DETAILED METHODS

for

"Dominant negative mutation in oxalate transporter *SLC26A6* associated with enteric hyperoxaluria and nephrolithiasis" by Nicolas Cornière, R. Brent Thomson, Stéphanie Thauvin, Bruno O. Villoutreix, Sophie Karp, Diane W. Dynia, Sarah Burlein, Lennart Brinkmann, Alaa Badreddine, Aurélie Dechaume, Mehdi Derhourhi, Emmanuelle Durand, Emmanuel Vaillant, Philippe Froguel, Régine Chambrey, Peter S. Aronson, Amélie Bonnefond, & Dominique Eladari

Patients testing & ethics

Biochemical and hormone measurements were performed using standard techniques on an Alinity C auto-analyzer (Abbott Diagnostics, Rungis, France). Urinary oxalate excretion was measured by an enzymatic method using oxalate oxidase. Normal values for oxalate excretion were determined on healthy volunteers in our service and define as the lower and upper limits of the 95% confidence interval. Creatinine was measured using an isotope dilution-mass spectrometry (IDMS)-traceable enzymatic creatinine assay. Estimated GFR (eGFR) was determined using plasma creatinine value by the CKD Epidemiology Collaboration formula.[1] The relationship between parathormone (PTH) secretion and calcemia was tested using the calcium-loading test as described by Pak and Broadus.[2,3] This test also allowed assessing apparent digestive calcium absorption and net bone resorption.

All participants gave written informed consent for genetic study, and genetic testing was performed in accordance with French legislation on genetics diagnostics tests (French bioethics law no. 2004-800). The study was approved by the institutional review board of the University Hospital of La Réunion, and conducted in accordance with the Declaration of Helsinki. According to the French law, the study was declared and registered in the public database of the French National Institute of Health Data (N° MR 2105161219), and all data were collected and protected as requested by the Commission Nationale de l'Informatique et des Libertés (CNIL), collection of data was registered by the CNIL (N° 220739 v.0).

Participants from UK Biobank. We analyzed up to 200,619 samples, with available exome sequencing data. This research is part of UK Biobank research application #67575. The participants were recruited by UK Biobank from across the United

Kingdom between 13 March 2006 and 1 October 2010. People with kidney stone were defined as follows: 'at least one ICD-10 code starting with N20' (Calculus of kidney and ureter) or 'at least one ICD-10 code starting with N21' (Calculus of lower urinary tract) or 'Calculus of urinary tract in other diseases classified elsewhere' (ICD10-N22.8) or 'Unspecified renal colic' (ICD10-N23) or 'Unspecified renal colic' (ICD10-N23) or 'Other specified disorders of carbohydrate metabolism (oxaluria)' (ICD10-E748) or 'Other disorders of calcium metabolism (nephrocalcinosis)' (ICD10-E8359) or 'self-reported kidney stone' (field #20002).

DNA sequence analyses

The DNA proband was assessed through whole-exome sequencing. For this purpose, we used the Human Core Exome EF Multiplex Complete kit (Twist Bioscience) in combination with Illumina next-generation sequencing. Briefly, 50 ng DNA was fragmented through enzymatic fragmentation. The fragmented DNA samples were end-repaired, ligated to the Twist indexed adapters and purified through DNA purification beads. These samples were subsequently amplified by PCR using Illumina amplification primers and KAPA HiFi HotStart Ready mix, and purified through DNA purification beads. After size selection (between 375 bp and 425 bp) and sample quantification (Agilent BioAnalyzer), eight samples including the proband were combined in a single pool of at least 1.5 µg, and hybridized to the Twist biotin-labeled exome probes. After 16 hours at 70 °C, the hybridized targets were bound to streptavidin beads following the recommendations. After washing unspecific hybridization, the captures were subsequently amplified using the KAPA HiFi HotStart ReadyMix and quantified by both Perkin Elmer LabChip GX and Agilent BioAnalyzer. Then, the samples were sequenced on the Illumina NovaSeq 6000 system, using a paired-end 2×100 bp protocol. The demultiplexing of sequence data was performed using bcl2fastq Conversion Software v2.19.1 (Illumina). Subsequently, sequence reads were mapped to the human genome (hg38/GRCh38) using Burrows-Wheeler Aligner v0.7.17.[4] The mean depth of coverage of the target was 44.9×, with 99.08% of the target covered with more than 8 reads. The variant calling was performed using Genome Analysis ToolKit v3.8-1 (GATK).[5] Only variants with coverage higher than 8 reads were kept for further analyses. The annotation of variants was performed using the Ensembl Variant Effect Predictor tool (VEP) v100 with data from COSMIC (v90),

ClinVar (v201912), dbSNP (v153), dbNSFP (v4.0a), GENCODE (v34), gnomAD (v2.1) databases.[6]

The *SLC26A6* mutation was then confirmed in the proband, and then assessed in the family members through Sanger sequencing. Primer sequences and PCR conditions are available upon request. Fragments were sequenced in both directions, and subsequently analyzed using the 3730xl DNA Analyzer (Applied Biosystems). Electropherogram reads were assembled and examined using the SeqScape software (Applied Biosystems).

Detection of variants in UK Biobank. We used exome data from pVCF format (field #23156). Only variants with a coverage higher than 10 reads and quality GQ score higher than 20 were kept for further analyses. Annotation of variants was done using the Ensembl Variant Effect Predictor (VEP) tool version 103 (RefSeq).

SLC26A6 Expression Constructs. The generation of the HA-tagged wild-type human SLC26A6 expression construct (HA-WT) was described previously.[7] Briefly, a commercially available cDNA clone (Origene) encoding the entire transcript variant 1 isoform of human *SLC26A6* was subcloned into the retroviral mammalian expression vector pLNCX2 (Clontech). This construct was used as the template to generate the *N*-terminal HA-tagged p.R507W human SLC26A6 mutant (HA-MT) with a Q5® site-directed mutagenesis kit (New England Biolabs). The forward and reverse primers used to generate HA-MT were 5'-CGTGGTGGTCTGGACACAGAT and 5'-AGCAGCAGGGAGAAGATG respectively. HA-WT was also used as the template to develop an *N*-terminal *myc*-tagged variant (*myc*-WT). The HA-epitope tag was replaced with the *myc*-epitope tag in three steps with the Q5® site-directed mutagenesis kit with primers:

F1:5'-GAAAGATGTTCCAGATTACGCTCTGCGGAG;
R1:5'TGCTCGTCCATGGCGGCCGCCTC;
F2:5'-TCTCAGATTACGCTCTGCGGA;
R2:5'-TGAGTTTCTGCTCGTCCATGGC;
F3:5'-GATCTGCGGAGGCGAGACTAC;
R3:5'-CTCTTCTGAGATGAGTTTCTGCTCG.

PCR cycling conditions and quantities of template and primers used for each reaction were exactly as specified in the Q5® kit guidelines (New England Biolabs). HA-WT, *myc*-WT, and HA-MT were sequence validated in their entirety by the Yale Keck DNA Sequencing Facility.

Cell Culture and Transfections. OKP cells (ATTC) were utilized for all transfection experiments. They are a well characterized polarized epithelial cell line that has been used extensively to study substrate transport, protein biosynthesis, membrane trafficking, and post-translational modification of a variety of plasma membrane proteins. They are extremely amenable to standard transient transfection techniques, they do not produce significant quantities of mucus that could potentially alter efficacy of cell-surface biotinylation, and, most importantly, they do not express significant quantities of endogenous SLC26A6. OKP cells (ATTC) were maintained in high glucose DMEM (Gibco) supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 10% FBS, and penicillin-streptomycin (100 U/ml) at 37°C in 5% CO_2. Cells were monitored routinely for mycoplasma infection by a PCR-based mycoplasma detection assay (ATTC). $1x10^5$ OKP cells per well of a 24 wells cell culture dish were reverse transiently transfected with 2 µl of Lipofectamine 2000 (Thermo Scientific) and 0.1-0.25 µg of HA-WT, *myc*-WT, or HA-MT cDNA as indicated in the respective figure legend for each experiment. Cells were assayed 72 hours post transfection.

Cell Surface Biotinylation. OKP cells were surface biotinylated 72 hours post transfection as described previously.[7] Biotinylations were performed at 4°C and all solutions were ice cold. Briefly, each well of a 24 well plate was washed 3X with complete PBS (containing 0.1 mM $CaCl_2$ and 1 mM $MgCl_2$), surface labeled with 0.25 mg of membrane impermeant EZ-Link Sulfo-NHS-LC-Biotin labeling reagent (Thermo Scientific) in a 10 mM triethanolamine buffer containing 2 mM $CaCl_2$ and 150 mM NaCl, pH 7.4 for 20 minutes (2X), washed 1X with complete PBS, washed 3X (5 minutes each) with 100 mM glycine in complete PBS to quench any unbound biotinylation reagent, washed 1X with complete PBS, and solubilized for 30 minutes at 4°C in normal salt RIPA buffer (150 mM NaCl, 50 mM Tris-HCl, 5 mM EDTA, 0.5% sodium deoxycholate, 1% Triton X-100, 0.1% SDS, pH 7.4) containing protease inhibitors (Halt; Thermo Scientific). The RIPA lysate was cleared by centrifugation at 15k x g at 4°C for 10 minutes. 10% of the cleared lysate was saved for SDS-PAGE

and the remaining fraction was incubated with Neutravidin Plus Ultralink Resin (Thermo Scientific) for 2 hours at 4°C to capture biotin-labeled surface proteins. After incubation the Neutravidin resin was washed 3X with normal salt RIPA buffer, 2X with high salt RIPA buffer (containing 500 mM NaCl) and then 2X with normal salt RIPA buffer. All washes were 10 minutes in length and performed at 4°C with ice cold wash buffers containing protease inhibitors (Halt; Thermo Scientific). Biotin-labeled proteins were eluted from the Neutravidin resin by incubation with 100°C 2X SDS-PAGE loading buffer containing 100 mm DTT for 2 minutes. The total lysate and cell-surface biotinylated fractions were subjected to standard SDS-PAGE and then transferred to a PVDF membrane (Immobilon-P, 0.45 µm; EMD Millipore) as described previously for subsequent analysis.[7] Specificity of cell surface labeling was confirmed by absence of biotin-labeling of the abundant intracellular protein, β-actin (see Supplement Figure S5).

Western analysis. PVDF membranes were probed as described previously with a rabbit anti-SLC26A6 polyclonal antibody directed against a synthetic peptide derived from the *N*-terminal region of human SLC26A6 (R29; 1:50,000 dilution),[7] a commercially available rabbit anti-HA polyclonal antibody (71-5500; 1:250 dilution; Thermo Scientific), or a commercially available mouse anti-*myc* monoclonal antibody (R950-25; 1:5000 dilution; Thermo Scientific). Primary antibody labeling was detected with either a donkey anti-rabbit or a donkey anti-mouse HRP-labeled secondary antibody (711-035-152 and 715-035-150 respectively; 1:20,000 dilution; Jackson ImmunoResearch), visualized by enhanced chemiluminescence (Clarity Western ECL Substrate; Bio-Rad), and recorded on film. R29 specificity was previously confirmed by Western analysis of duodenal lysates from wild-type and *Slc26a6*-null mice.[7] Specificity of anti-HA and anti-*myc* antibodies and validation of epitope-tagged SLC26A6 expression constructs was verified by comparative Western analysis of lysates from untransfected and HA-WT, *myc*-WT, and HA-MT transfected OKP cells.

Co-immunoprecipitation.

OKP cells were transfected as above and cultured for 72 hours prior to processing for co-immunoprecipitation. Cells were washed 3X with cold PBS and then solubilized for 1 hr with cold normal salt lysis buffer containing 50mM MOPS, 100mM NaCl, pH 7.2, 1% digitonin (Sigma D141), protease inhibitors (Roche Complete Ultra), and

phosphatase inhibitors (Roche PhosSTOP). Insoluble material was removed by centrifugation at 15k x g for 10 min at 4°C. *myc*-tagged WT SLC26A6 was immunoprecipitated by incubation with the anti-*myc* antibody R950-25 (Thermo; 5 µg IgG/ml lysate) for 2 hr at 4°C. Immunoprecipitates were captured by incubation with Protein G Sepharose Fast Flow (Sigma P3296) for 1 hr at 4°C. Sepharose beads were then washed (10 min each) at 4°C 3X with normal salt lysis buffer, 2X with high salt lysis buffer containing 500mM NaCl, and then 2X with normal salt lysis buffer. Immunoprecipitates were released from sepharose beads by incubation with 2X SDS-PAGE sample buffer containing 100mM DTT for 2 min at 100°C. The immunoprecipitates were subjected to SDS-PAGE, transferred to a PVDF membrane, and then analyzed by Western blot with antibodies directed against the HA- and *myc*-epitope tags (71-5500 and R950-25 antibodies respectively).

Cl^--dependent ^{14}C-oxalate uptake. See reference[7] for a detailed description of the rationale and methodology behind the assessment of Cl^--gradient dependent ^{14}C-oxalate uptake mediated by the SLC26A6 expression constructs (HA-WT and HA-MT) when transfected into OKP cells. Briefly, uptakes were performed for each transfection condition in the presence and absence of an outwardly directed chloride gradient to provide an estimate of the proportion of the total ^{14}C-oxalate cell uptake that was specifically chloride-dependent. Uptakes were also performed in untransfected cells (exposed to transfection reagent alone) to provide an estimate of background levels of endogenous chloride-dependent ^{14}C-oxalate uptake. The background values were then used to normalize expression of chloride-dependent ^{14}C-oxalate uptake to that directly and solely attributable to the activity of the transfected SLC26A6 expression constructs. See Suppl. Fig. S3 for a detailed example of data handling for a representative transfection event.

All uptakes were performed at room temperature and with the exception of the ice-cold Stop Buffer all uptake-related solutions were held at room temperature for the duration of each uptake experiment. Briefly, OKP cells were maintained in normal chloride high-glucose DMEM (as above) for 72 hours post transfection before each uptake experiment. Cells were washed 1X for 2 min with a chloride-free K-gluconate buffer (130 mM K-gluconate, 5 mM glucose, 20 mM HEPES, pH 7.4) to minimize residual extracellular chloride before incubating with 0.2 µCi (8 µM) ^{14}C-oxalate for 2 minutes in either the presence (200 µl K-gluconate buffer; as above) or absence (200

µl KCl buffer; 130 mM KCl, 5 mM glucose, 20 mM HEPES, pH 7.4) of an outwardly directed chloride gradient. Cell uptake of ^{14}C-oxalate was terminated by a series (3X) of 2-minute washes with ice cold K-gluconate (as above) stop buffer. The stop buffer was removed and cells were lysed by sequential 5-minute incubations of each well with 200 µl 0.1 N NaOH followed by 200 µl 0.1 N HCl. The lysate was added to 5 ml of Opti-Fluor (PerkinElmer) scintillation cocktail and counted in a Tri-Carb 2910 TR (PerkinElmer) liquid scintillation analyzer.

Statistics. For the transfection events described in Figures 1 and 2 uptakes and companion biotinylations were performed in triplicate (3 wells each condition) and spot estimates were averaged to generate a single value. Each averaged value from each transfection event represents n=1 for subsequent statistical analyses. Therefore n=3 represents 3 independent transfection events with 3 wells per transfection for each condition for both ^{14}C-oxalate uptake and cell surface biotinylation assessment. Due to the more extensive nature of the Western analysis required for the cotransfection study described in Figure 3 each value represents data obtained from a single well. Therefore n=4 represents 4 independent transfections with 1 well per transfection for each condition. Data values are presented as mean ± SEM. Statistical significance was evaluated by unpaired two-tailed Student's t-test (GraphPad Prism).

In Silico analyses of the effects of the p.R507W mutation on SLC26A6 protein

The human SLC26A6 protein sequence (Q9BXS9) was downloaded from the UniProt database.[8] Eight other sequences from different species were selected (Pan troglodytes, Rattus norvegicus, Mus musculus, Canis familiaris, Gallus gallus, Xenopus tropicalis, Tetraodon nigroviridis, and Danio rerio) in order to investigate inter-species amino acid conservation. Multiple sequence alignment was performed with Clustal Omega,[9] via the EMBL-EBI web services.[10] The Consurf server was used to obtain amino acid conservation scores within the family by comparing 150 homologous sequences.[11] A comparative model for the human SLC26A6 protein was developed using the SwissModel server, [12] and the experimental structure of the mouse Slc26a9 homodimer anion transporter was used as structural template.[13] Alternative sequence to structure alignments were investigated using the PROMALS3D server to select the final alignment to use for the modeling procedure.[14] The SLC26A6 3D model structure was geometry optimized using the PREFMD structure refinement via molecular

dynamics simulation package[15] while the analysis of the electrostatic properties was carried out with our PCE server.[16]

The human SLC26A6 3D model was inserted in a virtual membrane using the PPM server.[17] A fast estimation of the protein flexibility was performed using the PredyFlexy server that classifies amino acid residues into three class: rigid, intermediate or flexible sites.[18] The PyMOL Molecular Graphics System (Version 1.8.2.2 Schrödinger, LLC) and Chimera were both used for the interactive structural analysis.[19] The assessment of the protein stability was carried out with two different approaches DUET,[20] and SAAFEC.[21] These tools compute ΔΔG values between the wild-type and the variant proteins. The Residue Interaction Network Generator (RING) software was used to gain additional insights into the structures of the human SLC26A6 protein 3D model through visualization of non-bonded interactions.[22] The RING-2.0 server was used in our study.[23] The generated RING network XML file was then analyzed with Cytoscape.[24] Twenty-six parameters that can help predict the impact of an amino acid substitution on the structure and function of a protein were investigated using some of the above-mentioned computations and after interactive structural analysis.[25-28] To condense the resulting information into a single figure, we developed a Python script that makes use of the Matplotlib plotting library so as to generate a flower-donut traffic light chart.

Supplementary References

1. Inker LA, Schmid CH, Tighiouart H, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. N Engl J Med 2012;367:20-9.
2. Pak CY, Kaplan R, Bone H, et al. A simple test for the diagnosis of absorptive, resorptive and renal hypercalciurias. N Engl J Med 1975;292:497-500.
3. Broadus AE, Dominguez M, Bartter FC. Pathophysiological studies in idiopathic hypercalciuria: use of an oral calcium tolerance test to characterize distinctive hypercalciuric subgroups. J Clin Endocrinol Metab 1978;47:751-60.
4. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.
5. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297-303.

6. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. Genome Biol 2016;17:122.
7. Thomson RB, Thomson CL, Aronson PS. N-glycosylation critically regulates function of oxalate transporter SLC26A6. Am J Physiol Cell Physiol 2016;311:C866-C73.
8. UniProt C. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 2019;47:D506-D15.
9. Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 2011;7:539.
10. Madeira F, Madhusoodanan N, Lee J, et al. Using EMBL-EBI Services via Web Interface and Programmatically via Web Services. Curr Protoc Bioinformatics 2019;66:e74.
11. Ashkenazy H, Abadi S, Martz E, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res 2016;44:W344-50.
12. Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018;46:W296-W303.
13. Walter JD, Sawicka M, Dutzler R. Cryo-EM structures and functional characterization of murine Slc26a9 reveal mechanism of uncoupled chloride transport. Elife 2019;8.
14. Pei J, Grishin NV. PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. Methods Mol Biol 2014;1079:263-71.
15. Heo L, Feig M. PREFMD: a web server for protein structure refinement via molecular dynamics simulations. Bioinformatics 2018;34:1063-65.
16. Miteva MA, Tuffery P, Villoutreix BO. PCE: web tools to compute protein continuum electrostatics. Nucleic Acids Res 2005;33:W372-5.
17. Lomize MA, Pogozheva ID, Joo H, et al. OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Res 2012;40:D370-6.
18. de Brevern AG, Bornot A, Craveur P, et al. PredyFlexy: flexibility and local structure prediction from sequence. Nucleic Acids Res 2012;40:W317-22.

19. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605-12.
20. Pires DE, Ascher DB, Blundell TL. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. Nucleic Acids Res 2014;42:W314-9.
21. Getov I, Petukh M, Alexov E. SAAFEC: Predicting the Effect of Single Point Mutations on Protein Folding Free Energy Using a Knowledge-Modified MM/PBSA Approach. Int J Mol Sci 2016;17:512.
22. Martin AJ, Vidotto M, Boscariol F, et al. RING: networking interacting residues, evolutionary information and energetics in protein structures. Bioinformatics 2011;27:2003-5.
23. Piovesan D, Minervini G, Tosatto SC. The RING 2.0 web server for high quality residue interaction networks. Nucleic Acids Res 2016;44:W367-74.
24. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498-504.
25. Kucukkal TG, Petukh M, Li L, et al. Structural and physico-chemical effects of disease and non-disease nsSNPs on proteins. Curr Opin Struct Biol 2015;32:18-24.
26. Thusberg J, Vihinen M. Pathogenic or not? And if so, then how? Studying the effects of missense mutations using bioinformatics methods. Hum Mutat 2009;30:703-14.
27. Villoutreix BO. Structural bioinformatics: methods, concepts and applications to blood coagulation proteins. Curr Protein Pept Sci 2002;3:341-64.
28. Wang Z, Moult J. SNPs, protein structure, and disease. Hum Mutat 2001;17:263-70.