Published in final edited form as:

Diabetes Res Clin Pract. 2017 September ; 131: 82–90. doi:10.1016/j.diabres.2017.06.023.

Functional analysis of *PCSK2* coding variants: a founder effect in the Old Order Amish population

Alexandra Winters^{a,b,#,1}, Bruno Ramos-Molina^{a,2,#}, Timothy S. Jarvela^a, Laura Yerges-Armstrong^b, Toni Pollin^b, and Iris Lindberg^{a,*}

^aDepartment of Anatomy and Neurobiology, University of Maryland School of Medicine, Baltimore, Maryland 21201

^bDivision of Endocrinology, Diabetes, and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland 21201

Abstract

Aims—In humans, noncoding variants of *PCSK2*, the gene encoding prohormone convertase 2 (PC2), have been previously associated with risk for and age of onset of type 2 diabetes (T2D). The aims of this study were to identify coding variants in *PCSK2*; to determine their possible association with glucose handling; and to determine functional outcomes for coding variants in biochemical studies.

Methods—Exome-wide genotyping was performed on 1725 Old Order Amish (OOA) subjects. *PCSK2* coding variants were tested for association with diabetes-related phenotypes. *In vitro* analyses using transfected human PC2-encoding constructs were performed to determine the impact of each mutation on PC2 activity.

Results—We identified 10 rare missense coding variants in *PCSK2* in various genomic databases. R430W (rs200711626) is greatly enriched in the OOA population (MAF 4.3%). This variant is almost twice as common (MAF 7.4%) in OOA individuals with T2D as in OOA individuals with normal or with normal/impaired glucose tolerance (MAF 3.9% and 2.9%, respectively; $p = 0.25$ and $p = 0.10$). *In vitro* experiments revealed a broadening of the pH optimum for the R430W variant, which may result in increased activity against *PCSK2* substrates.

Conclusions—Although the association of the R430W variation with T2D in the OOA population did not reach significance, based upon the broadened pH profile of R430W PC2, we

*To whom correspondence should be addressed: Dr. Iris Lindberg, Department of Anatomy and Neurobiology, Department of Anatomy and Neurobiology, School of Medicine, University of Maryland, Baltimore, Maryland 21201, ilindberg@som.umaryland.edu.

¹Current address: Division Allergy and Clinical Immunology, Johns Hopkins University, 3A.62, 5501 Hopkins Bayview Circle, Baltimore, Maryland 21224

²Current address: Department of Endocrinology and Nutrition, Institute of Biomedical Research in Malaga (IBIMA) and University of Malaga, Virgen de la Victoria University Hospital, Campus Teatinos S/N, 29010 Malaga, Spain; CIBER Physiopathology of Obesity and Nutrition (CIBEROBN), Institute of Health Carlos III, Madrid, Spain.

[#]These authors contributed equally to this manuscript

Conflicts of interest: None.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

speculate that the presence of this substitution may result in altered processing of *PCSK2* substrates, ultimately leading to increased conversion to diabetes.

Keywords

prohormone convertase 2; *PCSK2*; diabetes; glucagon; glucose

1. Introduction

Prohormone processing of large precursors to bioactive species represents a carefully controlled series of catalytic events that culminates in the secretion of appropriate amounts of peptide hormones into the bloodstream. Any disturbance in the proteolytic processing of prohormones can result in endocrine pathology. For example, obesity, diabetes, and gastrointestinal dysfunction are associated with human and mouse prohormone convertase 1/3 (PC1/3) mutations and polymorphisms [1, 2]; reviewed in [3, 4]; and loss of the trimming carboxypeptidase E (CPE) results in obesity (= *fat/fat* mice) [5]. A rare human CPE variant with reduced activity, R283W, is associated with earlier age of onset of diabetes [6], while other CPE variants are associated with increased body mass index [7]. In mice, complete knockout of the prohormone convertase PC2 in the C57Bl/6j strain results in hypoglycemia due to the total absence of glucagon [8, 9]; PC2 knockout in the 129Sv background results in even larger hypoglycemic effects [10]. These mouse studies indicate that unlike other processing enzyme mutations, loss of PC2 activity may be associated with protection from diabetes and obesity rather than with susceptibility [8, 11, 12].

Increasing evidence from genome-wide association studies suggests that at least 9 common variants in the gene encoding human PC2, *PCSK2*, are associated with diabetes risk and related traits in human populations ([13–15]; summarized in [16]). The *PCSK2* intronic variant rs2208203 was associated with lower fasting glucagon levels, as well as reduced insulin secretion, in a Scandinavian population [17]. A study of African-Americans with type 2 diabetes (T2D), showed association between diabetes and four intronic *PCSK2* variants [18]. Work in a Chinese population confirmed association between rs2021785 and T2D risk, and a study in a Japanese population also found clear association between an intronic *PCSK2* microsatellite and diabetes risk [13, 14]. In a longitudinal study, Chang et al. [16] found that individuals carrying certain intronic *PCSK2* variants (rs6044695 and rs2284912) had a lower incidence of diabetes during a 5-year follow-up, suggesting a protective effect of these particular variants (or linked genes). The repeated association between *PCSK2* variants and diabetes risk, either positive or negative, in a variety of human populations suggests that differences in PC2 activity, potentially affecting peptide hormone cleavage and thus glycemic control, may represent an important factor in the progression to T2D. However, the underlying biochemical mechanism for this increased risk- increased or decreased PC2 function- is as yet unclear, since no exonic variants of *PCSK2* have been studied thus far.

Based on the mouse data showing that loss of *PCSK2* is protective against the development of diabetes, as well as the human study showing a similar protective effect of the intronic variants rs6044695 and rs2284912, we hypothesized that coding variants associated with

decreased enzyme activity might be beneficial, i.e. associated with protection from diabetes, while coding variants associated with increased activity might be deleterious, i.e. associated with increased conversion to T2D. However, human *PCSK2* exonic variants are extremely rare, with the most frequent SNPs exhibiting minor allele frequencies (MAF) of less than 0.05% (1000Genomes and Exome Aggregation Consortium [ExAC] databases). Given this rarity, no associations have been found between any *PCSK2* coding variants and diabetes or related traits thus far.

In the work presented below, we have identified a *PCSK2* coding variant, R430W, which is present at a much higher frequency in the Old Order Amish (OOA) population than in the general population, and have evaluated its association with T2D and diabetes-related traits. Furthermore, in an effort to determine whether coding mutations cause a loss or gain of function, we have enzymatically characterized this variant as well as three other rare *PCSK2* missense coding variants.

2. Methods

2.1. Study Population

The Lancaster OOA are a genetically well-defined founder population of European descent that immigrated to the US from Western Europe over a 50-year period starting in 1727. Today's Lancaster Amish population of >30,000 individuals are descended from approximately 200 of these families. Because the OOA grew from a small number of founders with few outside individuals joining the population, certain variants that are rare in outbred populations have become common in the OOA, providing unique opportunities for genetic studies. In addition, the OOA have a high parity with an average sibship of 6–7; a relatively homogenous lifestyle; and a willingness to participate in research that make them an ideal population for genetic studies.

The Amish Family Diabetes Study (AFDS) was a study begun in 1995 designed with the aim of identifying genetic determinants of T2D and related traits [19]. Probands were OOA individuals with previously diagnosed T2D diagnosed between 35 and 65 years of age. These individuals and their first- and second-degree relatives of at least 18 years of age were recruited. Oral glucose tolerance tests (OGTTs; 75g) were only performed on individuals without prevalent diabetes. The AFDS consisted of two phases. In the first phase (n=904) all individuals without prevalent diabetes in the study underwent OGTTs. However, in phase 2 (n=383) the focus shifted to enrolling as many siblings as possible in order to perform transmission disequilibrium tests, which rely only on affected individuals. For this reason, in the second phase, participants had only a random glucose measurement to determine a binary outcome of T2D or unknown status. Thus, the quantitative OGTT data presented here includes only individuals from phase 1 of the AFDS, while the diabetes comparison includes all individuals from Phase 1 and diabetic individuals from phase 2. Diabetes, impaired glucose tolerance (IGT), and normal glucose tolerance (NGT) statuses were assigned according to American Diabetes Association criteria.

2.2. Variant Identification

Exome-wide genotyping was performed on 1725 Amish subjects using the Illumina Human Exome BeadChip (Illumina Inc., San Diego, CA). Genotyping was performed according to the manufacturer's instructions and genotypes were assigned using Illumina GenomeStudio software (Illumina Inc.). The Human Exome Chip *PCSK2* coding variants were examined for frequency in the OOA to determine if any variants are enriched in this population. In addition, whole exome data were analyzed for the presence of *PCSK2* coding variants in 154 individuals from a prior sequencing study of OOA individuals chosen for extremes in certain traits including platelet aggregation and cholesterol levels.

2.3. Genotyping

The *PCSK2* R430W variant was genotyped using a custom TaqMan assay, and HumanCore 500k Exome Chip. Of the AFDS participants with phenotype data for at least one of the traits analyzed here, there were 232 individuals with only TaqMan genotypes, 86 with only HumanCore 500k Exome Chip genotypes, and 914 individuals genotyped by both of these platforms. Of the individuals genotyped on both platforms, there were three individuals with discordant genotypes, a rate of 0.3%. These three individuals were excluded from analyses. In total, 1229 OOA individuals were analyzed. As noted above, OGTTs were only performed on individuals without prevalent diabetes in the first phase of the study; additionally, individuals diagnosed with diabetes by the OGTT were included as cases but excluded from the quantitative trait analysis. Finally, OGTT data were only available for 877 individuals.

2.4. Analysis

The R430W variant was evaluated for association with diabetes and with ten diabetes-related quantitative traits in the OOA. Insulin and HbA1c were analyzed as log-transformed traits, to correct for non-normal distribution. Analyses were performed using a mixed model that assesses the additive effect of 0, 1, or 2 copies of the variant and was adjusted for age at inclusion in the study, sex, and relatedness as modeled by a relationship matrix. All analyses were performed in MMAP (mixed models analysis for pedigrees and populations), a mixed model program that allows for a wide range of covariance structures. This program was chosen to allow the use of a relationship matrix to control for relatedness of the OOA population.

2.5. Expression vector construction and mutagenesis

The generation of a human PC2 (hPC2) wild-type expression vector (pcDNA3.1/hPC2) was previously described in Blanco *et al.* [20]. The pcDNA3.1/hPC2 vector was used as a cDNA template for site-directed mutagenesis of four rare coding variants present in 1000Genomes and ExAC: A77D, A267T, R430W and M525V. Mutagenesis was performed by GenScript (Piscataway, NJ); all constructs were verified by full insert sequencing. A eukaryotic expression plasmid encoding human proglucagon was obtained from FivePrime Therapeutics (San Francisco, CA).

2.6. Cell culture

CHO cells stably expressing rat 21 kDa 7B2 (CHO-7B2) via dihydrofolate reductase-coupled amplification were generated as previously described for proenkephalin [21], except that additional 7B2 was expressed in the methotrexate-amplified cells by supertransfection with a hygromycin-resistant 21 kDa 7B2-encoding vector and selection for the best expressor via Western blotting. Cells were grown in medium lacking nucleosides (Alpha Minus MEM medium; Invitrogen) containing 50 μ M methotrexate, 10% well-dialyzed fetal bovine serum (FBS), 100 μ g/ml hygromycin, 250 μ g/ml G418, 10 mM glucose, and 4 mM glutamine. Rat pituitary tumor GH4C1 cells were obtained from ATCC and were maintained in high glucose DMEM, supplemented with 10% FBS, in humidified 5% CO₂ at 37°C.

2.7. Fluorogenic PC2 enzyme assay

CHO-7B2, seeded in 12-well plates, were transfected with plasmids encoding different PC2 vectors (0.5 μ g/well) using FuGENE HD (Promega). After 24 h of incubation in 0.5 mL growth medium, serum-free medium (Opti-MEM; Invitrogen) containing 100 μ g/mL bovine aprotinin (Desert Biologicals) was added to each well. Cells and media were harvested the next day.

The enzymatic activity of secreted PC2 was measured in triplicate in 50 μ L reactions in a 96-well polypropylene plate containing 25 μ L of conditioned medium with final concentrations of 200 μ M fluorogenic substrate (pERTKR-aminomethylcoumarin; Peptides International), 100 mM sodium acetate (pH 5), 2 mM CaCl₂, 0.1% Brij 35, and a protease inhibitor cocktail (final concentrations: 1 μ M pepstatin, 0.28 mM N-p-tosyl-L-phenylalanine chloromethyl ketone, 10 μ M trans-epoxysuccinyl-L-leucylamido(4-guanidino)butane, and 0.14 mM N- α -tosyl-L-lysine chloromethyl ketone; all from Sigma). Reaction mixtures were incubated at 37°C and fluorescence measurements (380 nm excitation, 460 nm emission) were taken under kinetic conditions every min for 180 min in a SpectraMax M2 Microplate Reader. Experiments were repeated three times with similar results.

For pH experiments, CHO-7B2 cells were seeded in 6-well plates and co-transfected with plasmids encoding PC2s (or a control LacZ vector) at 3 μ g/well together with cDNA encoding 7B2-21kDa (2 μ g/well) using Xfect™ Transfection Reagent (Clontech) in triplicate. The inclusion of 7B2-encoding cDNA greatly increased enzyme activity, possibly because of insufficient 7B2 expression by CHO-7B2 cells due to the imbalance between the overexpression methods used for proPC2 (dihydrofolate reductase-coupled amplification) and 7B2 (antibiotic resistance). Overnight-conditioned Opti-MEM obtained from control and variant *PCSK2*-transfected CHO cells was diluted 1:1 with a 2x buffer mix (400 μ M fluorogenic substrate, 200 mM Tris, 200 mM sodium acetate, 4 mM CaCl₂, 0.2% Brij 35, and a 2x protease inhibitor cocktail (as described above), pre-adjusted to the desired pH. To fit an entire assay on one plate, a 384-well plate was used, with reactions scaled down to 20 μ L. The entire experiment was carried out independently 4 times, with similar results. Results are the means and SD of the four independent experiments.

2.8. Western blotting

Cells were lysed by adding Laemmli sample buffer directly to PBS-washed cell monolayers. Cell and media samples were subjected to SDS-PAGE (10% gels) followed by Western blotting. For the detection of PC2, a combination of two rabbit polyclonal anti-PC2 antisera directed to the mature amino terminus and the propeptide (LSU7 and LSU26) was used as previously described [22–24]. For 7B2 detection we used rabbit antisera directed against mouse 7B2 (13BF) [25] at a 1:1000 dilution. Visualization of immunoreactive proteins was accomplished using the Clarity enhanced chemiluminescent Western blotting substrate (Bio-Rad), and images were obtained with a ChemiDoc MP imaging system (Bio-Rad).

2.9. Glucagon radioimmunoassay

GH4C1 cells were plated in six-well plates to 60%–70% confluence, and each well was transfected with 1 µg of human proglucagon-encoding vector in the presence of 1 µg of either PC2-encoding vector (wild type or mutant) or with empty vector, using FuGENE (Promega). One day after transfection, 1 ml of Opti-MEM containing 100 µg/mL bovine aprotinin was added to each PBS-washed well. Cells were incubated for an additional 24 h before the conditioned medium was harvested. Media samples were centrifuged at 3000xg for 10 min and supernatants were kept frozen for later analysis.

The thawed samples were then assayed using a commercial glucagon radioimmunoassay (GL-32K; Millipore, Billerica, MA) according to the manufacturer's instructions. Radioactivity was determined using a Wallac 1470 Wizard gamma counter (PerkinElmer Life and Analytical Sciences, Waltham, MA). The experiment was repeated three times with similar results. This glucagon RIA exhibits less than 0.1% cross-reactivity with oxyntomodulin, the major glucagon-containing fragment; we have successfully used this assay to examine the liberation of glucagon from proglucagon in past work [26, 27].

3. Results

3.1. Identification of missense PCSK2 coding variants

While intronic *PCSK2* variants are present at minor allele frequencies (MAF) of up to 8% [16], exonic variants are extremely rare in the outbred population. Fourteen putative missense variants were identified in the NHLBI Exome Sequencing Project (ESP), T2D-Genes, and 1000Genomes phase 1 databases (C18G, R47S, E52K, A77D, R81C, D123N, N221D, P244Q, A267T, A385V, R430W, M525V, T569I, and Q595L). However, only ten of these (C18G, R47S, A77D, R81C, N221D, A267T, R430W, M525V, T569I, and Q595L) were found in the ExAC database, an online database containing genotypes from ~60,000 individuals compiled from a variety of sequencing efforts; this suggests that at least some of the initially identified variants represent sequencing errors. For example, the variant rs188499534 (A385V) reported in the 1000Genomes phase 1 database at a MAF of 12.7%, is clearly a sequencing artifact, as it was not identified in several other genomic databases including 1000Genomes phase 3, which contains more accurate sequencing data. All ten ExAC coding variants are very rare in all known outbred populations (MAFs between 0.0016% and 0.076%) (Table 1). The location of these ten variants, which span the five preproPC2 domains, is depicted in Fig. 1A.

The “genomic evolutionary rate profiling” (GERP) score, a measure of conservation using estimates of evolutionary constraint [28], indicates that the wild type alleles of all variants are conserved, given GERP scores between 3.3 and 5.93 (where values greater than 2 suggest evolutionary constraint), suggesting that all of these residues are important for PC2 function. The “sorting intolerant from tolerant” (SIFT) score is a homology-based tool that categorizes missense changes based on their likelihood to cause a phenotypic effect [29]; six *PCSK2* variants are also rated as deleterious, with a score of 0.02 or less. PolyPhen2 is a third algorithm that predicts the outcome of a nonsynonymous substitution [30]; here, A267T and R430W are rated as “probably damaging”, with scores of 0.99 and 0.97, respectively. The concordance of three prediction algorithms indicate that there is a large likelihood that at least three of the ten substitutions, A267T, R430W, and M525V, will alter PC2 activity.

With regard to the OAA population, direct exome sequencing of 154 OOA individuals showed that 13 copies of the rs200711626 variant allele encoding the missense mutation R430W were found, which represents a MAF of 4.2%. This is greatly enriched compared to other populations, indicating a likely founder effect; only 16 copies are reported in 60,000 individuals in the Exome Aggregation Consortium (equivalent to an MAF of 0.013%; Table 1). None of the other ten coding variants were found in the OAA.

Of the ten rare coding variants confirmed in the ExAc database, we selected four for further analysis which are predicted to have deleterious effects and which span different domains (Table 1). The enzymatic activity against a fluorogenic substrate, and the protein expression of these four *PCSK2* variants, A77D, A267T, R430W and M525V, is shown in Fig. 1B. The expression of 7B2, an obligate proPC2 binding protein [25], was also assessed by Western blotting in order to ensure that similar levels were present for each variant. Whereas the A77D, R430W and M525V variant proPC2 proteins were secreted normally by CHO-7B2 cells and exhibited catalytic activity similar to the wild-type protein after auto-activation in the reaction buffer, the substitution of Ala at residue 267 by Thr critically affected its secretion (Fig. 1C). As a result, neither protein nor enzymatic activity was detected in the conditioned medium for this variant (Fig. 1B). In agreement, higher PC2 protein levels were found in the A267T cell lysates, indicative of intracellular accumulation, most likely as a result of misfolding (Fig. 1C).

In order to assay PC2 variant activity against an actual prohormone substrate, we also evaluated the cleavage of co-transfected proglucagon to glucagon; the liberation of this peptide is known to be dependent on PC2 expression [8, 31]. For these studies, we used the somatotroph line GH4C1. Unlike CHO-7B2 cells, this regulated neuroendocrine cell line contains an acidic secretory granule compartment in which intracellular conversion of proPC2 to PC2, and therefore substrate cleavage, can occur. Again, all *PCSK2* variants except A267T were secreted normally (Fig. 1D); the fact that no A267T PC2 or proPC2 was found in the medium taken from GH4C1 cells may reflect better control of basal secretion in GH4C1 cells than in CHO-7B2 cells (see Fig. 1B vs. 1D). No differences in glucagon production from proglucagon were detected between the A77D, R430W and M525V variants and the wild-type enzyme (Fig. 1D). In contrast, glucagon synthesis was radically reduced in the presence of the A267T variant, supporting the idea that the A267T proPC2

protein does not efficiently traffic through the secretory pathway to secretory granules, the low-pH subcellular compartment in which PC2 is activated and is known to act. Interestingly, in GH4C1 cells, the A77D variant protein exhibited an additional higher molecular mass band. Western blot data using propeptide-specific antiserum [23] support the idea that this band corresponds to a PC2 species containing propeptide (data not shown). Secondary site propeptide cleavage is thought to take place at or within the KRKR consensus site located at residues 78–81 [23], close to the site of the Asp substitution at residue 77; thus, the Asp77 substitution might be expected to impact activation of proPC2, either positively or negatively. Surprisingly, however, this substitution did not impact the efficiency of proglucagon cleavage in our model system.

Based on a model of the PC2 protein [21], the R430W substitution is located at the very end of the catalytic domain of PC2, in an area which physically adjoins the P domain. Since the P domain influences pH sensitivity [22, 23], we surmised that this parameter might be affected by the R430W substitution. Figure 1E shows that the R430W mutant enzyme exhibited a broadened pH curve as compared to the wild-type enzyme. Specifically, R430W PC2 was considerably more active at pHs between 5.0 and 6.0 than the wild-type enzyme, suggesting a likely increase in cellular activity for the variant enzyme. For example, at pH 5.5, R430W PC2 was 30% more active, and at pH 5.75 50% more active, than the wild-type enzyme. However, increased glucagon production was not observed in GH4C1 cells co-transfected with R430W PC2 and proglucagon.

3.2. Association of the common PCSK2 coding variant rs200711626 (R430W) and type 2 diabetes risk in the OOA population

As described above, using direct exome sequencing the *PCSK2* coding variant R430W was found to exhibit a high frequency in an OOA population (MAF 4.2%). In agreement, using Human Exome chip data from 1909 OOA individuals that were participants in ongoing complex disease studies other than the AFDS, 187 copies of R430W were identified, yielding a MAF of 4.9%. Only two homozygotes were identified; one is known to have diabetes, while the diabetic status of the other individual is not known.

We then evaluated the allele frequency of R430W by glucose tolerance status in 1229 AFDS individuals. We found that R430W is almost twice as common (MAF 7.4%) among OOA individuals with T2D as OOA individuals with normal and/or with impaired glucose tolerance; however, these differences were not statistically significant (Table 2; $p=0.10$, $p=0.25$, and $p=0.25$). The further evaluation of ten related glucose, insulin, and body composition traits did not reveal any significant association of these traits with R430W carrier status (Table 3). We then studied the association between this variant and the insulin and glucose levels during a 75g oral glucose tolerance test (OGTT). (For ethical reasons, OGTTs were only performed on individuals without prevalent diabetes and were analyzed quantitatively only among individuals not diagnosed with diabetes.) There was no significant association between R430W and glucose or insulin levels during the OGTT (Figure 2). However, we observed a trend toward higher fasting glucose ($4.31 \text{ mg/dl} \pm 2.91$; $p=0.14$) and lower fasting insulin ($-0.08 \text{ } \mu\text{IU/mL} \pm 0.06$; $p=0.18$) in R430W carriers.

Lastly, we also evaluated the association of five previously reported intronic *PCSK2* SNPs (rs4814597, rs1609659, rs2208203, rs2021785, and rs2269023) with diabetes and OGTT glucose and insulin traits. No significant associations were seen with the standard additive model, though a *post-hoc* recessive model yielded a nominally significant association ($p = 0.01$ to 0.04) of homozygosity for the previously reported protective T allele in rs2021785 with decreased glucose at all time points except at 30 minutes ($p = 0.07$). The associated intronic SNP was not highly correlated with R430W ($r^2 = 0.05$).

4. Discussion

Previous studies have found both positive and negative associations between intronic variants in *PCSK2* and increased T2D risk [13, 14, 16–18]. Since coding variants in *PCSK2* are very rare in outbred populations, no information exists on whether variants confer a loss or gain of function for this enzyme. While the repeated independent findings of association of *PCSK2* variations with parameters involved in diabetes risk in a variety of different populations clearly implies functionality, whether these represent direct or indirect effects is not yet clear. Further work mapping possible chromatin interactions of intronic variants [32] represents an interesting avenue for further research which could lead to possible indirect gene targets. In this work we have attempted to approach this problem by examining coding variants for this enzyme, one of which, although rare in the general population, was found to be enriched in the Old Order Amish, suggesting a founder effect.

We enzymatically characterized four rare coding *PCSK2* variants (A77D, A267T, R430W and M525V). These variants were chosen for their presence in multiple exome databases, including 1000Genomes phase 3, ESP, and ExAC; their distribution within proPC2; and their predicted lack of functionality. Coding variations can theoretically affect PC2 function in numerous ways, for example by affecting specific activity; substrate specificity; enzyme stability; and/or trafficking through the secretory pathway to the secretory granule (the locus of precursor cleavage). Coding mutations in the related convertase *PCSK1* are known to affect many of these parameters, and nearly always result in impaired enzyme function (reviewed in [3]). Interestingly, however, a known human *PCSK1* mutation, S357G, found in a lung cancer patient, results in a significantly *more* active PC1/3 enzyme [22]; thus, convertase coding mutations are not necessarily loss-of-function.

Given their predicted deleterious functional consequences, it is surprising that three of the four *PCSK2* coding variants studied exhibited no gross alterations in secretion or catalytic activity under standard reaction conditions. Only the A267T substitution resulted in dramatic effects, with a profound reduction in medium PC2 protein and corresponding activity. Since cellular protein levels of this enzyme were increased relative to wild-type enzyme, this variant is likely subject to misfolding followed by retention within the endoplasmic reticulum, a known problem for secretory protein mutants such as PC1/3 [20, 33] or proinsulin (reviewed in [34]). Of note, this variant appears only twice in ExAC (Table 1) and twice in 1000Genomes phase 3, both in the heterozygous state.

With regard to the other PC2 variants, slight differences in activity, either positive or negative, can be difficult to detect in transfection experiments. For example, while a

common *PCSK1* variant, N221D (rs6232) results in a modest reduction in PC1/3 activity in similar transfection studies, the paired Q665E/S690T variant (MAFs of 15–30%) rs6234–rs6235, which is strongly associated with obesity risk in a variety of populations [3, 27, 28] exhibits enzyme activity identical to that of wild-type PC1/3 [20, 35]. In our experiments, variant PC2 enzymes which exhibit identical activity to wild-type enzyme could exhibit either increased or decreased activity *in vivo*. More active enzyme would likely lead to more efficient cleavage of prohormones. Since the phenotype of the PC2 null mouse is clearly dominated by the complete lack of glucagon [8, 36], changes in glucagon production might be expected to show the largest effect. Interestingly, the *Pcsk2* null mouse is protected from dietary-induced weight gain, while genetically obese mice are known to express higher levels of *Pcsk2*, suggesting that that susceptibility to either obesity and/or diabetes could be associated with *increased* rather than with decreased PC2 activity [11, 37]. This idea is supported by our results on the R430W variant enzyme, whose broadened pH optimum supports the idea of increased activity *in vivo*. However, we did not detect R430W PC2-mediated alterations in glucagon production in proglucagon-transfected GH4C1 cells; we speculate that this cell system may already operate at maximal glucagon production. Studies using *Pcsk2* gene-edited mice will likely be required to definitively establish the effects of rare coding variants on glucagon production and susceptibility to diabetes and obesity.

In our studies of the OOA population, we found that the *PCSK2* variant R430W is twice as common among OOA individuals with T2D as in those with normal or impaired glucose tolerance. A power analysis suggests that the lack of significant association between carriers of R430W *PCSK2* with glucose and insulin tolerance during the OGTT may be ascribed to the modest sample sizes (to have 80% power to detect a significant difference in R430W carrier frequency of 3.9% vs. 7.4% between NGT and diabetes groups with an alpha of 0.05, data from 230 diabetic individuals and 955 NGT individuals would be required). It is also possible that the effect of the R430W substitution is manifested only in those already exhibiting impaired glucose tolerance. Changes in PC2 activity, even if small, might be particularly important in individuals with impaired glucose tolerance. We speculate that individuals who are beginning to exhibit dysregulated glucose metabolism may be more likely to convert to T2D if they carry the R430W variation. Additional studies measuring fasting and postprandial glucagon levels in sufficient numbers of R430W carriers vs non-carriers are needed to better define the effect of this variant on peptide hormone synthesis, and could help to clarify how enzyme activity changes might underlie the observed trend of linkage of the R430W allele with diabetes risk. Of note, rs200711626/*PCSK2* R430W is not associated with any eQTLs in any of the >70 tissues profiled and indexed at the GTEx portal (gtexportal.org).

In conclusion, understanding how convertase missense variants affect the activity of PC2 against proglucagon and other prohormone substrates will improve our functional knowledge of the role of this enzyme in controlling blood sugar. While missense variants of *PCSK2* are rare in the general population, intronic variants are common and are clearly linked to metabolism. Since PC2 action is critical to glucose handling, achieving a full understanding how different *PCSK2* variations generate gain- or loss-of-function enzymes could have important implications for differential treatment strategies in metabolic disease.

Acknowledgments

We thank Dr. Robert Day for originally providing human proPC2-encoding cDNA and Dr. Elias Blanco for technical advice. We are extremely grateful to Jeffrey Kleinberger for performing the *in silico* analysis of coding mutants.

Funding: This work was supported by the National Institute of Health grants DA05084 and DK079637 to IL and TIP respectively.

Abbreviations

AFDS	Amish Family Diabetes Study
CPE	carboxypeptidase E
ExAC	Exome Aggregation Consortium
ESP	Exome Sequencing Project
GERP	genomic evolutionary rate profiling
IGT	impaired glucose tolerance
MAF	minor allele frequency
NGT	normal glucose tolerance
OOA	Old Order Amish
PC	prohormone convertase
SIFT	sorting intolerant from tolerant

References

1. Jackson RS, Creemers JW, Ohagi S, Raffin-Sanson ML, Sanders L, Montague CT, et al. Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. *Nat Genet.* 1997; 16:303–6. [PubMed: 9207799]
2. Martin MG, Lindberg I, Solorzano-Vargas RS, Wang J, Avitzur Y, Bandsma R, et al. Congenital proprotein convertase 1/3 deficiency causes malabsorptive diarrhea and other endocrinopathies in a pediatric cohort. *Gastroenterology.* 2013; 145:138–48. [PubMed: 23562752]
3. Ramos-Molina B, Martin MG, Lindberg I. PCSK1 Variants and Human Obesity. *Prog Mol Biol Transl Sci.* 2016; 140:47–74. [PubMed: 27288825]
4. Stijnen P, Ramos-Molina B, O’Rahilly S, Creemers JW. PCSK1 Mutations and Human Endocrinopathies: From Obesity to Gastrointestinal Disorders. *Endocr Rev.* 2016; 37:347–71. [PubMed: 27187081]
5. Naggert JK, Fricker LD, Varlamov O, Nishina P, Rouille Y, Steiner DF, et al. Hyperproinsulinemia in obese fat/fat mice associated with a point mutation in the carboxypeptidase R gene and reduced enzyme activity. *Nat Genet.* 1995:135–42. [PubMed: 7663508]
6. Chen H, Jawahar S, Qian Y, Duong Q, Chan G, Parker A, et al. Missense polymorphism in the human carboxypeptidase E gene alters enzymatic activity. *Hum Mutat.* 2001; 18:120–31. [PubMed: 11462236]
7. Li P, Tiwari HK, Lin WY, Allison DB, Chung WK, Leibel RL, et al. Genetic association analysis of 30 genes related to obesity in a European American population. *Int J Obes (Lond).* 2014; 38:724–9. [PubMed: 23900445]

8. Furuta M, Zhou A, Webb G, Ravazzola M, Orci L, Steiner DF. Severe defect in proglucagon processing in islet A-cells of prohormone convertase 2 null mice. *J Biol Chem.* 2001; 276:27197–202. [PubMed: 11356850]
9. Webb GC, Akbar MS, Zhao C, Swift HH, Steiner DF. Glucagon replacement via micro-osmotic pump corrects hypoglycemia and alpha-cell hyperplasia in prohormone convertase 2 knockout mice. *Diabetes.* 2002; 51:398–405. [PubMed: 11812747]
10. Peinado JR, Laurent V, Lee SN, Peng BW, Pintar JE, Steiner DF, et al. Strain-dependent influences on the hypothalamo-pituitary-adrenal axis profoundly affect the 7B2 and PC2 null phenotypes. *Endocrinology.* 2005; 146:3438–44. [PubMed: 15878971]
11. Anini Y, Mayne J, Gagnon J, Sherbafi J, Chen A, Kaefer N, et al. Genetic deficiency for proprotein convertase subtilisin/kexin type 2 in mice is associated with decreased adiposity and protection from dietary fat-induced body weight gain. *Int J Obes (Lond).* 2010; 34:1599–607. [PubMed: 20498660]
12. Jones HB, Reens J, Brocklehurst SR, Betts CJ, Bickerton S, Bigley AL, et al. Islets of Langerhans from prohormone convertase-2 knockout mice show alpha-cell hyperplasia and tumorigenesis with elevated alpha-cell neogenesis. *Int J Exp Pathol.* 2014; 95:29–48. [PubMed: 24456331]
13. Yoshida H, Ohagi S, Sanke T, Furuta H, Furuta M, Nanjo K. Association of the prohormone convertase 2 gene (PCSK2) on chromosome 20 with NIDDM in Japanese subjects. *Diabetes.* 1995; 44:389–93. [PubMed: 7698505]
14. Zheng X, Ren W, Zhang S, Liu J, Li S, Li J, et al. Association of type 2 diabetes susceptibility genes (TCF7L2, SLC30A8, PCSK1 and PCSK2) and proinsulin conversion in a Chinese population. *Mol Biol Rep.* 2012; 39:17–23. [PubMed: 21437630]
15. Bouatia-Naji N, Vatin V, Lecoœur C, Heude B, Proenca C, Veslot J, et al. Secretory granule neuroendocrine protein 1 (SGNE1) genetic variation and glucose intolerance in severe childhood and adult obesity. *BMC Med Genet.* 2007; 8:44. [PubMed: 17617923]
16. Chang TJ, Chiu YF, Sheu WH, Shih KC, Hwu CM, Quertermous T, et al. Genetic polymorphisms of PCSK2 are associated with glucose homeostasis and progression to type 2 diabetes in a Chinese population. *Sci Rep.* 2015; 5:14380. [PubMed: 26607656]
17. Jonsson A, Isomaa B, Tuomi T, Eriksson JG, Groop L, Lyssenko V. Effect of a common variant of the PCSK2 gene on reduced insulin secretion. *Diabetologia.* 2012; 55:3245–51. [PubMed: 23011353]
18. Leak TS, Keene KL, Langefeld CD, Gallagher CJ, Mychaleckyj JC, Freedman BI, et al. Association of the proprotein convertase subtilisin/kexin-type 2 (PCSK2) gene with type 2 diabetes in an African American population. *Mol Genet Metab.* 2007; 92:145–50. [PubMed: 17618154]
19. Hsueh WC, Mitchell BD, Aburomia R, Pollin T, Sakul H, Gelder Ehm M, et al. Diabetes in the Old Order Amish: characterization and heritability analysis of the Amish Family Diabetes Study. *Diabetes Care.* 2000; 23:595–601. [PubMed: 10834415]
20. Blanco EH, Ramos-Molina B, Lindberg I. Revisiting PC1/3 Mutants: Dominant-Negative Effect of Endoplasmic Reticulum-Retained Mutants. *Endocrinology.* 2015; 156:3625–37. [PubMed: 26207343]
21. Lindberg I, Shaw E, Finley J, Leone D, Deininger P. Posttranslational modifications of rat proenkephalin overexpressed in Chinese hamster ovary cells. *Endocrinology.* 1991; 128:1849–56. [PubMed: 2004604]
22. Blanco EH, Peinado JR, Martin MG, Lindberg I. Biochemical and cell biological properties of the human prohormone convertase 1/3 Ser357Gly mutation: a PC1/3 hypermorph. *Endocrinology.* 2014; 155:3434–47. [PubMed: 24932808]
23. Muller L, Cameron A, Fortenberry Y, Apletalina EV, Lindberg I. Processing and sorting of the prohormone convertase 2 propeptide. *J Biol Chem.* 2000; 275:39213–22. [PubMed: 10995742]
24. Shen FS, Seidah NG, Lindberg I. Biosynthesis of the prohormone convertase PC2 in Chinese hamster ovary cells and in rat insulinoma cells. *J Biol Chem.* 1993; 268:24910–5. [PubMed: 8227053]
25. Zhu X, Lindberg I. 7B2 facilitates the maturation of proPC2 in neuroendocrine cells and is required for the expression of enzymatic activity. *J Cell Biol.* 1995; 129:1641–50. [PubMed: 7790360]

26. Vivoli M, Caulfield TR, Martinez-Mayorga K, Johnson AT, Jiao GS, Lindberg I. Inhibition of prohormone convertases PC1/3 and PC2 by 2,5-dideoxystreptamine derivatives. *Mol Pharmacol*. 2012; 81:440–54. [PubMed: 22169851]
27. Kowalska D, Liu J, Appel JR, Ozawa A, Nefzi A, Mackin RB, et al. Synthetic small-molecule prohormone convertase 2 inhibitors. *Mol Pharmacol*. 2009; 75:617–25. [PubMed: 19074544]
28. Cooper GM, Stone EA, Asimenos G, Program NCS, Green ED, Batzoglou S, et al. Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res*. 2005; 15:901–13. [PubMed: 15965027]
29. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res*. 2001; 11:863–74. [PubMed: 11337480]
30. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010; 7:248–9. [PubMed: 20354512]
31. Dey A, Lipkind GM, Rouille Y, Norrbom C, Stein J, Zhang C, et al. Significance of prohormone convertase 2, PC2, mediated initial cleavage at the proglucagon interdomain site, Lys70-Arg71, to generate glucagon. *Endocrinology*. 2005; 146:713–27. [PubMed: 15528303]
32. Fang R, Yu M, Li G, Chee S, Liu T, Schmitt AD, et al. Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. *Cell Res*. 2016; 26:1345–8. [PubMed: 27886167]
33. Stijnen P, Brouwers B, Dirkx E, Ramos-Molina B, Van Lommel L, Schuit F, et al. Endoplasmic reticulum-associated degradation of the mouse PC1/3-N222D hypomorph and human PCSK1 mutations contributes to obesity. *Int J Obes (Lond)*. 2016; 40:973–81. [PubMed: 26786350]
34. Liu M, Hodish I, Haataja L, Lara-Lemus R, Rajpal G, Wright J, et al. Proinsulin misfolding and diabetes: mutant INS gene-induced diabetes of youth. *Trends Endocrinol Metab*. 2010; 21:652–9. [PubMed: 20724178]
35. Benzinou M, Creemers JW, Choquet H, Lobbens S, Dina C, Durand E, et al. Common nonsynonymous variants in PCSK1 confer risk of obesity. *Nat Genet*. 2008; 40:943–5. [PubMed: 18604207]
36. Furuta M, Yano H, Zhou A, Rouille Y, Holst JJ, Carroll R, et al. Defective prohormone processing and altered pancreatic islet morphology in mice lacking active SPC2. *Proc Natl Acad Sci U S A*. 1997; 94:6646–51. [PubMed: 9192619]
37. Chiu S, Kim K, Haus KA, Espinal GM, Millon LV, Warden CH. Identification of positional candidate genes for body weight and adiposity in subcongenic mice. *Physiol Genomics*. 2007; 31:75–85. [PubMed: 17536020]

Highlights

- The *PCSK2* coding variant R430W is greatly enriched in the Old Order Amish population
- The R430W variant is twice as common among Old Order Amish individuals with T2D
- The R430W substitution results in a broader pH range of enzyme activity
- The rare coding variant A267T is not secreted

Figure 1. Enzyme activity and proglucagon cleavage of the PCSK2 variants

Panel A) Diagram of the human PC2 protein indicating the positions of the *PCSK2* coding variants studied here (grey ovals) and other known coding variants (yellow rectangles). **Panel B) *In vitro* activity assay in CHO-7B2 cells** (a constitutively-secreting cell line). Conditioned medium obtained from transfected CHO-7B2 cells was subjected to *in vitro* enzymatic assays. **Panel C) Expression and secretion of PCSK2 variants in CHO-7B2 cells.** Immunoblot analysis of proPC2 and 7B2 protein levels in whole cell lysates and in the overnight conditioned medium obtained from transfected CHO-7B2 cells. Ponceau S staining was used as a loading control. Note the comparable levels of activity for all variants except A267T; this variant was largely retained within the cell. **Panel D) Cleavage of proglucagon by wild-type and variant enzymes in GH4C1 cells** (a neuroendocrine cell line). Glucagon levels were measured by RIA in the overnight conditioned medium of GH4C1 cells co-transfected with proglucagon and *PCSK2* variant cDNAs; an immunoblot of PC2 protein levels in the same conditioned medium samples is shown above. Note the lack of A267T PC2 in the medium, indicative of impaired secretion; glucagon production by other variants was similar to that of wild-type PC2. Enzyme activity and RIA data were analyzed using a one-way ANOVA followed by a *post hoc* Bonferroni's

test (***, $P < 0.001$); $n = 3$. **Panel E) The R430W variant enzyme exhibits a broadened pH curve.** The activity of wild-type PC2 and R430W-containing PC2 was measured following dilution into buffers designed to reach the pHs shown. The values have been normalized to the maximum activity at pH 5.0 to correct for differences in expression across different experiments. Note the increased activity of the R430W PC2 mutant at higher pHs. Enzyme activity was analyzed using a two-way ANOVA followed by a post-hoc Bonferroni's test (***, $P < 0.001$; mean $\pm$ SD for 4 independent experiments).

Figure 2. Association of rs200711626 R430W with glucose and insulin levels during an oral glucose tolerance test

An OGTT was performed in R430W carriers who do not have diabetes; and in non-carrier Amish individuals. The values are the mean $\pm$ SE of 50–57 R430W carriers (red) and 624–728 non-carriers (blue). **Panel A**, glucose; **Panel B**, insulin.

Table 1Properties of *PCSK2* coding variants in individuals in the ExAC database

SNP ID	Allele frequency in ExAC,%	Allele change	Amino acid change	Protein region	Appears in ExAC (#)	Polyphen2 HVAR	SIFT	GERP
rs149426043	0.0033	T>G	C18G	signal	4	0.00 (B)	0.169 (T)	5.40
rs143788251	0.0099	C>A	R47S	propeptide	12	0.013 (B)	0.331 (T)	5.46
rs201718679	0.0033	C>A	A77D	propeptide	4	0.01 (B)	0.020 (D)	3.65
rs372462573	0.0016	C>T	R81C	propeptide	2	0.45 (B)	0.036 (D)	5.58
rs201915698	0.0041	A>G	N221D	catalytic	5	0.083 (B)	0.369 (T)	4.56
rs144151196	0.0017	G>A	A267T	catalytic	2	0.99 (D)	0.000 (D)	5.38
rs200711626	0.013	C>T	R430W	catalytic	16	0.97 (D)	0.004 (D)	4.65
rs139619496	0.076	A>G	M525V	P domain	92	0.17 (B)	0.018 (D)	5.93
rs200384798	0.014	C>T	T569I	P domain	17	0.012 (B)	0.262 (T)	3.34
rs148983041	0.0034	A>T	Q595L	C-terminal	4	0.468 (P)	0.035 (D)	5.47

D: deleterious; B: benign; T: tolerated; P: probably damaging

Table 2

Minor allele frequency of rs200711626 (R430W) in the Amish population

	N	Allele Frequency (%)
Diabetes	156	7.4
IGT	156	2.9
NGT	648	3.9
NGT+IGT	804	3.7
Overall	957	4.3

P=0.10 comparing diabetes to NGT + IGT; P=0.25 comparing IGT to NGT; P=0.25 comparing diabetes to NGT; IGT – impaired glucose tolerance; NGT – normal glucose tolerance.

Table 3

Association of rs200711626 R430W with type 2 diabetes status and diabetes-related traits in the Amish population

	N	β	P
Lg HbA1C	1118	0.02 ± 0.02	0.17
Fasting glucose	987	4.07 ± 3.21	0.20
AUC glucose	772	7.40 ± 16.85	0.66
Log fasting insulin	820	-0.07 ± 0.05	0.18
AUC insulin	711	12.18 ± 11.56	0.29
HOMA-IR	819	-0.30 ± 0.21	0.16
BMI	1229	0.67 ± 0.53	0.21
Weight	1226	1.25 ± 1.51	0.41
Waist circumference	1206	0.81 ± 1.22	0.50
Waist to hip ratio	1204	0.005 ± 0.01	0.49