Localization and characterization of the novel protein encoded by C20orf3

Aysegul ILHAN†, Wolfgang GARTNER†, Anastasiya NABOKIKH*, Teodora DANEVA†, Otto MAJDIC‡, Gerald COHEN†, Georg A. BÖHMIG†, Wolfgang BASE†, Walter H. HÖRL† and Ludwig WAGNER†

*Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria; †Bulgarian Academy of Sciences, 73 Tzarigradsko Shosse Blud, 1113 Sofia, Bulgaria; ‡Institute of Immunology, Medical University of Vienna, Borschkehagasse 8a, A-1090 Vienna, Austria

In the present study, we characterized the gene product of open reading frame 3 encoded at human chromosome 20 (C20orf3), which represents a member of the lactonohydrolase superfamily. Multiple-tissue Northern blot analysis showed ubiquitous expression of the 2.4 kb transcript coding for 416 amino acids, with highest levels in human liver, placenta and kidney. After recombinant production of protein variants in Escherichia coli and insect cells, antibodies directed against different epitopes within the C20orf3 gene product were generated. Using these immunoreagents, protein expression was demonstrated in the liver, and glomerular and tubular structures of the kidney, as well as in endothelial cells and arterial wall. Positive staining was also observed at the pancreatic islets of Langerhans. Using immunoblotting, we identified three size variants. In line with the results of in silico analysis demonstrating a single transmembrane sequence (amino acids 40–61) at the N-terminus of the full-length protein, FACS cell-surface staining confirmed a mainly extracellular localization of the full-length protein. Sucrose density gradient cell fractionation revealed membrane association of the dominant 50 kDa variant in HepG2 and Rin-5F cells. The finding of a strong arylesterase activity with β-naphthyl acetate and phenyl acetate of the C20orf3 protein-containing fractions suggests potential involvement of this protein in enzymatic processes. C20orf3 promoter-driven reporter assays, which were verified by gene-specific RT-qPCR (real-time quantitative PCR) showed a strong inhibitory effect of human serum on transcription using the HEK-293 human embryonic kidney cell line. In conclusion, we characterized the structure and expression pattern of the C20orf3 gene product. According to a series of analogies with PON (paraoxonase) family members, we speculate that the C20orf3 gene product represents a new member of this important protein family present at the cellular level.

Key words: C20orf3, gene cloning, lactonohydrolase, paraoxonase (PON), protein purification, strictosidine synthase.

INTRODUCTION

The liver and the kidney represent important sites of drug metabolism and detoxification. In addition to the excretory capacity of the kidneys, the involvement of these organs in balancing and modifying the plasma protein pool as well as their responsibility for inactivation of bioactive endogenous compounds have been demonstrated.

The renal detoxification process depends strongly on the water solubility of toxins. Therefore enzymes involved in the attachment of small hydrophilic groups or in hydrolysis, thus inactivating ingested or endogenous potentially toxic compounds, are highly prevalent in the kidney. Several such enzymes have been identified, and their expression pattern has been characterized. One important group of such detoxification enzymes are PONs (paraoxonases), of which three members (PON1–PON3) have been defined in humans [1]. The highest expression of PON1 and PON3 was found in the liver [2–4], whereas PON2 is also expressed in other organs, e.g. the brain and the kidney [5,6]. Although the endogenous substrates of PON family members in mammals are still not defined, an important role of these enzymes is drug modification via their lactonase activity [7,8]. For example, spirinolactone and distinct statins can be hydrolysed by PON3 and, to a much lower extent, by PON1 [9]. Furthermore, dihydrocumarin has been reported to represent the substrate of all three members [7,8,10].

Interestingly, PON1 and PON3 were detectable in the blood [2,4], whereas PON2 was only detected at the cell surface and intracellularly, where it exerts its antioxidant properties [6]. An important function of secreted PONs is their association with lipoproteins [11–13]. Thereby it has become evident that PON1 exerts antioxidant effects on lipids, thus preventing atherosclerosis [14,15]. In addition, PON1 and PON2 polymorphisms which affect the primary structure of the protein and activity [16,17] are associated with changes in the intermediate traits of plasma lipoprotein metabolism [18].

The recombinant expression of PON family members has facilitated their structural and functional analysis [8,19]. Interestingly, this protein family uses a helix at the N-terminus either as a transmembrane domain or to associate with lipoproteins [13,20]. In line with the drug-modifying capacities of PON family members, the similarity of the crystal structures of PON1 and plant SS (strictosidine synthase), namely their six-bladed β-propeller fold structure, has to be mentioned [20,21]. Plant SS catalyses the condensation of secologanin and tryptamin to strictosidine, thus being involved in the generation of numerous biologically important alkaloids [22]. This enzyme shows similarity with the as yet unanalysed gene product of the C20orf3 (chromosome...
20 open reading frame 3) gene locus and might represent both members of a superfamily with lactonohydrolases and PONs [23].

In the present paper, we describe the cloning and molecular characterization of the C20orf3 gene product representing an until now undefined protein in humans. Using newly generated antibodies and multiple-tissue Northern blotting, the C20orf3 expression pattern has been deduced, showing a characteristic glomerular and tubular expression within the kidney and high levels of hepatocellular presence. Additionally, central nervous system expression and presence at distinct endocrine organs was revealed. Its homology with plant SS and its similarities with the PON family members suggest important functions of this protein in central metabolic processes.

METHODS

Cloning of the human C20orf3 gene using Gateway and Bac-to-Bac® technologies and antibody generation

Methods of cloning human C20orf3, purification of the protein variants and antibody generation are described in detail in the Supplementary material (http://www.BiochemJ.org/bj/414/bj4140485add.htm). Escherichia coli BL21 (Invitrogen) and insect cell lines SF9 and High Five (both from Invitrogen) were utilized for protein purification using Gateway and Bac-to-Bac® (Invitrogen) technologies respectively. The resultant human recombinant C20orf3 protein was injected into mice and rabbits for monoclonal and polyclonal antibody generation.

Northern blot analysis

For RNA extraction, tissues were homogenized with a Polytron PT 3100 tissue homogenizer (Kinematica) in TRIZol® (Gibco-BRL). The resultant solution was treated as indicated in the manufacturer’s instructions. In brief, TRIZol® homogenates were extracted with chloroform, and the aqueous phase was transferred into a separate tube for RNA precipitation using propan-2-ol. After washing in 75% ethanol, the RNA pellet was dissolved in DEPC (diethyl pyrocarbonate)-treated water and either used immediately for reverse transcription or frozen at −80 °C.

Multiple-tissue Northern blots (BD Biosciences) containing 2 μg of mRNA in each lane were hybridized with a nine-exon-spanning C20orf3-specific probe. Hybridization using ULTRAhyb solution was performed according to the manufacturer’s instructions (Ambion). Briefly, after a 2 h pre-hybridization step at 42 °C, blots were incubated overnight with the [32P]-labelled probe in 10 ml of ULTRAhyb solution and were then washed with NorthernMax™ Low Stringency Wash Solution #1 followed by washing with NorthernMax™ High Stringency Wash Solution #2 (both from Ambion). Blots were exposed to X-ray films at −80 °C overnight. In order to control for RNA loading, the same blots were re-hybridized using a [32P]-labelled β-actin probe.

RT-qPCR (real-time quantitative PCR)

cDNA generated from HEK-293 (human embryonic kidney) cells (A.T.C.C. CRL-1573), SH-SYSY cells (A.T.C.C. CRL-2266), HepG2 cells (A.T.C.C. HB-8065), placenta, kidney and liver were used as template in TaqMan gene expression assays. The original cDNA transcript (1.8 μl) was used together with 2.5 μl of the 20× gene-specific FAM (6-carboxyfluorescein)-labeled TAMRA (6-carboxytetramethylrhodamine)-labeled Assay-on-Demand C20orf3-specific probe (Applied Biosystems) and 25 μl of the 2× PCR MasterMix (Roche) diluted to a final volume of 50 μl with water. RT-qPCR was performed in duplicates using an ABI Prism 7000 Cycler (Applied Biosystems). C20orf3 values were normalized to 18S or GAPDH (glyceraldehyde-3-phosphate dehydrogenase) as endogenous control using an Assay-on-Demand 18S- or GAPDH-specific probes (Applied Biosystems).

TaqMan SNP (single nucleotide polymorphism) genotyping assay

Isolated genomic DNA (1–20 ng) was diluted in 11.25 μl of water and combined with 12.5 μl of TaqMan universal PCR Master Mix (2×) (Applied Biosystems) and 1.25 μl of 20× working stock of C20orf3 SNP genotyping assay. This was purchased from Applied Biosystems (C-25652369_10 and C-61038612_10) for the polymorphic site of residues 282 (rs28364786) and 374 (rs35097515) according to NCBi dbSNP cluster identification. Each genotyping assay set contains a probe labelled with VIC® (2-chloro-7-phenyl-1,4-dichloro-6-carboxyfluorescein) specific for one allele and a probe labelled with FAM specific for the second allele. Detailed information is provided in the manufacturer’s instructions. The reaction was set up in a 96-well optical plate, and PCR cycling and allele discrimination analysis was carried out using an ABI Prism 7000 cycler with ABI Prism 7000 SDS software.

Ion-exchange chromatography

The Ni-NTA (Ni2+-nitritoltriacetate) His-Bind® resin (Novagen)-purified insect-cell-derived protein (1 ml) was applied on to an FPLC system (AKTA Explorer; Amersham Biosciences). Ion-exchange chromatography was performed on a Q Sepharose column (HiTrap™ FF; 1 ml) using the Buffer Prep program AIEX. The separation was geared by a specific program chosen in accordance with the column using Unicorn software. The flow rate was 1 ml/min. Fractions of 500 μl were collected and the appropriate peak was used for further analysis and testing. All chromatographic media were purchased from Amersham Biosciences.

ELISA measurement

Reacti-Blind™ goat anti-mouse antibody-coated plates (Pierce) were incubated overnight at 4 °C with polyclonal mouse anti-C20orf3 antibody diluted 1/400 in tissue culture medium. After a wash with PBS or a serial dilution of human recombinant C20orf3 as standard (100 μl each), samples were incubated for 2 h at room temperature (22 °C) with constant shaking. Bound antigen was detected with rabbit anti-(human C20orf3) antibody (1/4000 dilution) with incubation for 90 min, followed by incubation for 90 min with horseradish-peroxidase-labelled goat anti-rabbit antibody (1/6000 dilution) (Dako). Following each incubation step, the plate was washed three times with PBS/0.1% Tween 20 (Bio-Rad) in the ELX Auto strip Washer (Biotec Instruments). The ELISA was developed using tetramethylbenzidine dihydrochloride two-component peroxidase substrate solution (Kirkegaard & Perry). After stopping the reaction with 1 M phosphoric acid, the absorbance was quantified at 450 nm using a Powerwave ELISA reader (Biotec Instruments), and concentrations were calculated according to the standard curve with human recombinant C20orf3.

Immunoblot analysis

Healthy kidney tissue obtained from tumour nephrectomy, healthy liver from tumour metastasectomy and placenta tissues were cut into small pieces which were homogenized using the Sample Grinding Kit (Amersham Biosciences) and Weinberg lysis buffer containing 50 mM Hepes (pH 7.0), 0.5% Nonidet P-40, 250 mM NaCl, 5 mM EDTA and protease inhibitors (Complete™ Mini;
Roche). Clarified samples were either frozen in 20 μl aliquots or used immediately. Cell lines HEK-293, SH-SY5Y and HepG2 were grown under standard culture conditions in RPMI 1640 medium supplemented with 10% fetal bovine serum at 37°C in 5% CO2 and 95% humidified air.

Protein aliquots (30 μg) were separated by SDS/PAGE (12% gels), which were run at 200 V using Tris-glycine running buffer. A semidry blotting device was used to electrophoretically transfer the protein on to nitrocellulose. The blotted membrane was blocked (40 min) with 2% non-fat dried skimmed milk powder in PBS. Antigen detection was performed overnight with constant shaking at 4°C using the polyclonal mouse/rabbit anti-C20orf3 antibody (1/1500 dilution). Horseradish-peroxidase-conjugated goat anti-mouse/rabbit antibody (1/3000 dilution) was used for detection of antibody-binding sites (Dako). Each incubation step was followed by two washes with PBS containing 0.1% Tween 20 for 10 min. Binding-site visualization was carried out by BM chemiluminescence reagent (Roche), and electronic images were recorded using Lumi-Imager F1 (Roche).

**Immunostaining**

Cryosections of 4 μm of kidney, liver and pancreas tissue were air-dried and fixed with 3% (w/v) paraformaldehyde in PBS (3 min) followed by washing in PBS. Mouse mAb56 (monoclonal antibody 56) and control (mouse Ig) were applied on the sections together with rabbit anti-SCGN (secretagogin) antibody (diluted 1/3000 in PBS) [24,25] for human pancreas and incubated overnight at 4°C. After 10 min of washing with PBS under constant stirring, Alexa Fluor® 488-conjugated donkey anti-mouse or goat anti-rabbit (diluted 1/200 in PBS) (Molecular Probes), and TRITC (tetramethylrhodamine β-isothiocyanate)-conjugated AffiniPure F(ab’2) labelled goat anti-mouse [H + L (heavy and light chains)] IgG (Accurate Chemical & Scientific Corporation) was incubated for 45 min at room temperature. Following a 10 min washing step, slides were mounted using Vectashield mounting medium with DAPI (4′,6-diamidino-2-phenylindole) (Vector Laboratories) and covered with a coverslip. Confocal images were recorded using a Zeiss Axiovert confocal microscope and processed further using Adobe Photoshop 6.0.

**In silico analysis**

In order to identify the hydropathy of the C20orf3 gene product, we entered the gene sequence into the bioinformatic software SeqVu 1.0 (Garvan Institute of Medical Research, Sydney, NSW, Australia). For transmembrane and signal peptide determination Signal P 3.0 Server (Technical University of Denmark, Kgs. Lyngby, Denmark) and the PHD prediction server (http://cubic.bioc.columbia.edu/) were used. The domain search was performed at ProDom (http://protein.toulouse.inra.fr/prodom.html). The three-dimensional structural image of C20orf3 was generated using Cn3D 4.1 software. This program was obtained from NCBI together with the co-ordinates of the well known three-dimensional structures of plant SS which were taken as a template for alignment with C20orf3 protein and definition of its co-ordinates. For sequence alignment, ClustalX 2 (Conway Institute, UCD, Dublin, Ireland) was used.

**FACS analysis**

Peripheral blood cells were incubated with red cell ammonium chloride-based haemolysis buffer solution on ice (20 min). For cell-surface staining, 1 × 10⁶ washed leucocytes were incubated with mAb56–FITC and control for 45 min on ice including Beriglobin® P (Aventis Behring) for blocking Fc binding. After two washing steps, fluorescence measurement was carried out on a FACS Calibur (BD Biosciences) by gating specifically for lymphocytes, granulocytes and monocytes.

Exponentially growing HepG2 hepatoma cells were obtained either by trypsinization or detachment from the flask by vigorous pipetting. Antibody staining was carried out using mAb56–FITC on ice. Cell-surface expression was evaluated using the FACS Calibur.

FACS dual-colour analysis was performed using either co-staining of mononuclear cells with the anti-HLA (human leucocyte antigen)-DR PerCP-specific antibody (BD Biosciences) or the PE (phycoerythrin)-conjugated anti-human CD14) antibody (BD Biosciences) and FITC-conjugated mAb56 including Beriglobin® P. Dual-colour profiles were recorded on the FACS Canto (BD Biosciences) and analysed with FACS Diva version 6.0 (BD Biosciences) software. FITC-conjugated mAb56 was diluted in all experiments 1/100 in Beriglobin® P (1 mg/ml).

**Sucrose-density-gradient cell fractionation**

Rin-5F (A.T.C.C. CRL-2058) and HepG2 cells were cultured in 75-cm² tissue culture flasks. At 75% confluence, both were washed twice with ice-cold PBS and then scraped into 2.5 ml of ice-cold homogenization buffer (250 mM sucrose and 3 mM imidazole, pH 7.4) supplemented with protease inhibitors (10 μg/ml aprotinin, 1 μg/ml pepstatin, 10 μg/ml leupeptin and 0.8 mM Pefabloc®). The cell suspension was then passed four times through a 22-gauge needle. After centrifugation at 2200 g for 10 min at 4°C, the post-nuclear supernatant was overlaid on a discontinuous sucrose density gradient (50 and 20% respectively) in 14 mm × 95 mm polyallomer centrifuge tubes (Beckman Instruments) and centrifuged at 40,000 rev./min for 2 h at 4°C using a SW40 Ti rotor of an L-80 ultracentrifuge (Beckman). The resultant density equilibrium was then fractionated using a peristaltic pump. Starting at the bottom, we collected 21 aliquots. The fractions were subjected to immunoblot analysis for C20orf3 gene product, SNAP-25 (25 kDa synaptosome-associated protein) or NSF (N-ethylmaleimide-sensitive factor) and arylesterase detection assays with β-naphthyl acetate and phenyl acetate.

**Detection of arylesterase activity**

Zymography

Samples were loaded on to a 10% PAGE gel omitting ionic detergents. The gel was run at 200 V using Tris-glycine. Enzyme detection was performed as indicated previously [8] by immersing the gel in a solution of β-naphthyl acetate.

The enzyme substrate solution consisted of 80 mg of β-naphthyl acetate and 40 mg of Fast Blue BB salt (both from Sigma) dissolved in 20 ml of ethylene glycol monomethyl ether. The resultant solution was diluted further in 50 ml of 20 mM Tris buffer (pH 8.0) including 5 mM CaCl₂.

Arylesterase activity with phenyl acetate

Measurement of arylesterase activity with phenyl acetate was performed in 20 mM Tris/HCl (pH 8.0) and 1 mM CaCl₂ (for Ca²⁺ free conditions, CaCl₂ was omitted in the absence of 5 mM EDTA) by monitoring the molar absorption value at 270 nm (Ultrospec™ 3000; GE Healthcare) for 3 min essentially as described by Connelly et al. [26]. The enzymatic activity of each fraction was given in units defined as hydrolysis of 1 μmol of phenyl acetate using a molar absorption coefficient of 1310 M⁻¹·cm⁻¹ for phenyl acetate.
A. Ilhan and others

Figure 1 cDNA sequence of the human C20orf3 protein and its translation into amino acids

C20orf3 consists of 1248 nucleotides which encode a protein of 416 amino acids. This comprises an intracellular part (underlined), a transmembrane domain (bold) and an extracellular domain. Polymorphic amino acids at positions 65, 282 and 374 are indicated by italics. Amino acids acting as glycosylation sites (residues 160 and 196) are marked with a circle.

Statistical analysis

Statistical analysis was performed using Student’s t test. Statistical significance was taken at \( P < 0.05 \). Results are means \( \pm \) S.D. for three or four experiments.

RESULTS

Cloning and C20orf3-transcript characterization

Inspired by the observation of high expression levels in various tissues, deduced from array-based expression screens, we cloned C20orf3. After choosing specific primers, cDNA of the coding sequence was amplified by PCR using a proofreading enzyme. The resultant product was ligated into a replication plasmid and sequenced.

In silico analysis confirmed an open reading frame coding for a protein of 416 residues (Figure 1). It consists of an N-terminal cytosolic tail (39 residues) and a single transmembrane domain (residues 40–61) and a cell-surface location of the main part of the protein (residues 62–416). Two cysteine residues with a high confidence level for intra- and inter-molecular disulfide bonding are located at the anterior part (residues 140 and 149) and...
Localization and characterization of the novel protein encoded by C20orf3

One additional cysteine residue is located at the C-terminal end (residue 409). Two potential glycosylation sites were identified at residues 160 and 196. Additionally, we found multiple similarities with the PON family of detoxifying enzymes. First, PON2 and C20orf3 share some degree of identity in their primary structure (Figure 2) together with plant SS and DFPase (di-isopropyl fluorophosphatase). Secondly, both proteins are characterized by an N-terminally located single transmembrane domain. Thirdly, the PON family members and the C20orf3 gene share a nine-exon structure, which indicates a common ontogenetic origin. Interestingly, a review of C20orf3-specific GEO (Gene Expression Omnibus) Profiles at NCBI showed that the gene undergoes a time-dependent down-regulation upon exposing serum-starved fibroblasts to serum (GDS85/6772) and that oxidized LDL (low-density lipoprotein) induces an up-regulation of this gene in retinal epithelial cells (GDS2307/228309). This indicates a metabolic regulation resembling that of PON family members.

Further comparative sequence analysis and conserved domain search revealed similarities between plant SS and C20orf3. This degree of similarity enabled tertiary structure modelling, using Cn3D 4.1 software which also provided similar co-ordinates for the C-terminal C20orf3 protein and the well-known six-bladed \( \beta \)-propeller structure of SS (see Supplementary Figure S1 at http://www.BiochemJ.org/bj/414/bj4140485add.htm).

Recombinant protein production and antibody generation

In order to obtain further information on the properties of the C20orf3 gene product, we focused on its recombinant expression. In an initial attempt the coding sequence of residues 45–416 was cloned into the expression vector pDEST\textsuperscript{TM}15. Notably, the protein had to be recovered from E. coli inclusion bodies, indicating its insolubility. The Superose chromatographically purified protein was used for monoclonal antibody production. Because of its insolubility, an additional transcript omitting the hydrophobic transmembrane and the cytosolic domain (residues 58–416) was recombinantly expressed using the pGEX-1\( \lambda \) expression vector. This resulted in complete solubility of the GST (glutathione transferase)-purified protein named GST–C20orf3–748, which could be used in further ELISA-based antibody-screening methods.

Using ELISA screening, we selected one antibody (mAb56) specifically recognizing the cell-surface domain of the C20orf3 gene product. Furthermore, the GST-purified protein was used for mouse and rabbit polyclonal antibody generation. Additionally, a polyclonal rabbit antibody was generated using an N-terminal cytosolic domain-specific synthetic peptide (as indicated in the Supplementary material).

Aiming at the recombinant expression of a more physiological C20orf3 variant, we changed the expression system towards a eukaryotic organism. For this purpose, insect cell lines Sf9 and High Five\textsuperscript{TM} were transfected with a baculovirus-based system, using the pDEST\textsuperscript{TM}10 baculo-competent Gateway vector encoding residues 45–416 of the C20orf3 gene. This sequence only lacked the cytosolic domain, but included a significant part of the transmembrane and full-length cell-surface domain, resulting in a 46 kDa protein (Figure 3). The N-terminal His\textsubscript{\textsuperscript{6}} tag enabled us to purify the protein using Ni-NTA His-Bind\textsuperscript{R} resin from Sf9 and High Five\textsuperscript{TM} cell lysates. Further purification was performed by ion-exchange chromatography. This resulted in a highly purified protein, which was soluble in buffer consisting of 300 mM NaCl and 50 mM sodium phosphate (pH 7.4).

---

**Figure 2** Similarity between C20orf3, PON 2, SS and DFPase

Similarity between the C20orf3 protein sequence and three members of the six-bladed \( \beta \)-propeller-folded proteins were aligned using ClustalX2 software.

**Figure 3** Recombinant expression and purification of the C20orf3 gene product in baculovirus-based High Five\textsuperscript{TM} cells

C20orf3 was expressed using a baculovirus-based expression system in High Five\textsuperscript{TM} insect cells and was subsequently purified by Ni-NTA His-Bind\textsuperscript{R} resin. After PAGE, the protein was detected by a polyclonal-specific monoclonal antibody (lane 1) and by a rabbit anti-(cell-surface domain) antibody (lane 2). For additional purification, the Ni-NTA His-Bind\textsuperscript{R} resin was subjected to Q Sepharose anion-exchange chromatography (lane 3, silver staining). Molecular masses are indicated in kDa.
Tissue expression pattern of the C20orf3 transcript

We investigated the tissue expression pattern of the C20orf3 transcript using multiple-tissue Northern blots and a C20orf3-specific cDNA probe. A dominant size transcript of approx. 2.4 kb was found ubiquitously in adult as well as in embryonic tissues (Figure 4). Interestingly, we found differences in the distribution when comparing the maturation status of the tissues. In adult tissue, highest expression was found in the liver, the placenta, and the heart. In embryonic tissue, the highest expression levels were found in the liver and the kidney.

To confirm and to extend the Northern blot analysis, RT-qPCR was carried out on human placenta, kidney, and liver tissues. As in vitro models, the HEK-293 human embryonic kidney cell line, the SH-SY5Y neuroblastoma cell line and the HepG2 human hepatoma cell line were included in the analysis. In line with the Northern blot results, highest expression levels were found in liver followed by the HepG2 cell line, with lower expression levels in kidney-derived tissue and in the HEK-293 cell line (Figure 5A).

Immunoblot analysis

To identify C20orf3 protein variants, we tested various cell lines and ex vivo tissues by immunoblotting using the polyclonal rabbit and mouse antibodies generated against the cell-surface domain (residues 58–416) respectively. In addition to the full-length protein of 50 kDa, the cell-surface-specific antibody detected two protein bands of 30 and 32 kDa (Figure 5B). It has to be stressed that the quantity as well as the ratio of the individual bands varied, depending on the cell lines. The 30 kDa band increased in repeated freeze–thaws, suggesting a degradation product. An additional 52 kDa size variant was observed in HepG2 and SH-SY5Y.

The cytosolic domain-specific antibody exhibited immunoreactivity against the 32 and 50 kDa proteins, confirming an intact N-terminus within these variants. In contrast, no detection of the 30 kDa moiety was found with the cytosolic domain-specific antibody (results not shown).

Total protein (30 μg) derived from ex vivo tissue (placenta, kidney, and liver) was loaded on to an SDS/PAGE gel, blotted on to nitrocellulose membranes and developed with polyclonal mouse anti-(cell-surface domain) antibody. This revealed the dominant 50 kDa size variant with a degradation product at small peptide size. Its high expression in human liver and placenta (Figure 5C) confirmed the mRNA results.

To identify the individual cell types expressing the C20orf3 gene product, confocal immunofluorescence and immunohistochemical analysis of renal, hepatic and pancreatic tissue sections were performed. For the immunohistochemical analysis, we used the polyclonal, N-terminal and cell-surface domain-specific rabbit antibodies (results not shown). For immunofluorescence analysis, we used the newly generated mAb56. Interestingly, this revealed a highly characteristic expression pattern of C20orf3 in these organs. Within the kidneys, we found strong immunostaining of the glomerulus, albeit with varying staining intensities (Figure 6A). The epithelial cells of distinct tubules were also characterized by a marked staining (Figure 6B). Additionally, expression of the protein was found in endothelial
In our efforts to identify the subcellular distribution of the C20orf3 protein variants in more detail, we performed a sucrose-density-gradient fractionation of Rin-5F cells. The individual fractions were subsequently analysed by immunoblotting. Using the well-established membrane marker protein SNAP-25, we identified that the high-molecular-mass C20orf3 protein appeared in the heavy membrane fractions (Figure 8A, a). Only a minor quantity of the smaller variant and some degradation products were also detected in these fractions. In addition, the 32 kDa variant was found within the cytosolic fractions, indicated by an absence of SNAP-25 (Figure 8A, b). This demonstrates different intracellular compartmentalization of these variants.

With the intention of comparing the subcellular location pattern of C20orf3 in Rin-5F cells with that of a human cell line, sucrose-density-gradient centrifugation of the human hepatoma cell line HepG2 was performed. Interestingly, we found the full-length protein in membrane fractions (Figure 8B, a). The 32 kDa variant was found in the membrane fractions in scant quantity. In these experiments, the SNAP (soluble NSF-attachment protein receptor) family member NSF was used as membrane marker.

Expression regulation of the C20orf3 gene product

To elucidate in further detail the C20orf3 gene expression regulation, we cloned the gene promoter (1081 bp) in front of the Renilla luciferase gene using the pMLuc3 reporter vector. This region includes 936 bp upstream of the defined transcription start site and consists of 62 well-defined transcription regulatory elements [GATA1, 2 and 3, Sp1 (specificity protein 1), AP-1 (activator protein 1), NF-κB (nuclear factor κB), Pbx1, Ik-2 and others]. The reporter gene assay was performed in HEK-293 cells 24 h after stimulation with human serum. Under basal conditions, the HEK-293 cells exhibited a considerable expression efficacy for the tested sequence when compared with the vector omitting the promoter insert (background expression <10%). Luciferase light units were normalized to 10,000 cells. HEK-293 cell exposure to human serum resulted in a marked and significant reduction (42 ± 11%; P < 0.003; n = 4) of C20orf3 transcription when compared with control (without serum).

Additionally, we performed FACS analysis using mAb56 and RT-qPCR experiments comparing the C20orf3 expression in unstimulated and serum-stimulated HEK-293 cells. In line with the luciferase reporter expression analysis, the serum treatment resulted in a marked decrease in C20orf3 cell-surface detection (56 ± 21%; P < 0.01; n = 4), as well as in a down-regulation of the C20orf3 gene expression (80 ± 6%; P < 0.0003; n = 3).

Expression activity of the C20orf3 gene product-containing fractions

To screen for an enzymatic activity of the C20orf3 gene product, we performed arylesterase staining of the individual fractions with β-naphthyl acetate obtained by sucrose-density-gradient cell fractionation of Rin-5F (Figure 8A, c) and HepG2 (Figure 8B, b) cells, followed by PAGE and zymography. This revealed arylesterase activity in samples which tested positive for the C20orf3 gene product in parallel immunoblot analysis.

The same fractions were evaluated using spectrophotometric analysis for arylesterase activity with phenyl acetate (Figures 8A, d and 8B, c). The fractions containing C20orf3 gene product exhibited a high arylesterase activity. Under Ca²⁺-chelated conditions, enzyme activity was reduced by approx. 1/3 (Figures 8A, e and 8B, d).

Subcellular localization of the protein

To confirm the cell-surface localization of the C20orf3 gene product, which was implicated by the comparative sequence analysis, FACS measurement of dispersed HepG2 cells using the newly prepared mAb56–FITC was performed (Figure 7A). HepG2 cells were detached either with trypsin or mechanically (vigorous pipetting). As expected, HepG2 cells exhibited high-surface-staining intensities. Remarkably, trypsin treatment resulted in a marked reduction of the immunostaining intensities, indicating loss of the cell-surface part due to C20orf3 trypsin fragmentation. In analogy to PON family members, we additionally found high-C20orf3-immunostaining intensities of monocytes, which were identified by co-staining for CD14 (Figure 7B). In contrast, granulocytes and lymphocytes did not show any mAb56-specific cell-surface staining.

Expression and characterization of the novel protein encoded by C20orf3

Localization and characterization of the novel protein encoded by C20orf3

In order to elucidate in further detail the C20orf3 gene expression regulation, we cloned the gene promoter (1081 bp) in front of the Renilla luciferase gene using the pMLuc3 reporter vector. This region includes 936 bp upstream of the defined transcription start site and consists of 62 well-defined transcription regulatory elements [GATA1, 2 and 3, Sp1 (specificity protein 1), AP-1 (activator protein 1), NF-κB (nuclear factor κB), Pbx1, Ik-2 and others]. The reporter gene assay was performed in HEK-293 cells 24 h after stimulation with human serum. Under basal conditions, the HEK-293 cells exhibited a considerable expression efficacy for the tested sequence when compared with the vector omitting the promoter insert (background expression <10%). Luciferase light units were normalized to 10,000 cells. HEK-293 cell exposure to human serum resulted in a marked and significant reduction (42 ± 11%; P < 0.003; n = 4) of C20orf3 transcription when compared with control (without serum).

Additionally, we performed FACS analysis using mAb56 and RT-qPCR experiments comparing the C20orf3 expression in unstimulated and serum-stimulated HEK-293 cells. In line with the luciferase reporter expression analysis, the serum treatment resulted in a marked decrease in C20orf3 cell-surface detection (56 ± 21%; P < 0.01; n = 4), as well as in a down-regulation of the C20orf3 gene expression (80 ± 6%; P < 0.0003; n = 3).
A. Ilhan and others

Figure 6 Confocal microscopy of human kidney, human liver and human pancreatic islet cryosections

Kidney and liver sections were stained with the newly generated mouse C20orf3-specific mAb56. Pancreatic islets were co-stained with rabbit anti-SCGN and mouse anti-C20orf3 mAb56. Nuclear staining was performed with DAPI. (A) Kidney glomerulus stained positive for C20orf3 protein (×400 magnification). (B) Distinct kidney tubular epithelia were frequently observed to express the protein (×400 magnification). (C) The C20orf3 gene product was expressed in arterial wall (×400 magnification). (D) Human liver cells stained highly positive for C20orf3 gene product (×400 magnification). (E) The C20orf3 gene product was found in pancreatic islet cells (red). The localization of the C20orf3 gene product in the endocrine parts of the pancreas was demonstrated by co-immunostaining for the neuroendocrine-specific marker protein SCGN (green).

Genotyping

SNP genotyping using specifically purchased TaqMan probes to evaluate polymorphic residues 282 and 374 identified homozygosity for Arg282 and Arg374 in 100 individuals originating from a middle European population.

ELISA measurement of the C20orf3 transcript in different biological fluids

In order to elucidate whether the protein is secreted or whether parts of it can be shed into different body fluids, we performed ELISA using serum, pleural effusions and urine. The protein

In the present paper, we detail the characterization of the C20orf3 gene product and its expression topology, as well as potential functional aspects. According to structural characteristics, C20orf3 is related to PONs, especially to PON2.

Structural analysis revealed that C20orf3 contains a hydrophobic region at the N-terminus (residues 40–61), which is typical for transmembranous regions. According to in silico analysis, the C20orf3 gene product represents a 46 kDa protein, consisting of 416 residues. However, our immunoblot data showed a protein presenting with two bands of approx. 50 and 52 kDa in cell lines. This is in analogy to PON1 and was attributed to glycosylation [27]. In this context, the glycosylation sites at residues 160 and 196 exhibited by the C20orf3 sequence seem of interest. The fact that the C20orf3 protein is prone to glycosylation was confirmed recently by another research group, who found peptides out of this protein in the course of N-glycoproteome analysis by hydrazide chemistry [28]. This indicates that the differences between the expected molecular mass and that observed in the immunoblot analysis might be due to glycosylation. Interestingly, we found two additional bands in our immunoblots of 30 and 32 kDa. The detection of the 32 kDa band with the N-terminal- peptide-specific antibody indicates that this variant contains the cytosolic domain. Looking at the EST (expressed sequence tag) database, we found splice variants lacking exons 3, 4 and 5 (BI117692). This results in a truncated C20orf3 variant containing the cytosolic part and finds parallels in splice variants described for the PON2 enzyme [5]. The smaller protein band of approx. 30 kDa is not detected by the cytosolic-domain-specific antibody and increased after repeated freeze–thaws of the cell lysates, which supports the assumption that it represents a degradation product. Small peptide degradation products were detected in all ex vivo tissue lysates.

Tests for C20orf3 gene regulation point towards serum factors influencing its expression. This is underlined by the fact that the protein cell-surface expression is reduced upon exposure to human serum using the embryonic kidney cell line HEK-293. Although expression suppression of C20orf3 after serum exposure seems to be the main contributor to this phenomenon, internalization of the protein upon ligand or substrate association might also play a role. This is underlined by the finding that the C20orf3 suppression detected at the mRNA level was less intense than that at the protein level.

It represented one of our goals to define a substrate for this molecule. It is important to note that, in the mouse, this protein has been identified as adipocyte plasma-membrane-associated protein and has been shown in 3T3-L1 cells induced to differentiate towards adipocytes [29]. Although these important data provide evidence that, in rodents, the protein might have some functional relevance in adipocyte differentiation, our data of ubiquitous expression in humans demonstrate a functional relevance beyond adipose tissue, most strikingly in the liver and the central nervous system, as well as the kidney and in blood vessels.

Although a detailed functional analysis of the C20orf3 gene product is still missing, high arylesterase activity was found in the C20orf3 gene product containing cell fractions. Furthermore, the sequence homology with PON family members within the functionally active protein parts points towards an enzymatic activity of this protein. The definition of its substrate seems of particular interest as the protein is expressed in organs that are involved in detoxification processes. It is additionally present at the surface of monocytes and of endothelial cells. Moreover, C20orf3 gene expression is influenced by human serum. All of these characteristics are in line with human PON family members [30,31], which are known to be involved in antioxidant processes [6,10,32]. Oxidative radicals represent important contributors to atherosclerosis development. There are several indicators for C20orf3 involvement in antioxidant processes. The results of a GEO profile review revealed C20orf3 transcript up-regulation in skeletal muscle of children suffering from Duchenne’s muscular dystrophy. This disease due to impaired dystrophin function is thought to be the result of calcium overflow and oxidative radical production, finally leading to muscle cell death. Furthermore, retinal epithelial cells respond with up-regulation of the gene following treatment with oxidized LDL (GEO Profile, GDS2307). Additionally, the demonstrated underexpression in placentas from patients with pre-eclampsia might also point in this direction [33].

In conclusion, we have identified the protein encoded by the C20orf3 and its protein variants as well as the sites of expression in humans. Owing to sequence similarities and parallels of the expression pattern with PON family members, we hypothesize that the C20orf3 gene product represents a novel member of this enzyme family detected at the cellular level. Its characteristic localization and characterization of the novel protein encoded by C20orf3

493

© The Authors Journal compilation © 2008 Biochemical Society

Figure 7  FACS analysis of human hepatoma cells (HepG2) using mAb56

(A) HepG2 cells stained after trypsinization. The filled curve represents isotype control, the line represents staining with mAb56 after trypsin treatment (7 min), and the hatched line shows HepG2 cells stained with mAb56 after suspending cells via vigorous pipetting and omitting trypsin treatment. (B) Human peripheral blood monocytes co-stained with FITC-labelled mAb56 and PE-A-labelled monocyte marker CD14.

concentration in serum was below the detection limit of our ELISA (35 pg/ml). Surprisingly, the protein was found in urine of some individuals and could be measured in the supernatant of thrombin-treated HepG2 cells (~ 60 pg/ml).
Figure 8  Immunoblot analysis of C20orf3 gene product, zymography and arylesterase activity with phenyl acetate in the presence and absence of Ca²⁺

The rat insulinoma cell line Rin-5F (A) and HepG2 cells (B) were subjected to sucrose density gradient cell fractionation. The individual fractions were loaded on to SDS/PAGE gels, transferred on to nitrocellulose membranes and developed by the C20orf3-specific antibody (A, a, and B, a). For identification of the cellular compartment present at the individual fractions, the blot was redeveloped using a SNAP-25-specific monoclonal antibody (for Rin-5F, A, b) and a NSF-specific monoclonal antibody (for HepG2, B, a). In parallel, the same fractions were loaded on to native PAGE gels and stained for arylesterase activity using β-naphthyl acetate as substrate (A, c, and B, b). The same fractions of Rin-5F (A, d) and HepG2 (B, c) cells were tested for arylesterase activity with phenyl acetate using the spectrophotometric method. The same analysis was performed under Ca²⁺-chelated conditions (A, e, and B, d). Enzyme activity was demonstrated in units (U) defined as 1 μmol of phenyl acetate hydrolysed in 3 min. Fraction numbers are indicated at the bottom, and molecular-mass markers are shown on the left for immunoblot data.
renal, hepatic and endothelial cell expression patterns point towards important biological functions probably based on its substrate specificity or ligand interaction.

REFERENCES


SUPPLEMENTARY ONLINE DATA

Localization and characterization of the novel protein encoded by C20orf3

Aysegul ILHAN*,†, Wolfgang GARTNER*,†, Anastasiya NABOKIKH*, Teodora DANEVA‡, Otto MAJDIC*‡, Gerald COHEN*†, Georg A. BÖHMIG*†, Wolfgang BASE*†, Walter H. HÖRL*† and Ludwig WAGNER*†

*Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria, †Department of Internal Medicine III, Division of Nephrology and Dialysis, Medical University Vienna, Währinger Gürtel 18-20, A-1090 Vienna, Austria, ‡Bulgarian Academy of Sciences, 73 Tzarigradsko Shosse Blvd, 1113 Sofia, Bulgaria, and §Institute of Immunology, Medical University of Vienna, Borschkegasse 8a, A-1090 Vienna, Austria

MATERIALS AND METHODS

RT-qPCR and PCR amplification of C20orf3

After pre-treatment with DNase I for 15 min at room temperature, 1 μg of total RNA was subjected to reverse transcription for 1 h at 40°C using Superscript II (Invitrogen). cDNA synthesis was stopped by heating at 65°C for 15 min. Primers (forward, 5′-ATGAGCCAGCCGACCG-3′; reverse, 5′-TATCTGGAGGGCTAAACAGC-3′) were designed to amplify the coding part of the human gene (1248 bp). PCR was performed using the Expand® High Fidelity PCR System (Boehringer Mannheim) with the following cycling parameters: 35 cycles of 94°C for 30 s for denaturation, 62°C for 30 s for annealing and 72°C for 30 s for synthesis, with a final extension period of 72°C for 7 min.

Cloning of the human C20orf3

Overhangs were added to the gel-purified PCR product using Taq polymerase (Invitrogen). The resultant inserts were ligated into the pGEM®-T-easy vector using T4 ligase (both from Promega) according to the manufacturer’s instructions. Ligated plasmids were transformed into MAX Efficiency® DH5α competent E. coli (Invitrogen) by the heat-shock method. Transformants were plated on to LB (Luria–Bertani)/Amp (ampicillin)/X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) plates. Colonies with the correct insert were screened further by colony PCR using the gene-specific primers indicated above. Colonies containing the correct insert were grown overnight in LB/Amp agar plates and isolated using the Plasmid Mini Kit (Qiagen). Isolated plasmids were subjected to restriction enzyme digestion and DNA sequencing.

Directional cloning of the human C20orf3 protein using Gateway technology

Four nucleotides (CCAC) were added at the 5′-end of the coding region in front of the start codon. For this purpose, it was necessary to choose a second forward primer (5′-CACCATGAGCCAGCCGACCG-3′). The PCR product was obtained as described above using the Expand® High Fidelity PCR System. The amplification product was cloned into the pENTR/D-TOPO vector. The ligated vector was transformed into MAX Efficiency® DH5α competent E. coli, which were plated on to LB/Kan (kanamycin) plates. Colonies with the correct orientation (identified by PCR) were chosen for amplification and plasmids were purified using the Qiagen Plasmid Mini Kit.

Transfer of C20orf3 gene from pENTR/D-TOPO to pcDNA-DEST plasmids

The clonase reaction was performed as indicated in the supplier’s product manual (Invitrogen). Briefly, 300 ng of pENTR/D-TOPO and 300 ng of pcDNA-DEST 15 were combined and incubated with clonase for 1 h at 25°C followed by transformation into MAX Efficiency® DH5α competent E. coli and subsequent selection on LB/Amp agar plates.

Baculovirus preparation and insect cell culture

Preparation of recombinant baculovirus

The cDNA of C20orf3 coding for residues 16–416 was cloned into the baculovirus-compatible pDEST™10 expression vector (Invitrogen). After insert verification, the Bac-to-Bac® baculovirus expression system was chosen following the manufacturer’s instructions (Invitrogen, 10359, version D). Briefly, MAX Efficiency® DH10Bac™ competent E. coli cells containing wild-type bacmid and helper plasmid were transformed with pDEST™10 containing the C20orf3 using a 45 s heat shock at 42°C followed by 4 h of incubation at 37°C under constant rotation (225 rev./min). Recombinants were selected on 7 μg/ml gentamycin, 10 μg/ml tetracycline, 50 μg/ml kanamycin, 300 μg/ml S-Gal® (3,4-cyclohexenoeculetin β-D-galactopyranoside) (Sigma) and 30 μg/ml IPTG (isopropyl-β-D-thiogalactoside)/LB agar plates. White colonies were selected and tested by colony PCR after 48 h of incubation at 37°C and inoculated in LB containing the same combination of antibiotics as described above. Recombinant bacmid was purified under standard conditions such as described in the manufacturer’s instructions.

Sf9 and High Five™ insect cell transfection

Cellfectin®-bacmid transfection mixture was prepared by co-incubation of Cellfectin® reagent (Invitrogen) and bacmid DNA in 200 μl of unsupplemented Grace’s insect medium (Invitrogen) for 20 min at room temperature. Exponentially growing Sf9 and High Five™ cells were exposed to transfection mixture for 4 h which was replaced then by Insect-Xpress (BioWhittaker) for protein expression and baculovirus collection. For maintaining virus titre stability, the resultant fluid was supplemented with 2% calf serum (final concentration). Verification of protein translation and expression in insect cells was performed using the anti-(His6 tag) monoclonal antibody clone HIS.H8 (Upstate Biotechnology) in immunoblotting techniques on Sf9 and High Five™ cell lysates.
A. Ilhan and others

Figure S1  Three-dimensional structure of C20orf3

On the basis of in silico analysis, we deduced the main parts of the tertiary structure of the C20orf3 gene product by alignment with SS. Homologous sequence parts between SS and the C20orf3 gene product are marked in red. The arrows represent $\beta$-sheets. Note the six-bladed $\beta$-propeller structure of the protein.

Purification of Sf9 and High FiveTM cell-expressed C20orf3 protein

At 3 days after infection with recombinant baculovirus, Sf9 or High FiveTM cells were harvested by centrifugation at 2000 $g$ for 5 min followed by lysis in native lysis buffer (300 mM NaCl, 50 mM sodium phosphate buffer, pH 8.0, and 20 mM imidazole). After a brief sonication with eight strokes, the insoluble material was pelleted in a Beckmann J2-MC centrifuge using the JA20 rotor at 5000 rev./min for 15 min at 4°C. The supernatant was incubated with pre-equilibrated Ni-NTA His-Bind® resin (Novagen) for 30 min at 4°C for His6 tag protein binding. The resultant Ni-NTA His6–C20orf3 resin was washed three times using lysis buffer as described above. Elution of the recombinant protein was carried out using the 1× Ni-NTA elution buffer (300 mM NaCl, 50 mM sodium phosphate buffer, pH 8.0, and 250 mM imidazole).

Purification of inclusion bodies

E. coli bacteria (BL21) transformed with pDESTTM15-C20orf3 vector were grown in TY medium [1.6% (w/v) tryptone, 1% (w/v) yeast extract and 0.5% (w/v) NaCl]. After induction for 4 h at 31°C using 100 $\mu$M IPTG, bacteria were harvested by centrifugation at 4000 $g$ for 10 min. Lysis was performed in 50 mM Tris/HCl (pH 8.0), 100 mM NaCl, 5 mM EDTA, 0.1% sodium azide and 0.1 mM PMSF. After sonication, EDTA was chelated by MgSO4 (10 mM final concentration) and DNase (0.01 mg/ml final concentration) and lysozyme (0.1 mg/ml final concentration) were added. After 20 min of incubation at room temperature, inclusion bodies were pelleted by centrifugation at 12000 rev./min for 14 min in 14 mm × 95 mm polyallomer centrifuge tubes (Beckman Instruments) using the SW 40 Ti rotor (model L-80, Beckman). The resultant inclusion body pellet was resuspended in E. coli lysis buffer as indicated above. Washing of inclusion bodies was repeated four times. Purity testing was carried out by SDS/PAGE (12% gels) analysis. Particles were dissolved in 8 M urea for further purification by Superose 6 (GE Healthcare) column chromatography using 6 M urea as mobile phase. The collected fractions containing the fusion protein were selected after SDS/PAGE (12% gels) and Coomassie Blue staining. Pooled fractions were dialysed against PBS, which resulted in precipitation of the fusion protein. The precipitate was collected and frozen at −80°C until further use.

Generation of immunoreagents

The Superose 6 column-purified GST–C20orf3 fusion protein was emulsified in Freund’s adjuvant and used for immunization of 6-week-old Balb/c mice by subcutaneous injection. The monoclonal antibody was generated as described previously [1]. In brief, the spleen of the immunized mouse was removed 4 days after the last immunization, and separated spleen cells were fused with the P3 cells growing in exponential phase under standard conditions. Outgrowing clones were screened using a direct ELISA technique. Plates were coated overnight with recombinant protein (residues 58–416). Additionally, all clones were tested in parallel by immunoblotting using MilliBlot-MP membrane processor (Millipore Corporation). Poyclonal mouse antibody was generated by three immunizations at 2 week intervals using GST–C20orf3 fusion protein. The antiserum obtained was pre-adsorbed to recombinant GST protein-covered glutathione–Sepharose 4B. Monospecificity of the resultant antibody was tested by parallel immunoblotting against GST and the fusion protein. Antibody not detected by GST was used for further procedures. Polyonal rabbit antibody was produced by immunizing 6-week-old rabbit at 22 day intervals using the GST–C20orf3 fusion protein. In addition, a peptide-specific antibody was generated against the C-terminally KLH (keyhole-limpet haemocyanin)-conjugated C20orf3-specific peptide GLRQRRPLRPQVVTC (amino acids 6–19). The antibody directed against the cytosolic part (N-terminus) of the protein was used for characterization of protein variants containing the cytosolic domain.

REFERENCE