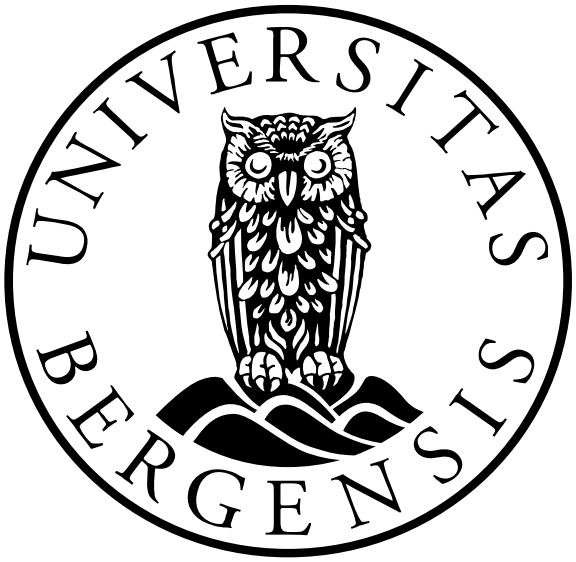
# Functional characterization of Carboxyl Ester Lipase variants causing diabetes and exocrine dysfunction

**Anny Gravdal**

This thesis is submitted in partial fulfilment of the requirements for the degree of Master of Science




Department of Molecular Biology and
KG Jebsen Centre for Diabetes Research
University of Bergen
Center for Medical Genetics and Molecular Medicine
Haukeland University Hospital

November 2016

# Table of Contents

**Acknowledgements** ... **3**
**Abbreviations** ... **4**
**Abstract** ... **5**
**1. Introduction** ... **6**
1.1. The pancreas ... 6
1.1.1. The exocrine pancreas ... 6
1.1.2. The endocrine pancreas ... 7
1.2. Diseases of the pancreas ... 8
1.2.1. Diabetes mellitus ... 8
1.2.2. Pancreatitis ... 9
1.2.3. Pancreatic cancer ... 10
1.3. Carboxyl ester lipase ... 10
1.3.1. The *CEL* gene ... 11
1.3.2. CEL protein structure and secretion ... 12
1.3.3. The polymorphic nature of *CEL* ... 14
1.4. Pathogenic variants of CEL ... 16
1.4.1. CEL-DEL1 (CEL-MODY) ... 16
1.4.2. CEL-HYB ... 17
1.4.3. CEL VNTR length polymorphisms in disease ... 19
1.5. The use of epitope tags in experimental research ... 19
**2. Aims of the study** ... **21**
**3. Materials** ... **22**
**4. Methods** ... **28**
4.1. Plasmid preparation and evaluation of DNA quality ... 28
4.1.1. Transformation of OneShot® TOP10 chemically competent *E.coli* cells ... 28
4.1.2. Plasmid preparation ... 28
4.1.3. Determination of plasmid yield and quality ... 28
4.1.3.1. OD measurements ... 28
4.1.3.2. Agarose gel electrophoresis ... 29
4.2. Constructing a *CEL-HYB* plasmid without an epitope tag ... 29
4.2.1. Primer design and mutagenesis ... 29
4.2.2. *DpnI* digestion ... 30
4.2.3. Sequencing of *CEL-HYB* ... 31
4.3. Cell cultures and transfection ... 32
4.3.1. Culturing ... 32
4.3.2. Sub-culturing and seeding ... 32
4.3.3. Freezing and thawing ... 32
4.3.4. Transient transfection of HEK293 cells ... 32

4.4. Preparation of analytical fractions: cell lysate, cell pellet and medium ..... 33
4.4.1. Protein concentration determination ..... 33
4.5. SDS-PAGE and immunoblotting ..... 34
4.6. Immunostaining ..... 34
4.7. Cell viability assay ..... 35
4.8. Apoptosis assay ..... 36
4.9. Statistical analysis ..... 37
4.10. De-glycosylation assay ..... 37
4.11. Isoelectric Focusing ..... 37
**5. Results ..... 38**
5.1. Construction of a *CEL-HYB* plasmid without the V5/His tag ..... 38
5.2. Plasmid purification and determination of DNA purity ..... 39
5.3. The effect of the C-terminal V5/His-tag on CEL protein expression and secretion ..... 40
5.3.1. Western blot analysis ..... 40
5.3.2. Immunostaining and confocal analysis ..... 43
5.4. The effect of CEL VNTR length on protein expression and secretion ..... 46
5.4.1. Western blot analysis ..... 46
5.5. The effect of CEL on cell viability and apoptosis ..... 48
5.5.1. Cell viability ..... 48
5.5.2. Apoptosis ..... 50
5.6. De-glycosylation of the CEL DEL1, -DEL4 and -DEL8 proteins ..... 54
5.7. Isoelectric focusing ..... 56
**6. Discussion ..... 58**
6.1. Comparison of V5/His tagged- versus untagged CEL protein variants ..... 58
6.2. The effect of CEL VNTR lengths on protein expression, localisation and secretion ..... 59
6.3. The effect of CEL variants on cell viability and apoptosis in HEK293 cells ..... 60
6.4. CEL VNTR sequence and length alter glycosylation and protein charge ..... 61
6.4.1. De-glycosylation of CEL-DEL1, -DEL4 and -DEL8 ..... 61
6.4.2. Isoelectric focusing ..... 62
6.5. HEK293 cells as a research tool for studying CEL ..... 64
**7. Conclusions ..... 65**
**8. Future perspectives ..... 66**
**References ..... 67**
**Appendix ..... 74**
I: CEL variants exon 11 DNA sequence ..... 74
II: CEL variants exon 11 amino acid sequence ..... 77

# Acknowledgements

Firstly, I would like to express my sincere gratitude to my incredible supervisors, Dr. Karianne Fjeld and Dr. Bente Berg Johansson. Thank you for everything that you have taught me this year, and for your enthusiasm throughout the learning process of my master project, great counselling, advice and support along the way. Thank you for being patient and kind, but also strict when it was needed. I am grateful for all comments through the writing process and for stimulating me to become a more independent student. I could not have imagined better supervisors for my master project.

My sincere thanks also goes to my co-supervisors, Prof. Anders Molven and Prof. Pål Rasmus Njølstad. In particular, I wish to thank you, Anders, for your insightful comments and remarks, and for always asking me the difficult questions. I am also indebted to you for valuable help in the writing process, and for always motivating me.

I would like to thank PhD students Monica Dalva and Khadija El Jellas. Thank you Monica for all the help in the lab, and your engagement in my thesis during this year, especially during the writing process. I am grateful for the many conversations and all the laughter. Also, thank you, Khadija, for your useful comments and remarks during the writing of this thesis, and help with experiments. It was truly appreciated.

My sincere thanks are also directed to Prof. Mark Lowe and his research group in Pittsburgh, USA for kindly providing us with many of the *CEL* variants that were used throughout this project. I am also grateful to the entire MODY group for being so welcoming and kind. I feel immensely lucky that I ended up in this research group. I would also like to thank Center for Medical Genetics and Molecular Medicine (MGM) for hosting me and providing laboratory space, and in particular Monika Ringdal.

Lastly, I would like to thank my loved ones. Thank you to my closest friends for being a distraction from the work when I needed it most.

Thank you to my entire family, for always believing in me, and always pushing me keep going, even when I wanted to give up. I would never have managed to do this without your love and support. Alex, I will forever be grateful for your support throughout this entire process.

Bergen, November 2016

Anny Gravdal

# Abbreviations

| | |
|---|---|
| ATP | Adenosine Triphosphate |
| C | Celsius |
| CEL/*CEL* | Carboxyl ester lipase/gene |
| *CELP* | Carboxyl ester lipase pseudogene |
| $ddH_2O$ | Double deionized water |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| *E.coli* | *Escherichia coli* |
| EDTA | Ethylenediaminetetraacetic acid |
| EV | Empty vector |
| g | Gram |
| GAPDH | Glyceraldehyde 3-phosphate dehydrogenase |
| h | Hours |
| HEK293 | Human embryonic kidney cell line 293 |
| HYB | Hybrid |
| IEF | Isoelectric focusing |
| kDa | Kilo Daltons |
| M | Molar |
| min | Minutes |
| mg | Milligram |
| mL | Milliliter |
| MUT | Mutant |
| o/n | overnight |
| *pI* | Isoelectric point |
| RCF | Relative centrifugal force |
| rpm | Revolutions per minute |
| sec | Seconds |
| v/v | Volume/volume |
| w/v | Weight/volume |
| µg | Microgram |
| µl | Microliter |
| WT | Wild-type |
| ~ | Approximately |

# Abstract

The digestive enzyme carboxyl ester lipase (CEL) is mainly expressed in the acinar cells of the pancreas. We have previously identified disease-causing mutations in a variable number of tandem repeats (VNTR) domain, localised in the last exon of the *CEL* gene. Single base deletions (*DEL1* or *DEL4*) in the VNTR lead to frameshifts and a truncated C-terminal of the CEL protein. Patients with these mutations suffer from exocrine pancreatic dysfunction and diabetes. In addition, a copy number variant of the *CEL* gene (*CEL-HYB*) with only 3 VNTR repeats has shown to be a novel risk factor for chronic pancreatitis.

In this study, we wanted to further explore the role of the C-terminal domain of CEL in human pancreatic disease. We aimed to investigate the functional properties of not only previously characterized pathogenic variants, but also CEL variants identified in healthy controls. We also wanted to investigate whether a C-terminal epitope-tag could have an impact on our CEL protein studies.

In HEK293 cells, we found that the expression and secretion of CEL variants varied with VNTR length and composition. Both pathogenic and normal CEL variants were observed in lysate and in insoluble fractions of the cell. However, secretion of the disease-causing variants was reduced compared to secretion of the normal protein. When comparing CEL variants with and without tag, we observed a higher amount of untagged proteins intracellularly. Additionally, we found that several CEL variants induced apoptosis in HEK293 cells, but unexpectedly, some of these variants also led to increased cell viability.

The CEL VNTR is known to be heavily O-glycosylated, which is thought to be important for proper folding, secretion and stability of the protein. Here, we obtained evidence that the three deletion variants (DEL1, DEL4 and DEL8), all with the same theoretical molecular mass (73 kDa) but different VNTR composition, are likely to contain different patterns of O-glycosylation. Furthermore, the variants differed in secretion with only the pathogenic proteins (DEL1, DEL4 and HYB) showing impaired secretion. We also confirmed that the *pI* increased from CEL-WT to CEL-DEL1 by isoelectric focusing (*pI* 5.1 and 9.5 respectively).

In summary, we have shown that the cellular fate of CEL protein variants is influenced by the C-terminal of the protein. In addition, we found that epitope-tags most probably increase the solubility and stability of CEL proteins HEK293 cells, which may have led to an underestimation of the pathogenic effect of CEL-DEL1, CEL-DEL4 and CEL-HYB shown in previous studies.

# 1. Introduction

## 1.1. The pancreas

The human pancreas is an organ located in the upper part of the abdominal cavity, behind the stomach (Fig. 1.1 A). The organ has an elongated structure, it normally weighs 85-100 g, and can be divided into a head, body and tail region (Putte et al., 2013). The head of the pancreas is surrounded by the duodenum (the first part of the small intestine) while the tail is located next to the spleen (Fig. 1.1 B). The two main blood vessels *aorta* and *vena cava inferior* are located behind the pancreas, providing the organ with a rich blood supply (Holck, 2015). The pancreas comprises mainly of glandular tissue, which serves both exocrine and endocrine functions.

### 1.1.1. The exocrine pancreas

The exocrine part of the pancreas constitutes approximately 95 % of the organ, and consists of acini and the ductal system (Fig. 1.1 B). The acini are built of clusters of acinar cells, which produce and secrete digestive enzymes into the small ductal tracts. The ductal cells produce and secrete bicarbonate and water, which together with the digestive enzymes form the pancreatic juice. The latter is transported to the duodenum via the ductal system, and serves important roles in digestion of food (Putte et al., 2013). The production of juice is regulated through the hormones secretin and pancreozymin from the small intestine, but also through nerve impulses. As much as one litre of pancreatic juice can be produced per day (Holck, 2015).

The digestive enzymes are secreted as inactive forms (zymogens) to the intestine, to prevent self-digestion of the pancreas (Putte et al., 2013). The three main types of enzymes are proteases, lipases and amylases (Table 1.1). The acidic content of the stomach is partly neutralized by the alkaline bicarbonate, providing the appropriate pH for the digestive enzymes to work efficiently (Holck, 2015).

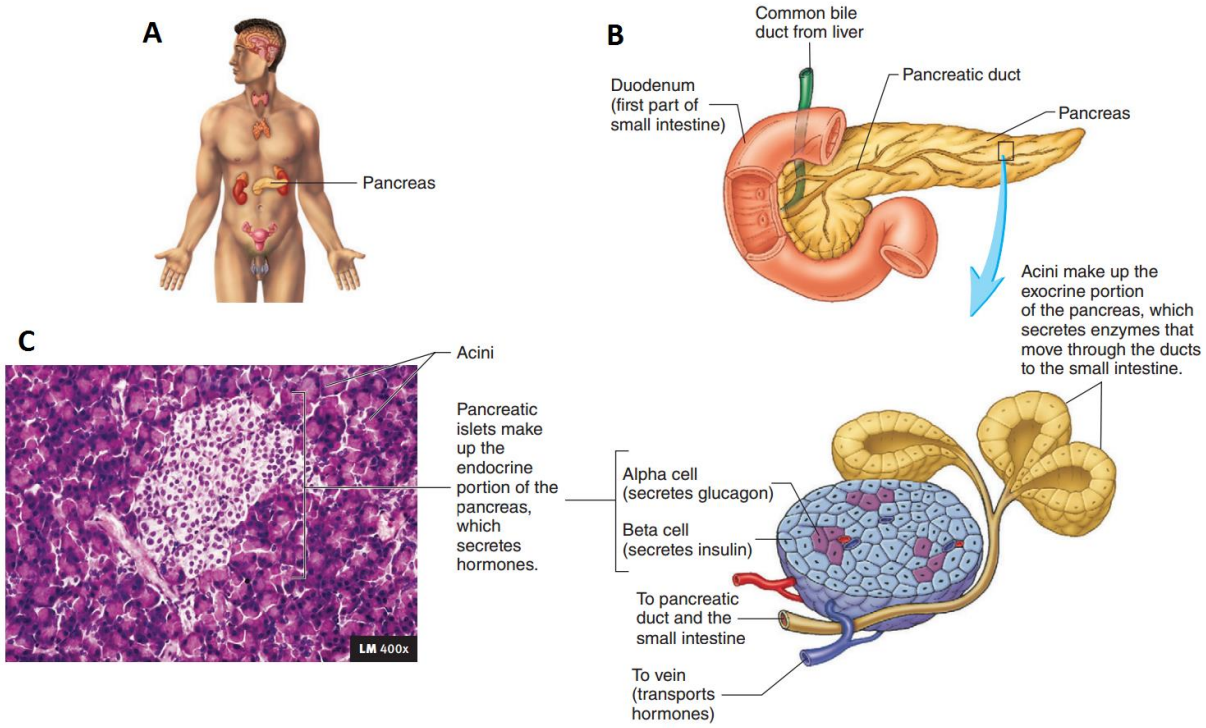



**Figure 1.1 Structure and histology of the human pancreas. A)** The pancreas (yellow) is located behind the stomach, in the upper part of the abdominal cavity and is surrounded by other organs including the small intestine, spleen and liver. **B)** The exocrine pancreas secretes digestive enzymes to the small intestine via the pancreatic ducts, whereas the alpha and beta cells of the islets of Langerhans produce metabolic hormones that are secreted to the circulatory system. **C)** Histological section of the human pancreas. (The illustrations are adapted from Seeley`s anatomy and physiology (Putte et al., 2013)).

### 1.1.2. The endocrine pancreas

The endocrine part of the pancreas constitutes only 1-2 % of the total organ. It consists of specialized cells located in the islets of Langerhans, all of which produce and secrete their specific metabolic hormones to the circulatory system (Pocock et al., 2013). The pancreatic islets are embedded in the exocrine tissue (Fig. 1.1 B, C). There are 500,000 - 1 million islets, each made up of mainly alpha (20%) and beta cells (75%) (Holck, 2015). The alpha cells produce and secrete glucagon, while the beta cells produce and secrete insulin (Putte et al., 2013). Insulin and glucagon have opposite effects, and the production of these hormones is regulated by the glucose levels in the blood. A rise in blood-glucose triggers production of insulin, which increases glucose-uptake in the tissues. This, in turn, promotes storage of glucose as glycogen in the liver, and thus the glucose level in the blood is reduced. Glucagon production is activated upon low levels of blood-glucose, which promotes breakdown of the glycogen storages (Holck, 2015).

The remaining endocrine tissue consists of delta-, gamma- and epsilon cells that produce somatostatin, pancreatic polypeptides and ghrelin, respectively (Table 1.1) (Putte et al., 2013).

**Table 1.1 Overview of secreted endocrine and exocrine products from the pancreas.**

| Secretion | Activated response |
|---|---|
| **Exocrine product** | |
| Carboxypeptidase A | Digestion of proteins |
| Carboxypeptidase B | Digestion of proteins |
| Carboxyl ester lipase (CEL) | Digestion of lipids |
| Chymotrypsin | Digestion of proteins |
| Colipase | Facilitates lipase function |
| Deoxyribonuclease | Digestion of DNA |
| Elastase | Digestion of proteins |
| $HCO_3^-$ | Neutralization of acidic chyme |
| Pancreatic α-amylase | Digestion of starch |
| Phospholipase $A_2$ | Digestion of lipids |
| Ribonuclease | Digestion of RNA |
| Triacylglycerol lipase | Digestion of lipids |
| Trypsin | Digestion of proteins |
| **Endocrine product** | |
| Ghrelin | Regulates appetite |
| Glucagon | Increases blood glucose levels and stimulates secretion of insulin and somatostatin |
| Insulin | Reduces blood glucose levels and inhibits glucagon secretion |
| Pancreatic polypeptide | Inhibits pancreatic bicarbonate- and protein secretion |
| Somatostatin | Inhibits secretion of insulin, glucagon and pancreatic polypeptides |

## 1.2. Diseases of the pancreas

The most common diseases of the pancreas are diabetes mellitus, inflammation of the pancreas (pancreatitis) and pancreatic cancer (Holck, 2015). Whereas diabetes affects the endocrine part of the gland, the other diseases are primarily affecting the exocrine tissue.

### 1.2.1. Diabetes mellitus

Diabetes mellitus, commonly referred to as diabetes, is a group of metabolic diseases in which the body fails to regulate blood glucose levels properly, rendering them chronically elevated. Diabetes is considered to be the fifth leading cause of deaths worldwide (Leslie et al., 2013).

Type 1 diabetes (T1D) represents between 5-10 % of all diabetes cases. T1D is an autoimmune disease, caused by the immune system attacking the beta cells (Leslie et al., 2013). This leads to reduced beta cell mass and reduced insulin production. As T1D is incurable, patients suffering from this disease require insulin replacements to regulate blood glucose levels (Leslie et al., 2013, Atkinson et al., 2014).

The most common type of diabetes is Type 2 diabetes (T2D), which accounts for approximately 90% of all diabetes cases. T2D is considered to be a multifactorial disease caused by a combination of genetic

susceptibility and an unhealthy lifestyle. It is highly associated with obesity (Leslie et al., 2013), and the prevalence of T2D also increases with age. Furthermore, T2D is often linked to risk factors such as high blood pressure, high levels of fatty acids in the blood, thrombosis and abdominal obesity. T2D is caused by impaired insulin production as well as impaired insulin sensitivity, so that tissues do not fully respond to the insulin which is produced. A proper diet and physical activity are the two main actions against this disease. However, various drugs to lower the blood glucose content are also given to patients (Leslie et al., 2013, Kahn and Davidson, 2014).

Other forms of diabetes include gestational diabetes and monogenic diabetes. Whereas T1D, T2D and gestational diabetes are complex diseases, monogenic diabetes is caused by mutations in one single gene (Murphy et al., 2008). Monogenic diabetes can be sub-divided into maturity onset diabetes of the young (MODY), neonatal diabetes and mitochondrial diabetes, where MODY is the most common form. Monogenic diabetes is a rare disease, accounting for approximately 2 % of all diabetes cases. There are more than 20 different genes that have been linked to this disease (Molven and Njolstad, 2011), however, patients suffering from monogenic diabetes are often misdiagnosed with T1D or T2D (Kavvoura and Owen, 2014). It has been reported that as many as 80% of MODY cases may be misclassified (Shields et al., 2010).

### 1.2.2. Pancreatitis

Pancreatitis is divided into an acute and a chronic form. Acute pancreatitis is characterized by fluid retention and swelling of the pancreas, caused by impaired drainage of pancreatic enzymes to the intestine. The disease mechanism is related to inappropriate activation of digestive enzymes in the pancreas, characterized by elevated levels of serum amylase and lipase. This can cause an inflammatory response leading to pancreatic tissue damage. The most common cause of acute pancreatitis is alcohol abuse and gallstones. Patients suffering from acute pancreatitis must receive medical care, and one way to treat them is to let the patients undergo fasting to prevent the pancreas from producing more pancreatic juice than what is needed (Whitcomb, 2006, Mounzer and Whitcomb, 2013, Aabakken, 2016).

If acute pancreatitis is not reversed, or if it recurs, it can progress into chronic pancreatitis. This condition is characterized by fibrosis, atrophy and further impaired function of the pancreas. As a result, food digestion will become increasingly affected, resulting in weight loss and diarrhoea. The patients often suffer from chronic pain, which can be difficult to treat. Some patients also develop diabetes as a complication. To prevent relapse of acute pancreatitis and subsequent development of chronic pancreatitis, the cause of the inflammation must be treated if possible. All patients must avoid alcohol, and they can be given perioral pancreatic enzyme replacement treatment. In very severe cases, surgical treatment is an option.

Pancreatitis is a complex disease and in recent years, genetics has also shown to play an important role in the disease development. In patients where environmental or metabolic factors do not explain the disease, genetic factors are considered particularly important (Whitcomb, 2013, Aabakken, 2016). Variants of *PRSS1* (cationic trypsinogen), *CFTR* (cystic fibrosis transmembrane conductance regulator), *SPINK1* (serine protease), *CTRC* (chymotrypsinogen C), *CPA1* (carboxypeptidase A1) and *CLDN2-MORC4* locus have been identified at risk factors (Whitcomb et al., 1996, Sharer et al., 1998, Witt et al., 2000, Le Marechal et al., 2006, Witt et al., 2006, Rosendahl et al., 2008, Whitcomb et al., 2012).

### 1.2.3. Pancreatic cancer

Pancreatic cancer can originate both from the endocrine and exocrine tissue of the organ. However, approximately 80 % of the tumours are adenocarcinomas, originating from the exocrine tissue. While the most common risk factor for pancreatic cancer is smoking, other risk factors are diabetes, chronic pancreatitis and genetics (Lowenfels et al., 1993, Bond-Smith et al., 2012).

## 1.3. Carboxyl ester lipase

Carboxyl ester lipase (CEL), also known as bile salt-stimulated lipase (BSSL) (Hernell and Olivecrona, 1974) or bile salt-dependent lipase (BSDL) (Abouakil and Lombardo, 1989), is a digestive enzyme mainly expressed in acinar cells of the pancreas and in lactating mammary glands as (Lombardo, 2001). From the mammary glands, the enzyme is secreted as part of the breast milk (Blackberg et al., 1981, Blackberg et al., 1987) and plays an important role in the digestion of fat in new-borns (Hui and Howles, 2002). Low levels of the enzyme has also been detected in fetal liver (Roudani et al., 1995), macrophages (Kodvawala et al., 2005), eosinophils (Holtsberg et al., 1995) and endothelial cells (Li and Hui, 1998, Lombardo, 2001).

CEL is one of three lipases secreted from the exocrine part of the pancreas and comprise about 4% of the total proteins detected in the pancreatic juice (Lombardo et al., 1978, Lombardo et al., 1980b). The protein is secreted to the digestive tract as an inactive enzyme. Once activated by bile-salts in the duodenum, the enzyme has wide substrate specificity and hydrolyses substrates such as cholesteryl esters, triacylglycerids, phospholipids, and fat-soluble vitamins (Lombardo et al., 1980b, Lombardo et al., 1980a, Lombardo and Guy, 1980, Blackberg et al., 1981, Hui and Howles, 2002, Kolar et al., 2016). It has been reported that CEL can be transported from the duodenum to the blood by transcytosis in intestinal cells (Bruneau et al., 2003). In the blood, CEL can further associate with low density lipoproteins (Caillol et al., 1997). As CEL is cleared from the circulation by renal glomerulus filtration, it is detected in the urine of healthy individuals (Comte et al., 2006).

### 1.3.1. The *CEL* gene

The human *CEL* gene is approximately 10 kb in size and consists of 11 exons. The gene is located on chromosome 9q34.3. The last exon of *CEL* includes a variable number of tandem repeats (VNTR) region (Fig. 1.2) (Taylor et al., 1991, Lidberg et al., 1992). These repeats consist of nearly identical 33 base pair segments, thus encoding 11 amino acids each. In the general population, the number of VNTR repeats varies from 3 to 23, and several studies have shown that the most common allele contains 16 repeats. (Lindquist et al., 2002, Higuchi et al., 2002, Raeder et al., 2006, Torsvik et al., 2010, Ragvin et al., 2013). This corresponds to a protein of 745 amino acids.

The human *CEL* locus also includes a nearby *CEL* pseudogene, a tandemly arranged gene designated *CELP* (Fig. 1.2). It has been proposed that *CEL* evolved from *CELP* through gene duplication, and that the original gene has lost its function (Lidmer et al., 1995, Madeyski et al., 1998, Lombardo, 2001, Vesterhus et al., 2010) Compared to *CEL*, *CELP* is missing exons 2-7 (Taylor et al., 1991, Lidberg et al., 1992). The remaining sequence however, shows 97 % similarity to *CEL* (Madeyski et al., 1998).

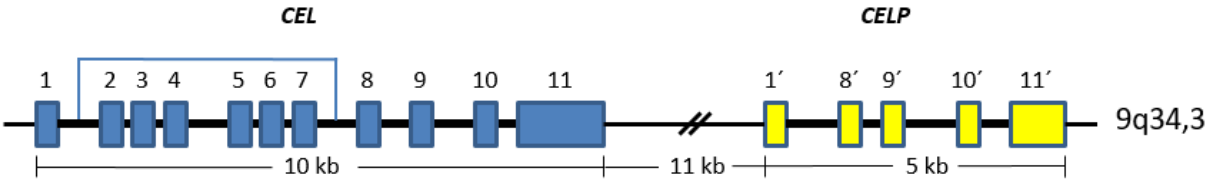



**Figure 1.2 Structure of the human CEL locus.** The *CEL* gene (blue) and the *CEL* pseudogene *CELP* (yellow), located on chromosome 9q34.3.

Furthermore, the *CEL* gene is highly conserved among vertebrates. However, while the gene is organized and positioned similarly with regard to introns and exons, it differs dramatically in the number of VNTR repeats. There are no repeats in fish, three in mouse, four in rat and as many as 39 in gorilla (Holmes and Cox, 2011). This is likely a result of internal duplications in the last exon during evolution, and indicates that the number of repeats is not critical for the function of the CEL enzyme (Kissel et al., 1989, Lidberg et al., 1992, Holmes and Cox, 2011).

### 1.3.2. CEL protein structure and secretion

CEL belongs to the α/β hydrolase fold family of enzymes, composed of a beta-sheet core surrounded by alpha-helices. The protein is composed of two structural domains; an N-terminal globular core domain with both the catalytic site and a signalling peptide, and a C-terminal proline-rich VNTR region (Reue et al., 1991, Terzyan et al., 2000, Hui and Howles, 2002). The bile salt-binding sites are located in exons 2, 3 and 10 (DiPersio et al., 1994, Aubert et al., 2002, Aubert-Jousset et al., 2004), as illustrated in Figure 1.3.

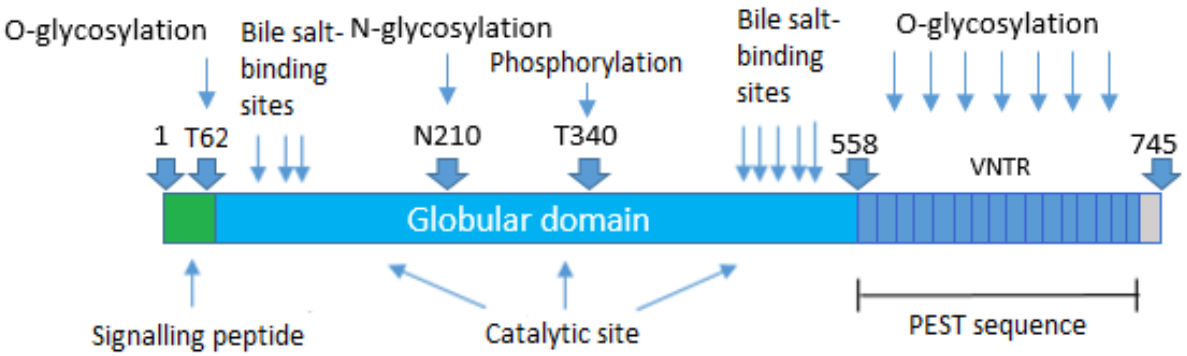



**Figure 1.3 Schematic overview of the CEL protein.** The protein is composed of two structural domains; the globular core domain with the catalytic site, and the C-terminal VNTR region. Also shown in this schematic figure is the N-terminal signalling peptide (green), bile salt-binding sites, site of phosphorylation, N- and O-glycosylation sites and the PEST sequence. The CEL protein presented here has 16 VNTR repeats, corresponding to the most common CEL variant in the general population.

CEL is secreted from the acinar cells and is thought to follow the classical pathway of secretory proteins (Palade, 1975, Lombardo, 2001). The enzyme is appropriately folded, assembled and undergoes post-translational modification during the secretion process from endoplasmic reticulum (ER). The protein is directed to the ER by its N-terminal signalling peptide. Here, asparagine at position 210 is N-glycosylated, a modification that is thought to be essential for the secretion and expression of CEL (Abouakil et al., 1993, Bruneau and Lombardo, 1995, Lombardo, 2001). Next, the enzyme is transported to the Golgi apparatus with the help of a multiprotein complex that contains the chaperone GRP94 (glucose-regulated protein with a size of 94 kDa) (Fig. 1.4 A) (Bruneau et al., 1995, Bruneau et al., 2000). In the Golgi, CEL is heavily O-glycosylated at several threonine and serine positions (Rogers et al., 1986, Wang et al., 1995, Bruneau et al., 1997). Actually, the most common CEL protein with 16 VNTR repeats is predicted to have a total of 36 O-glycosylation sites and, furthermore, 35 of these sites are located in the VNTR (NetOGlyc3.1). Once in the trans-Golgi area, CEL is released from the multiprotein folding complex upon phosphorylation, and further stored in zymogen granules (Fig.1.4 B) (Pasqualini et al., 2000, Verine et al., 2001). Incomplete or misfolded proteins are retained inside the

ER (Lodish et al., 2013). The N-terminal signalling sequence is cleaved from the protein during elongation, and is not present in the mature protein (Lodish et al., 2013).

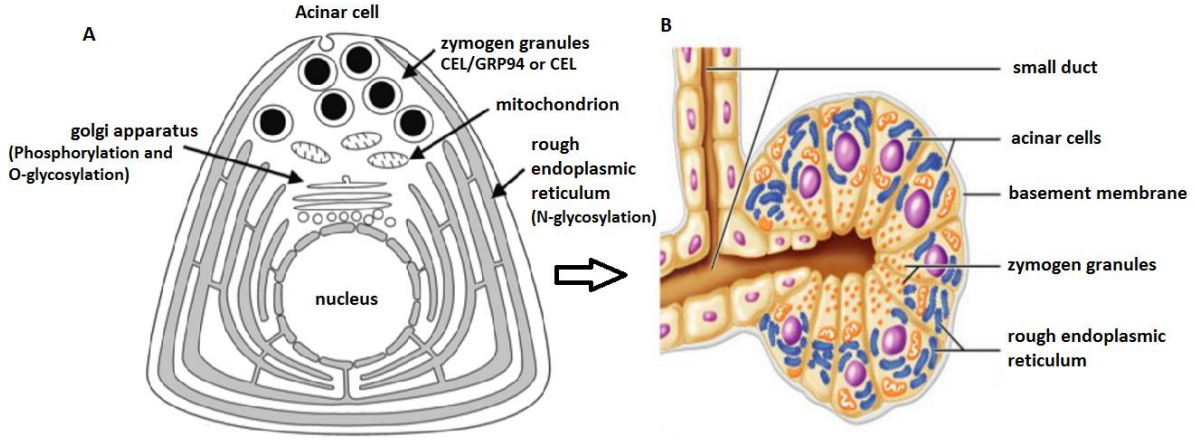



**Figure 1.4 Acinar cell secretion of CEL. A)** CEL is N-glycosylated in the ER lumen, and forms a complex with chaperone GRP94. The multiprotein complex is transported to the Golgi compartment where it is heavily O-glycosylated and phosphorylated. **B)** CEL is stored alone or in complex with GRP94 in zymogen granules until stimulation to be released into the small ducts. (Figures are adapted from, **A**: *Secretory Granule Exocytosis* (Burgoyne and Morgan, 2003), **B**: *Digestive System* from http://www.austincc.edu/apreview/PhysText/Digestive.html).

The O-glycosylation plays an important role for the integrity of the CEL protein. It has been reported that proteins without O-glycosylation are less secreted and rapidly degraded (Bruneau et al., 1997). Furthermore, the VNTR is enriched in regions of proline, glutamate, serine and threonine, known as PEST sequences, and it has been suggested that O-glycosylation has a protective role by masking the PEST sequence, thereby prohibiting degradation of the enzyme (Rogers et al., 1986, Loomes et al., 1999, Lombardo, 2001, Raeder et al., 2006). This modification may also serve to protect the enzyme from proteolytic degradation after secretion to the intestine (Loomes et al., 1999).

It is also possible that O-glycosylation is important for the solubility of the protein, by preventing non-specific intramolecular interactions between hydrophobic areas exposed on the globular core, thus protecting the protein from self-associating (Loomes and Senior, 1997). Surprisingly, a complete truncation of the VNTR does not compromise or interfere with the enzymatic activity of CEL (Downs et al., 1994), but it may affect its interaction with bile salt (Hui and Howles, 2002).

The enzyme activity of CEL is conferred by a catalytic amino acid triad of serine-histidine-aspartate, located within the globular domain of the enzyme (DiPersio et al., 1990). The enzyme activity is regulated by covering the active site with a loop structure, which associates with the protruding C-terminal by hydrogen binding. When bile salt is present, it will bind to the loop, thereby exposing the active site for substrate binding (Wang et al., 1997, Moore et al., 2001, Hui and Howles, 2002).

### 1.3.3. The polymorphic nature of *CEL*

Despite being conserved between species, the *CEL* gene has shown to be highly polymorphic. In addition to common single nucleotide polymorphisms (SNPs) that are present throughout the locus, and the VNTR length variation described above, mutations and copy number variants of the *CEL* locus have been reported (Raeder et al., 2006, McCarroll et al., 2008, Kidd et al., 2008, Torsvik et al., 2010, Fjeld et al., 2015). Some of these variants are described below and their protein structures illustrated in Figure 1.5. The disease-causing variants are further described in section 1.4.

A single-base deletion in the first repeat of the VNTR (*CEL-DEL1*) leads to diabetes and pancreatic exocrine dysfunction (Raeder et al., 2006). Another single-base deletion in the fourth repeat (*CEL-DEL4*) has a similar but milder phenotype (Raeder et al., 2006). Both variants lead to frameshifts resulting in premature stop codons, and thus truncated protein variants with new C-terminal ends (Raeder et al., 2007). However, not all *CEL* deletions are pathogenic. A single-base deletion in the eighth repeat of the VNTR (*CEL-DEL8*) has been observed in the normal population.

Single-base insertions are also commonly found in the *CEL* gene. These insertions result in truncated proteins, but has not yet been associated with disease (Raeder et al., 2006). One example is *CEL-INS9*, which has a single-base insertion in repeat 9. This leads to a frameshift and premature stop codon, also in repeat 9.

During a study on suspected MODY families, one Danish patient with a *CEL* allele containing only three VNTR repeats was identified (*CEL-3R*) (Torsvik et al., 2010).

*CEL-HYB* is a *CEL* deletion variant that recently was identified as a novel risk factor for chronic pancreatitis. *CEL-HYB* is formed by a fusion between *CEL* and *CELP*, and includes a VNTR from the pseudogene that only contains three repeats (Fjeld et al., 2015, Molven et al., 2016a).

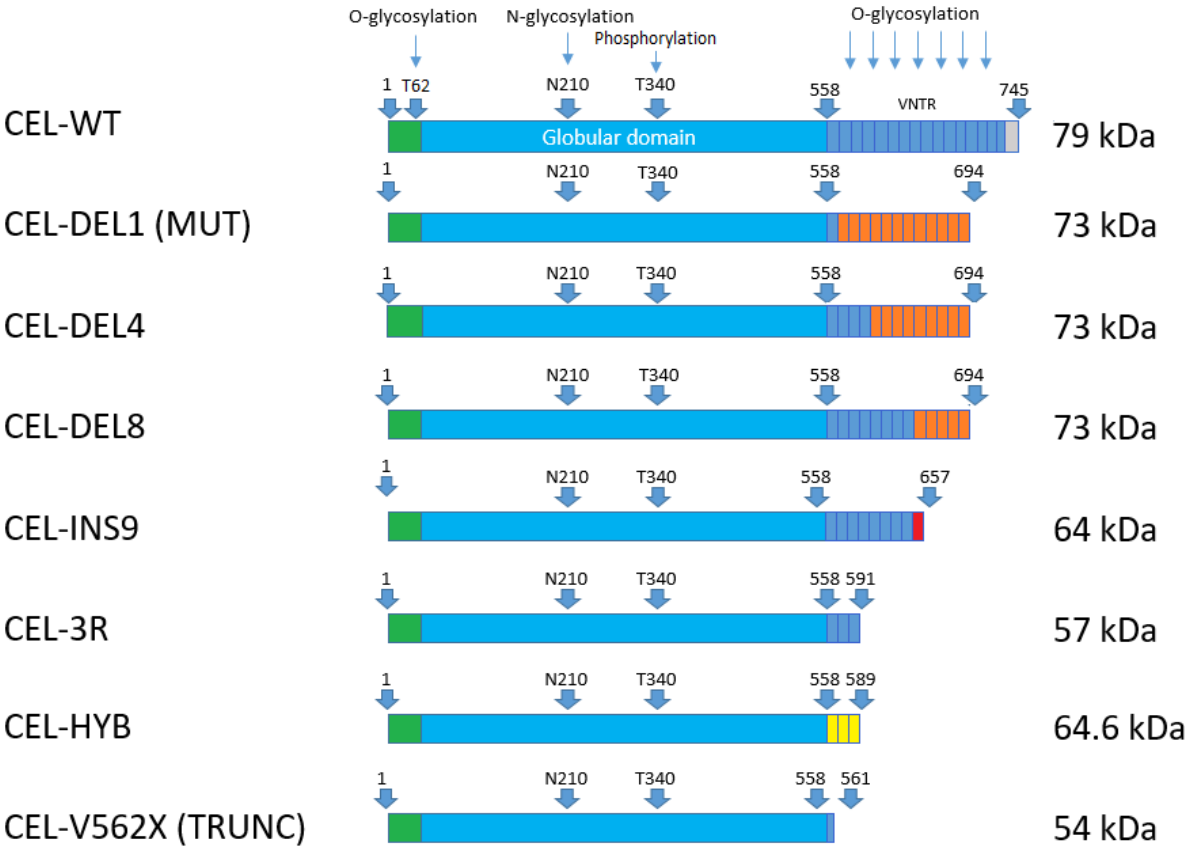



**Figure 1.5 Schematic overview of CEL protein variants.** The illustration shows several CEL variants, their theoretical size and the length/composition of their VNTR domains. Blue boxes of the deletion variants illustrate the normal repeats, while orange boxes illustrate the downstream repeats as result of the frame-shifts. The red box in CEL-INS9 illustrates the downstream protein segment as result of a new reading frame. The yellow boxes in CEL-HYB illustrates the three VNTR repeats originating from *CELP*.

All CEL variants shown in Fig. 1.5 have different amino acid compositions compared to the wild-type protein. All variants have also been observed in humans, except for CEL-TRUNC. The CEL-TRUNC variant was artificially constructed to analyse a CEL protein completely lacking the VNTR. This variant consists only of the first four amino acids of the first repeat (Torsvik et al., 2014).

## 1.4. Pathogenic variants of CEL

### 1.4.1. CEL-DEL1 (CEL-MODY)

CEL-MODY follows an autosomal dominant pattern of inheritance and age of onset is typically childhood to young adulthood. The disease is characterized by pancreatic exocrine dysfunction and beta cell dysfunction (Raeder et al., 2006, Murphy et al., 2008, Torsvik et al., 2014). Patients suffering from CEL-MODY develop morphological abnormalities of the pancreas, such as lipomatosis, at an early age. These abnormalities may occur before the disease is noted on a clinical level (Raeder et al., 2006, Raeder et al., 2007, Bjorlykke et al., 2015). At later stages, the patients develop monogenic diabetes due to beta cell failure and pancreatic cyst formation (Raeder et al., 2014 ). Since CEL-MODY is a disease developing over time, studies on biomarkers can provide important information for prediction and treatment of the disease in the future (Raeder et al., 2007, Bjorlykke et al., 2015).

CEL-MODY is caused by a single-base deletion in the VNTR region of the *CEL* gene. Such mutations were identified in two independent families with two separate single-base deletions in the first and the fourth repeat, respectively. As described before, these single-base deletions lead to frameshifts and premature stop codons in repeat 13, resulting in two truncated proteins denoted CEL-DEL1 and CEL-DEL4 (Fig. 1.6) (Raeder et al., 2006).

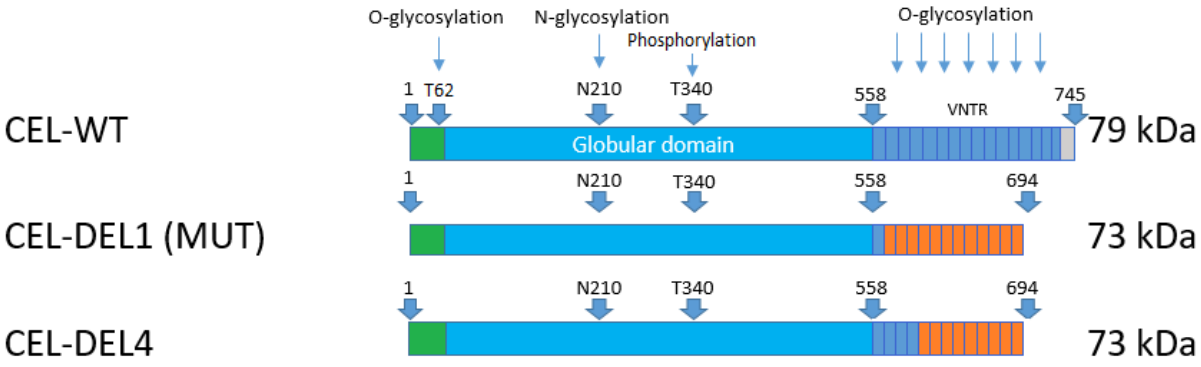



**Figure 1.6 Overview of pathogenic MODY variants.** The pathogenic variants CEL-DEL1 and CEL-DEL4 have the same globular domain as CEL-WT, but different VNTRs. The blue boxes illustrate normal VNTR repeats, while orange boxes illustrate the downstream repeats as a result of the frame-shifts.

Compared to the normal protein with a theoretical size of 79 kDa, both CEL-DEL1 and CEL-DEL4 are reduced in size (73 kDa) (Fig. 1.6). Furthermore, there is a significant increase in the *pI* of CEL-DEL1 compared to CEL-WT: *pI* = 9.5 vs *pI* = 5.2, respectively (Johansson et al., 2011). In a study performed by our group, CEL-DEL1 exhibited similar glycosylation, ubiquitination and secretion as CEL-WT, but displayed a high propensity to form both intra- and extracellular aggregates (Johansson et al., 2011). We have also found that after CEL-DEL1 secretion, there is a high degree of cellular reuptake and further

lysosomal degradation. Endocytosis of CEL-DEL1 resulted in reduced cell viability of both pancreatic acinar and beta cells, (Torsvik et al., 2014). Thus, based on these results, we have proposed that CEL-MODY is a protein-misfolding disease due to a negative gain-of-function effect (Johansson et al., 2011).

Interestingly, in a recent report they found that CEL-DEL1 causes ER stress, induction of the unfolded protein response (UPR), activation of NF-κB, and subsequent apoptosis (Xiao et al., 2016). In that study, the author suggests that intracellular accumulation of CEL-DEL1 overwhelms the cells` ER responsive stress mechanisms, and that apoptosis serves as a protection for the surrounding cells without activating further inflammation (Xiao et al., 2016).

### 1.4.2. CEL-HYB

Several copy number variants of *CEL* have been detected by our research group and others (McCarroll et al., 2008, Torsvik et al., 2010, Ragvin et al., 2013, Fjeld et al., 2015). However, we were the first to report that a hybrid variant of the *CEL* gene is a risk factor for chronic pancreatitis (Fjeld et al., 2015). The *CEL* hybrid is a deletion variant, which most likely has originated from non-allelic homologous recombination between *CEL* and its nearby pseudogene *CELP*, due to the high sequence similarity between the two genes (Fjeld et al., 2015). A proposed model of the recombination event is presented in Figure 1.7.

Compared with healthy blood donor controls, we found that the carrier frequency of *CEL-HYB* was increased by >5-fold in patients with idiopathic CP from Germany and France. The *CEL-HYB* allele was also enriched in alcoholic chronic pancreatitis patients (Fjeld et al., 2015).

The *CEL-HYB* gene includes only three VNTR repeats that originates from the *CEL* pseudogene (Fig. 1.7 A, B). When expressed in cellular models, the CEL-HYB protein showed impaired secretion, intracellular accumulation and induced autophagy (Fjeld et al., 2015). The catalytic activity of CEL-HYB was also reduced to about 40 % of the normal CEL enzyme activity. In contrast, the artificial CEL variant lacking the entire VNTR region (CEL-TRUNC) showed 90 % of wild-type enzyme activity. This implies that the reduction of CEL-HYB activity is due to an altered amino acid composition, rather than a truncated VNTR, and that increased enzyme activity is not part of the disease mechanism (Fjeld et al., 2015).

CEL-HYB represents a new disease mechanism for chronic pancreatitis, as most pancreatitis genes are linked to the protease-antiprotease system of the pancreas (Ravi Kanth and Nageshwar Reddy, 2014).

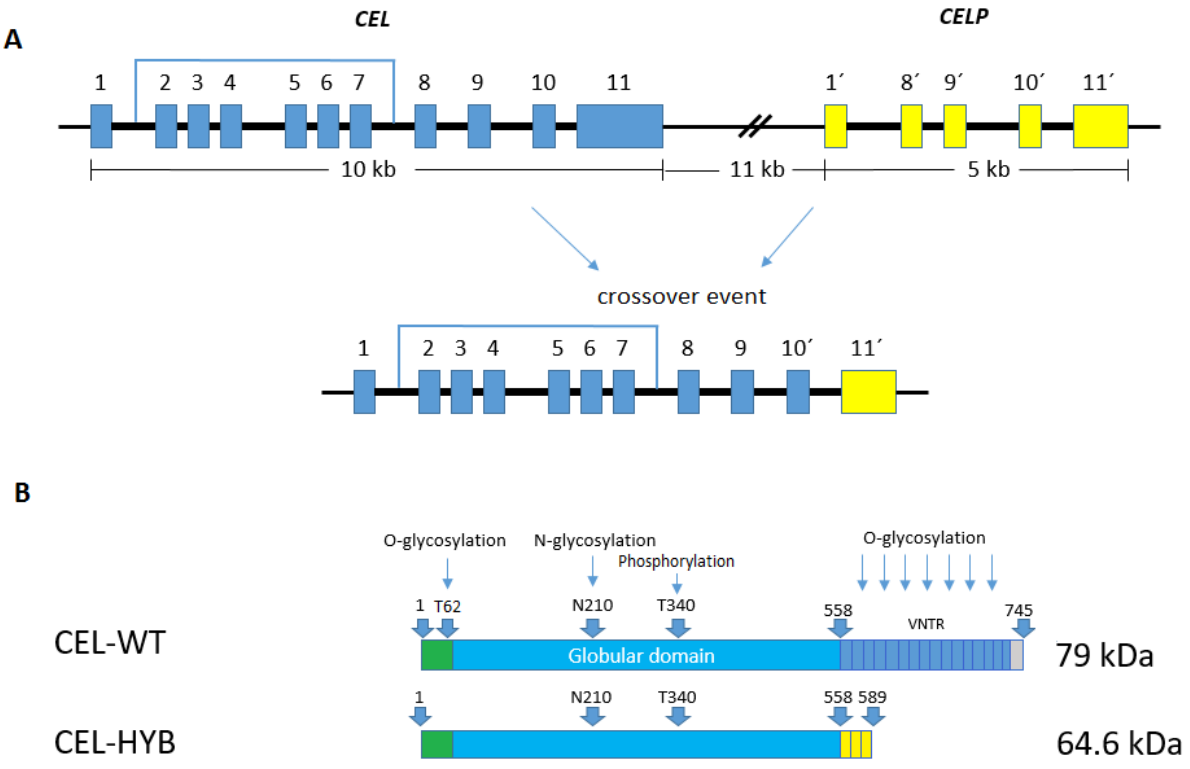



**Figure 1.7 Schematic overview of CEL-HYB. A)** A proposed model for how the *CEL-HYB* gene variant originated. Non-allelic homologous recombination has taken place between *CEL* and *CELP* and the recombination region includes intron 10 as well as adjacent exon boundaries. **B)** Overview of CEL-HYB protein compared to CEL-WT protein. CEL-HYB has a truncated VNTR with only three repeats, originating from *CELP*.

When comparing the disease mechanisms of CEL-HYB and CEL-MODY, they are different, as also indicated by their roles in pancreatic disease. CEL-MODY causes a dominantly inherited syndrome of exocrine and endocrine pancreatic dysfunction affecting all carriers with the mutation (Raeder et al., 2006). In contrast, CEL-HYB is a risk factor for the development of chronic pancreatitis, although most carriers in the general population will remain unaffected (Fjeld et al., 2015, Molven et al., 2016a).

### 1.4.3. CEL VNTR length polymorphisms in disease

CEL VNTR length polymorphism has been associated with cholesterol profile of serum (Bengtsson-Ellmark et al., 2004), rate of HIV-1 disease progression (Stax et al., 2012) and risk for alcohol-induced pancreatitis (Miyasaka et al., 2005). The latter association, however, could not be replicated in more recent and larger studies (Ragvin et al., 2013, Fjeld et al., 2016). Actually, one study is suggesting that CEL VNTR lengths are associated with alcoholic liver cirrhosis, and not alcohol induced pancreatitis (Fjeld et al., 2016). Also, one study set out to analyse the association between VNTR lengths and pancreatic cancer, but no association was found (Dalva et al., 2016).

Given the unusually polymorphic nature of CEL, several unknown variants are likely to exist, and some of them possibly disease-causing. Also, it could be that some disease-associated *CEL* alleles are confined to certain populations (Molven et al., 2016b, Zou et al., 2016).

## 1.5. The use of epitope tags in experimental research

Epitope tags are easily detectable peptide sequences that can be introduced to proteins, simplifying experimental studies and protein purification (Hengen, 1995). The use of epitope tags can be exploited in immunoprecipitation, Western blot and immunocytochemistry. They are commonly used in studies of physical and functional properties of proteins, as well as their intracellular localisation. A protein is fused with an epitope tag by modifying the corresponding endogenous gene by recombinant DNA techniques (Kim et al., 2008).

An epitope-containing protein is often referred to as a fusion protein, which refers to the end-to-end joining between the target protein and tag. Epitope tags can be fused with either the N- or C-terminal end of the protein, but can sometimes also be placed in the middle of the protein. Epitope tags are often divided into different classes, depending on their functions. They can roughly be divided into affinity and solubility tags. While affinity tags allow rapid and efficient purification, solubility tags enhance the solubility of a protein, and also proper folding (Malhotra, 2009).

Several vectors containing epitope tag sequences exist, which can be fused with almost any protein. These tags range from just a few amino acids, to whole proteins or domains (Martinez-Ceron et al., 2012). Some common short tags are: poly-His, poly-Arg, FLAG, c-myc, and strep-tag, while some common long tags are Glutathione S-transferase (GST) and Maltose-binding protein (MBP) (Table 1.2). In the case of short fusion tags, it is for most studies not needed to remove the tag after purification, as it will in most cases not affect the structure or function of the protein (Kosobokova et al., 2016). Proteins fused to large affinity tags such as Glutathione S-transferase and Maltose-binding proteins (Table 1.2)

can easily be purified from their biological source with affinity chromatography. These large tags do however need to be removed by proteases after purification (Terpe, 2003).

Polypeptides consisting of several histidine residues are often used as epitope tags. They are small and have little effect on the charge or folding properties in the cell, or the structure and function of the mature protein (Mason et al., 2002, Carson et al., 2007, Li et al., 2013, Pajęcka et al., 2014). The 6xHis-tag has been constructed as an affinity tag. A recent study found that a 6xHis tag increased the expression level of their target protein compared to the untagged variant (Kosobokova et al., 2016). Another study showed that the fusion of an N-terminal epitope tag could alter the intracellular trafficking of human BST-2 (Lv et al., 2014). Furthermore, a study found that a His tag had a small negative effect on enzyme activity, and also an impact on the secondary structure of the protein (Chen et al., 2015).

An ideal epitope tag should preferably be large and hydrophilic, and should, for easy detection, have an antibody recognition site in an area of the protein which is exposed. This often complicates the choice of epitope tag; as large epitope tags can interfere with the structure of the protein. In addition, exposed areas on the protein surface are often involved in the function of the protein (Sigma-Aldrich.).

**Table 1.2 Overview of common tags used in experimental procedures.** Table is adapted from; *Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems* (Terpe, 2003).

| Tag | Sequence | Size (kDa) |
|---|---|---|
| Poly-Arg | RRRRR | 0.80 |
| Poly-His | HHHHHH | 0.84 |
| FLAG | DYKDDDDK | 1.01 |
| Strep-tag II | WSHPQFEK | 1.06 |
| c-myc | EQKLISEEDL | 1.20 |
| V5 | GKPIPNPLLGLDST | 1.42 |
| S- | KETAAAKFERQHMDS | 1.75 |
| HAT- | KDHLIHNVHKEFHAHANK | 2.31 |
| 3x FLAG | DYKDHDGDYKDHDIDYKDDDDK | 2.73 |
| Calmodulin-binding peptide | KRRWKKNFIAVSAANRFKKISSSGAL | 2.96 |
| Glutathione S-transferase | Protein | 26 |
| Maltose-binding protein | Protein | 40 |

# 2. Aims of the study

The overall objective of this study was to understand more of the role of the CEL protein C-terminal in pancreatic disease.

The specific aims of the project were:

1. To investigate whether the presence of a C-terminal epitope-tag influences CEL protein expression, secretion and intracellular localisation in HEK293 cells
2. To compare the expression, secretion and intracellular localisation of CEL variants differing in VNTR-length
3. To investigate the effect of these CEL variants on cell viability and apoptosis in HEK293 cells
4. To investigate the extent of putative O-glycosylation sites on CEL variants with varying VNTR-lengths
5. To compare the predicted changes in *pI* with the actual charge of CEL-WT and CEL-DEL1

# 3. Materials

**Table 3.1. DNA techniques**

| Material | Catalog number | Supplier |
|---|---|---|
| Ampicillin | A9518-5G | Sigma-Aldrich |
| *E.coli* OneShot® TOP10 chemically competent cells | C404010-Aug2015 | Invitrogen |
| Ethidium bromide (10 mg/ml) | 161-0433 | Life Technologies |
| Gel loading dye blue | 70215S | BioLabs |
| Lysogeny broth (LB) agar-ampicillin | Q60120 | Invitrogen |
| Lysogeny broth (LB) medium | L7275-500TAB | Sigma-Aldrich |
| Plasmid midi Kit (100) | 12245 | Qiagen |
| SeaKem® LE Agarose | 500004 | Lonz |
| Super optimal broth with catabolite repression (SOC) medium | 15544-034 | Invitrogen |
| 1 Kb DNA Ladder | N3232S | BioLabs |

**Table 3.12. Plasmids**

| Plasmids | Description | Supplier |
|---|---|---|
| pcDNA 3 | CEL-DEL1 | Gift from Prof M.Lowe, Children`s hospital of Pittsburgh, USA |
| pcDNA 3 | CEL-DEL4 | Gift from Prof M.Lowe, Children`s hospital of Pittsburgh, USA |
| pcDNA 3 | CEL-DEL8 | Gift from Prof M.Lowe, Children`s hospital of Pittsburgh, USA |
| pcDNA 3 | Empty vector | Addgene |
| pcDNA 3 | CEL HYB | Made in Bergen |
| pcDNA 3 | CEL-INS9 | Gift from Prof M.Lowe, Children`s hospital of Pittsburgh, USA |
| pcDNA 3 | CEL-V562X | Gift from Prof M.Lowe, Children`s hospital of Pittsburgh, USA |
| pcDNA 3 | CEL-WT | Gift from Prof M.Lowe, Children`s hospital of Pittsburgh, USA |
| pcDNA 3 | CEL-3R | Gift from Prof M.Lowe, Children`s hospital of Pittsburgh, USA |
| pcDNA 3.1/V5-His B | CEL-DEL1 | Made in Bergen |
| pcDNA 3.1/V5-His B | CEL-HYB | Made in Bergen |
| pcDNA 3.1/V5-His B | CEL-TRUNC | Made in Bergen |
| pcDNA 3.1/V5-His B | CEL-WT | Made in Bergen |
| pcDNA 3.1/V5-His B | Empty vector | Invitrogen |

**Table 3.3. Cell lines**

| Cell line | Description | Catalog number | Supplier |
|---|---|---|---|
| HEK293 | Human embryonic kidney cells | 632180 | Clontech |

**Table 3.4. Reagents for cell culturing**

| Material | Catalog number | Suppier |
|---|---|---|
| Antibiotic Antimycotic | 15240062 | Invitrogen |
| Dimethyl sulfoxide (DMSO) | D8418-250ml | Sigma-Aldrich |
| Dulbecco`s Modified Eagle`s Medium (DMEM) | 41966-029 | Gibco by Life Technologies |
| Fetal Bovine Serum (FBS) | F7524-500ml | Sigma-Aldrich |
| Lipofectamine 2000 Transfection reagent | 11668-019 | Invitrogen |
| OPTI-MEM (1x) Reduced Serum Medium | 31985-062-100ml | Gibco by Life Technologies |
| Dulbecco`s Phosphate Buffered Saline (PBS) | D8537-500ml | Sigma-Aldrich |
| Poly-L-Lysin | P4832-50ml | Sigma-Aldrich |
| 0.05% Trypsin-EDTA | 25300-054-100ml | Gibco by Life Technologies |

**Table 3.5. Antibodies**

| Antibody | Catalog number | Supplier | Classification |
|---|---|---|---|
| Anti-CEL antibody (rabbit polyclonal) | Gift | Gift from Prof M.Lowe, Children`s hospital of Pittsburgh, USA | Primary |
| As20 anti-CEL antibody (mouse monoclonal) | Gift | Gift from Prof O.Hernell, University of Umeå, Sweden | Primary |
| GAPDH antibody (V18) (goat polyclonal) | sc-20357 | Santa Cruz Biotechnologies | Primary |
| Vanko anti-CEL antibody (rabbit polyclonal) | Gift | Gift from Prof O.Hernell, University of Umeå, Sweden | Primary |
| V5 epitope tag antibody (mouse monoclonal) | R960-25 | Invitrogen | Primary |
| Donkey anti-goat IgG-HRP | sc-2020 | Santa Cruz Biotechnologies | Secondary |
| Donkey anti-mouse IgG-HRP | sc-2318 | Santa Cruz Biotechnologies | Secondary |
| Donkey anti-rabbit IgG-HRP | sc-2305 | Santa Cruz Biotechnologies | Secondary |
| F(ab`)2 Goat anti-mouse IgG (H+L) secondary antibody Alexa Flour @ 488 conjugate | A11017 | Invitrogen | Secondary |
| F(ab`)2 Goat anti-rabbit IgG (H+L) secondary antibody Alexa Flour @ 594 conjugate | A11032 | Invitrogen | Secondary |

**Table 3.6. Cell lysis, SDS-PAGE and Immunoblotting – products**

| Materials | Catalog number | Supplier |
|---|---|---|
| Amersham™ ECL™ prime western blotting detection reagent | RPN2232 | GE Healthcare |
| Blocking grade blocker nonfat-dry milk | 170-6404 | BioRad |
| BSA protein assay kit | 23227 | Pierce |
| Complete™ Mini EDTA-free protease inhibitors cocktail tablets | 11836170001 | Roche Diagnostics |
| Falcon™ Polystyrene Microplates (6 well plate) | 08-772-29 | Fisher-Scientific |
| Immobilion-P membrane, PVDF, 0.45 µm | IPVH07850 | Merck Millipore |
| Magic Mark XP western protein standard | LC5603 | Invitrogen |
| Methanol | 67-56-1 | Merck Millipore |
| NuPAGE® LDS sample buffer (4x) | NP0007 | Invitrogen |
| NuPAGE® MOPS SDS running buffer (20x) | NP0001-02 | Invitrogen |
| NuPAGE® Novex 10% Bis-Tris protein gels 1.0 mm | NP0301BOX | Invitrogen |
| NuPAGE® Sample reducing agent (10x) | NP0009 | Invitrogen |
| NuPAGE® Transfer buffer (20x) | NP0006-1 | Thermo Scientific |
| Phosphate buffer saline (PBS) Tablets | 18912-014 | Gibco |
| Pierce™ BCA protein assay kit | 23227 | Thermo Scientific |
| Precision Plus Protein™ dual colour standard | 161-0374 | BioRad |
| Restore™ Western blot stripping buffer | 21059 | Thermo Scientific |
| RIPA lysis buffer (10x) | 20-188 | Merck Millipore |
| Tween®20 | 8.22184.0500 | Merck Millipore |

**Table 3.7. Immunostaining**

| Material | Catalog number | Supplier |
|---|---|---|
| Bovine serum albumin (BSA) | A4503-100G | Sigma-Aldrich |
| Cover slips (18mm) | 117580 | Werderop Werd |
| Falcon™ Polystyrene Microplates (12 well plate) | 08-772-29 | Fisher-Scientific |
| Gold antifade mounting reagent with DAPI | P36935 | Molecular Probes |
| Superfrost microscope slides | Z692255-100EA | Sigma-Aldrich |
| Normal goat serum | 1000C | Invitrogen |
| Paraformaldehyde (PFA) | 818715 | Merck Millipore |
| Phosphate buffered saline (PBS) | D8537 | Sigma-Aldrich |
| Saponin | 47036-50G-F | Sigma-Aldrich |

**Table 3.8. Mutagenesis**

| Material | Catalog number | Supplier |
|---|---|---|
| *DpnI* (10 U/µl) | 200518-52 | Agilent Technologies |
| PfuUltra HF DNA polymerase | 200524-51 | Agilent Technologies |
| QuikChange XL dNTP Mix | 200516-52 | Agilent Technologies |
| QuickSolution | 200516-51 | Agilent Technologies |
| XL 10-Gold Ultracompetent cells | 200516-4-Aug2016 | Agilent Technologies |
| XL 10-Gold 2-Mercaptoethanol | 200314-43 | Agilent Technologies |
| 10X Reaction Buffer | 200518-50 | Agilent Technologies |

**Table 3.9. Cell viability and apoptosis assay**

| Material | Use | Catalog number | Supplier |
|---|---|---|---|
| CellEvent Caspase-3/7 Green Detection Reagent | Apoptosis assay | C1042 | Invitrogen |
| Cover slips (18mm) | Apoptosis assay | 117580 | Werderop Werd |
| Falcon™ Polystyrene Microplates (12 well plate) | Apoptosis assay | 08-772-29 | Fisher-Scientific |
| Superfrost microscope slides | Apoptosis assay | Z692255-100EA | Sigma-Aldrich |
| CellTitre-Glo® Buffer | Cell viability assay | G756A | Promega |
| CellTitre-Glo® Substrate | Cell viability assay | G755A | Promega |
| 96-Well Clear Bottom Plates | Cell viability assay | 29444-010 | Corning® |

**Table 3.10. De-glycosylation assay**

| Material | Catalog number | Supplier |
|---|---|---|
| Fisher chemical acetone | 11369773 | Fisher Scientific |
| O-Glycosidase (25 mU) | 11365185001 | Sigma-Aldrich |
| N-Glycosidase F (100 U) | 11347101001 | Sigma-Aldrich |
| Neuramidase (Sialidase) (1 U) | 11080725001 | Sigma-Aldrich |

**Table 3.11. Isoelectric focusing**

| Material | Catalog number | Supplier |
|---|---|---|
| Novex® IEF Sample Buffer pH 3-10 (2X) | LC5311 | ThermoFisher Scientific |
| Novex™ pH 3-10 IEF Protein Gels, 1.0 mm, 10-well | EC6655BOX | ThermoFisher Scientific |
| Novex® IEF Anode Buffer (50X) | LC5300 | ThermoFisher Scientific |
| Novex® IEF Cathode Buffer pH 3-10 (10X) | LC5310 | ThermoFisher Scientific |

**Table 3.12. Buffers and solutions**

| Buffers and solutions | Use | Composition |
|---|---|---|
| 1x RIPA lysis buffer | Cell lysis | 10x RIPA lysis buffer (0.5M Tris-HCl, pH 7.4, 1.5M NaCl, 2.5% deoxycholic acid, 10% NP-40, 10mM EDTA.) in 10 mlddH2O + 1 tablett Complete™ Mini EDTA-free protease inhibitors cocktail tablet |
| Agarose gel | Gel electrophoresis | 1% (w/v) Seakem/agarose dissolved in 1xTBE buffer pH 8.3 with EtBr (10 mg/ml) |
| TBE buffer (pH 8.3) | Gel electrophoresis | Tris-borate (45 mM)/EDTA (1 mM) |
| NuPAGE® transfer buffer (1x) | Immunoblotting | NuPAGE® transfer buffer (20x) in $ddH_2O$ and 10 % Methanol |
| PBS-tween (0.05 %) | Immunoblotting | Phosphate buffered saline tablet (pH 7.45) and Tween® 20 dissolved in $ddH_2O$ |
| Blocking buffer | Immunofluorescence | 0.2% BSA, 0.2% Saponin, 5% normal goat serum in |
| Fixative | Immunofluorescence | 6% PFA + 0.2 M PB pH 7.2 (1:1) |
| Paraformaldehyde (PFA) solution | Immunofluorescence | 6% (w/v) PFA solution, 3 g PFA dissolved in 50 ml $ddH_2O$ + 5 drops 1 M NaOH |
| Quenching solution | Immunofluorescence | 50 mM $NH_4Cl$ in PBS |
| Washing buffer | Immunofluorescence | 0.2% BSA in PBS |
| Washing buffer saponine | Immunofluorescence | 0.2% BSA and 0.2% Saponin in PBS |
| 0.2 M Phosphate buffer (PB) | Immunofluorescence | 0.2 M $Na_2HPO_4$ and 0.2 M $NaH_2PO_4$ adjusted to pH 7.2 |
| Denaturation buffer | Isoelectric focusing | 20 mM phosphate buffer, 1% (w/v) SDS, 1% (v/v) beta-mercaptoethanol in $ddH_2O$ |
| Incubation buffer | Isoelectric focusing | 20mM EDTA, 100mM phosphate buffer, 0.5% (v/v) Nonidet P-40, 1% (v/v) beta-mercaptoethanol in ddH2O |
| NuPAGE® MOPS SDS running buffer (1x) | SDS-PAGE | NuPAGE® MOPS buffer (20x, 50 mM MOPS, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.7) in $ddH_2O$ |
| NuPAGE® Transfer buffer (1x) | SDS-PAGE | 50 ml NuPAGE® Transfer buffer (20x) in 950 ml $ddH_2C$ |

**Table 3.13. Technical equipment**

| Instrument | Manufacturer |
|---|---|
| Eppendorf Centrifuge 5417C | A/B Phil |
| Heragus Multifuge 3S-R | Thermo Electron Corporation |
| Infinite M200Pro luminomenter | Tecan |
| Innova 4300 Incubation shaker Dipling HOUM | New Brunswick Scientific |
| LAS-1000 imager | Fujifilm |
| Leica Confocal SP5 | Leica Microsystems |
| Megafuge 1.0 R | Heraeus Seoatech |
| Molecular Imager® Gel Doc™ XR system | BioRad |
| NanoDrop ND-1000 | Saveen Werner |
| Scepter handheld automated cell counter | Merck Millipore |

**Table 3.14. Analytical Software**

| Software | Supplier |
|---|---|
| Image Reader LAS-100 Pro v2.6 | Fujifilm |
| Leica Application Suite v2.0 | Leica Microsystems |
| Microsoft Office 365 | Microsoft Corporation |
| Photoshop | Adobe Photoshop CC |
| QuantityOne | Biorad |
| Infinite® M200 PRO | Tecan |

# 4. Methods

## 4.1. Plasmid preparation and evaluation of DNA quality

### 4.1.1. Transformation of OneShot® TOP10 chemically competent *E.coli* cells

Transformation was performed by applying the heat shock method as described by the manufacturer's instructions (Invitrogen). cDNAs encoding the CEL variants to be analysed in this study, were cloned into either a pcDNA3 (Addgene) or pcDNA3.1/V5-His (Invitrogen) vector backbone. In short, 1 µl (300-1000 ng) of plasmid was added to a vial (50 µl) of OneShot® TOP10 chemically competent *E.coli* cells. After incubation on ice for 30 min, the cells were subjected to heat shock (42°C) for 30 s and directly transferred to ice. SOC medium (250 µl) was added to each vial, and incubated 1 h with shaking (200 rpm at 37°C). The cells were spread onto agar plates containing ampicillin (100 µg/ml), and incubated o/n at 37°C.

### 4.1.2. Plasmid preparation

A single colony was inoculated in starting cultures with LB medium (5 ml) containing ampicillin (100 µg/ml). The cultures were incubated for 8 h at 37°C with shaking at 250 rpm. Next, starting cultures were diluted according to the protocol for high copy plasmids; 1:500 in LB medium containing 100 µg/ml ampicillin and grown o/n. Plasmid purification was performed according to the Qiagen® Plasmid Midi Kit (100) protocol.

### 4.1.3. Determination of plasmid yield and quality

The concentration and quality of the purified plasmids were determined by OD measurements and agarose gel electrophoresis.

#### 4.1.3.1. OD measurements

We measured the absorbance of 1 µl DNA at 260 nm by a NanoDrop spectrophotometer. To assess the purity of the DNA, the 260/280 nm absorbance ratio was considered. This ratio should be approximately 1.8. Free nucleotides, RNA, single stranded DNA, and double-stranded DNA all absorb at of 260 nm and will contribute to the total absorbance measured in the sample. A ratio of approximately 1.8 is considered as a pure DNA sample. A lower ratio could indicate a contamination of protein, phenol or other molecules that absorb near 280 nm, while a higher ratio indicates RNA contamination.

The 260/230 absorbance ratio is a secondary measurement of the quality of the DNA. This value is often higher than the 260/280 ratio, and should range between 2.0-2.2 to be considered pure. A lower absorbance ratio could indicate contaminants that absorb near 230 nm, such as residual phenol- or guanidine residues.

#### 4.1.3.2. Agarose gel electrophoresis

For verification of plasmid quality, all CEL variants were separated on a 1% agarose gel, containing 0.5 µg/ml ethidium bromide, to visualize the DNA migration in the gel. A total of 1 µl (300-1000 ng) of each plasmid was used. This volume was adjusted to 10 µl by adding $ddH_2O$, and 2 µl of 6x loading buffer. The samples were run adjacent to a molecular-weight size marker at 100 V in TBE buffer for 1.5 h. Molecular Imager® Gel Doc™ XR system was used for documentation.

## 4.2. Constructing a *CEL-HYB* plasmid without an epitope tag

A stop codon was introduced into the pcDNA3.1/V5-His-CEL-HYB plasmid using the QuikChange site directed mutagenesis kit (Agilent Technologies). This plasmid contains an XhoI restriction site directly after the *CEL-HYB* cDNA, before the epitope tag sequence. The stop codon was created in this XhoI restriction site (see Fig. 5.1).

### 4.2.1. Primer design and mutagenesis

The primers needed for mutagenesis were constructed using Agilent Technologies QuikChange Primer Design.

Both primers were diluted to a 100 µM stock solution in TE-buffer. The primers were kept at -20°C until use. The mutagenesis was performed according to the protocol included in the QuikChange II XL Site-Directed Mutagenesis kit.

A PCR reaction mix was prepared in a tube as described in Table 4.1. The tube was mixed and briefly centrifuged, and the PCR was run as shown in Table 4.2.

**Table 4.1. Master mix components for the site directed mutagenesis kit.**

| Component | Volume (µl) |
|---|---|
| $ddH_2O$ | 35.6 |
| DNA template (10 ng/µl) | 4 |
| dNTP mix | 1 |
| Forward primer (20 µM) | 0.7 |
| Reverse primer (20 µM) | 0.7 |
| Pfu ultra HF DNA-polymerase | 1 |
| QuickSolution | 3 |
| 10x reaction buffer | 5 |
| Total volume | 51 |

**Table 4.2 Thermal amplification cycles for mutagenesis.**

| Step | Temperature (°C) | Time | Number of cycles |
|---|---|---|---|
| Initial denature | 95 | 1 min | 1 |
| Denature | 95 | 50 sec | |
| Annealing | 69 | 50 sec | 18 |
| Extension | 68 | 7.5 min | |
| Final extension | 68 | 7 min | 1 |
| | 4 | ∞ | |

### 4.2.2. *DpnI* digestion

To remove parental non-mutated supercoiled dsDNA, *DpnI* digestion was performed. *DpnI* restriction enzyme (1 µl, 10 U/µl) was added directly to the PCR reaction tube. The tube was mixed, spun down for 1 min and immediately incubated for 1 h at 37°C. The plasmid was further transformed according to the manufacturer's (Agilent technologies) protocol for transformation of XL-10 GOLD Ultra competent cells.

The competent cells (45 µl) were thawed on ice before being added 2-mercaptoethanol (2 µl). The tube was incubated on ice for 10 min, with gentle swirling every 2 min. The *DpnI* treated DNA was added (2 µl) to the competent cells and incubated for 30 min on ice, before heat shocked at 42°C for 30 sec. Preheated LB medium (500 µl) was added to the tube and incubated for 1 h at 37°C. Subsequently, the cells were centrifuged at 2000 rpm for 2 min, and 300 µl of the supernatant removed. The cells were resuspended, plated on agar plates containing ampicillin (100 µg/ml) and incubated o/n at 37°C.

Following transformation, plasmid purification was performed as described in section 4.1.2.

### 4.2.3. Sequencing of *CEL-HYB*

The new *CEL-HYB* construct without tag was verified by Sanger sequencing. Sequencing was performed using 0.5 µg of DNA in a reaction mix (Table 4.3) with various primers for the *CEL* gene (Table 4.4). Thermocycling was performed on a GeneAmp PCR System 9700 from Life Technologies according to conditions listed in Table 4.5.

**Table 4.3 Components for sequencing of *CEL* variants.**

| Component | Sample (µl) |
|---|---|
| Betain (5 µM) | 2 |
| BigDye® terminator v1.1 ready reaction mix | 1 |
| Primer (20 µM) | 0.25 |
| Template/dd$H_2O$ | 4.75 |
| 5x sequencing buffer | 2 |
| Total volume | 10 |

**Table 4.4 Primers used for sequencing of *CEL* variants.**

| Primer | Target DNA | DNA target sequence |
|---|---|---|
| BF (FWD) | Exon 2 | GAAGCTGGGCGCC |
| BGH (REV) | BGH polyadenylated sequence | ATCTTCCGTGTCAGCTCC |
| BR (REV) | Exon 6 | CATCCGGCGA |
| CF (FWD) | Exon 4 | TCACCTTCAA |
| CR (REV) | Exon 8 | GAAAGTCACGGAGGAG |
| DF (FWD) | Exon 8 | CCGCCGACATCGACTA |
| EF (FWD) | Exon 11 | ACACACTGGGAACCCT |
| T7 (FWD) | T7 promotor | ATTATGCTGAGTGATATCCC |

**Table 4.5 PCR thermal amplification conditions.**

| Step | Temperature (°C) | Time | Number of cycles |
|---|---|---|---|
| Initial denature | 96 | 10 min | 1 |
| Denature | 96 | 10 sec | |
| Annealing | 58 | 5 sec | 25 |
| Extension | 60 | 4 min | |
| | 4 | ∞ | |

## 4.3. Cell cultures and transfection

### 4.3.1. Culturing

Human embryonic kidney, HEK293 cells were cultured in Dulbeccos Modified Eagle`s Medium (DMEM) with high glucose (4500 mg/L) supplemented with 10% Fetal Bovine Serum (FBS) and 100 U/ml of Antibiotics Antimycotic (AA) in humidified atmosphere with 5% $CO_2$ at 37°C. Cell culturing was done under sterile conditions, and the cells were grown in T75 $cm^2$ flasks during all experiments unless stated otherwise.

### 4.3.2. Sub-culturing and seeding

For passaging of cells, growth medium was removed and the cells were washed with pre-warmed PBS. The cells were added 0.5 ml (T25 $cm^2$) or 1 ml (T75 $cm^2$) of 0.05 % Trypsin-EDTA until they detached from the surface. The cells were re-suspended in DMEM, and an appropriate dilution of cells were transferred to a new T75 $cm^2$ flask, and filled up to 13 ml with medium. For experiments requiring an exact number of cells, the re-suspended cells were transferred to a 15 ml Falcon tube, diluted in PBS and counted with a handheld Sceptre automated cell counter. For all experiments presented in this study, the cell passage number was kept from 11 to 30.

### 4.3.3. Freezing and thawing

A vial of 1 ml frozen HEK293 cells was rapidly thawed and the cells were resuspended in 5 ml of pre-warmed DMEM before being transferred to a T25 $cm^2$ flask. The following day, medium was removed and replaced with fresh growth medium to remove debris, dead cells and traces of DMSO.

To freeze the cells, they were grown to 80-90% confluency in a T75 flask. Next, the cells were washed with PBS, trypsinated and re-suspended in 5 ml of DMEM. The cells were centrifuged for 4 min at 1000 rpm and the medium (supernatant) removed. The cell pellet was re-suspended in 5 ml freezing medium (10% DMSO, 50% FBS and 40% DMEM) and aliquoted into cryo vials. The vials were placed in a container with isopropanol at -80°C for minimum 24 h and transferred to a nitrogen tank for long term storage.

### 4.3.4. Transient transfection of HEK293 cells

Cells were seeded in 6-well plates (4 x $10^5$ cells per well), 12-well plates (2 x $10^5$ cells per well) or 96-well plates (1500 cells per well) and grown to 60-70 % confluency prior to transfection. Transfection

was performed using Lipofectamine® 2000 according to the manufacturer's instructions (Fig. 4.1). For each transfection, the total amount of DNA was 4 µg in 6-wells, 0.5 µg in 12-wells and 0.1 µg in 96-wells. Both Lipofectamine® 2000 and DNA was diluted in OptiMem medium and incubated for 5 min at room temperature (RT) before mixed at a ratio of 1:1. The DNA-lipid mix was incubated for 20 min at RT before added to the cells. The cells were incubated for 4-6 h before the medium was replaced with fresh growth medium and left to grow for 24, 48 or 72 h before further analysis.

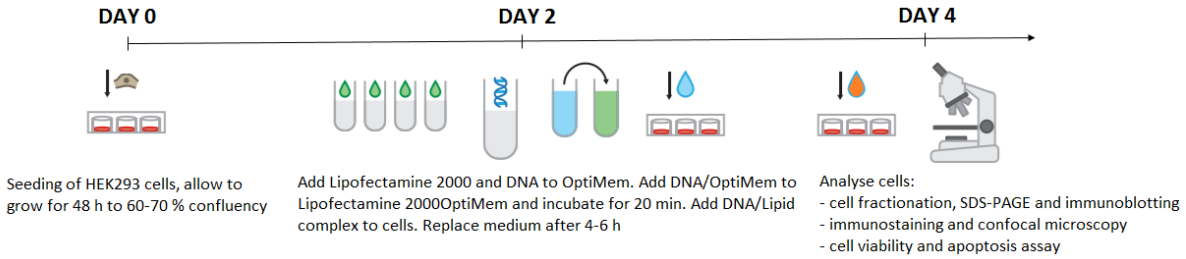



**Figure 4.1 Timeline for transfection with Lipofectamine 2000 of HEK293 cells.** Figure adapted from Lipofectamine® 2000 Reagent protocol.

## 4.4. Preparation of analytical fractions: cell lysate, cell pellet and medium

48 or 72 h post transfection, cells grown in 6-wells were prepared for cell fractionation. The medium was collected and centrifuged at 14 000 g for 5 min. The supernatant was removed and kept for further analysis (medium fraction). The cells were washed once with PBS and added 150 µl of ice-cold RIPA lysis buffer containing a cocktail of protease inhibitors (1 tablet per 10 ml of 1x RIPA buffer). The cells were collected using a cell scraper and the lysate incubated for 30 min on ice before centrifugation at 20 000 g for 15 min. Again, the supernatant was removed and kept for further analysis (lysate fraction). The pellet was washed with 150 µl of PBS and centrifuged at 20 000 g for 15 min two times. Finally, the pellet was added 100 µl of 2x LDS sample buffer with reducing agent and stored at -80°C until further analysis (pellet fraction).

### 4.4.1. Protein concentration determination

The protein concentration was determined for all samples prior to SDS-PAGE and immunoblotting, using the Pierce™ BCA Protein Assay Kit. A total of 5 µg protein from the lysate fraction was used for SDS-PAGE sample loading. The equivalent volume was used from the medium fraction, assuming a correlation between the amount of protein in lysate and medium fractions.

## 4.5. SDS-PAGE and immunoblotting

The cell lysate and medium samples were prepared with a total volume of 20 µl (Table 4.6). The samples were denatured on a heating block at 56°C for 15 min, while the pellet samples were denatured for 95°C for 15 min. The samples were loaded on a 10% NuPage Bis-Tris gel (1 mm) and separated by electrophoresis. For all medium and lysate samples, 20 µl was loaded, while 3-5 µl was loaded for the pellet samples. Four µl of Precision Plus Protein™ Dual Colour Standards marker and 2 µl of Magic Mark XP Western Protein Standard were loaded on the gels. The gels were run in 1x NuPage MOPS buffer at 90 V for 15 min, and then at 180 V for 1 h.

**Table 4.6 Components for SDS-PAGE sample preparation for lysate and medium samples.**

| Component | Volume (µl) |
|---|---|
| Reducing agent (500 mM dithiothreitol (DTT)) | 2 |
| Sample/dd$H_2O$ | 13 |
| 4x LDS sample buffer | 5 |
| Total volume | 20 |

Western Blotting was performed with semi-wet blotting of NuPage® gels in XCell Blot Module chambers according to the manufacturer's instructions. Proteins were transferred to a Polyvinylidene fluorid (PVDF) membrane, which was activated in methanol for 1 min prior to use. The electro blotting procedure was performed at 30 V for 1 h in NuPage transfer buffer with 10% methanol. The membrane was then blocked in 5% dry milk either overnight at 4°C or for 1 h at RT. The membrane was incubated with the primary antibodies anti-V5 (1:20 000) or anti-CEL (1:10 000) at 1 h at RT or o/n at 4°C to detect the recombinant CEL proteins. The anti-GAPDH (1:1000) antibody was used for determining the amount of GAPDH protein as loading control. All antibodies were diluted in 1% dry milk in PBS-Tween. The membrane was washed 3 x 5 min with PBS-Tween (0.05 %) and further incubated with HRP-conjugated secondary antibodies for 1 h at RT. All secondary antibodies were diluted 1:5000. The membrane was washed 1x 15 min and 3x 5 min with PBS-Tween before being developed with an ECL Western Blotting kit. The signals were analysed using a Las 1000 imager.

## 4.6. Immunostaining

Cells seeded in 12-well plates on coverslips coated with Poly-L-lysine were transiently transfected and analysed with immunostaining 48 h post transfection. The cells were washed twice with pre-warmed PBS and fixed in pre-warmed 3% paraformaldehyde (pH 7.2) for 30 min. Then, the cells were washed twice with washing buffer (WB), before treatment with quench solution for 15 min to remove unreacted

aldehyde groups. For permeabilization, the cells were treated with blocking buffer (BB) for 30 min. Next, the cells were washed three times with WB-saponin and incubated with primary antibody (1:200 in BB) for 2 h at RT. The cells were further washed quickly three times with WB-saponin, followed by 2 h washing at RT on a shaker with subsequent buffer changes. The samples were incubated with Alexa Fluor 488 or Alexa Fluor 594 conjugated secondary antibodies for 1 h at RT. After a final triple washing with WB-saponin, the cells were washed twice with PBS, and mounted on glass slides with 10 µl of Prolong Gold Antifade Solution with DAPI. The results were evaluated and images collected with an SP5 AOBS confocal microscope.

For double staining, the protocol described above was repeated. Thus, after the last wash with WB-saponin, the cells were incubated with blocking buffer again for 30 min – then incubated with a new primary and secondary antibody.

### 4.7. Cell viability assay

HEK293 cells were seeded in 96-well plates and transfected as described above (section 4.3.4), and cell viability was evaluated 48 h post transfection. To determine the number of cells needed for 60-70 % confluency at the day of transfection, a cell titration was carried out with a 96-well plate. Cells were seeded from 500 to 2000 cells per well, to evaluate the optimal number of cells to use for detection with CellTitre-Glo®. The wells containing 1500 cells was found to be the most optimal after 48 h of growth and was used for further experiments. Measurements showed that 20 µl of medium would evaporate after 48 h, and so 120 µl of medium was added to the cells, leaving 100 µl which was required for detection 48 h post transfection.

Cell viability was measured using the CellTiter-Glo® Luminescent Cell Viability kit (Promega), following the manufacturer's instructions (Fig. 4.2). Both cells and reagents were left in RT for 30 min before treatment with equal amounts of reagents, 100 µl of CellTitre-Glo®-solution was added to the cells. Before detection, the 96-well plates were subjected to 2 min of shaking and 10 min of destabilization. The luminescent signal was detected with a Tecan Infinite M200PRO luminometer.

The number of viable cells in the plate is determined by measuring the presence of ATP, which indicates metabolically active cells. Luciferase generates a luminescent signal in the presence of ATP, $Mg^{2+}$ and molecular oxygen by converting luciferin to oxyluciferin and light. By addition of the CellTitre-Glo® reagent, the luminescent signal is proportional to the ATP present, which in turn is directly proportional to the number of viable cells.

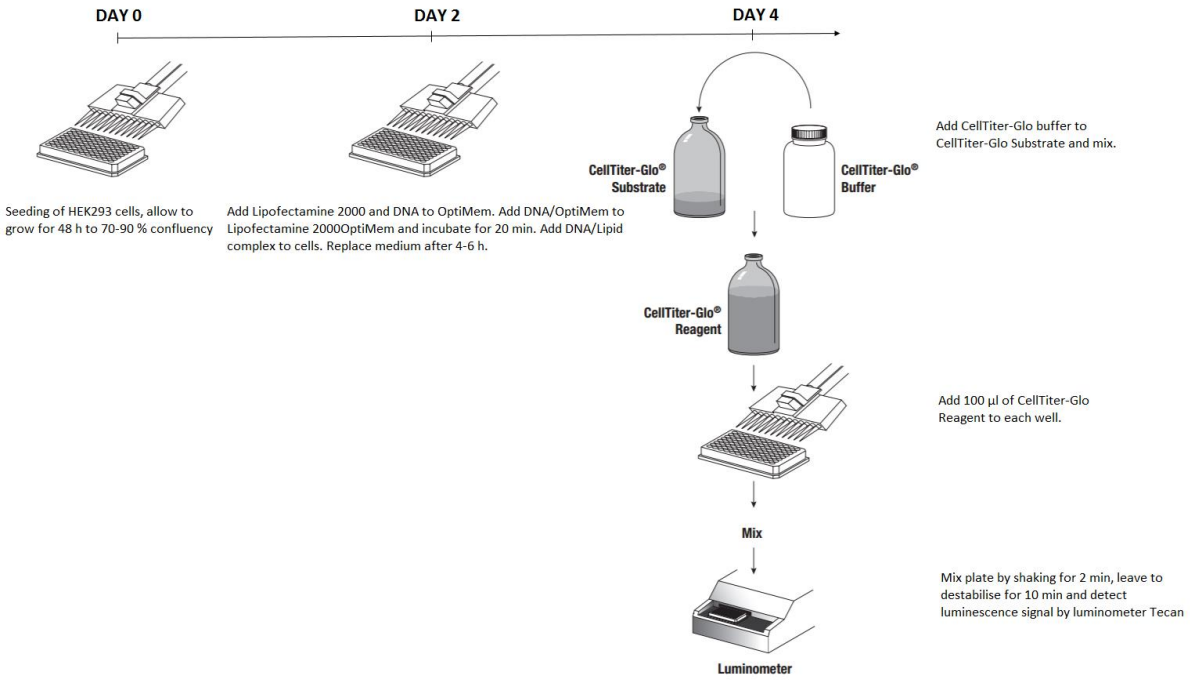



**Figure 4.2 Timeline for cell viability assay**. Figure adapted from CellTiter-Glo® Luminescent Cell Viability Assay protocol.

## 4.8. Apoptosis assay

HEK293 cells were seeded and transfected as described above in section 4.3.4. After 48 h, growth medium was removed to leave a volume of 250 µl per well. One µl of CellEvent™ Caspase-3/7 Green Detection Reagent (final concentration 7.5 µM) was added to the cells, and incubated at 37°C for 30 min. The cells were washed twice with PBS, fixed in 3 % PFA and washed twice. The coverslips were mounted in 10 µl of Prolong Gold Antifade Solution with DAPI and analysed by confocal microscopy as described above. Apoptotic cells were visualized and identified by their green nuclei.

The Caspase-3/7 Green Detection Reagent is a four-amino acid peptide (DEVD), containing a cleaving site for Caspase-3/7, and is bound to a DNA-binding dye with an absorption maximum of 502-530 nm. The DNA-binding dye is non-fluorescent as it cannot bind to DNA due to inhibition from the DEVD peptide. However, in apoptotic cells, Caspase-3/7 is activated and the DEVD peptide is cleaved off, allowing the DNA-binding dye to bind to DNA and produce a green luminescent signal.

### 4.9. Statistical analysis

All statistical analyses were conducted by two-tailed student t-tests. The statistical calculations were all carried out using the Excel program of the Microsoft Office software package. *P*-values < 0.05 were considered as statistically significant.

### 4.10. De-glycosylation assay

Cell growth medium (100 µl) was collected from transiently transfected HEK293 cells, 48 h post transfection. The medium was acetone-precipitated (medium/acetone 1:5) at -20°C for 30 min, and the proteins collected by centrifugation at 18000 g for 30 min at 4°C. The acetone was removed by carefully pipetting off the solution, and the pellet was allowed to dry at 37°C for 5 min. The proteins were re-suspended in 25 µl of denaturation buffer, and incubated at 100°C for 10 min. Next, 125 µl incubation buffer was added to the proteins. The proteins were treated with three different glycosidase enzymes; 2 units of N-glycosidase F and/or 2.5 milliunits of O-glycosidase and/or 5 milliunits of α2-3,6,8,9 Neuraminidase (Sialidase). The samples were digested on shaking (300 rpm) overnight at 37°C, and 1:5 acetone-precipitation was repeated. The pellets were resuspended in 2x LDS loading buffer and denatured at 95°C for 10 min, before analysis by SDS-PAGE and immunoblotting as described above.

### 4.11. Isoelectric Focusing

Cell growth medium (5 µl) was collected from transiently transfected HEK293 cells, 48 h post transfection, and 10 µl samples were prepared by adding 5 µl of 2x IEF Sample buffer (pH 3-10). To prepare 1x IEF Anode- and Cathode buffer, 20 ml of 50x IEF Anode buffer was added to 980 ml of $ddH_2O$, and 20 ml of IEF Cathode buffer was added to 180 ml of $ddH_2O$. Both buffers were kept at 4°C until use. The 10 µl samples were loaded on the gels, and the inner buffer chamber was filled with chilled 1x IEF Cathode buffer, while the outer chamber was completely filled with chilled 1x IEF Anode buffer. The gel was run for 1 h at 100 V, 1 h at 200 V and for 30 min at 500 V. The gel was further equilibrated in 0.7 % acetic acid for 10 min, and Western blot was performed. Western blot was performed as described previously, however with 0.7 % acetic acid as transfer buffer, and the gel/membrane sandwich was assembled in reverse order, so that the proteins migrated towards the cathode during the transfer.

# 5. Results

## 5.1. Construction of a *CEL-HYB* plasmid without the V5/His tag

To make a plasmid that expressed the human *CEL-HYB* gene without a tag, we introduced a stop codon into our already existing pcDNA3.1/V5-His-CEL-HYB construct. The stop codon was introduced by site-directed mutagenesis after CEL exon 11 but before the V5/His sequence, as illustrated in Fig. 5.1. The new stop codon was verified by sequencing. Furthermore, we sequenced through the entire new *CEL-HYB* cDNA using primers covering *CEL* as well as vector sequence boundaries to make sure that the construct was correct and in frame.


**Figure 5.1 Sequence of *CEL-HYB* after introduction of stop codon in *CEL-HYB-V5/His*.** Depicted in the figure is *CEL-HYB* exon 11. The XhoI site (blue) and V5/His sequence (green) are shown, followed by a stop codon (purple). A two-base mutation in the XhoI site creates a new stop codon prior to the V5/His sequence.

## 5.2. Plasmid purification and determination of DNA purity

In addition to OD measurements for estimating DNA concentration and quality, all plasmids used in this project were separated on a 1 % agarose gel to confirm that the preparation procedure had been successful, and that the plasmids were intact. The high molecular bands seen in Figure 5.2 are likely open circular plasmids while the low molecular bands are the intact, supercoiled plasmids.

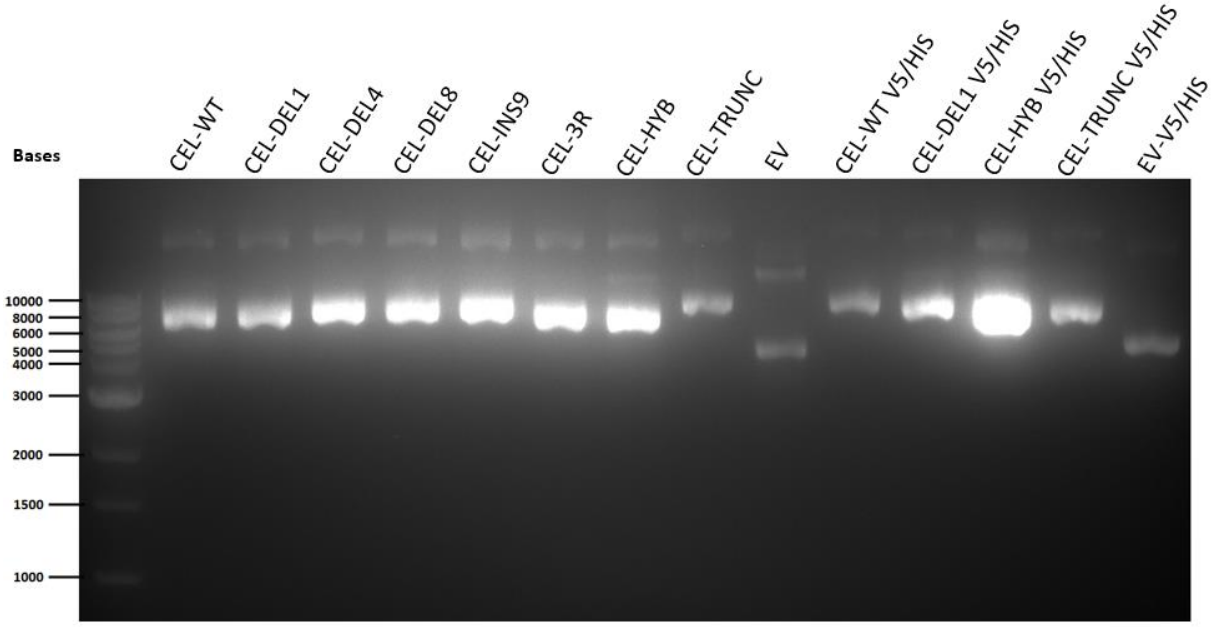



**Figure 5.2 Plasmid verification by agarose gel electrophoresis.** A total of 1 µl (300-1000 ng) of the different CEL plasmids were loaded and separated on a 1 % agarose gel stained with ethidium bromide.

## 5.3. The effect of the C-terminal V5/His-tag on CEL protein expression and secretion

### 5.3.1. Western blot analysis

To investigate whether the C-terminal V5/His tag has an effect on CEL protein expression and secretion, HEK293 cells were transiently transfected with plasmids expressing CEL-WT, CEL-DEL1, or CEL-TRUNC with and without tag. Cells transfected with empty vector (EV) was included as a negative control. After 48 h of transfection, cell medium, lysate and pellet fractions were isolated and analysed by SDS-PAGE and Western blotting. Immuno-detection was performed using either the anti-CEL (Pittsburgh) or anti-V5 antibodies. GAPDH was used as loading control.

Detection with the anti-CEL (Pittsburgh) antibody is seen in the left panel of Figure 5.3. An unspecific band appeared for all variants at approximately 70 kDa in the same fraction (this unspecific band has been observed by others in our research group as well as by our collaborators in Pittsburgh who developed the antibody). For several of the variants, we also detected two or more bands. These bands are likely to be differently matured forms of CEL, either fully glycosylated (bands with higher molecular weight) or partially glycosylated (lower bands) proteins (Johansson et al., 2011, Fjeld et al., 2015). For CEL-TRUNC, these are likely different isomers of N-glycosylation, as CEL-TRUNC contains only the four first amino acids of the VNTR and no C-terminal O-glycans.

All variants were detected in the lysate fraction, however, the untagged CEL variant had stronger signals compared to the V5/His tagged variants. CEL-WT displayed a slightly stronger band compared to CEL-DEL1 and CEL-TRUNC.

The trend in the pellet fraction was similar to the lysate: all V5/His tagged variants showed a lower expression than their untagged counterparts. However, we did not observe any band for CEL-WT-V5/His, and CEL-TRUNC had a very strong signal, both when compared to its tagged counterpart and to the other variants. We also detected different glycosylated variants of CEL in the pellet.

All variants could be detected in the medium. Both untagged and tagged variants of CEL-WT and CEL-DEL1 were secreted at similar levels, while both variants of CEL-TRUNC showed weak bands. Except for CEL-TRUNC, all variants showed only one band, which is likely to correspond to the fully glycosylated form of the proteins.

Detection with the anti-V5 antibody (right panel) resulted in a faint band for CEL-WT-V5/His in the pellet fraction. CEL-DEL1-V5/His and CEL-TRUNC-V5/His clearly had weaker bands than CEL-WT-V5/His in medium fraction. This blot also confirmed that only the V5/His-tagged variants of CEL were recognized by the anti-V5 antibody.

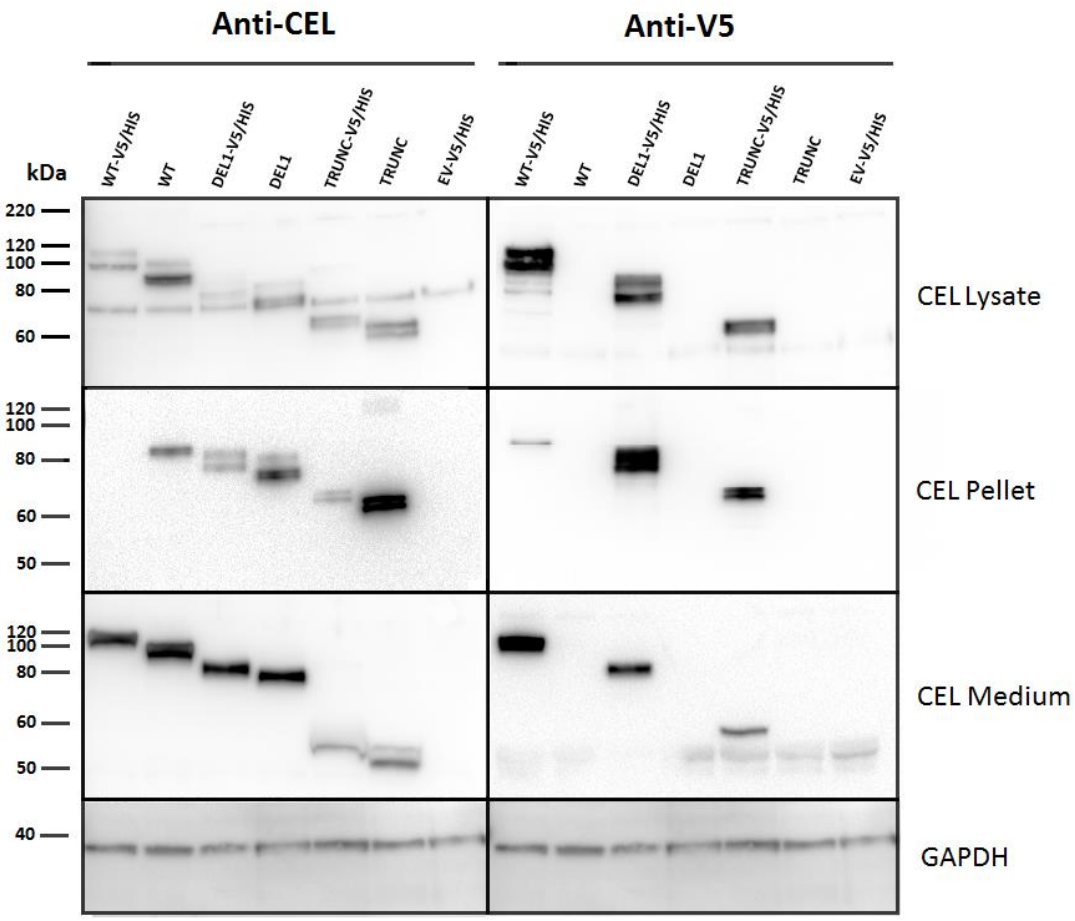



**Figure 5.3 Expression, distribution and secretion of tagged versus untagged CEL-variants in various fractions of HEK293 cells.** Cells were transiently transfected with plasmids encoding CEL-WT, CEL-DEL1 and CEL-TRUNC with or without the V5/His tag. Lysate, pellet and medium fractions were analysed by SDS-PAGE and Western blotting using a CEL-specific antibody (Pittsburgh) or anti-V5 antibody. GAPDH was used as loading control. The results shown are representative of three independent experiments.

We also analysed the effect of the V5/His tag on the CEL-HYB variant. Due to the need of constructing a CEL-HYB plasmid without tag, this experiment was performed at a later stage than for the variants above. However, we also included CEL-WT in this experiment, as a control. The result can be seen in Figure 5.4.

With the anti-CEL antibody (Pittsburgh - left panel) we observed the same unspecific band in the lysate (70 kDa) as shown in Figure 5.3. Both CEL-HYB variants were detected in this fraction and the tagged HYB protein appeared in different modified forms. In the pellet, we could observe strong bands for the CEL-HYB variants, as well as higher molecular forms of the proteins. In contrast, hardly any CEL-HYB was seen in the medium. Actually, no band was observed for the CEL-HYB variant without a tag.

With the anti-V5 antibody, CEL-HYB was detected in all fractions, but the amount was reduced in the lysate compared to the pellet and medium fractions. Compared to the anti-CEL antibody, the anti-V5 antibody recognized the CEL-HYB-V5/His protein to a much higher degree. As for CEL-WT proteins, the expression patterns observed were basically the same as in Figure 5.3, except that CEL-WT-V5/His was now detected in the pellet. This may be due to less protein loaded to the SDS-PAGE represented in Figure 5.3 compared to Figure 5.4 (3 µl vs 5 µl, respectively).

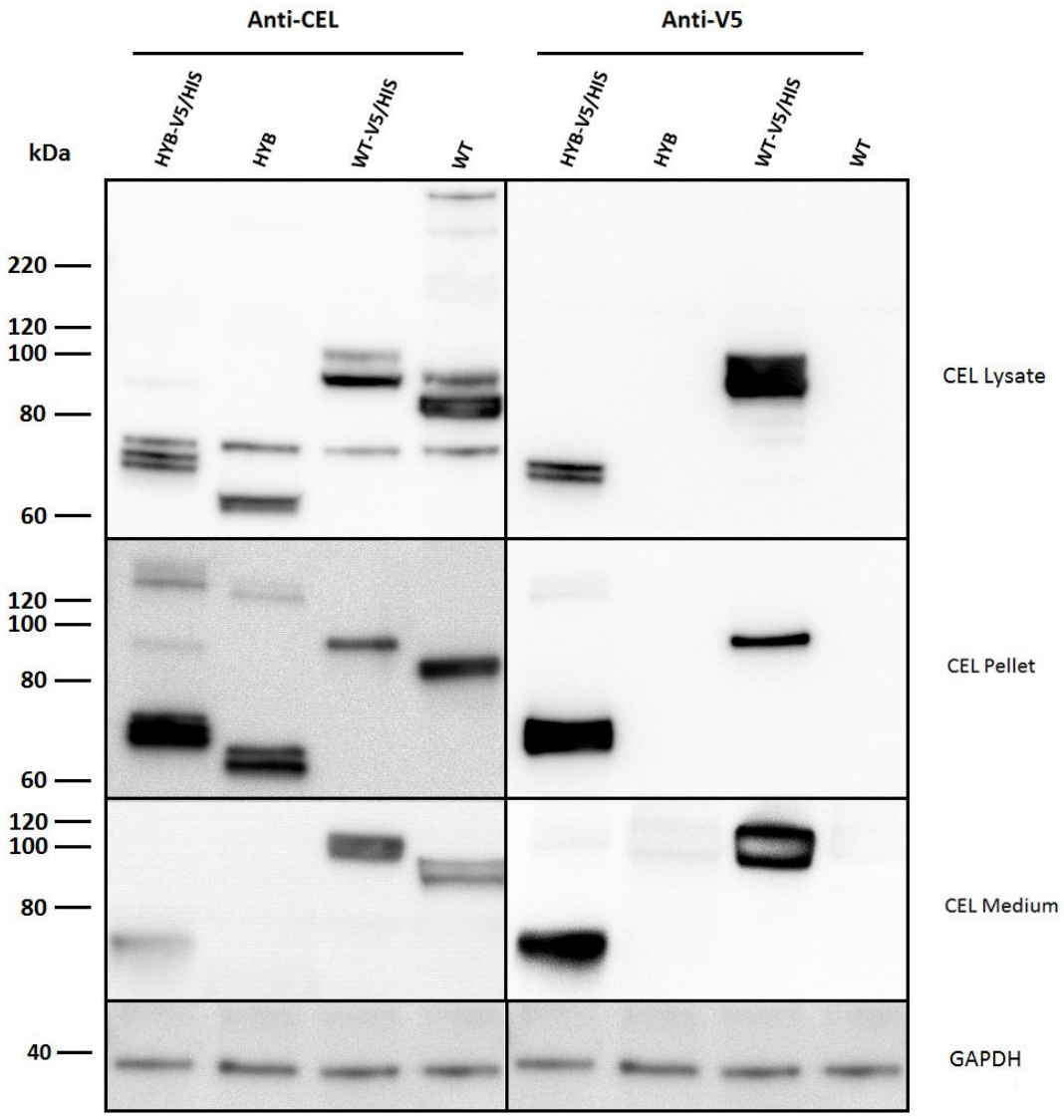
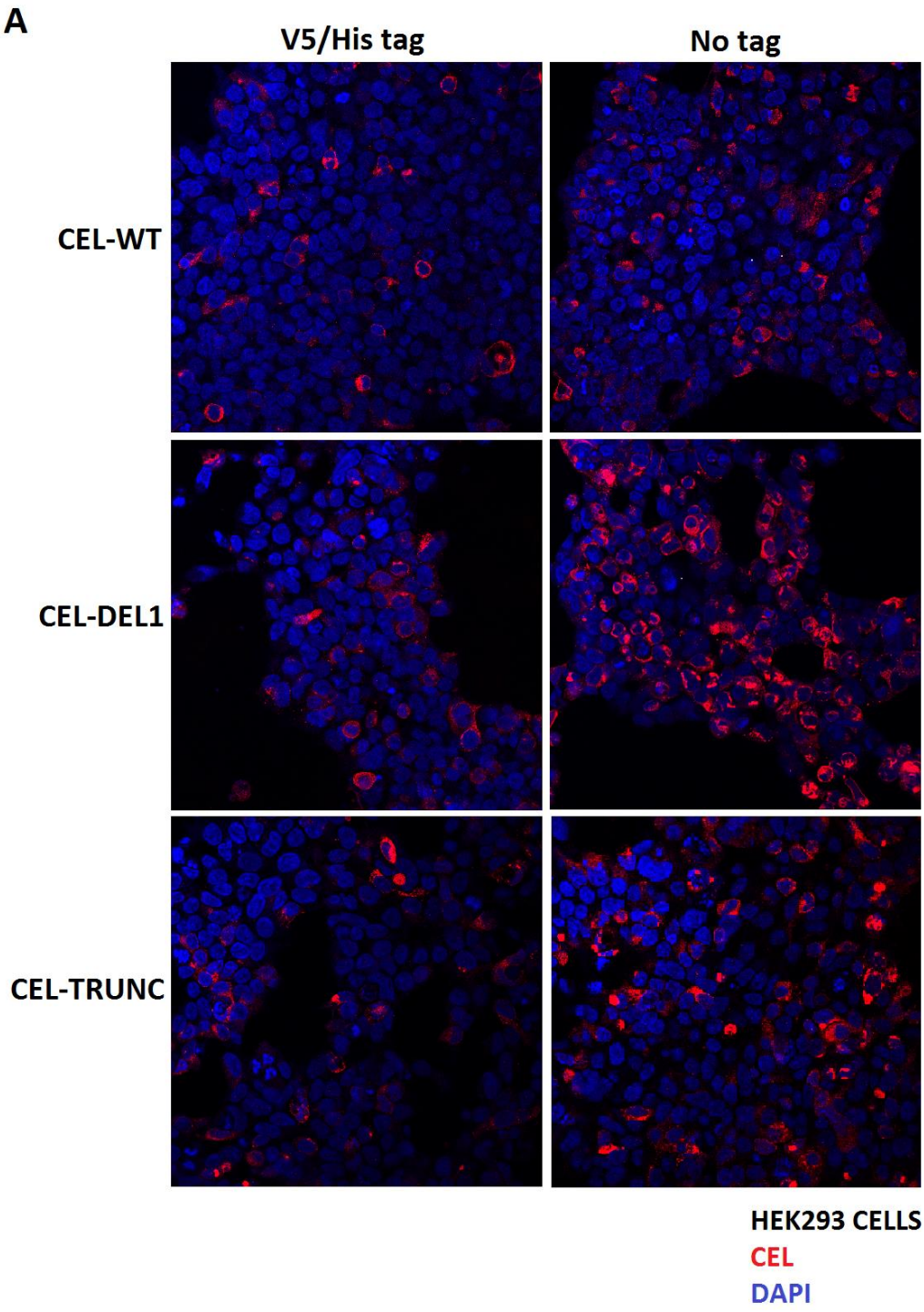



**Figure 5.4 Expression, distribution and secretion of epitope tagged versus untagged CEL-HYB and CEL-WT variants in various fractions of HEK293 cells.** Cells were transiently transfected with CEL-WT and CEL-HYB with or without the V5/His tag. Lysate, pellet and medium fractions were analysed by SDS-PAGE and Western blotting using the CEL-specific antibody (Pittsburgh) and anti-V5 antibody. GAPDH was used as loading control. The results shown are representative of two independent experiments.

### 5.3.2. Immunostaining and confocal analysis

Immunostaining and confocal imaging were performed to study the intracellular localisation of CEL-WT, CEL-DEL1, CEL-TRUNC and CEL-HYB, and to compare possible effects of the conjugated V5/His tag. HEK293 cells were transiently transfected and the cells subjected to immunostaining 48 h post transfection. Both tagged and untagged variants of CEL were detected with the CEL-specific antibodies. Cells transfected with empty vector were used as a negative control (data not shown).

In summary, we observed a stronger intracellular signal for CEL-WT, CEL-DEL1 and CEL-TRUNC without tag compared to their tagged counterparts (Fig. 5.5 A). We also found that there was a much stronger intracellular signal of CEL-DEL1 compared to both CEL-WT and CEL-TRUNC, whereas CEL-TRUNC showed a stronger signal than CEL-WT (Fig. 5.5 A). As for CEL-HYB, it was observed that the intracellular staining was equally strong for both the untagged and tagged variant (Fig. 5.5 B). However, no conclusions can be drawn about subcellular localisation based on the results presented in these figures, as no markers for intracellular compartments were used.


A
V5/His tag
No tag
CEL-WT
CEL-DEL1
CEL-TRUNC
HEK293 CELLS
CEL
DAPI

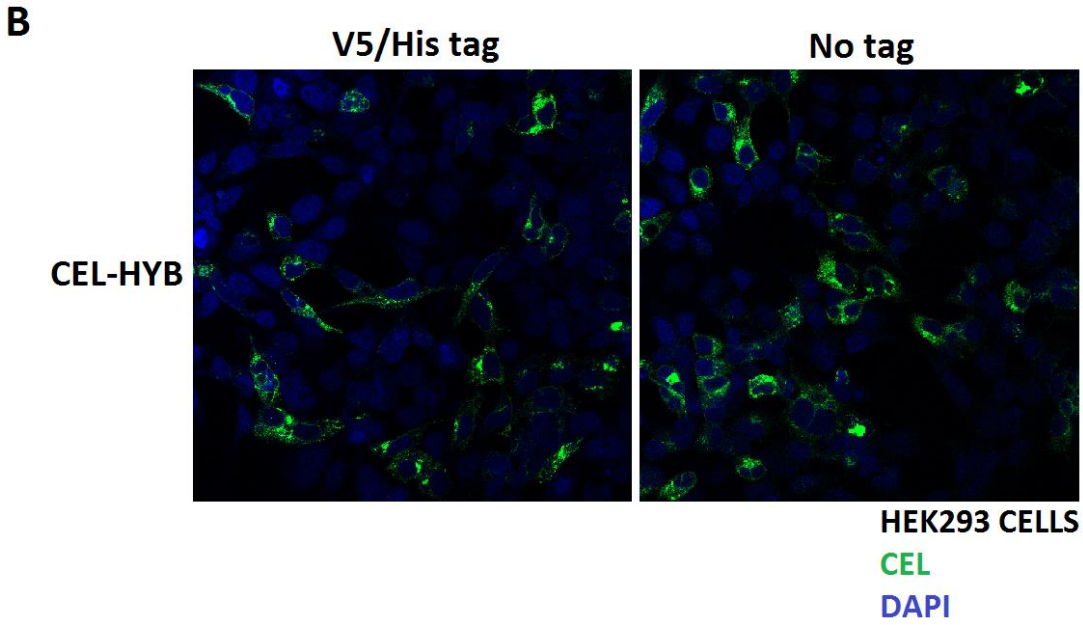



**Figure 5.5 Intracellular localisation of epitope tagged and untagged CEL variants in HEK293 cells.** The cells were subjected to immunostaining 48 h post transfection. The As20 antibody was used to detect the CEL-HYB variants whereas the Vanko antibody was used for all other variants. The secondary antibodies used were AlexaFlour 488 (green) and AlexaFlour 594 (red), respectively. **A)** Cells transfected with plasmids encoding CEL-WT, CEL-DEL1 and CEL-TRUNC with and without the V5/His tag. The results shown are representative of three independent experiments. **B)** Cells transfected with plasmids encoding CEL-HYB and CEL-HYB-V5/His.

## 5.4. The effect of CEL VNTR length on protein expression and secretion

### 5.4.1. Western blot analysis

In this study, we transfected cells for both 48 h and 72 h. The latter transfection time was selected to compare our results more directly with the data of Xiao et al., 2016. We only used untagged constructs. HEK293 cells were transiently transfected with plasmids encoding CEL-WT, CEL-DEL1, CEL-DEL4, CEL-DEL8, CEL-TRUNC, CEL-3R, CEL-INS9 and EV, to investigate and compare the protein secretion of the different variants. Cell fractionation was performed as described in section 5.3.1. The fractions were analysed by SDS-PAGE and Western blot, and detected with the anti-CEL antibody (Pittsburgh) and GAPDH as a loading control (Fig. 5.6).

For both 48 and 72 h transfection, all variants could be detected in the different fractions. However, the amount of CEL protein in each fraction varied between variants, and with the transfection time. When detecting with the anti-CEL antibody (Pittsburgh), the same unspecific band seen in Figure 5.3 appeared for all variants at approximately 70 kDa in the lysate fraction. In addition, we detected two or more bands for several variants. Again, these are likely to be differently post-translational N- and O-glycosylated forms, as previously reported (Johansson et al., 2011, Fjeld et al., 2015).

All variants were detected in the lysate fraction. We also observed that all variants except CEL-DEL1 were present as two bands. Also, we observed weaker bands for the disease-causing variants CEL-DEL1 and CEL-DEL4 in the lysate fraction. These bands were weaker after 72 h transfection compared to the 48 h transfection. CEL-TRUNC appeared as equally weak as CEL-DEL1 and CEL-DEL4 after 72 h, despite a very strong band in the 48 h transfection. In addition, the unspecific band appearing for all variants at approximately 70 kDa in the lysate fraction was generally observed as weaker or absent after 72 h of transfection.

Although not as prominent as for the lysate fraction, we did observe stronger band for all variants compared to CEL-WT in the pellet fraction, after both 48 and 72 h transfection. Also, we observed that some of the variants migrated as different glycosylated variants of CEL in the pellet.

Less amounts of CEL-TRUNC and disease-causing variants CEL-DEL1 and CEL-DEL4 were observed in the medium fraction, while variants CEL-DEL8, CEL-3R and CEL-INS9 seemed to be secreted to the same extent as CEL-WT. After 72 h of transfection, the secretion of these variants had decreased even more, with very faint bands for CEL-DEL1 and CEL-TRUNC. Interestingly, for the 72 h transfection, CEL-DEL8 exhibited a stronger band in the medium than CEL-WT. Except for CEL-TRUNC, all variants appeared as only one band, which probably corresponds to the fully glycosylated form of the proteins.

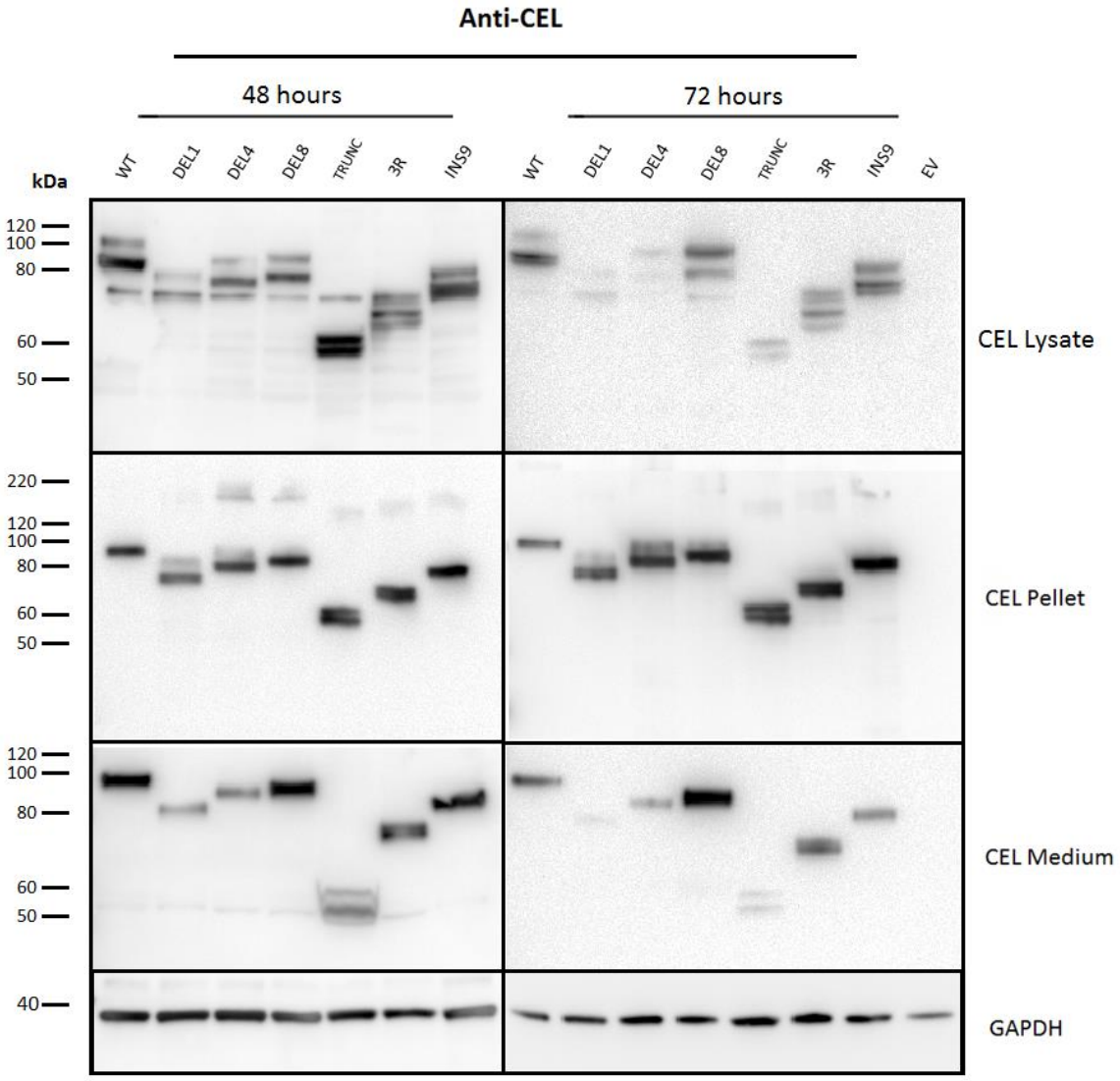



**Figure 5.6 Comparison of expression, distribution and secretion in cellular fractions of CEL-variants without tags at 48 and 72 h post-transfection.** HEK293 cells were transiently transfected with plasmids encoding different CEL variants CEL-WT, CEL-DEL1, CEL-DEL4, CEL-DEL8, CEL-TRUNC, CEL-3R and CEL-INS9. Lysate-, pellet- and medium fractions were collected 48 and 72 h post-transfection, and analysed by SDS-PAGE and Western blotting using CEL specific antibody (Pittsburgh). GAPDH was used as loading control. The results shown are representative of three independent experiments.

## 5.5. The effect of CEL on cell viability and apoptosis

### 5.5.1. Cell viability

We examined all CEL variants to investigate the effect of both the epitope tag and VNTR length on cell viability. The cell viability experiments were performed in transiently transfected HEK293 cells. The cell viability was analysed 48 h post transfection by detecting the luminescence produced when luciferase converts luciferin to oxyluciferin and light. This takes place in the presence of ATP, $Mg^{2+}$ and molecular oxygen, thereby indirectly measuring the ATP levels in the cells, which indicate metabolically active cells.

The results are shown in Figure 5.7. Surprisingly, cells transfected with untagged CEL variants showed a higher overall cell viability compared to CEL-WT, except for CEL-TRUNC and EV. The experiment also indicated that some variants resulted in a higher viability than both untransfected control cells and CEL-WT, observed for both tagged and untagged.

Interestingly, untagged CEL-DEL1 displayed a stronger impact on viability compared to CEL-WT, but this was not observed for the V5/His tagged variant. Variants that resulted in significant increased or reduced viability of HEK293 cells were: CEL-WT-V5/His, CEL-DEL1-V5/His, CEL-HYB-V5/His and both variants of CEL-TRUNC.

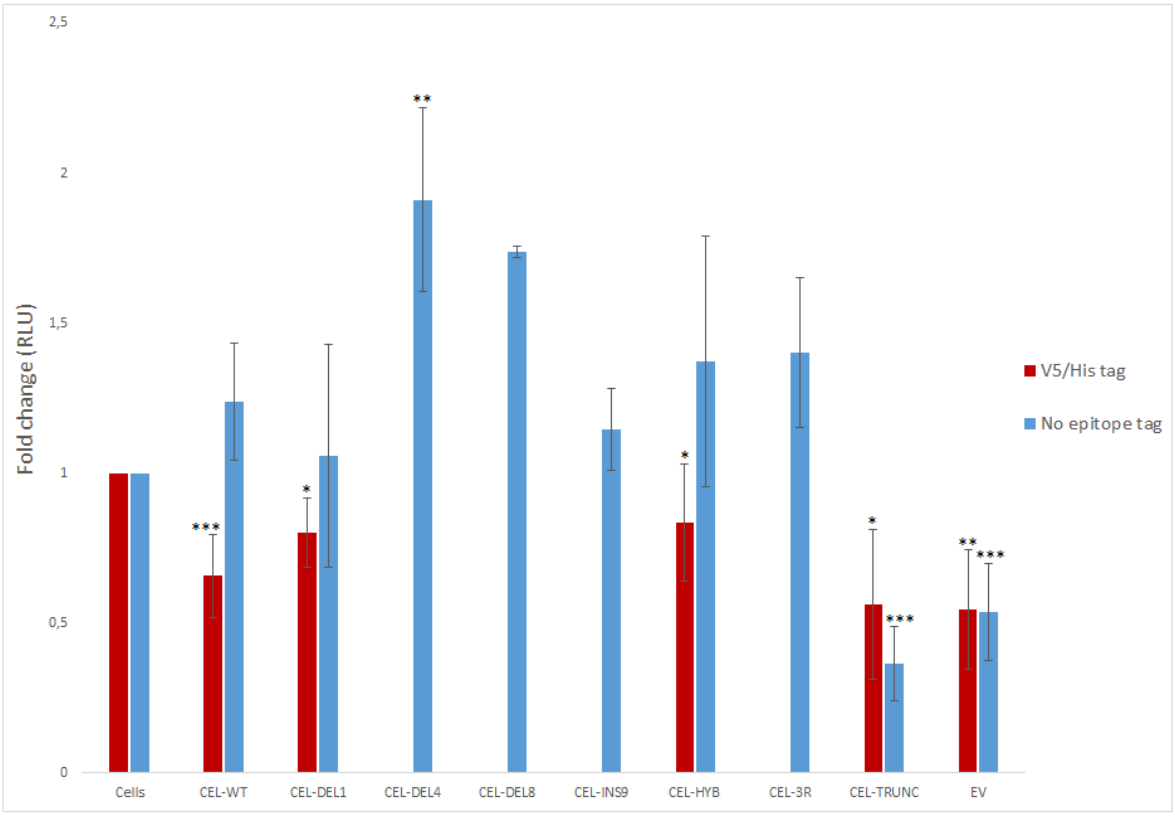



**Figure 5.7 The effect of CEL variants of HEK293 cell viability.** Cell viability was measured as intracellular ATP levels. Results are presented as fold change of relative light units (RLU) for all variants compared to untransfected cells. The data shown are the mean of three experiments with ten technical replicates. ±SEM and *p*-values of statistical significance; *p≤0.05, **p≤0.01, ***p≤0.001. Red bars represent the V5/His tagged variants, while blue bars represent untagged variants. For variants CEL-DEL4, CEL-DEL8, CEL-INS9 and CEL-3R, only untagged variants were used.

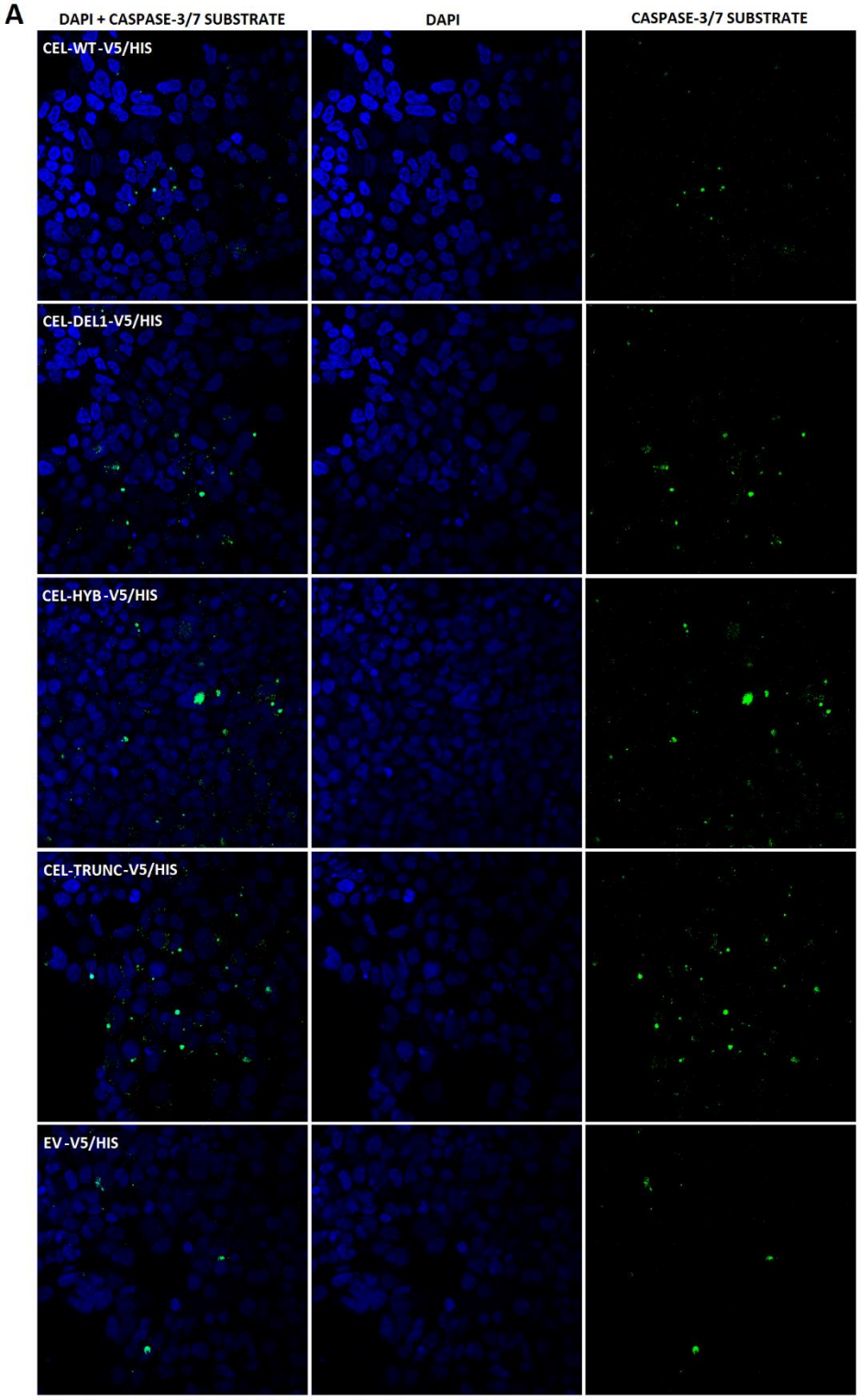
### 5.5.2. Apoptosis

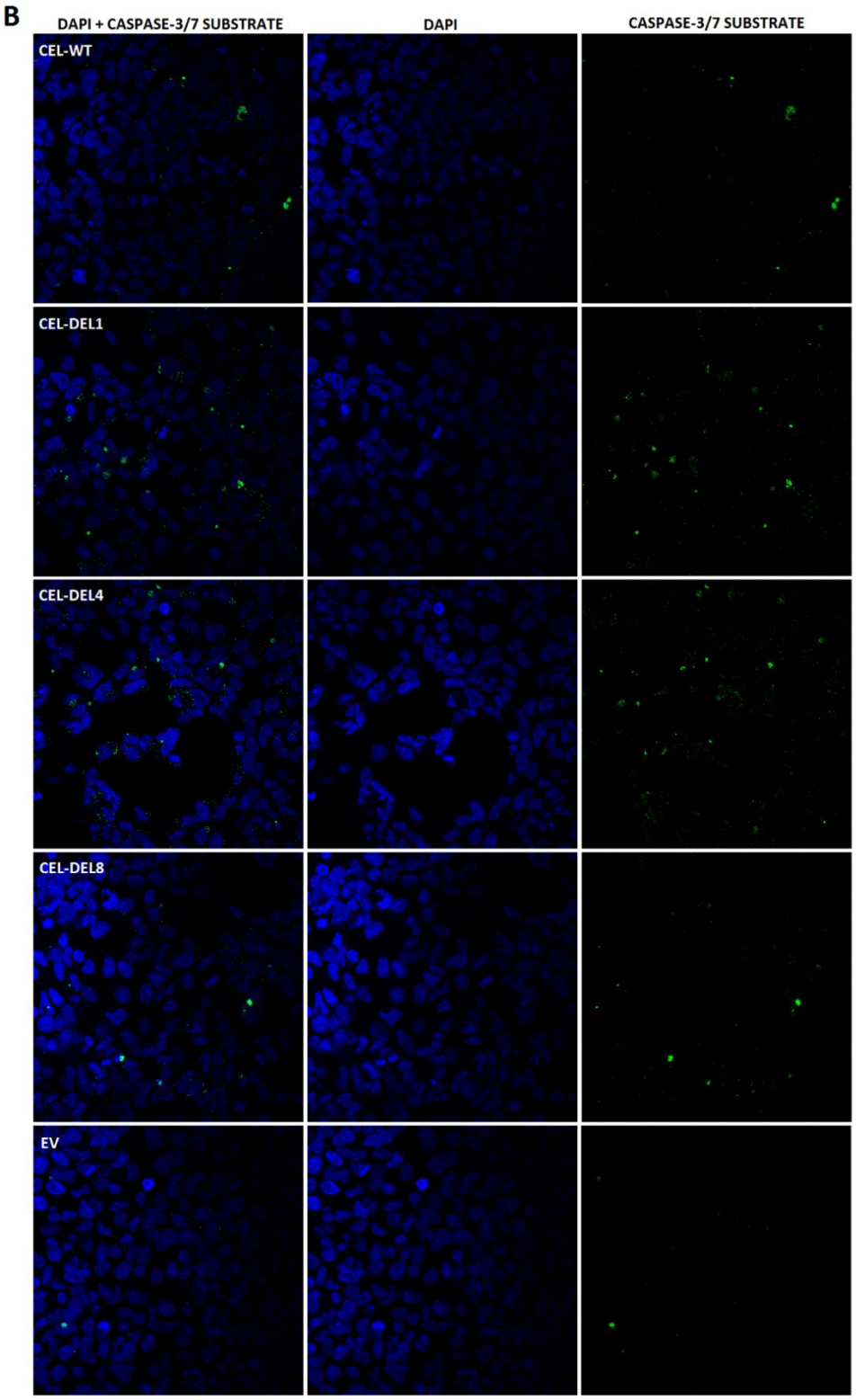
We investigated the effect of all CEL variants on apoptosis in HEK293 cells. The cells were transiently transfected and analysed 48 h post transfection using the CellEvent Caspase-3/7 Green Detection Agent that detects Caspase 3/7 activity, a marker for apoptosis. Apoptosis was observed by confocal imaging as green cells or green nuclear fragments (Fig. 5.8 A, B and C). The number of apoptotic cells were quantified for each CEL variant and means of two independent experiments are presented in Figure 5.8 D.

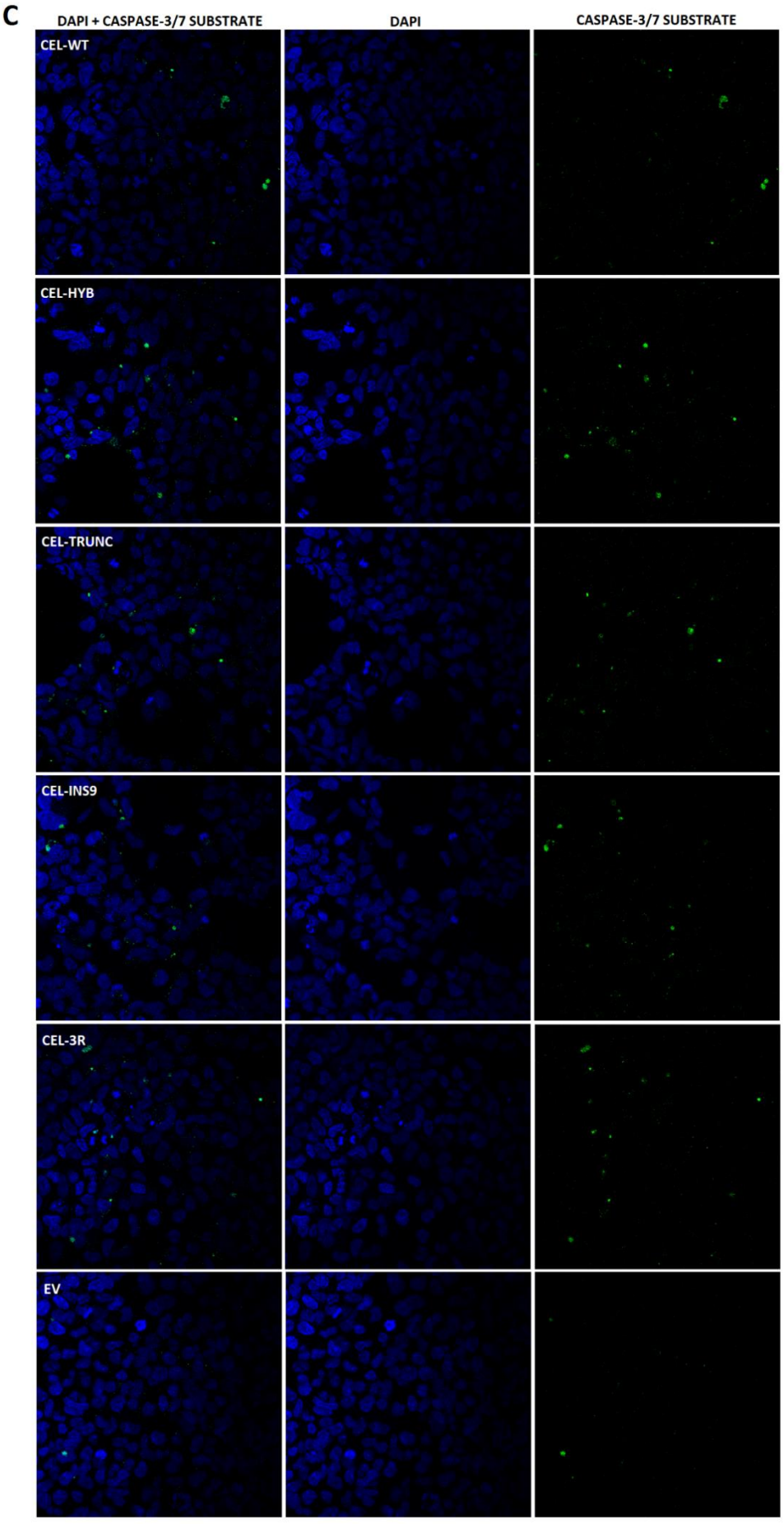
Compared to CEL-WT-V5/His, showing 5.3 % apoptotic cells, there were an increase in apoptosis in HEK293 cells transfected with CEL-DEL1-V5/His, CEL-HYB-V5/His and CEL-TRUNC-V5/His, with total percentage of apoptotic cells being 10.7 %, 6.6 % and 7.7 %, respectively (Fig. 5.8 A, D). Cells transfected with empty vector showed 4.3 % of apoptosis, indicating that the transfection in itself was harmful to the cells.

For the untagged CEL proteins, all variants except CEL-DEL8 had higher numbers of apoptotic cells than CEL-WT. CEL-DEL1, CEL-DEL4 and CEL-HYB showed considerably higher levels of apoptosis compared to CEL-WT, with 9.2 %, 8.2 % and 11.7 % against 4.3 %, respectively (Fig. 5.8 B, D).

Slightly lower level of apoptosis was observed for CEL-TRUNC, CEL-3R and CEL-INS9, with 5.3 %, 4.9 % and 4.7 %, respectively (Fig. 5.8 C, D).


A
DAPI + CASPASE-3/7 SUBSTRATE
DAPI
CASPASE-3/7 SUBSTRATE
CEL-WT-V5/HIS
CEL-DEL1-V5/HIS
CEL-HYB-V5/HIS
CEL-TRUNC-V5/HIS
EV-V5/HIS


B
DAPI + CASPASE-3/7 SUBSTRATE
DAPI
CASPASE-3/7 SUBSTRATE
CEL-WT
CEL-DEL1
CEL-DEL4
CEL-DEL8
EV


C
DAPI + CASPASE-3/7 SUBSTRATE
DAPI
CASPASE-3/7 SUBSTRATE
CEL-WT
CEL-HYB
CEL-TRUNC
CEL-INS9
CEL-3R
EV

D

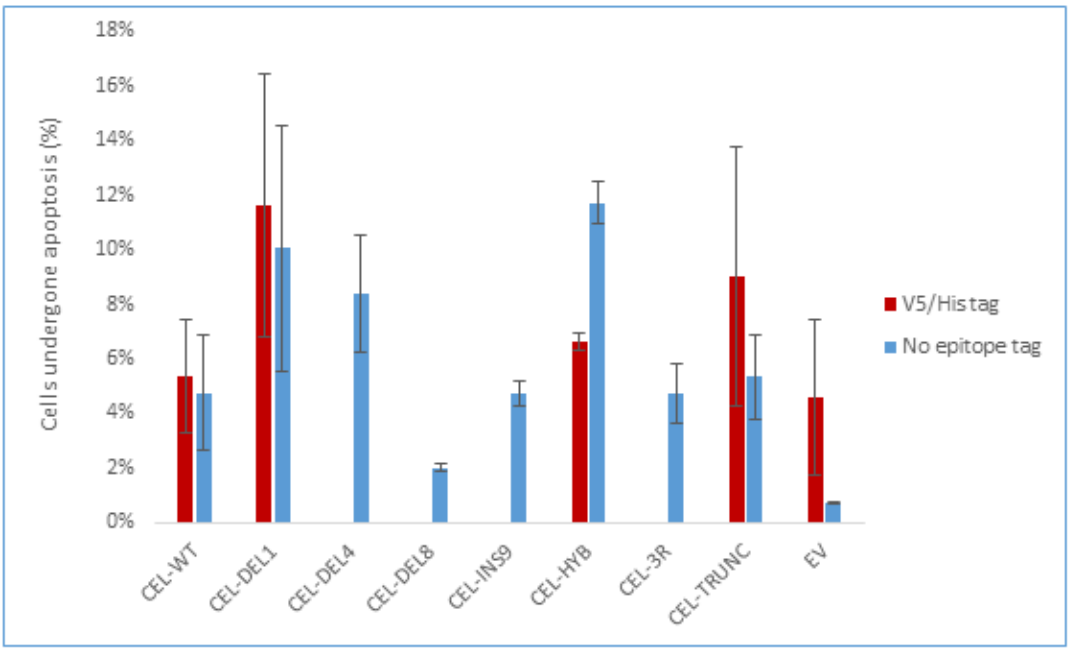



**Figure 5.8 The effect of CEL variants on apoptosis.** HEK293 cells were transiently transfected with plasmids encoding the different CEL variants. Apoptotic cells were detected using the apoptosis marker Caspase-3/7 Green Detection Agent. Confocal images of representative cells are shown in A, B, and C, and apoptotic cells are seen in green. **A)** Cells transfected with CEL-WT, CEL-DEL1, CEL-HYB, CEL-TRUNC and EV containing the V5/His tag. **B)** Cells transfected with untagged CEL-WT, CEL-DEL1, CEL-DEL4, CEL-DEL8 and EV. **C)** Cells transfected with untagged CEL-WT, CEL-HYB, CEL-TRUNC, CEL-INS9, CEL-3R and EV. **D)** The results presented are percentage of apoptotic cells of total number of cells (mean of two independent experiments ±SEM). *P*-values of significance was not performed as results were only available from two experiments.

## 5.6. De-glycosylation of the CEL DEL1, -DEL4 and -DEL8 proteins

Even though the three CEL deletion variants CEL-DEL1, -DEL4 and -DEL8 have different VNTRs, the expressed proteins are predicted to have the same molecular weight (around 73 kDa). This is illustrated in Figure 1.5 in the Introduction. However, when we analysed the expression of these three protein variants by Western blotting we observed bands with different molecular weight (see Fig. 5.6). Given that the CEL VNTR is heavily O-glycosylated, we wanted to examine if the difference in size could be due to such post-translational modification.

To investigate the glycosylation patterns of the three deletion variants, medium from HEK293 cells transfected with CEL-DEL1, CEL-DEL4 or CEL-DEL8 were subjected to acetone precipitation and denaturation, before being treated with O- and N-glycosidases. Medium from cells transfected with CEL-WT was used as a positive control. The samples were analysed by SDS-PAGE and Western blotting and the result is presented in Figure 5.9.

The experiment revealed that exposure to N-glycosidase resulted in a small shift, while combined treatment with N- and O-glycosidase resulted in a relatively larger shift for all variants (Fig. 5.9). Interestingly, the three deletion variants did not display the same molecular size after both N- and O-deglycosylation. When treated with both N- and O-glycosidase, CEL-DEL4 and DEL8 showed proteins of higher molecular weight than the CEL-DEL1 variant.

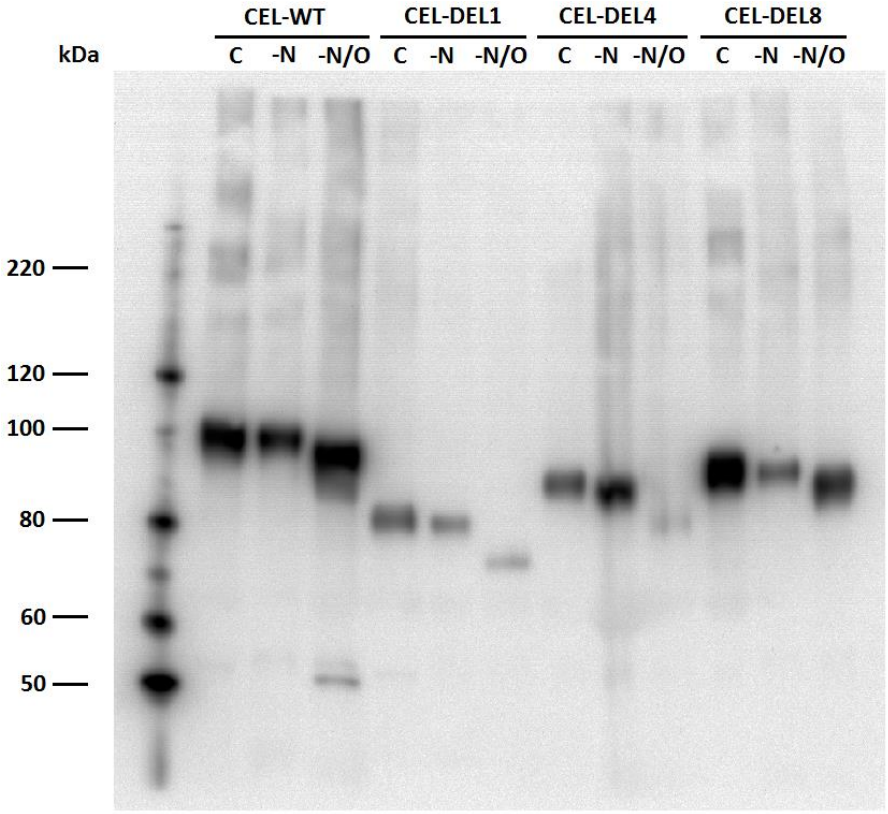



**Figure 5.9 De-glycosylation of CEL variants.** Acetone precipitated proteins from medium of transiently transfected HEK293 cells, expressing untagged CEL-WT CEL-DEL1, CEL-DEL4 and CEL-DEL8 were subjected to N-glycosidase (-N), to N- and O-glycosidase (-N/O) or no glycosidase treatment (-C). The samples were analysed with SDS-PAGE and Western blot. The proteins were detected with a CEL specific antibody (Pittsburgh). This experiment was performed once.

To examine the glycosylation pattern in more detail, we used the NetOGlyc 3.1 server to estimate the numbers of theoretical glycosylation sites of each of the variants in Figure 5.9. The VNTR of CEL-WT was predicted to have a total of 36 O-glycosylation sites, while CEL-DEL1, -DEL4, and –DEL8 showed 13, 18 and 22 sites, respectively (Figure 5.10). As well as the C-terminal O-glycosylation sites, all variants have an O-glycosylation of threonine at position 62 and a N-glycosylation of asparagine at position 210 in the globular domain. Taken together, these results indicate that the number of theoretical O-glycosylation sites increases with the length of the altered VNTR (Fig. 1.5).

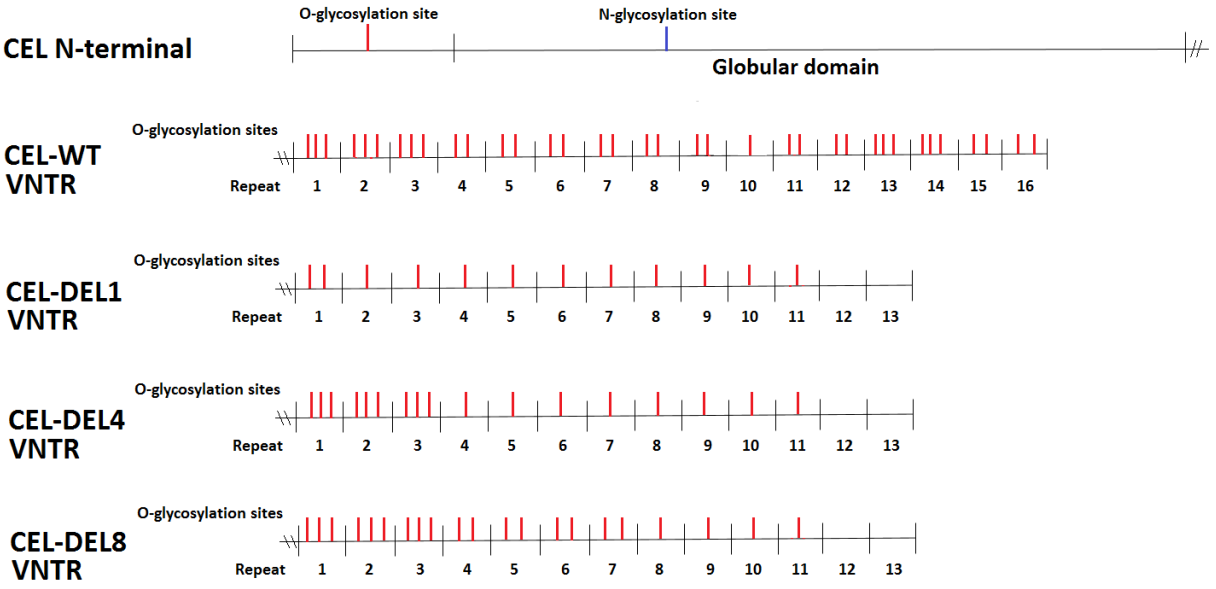



**Figure 5.10 Theoretical N- and O-glycosylation sited predicted by the NetOGlyc3.1 server.** The putative O- glycosylation sites for CEL-WT, CEL-DEL1, CEL-DEL4 and CEL-DEL8 were predicted to be a total 36, 13, 18 and 22, respectively. Also predicted was one N-glycosylation site positioned in the globular domain, position 210. O-glycosylation (red) and N-glycosylation (blue) are illustrated for N-terminal globular domain in the upper panel and for the VNTR of CEL-WT and deletion variants in the lower panel.

## 5.7. Isoelectric focusing

Due to the deletion mutation, CEL-DEL1 has a different charge than the normal protein. CEL-WT consists of 22 negatively charged amino acids (17 Asp and 5 Glu), while CEL-DEL1 has 19 positively charged amino acids (19 Arg), which results in a higher theoretical isoelectric point (*pI*) (Johansson et al., 2011). The theoretical *pI* for CEL-WT protein is predicted to be 5.1, while CEL-DEL1 has a *pI* of 9.5, not taking post-translational modifications or the epitope tag into account (ExPASy *pI* calculator). When analysing only the VNTR, CEL-WT has a *pI* of 3.3, while CEL-DEL1 has a *pI* of 11.8

We wanted to further investigate the charge and isoelectric point of CEL variants due to differences in VNTR length and composition. Isoelectric focusing was performed on medium secreted from cells expressing CEL-WT or CEL-DEL1 with and without V5/His tag. The results were analysed by Western blot using both the anti-CEL (Pittsburgh) and anti-V5 antibodies.

Although several of the bands appear as smears, CEL-WT and CEL-WT-V5/His have clear bands at approximately *pH* 6 (Fig. 5.11), which is in concordance with what was found using the *pI* predictor (*pI* 5.1). We also observed that the V5/His tag only showed a minor shift in the *pI* of CEL-WT, suggesting that the effect of the V5/His tag on the isoelectric point of CEL is minor.

CEL-DEL1 exhibited no distinct bands, even when overexposing the blot. Instead, we observed a weak smear in the blot, located mainly in the alkaline part of the *pH* gradient, which is also in concordance with what was found using the *pI* predictor ExPASy. The same result has been observed by others in our group *(El Jellas. pers. comm).* For CEL-DEL1-V5/His, there was a distinct smear ranging from *pH* 8.5 to *pH* 10 (Fig. 5.11), which could be due to differences in glycosylation patterns of the protein, leading to differently charged molecules of the same variant.

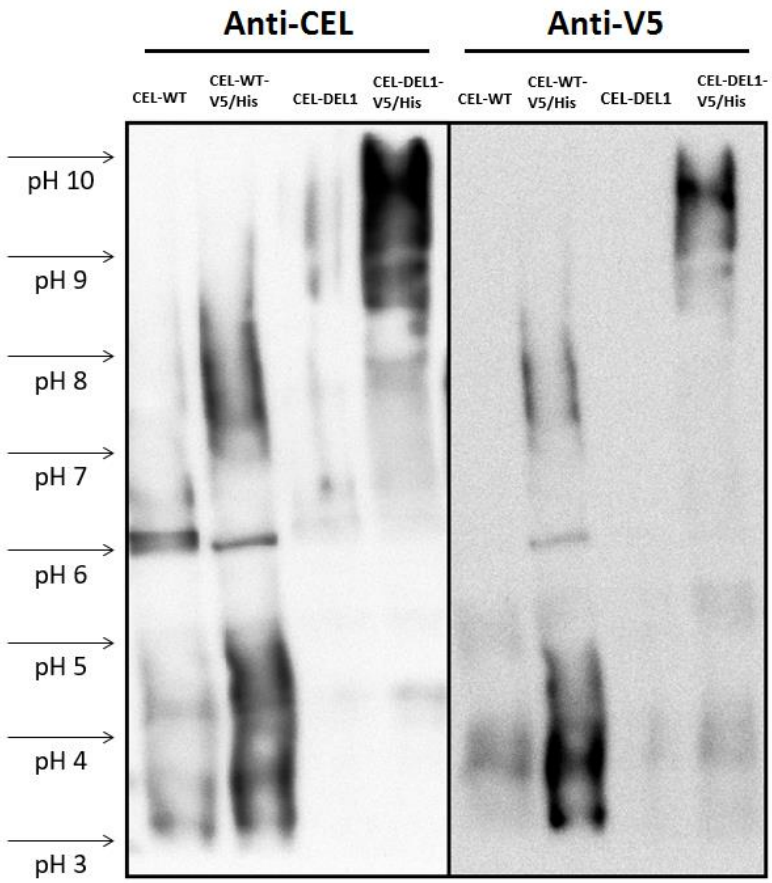



**Figure 5.11 Isoelectric focusing of V5/His tagged and untagged CEL-WT and CEL-DEL1.** Medium samples from HEK293 cells expressing the variant were subjected to isoelectric focusing by separation on a pH gradient, and analysed by Western blot using a CEL specific antibody (Pittsburgh) or the anti-V5 antibody. The experiment was performed once.

# 6. Discussion

The digestive enzyme carboxyl ester lipase has been studied for decades (Hernell and Olivecrona, 1974, Lombardo et al., 1980b, Lombardo, 2001, Hui and Howles, 2002). Two of the CEL variants used in this project, CEL-DEL1 and CEL-HYB have been analysed extensively and are known to be associated with monogenic diabetes and chronic pancreatitis, respectively (Raeder et al., 2006, Johansson et al., 2011, Tjora et al., 2013, Fjeld et al., 2015, Xiao et al., 2016). When the CEL-DEL1 variant was published in 2006, Ræder et al. also reported another disease-causing deletion variant, namely CEL-DEL4. Compared to CEL-DEL1, this mutation showed a seemingly milder phenotype of endocrine and exocrine dysfunction. Other CEL-variants normally occurring in the population were also observed, all showing varying lengths of the VNTR-region in CEL exon 11 (Raeder et al., 2006).

In the present study, we have investigated whether a V5/His tag could influence the functional properties of CEL proteins expressed in cellular models. Moreover, we have examined the effect of different CEL-VNTR lengths with regard to protein expression, secretion, modification and cell viability, to learn more about the role of the protein C-terminal in endocrine and exocrine pancreatic disease. Our main findings are discussed here. A complete overview of the variants used in this study is attached in Appendix I: CEL variants exon 11 DNA sequence and II: CEL variants exon 11 amino acid sequence.

## 6.1. Comparison of V5/His tagged- versus untagged CEL protein variants

In our previous studies analysing the disease mechanism of CEL-DEL1 and CEL-HYB in cellular systems, most results were based on the pcDNA3.1/V5-His vector from Invitrogen for protein expression (Johansson et al., 2011, Torsvik et al., 2014, Fjeld et al., 2015). This V5/His tag consists of a 14 amino acid V5 segment followed by a positively charged poly-histidine (6xHis) tag, which is fused to the C-terminal of the CEL protein.

In this study, we compared the effect of the V5/His tag on CEL expression, secretion and localisation in HEK293 cells by Western blotting and immunostaining. The Western blots clearly showed stronger bands in the lysate- and pellet fractions for untagged CEL-WT, CEL-DEL1 and CEL-TRUNC (Fig. 5.3) as compared to the tagged variants. The exception was CEL-HYB, which showed more equal expression in the same fractions (Fig. 5.4). These results were confirmed by immunostaining: stronger signals were detected in cells transfected with CEL plasmids without V5/His, except for CEL-HYB. Thus, our findings indicate that the cellular properties of the CEL proteins are influenced by the epitope tag, and that the epitope tag may indeed render CEL variants more soluble.

One study has been performed on CEL-DEL1 without V5/His tag (Xiao et al., 2016). Here, the authors observed intracellular aggregates and activation of the unfolded protein response (UPR) in transfected

cells, similar to our previous findings with the tag (Johansson et al., 2011, Xiao et al., 2016). However, they also observed impaired protein secretion and induced apoptosis, suggesting an even stronger pathogenic effect without the V5/His tag.

It is also interesting to see that CEL-HYB and CEL-TRUNC, the two variants with the shortest VNTR domains, were detected as the most abundant in the pellet fraction (Fig. 5.3, 5.4). In addition, these variants showed dramatically impaired secretion compared to CEL-WT, suggesting that the size and/or the composition of the C-terminal domains has an effect on the secretion of CEL. As for CEL-HYB, this is in line our previous report showing intracellular retention and impaired secretion of this protein (Fjeld et al., 2015). However, the fact that CEL-HYB V5/His, but not CEL-HYB was detected in the medium (Fig. 5.4), further confirms that the V5/His tag has an effect on the functional properties of CEL.

## 6.2. The effect of CEL VNTR lengths on protein expression, localisation and secretion

Next, we investigated the effect of CEL VNTR lengths (without V5/His tag) on protein expression, secretion and localisation in HEK293 cells. By cell fractionation and Western blot analysis, we found the same trend after both 48h and 72h transfection (Fig 5.6). Most interestingly, CEL-DEL1 and CEL-DEL4 were less secreted compared CEL-WT. Thus, for the first time we show reduced secretion for both CEL-MODY variants. Furthermore, we observe that variants found in the normal population are secreted similarly to the normal protein (CEL-DEL8, CEL-3R and CEL-INS9).

Another interesting finding is that the three deletion variants (DEL1, DEL4 and DEL8) migrate with different sizes in all cell fractions (Fig. 5.6). As these variants have the same theoretical size, but are likely to differ in VNTR glycosylation patterns (Fig. 5.10), suggests that the shift in band size is due to different post-modification. This is further discussed in section 6.4.

We also observe higher molecular bands for several of the CEL variants in both lysate and pellet fractions (Fig. 5.3, 5.4 and 5.6). We know that CEL-WT, CEL-DEL1 and CEL-HYB are cleared from transfected HEK293 cells by proteosomal degradation (Johansson et al., 2011, Torsvik et al., 2014) (*Fjeld. pers. comm.*), indicating that the higher molecular forms could be ubiquitinated proteins targeted for degradation. We also speculate that there could be oligomeric/dimeric forms of CEL but this needs to be further investigated.

### 6.3. The effect of CEL variants on cell viability and apoptosis in HEK293 cells

Given the pathogenic nature of CEL-DEL1, -DEL4 and -HYB, it was surprising to see that all three variants, both untagged and V5/His tagged, led to an increase in cell viability compared to CEL-WT (Fig. 5.7). Actually, of all the variants, only cells transfected with CEL-TRUNC or EV showed a reduced viability compared to the normal CEL protein.

Previous reports have found that CEL-DEL1 has a negative impact on cell viability (Johansson et al., 2011, Xiao et al., 2016). The most recent article showed that CEL-DEL1 increases cell death in transiently transfected HEK293T cells, however, the effect was most prominent after 72h and 92h transfection (Xiao et al., 2016). It has also been shown that cells may have an increase in ATP before undergoing apoptosis, perhaps due to activating all survival mechanisms (Zamaraeva et al., 2005). Thus, given that our cell viability assay relies on measuring ATP and that we analysed our cells 48 h post transfection, we speculate that our results could have been different if the measurement had been performed at a later time point.

Compared to the cell viability results, we observe a different trend when analysing our transfected cells for apoptosis. Now, elevated levels of apoptosis were observed for cells transfected with disease-causing variants (Fig. 5.8). This is in line with a previous report (Xiao et al., 2016) and suggests that programmed cell death is a result of the altered physicochemical properties of the pathogenic variants. It was interesting to observe that cells transfected with CEL-DEL8, CEL-INS9 and CEL-3R all displayed lower or similar levels of apoptosis compared to CEL-WT, indicating no toxic effect of CEL variants detected in the healthy population.

In addition, for the apoptosis assay, most variants with tag led to a higher percentage of apoptotic cells compared to the untagged variants. This leaves the question of whether the tag could be harmful in any way. One explanation could be that a higher secretion rate of the tagged variants could lead to higher levels of endocytosis, followed by induced apoptosis. The exception was CEL-HYB, which showed a higher percentage of apoptotic cells for the untagged variant. Perhaps this is caused by the strong tendency of CEL-HYB to be retained inside the cell (Fjeld et al., 2015), and that the untagged variant is more toxic to the cells. To support this hypothesis, the untagged CEL-HYB also showed reduced cell viability (Fig. 5.7) and reduced secretion (Fig. 5.4) compared to the tagged variant.

Another important observation in this study is that transient transfection in itself can be harmful to cells (Juckem, 2012), indicated by the low viability and a relative high level of apoptosis in cells transfected with EV (Fig. 5.7, 5.8).

The fact that results on viability and apoptosis show opposite trends are difficult to explain. Long term exposure of CEL variants could be resembling the late onset disease of CEL-MODY and CEL-HYB more. However, viability assays performed by our group on cells stably expressing pathogenic CEL

variants did not show reduced cell viability. It would be interesting to perform a virus-mediated transfection, which would increase the time period of expression, thus allowing us to study the effect of the variants for a longer time (Durocher et al., 2002). The advantages of using virus-mediated transfection are plentiful; it is highly efficient, easy to use, and results in sustainable transgene expression after integration in the host genome (Hacein-Bey-Abina et al., 2002, Roesler et al., 2002). However, one must keep in mind that viral transfection can result in cytotoxicity (Woods et al., 2003, Kim and Eberwine, 2010).

## 6.4. CEL VNTR sequence and length alter glycosylation and protein charge

### 6.4.1. De-glycosylation of CEL-DEL1, -DEL4 and -DEL8

Glycosylation is a very complex post-translational modification characterized by the addition of a monosaccharide, or an oligosaccharide chain in a specific manner to macromolecules such as proteins. Two predominantly classes of glycosylation exist; N- and O-glycosylation (Nelson and Cox, 2008, Varki, 2009). N-glycosylation occurs co-translationally in the ER at the consensus sequence N-X-T/S (where N is an asparagine residue, X is any residue except proline, and T/S are serine or threonine residues). O-glycosylation however, occurs post-translationally in the Golgi apparatus on serine or threonine residues. During this complex process, different glycosidases participates in assembling glycosylation structures on the protein (Varki, 2009). The different residues can be assembled in different configurations, with different types of glycosidic bonds, which subsequently provides a high amount of different structural moieties that can affect the biochemical properties of a protein (Nelson and Cox, 2008, Varki, 2009).

As mentioned before in this discussion (section 6.2), we found it very interesting that these three CEL deletion variants migrated with different sizes (Fig. 5.6), given that they all have an estimated weight of 73 kDa (Fig 1.5). However, when looking more closely into the theoretical glycosylation pattern of CEL-DEL1, -DEL4 and -DEL8, we found them to differ greatly. The number of theoretical O-glycosylation sites actually decreased with the length of the altered VNTR (Fig. 5.10). Thus, the more proximal the deletion appeared in the VNTR, the more glycosylation sites were lost in the resulting protein. The differences in the number of O-glycosylation sites, are however only predicted, since it is not yet known how many of these theoretical sites are actually occupied by glycans, or what their actual structure is.

To look more into this, we treated the different variants with N- and O-glycosidases. The results showed that CEL-DEL1, -DEL4 and -DEL8 did not result in the same molecular size after treatment with both N- and O-glycosidases (Fig. 5.9).

Based on the shifts we observed, one could speculate the more glycans the variants consist of, the fewer glycans have actually been removed. This could be due to limited access for the glycosidases when the protein is heavily modified. To support this, we observe that CEL-WT, being the most O-glycosylated form of the CEL variants with 37 sites, has a shift more similar to CEL-DEL8 (22 sites) than CEL-DEL1 (13 sites).

Given that O-glycosylation of the CEL-VNTR is thought to be important for both stability and secretion of the protein (Rogers et al., 1986, Bruneau and Lombardo, 1995, Bruneau et al., 1997), it is interesting to observe that CEL-DEL1, CEL-DEL4 and CEL-HYB, which all are low in the numbers of glycosylation sites (13, 18 and 3, respectively) are variants found less secreted (Fig. 5.4 and Fig. 5.6).

However, both N- and O-glycosylation are complex modifications, varying depending on the cell/tissue from which it is secreted (Landberg et al., 1997). Furthermore, CEL has O-glycans with an average of 7-8 sugars per O-glycan molecule, each with a high degree of structural diversity (Wang et al., 1995, Landberg et al., 1997, Landberg et al., 2000). Also, one study found that CEL in mother's milk had dramatically different compositions and levels of sialyated O-glycan and fucosylated N-glycan structures in different stages of lactation, and suggested that glycosylation of CEL depends on several factors including blood group phenotype and hormonal regulation of glycosylic transferases (Landberg et al., 2000). Thus, the role of CEL glycosylation in pancreatic disease is complex and further studies are needed to analyse this in more depth.

### 6.4.2. Isoelectric focusing

As mentioned above, the frame shift mutation in CEL-DEL1 leads to more positively charged amino acids, and an additionally predicted increase in *pI* from 5.2 to 9.5 compared to the normal protein. To investigate if the theoretical predictions of charge of CEL-WT and CEL-DEL1 are correct, we performed isoelectric focusing on these variants. Glycans may also play an important role in altering the charge, as both N- and O-glycosylation often contain the negatively charged sialic acid sugar residues (Eylar et al., 1962). The abundance of O-glycosylation of the VNTR should in theory render the C-terminal hydrophilic and charged (Wang et al., 1995).

Although no information could be obtained from untagged CEL-DEL1, CEL-DEL1-V5/His had a strong band located in the blot, albeit quite smeared, as did the other samples. These smeared bands observed reflect the complexity of working with glycosylated proteins. The glycans are often heterogeneous, which would result in differently glycosylated CEL molecules, due to micro-heterogeneity (Fig. 5.11). The glycans do not necessarily have the same conformation in each CEL molecule, and each CEL molecule may have a different pattern of glycans, which would give the protein molecules different *pI*

values, and thus present at different points along the pH gradient (Landberg et al., 1997, McKillop et al., 2004).

CEL-DEL1-V5/His is located mainly in the alkaline part of the pH gradient, in concordance with the *pI* predictor, and with results from a two dimensional gel *(El Jellas. pers. comm).* The smear may correspond to fully glycosylated proteins at the upper part, and partly glycosylated proteins at the lower part. If the CEL-DEL1 variant is less glycosylated, a possible explanation could be related to the high number of cysteine residues. It has been reported that the excess of cysteines residues that appear in the new C-terminal could lead to aggregation of the protein via formation of disulphide bridges, and subsequently large protein aggregates (Xiao et al., 2016). If this is the case, this may compromise the ability and access of glycosidases to the sites that could be exploited for glycosylation (i.e. a serine or a threonine), leading to even fewer O-glycans for CEL-DEL1. In any case, we cannot exclude the possibility of other post-translational modifications could occur simultaneously in different variants such as protein phosphorylation, or acetylation, and none, except glycosylation have been examined in this work.

For both CEL-WT and CEL-WT His/V5, the bands at approximately pH 6-8 are likely partly- and fully-glycosylated forms, as that will render the protein more acidic. The smear at pH ~ 5 may be non-glycosylated forms, correlating to the predicted *pI* values for CEL-WT (NetOGlyc3.1), with a slight shift from tagged to untagged variant. However, as the bands at approximately pH 4 can also be seen for both variants of CEL-DEL1, although weakly, one could speculate if these are non-glycosylated forms of the globular domain of the CEL variants,

Detecting the variants with the anti-V5 antibody serves as a control as it only detects the C-terminal of the variants (Johansson et al., 2011). CEL-WT-V5/His occurs in the acidic part of the pH gradient at pH 3-5 corresponding to the predicted *pI* value of the VNTR at 3.3, while CEL-DEL1-V5/His has a strong band in the more alkaline part at pH 9-10 (Fig. 5.11). The smears could be O-glycosylated forms with fewer/other glycans, as the *pI* corresponds to the *pI* of the two variants VNTR domains (Johansson et al., 2011).

It would be interesting to include the other CEL variants, especially the deletion variants, and also variants devoid of N-glycans, to conclude if the upper alkaline smear is the N-glycan. For further investigation, all CEL variants could be run on a two dimensional gel to compare both charge and size.

## 6.5. HEK293 cells as a research tool for studying CEL

The HEK293 cells are a specific cell line that originally derived from Human Embryonic Kidney cells and they are a well-known tool in cell biology research. The reasons behind this is that they are easy to culture, easy to transfect, and they proliferate rapidly. For our studies, it is also an advantage that they are secretory cells and that they do not express the CEL protein endogenously. Additionally, the cells are of human origin, which is essential when studying post-translational modifications of our CEL variants.

However, there are limitations when choosing the HEK293 cell line for CEL-related projects as these cells do not give a representative physiological presentation of how the CEL variants would behave in pancreatic cells. For example, even though they are secretory cells, the HEK293 cells do not have the characteristic zymogen granules of the acinar cells. The best option would be to use primary acinar cells, but they are hard to get hold of, difficult to obtain in long-term culture models, sensitive to work with, and often loses their secretory differentiation (Blinman et al., 2000, Lardon and Bouwens, 2005). There are two commercially available acinar cell lines, one deriving from mouse (266-6 cells) and one from rat (AR42J). Our research group have used the 266-6 cells for endocytosis experiments (Torsvik et al., 2014), however we have had no success with liposomal transfection or electroporation of the cells. Thus, we need to start using virus transfection, a method that has shown to be very successful for studying these acinar cell lines (Szmola and Sahin-Toth, 2010, Xiao et al., 2016).

# 7. Conclusions

The pathogenic variants of CEL are associated with diabetes (CEL-DEL1 and CEL-DEL4) and pancreatitis (CEL-HYB). Here, the overall objective was to understand more of the role of the CEL protein C-terminal in pancreatic disease.

The findings in this study suggest that the properties of the CEL protein variants expressed in HEK293 cells are influenced by the presence of an epitope tag, which leads to an underestimation of their propensity to accumulate and aggregate. This could be relevant for the interpretation of previous results from our group regarding the pathogenic variants CEL-DEL1 and CEL-HYB. In addition, we observed that CEL protein variants are influenced by the VNTR lengths and composition, and that O-glycosylation possibly has an important impact on CEL variant expression, secretion and cellular fate. The disease-causing protein variants CEL-DEL1, CEL-DEL4 and CEL-HYB resulted in elevated intracellular accumulation and increased apoptosis compared to the CEL-WT. Finally, we found evidence that of the three deletion variants, the first VNTR repeat contains more accessible O-glycosylation sites, resulting in differently glycosylated proteins, depending on the site of mutation.

# 8. Future perspectives

We have the following plan for further investigations:

- The experiments performed in this study will have to be repeated three times with similar results before publication.
- The *CEL* mutation carriers are heterozygous. It would therefore be interesting to study all variants investigated here when co-expressed with CEL-WT, to see the effect of the disease related variant on the normal protein.
- It would be beneficial to perform the experiments with stably transfected cells to obtain long time expression. This would better fit the disease mechanism of late onset disease.
- Further studies on the glycosylation patterns and *pI* of the different CEL variants to understand more of the role of glycans and charge in molecular interactions and disease mechanism.
- Characterization of the glycosylation-profiles of CEL-variants in duodenal juice from affected patients or healthy individuals with varying VNTR-lengths.
- It would be beneficial and more realistic in terms of physiology to investigate the CEL variants in a pancreatic acinar cell line, or in animal models. Transgenic mouse expressing human CEL-WT, CEL-DEL1 and CEL-HYB are now being made by our group.
- In addition, measuring the enzymatic activity of each variant would be interesting, as such experiments have not been performed on the novel CEL variants.

# References


AABAKKEN, L. 2016. *Bukspyttkjertelbetennelse* [Online]. Store medisinske leksikon. Available: https://sml.snl.no/bukspyttkjertelbetennelse [Accessed 21.09.2016].

ABOUAKIL, N. & LOMBARDO, D. 1989. Inhibition of human milk bile-salt-dependent lipase by boronic acids. Implication to the bile salts activator effect. *Biochim Biophys Acta,* 1004**,** 215-20.

ABOUAKIL, N., MAS, E., BRUNEAU, N., BENAJIBA, A. & 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**,** 25755-63.

ATKINSON, M. A., EISENBARTH, G. S. & MICHELS, A. W. 2014. Type 1 diabetes. *The Lancet,* 383**,** 69-82.

AUBERT-JOUSSET, E., SBARRA, V. & LOMBARDO, D. 2004. Site-directed mutagenesis of the distal basic cluster of pancreatic bile salt-dependent lipase. *J Biol Chem,* 279**,** 39697-704.

AUBERT, E., SBARRA, V., LE PETIT-THEVENIN, J., VALETTE, A. & LOMBARDO, D. 2002. Site-directed mutagenesis of the basic N-terminal cluster of pancreatic bile salt-dependent lipase. Functional significance. *J Biol Chem,* 277**,** 34987-96.

BENGTSSON-ELLMARK, S. H., NILSSON, J., ORHO-MELANDER, M., DAHLENBORG, K., GROOP, L. & BJURSELL, G. 2004. Association between a polymorphism in the carboxyl ester lipase gene and serum cholesterol profile. *Eur J Hum Genet,* 12**,** 627-32.

BJORLYKKE, Y., VETHE, H., VAUDEL, M., BARSNES, H., BERVEN, F. S., TJORA, E. & RAEDER, H. 2015. Carboxyl-ester lipase maturity-onset diabetes of the young disease protein biomarkers in secretin-stimulated duodenal juice. *J Proteome Res,* 14**,** 521-30.

BLACKBERG, L., ANGQUIST, K. A. & HERNELL, O. 1987. Bile-salt-stimulated lipase in human milk: evidence for its synthesis in the lactating mammary gland. *FEBS Lett,* 217**,** 37-41.

BLACKBERG, L., LOMBARDO, D., OLLE HERNELL, O., GUY, O. & OLIVECRONA, T. 1981. Bile-salt stimulated lipase in human milk and carboxyl ester hydrolase in pancreatic juice. Are they identical enzymes? . *Federation of european biochemical societies (FEBS) Letters,* 136**,** 284-288.

BLINMAN, T. A., GUKOVSKY, I., MOURIA, M., ZANINOVIC, V., LIVINGSTON, E., PANDOL, S. J. & GUKOVSKAYA, A. S. 2000. Activation of pancreatic acinar cells on isolation from tissue: cytokine upregulation via p38 MAP kinase. *Am J Physiol Cell Physiol,* 279**,** C1993-2003.

BOND-SMITH, G., BANGA, N., HAMMOND, T. M. & IMBER, C. J. 2012. Pancreatic adenocarcinoma. *BMJ,* 344**,** e2476.

BRUNEAU, N., DE LA PORTE, P. L., SBARRA, V. & LOMBARDO, D. 1995. Association of bile-salt-dependent lipase with membranes of human pancreatic microsomes. *Eur J Biochem,* 233**,** 209-18.

BRUNEAU, N. & 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**,** 13524-33.

BRUNEAU, N., LOMBARDO, D., LEVY, E. & BENDAYAN, M. 2000. Roles of molecular chaperones in pancreatic secretion and their involvement in intestinal absorption. *Microscopy research and technique,* 49**,** 329-345.

BRUNEAU, N., NGANGA, A., FISHER, E. A. & 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**,** 27353-61.

BRUNEAU, N., RICHARD, S., SILVY, F., VERINE, A. & LOMBARDO, D. 2003. Lectin-like Ox-LDL receptor is expressed in human INT-407 intestinal cells: involvement in the transcytosis of pancreatic bile salt-dependent lipase. *Mol Biol Cell,* 14**,** 2861-75.

BURGOYNE, R. D. & MORGAN, A. 2003. Secretory granule exocytosis. *Physiol Rev,* 83**,** 581-632.

CAILLOL, N., PASQUALINI, E., MAS, E., VALETTE, A., VERINE, A. & LOMBARDO, D. 1997. Pancreatic bile salt-dependent lipase activity in serum of normolipidemic patients. *Lipids,* 32**,** 1147-53.

CARSON, M., JOHNSON, D. H., MCDONALD, H., BROUILLETTE, C. & DELUCAS, L. J. 2007. His-tag impact on structure. *Acta Crystallogr D Biol Crystallogr,* 63**,** 295-301.

CHEN, Z., LI, Y. & YUAN, Q. 2015. Study the effect of His-tag on chondroitinase ABC I based on characterization of enzyme. *Int J Biol Macromol,* 78**,** 96-101.

COMTE, B., FRANCESCHI, C., SADOULET, M. O., SILVY, F., LAFITTE, D., BENKOEL, L., NGANGA, A., DANIEL, L., BERNARD, J. P., LOMBARDO, D. & MAS, E. 2006. Detection of bile salt-dependent lipase, a 110 kDa pancreatic protein, in urines of healthy subjects. *Kidney Int,* 69**,** 1048-55.

DALVA, M., JELLAS, K. E., 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. & MOLVEN, A. 2016. Copy number variants and VNTR length polymorphisms of the carboxyl-ester lipase (CEL) gene as risk factors in pancreatic cancer. *In:* KG JEBSEN CENTER FOR DIABETES RESEARCH, D. O. C. S., UNIVERSITY OF BERGEN (ed.).

DIPERSIO, L. P., CARTER, C. P. & HUI, D. Y. 1994. Exon 11 of the rat cholesterol esterase gene encodes domains important for intracellular processing and bile salt-modulated activity of the protein. *Biochemistry,* 33**,** 3442-8.

DIPERSIO, L. P., FONTAINE, R. N. & HUI, D. Y. 1990. Identification of the active site serine in pancreatic cholesterol esterase by chemical modification and site-specific mutagenesis. *J Biol Chem,* 265**,** 16801-6.

DOWNS, D., XU, Y.-Y., TANG, J. & 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.

DUROCHER, Y., PERRET, S. & KAMEN, A. 2002. High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res,* 30**,** E9.

EYLAR, E. H., MADOFF, M. A., BRODY, O. V. & ONCLEY, J. L. 1962. The contribution of sialic acid to the surface charge of the erythrocyte. *J Biol Chem,* 237**,** 1992-2000.

FJELD, K., BEER, S., JOHNSTONE, M., ZIMMER, C., MOSSNER, J., RUFFERT, C., KREHAN, M., ZAPF, C., NJOLSTAD, P. R., JOHANSSON, S., BUGERT, P., MIYAJIMA, F., LILOGLOU, T., BROWN, L. J., WINN, S. A., DAVIES, K., LATAWIEC, D., GUNSON, B. K., CRIDDLE, D. N., PIRMOHAMED, M., GRUTZMANN, R., MICHL, P., GREENHALF, W., MOLVEN, A., SUTTON, R. & ROSENDAHL, J. 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.

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., GRUTZMANN, R., MAYERLE, J., MOSSNER, J., RINGDAL, M., SCHULZ, H. U., SENDLER, M., SIMON, P., SZTROMWASSER, P., TORSVIK, J., SCHOLZ, M., TJORA, E., FEREC, C., WITT, H., LERCH, M. M., NJOLSTAD, P. R., JOHANSSON, S. & MOLVEN, A. 2015. A recombined allele of the lipase gene CEL and its pseudogene CELP confers susceptibility to chronic pancreatitis. *Nat Genet,* 47**,** 518-22.

HACEIN-BEY-ABINA, S., LE DEIST, F., CARLIER, F., BOUNEAUD, C., HUE, C., DE VILLARTAY, J. P., THRASHER, A. J., WULFFRAAT, N., SORENSEN, R., DUPUIS-GIROD, S., FISCHER, A., DAVIES, E. G., KUIS, W., LEIVA, L. & CAVAZZANA-CALVO, M. 2002. Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N Engl J Med,* 346**,** 1185-93.

HENGEN, P. 1995. Purification of His-Tag fusion proteins from Escherichia coli. *Trends Biochem Sci,* 20**,** 285-6.

HERNELL, O. & OLIVECRONA, T. 1974. Human milk lipases. II. Bile salt-stimulated lipase. *Biochim Biophys Acta,* 369**,** 234-44.

HIGUCHI, S., NAKAMURA, Y. & SAITO, S. 2002. Characterization of a VNTR polymorphism in the coding region of the CEL gene. *J Hum Genet,* 47**,** 213-5.

HOLCK, P. 2015. *Bukspyttkjertelen* [Online]. Store medisinske leksikon. Available: https://sml.snl.no/bukspyttkjertelen [Accessed 26.08.2016].

HOLMES, R. S. & COX, L. A. 2011. Comparative Structures and Evolution of Vertebrate Carboxyl Ester Lipase (CEL) Genes and Proteins with a Major Role in Reverse Cholesterol Transport. *Cholesterol,* 2011**,** 781643.

HOLTSBERG, F. W., OZGUR, L. E., GARSETTI, D. E., MYERS, J., EGAN, R. W. & CLARK, M. A. 1995. Presence in human eosinophils of a lysophospholipase similar to that found in the pancreas. *The Biochemical journal,* 309**,** 141-144.

HUI, D. Y. & HOWLES, P. N. 2002. Carboxyl ester lipase: structure-function relationship and physiological role in lipoprotein metabolism and atherosclerosis. *J Lipid Res,* 43**,** 2017-30.

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., MOLVEN, M. & NJØLSTAD, P. R. 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. *The Journal of Biological Chemistry,* 286**,** 34593-34605.

JUCKEM, L. 2012. Cellular Toxicity Caused by Transfection: Why is it important? Available: http://www.biocompare.com/Bench-Tips/121111-Cellular-Toxicity-Caused-by-Transfection-Why-is-it-important/ [Accessed 31.10.2016].

KAHN, R. & DAVIDSON, M. B. 2014. The reality of type 2 diabetes prevention. *Diabetes Care,* 37**,** 943-9.

KAVVOURA, F. K. & OWEN, K. R. 2014. Monogenic diabetes. *Medicine,* 42**,** 692–697.

KIDD, J. M., COOPER, G. M., DONAHUE, W. F., HAYDEN, H. S., SAMPAS, N., GRAVES, T., HANSEN, N., TEAGUE, B., ALKAN, C., ANTONACCI, F., HAUGEN, E., ZERR, T., YAMADA, N. A., TSANG, P., NEWMAN, T. L., TUZUN, E., CHENG, Z., EBLING, H. M., TUSNEEM, N., DAVID, R., GILLETT, W., PHELPS, K. A., WEAVER, M., SARANGA, D., BRAND, A., TAO, W., GUSTAFSON, E., MCKERNAN, K., CHEN, L., MALIG, M., SMITH, J. D., KORN, J. M., MCCARROLL, S. A., ALTSHULER, D. A., PEIFFER, D. A., DORSCHNER, M., STAMATOYANNOPOULOS, J., SCHWARTZ, D., NICKERSON, D. A., MULLIKIN, J. C., WILSON, R. K., BRUHN, L., OLSON, M. V., KAUL, R., SMITH, D. R. & EICHLER, E. E. 2008. Mapping and sequencing of structural variation from eight human genomes. *Nature,* 453**,** 56-64.

KIM, J. S., BONIFANT, C., BUNZ, F., LANE, W. S. & WALDMAN, T. 2008. Epitope tagging of endogenous genes in diverse human cell lines. *Nucleic Acids Res,* 36**,** e127.

KIM, T. K. & EBERWINE, J. H. 2010. Mammalian cell transfection: the present and the future. *Anal Bioanal Chem,* 397**,** 3173-8.

KISSEL, J. A., FONTAINE, R. N., TURCK, C. W., BROCKMAN, H. L. & HUI, D. Y. 1989. Molecular cloning and expression of cDNA for rat pancreatic cholesterol esterase. *Biochim Biophys Acta,* 1006**,** 227-36.

KODVAWALA, A., GHERING, A. B., DAVIDSON, W. S. & HUI, D. Y. 2005. Carboxyl ester lipase expression in macrophages increases cholesteryl ester accumulation and promotes atherosclerosis. *J Biol Chem,* 280**,** 38592-8.

KOLAR, M. J., KAMAT, S. S., PARSONS, W. H., HOMAN, E. A., MAHER, T., PERONI, O. D., SYED, I., FJELD, K., MOLVEN, A., KAHN, B. B., CRAVATT, B. F. & SAGHATELIAN, A. 2016. Branched Fatty Acid Esters of Hydroxy Fatty Acids Are Preferred Substrates of the MODY8 Protein Carboxyl Ester Lipase. *Biochemistry,* 55**,** 4636-41.

KOSOBOKOVA, E. N., SKRYPNIK, K. A. & KOSORUKOV, V. S. 2016. Overview of Fusion Tags for Recombinant Proteins. *Biochemistry (Mosc),* 81**,** 187-200.

LANDBERG, E., HUANG, Y., STROMQVIST, M., MECHREF, Y., HANSSON, L., LUNDBLAD, A., NOVOTNY, M. V. & PAHLSSON, P. 2000. Changes in glycosylation of human bile-salt-stimulated lipase during lactation. *Arch Biochem Biophys,* 377**,** 246-54.

LANDBERG, E., PAHLSSON, P., KROTKIEWSKI, H., STROMQVIST, M., HANSSON, L. & LUNDBLAD, A. 1997. Glycosylation of bile-salt-stimulated lipase from human milk: comparison of native and recombinant forms. *Arch Biochem Biophys,* 344**,** 94-102.

LARDON, J. & BOUWENS, L. 2005. Metaplasia in the pancreas. *Differentiation,* 73**,** 278-86.

LE MARECHAL, C., MASSON, E., CHEN, J. M., MOREL, F., RUSZNIEWSKI, P., LEVY, P. & FEREC, C. 2006. Hereditary pancreatitis caused by triplication of the trypsinogen locus. *Nat Genet,* 38**,** 1372-4.

LESLIE, D. R., LANSANG, M. C., COPPACK, S. & KENNEDY, L. 2013. *Diabetes,* London, Manson Publishing.

LI, D. F., FENG, L., HOU, Y. J. & LIU, W. 2013. The expression, purification and crystallization of a ubiquitin-conjugating enzyme E2 from Agrocybe aegerita underscore the impact of His-tag location on recombinant protein properties. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 69**,** 153-7.

LI, F. & HUI, D. Y. 1998. Synthesis and secretion of the pancreatic-type carboxyl ester lipase by human endothelial cells. *Biochem J,* 329 ( Pt 3)**,** 675-9.

LIDBERG, U., NILSSON, J., STROMBERG, K., STENMAN, G., SAHLIN, P., ENERBACK, S. & BJURSELL, G. 1992. Genomic organization, sequence analysis, and chromosomal localization of the human carboxyl ester lipase (CEL) gene and a CEL-like (CELL) gene. *Genomics,* 13**,** 630-40.

LIDMER, A. S., KANNIUS, M., LUNDBERG, L., BJURSELL, G. & NILSSON, J. 1995. Molecular cloning and characterization of the mouse carboxyl ester lipase gene and evidence for expression in the lactating mammary gland. *Genomics,* 29**,** 115-22.

LINDQUIST, S., BLACKBERG, L. & HERNELL, O. 2002. Human bile salt-stimulated lipase has a high frequency of size variation due to a hypervariable region in exon 11. *European Journal of Biochemistry,* 269**,** 759-767.

LODISH, H., BERK, A., KAISER, C. A., KRIEGER, M., BRETSCHER, A., PLOEGH, H., AMON, A. & SCOTT, M. P. 2013. *Molecular Cell Biology (seventh edition),* New York, W. H. Freeman and Company.

LOMBARDO, D. 2001. Bile salt-dependent lipase: its pathophysiological implications. *Biochim Biophys Acta,* 1533**,** 1-28.

LOMBARDO, D., DEPREZ, P. & GUY, O. 1980a. Esterification of cholesterol and lipid-soluble vitamins by human pancreatic carboxyl ester hydrolase. *Biochimie,* 62**,** 427-32.

LOMBARDO, D., FAUVEL, J. & GUY, O. 1980b. Studies on the substrate specificity of a carboxyl ester hydrolase from human pancreatic juice. I. Action on carboxyl esters, glycerides and phospholipids. *Biochim Biophys Acta,* 611**,** 136-46.

LOMBARDO, D. & GUY, O. 1980. Studies on the substrate specificity of a carboxyl ester hydrolase from human pancreatic juice. II. Action on cholesterol esters and lipid-soluble vitamin esters. *Biochim Biophys Acta,* 611**,** 147-55.
LOMBARDO, D., GUY, O. & FIGARELLA, C. 1978. Purification and characterization of a carboxyl ester hydrolase from human pancreatic juice. *Biochim Biophys Acta,* 527**,** 142-9.
LOOMES, K. M., SENIOR, H. E., WEST, P. M. & ROBERTON, A. M. 1999. Functional protective role for mucin glycosylated repetitive domains. *Eur J Biochem,* 266**,** 105-11.
LOOMES, K. M. & SENIOR, H. E. J. 1997. Bile salt activation of human cholesterol esterase does not require protein dimerisation. *FEBS Letters,* 495**,** 369-372.
LOWENFELS, A. B., MAISONNEUVE, P., CAVALLINI, G., AMMANN, R. W., LANKISCH, P. G., ANDERSEN, J. R., DIMAGNO, E. P., ANDREN-SANDBERG, A. & DOMELLOF, L. 1993. Pancreatitis and the risk of pancreatic cancer. International Pancreatitis Study Group. *N Engl J Med,* 328**,** 1433-7.
LV, M., WANG, J., ZHANG, J., ZHANG, B., WANG, X., ZHU, Y., ZUO, T., LIU, D., LI, X., WU, J., ZHANG, H., YU, B., WU, H., ZHAO, X., KONG, W. & YU, X. 2014. Epitope tags beside the N-terminal cytoplasmic tail of human BST-2 alter its intracellular trafficking and HIV-1 restriction. *PLoS One,* 9**,** e111422.
MADEYSKI, K., LIDBERG, U., BJURSELL, G. & NILSSON, J. 1998. Structure and organization of the human carboxyl ester lipase locus. *Mamm Genome,* 9**,** 334-8.
MALHOTRA, A. 2009. *Methods in Enzymology*, Miami, Florida, USA, Department of Biochemistry and Molecular Biology, University of Miami Miller School of Medicine.
MARTINEZ-CERON, M. C., TARGOVNIK, A. M., URTASUN, N., CASCONE, O., MIRANDA, M. V. & CAMPER, S. A. 2012. Recombinant protein purification using complementary peptides as affinity tags. *New BiotechnoIogy* 29**,** 206-210.
MASON, B. A., HE, Q., HALBROOKS, J. P., EVERSE, J. S., GUMEROV, R. D., KALTASHOV, A. I., SMITH, C. V., HEWITT, J. & MACGILLIVRAY, T. A. R. 2002. Differential effect of a His tag at the N-and C-termini: functional studies with recombinant human serum transferrin. *Biochemistry,* 41**,** 9448-9454.
MCCARROLL, S. A., KURUVILLA, F. G., KORN, J. M., CAWLEY, S., NEMESH, J., WYSOKER, A., SHAPERO, M. H., DE BAKKER, P. I., MALLER, J. B., KIRBY, A., ELLIOTT, A. L., PARKIN, M., HUBBELL, E., WEBSTER, T., MEI, R., VEITCH, J., COLLINS, P. J., HANDSAKER, R., LINCOLN, S., NIZZARI, M., BLUME, J., JONES, K. W., RAVA, R., DALY, M. J., GABRIEL, S. B. & ALTSHULER, D. 2008. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet,* 40**,** 1166-74.
MCKILLOP, A. M., O'HARE, M. M., CRAIG, J. S. & HALLIDAY, H. L. 2004. Characterization of the C-terminal region of molecular forms of human milk bile salt-stimulated lipase. *Acta Paediatr,* 93**,** 10-6.
MIYASAKA, K., OHTA, M., TAKANO, S., HAYASHI, H., HIGUCHI, S., MARUYAMA, K., TANDO, Y., NAKAMURA, T., TAKATA, Y. & FUNAKOSHI, A. 2005. Carboxyl ester lipase gene polymorphism as a risk of alcohol-induced pancreatitis. *Pancreas,* 30**,** e87-e91.
MOLVEN, A., FJELD, K. & LOWE, M. E. 2016a. Lipase Genetic Variants in Chronic Pancreatitis: When the End Is Wrong, All's Not Well. *Gastroenterology,* 150**,** 1515-8.
MOLVEN, A. & NJOLSTAD, P. R. 2011. Role of molecular genetics in transforming diagnosis of diabetes mellitus. *Expert Rev Mol Diagn,* 11**,** 313-20.
MOLVEN, A., RÆDER, H., JOHANSSON, B. B., FJELD, K., JOHANSSON, S., KULKARNI, R. N. & NJØLSTAD, P. R. 2016b. The role of the carboxyl ester lipase (CEL) gene in pancreatic disease. *Pancreatology*.
MOORE, S. A., KINGSTON, R. L., LOOMES, K. M., HERNELL, O., BLACKBERG, L., BAKER, H. M. & BAKER, E. N. 2001. The structure of truncated recombinant human bile salt-stimulated lipase reveals bile salt-independent conformational flexibility at the active-site loop and provides insights into heparin binding. *J Mol Biol,* 312**,** 511-23.
MOUNZER, R. & WHITCOMB, D. C. 2013. Genetics of acute and chronic pancreatitis. *Curr Opin Gastroenterol,* 29**,** 544-51.
MURPHY, R., ELLARD, S. & HATTERSLEY, A. T. 2008. Clinical implications of a molecular genetic classification of monogenic beta-cell diabetes. *Nat Clin Pract Endocrinol Metab,* 4**,** 200-13.
NELSON, L. D. & COX, M. M. 2008. *Lehninger Principles of Biochemistry,* New York, W. H. Freeman and Company.
NETOGLYC3.1 Center for biological sequence analysis Denmark: Technical University of Denmark.
PAJĘCKA, K., NIELSEN, W., C., HAUGE, A., ZAGANAS, I., BAK, K., L., SCHOUSBOE, A., PLAITAKIS, A. & WAAGEPETERSEN, S., H. 2014. Glutamate dehydrogenase isoforms with N-terminal (His) 6-or

FLAG-tag retain their kinetic properties and cellular localization. *Neurochemical research,* 39**,** 487-499.
PALADE, G. 1975. Intracellular Aspects of the Process of Protein Synthesis. *Science,* 189**,** 347-358.
PASQUALINI, E., CAILLOL, N., VALETTE, A., LLOUBES, R., VERINE, A. & LOMBARDO, D. 2000. Phosphorylation of the rat pancreatic bile-salt-dependent lipase by casein kinase II is essential for secretion. *Biochemical Journal* 345**,** 121-128.
POCOCK, G., RICHARDS, C. D. & RICHARDS, D. A. 2013. *Human Physiology,* United Kingdom, Oxford University Press.
PUTTE, C. V., REGAN, J., SEELEY, R., STEPHENS, T., TATE, P. & RUSSO, A. 2013. *Seeley`s anatomy and physiology*.
RAEDER, H., HALDORSEN, I. S., ERSLAND, L., GRUNER, R., TAXT, T., SOVIK, O., MOLVEN, A. & NJOLSTAD, P. R. 2007. Pancreatic lipomatosis is a structural marker in nondiabetic children with mutations in carboxyl-ester lipase. *Diabetes,* 56**,** 444-9.
RAEDER, H., JOHANSSON, S., HOLM, P. I., HALDORSEN, I. S., MAS, E., SBARRA, V., NERMOEN, I., EIDE, S. A., GREVLE, L., BJORKHAUG, L., SAGEN, J. V., AKSNES, L., SOVIK, O., LOMBARDO, D., MOLVEN, A. & NJOLSTAD, P. R. 2006. Mutations in the CEL VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. *Nat Genet,* 38**,** 54-62.
RAEDER, H., MCALLISTER, F. E., TJORA, E., BHATT, S., HALDORSEN, I., HU, J., WILLEMS, S. M., VESTERHUS, M., OUAAMARI, A. E., LIU, M., RÆDER, M. B., IMMERVOLL, H., HOEM, D., DIMCEVSKI, G., NJØLSTAD, P. R., MOLVEN, A., GYGI, S. P. & KULKARNI, R. N. 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.
RAGVIN, A., FJELD, K., WEISS, F. U., TORSVIK, J., AGHDASSI, A., MAYERLE, J., SIMON, P., NJØLSTAD, P. R., LERCH, M. M., JOHANSSON, S. & 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.
RAVI KANTH, V. & NAGESHWAR REDDY, D. 2014. Genetics of acute and chronic pancreatitis: An update. *World J Gastrointest Pathophysiol,* 5**,** 427-37.
REUE, K., ZAMBAUX, J., WONG, H., LEE, G., LEETE, T. H., RONK, M., SHIVELY, J. E., STERNBY, B., BORGSTROM, B., AMEIS, D. & ET AL. 1991. cDNA cloning of carboxyl ester lipase from human pancreas reveals a unique proline-rich repeat unit. *J Lipid Res,* 32**,** 267-76.
ROESLER, J., BRENNER, S., BUKOVSKY, A. A., WHITING-THEOBALD, N., DULL, T., KELLY, M., CIVIN, C. I. & MALECH, H. L. 2002. Third-generation, self-inactivating gp91(phox) lentivector corrects the oxidase defect in NOD/SCID mouse-repopulating peripheral blood-mobilized CD34+ cells from patients with X-linked chronic granulomatous disease. *Blood,* 100**,** 4381-90.
ROGERS, S., WELLS, R. & RECHSTEINER, M. 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science,* 234**,** 364-8.
ROSENDAHL, J., WITT, H., SZMOLA, R., BHATIA, E., OZSVARI, B., LANDT, O., SCHULZ, H. U., GRESS, T. M., PFUTZER, R., LOHR, M., KOVACS, P., BLUHER, M., STUMVOLL, M., CHOUDHURI, G., HEGYI, P., TE MORSCHE, R. H., DRENTH, J. P., TRUNINGER, K., MACEK, M., JR., PUHL, G., WITT, U., SCHMIDT, H., BUNING, C., OCKENGA, J., KAGE, A., GRONEBERG, D. A., NICKEL, R., BERG, T., WIEDENMANN, B., BODEKER, H., KEIM, V., MOSSNER, J., TEICH, N. & SAHIN-TOTH, M. 2008. Chymotrypsin C (CTRC) variants that diminish activity or secretion are associated with chronic pancreatitis. *Nat Genet,* 40**,** 78-82.
ROUDANI, S., MIRALLES, F., MARGOTAT, A., ESCRIBANO, M. J. & LOMBARDO, D. 1995. Bile salt-dependent lipase transcripts in human fetal tissues. *Biochim Biophys Acta,* 1264**,** 141-50.
SHARER, N., SCHWARZ, M., MALONE, G., HOWARTH, A., PAINTER, J., SUPER, M. & BRAGANZA, J. 1998. Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. *N Engl J Med,* 339**,** 645-52.
SHIELDS, B. M., HICKS, S., SHEPHERD, M. H., COLCLOUGH, K., HATTERSLEY, A. T. & ELLARD, S. 2010. Maturity-onset diabetes of the young (MODY): how many cases are we missing? *Diabetologia,* 53**,** 2504-8.
SIGMA-ALDRICH. *Epitope tags in protein research. Tag selection & immunotechniques* [Online]. Sigma Life Sciences Available: http://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma-Aldrich/Brochure/1/epitope-tags-in-protein-research.pdf [Accessed 29.10.2016].
STAX, M. J., KOOTSTRA, N. A., VAN 'T WOUT, A. B., TANCK, M. W., BAKKER, M., POLLAKIS, G. & PAXTON, W. A. 2012. HIV-1 disease progression is associated with bile-salt stimulated lipase (BSSL) gene polymorphism. *PLoS One,* 7**,** e32534.

SZMOLA, R. & SAHIN-TOTH, M. 2010. Pancreatitis-associated chymotrypsinogen C (CTRC) mutant elicits endoplasmic reticulum stress in pancreatic acinar cells. *Gut,* 59**,** 365-72.
TAYLOR, A. K., ZAMBAUX, J. L., KLISAK, I., MOHANDAS, T., SPARKES, R. S., SCHOTZ, M. C. & LUSIS, A. J. 1991. Carboxyl ester lipase: a highly polymorphic locus on human chromosome 9qter. *Genomics,* 10**,** 425-31.
TERPE, K. 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol,* 60**,** 523-33.
TERZYAN, S., WANG, C. S., DOWNS, D., HUNTER, B. & ZHANG, X. C. 2000. Crystal structure of the catalytic domain of human bile salt activated lipase. *Protein Sci,* 9**,** 1783-90.
TJORA, E., WATHLE, G., ENGJOM, T., ERCHINGER, F., MOLVEN, A., AKSNES, L., HALDORSEN, I. S., DIMCEVSKI, G., NJOLSTAD, P. R. & RAEDER, H. 2013. Severe pancreatic dysfunction but compensated nutritional status in monogenic pancreatic disease caused by carboxyl-ester lipase mutations. *Pancreas,* 42**,** 1078-84.
TORSVIK, J., JOHANSSON, B. B., DALVA, M., MARIE, M., FJELD, K., JOHANSSON, S., BJORKOY, G., SARASTE, J., NJOLSTAD, P. R. & MOLVEN, A. 2014. Endocytosis of secreted carboxyl ester lipase in a syndrome of diabetes and pancreatic exocrine dysfunction. *J Biol Chem,* 289**,** 29097-111.
TORSVIK, J., JOHANSSON, S., JOHANSEN, A., EK, J., MINTON, J., RAEDER, H., ELLARD, S., HATTERSLEY, A., PEDERSEN, O., HANSEN, T., MOLVEN, A. & NJOLSTAD, 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.
VARKI, A. 2009. Essentials of glycobiology. Cold Spring Harbor Laboratory Press.
VERINE, A., LE PETIT-THEVENIN, J., PANICOT-DUBOIS, L., VALETTE, A. & LOMBARDO, D. 2001. Phosphorylation of the oncofetal variant of the human bile salt-dependent lipase. identification of phosphorylation site and relation with secretion process. *J Biol Chem,* 276**,** 12356-61.
VESTERHUS, M., RAEDER, H., KURPAD, A. J., KAWAMORI, D., MOLVEN, A., KULKARNI, R. N., KAHN, C. R. & NJOLSTAD, P. R. 2010. Pancreatic function in carboxyl-ester lipase knockout mice. *Pancreatology,* 10**,** 467-76.
WANG, C. S., DASHTI, A., JACKSON, K. W., YEH, J. C., CUMMINGS, R. D. & TANG, J. 1995. Isolation and characterization of human milk bile salt-activated lipase C-tail fragment. *Biochemistry,* 34**,** 10639-44.
WANG, X., WANG, C.-S., TANG, J., DYDA, F. & ZHANG, X. C. 1997. The crystal structure of bovine bile salt activated lipase: insights into the bile salt activation mechanism. *Structure,* 5**,** 1209-1218.
WHITCOMB, D. C. 2006. Clinical practice. Acute pancreatitis. *N Engl J Med,* 354**,** 2142-50.
WHITCOMB, D. C. 2013. Genetic risk factors for pancreatic disorders. *Gastroenterology,* 144**,** 1292-302.
WHITCOMB, D. C., GORRY, M. C., PRESTON, R. A., FUREY, W., SOSSENHEIMER, M. J., ULRICH, C. D., MARTIN, S. P., GATES, L. K., AMANN, S. T., TOSKES, P. P., LIDDLE, R., MCGRATH, K., UOMO, G., POST, J. C. & EHRLICH, G. D. 1996. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nature Genetics,* 14**,** 141-145.
WHITCOMB, D. C., LARUSCH, J., KRASINSKAS, A. M., KLEI, L., SMITH, J. P., BRAND, R. E., NEOPTOLEMOS, J. P., LERCH, M. M., TECTOR, M., SANDHU, B. S., GUDA, N. M., ORLICHENKO, L., ALZHEIMER'S DISEASE GENETICS, C., ALKAADE, S., AMANN, S. T., ANDERSON, M. A., BAILLIE, J., BANKS, P. A., CONWELL, D., COTE, G. A., COTTON, P. B., DISARIO, J., FARRER, L. A., FORSMARK, C. E., JOHNSTONE, M., GARDNER, T. B., GELRUD, A., GREENHALF, W., HAINES, J. L., HARTMAN, D. J., HAWES, R. A., LAWRENCE, C., LEWIS, M., MAYERLE, J., MAYEUX, R., MELHEM, N. M., MONEY, M. E., MUNIRAJ, T., PAPACHRISTOU, G. I., PERICAK-VANCE, M. A., ROMAGNUOLO, J., SCHELLENBERG, G. D., SHERMAN, S., SIMON, P., SINGH, V. P., SLIVKA, A., STOLZ, D., SUTTON, R., WEISS, F. U., WILCOX, C. M., ZARNESCU, N. O., WISNIEWSKI, S. R., O'CONNELL, M. R., KIENHOLZ, M. L., ROEDER, K., BARMADA, M. M., YADAV, D. & DEVLIN, B. 2012. Common genetic variants in the CLDN2 and PRSS1-PRSS2 loci alter risk for alcohol-related and sporadic pancreatitis. *Nat Genet,* 44**,** 1349-54.
WITT, H., LUCK, W., HENNIES, H. C., CLASSEN, M., KAGE, A., LASS, U., LANDT, O. & BECKER, M. 2000. Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis. *Nat Genet,* 25**,** 213-6.
WITT, H., SAHIN-TOTH, M., LANDT, O., CHEN, J. M., KAHNE, T., DRENTH, J. P., KUKOR, Z., SZEPESSY, E., HALANGK, W., DAHM, S., ROHDE, K., SCHULZ, H. U., LE MARECHAL, C., AKAR, N., AMMANN, R. W., TRUNINGER, K., BARGETZI, M., BHATIA, E., CASTELLANI, C., CAVESTRO, G. M., CERNY, M., DESTRO-BISOL, G., SPEDINI, G., EIBERG, H., JANSEN, J. B., KOUDOVA, M., RAUSOVA, E., MACEK, M., JR., MALATS, N., REAL, F. X., MENZEL, H. J., MORAL, P., GALAVOTTI, R., PIGNATTI, P. F., RICKARDS, O., SPICAK, J., ZARNESCU, N. O.,

BOCK, W., GRESS, T. M., FRIESS, H., OCKENGA, J., SCHMIDT, H., PFUTZER, R., LOHR, M., SIMON, P., WEISS, F. U., LERCH, M. M., TEICH, N., KEIM, V., BERG, T., WIEDENMANN, B., LUCK, W., GRONEBERG, D. A., BECKER, M., KEIL, T., KAGE, A., BERNARDOVA, J., BRAUN, M., GULDNER, C., HALANGK, J., ROSENDAHL, J., WITT, U., TREIBER, M., NICKEL, R. & FEREC, C. 2006. A degradation-sensitive anionic trypsinogen (PRSS2) variant protects against chronic pancreatitis. *Nat Genet,* 38**,** 668-73.

WOODS, N. B., MUESSIG, A., SCHMIDT, M., FLYGARE, J., OLSSON, K., SALMON, P., TRONO, D., VON KALLE, C. & KARLSSON, S. 2003. Lentiviral vector transduction of NOD/SCID repopulating cells results in multiple vector integrations per transduced cell: risk of insertional mutagenesis. *Blood,* 101**,** 1284-9.

XIAO, X., JONES, G., SEVILLA, W. A., STOLZ, D. B., MAGEE, K. E., HAUGHNEY, M., MUKHERJEE, A., WANG, Y. & 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.

ZAMARAEVA, M. V., SABIROV, R. Z., MAENO, E., ANDO-AKATSUKA, Y., BESSONOVA, S. V. & OKADA, Y. 2005. Cells die with increased cytosolic ATP during apoptosis: a bioluminescence study with intracellular luciferase. *Cell Death Differ,* 12**,** 1390-7.

ZOU, W. B., BOULLING, A., MASAMUNE, A., ISSARAPU, P., MASSON, E., WU, H., SUN, X. T., HU, L. H., ZHOU, D. Z., HE, L., FICHOU, Y., NAKANO, E., HAMADA, S., KAKUTA, Y., KUME, K., ISAYAMA, H., PALIWAL, S., MANI, K. R., BHASKAR, S., COOPER, D. N., FEREC, C., SHIMOSEGAWA, T., CHANDAK, G. R., CHEN, J. M., LI, Z. S. & LIAO, Z. 2016. No Association Between CEL-HYB Hybrid Allele and Chronic Pancreatitis in Asian Populations. *Gastroenterology,* 150**,** 1558-1560 e5.

# Appendix

## I: CEL variants exon 11 DNA sequence

| | |
|---|---|
| Green: | V5-tag |
| Purple: | Poly-histidine tag |
| Blue: | XhoI seat |
| Bold: | Mutation (deletion/insertion) |

**CEL-WT-V5/HIS**

(1) GAG GCC ACC CCT GTG CCC CCC ACA GGG GAC TCC
(2) GAG GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC
(3) GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC
(4) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(5) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(6) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(7) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(8) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(9) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(10) GGC GCC CCC CCC GTG CCG CCC ACG GGT GAC GCC
(11) GGG CCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(12) GGC GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(13) GGG GCC CCC CCC GTG ACC CCC ACG GGT GAC TCC
(14) GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC
(15) GGG GCC CCC CCT GTG CCC CCC ACG GGT GAC TCT
(16) GAG GCT GCC CCT GTG CCC CCC ACA GAT GAC TCC

**XhoI**
AAGGAAGCTCAGATGCCTGCAGTCATTAGGTTTTCCTCGAGTCTAGAGGGCCCGCGGTTCGAAGGTA
**V5-epitope** **poly His** **STOP**
AGCCTATCCCTAACCCTCTCCTCGGTCTCG ATTCTACGCGTACCGGTCATCATCACCATCACCATTGA

**CEL-DEL1-V5/HIS (1686DelT)**

(1) GAG GCC ACC CC**G** TGC CCC CCA CAG GGG ACT CC
(2) G AGG CCA CTC CCG TGC CCC CCA CGG GTG ACT CC
(3) G AGA CCG CCC CCG TGC CGC CCA CGG GTG ACT CC
(4) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(5) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(6) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(7) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(8) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(9) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(10) G GCG CCC CCC CCG TGC CGC CCA CGG GTG ACG CC
(11) G GGC CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(12) G GCG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(13) G GGG CCC CCC CCG

**XhoI**
AAGGAAGCTCAGATGCCTGCAGTCATTAGGTTTTCCTCGAGTCTAGAGGGCCCGCGGTTCGAAGGTA
**V5-epitope** **poly His** **STOP**
AGCCTATCCCTAACCCTCTCCTCGGTCTCG ATTCTACGCGTACCGGTCATCATCACCATCACCATTGA

**CEL-HYB-V5/HIS**

(1) CGA CCA GGA GGC CAG TTC CAT GCC CTC CAC AGG
(2) GGA CTC TGA GGC CAC TCC CGT CTC CCC CGA CAG
(3) GCA ACT CCG AGT CTG CCC CCG TCC CTG CAA CGG

**XhoI**
AAGGAAGCTCAGATGCCTGCAGTCATTAGGTTTTCCTCGAGTCTAGAGGGCCCGCGGTTCGAAGGTA
**V5-epitope** **poly His** **STOP**
AGCCTATCCCTAACCCTCTCCTCGGTCTCG ATTCTACGCGTACCGGTCATCATCACCATCACCATTGA

**CEL-TRUNC-V5/HIS**

(1) GAG GCC ACC CCT

**XhoI**

AAGGAAGCTCAGATGCCTGCAGTCATTAGGTTTTCCTCGAGTCTAGAGGGCCCGCGGTTCGAAGGTA

**V5-epitope** **poly His** **STOP**

AGCCTATCCCTAACCCTCTCCTCGGTCTCG ATTCTACGCGTACCGGTCATCATCACCATCACCAT**TGA**

**CEL-WT**

(1) GAG GCC ACC CCT GTG CCC CCC ACA GGG GAC TCC
(2) GAG GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC
(3) GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC
(4) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(5) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(6) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(7) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(8) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(9) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(10) GGC GCC CCC CCC GTG CCG CCC ACG GGT GAC GCC
(11) GGG CCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(12) GGC GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(13) GGG GCC CCC CCC GTG ACC CCC ACG GGT GAC TCC
(14) GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC
(15) GGG GCC CCC CCT GTG CCC CCC ACG GGT GAC TCT
(16) GAG GCT GCC CCT GTG CCC CCC ACA GAT GAC TCC

**STOP**

AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT **TGA**

**CEL-DEL1 (1686DelT)**

(1) GAG GCC ACC CC**G** TGC CCC CCA CAG GGG ACT CC
(2) G AGG CCA CTC CCG TGC CCC CCA CGG GTG ACT CC
(3) G AGA CCG CCC CCG TGC CGC CCA CGG GTG ACT CC
(4) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(5) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(6) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(7) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(8) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(9) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(10) G GCG CCC CCC CCG TGC CGC CCA CGG GTG ACG CC
(11) G GGC CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(12) G GCG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(13) G GGG CCC CCC CCG

**STOP**

AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT **TGA**

**CEL-DEL4 (1785DelC)**

(1) GAG GCC ACC CCT GTG CCC CCC ACA GGG GAC TCC
(2) GAG GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC
(3) GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC
(4) GGG GCC CCC CC**G** TGC CGC CCA CGG GTG ACT CC
(5) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(6) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(7) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(8) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(9) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(10) G GCG CCC CCC CCG TGC CGC CCA CGG GTG ACG CC
(11) G GGC CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(12) G GCG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(13) G GGG CCC CCC CCG

**STOP**

AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT **TGA**

**CEL-DEL8 (1917DelC)**

(1) GAG GCC ACC CCT GTG CCC CCC ACA GGG GAC TCC
(2) GAG GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC
(3) GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC
(4) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(5) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(6) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(7) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(8) GGG GCC CCC CC**G** TGC CGC CCA CGG GTG ACT CC
(9) G GGG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(10) G GCG CCC CCC CCG TGC CGC CCA CGG GTG ACG CC
(11) G GGC CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(12) G GCG CCC CCC CCG TGC CGC CCA CGG GTG ACT CC
(13) G GGG CCC CCC CCG

**STOP**

AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT **TGA**

**CEL-INS9 (1952InsT)**

(1) GAG GCC ACC CCT GTG CCC CCC ACA GGG GAC TCC
(2) GAG GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC
(3) GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC
(4) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(5) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(6) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(7) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(8) GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC
(9) GGG GCC CCC CCC GT**T** GCC GCC CAC GGG

**STOP**

AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT **TGA**

**CEL-3R**

(1) GAG GCC ACC CCT GTG CCC CCC ACA GGG GAC TCC
(2) GAG GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC
(3) GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC

**STOP**

AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT **TGA**

**CEL-HYB**

(1) CGA CCA GGA GGC CAG TTC CAT GCC CTC CAC AGG
(2) GGA CTC TGA GGC CAC TCC CGT CTC CCC CGA CAG
(3) GCA ACT CCG AGT CTG CCC CCG TCC CTG CAA CGG

**STOP**

CGG AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT TCC **TGA**

**CEL-TRUNC (V562X)**

(1) GAG GCC ACC CCG

**STOP**

AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT **TGA**

# II: CEL variants exon 11 amino acid sequence

| | |
|---|---|
| Yellow: | Signalling peptide |
| Blue: | VNTR region |
| Purple: | V5-tag |
| Green: | Poly-histidine tag |
| T: | Threonine position 62 – O-glycosylation |
| N: | Asparagine position 210 – N -glycosylation |
| Underlined and bold: | VNTR O-glycosylation sites |

**CEL-WT-V5/His**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLTYLALPTVTDQEA**T**PVPPTGD**S**EA**T**PVPP**T**GD**S**E**T**APVPP**T**GD**S**GAPPVP
P**T**GD**S**GAPPVPP**T**GD**S**GAPPVPP**T**GD**S**GAPPVPP**T**GD**S**GAPPVPP**T**GD**S**GAPPVPP**T**GD**S**GAPPVPP**T**GDAGPPPVPP**T**GD**S**GAPPVP
P**T**GD**S**GAPPV**T**P**T**GD**S**E**T**APVPP**T**GD**S**GAPPVPP**T**GD**S**EAAPVPP**T**DD**S**KEAQMPAVIRFSSSLEGPRFEGKPIPNPLLGLDSTRTGH
HHHHHH

**CEL-DEL1-V5/His**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLTYLALPTVTDQEA**T**PCPPQG**T**PRPLPCPPRV**T**PRPPPCRPRV**T**PGPPPCRP
RV**T**PGPPPCRPRV**T**PGPPPCRPRV**T**PGPPPCRPRV**T**PGPPPCRPRV**T**PGPPPCRPRV**T**PAPPPCRPRV**T**PGPPPCRPRV**T**PAPPPCRPRV
**T**PGPPPSSSLEGPRFEGKPIPNPLLGLDSTRTGHHHHHHH

**CEL-HYB-V5/His**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLIYLALPTVTDQEASSMPS**T**GDSEA**T**PVSPDRQLRVCPRPCNGSSSLEGPR
FEGKPIPNPLLGLDSTRTGHHHHHH

**CEL-TRUNC-V5/His**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLIYLALPTVTDQEATSSSLEGPRFEGKPIPNPLLGLDSTRTGHHHHHH

**CEL-WT**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLIYLALPTVTDQEATPVPPTGDSEATPVPPTGDSETAPVPPTGDSGAPPVP
PTGDSGAPPVPPTGDSGAPPVPPTGDSGAPPVPPTGDSGAPPVPPTGDSGAPPVPPTGDSGAPPVPPTGDAGPPPVPPTGDSGAPPVP
PTGDSGAPPVTPTGDSETAPVPPTGDSGAPPVPPTGDSEAAPVPPTDDSKEAQMPAVIRF

**CEL-DEL1**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLIYLALPTVTDQEATPCPPQGTPRPLPCPPRVTPRPPPCRPRVTPGPPPCRP
RVTPGPPPCRPRVTPGPPPCRPRVTPGPPPCRPRVTPGPPPCRPRVTPGPPPCRPRVTPAPPPCRPRVTPGPPPCRPRVTPAPPPCRPRV
TPGPPPKEAQMPAVIRF

**CEL-DEL4**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLIYLALPTVTDQEATPVPPTGDSEATPVPPTGDSETAPVPPTGDSGAPPCR
PRVTPGPPPCRPRVTPGPPPCRPRVTPGPPPCRPRVTPGPPPCRPRVTPGPPPCRPRVTPAPPPCRPRVTPGPPPCRPRVTPAPPPCRPR
VTPGPPPKEAQMPAVIRF

**CEL-DEL8**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVGP
LGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNPL
FWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLIYLALPTVTDQEATPVPPTGDSEATPVPPTGDSETAPVPPTGDSGAPPVP
PTGDSGAPPVPPTGDSGAPPVPPTGDSGAPPVPPTGDSGAPPCRPRVTPGPPPCRPRVTPAPPPCRPRVTPGPPPCRPRVTPAPPPCRP
RVTPGPPPKEAQMPAVIRF

**CEL-INS9**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLIYLALPTVTDQEATPVPPTGDSEATPVPPTGDSETAPVPPTGDSGAPPVP
PTGDSGAPPVP PTGDSGAPPVPPTGDSGAPPVPPTGDSGAPPVPPTGDS GAPPRAAHGKEAQMPAVIRF

**CEL-3R**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLIYLALPTVTDQEATPVPPTGDSEATPVPPTGDSETAPVPPTGDSKEAQM
PAVIRF

**CEL-HYB**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLIYLALPTVTDQEASSMPSTGDSEATPVSPDRQLRVCPRPCNGKEAQMPA
VIRF

**CEL-TRUNC**

MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFK
KRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEIATRGNVIVVTFNYRVG
PLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNP
LFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNN
MDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQH
RANAKSAKTYAYLFSHPSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMGDSAVPTH
WEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLIYLALPTVTDQEATPKEAQMPAVIRF