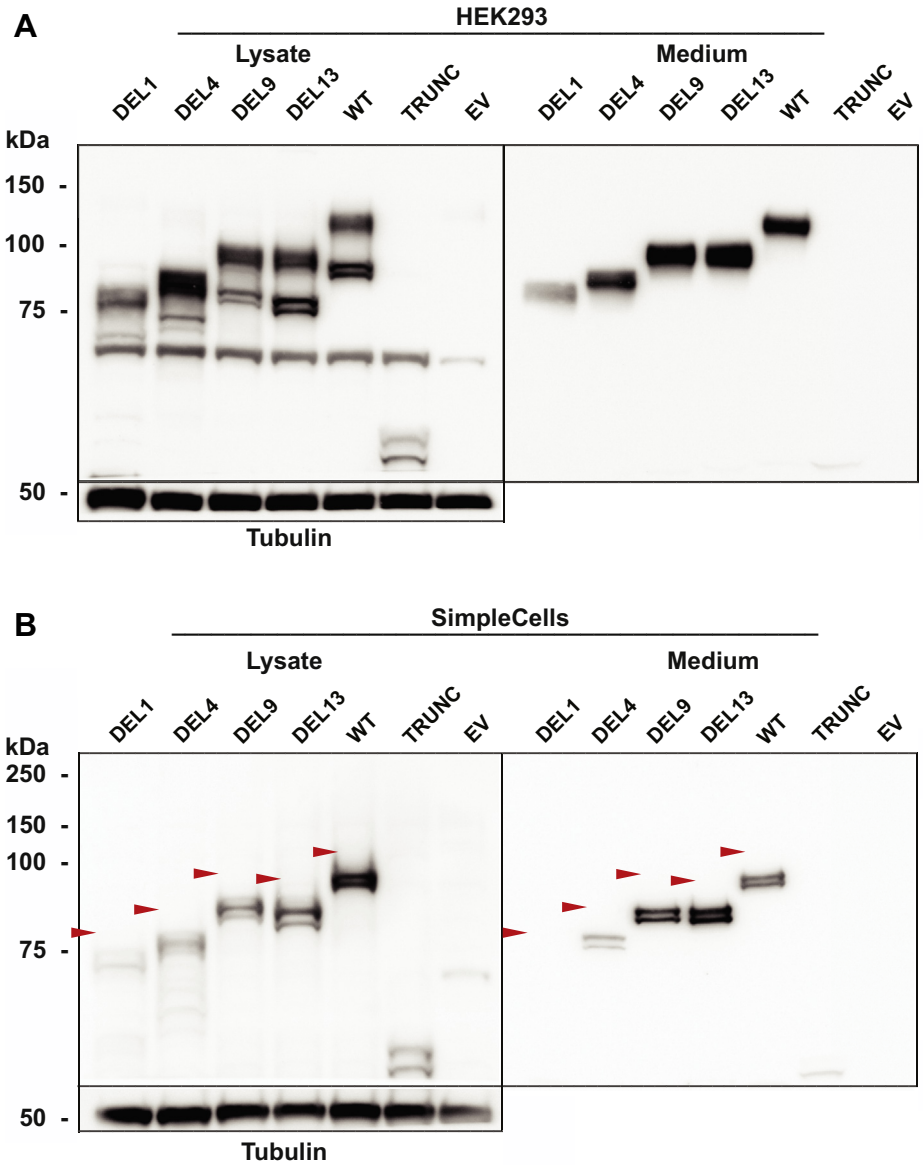

Figure 4. Expression of CEL variants in an O-glycosylation-deficient cellular model. HEK293 cells (*A*) and HEK293 cells with Cosmc KO (SimpleCells) (*B*) were transiently transfected with plasmids encoding the CEL variants shown in Figure 1 and analyzed by Western blotting. Cells transfected with the empty vector (EV) were used as negative control. Tubulin expression was monitored for control of loading. C-terminal O-glycosylation was assessed as shifts in protein migration in the SimpleCells when compared with expression in nonaltered HEK293 cells. The positions of the heaviest protein bands in each lane (except TRUNC), present in unmodified HEK293 cells but missing in the SimpleCells, are indicated in the lower image by *red arrowheads*. This reflects the difference in CEL O-glycosylation level between the two cell lines. Secretion was evaluated based on loading equal amounts of protein from the conditioned media. The band observed at ∼60 kDa in all lanes is an unspecific band occasionally observed in the lysate fraction with the antibody used. Representative images from n = 3 independent experiments are shown. HEK293, human embryonic kidney 293.

in SimpleCells (*i.e.*, attachment of GalNAc-α1 onto Ser/Thr residues), which probably explains why size differences still are present after expressing the DEL variants in this cellular model. Taken together, our results indicate that full O-glycosylation is important for efficient expression and secretion of CEL and that the level of O-glycosylation varies considerably between the DEL variants.

Expression of DEL1 and DEL4 in HEK293 cells increases ER stress

Expression of the DEL1 variant induces ER stress in both HEK293 cells and in the pancreatic acinar cell line AR42J (20). In the present study, we wanted to confirm this finding and to compare the ability of the different DEL variants to cause ER stress, both at the mRNA and the protein levels. In addition to analyzing an ER stress master regulator (the chaperone GRP78), all three branches of the unfolded protein response were investigated, namely the PERK, IRE1a, and ATF6 pathways (27). HEK293 cells were transiently transfected with our set of CEL variants, followed by expression analysis of six marker genes (Fig. 5). Cells expressing DEL1 had a 1.3-fold increase in mRNA levels for *XBP1s*, belonging to the IRE1a arm. For DEL1, we also observed increased levels of borderline significance for

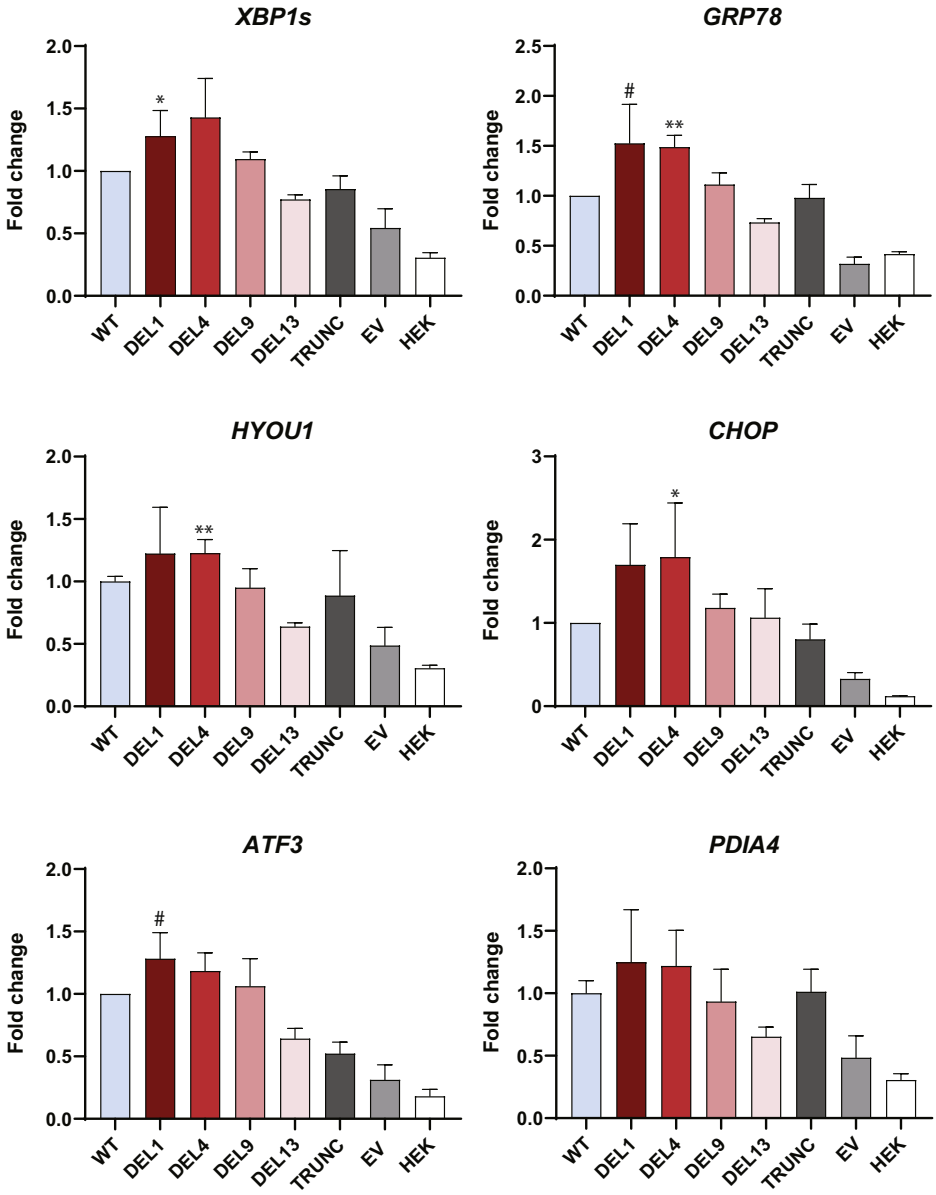

Figure 5. Effect of CEL variants on ER stress markers at the mRNA level. HEK293 cells were transiently transfected with plasmids encoding the CEL variants shown in Figure 1, and mRNA levels were analyzed by real-time quantitative PCR (n = 3). Cells transfected with the empty vector (EV), and untransfected cells (HEK) were used as negative controls. For each marker, the mRNA expression level was corrected to the geometric mean of three reference genes (*OAZ1*, *GAPDH*, and *ACTB*) and then normalized to the marker's expression in cells transfected with CEL-WT. Error bars are SD. Statistical significance is indicated as follows for the mRNA level different from that of the CEL-WT experiment: *$p \le 0.05$; **$p \le 0.01$. Borderline significance (#) for two markers was observed for the DEL1 variant (*GRP78*, $p = 0.055$; *ATF3*, $p = 0.059$). DEL, deletion; ER, endoplasmic reticulum; HEK293, human embryonic kidney 293.

GRP78 and *ATF3* (1.5-fold and 1.3-fold, respectively). In addition, cells expressing the DEL4 variant showed a 1.5-fold higher level for *GRP78*, a 1.2-fold higher expression of the ER folding chaperone *HYOU1* belonging to the ATF6 arm, and 1.8-fold higher expression of the proapoptotic factor *CHOP* belonging to the PERK arm (Fig. 5). DEL9 and DEL13 did not result in significantly elevated levels of any of the tested marker genes.

At the protein level, we examined the effect of CEL variants on the ER chaperones GRP78 and calnexin, and the ER stress transducer PERK. Clear differences between the CEL variants were observed in the insoluble pellet fraction but not in the soluble lysate (Fig. 6*A*). Compared with WT, DEL1 induced significantly higher levels of GRP78 (4.2-fold), calnexin (36-fold), and PERK (8.1-fold) (Fig. 6*B*). For DEL4, we observed a 2- to 10-fold increase in expression of the same ER stress markers although the differences did not quite reach statistical significance because of large experimental variation (Fig. 6*B*).

When the data of Figs. 5 and 6 were considered together, the position of the deletion within the VNTR showed an impact on ER stress signaling: A stress response was induced by both

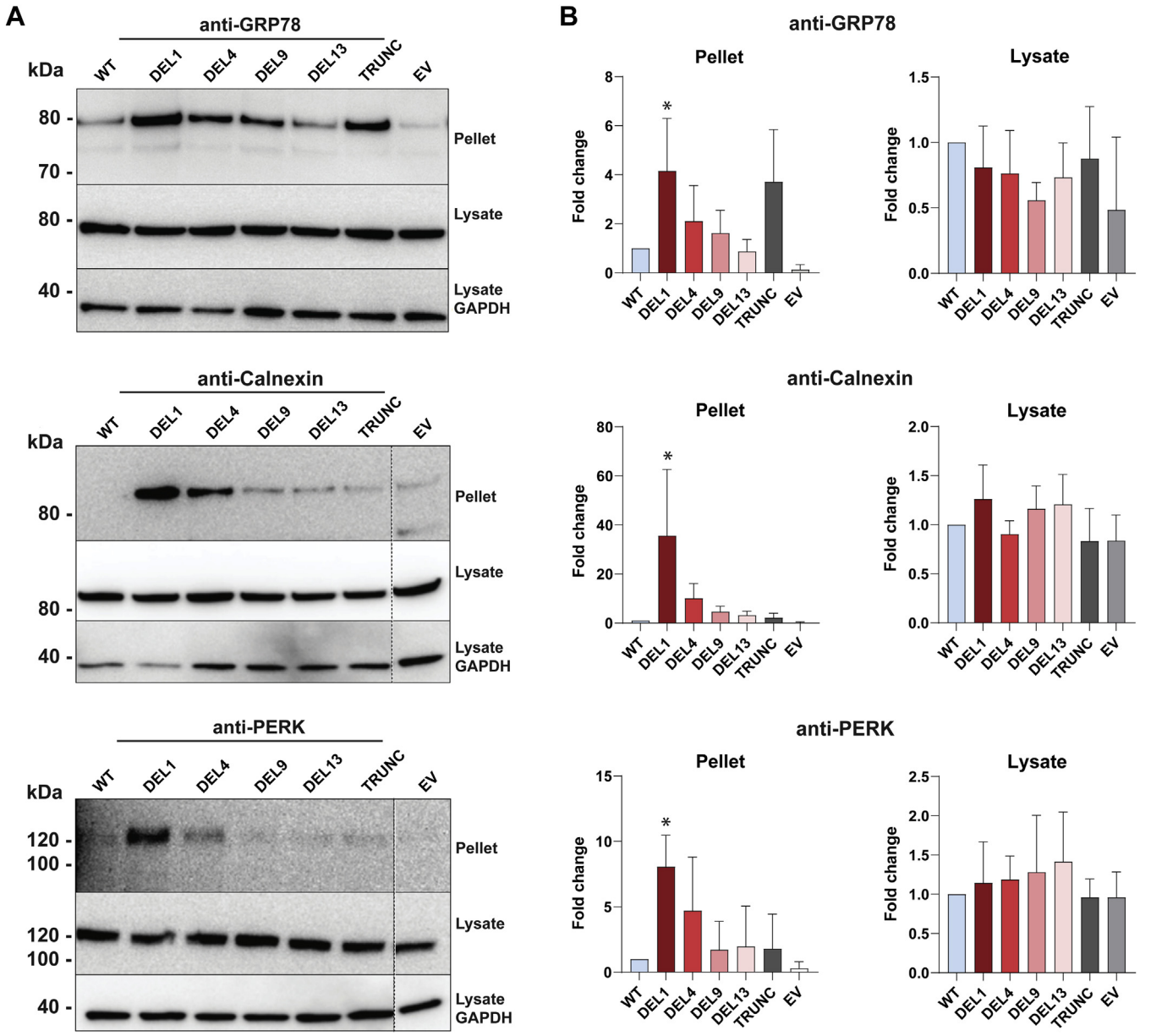

Figure 6. Effect of CEL variants on ER stress markers at the protein level. *A*, HEK293 cells were transiently transfected with plasmids encoding the CEL variants shown in Figure 1, and expression of ER stress markers was assessed by Western blotting of the soluble (lysate) and insoluble (pellet) fractions. Cells transfected with the empty vector (EV) were used as negative control, whereas GAPDH expression was monitored for control of loading. For each of the ER stress markers GRP78, calnexin, and PERK, representative images from n = 3 independent experiments are shown. *Stippled lines* for two of the markers indicate unrelated lanes removed from the images. *B*, quantification of marker band intensities after adjustment to the GAPDH levels and normalization to the expression in CEL-WT–transfected cells. Error bars are SD. The *asterisk* denotes statistical significance ($p \leq 0.05$) for band intensities different from that of the CEL-WT experiment. ER, endoplasmic reticulum; HEK293, human embryonic kidney 293.

DEL1 and DEL4, and this effect leveled off as the deletions were positioned more distally.

DEL variants display lipolytic activity comparable with that of normal CEL

To test whether the single-base deletions could impair the protein's primary function, CEL enzymatic activity was measured. Conditioned medium from HEK293T cells expressing the CEL variants was collected, and the different variants were purified to near homogeneity as demonstrated by SDS-PAGE (Fig. 7*A*). Lipolytic activity in the medium was then measured with trioctanoate as the substrate (Fig. 7*B*). To exclude that any changes in observed activity simply reflected altered levels of secretion, activities were first normalized against the molecular weight of the variant and then plotted after adjusting for the relative band intensities visualized in Figure 7*A*. All variants exhibited enzymatic activities comparable with that of the normal CEL protein (Fig. 7*B*). The relative activities ranged from 76% (DEL1) to 91% (DEL13). The TRUNC variant had an enzyme activity almost identical to that of CEL-WT (95%), suggesting that a complete loss of the VNTR region does not affect the lipolytic activity.

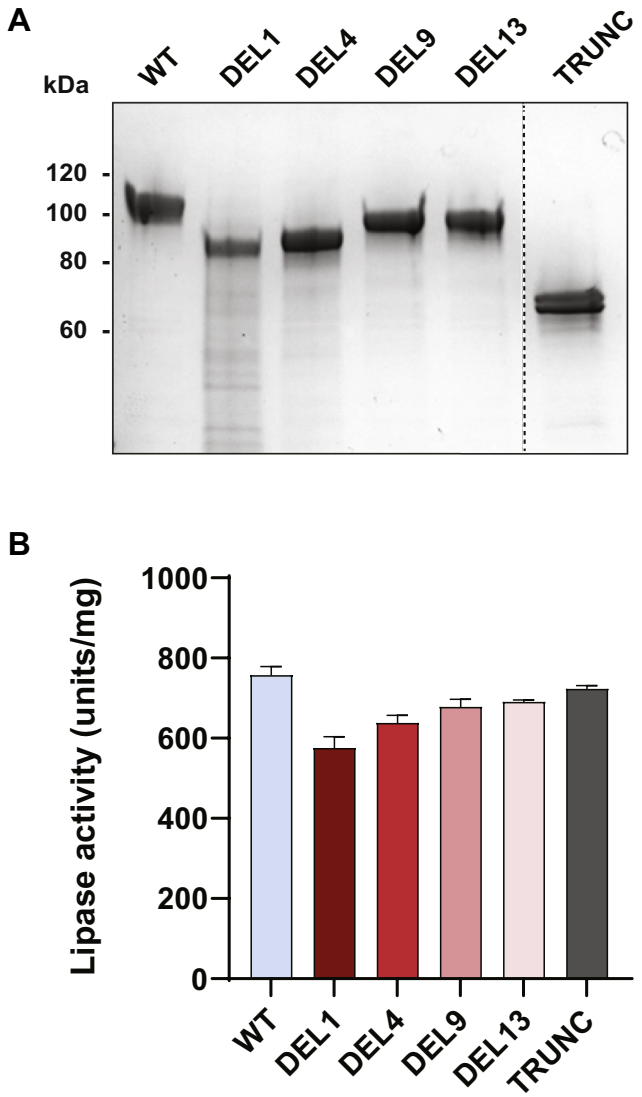

Figure 7. Enzymatic activity of the studied CEL variants. HEK293T cells were transiently transfected with plasmids encoding the CEL variants shown in Figure 1. *A*, SDS-PAGE of CEL variants purified from the conditioned medium and stained with Gel-Code Blue. The same amount of protein was loaded in each lane. The *stippled line* indicates an unrelated lane removed from the image. A representative image from n = 3 independent experiments is shown. *B*, CEL lipolytic activity in the medium measured with trioctanoate as the substrate. The activity was adjusted for band density (as in *panel A*) and molecular weight of each variant. Error bars are SD. HEK293, human embryonic kidney 293.

Discussion

The inherent variation and instability of VNTR sequences imply that these regions are excellent candidates for associations with human disease. We have previously investigated whether the number of VNTR segments in the *CEL* gene influences the risk for exocrine pancreatic disease, however, without finding any evidence that this is the case (9–11). In the current report, we have therefore focused on single–base deletions within the VNTR. These variants are extremely rare and are present in MODY8 families (16) or may occasionally be revealed when screening large cohorts of human DNA samples. As some cellular properties of the DEL1 variant are clearly altered from those of the normal CEL protein (16–20), we here investigated the impact of varying deletion positions within the VNTR.

Owing to the frameshift effect, each of the four DEL variants analyzed in this study has a C-terminal amino acid sequence that differs substantially from that of normal CEL. Moreover, between the DEL proteins, there is also considerable variation in the extension of normal *versus* aberrant sequence repeats (Fig. 1*B*). Discrepancies in cellular properties of the tested variants are therefore likely to reflect biochemical effects arising from their dissimilar amino acid compositions.

Notably, the frameshifts alter the *pI* of the protein variants. The consequence is that the overall acidic charge of the normal CEL protein transforms to the basic side for DEL1, DEL4, and DEL9 (Fig. 1*B*). The positively charged VNTR region observed for DEL1 and DEL4, and to some degree for DEL9, could contribute to aggregation and pathogenicity through change of short- and long-range interactions between CEL protein monomers and between CEL and membrane surfaces (18). It is therefore worth noting that high-molecular-weight protein bands were present only for the DEL1 and DEL4 variants (weak bands in the pellet fraction shown in Fig. 3*A*).

Moreover, the frameshift caused by the deletion mutations introduces several cysteine residues in the aberrant tails of DEL1, DEL4, and DEL9 (12, nine, and four cysteines, respectively; Fig. S2). In contrast, the tails of DEL13 and CEL-WT have no cysteines. In addition to charge effects, the pronounced tendency to aggregate observed for DEL1 and DEL4 might therefore be attributed to formation of intramolecular and intermolecular disulfide bridges between opposing cysteine residues, as discussed previously (20) and also reported in other pathological proteins (28, 29). The aggregation effect may be important for understanding the pathogenicity of the DEL variants, and it could be of interest to analyze effects

of the DEL-specific sequences also outside the context of the CEL protein, for example, by fusing the aberrant protein tails to various reporter proteins.

The DEL13 variant may be protected against aggregation because it keeps an overall acidic charge and has no cysteine residues in its VNTR region. Regarding the observed high-molecular-weight forms of CEL, we speculate that they also could be protein complexes with binding partners such as the GRP94 chaperone (30) or ubiquitinated proteins targeted for proteasomal degradation.

The normal CEL VNTR region is heavily O-glycosylated, a post-translational modification considered important for proper secretion and stability of the protein (13, 24). For several of the DEL variants, two or more bands were detected in the insoluble and soluble cell fractions (Fig. 3*A*), likely to represent fully glycosylated and partially glycosylated CEL, as observed previously (18). Although Western blotting is not an accurate way of estimating the molecular mass of a protein, none of the tested variants migrated with a band size that corresponded to their unmodified amino acid composition. Notably, DEL1 and DEL4, both variants with dramatically fewer O-glycosylation sites, were found to be less well secreted (Fig. 3*A*).

The rationale behind expressing the DEL variants in HEK293 SimpleCells, a cellular model with deficient O-glycosylation, was to obtain variants with an even and truncated glycosylation profile. Ser and Thr residues in the VNTR region would then contain at most one GalNAc residue (with or without a terminal sialyl group). This might enable us to evaluate whether there was a substantial difference in the basic glycosylation pattern among the variants. When expressed in SimpleCells, the CEL proteins displayed a dramatic shift in the band size. All variants migrated with a band size closer to that predicted from their amino acid composition (Fig. 4*B*), except for TRUNC, which has no VNTR sequence and only one theoretical site for O-glycosylation (Fig. S1). Of note, TRUNC was the most abundant variant in the pellet fraction (Fig. 3). Also, when expressed in bacteria, truncated (and nonglycosylated) CEL was found as large aggregates recovered in inclusion bodies (31), further proving the importance of this post-translational modification on CEL secretion. The fact that the band sizes of the DEL variants still varied when expressed in SimpleCells despite very similar predicted molecular masses (Fig. 4*B*) could indicate a large difference in glycan site occupancy. This might be due to the inability of the GalNAc transferases (and sialidases) to reach or recognize certain areas in the aberrant tail, or a consequence of extensive intramolecular or intermolecular disulfide bonds.

The morphological changes of the pancreatic parenchyma observed in MODY8 are consistent with a phenotype of chronic pancreatitis (16, 32). The classic model for understanding the pathogenesis of pancreatitis involves inappropriate intracellular activation of trypsinogen or failure of intracellular trypsin inactivation (33). However, it is becoming increasingly clear that there are also trypsin-independent disease mechanisms, caused by digestive enzyme misfolding and ER stress (21, 34–36). We therefore investigated the effect of DEL variant expression on well-known ER stress markers at the RNA and protein level (Figs. 5 and 6). There was a clear trend for DEL1 and DEL4 to upregulate ER stress markers, both when compared with WT and with the distal DEL variants DEL9 and DEL13. Increased signaling was detected in the three canonical branches of the unfolded protein response that are under the control of PERK, IRE1a, and ATF6. This is consistent with ER stress caused by DEL1 and DEL4 protein aggregation, intramolecular and intermolecular disulfide bridge formation, reduced O-glycosylation and/or impaired transit through the secretory pathway.

As the variants DEL1 and DEL4 led to mRNA upregulation of the proapoptotic marker *CHOP*, we speculate that increased cell death may be implicated in the disease mechanism. This is supported by a previous study, showing that DEL1 induced apoptosis in rat pancreatic acinar cells (20). At the protein level, we found significant differences only in the insoluble pellet fraction and not in the soluble lysate (Fig. 6*B*). For the folding chaperone GRP78, this could be explained by the protein binding tightly to misfolded CEL variants, thereby being pulled down in the pellet fraction as a protein complex.

All DEL variants displayed lipolytic activity comparable with that of the normal CEL protein (Fig. 7). There was a slight, gradual decrease in relative activity from DEL13 (91%) to DEL1 (76%), suggesting that the aberrant protein repeats might have a minor effect on CEL enzymatic function. The TRUNC variant displayed almost normal lipolytic activity (95%), a result in line with previous studies showing that complete truncation of the VNTR region does not compromise the enzymatic ability of CEL (37, 38).

Under any circumstances, we find it unlikely that reduced lipolytic activities can explain pathogenicity of the DEL variants. First, the observed differences do not appear large enough to alter physiological function of CEL, although the possibility remains that activity against other substrates than trioctanoate could be more severely affected. Second, a mouse model with a full-body KO of CEL did not recapitulate any aspect of the MODY8 pancreatic phenotype (39). Third, the cytotoxic effects observed for different variants of the pathogenic CEL-HYB1 protein (40) could not be attributed to differences in the lipolytic activity (41). Finally, the presence of a normal gene in the heterozygous state is expected to compensate for any loss of activity from a rare *CEL* mutation. Taken together, all available data favor the hypothesis that pathogenic CEL variants initiate pancreatic disease through a dominant-negative gain-of-function effect, rather than *via* a loss-of-function mechanism.

A limitation of our study is that pancreatic acinar cells were not used as model system. Still, our data demonstrate that the position of single-base deletions in the *CEL* VNTR has great impact on protein cytotoxicity. When analyzed in transfected HEK293 cells, the pathogenic DEL1 and DEL4 variants found in MODY8 families had the strongest effect on protein secretion, protein aggregation, and ER stress. On the other hand, DEL13 behaved more like normal CEL and must be considered a benign variant. We evaluate DEL9 to have intermediate effects, but given that this variant so far has not been reported in any MODY family, we conclude that it is

likely benign in this disease context. The *CEL* VNTR is a region very prone to mutational events (5). Should single-base deletions in the second or third VNTR repeat (*i.e.*, DEL2 and DEL3) be identified, we predict that they will be pathogenic. For any DEL5–DEL8 variant that might be observed, genetic segregation analysis and functional evaluation will be necessary to determine their pathogenic potential.

Experimental procedures

Plasmid constructions

Previously, we made the plasmid pcDNA3/CEL-WT 14R (with 14 VNTR segments) for expression of the normal CEL protein in mammalian cells (20). For this study, a *Psh*AI/*Xho*I cDNA fragment encoding normal CEL with 16 VNTR segments was synthesized by GenScript and used to replace the corresponding fragment in pcDNA3/CEL-WT 14R. All four *CEL* DEL constructs were then made through mutagenesis service provided by GenScript using the newly created pcDNA3/CEL-WT with 16 repeated segments as template. The CEL-TRUNC construct was generated by QuikChange II XL site-directed mutagenesis kit (Agilent Technologies) using pcDNA3/CEL-WT 14R as template (20).

For experiments comparing expression of epitope-tagged *versus* untagged CEL variants, the tagged plasmids have been described before (19). In short, the constructs were all based on a pcDNA3.1 vector containing a V5/His-tag (Invitrogen). Untagged constructs for comparison were prepared by introducing a stop codon into the pcDNA3.1/V5-His constructs using the same site-directed mutagenesis kit as above. The plasmid contains an *Xho*I restriction site directly before the epitope tag sequence, in which the stop codon was created.

Antibodies and reagents

Two anti-CEL antibodies were used: a rabbit polyclonal antibody for immunoblotting (against the truncated CEL variant pV562Δ, as described (20)), and a mouse mAb for immunocytochemistry (As20.1, against the CEL globular domain), kindly provided by Prof. Olle Hernell (Department of Clinical Sciences, Umeå University, Umeå, Sweden). Additional primary antibodies were rabbit polyclonal anti-GRP78 antibody (cat. 21685, Abcam), rabbit monoclonal anti-calnexin (C5C9, cat. 2679) and mouse monoclonal anti-PERK (D11A8, cat. 5683) from Cell Signaling Technology, and mouse monoclonal anti-GAPDH (sc-47724, Santa Cruz Biotechnology,). Secondary antibodies were HRP-goat anti-rabbit (cat. 656120, Invitrogen), HRP-donkey anti-mouse (sc-2318, Santa Cruz Biotechnology), and Alexa Fluor 488 anti-mouse (cat. A11017, Invitrogen).

RNeasy Mini Kit was from Qiagen. Polyethyleneimine (PEI) was from Polysciences. High-Capacity cDNA Reverse Transcription Kit was from Applied Biosystems. Amersham Hybond P Western Blotting polyvinylidene difluoride membranes were from GE Healthcare. TGX Precast protein gel (4–15%) and Blotting-Grade Blocker were from Bio-Rad. Poly-L-lysine, fetal bovine serum (FBS), PBS, Penicillin-Streptomycin, and cOmplete Mini Protease inhibitor cocktail tablets were from Sigma-Aldrich. ProLong Gold Antifade Solution containing 4′,6-diamidino-2-phenylindole nuclear stain, Antibiotic Antimycotic, and Lipofectamine 2000 were from Invitrogen. Radioimmunoprecipitation assay buffer lysis buffer (10×), Triton X-100, and Tween-20 were from Merck Millipore. NuPAGE Sample Reducing Agent (10×), NuPAGE lithium dodecyl sulphate (LDS) Sample Buffer (4×), Dulbecco's modified Eagle's medium (DMEM) with high glucose (4500 mg/L), NuPAGE Bis-Tris Protein gels (10% or 4–12%, 1.0 mm, ten-well), Pierce BCA Protein Assay, Pierce ECL Plus Western Blotting Substrate, Opti-MEM I Reduced Serum Medium, and GelCode Blue Stain Reagent were obtained from Thermo Fisher Scientific.

Cell culture of HEK293 and HEK293T cells

HEK293 cells (Clontech Laboratories) were used for most experiments described in this study. For the enzymatic activity assay, HEK293T cells from American Type Culture Collection were used. HEK293 Cosmc-KO cells (referred to as SimpleCells throughout the text) and their unaltered counterpart were obtained from the laboratory of Prof. Henrik Clausen (Copenhagen Center for Glycomics, Denmark). All cell lines were maintained in a growth medium with high glucose (4500 mg/L) supplemented with 10% FBS and either 100 U/ml of Antibiotic Antimycotic or 10 mg/ml of Penicillin-Streptomycin in humidified atmosphere with 5% CO_2 at 37 °C.

Transient transfections

For transfection of HEK293 cells, Lipofectamine 2000 was used according to the manufacturer's instructions. Cells were transfected with plasmids encoding CEL-WT, single-base DEL variants (DEL1, DEL4, DEL9, or DEL13), or CEL-TRUNC, either with or without the V5/His tag. The empty vector (either pcDNA3 or pcDNA3.1-V5/His) was included as a negative control. Samples were harvested for analysis 48 h after transfection for all experiments.

For the enzymatic activity assay, HEK293T cells were used. Cells were grown in 10-cm culture dishes in a regular medium until reaching 70 to 90% confluence. Next, the cells were transfected with plasmids expressing different CEL variants using PEI, which had been dissolved in ddH_2O at a concentration of 1 μg/μl (pH 7.4) and sterilized using a 0.22-μm filter. Before transfection, cells were switched to 10 ml DMEM without FBS. For one dish, 15 μg plasmid DNA was mixed with 45 μl PEI solution in 500 μl Opti-MEM I Reduced Serum Medium. The mixture was added to the culture medium of HEK293T cells after 30-min incubation at room temperature (RT). The cells were grown for 16 to 20 h before the medium was switched to 10 ml Opti-MEM I Reduced Serum Medium containing antibiotics. The cells were grown for another 24 h, the conditioned medium was harvested, and the cells were refreshed with another 10 ml of Opti-MEM I Reduced Serum Medium. The conditioned medium was again collected after 24-h incubation. The harvested media were pooled and filtered with 0.45-μm Steritop bottle top filters (Millipore) to remove debris before protein purification.

Cell fractionation and collection of conditioned media

The conditioned medium was collected and analyzed as the medium fraction. The cells were lysed in ice-cold radioimmunoprecipitation assay buffer supplemented with the protease inhibitor. After 30-min incubation on ice, the lysate was centrifuged at 14,000*g* for 15 min at 4 °C. The supernatant was isolated and analyzed as the soluble lysate fraction. The pellet was washed twice in PBS, boiled at 95 °C in 2× LDS sample buffer and reducing agent, and analyzed as the insoluble pellet fraction. All fractions were analyzed by Western blotting.

Western blotting

For detection of CEL protein, 4 μg of the soluble lysate fraction was used. For the medium and insoluble pellet fractions, the volume loaded on the SDS gel was the same as for the corresponding lysate. LDS sample buffer and reducing agent were added to all samples. Before loading, the medium and soluble lysate fractions were denatured on a heating block at 56 °C for 15 min, whereas the pellet fractions were denatured at 95 °C for 15 min. The samples were loaded and separated by 10% SDS-PAGE at 90 V for 15 min and then at 180 V for 1 h. Proteins were transferred to a polyvinylidene difluoride membrane by semiwet blotting using XCell Blot Module chambers (Invitrogen) according to manufacturer's instructions. The proteins were detected using standard methods, with primary antibody incubation overnight at 4 °C and secondary antibody incubation for 1 h at RT. Blots were developed using Pierce ECL Plus Western Blotting Substrate, and signals were analyzed using a Las 1000 Pro v2.6 software imager (Fujifilm). Band quantification was performed using the Multi Gauge v3.0 software (Fujifilm). For ER stress markers, 20 μg (GRP78 and calnexin) or 40 μg (PERK) of the insoluble pellet fraction was loaded and separated by 4 to 12% SDS-PAGE. For ER stress analyses of the soluble lysate fractions, 10 μg protein was loaded on the gel. For analyses of CEL expression in transfected SimpleCells, 10 μg protein was loaded on the gel and separated by 10% SDS-PAGE.

Immunocytochemistry and confocal microscopy

HEK293 cells were seeded onto coverslips (18 mm) coated with poly-L-lysine, transfected as described above and fixed in 3% paraformaldehyde for 30 min. Immunostaining was performed as described previously (17). Results were evaluated with an SP5 AOBS confocal microscope (Leica Microsystems) with 63×/1.4 NA NAHCX Plan-Apochromat oil immersion objective, ∼1.2 airy unit pinhole aperture, and appropriate filter combinations. Images were acquired with 405 diode and argon ion/argon krypton lasers (Leica) and processed using LAS AF Lite (Leica), Photoshop CC, and Adobe Illustrator CC (Adobe Systems).

RNA extraction and real-time quantitative PCR

The medium was removed, and the cells were washed before total RNA was extracted using the RNeasy Mini Kit according to the manufacturer's protocol. For cDNA synthesis, the High-Capacity cDNA Reverse Transcription Kit was used with total RNA input of 20 ng for each cDNA synthesis reaction. The real-time PCR was performed with SyBR Green on a Rotor-Gene Q cycler (Qiagen). Gene expression was calculated as copies/μl using the standard curve approach. The standards were made with suitable primers in a conventional PCR. *GAPDH*, *OAZ1*, and *ACTB* were used as reference genes.

Protein purification and enzymatic activity assay

CEL protein variants were purified by HiTrap Heparin HP Affinity Column (GE Healthcare Life Sciences) controlled by an ÄKTA pure protein purification system as described previously (20). Protein concentration was determined by measuring absorbance at 280 nm using the corresponding extinction coefficient for each CEL variant. Five micrograms of each purified protein was resolved by 4 to 15% TGX Precast protein gel and stained with GelCode Blue stain reagent to assess molecular mass, homogeneity, and integrity. Functional characterization of recombinant CEL protein variants was carried out by measuring lipase activity against trioctanoate in the presence of 12 mM sodium cholate using a pH-stat (Radiometer Analytical) at pH 8.0. For each assay, 5 μg of purified CEL protein was used. The lipolytic activities were expressed in international lipase units/mg of the enzyme (1 unit corresponds to 1 μmol of fatty acid released/min). The activities were adjusted according to densitometry and to the molecular weight of each variant.

Statistical analysis

The two-tailed Student's *t* test was used for all statistical analyses. Statistical calculations were carried out using Microsoft Excel 2016. *p*-values ≤ 0.05 were considered as statistically significant. Results are given as the mean ± SD.

Data availability

All data that support the findings of this study are included within the article.

Supporting information—This article contains supporting information.

Acknowledgments—The authors thank Olle Hernell, Umeå University, Sweden, and Henrik Clausen, Copenhagen Center for Glycomics, University of Copenhagen, for kindly providing the As20.1 antibody and HEK293 SimpleCells, respectively.

Author contributions—A. M., M. E. L., and K. F. conceived, designed, and directed the study. A. G., X. X., M. C., K. E. J., B. B. J., and K. F. performed experiments and interpreted data. S. J. provided genetic data. A. G. and A. M. wrote the manuscript with substantial contributions from X. X., M. C., P. R. N., M. E. L., B. B. J., and K. F. All authors contributed to the revision of the manuscript and approved the final version.

Funding and additional information—This research was funded by a PhD fellowship to A. G. from the University of Bergen and by grants to A. M. from Western Norway Regional Health Authority (Helse Vest 912057) and the Research Council of Norway

(FRIMEDBIO program 289534). This research was also supported by a Translational Medical Center grant from Stiftelsen Kristian Gerhard Jebsen to P. R. N., by a grant to M. E. L. from the NIH/National Institute of Diabetes and Digestive and Kidney Diseases (R01DK124415), and by grants from the Fonds National de la Recherche Scientifique (FNRS), the Brussels Region Innoviris project DiaType, and the Fonds Erasme for Medical Research to M. C. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: ER, endoplasmic reticulum; DEL, deletion; FBS, fetal bovine serum; HEK293, human embryonic kidney 293; LDS, lithium dodecyl sulphate; MODY8, maturity-onset diabetes of the young, type 8; VNTR, variable number of tandem repeat.

References

1. Brookes, K. J. (2013) The VNTR in complex disorders: The forgotten polymorphisms? A functional way forward? *Genomics* **101**, 273–281
2. Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985) Individual-specific 'fingerprints' of human DNA. *Nature* **316**, 76–79
3. Cai, C. Q., Zhang, T., Breslin, M. B., Giraud, M., and Lan, M. S. (2011) Both polymorphic variable number of tandem repeats and autoimmune regulator modulate differential expression of insulin in human thymic epithelial cells. *Diabetes* **60**, 336–344
4. Simpson, J., Vetuz, G., Wilson, M., Brookes, K. J., and Kent, L. (2010) The DRD4 receptor Exon 3 VNTR and 5' SNP variants and mRNA expression in human post-mortem brain tissue. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **153**, 1228–1233
5. Johansson, B. B., Fjeld, K., El Jellas, K., Gravdal, A., Dalva, M., Tjora, E., Ræder, H., Kulkarni, R. N., Johansson, S., Njølstad, P. R., and Molven, A. (2018) The role of the carboxyl ester lipase (CEL) gene in pancreatic disease. *Pancreatology* **18**, 12–19
6. Higuchi, S., Nakamura, Y., and Saito, S. (2002) Characterization of a VNTR polymorphism in the coding region of the CEL gene. *J. Hum. Genet.* **47**, 213–215
7. Lindquist, S., Blackberg, L., and Hernell, O. (2002) Human bile salt-stimulated lipase has a high frequency of size variation due to a hypervariable region in exon 11. *Eur. J. Biochem.* **269**, 759–767
8. Torsvik, J., Johansson, S., Johansen, A., Ek, J., Minton, J., Ræder, H., Ellard, S., Hattersley, A., Pedersen, O., Hansen, T., Molven, A., and Njølstad, P. R. (2010) Mutations in the VNTR of the carboxyl-ester lipase gene (CEL) are a rare cause of monogenic diabetes. *Hum. Genet.* **127**, 55–64
9. Ragvin, A., Fjeld, K., Weiss, F. U., Torsvik, J., Aghdassi, A., Mayerle, J., Simon, P., Njølstad, P. R., Lerch, M. M., Johansson, S., and Molven, A. (2013) The number of tandem repeats in the carboxyl-ester lipase (CEL) gene as a risk factor in alcoholic and idiopathic chronic pancreatitis. *Pancreatology* **13**, 29–32
10. Fjeld, K., Beer, S., Johnstone, M., Zimmer, C., Mossner, J., Ruffert, C., Krehan, M., Zapf, C., Njølstad, P. R., Johansson, S., Bugert, P., Miyajima, F., Liloglou, T., Brown, L. J., Winn, S. A., *et al.* (2016) Length of variable numbers of tandem repeats in the carboxyl ester lipase (CEL) gene may confer susceptibility to alcoholic liver cirrhosis but not alcoholic chronic pancreatitis. *PLoS One* **11**, e0165567
11. Dalva, M., El Jellas, K., Steine, S. J., Johansson, B. B., Ringdal, M., Torsvik, J., Immervoll, H., Hoem, D., Laemmerhirt, F., Simon, P., Lerch, M. M., Johansson, S., Njølstad, P. R., Weiss, F. U., Fjeld, K., *et al.* (2017) Copy number variants and VNTR length polymorphisms of the carboxyl-ester lipase (CEL) gene as risk factors in pancreatic cancer. *Pancreatology* **17**, 83–88
12. Abouakil, N., Mas, E., Bruneau, N., Benajiba, A., and Lombardo, D. (1993) Bile salt-dependent lipase biosynthesis in rat pancreatic AR 4-2 J cells. Essential requirement of N-linked oligosaccharide for secretion and expression of a fully active enzyme. *J. Biol. Chem.* **268**, 25755–25763
13. Wang, C. S., Dashti, A., Jackson, K. W., Yeh, J. C., Cummings, R. D., and Tang, J. (1995) Isolation and characterization of human milk bile salt-activated lipase C-tail fragment. *Biochemistry* **34**, 10639–10644
14. Loomes, K. M., Senior, H. E., West, P. M., and Roberton, A. M. (1999) Functional protective role for mucin glycosylated repetitive domains. *Eur. J. Biochem.* **266**, 105–111
15. El Jellas, K., Johansson, B. B., Fjeld, K., Antonopoulos, A., Immervoll, H., Choi, M. H., Hoem, D., Lowe, M. E., Lombardo, D., Njølstad, P. R., Dell, A., Mas, E., Haslam, S. M., and Molven, A. (2018) The mucinous domain of pancreatic carboxyl-ester lipase (CEL) contains core 1/core 2 O-glycans that can be modified by ABO blood group determinants. *J. Biol. Chem.* **293**, 19476–19491
16. Ræder, H., Johansson, S., Holm, P. I., Haldorsen, I. S., Mas, E., Sbarra, V., Nermoen, I., Eide, S.Å., Grevle, L., Bjørkhaug, L., Sagen, J. V., Aksnes, L., Søvik, O., Lombardo, D., Molven, A., *et al.* (2006) Mutations in the *CEL* VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. *Nat. Genet.* **38**, 54–62
17. Dalva, M., Lavik, I. K., El Jellas, K., Gravdal, A., Lugea, A., Pandol, S. J., Njølstad, P. R., Waldron, R. T., Fjeld, K., Johansson, B. B., and Molven, A. (2020) Pathogenic carboxyl ester lipase (CEL) variants interact with the normal CEL protein in pancreatic cells. *Cells* **9**, 244
18. Johansson, B. B., Torsvik, J., Bjørkhaug, L., Vesterhus, M., Ragvin, A., Tjora, E., Fjeld, K., Hoem, D., Johansson, S., Ræder, H., Lindquist, S., Hernell, O., Cnop, M., Saraste, J., Flatmark, T., *et al.* (2011) Diabetes and pancreatic exocrine dysfunction due to mutations in the carboxyl ester lipase gene-maturity onset diabetes of the young (CEL-MODY): A protein misfolding disease. *J. Biol. Chem.* **286**, 34593–34605
19. Torsvik, J., Johansson, B. B., Dalva, M., Marie, M., Fjeld, K., Johansson, S., Bjørkøy, G., Saraste, J., Njølstad, P. R., and Molven, A. (2014) Endocytosis of secreted carboxyl ester lipase in a syndrome of diabetes and pancreatic exocrine dysfunction. *J. Biol. Chem.* **289**, 29097–29111
20. Xiao, X., Jones, G., Sevilla, W. A., Stolz, D. B., Magee, K. E., Haughney, M., Mukherjee, A., Wang, Y., and Lowe, M. E. (2016) A carboxyl ester lipase (CEL) mutant causes chronic pancreatitis by forming intracellular aggregates that activate apoptosis. *J. Biol. Chem.* **291**, 23224–23236
21. Sahin-Toth, M. (2017) Genetic risk in chronic pancreatitis: The misfolding-dependent pathway. *Curr. Opin. Gastroenterol.* **33**, 390–395
22. Amarasinghe, C., and Jin, J. P. (2015) The use of affinity tags to overcome obstacles in recombinant protein expression and purification. *Protein Pept. Lett.* **22**, 885–892
23. Kosobokova, E. N., Skrypnik, K. A., and Kosorukov, V. S. (2016) Overview of fusion tags for recombinant proteins. *Biochemistry (Mosc.)* **81**, 187–200
24. Bruneau, N., Nganga, A., Fisher, E. A., and Lombardo, D. (1997) O-Glycosylation of C-terminal tandem-repeated sequences regulates the secretion of rat pancreatic bile salt-dependent lipase. *J. Biol. Chem.* **272**, 27353–27361
25. Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T. B. G., Lavrsen, K., Dabelsteen, S., Pedersen, N. B., Marcos-Silva, L., Gupta, R., Bennett, E. P., Mandel, U., Brunak, S., Wandall, H. H., *et al.* (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* **32**, 1478–1488
26. Ju, T., Aryal, R. P., Stowell, C. J., and Cummings, R. D. (2008) Regulation of protein O-glycosylation by the endoplasmic reticulum–localized molecular chaperone Cosmc. *J. Cell Biol.* **182**, 531–542
27. Hetz, C., and Papa, F. R. (2018) The unfolded protein response and cell fate control. *Mol. Cell* **69**, 169–181
28. Lee, S., and Eisenberg, D. (2003) Seeded conversion of recombinant prion protein to a disulfide-bonded oligomer by a reduction-oxidation process. *Nat. Struct. Biol.* **10**, 725–730
29. Rabdano, S. O., Izmailov, S. A., Luzik, D. A., Groves, A., Podkorytov, I. S., and Skrynnikov, N. R. (2017) Onset of disorder and protein aggregation due to oxidation-induced intermolecular disulfide bonds: Case study of RRM2 domain from TDP-43. *Sci. Rep.* **7**, 11161

30. Bruneau, N., and Lombardo, D. (1995) Chaperone function of a Grp 94-related protein for folding and transport of the pancreatic bile salt-dependent lipase. *J. Biol. Chem.* **270**, 13524–13533
31. Terzyan, S., Wang, C. S., Downs, D., Hunter, B., and Zhang, X. C. (2000) Crystal structure of the catalytic domain of human bile salt activated lipase. *Protein Sci.* **9**, 1783–1790
32. Ræder, H., McAllister, F. E., Tjora, E., Bhatt, S., Haldorsen, I., Hu, J., Willems, S. M., Vesterhus, M., El Ouaamari, A., Liu, M., Ræder, M. B., Immervoll, H., Hoem, D., Dimcevski, G., Njølstad, P. R., *et al.* (2014) Carboxyl-ester lipase maturity-onset diabetes of the young is associated with development of pancreatic cysts and upregulated MAPK signaling in secretin-stimulated duodenal fluid. *Diabetes* **63**, 259–269
33. Mayerle, J., Sendler, M., Hegyi, E., Beyer, G., Lerch, M. M., and Sahin-Toth, M. (2019) Genetics, cell biology, and pathophysiology of pancreatitis. *Gastroenterology* **156**, 1951–1968.e1951
34. Kereszturi, E., Szmola, R., Kukor, Z., Simon, P., Weiss, F. U., Lerch, M. M., and Sahin-Tóth, M. (2009) Hereditary pancreatitis caused by mutation-induced misfolding of human cationic trypsinogen: A novel disease mechanism. *Hum. Mutat.* **30**, 575–582
35. Pandol, S., Gorelick, F., and Lugea, A. (2011) Environmental and genetic stressors and the unfolded protein response in exocrine pancreatic function – a hypothesis. *Front. Physiol.* **2**, 8
36. Hegyi, E., and Sahin-Tóth, M. (2019) Human *CPA1* mutation causes digestive enzyme misfolding and chronic pancreatitis in mice. *Gut* **68**, 301–312
37. Downs, D., Xu, Y.-Y., Tang, J., and Wang, C.-S. (1994) Proline-rich domain and glycosylation are not essential for the enzymic activity of bile salt-activated lipase. Kinetic studies of T-BAL, a truncated form of the enzyme, expressed in Escherichia coli. *Biochemistry* **33**, 7979–7985
38. Bläckberg, L., Strömqvist, M., Edlund, M., Juneblad, K., Lundberg, L., Hansson, L., and Hernell, O. (1995) Recombinant human-milk bile-salt-stimulated lipase. Functional properties are retained in the absence of glycosylation and the unique proline-rich repeats. *Eur. J. Biochem.* **228**, 817–821
39. Vesterhus, M., Ræder, H., Kurpad, A. J., Kawamori, D., Molven, A., Kulkarni, R. N., Kahn, C. R., and Njølstad, P. R. (2010) Pancreatic function in carboxyl-ester lipase knockout mice. *Pancreatology* **10**, 467–476
40. Fjeld, K., Weiss, F. U., Lasher, D., Rosendahl, J., Chen, J.-M., Johansson, B. B., Kirsten, H., Ruffert, C., Masson, E., Steine, S. J., Bugert, P., Cnop, M., Grützmann, R., Mayerle, J., Mössner, J., *et al.* (2015) A recombined allele of the lipase gene *CEL* and its pseudogene *CELP* confers susceptibility to chronic pancreatitis. *Nat. Genet.* **47**, 518–522
41. Cassidy, B. M., Zino, S., Fjeld, K., Molven, A., Lowe, M. E., and Xiao, X. (2020) Single nucleotide polymorphisms in *CEL-HYB1* increase risk for chronic pancreatitis through proteotoxic misfolding. *Hum. Mutat.* **41**, 1967–1978