Recessive osteogenesis imperfecta caused by LEPRE1 mutations: clinical documentation and identification of the splice form responsible for prolyl 3-hydroxylation

A Willaert, F Malfait, S Symoens, K Gevaert, H Kayserili, A Megarbane, G Mortier, J G Leroy, P J Coucke and A De Paepe

Updated information and services can be found at:
http://jmg.bmj.com/cgi/content/full/46/4/233

These include:

Data supplement	*"web only appendix"* **http://jmg.bmj.com/cgi/content/full/46/4/233/DC1**
References	This article cites 22 articles, 7 of which can be accessed free at: **http://jmg.bmj.com/cgi/content/full/46/4/233#BIBL**
Rapid responses	You can respond to this article at: **http://jmg.bmj.com/cgi/eletter-submit/46/4/233**
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Notes

To order reprints of this article go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to *Journal of Medical Genetics* go to:
http://journals.bmj.com/subscriptions/

Recessive osteogenesis imperfecta caused by *LEPRE1* mutations: clinical documentation and identification of the splice form responsible for prolyl 3-hydroxylation

A Willaert,[1] F Malfait,[1] S Symoens,[1] K Gevaert,[2,3] H Kayserili,[4] A Megarbane,[5] G Mortier,[1] J G Leroy,[1] P J Coucke,[1] A De Paepe[1]

[1] Department of Medical Genetics, Ghent University Hospital, Ghent, Belgium; [2] Department of Biochemistry, Ghent University, Ghent, Belgium; [3] Department of Medical Protein Research, VIB, Ghent, Belgium; [4] Medical Genetics Department, Istanbul Medical Faculty, Istanbul University, Istanbul, Turkey; [5] Medical Genetics Unit, Saint-Joseph University, Beirut, Lebanon

Correspondence to:
Dr A De Paepe, Department of Medical Genetics, Ghent University Hospital, De Pintelaan 185, B-9000 Ghent, Belgium; anne.depaepe@UGent.be

AW and FM contributed equally to this work

ABSTRACT

Background: Recessive forms of osteogenesis imperfecta (OI) may be caused by mutations in *LEPRE1*, encoding prolyl 3-hydroxylase-1 (P3H1) or in *CRTAP*, encoding cartilage associated protein. These proteins constitute together with cyclophilin B (CyPB) the prolyl 3-hydroxylation complex that hydroxylates the Pro986 residue in both the type I and type II collagen α1-chains.

Methods: We screened *LEPRE1*, *CRTAP* and *PPIB* (encoding CyPB) in a European/Middle Eastern cohort of 20 lethal/severe OI patients without a type I collagen mutation.

Results: Four novel homozygous and compound heterozygous mutations were identified in *LEPRE1* in four probands. Two probands survived the neonatal period, including one patient who is the eldest reported patient ($17^{7/12}$ years) so far with P3H1 deficiency. At birth, clinical and radiologic features were hardly distinguishable from those in patients with autosomal dominant (AD) severe/ lethal OI. Follow-up data reveal that the longer lived patients develop a severe osteochondrodysplasia that overlaps with, but has some distinctive features from, AD OI. A new splice site mutation was identified in two of the four probands, affecting only one of three *LEPRE1* mRNA splice forms, detected in this study. The affected splice form encodes a 736 amino acid (AA) protein with a ''KDEL'' endoplasmic reticulum retention signal. While western blotting and immunocytochemical analysis of fibroblast cultures revealed absence of this P3H1 protein, mass spectrometry and SDS-urea-PAGE data showed severe reduction of α1(I)Pro986 3-hydroxylation and overmodification of type I (pro)collagen chains in skin fibroblasts of the patients.

Conclusion: These findings suggest that the 3-hydroxylation function of P3H1 is restricted to the 736AA splice form.

Osteogenesis imperfecta (OI) comprises a heterogeneous group of conditions characterised by a variable degree of bone fragility with susceptibility to pathological fractures, short stature, bowing of the long bones, wormian bones and moderate to severe kyphoscoliosis.[1] Dentinogenesis imperfecta, blue sclerae, hearing loss and joint hyperlaxity are variably associated features. OI is usually inherited as an autosomal dominant (AD) disorder, caused by a heterozygous mutation in either *COL1A1* (OMIM *120150) or *COL1A2* (OMIM *120160), encoding the α1 and the α2 chain of type I (pro)collagen, respectively. Type I (pro)collagen

is the major extracellular matrix protein of both bone and skin. It is composed of two α1(I) chains and one α2(I) chain, each consisting of 338 uninterrupted repeats of the Gly-Xaa-Yaa motif, in which Xaa is mostly proline (Pro) and Yaa often hydroxyproline (Hyp). This constellation allows the formation of a triple helix structure.[2] To facilitate folding and stability of the triple helix, different post-translational modifications are required in the endoplasmic reticulum (ER).[3] Throughout the helical region of the collagen chains, multiple lysine and proline residues at the Yaa position in the Gly-Xaa-Yaa motif are hydroxylated by lysyl hydroxylase and prolyl 4-hydroxylase, respectively. Only one single proline residue, Pro986, at the Xaa position in the α1 chain of type I collagen, is subjected to 3-hydroxylation. This is accomplished by the prolyl 3-hydroxylation complex, consisting of the enzyme prolyl 3-hydroxylase-1 (P3H1, encoded by *LEPRE1*) (OMIM *610339), its helper protein cartilage associated protein (CRTAP) (OMIM *605497), and the prolyl cis/trans isomerase cyclophilin B (or peptidylprolyl isomerase B) (CyPB or PPIB) (OMIM *123841), in a 1:1:1 ratio.[4]

Although some recurrences in sibs of unaffected parents have been explained by parental mosaicism for a mutation in one of the type I collagen genes, the existence of recessive forms of OI has long been suspected.[5] Moreover, several new types of OI, outside the Sillence classification into four subtypes (OI type I-IV), have been delineated, in some of which autosomal recessive (AR) inheritance has been documented.[6] Overmodification of the proα-chains of type I collagen has been demonstrated in some of these recessive forms, despite the absence of mutations in either *COL1A1* or *COL1A2*. Recent reports of patients with homozygous or compound heterozygous mutations in the *CRTAP* or *LEPRE1* gene and severely reduced 3-hydroxylation of Pro986 in the α1(I)-collagen chain have proved the existence of AR inheritance in a small fraction of patients with severe/lethal types of OI.[7-10]

We present the clinical and molecular findings of four unrelated probands carrying novel mutations in *LEPRE1*. The phenotype varies from perinatal lethal to ''surviving'' phenotypes with the currently longest living patient aged 17 years 7 months.

Original article

SUBJECTS AND METHODS

Subjects

The patients presented here belong to a larger study cohort of 20 probands with severe or lethal OI in whom no COL1A1 or COL1A2 mutation was identified. Hence, they were screened for LEPRE1, CRTAP and PPIB mutations. The phenotype of each patient was clinically and radiographically compatible with OI type II or III. Cultured skin fibroblasts were obtained from 15 probands. Nine of them showed overmodified type I collagen chains, whereas six probands showed a normal electrophoretic pattern for type I collagen. In the remaining five probands, only a blood sample was available. Informed consent to perform the biochemical and molecular studies was obtained from the patients and/or their parents.

Mutation and haplotype analysis

Genomic DNA (gDNA) and mRNA were extracted from fibroblast cultures using the Puregene DNA Isolation kit (Gentra Systems, Minneapolis, Minnesota, USA) and the RNeasy Mini Kit (Qiagen, Frankfurt, Germany), respectively. cDNA was synthesised using the SuperScript II reverse transcriptase kit with random hexamer primers (Invitrogen, Carlsbad, California, USA). For gDNA screening touchdown polymerase chain reaction (PCR) amplification of all exons of the LEPRE1, CRTAP and PPIB genes was performed using forward and reverse primers located in the flanking introns, while additional exonic primers were used for cDNA screening (primer sequences and PCR conditions available upon request). All fragments were directly sequenced on the ABI PRISM 3730XL automated sequencer (Applied Biosystems, Foster City, California, USA) using the BigDye terminator cycle sequencing chemistry and compared to the wild-type sequence as submitted to GenBank (LEPRE1: gene, NM_022356.3; protein, NP_071751.3 – CRTAP: gene, NM_006371.4; protein, NP_006362.1 – PPIB: gene, NM_000942.4; protein, NP_000933.1). Nucleotides were numbered starting from the first base in the start codon (ATG) in the cDNA reference sequence. Amino acids were numbered from the first methionine residue. Compound heterozygous mutations were verified to be in trans position by sequencing the gDNA in both parents. Control alleles from 50 non-OI individuals were sequenced to check the causality of the encountered splice site mutation. Microsatellite markers and SNPs in the LEPRE1 region were taken from the Marshfield map and the HAPMAP, respectively. Genotyping was performed on an Applied Biosystems Prism 3130XL automated sequencer (Applied Biosystems). The data were processed using Genemapper software (Applied Biosystems).

Cloning of PCR fragments

PCR products comprising the three LEPRE1 intron 14 splice forms were cloned using the TOPO TA Cloning system (Invitrogen) following the instructions of the One Shot Chemical Transformation Protocol supplied by the manufacturer.

Real-time quantitative PCR (qPCR)

Skin fibroblasts were grown in 6 cm dishes to 80% confluency. RNA was isolated with the RNeasy Mini Kit (Qiagen) and cDNA synthesis was performed using the SuperScript II reverse transcriptase kit with random hexamer primers (Invitrogen) in a total volume of 20 µl. First, an in silico assay evaluation for LEPRE1 qPCR primers was performed with RTPrimerDB.[11] Amplification efficiency (E) for each primer set was determined on the basis of a sixfold cDNA dilution series from Human qPCR Reference Total RNA (Stratagene, Cedar Creek, Texas, USA). Only primer pairs with E ⩾85% were included for analysis. PCR reaction mixtures contained Roche LightCycler SYBR Green Master Mix, 5 µM of each forward and reverse primer and 20 ng cDNA. Cycling conditions were as follows: 5 min at 95°C, 45 cycles at 95°C for 10 s, 60°C for 30 s and 72°C for 1 s. Subsequently, a melting curve (60–95°C) was generated for every amplicon to check PCR reaction specificities. qPCR analysis was performed on the Lightcycler 480 (Roche, Pleasanton, California, USA) device. All reactions were carried out in duplicate and normalised to the geometric mean of two stable reference genes, GAPDH and YWHAZ, using qBasePlus software.[12] Expression levels were determined in three independent experiments for each cell line. Differential gene expression was considered significant if the means differed by at least 50% and the 95% confidence intervals of the means did not overlap (equivalent with $p < 0.05$).

Western blotting

Dermal fibroblasts were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum, antibiotics, 50 µg/ml L-ascorbic acid (Sigma, St Louis, Missouri, USA) and 230 µg/ml L-proline (Sigma). Confluent cultures were lysed in RIPA buffer (Sigma), supplemented with a protease inhibitor cocktail (Roche). Protein concentrations of the lysates were measured using the BCA Protein Assay (Pierce). Samples (30 µg protein) were subjected to 10% SDS-PAGE and then transferred onto PVDF (polyvinylidene fluoride) membranes (Biorad, Hercules, California, USA). The membranes were blocked overnight at 4°C in 2% milk powder solution and incubated with primary antibody for 90 min: mouse anti-LEPRE1 MaxPab polyclonal antibody (Abnova, Taipei, Taiwan) at a 1:1000 dilution and mouse anti-GAPDH monoclonal antibody (Sigma) at a 1:6000 dilution, both in a 0.5% milk powder solution. After washing, the membranes were incubated with secondary antibody for 60 min: goat anti-mouse HRP-conjugated (Pierce, Rockford, Illinois, USA), diluted 1:1000 in a 0.5% milk powder solution. The membranes were washed again before detection of bound antibodies by enhanced chemiluminescence with the SuperSignal West Dura Extended Duration Substrate Kit (Pierce). To identify P3H1 proteoglycan core proteins, fibroblast cell extracts and concentrated medium fractions were treated with chondroitinase ABC (Sigma)[13] before western blot analysis.

Immunocytochemistry

We grew 3.10^4 skin fibroblasts from patients and controls in 8 well chamber slides (Nunc) for 72 h in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum, antibiotics, 50 µg/ml L-ascorbic acid (Sigma) and 230 µg/ml L-proline (Sigma). After washing in TBS, the cells were permeabilised and fixed in ice cold acetone (−20°C) for 10 min and blocked in 10% goat serum/TBS for 30 min. Next, the cells were incubated with 1:100 LEPRE1 MaxPab polyclonal antibody (H00064175-B01, Abnova), 1:500 collagen I polyclonal antibody (Ab292-100, Abcam, Cambridge, UK) or 1:200 calnexin polyclonal antibody (Ab13504, Abcam) in 10% goat serum/TBS for 150 min at room temperature. The cells were then washed three times in TBS and incubated with fluorochrome conjugated secondary antibody in 10% goat serum/TBS for 1 h (1:1000
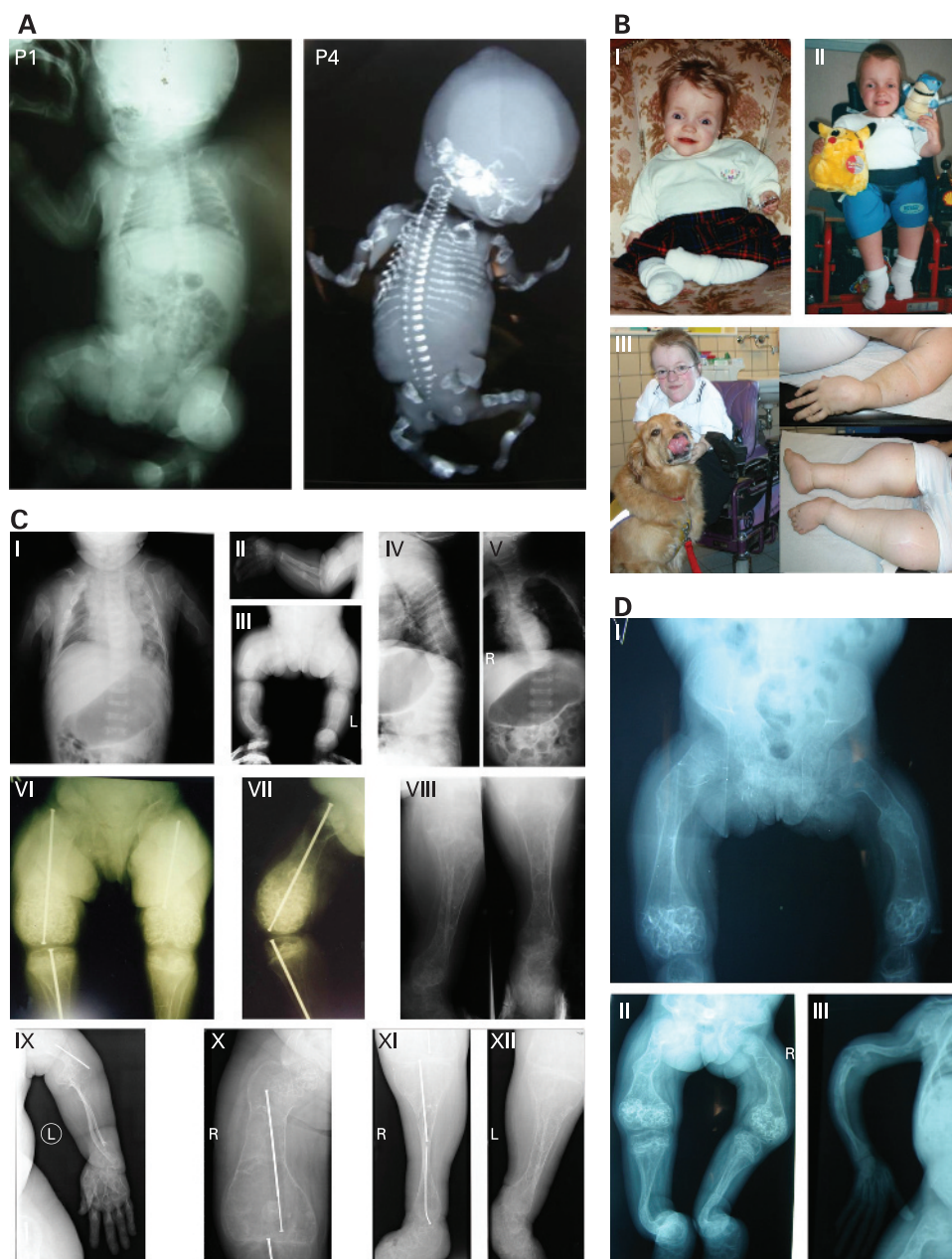
Figure 1 Phenotype of autosomal recessive osteogenesis imperfecta (AR OI) due to *LEPRE1* mutations. (A) Total body radiographs of neonates P1 and P4 supporting diagnosis of severe OI: lack of calvarial ossification; gracile beaded ribs with multiple fractures; platyspondyly; shortened, wide, bowed and fractured large tubular bones. (B) Clinical pictures of P3. (I) Age 13 months: unable to lift head and sit unaided. (II) Age 10 years: wheelchair bound patient propped into special chair and unable to sit or stand unaided. (III) Age 17 $^{4/12}$ years: immobile lower limbs; relatively long hands with hyperlax finger joints. (C) Radiographic studies in P3, showing the progression of the OCD throughout childhood and adolescence. (I) Age 3 months; (II-V) Age 13 months; (VI-VII) 5 $^{4/12}$ years; (VIII) 15 years; (IX-XII) 17 $^{4/12}$ years. (D) Skeletal radiographs of P4. (I) Age 5 years; (II-III) Age 8 years. (Informed consent was obtained from the patients for publication of this report and these photographs).

Alexa fluor 488 conjugated goat anti-rabbit IgG, A11008, Invitrogen; 1:100 Cy3 conjugated goat anti-mouse IgG, Ab6945-100, Abcam). Finally, the slides were washed three times in TBS, mounted in Vectashield-DAPI solution and photographed with a Zeiss-Axio-Plan 2 epifluorescence microscope.

Steady state collagen labelling

At confluency, fibroblasts were labelled with ^{14}C-proline as described previously.[14] After SDS-urea-PAGE separation, the gels were processed for fluorography, dried and visualised by autoradiography.[15]

(Tandem) mass spectrometric (MS) analysis

Protein bands derived from dermal fibroblasts were excised out of the Coomassie stained SDS-urea-PAGE gel, transferred to Biopure Eppendorf tubes and in-gel digested with trypsin as described.[16] The resulting peptide mixtures were dried in a centrifugal vacuum concentrator and re-dissolved in 20 µl of 0.1% formic acid in water. Half of this mixture was used for nanoLC-MS/MS analysis using a Waters nano-Acquity high performance liquid chromatography (HPLC) in-line coupled to a Q-TOF Premier mass spectrometer, operated as described previously.[17] The remainder of each sample was also spectrometrically analysed but only in MS mode (without data dependent MS/MS analysis).
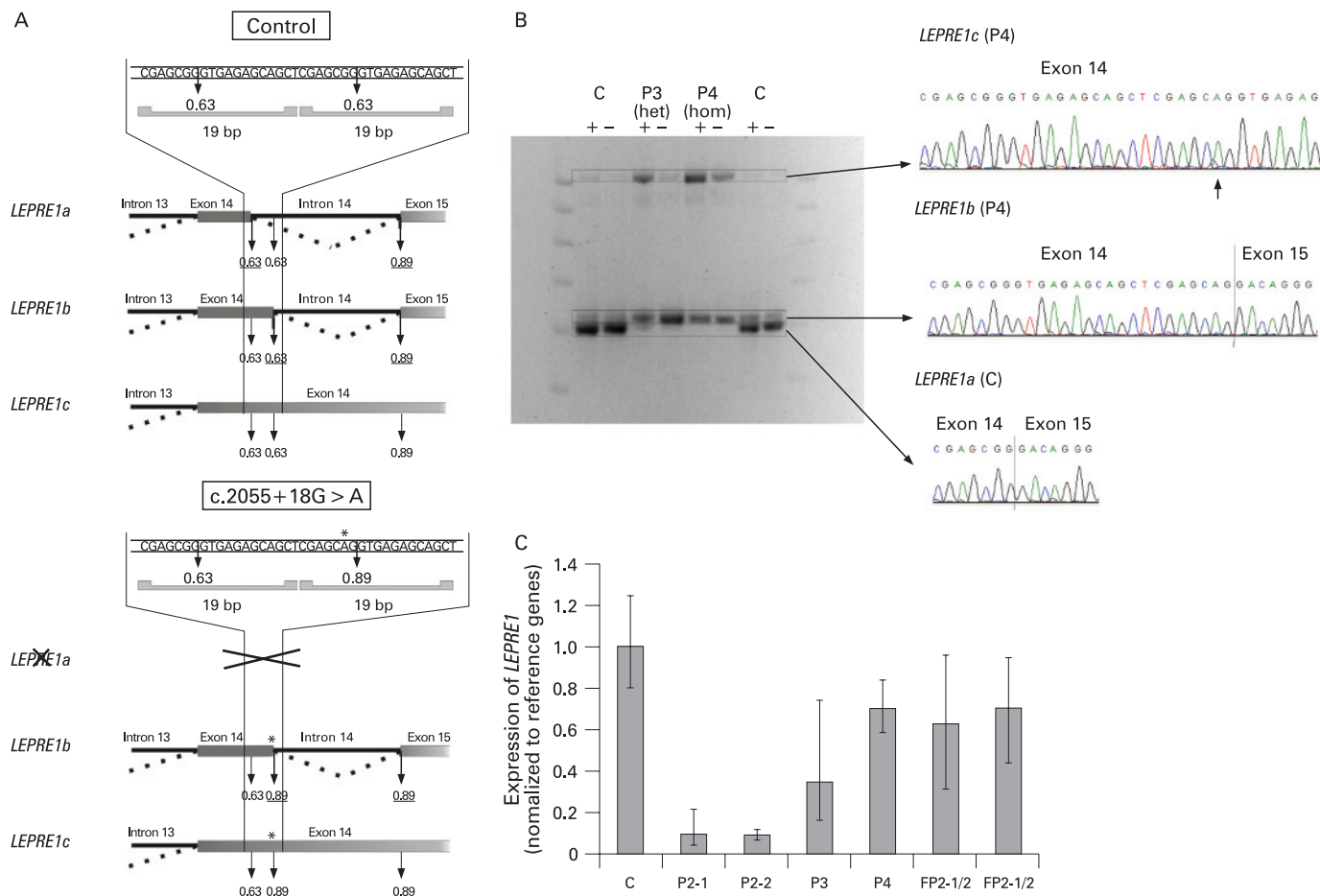
Figure 2 Alternative splicing and quantitative polymerase chain reaction (qPCR) of *LEPRE1* mRNA. (A) Schematic illustration of alternative splicing of *LEPRE1* intron 14 in a wild type (top) and the splice site mutation c.2055+18G>A allele (bottom). Enlargement of the splicing region shows the presence of a 19 bp repeat at the exon 14/intron 14 boundary. Potential splice sites and concomitant prediction scores are marked by arrows, which are bold in case the splice site is used in the corresponding splice form. Dotted lines point at splicing events. The splice site mutation is marked by an asterisk in the bottom part of (A). Splice form *LEPRE1a* is expected not to be formed by splice site mutant alleles, while the splice site mutation is silent in *LEPRE1c*. (B) Reverse transcriptase (RT) PCR of intron 14 and flanking exonic sequence with mRNA from c.2055+18G>A mutant fibroblasts (P3 and P4) versus mRNA from controls (C), followed by cloning and sequencing of the resulting amplicons. Fibroblasts were cultured in presence and absence of translational inhibitor cycloheximide (+/−). In control samples three different mRNA splice forms could be confirmed, whereas patient samples lack *LEPRE1a* in the absence of cycloheximide. The depicted sequences are derived from P4 (*LEPRE1c* and *LEPRE1b*) and from a control (*LEPRE1a*). An arrow denotes the position of the splice site mutation. (C) Relative mRNA expression of *LEPRE1* determined by qPCR in fibroblasts. In patients P2-1, P2-2 and P3 the level of *LEPRE1* mRNA is significantly decreased (p<0.05) compared with the control (mean of five control fibroblast cell lines), while P4 and parents of P2-1 and P2-2 show no significant lower *LEPRE1* mRNA expression compared to this control. Bars indicate the 95% confidence interval of the mean expression.

RESULTS

Identification of *LEPRE1* mutations

While no mutations were found in *CRTAP* and *PPIB*, novel homozygous and compound heterozygous *LEPRE1* mutations were identified in four probands (supplemental fig 1A). Proband 1 (P1) was homozygous for a frameshift mutation (c.1365_1366delAGinsC) in exon 9, resulting in a premature termination codon (PTC) in the same exon. Proband 2-1 (P2-1) and his sister P2-2 were homozygous for a nonsense mutation (c.628C>T) in exon 3. Proband 3 (P3) was found to be a compound heterozygote for a nonsense mutation (c.1102C>T) in exon 6 and a splice site mutation (c.2055+18G>A) in intron 14. The latter mutation was also identified in homozygous state in proband 4 (P4). This splice site mutation was not detected in 100 control alleles. The different ethnicity of P3 and P4 and the different haplotype backgrounds of the mutation in P3 and P4 argue against a founder mutation (supplemental fig 1B).

Clinical findings

Patients P1, P2-1, P2-2 and P4 were children of consanguineous parents of, respectively, Lebanese (P1) and Turkish (P2-1, P2-2 and P4) origin. P3 was the only child of a non-consanguineous Belgian couple. In all of them, the diagnosis of severe OI, reminiscent of OI type II/III, was established in utero or at birth (*x* rays at birth of P1 and P4 are shown in fig 1A). P1 died at 10 months of age due to respiratory failure, and the pregnancies of P2-1 and P2-2 were electively terminated at 20 and 18 weeks of gestation, respectively. P3 and P4 survived and are now 17 7/12 and 8 years of age, respectively.

Clinical and radiographic features of P3 at various ages are illustrated in figs 1B and 1C, respectively. P3 was born at 38 weeks gestation by caesarean section for breech presentation following an uneventful pregnancy. Birth weight was 3170 g (p50), length 43 cm (<p3) and occipitofrontal circumference (OFC) 33 cm (p10–p50). By 3 months of age severe radiological abnormalities were documented, including a narrow thorax,

short irregular ribs with large calluses, pronounced platyspon-dyly, and broad, short and fractured humeri (fig 1C.I). Nine fractures occurred in the first 9 months of life. At 13 months body length was hardly measurable because of painful lower limb deformities (fig 1B.I). The fontanel was still large. The girl had a tall forehead with frontal bossing and a narrow head, a rather prominent chin and slightly blue sclerae. The chest was short and narrow. At that time a skeletal survey (fig 1C.II-V) showed poor ossification of the calvarium and wormian bones. The humeri and the long bones in the lower limbs were much shortened with widening and bowing of the diaphyses (fig 1C.II-III). Extreme osteopenia and osteoporosis of the entire skeleton was particularly pronounced in the pelvic bones that had become barely visible. There was generalised, severe platyspon-dyly (fig 1C.IV-V) and already a thoracic dextroconvex scoliosis (fig 1C.V). Orthopaedic surgery at 2 $^{8/12}$ years for placing of supporting metal rods in some of the long bones did neither alleviate the patient's suffering nor favourably alter the natural course. Later similar surgery was also to no avail. Radiographs at the age of about 5 years showed, besides relentless progression of her OI, remarkable qualitative changes in and near the distal femoral and proximal tibial metaphyses (fig 1C.VI-VII). Besides progressive diaphyseal widening, the osseous trabeculae had coalesced into popcorn shaped concentric condensations leaving cyst-like translucent open spaces in between them. The distal femoral epiphyses were hardly visible. Trabeculation in the diaphyses was sparse and coarse (fig1 C.VI–VII). The patient has suffered dozens of long bone and rib fractures, but remarkably none since the age of 15 years. Radiographs at the age of 15 years showed funnel-type widening of both diaphyseal ends of the tibiae, where the irregular honeycomb-like config-uration of trabeculae had decreased considerably.

This patient, who is the longest survivor reported so far with OI due to *LEPRE1* mutations, was re-examined at age 17 $^{4/12}$ years. The morbidity of her OI had still increased with nearly continuous back pain, necessitating daily use of morphine analogues, severe thoracic scoliosis, life threatening pulmonary infections associated with restrictive lung disease, extreme lumbosacral kyphosis, and nearly immobile legs. The hands were relatively long with hyperlaxity of the finger joints (fig 1B.III). Hearing had remained normal. The skeletal survey confirmed the disorder's progress. Remarkably, the humeral and femoral ends had lost the popcorn-type dense areas and had become even more strikingly translucent. Whereas the rhizo-melic long bones had widened considerably, the mesomelic diaphyses had become narrower with only widening at both ends and a twisted ribbon-like aspect. The hands presented rather short metacarpals with diaphyseal constriction and widened ends, and phalanges that were similarly but more mildly abnormal. In particular the knee joint space had been reduced to a mere linear adjacent border between femur and tibia (fig 1C.VIII-XII).

At age 8 years P4 was able to sit without support but unable to stand. An older sister, who was also affected with OI, had died at 4 $^{1/2}$ years of age. P4 had a high forehead and narrow skull, bluish sclerae, downslanting palpebral fissures, a high arched palate, but no sign of dentinogenesis imperfecta. The chest was asymmetric and barrel shaped with severe thoracic kyphoscoliosis. She had an umbilical hernia. All limbs were very short and considerably bowed and mildly twisted. The hands were relatively long with pronounced hyperlaxity of the small joints. The bone density score was −11 at 6 $^{2/12}$ years but improved to −6.8 after "pamidronate treatment" (details of treatment not documented). Radiographs at the age of 5 years

(fig 1D.I) showed, like in P3, severely osteopenic bones. The femoral diaphyses were bowed with widening above the blown up fracture calluses. The proximal epiphyses, metaphyses and femoral necks were hardly discernable and the distal metaphy-seal ends were widened with trabecular aggregation into similar popcorn shaped abnormalities as in P3. At age 8 years (fig 1D.II–III) radiographs revealed extreme distal bowing of the tibiae with narrow diaphyses. Both proximal and distal tibial ends were widened with only partial trabecular condensation and the visible tarsal bones were severely osteopenic. The arm bone diaphyses were also severely bowed and narrow with widened ends, while epiphyses and metaphyses were hardly discernable with only a few areas of trabecular condensation. The humerus was significantly more shortened than the narrow, twisted ulna and radius.

Effects of *LEPRE1* mutations on transcription and splicing

The effects of the identified *LEPRE1* mutations on transcription, translation and type I collagen modification were evaluated in P2-1, P2-2, P3 and P4. For P1 no fibroblasts were available for further studies.

The c.2055+18G>A splice site mutation, identified in P3 and P4, is located in intron 14 of the *LEPRE1* gene. *LEPRE1* contains two identical 19 bp repeats located at the exon 14/intron 14 boundary, each containing a donor splice site with a score of 0.63, according to the NNSPLICE version 0.9 splice site prediction algorithm (http://www.fruitfly.org/seq_tools/splice.html) (fig 2A). Use of the donor splice site in the first and second 19 bp repeat generates two different transcripts: *LEPRE1a* (Genbank: AK075418) and *LEPRE1b* (Genbank: BC108311), encoding a 736AA (P3H1a) and 697AA (P3H1b) protein, respectively. However, the SwissProt protein database contains experimental evidence for the P3H1a protein only.[18] Moreover, according to the Genbank database (AK027648), a third transcript (*LEPRE1c*) with full retention of intron 14 may also be produced that, if translated, encodes a 804AA protein (P3H1c). The c.2055+18G>A splice site mutation increases the score of the donor splice site in the second repeat to 0.89. This suggests that the second repeat donor splice site will be preferably used, resulting in the formation of *LEPRE1b* and *LEPRE1c* transcripts only.

In order to explore the splicing pattern in patients carrying the c.2055+18G>A mutation, we have amplified intron 14 sequence and parts of the flanking exons 14 and 15 on reverse transcribed mRNA. By cloning and sequencing the different bands, three different mRNA splice forms were identified in controls, including an amplicon corresponding to the *LEPRE1a* transcript, a faint band representing the *LEPRE1b* transcript, and a very faint band corresponding to *LEPRE1c* (fig 2B). In patient P4, homozygous for the c.2055+18G>A mutation, only *LEPRE1b* and *LEPRE1c* transcripts were identified, prompting the conclusion that the donor splice site in the first repeat was completely ignored. A similar transcription pattern was observed in the compound heterozygous patient, P3. In this patient, the transcript from the nonsense allele is subjected to nonsense mediated decay (NMD). *LEPRE1a* transcripts from this allele can only be observed by adding cycloheximide, an agent that is known to substantially prevent NMD. qPCR in P2-1 and P2-2 fibroblast samples demonstrated a severe reduction of *LEPRE1* mRNA as expected from the degradation of mutant transcripts by the NMD pathway (fig 2C). P3 showed a significantly lower *LEPRE1* mRNA expression than controls. The residual *LEPRE1* transcripts are most likely derived solely from the allele carrying the splice site mutation
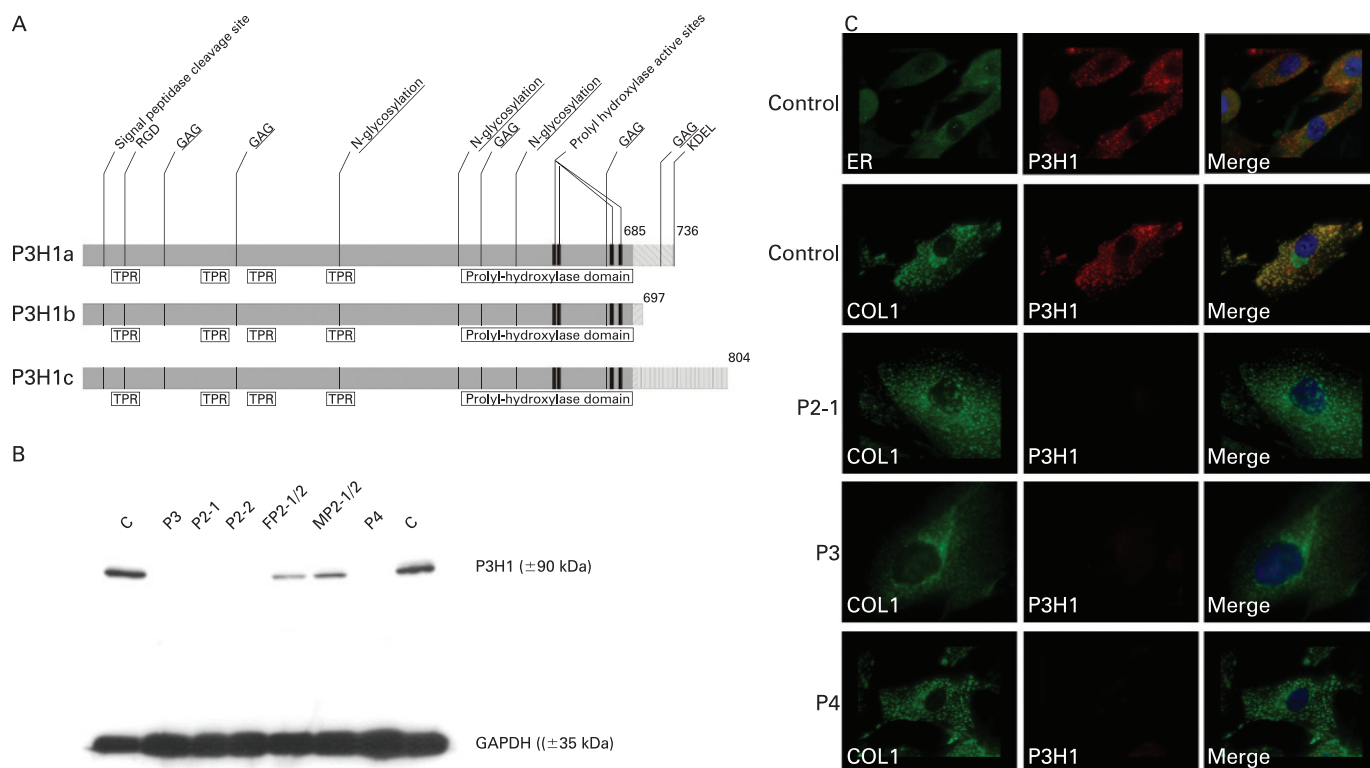
Figure 3 In silico protein structure analysis and protein expression of P3H1. (A) Schematic diagram of P3H1 splice forms. Common protein sequence is illustrated by grey bars. Differentially striped bars show variable sequence between the three splice forms. Four glycosaminoglycan acceptor sites (GAG), three N-glycosylation sites, an RGD motif and an N-terminal signal peptide with signal peptidase cleavage site, common to all P3H1 splice forms, are indicated. One additional C-terminal GAG is present in P3H1a. The prolyl hydroxylase domain containing prolyl 3-hydroxylase active sites, in addition to the C-terminal endoplasmic reticulum (ER) retention signal, KDEL, present in P3H1a only, predicts a prolyl 3-hydroxylation functioning of the P3H1a splice form in the ER. TPR motifs are known to be involved in protein–protein interaction. (B) Western blot of fibroblast cell lysates probed with anti-P3H1 antibody shows complete absence of P3H1 in P2-1, P2-2, P3 and P4 and slightly reduced in parents of sibs P2-1 and P2-2. The expression of the GAPDH protein is used as a loading control. (C) Immunofluorescence analysis of P3H1 in fibroblasts. Double staining of calnexin (alexa fluor 488: green) and P3H1 (Cy3: red) shows restriction of P3H1 to the ER. Double staining of type I collagen (alexa fluor 488: green) and P3H1 (Cy3: red) in control fibroblasts shows co-localisation of P3H1 and type I collagen (Ab292–100) in the ER. No P3H1 can be detected in any of the patient fibroblasts.

(LEPRE1b/c) and perhaps from some transcripts escaping NMD. Patient P4, who is homozygous for the splice site mutation, produces only LEPRE1b/c transcripts and shows hereby no significantly lower expression than in controls. Also in heterozygous carriers of nonsense LEPRE1 mutations (FP2–1/2 and MP2–1/2), mRNA expression was not significantly different from the control samples.

Effects of *LEPRE1* mutations on translation
In silico analysis of the expected protein products of the three observed *LEPRE1* mRNA splice forms predicted the presence of an N-terminal hydrophobic signal peptide that directs the proteins to the secretory pathway for putative functioning in either the plasma membrane or the extracellular matrix (fig 3A). The presence of N-glycosylation and GAG attachment sites suggests a function as extracellular matrix proteoglycan. In addition, the RGD cell attachment site implies a role in cell–matrix interaction. The C terminal "KDEL" ER retention signal is present only in P3H1a, restricting it to the ER and Golgi circuit. Western blot analysis of fibroblast lysates (fig 3B) has confirmed the absence of intracellular P3H1 in all patients and showed slightly reduced levels in the parents of P2-1/2, compared to the amount detected in two control cell lines. The protein detected in controls and parents, migrating at ±90 kDa, most likely corresponds to P3H1a, as the P3H1a encoding transcript is the only transcript absent in P3 and P4

(fig 2B). Although P3H1b/c are expected to be present in P3 and P4 and in control cell lines, none of these proteins were detected in fibroblast cell extracts, not even after treatment with chondroitinase ABC enabling the identification of proteoglycan core proteins (results not shown). Immunofluorescence (IF) analysis shows that in control skin fibroblasts, P3H1 and type I collagen co-localise exclusively in the ER, whereas none of the P3H1 protein splice forms was detectable in any patient cell line, consistent with the western blot results (fig 3C).

Secretion of P3H1b/c into the extracellular environment, expected by secretory motifs in the deduced protein structure, could not be demonstrated by western blotting of concentrated fibroblast medium either in the patients, their parents or controls (with or without prior treatment with chondroitinase ABC, results not shown), although we used a polyclonal antibody raised against a full length human P3H1 protein with which we should be able to detect all three protein splice forms.

Effects of P3H1 deficiency on type I collagen modification
Mass spectrometry analysis of tryptic peptides of secreted α1(I)-collagen chains revealed the severely reduced prolyl 3-hydroxylation at Pro986 in the α1(I) chain by the mutant P3H1 protein in patients P2-1, P3 and P4 when compared to that in a control and parental sample (fig 4A).

As shown in fig 4B, electrophoresis of the type I collagen chains showed only over-modified chains in P2-1, P3 and P4,
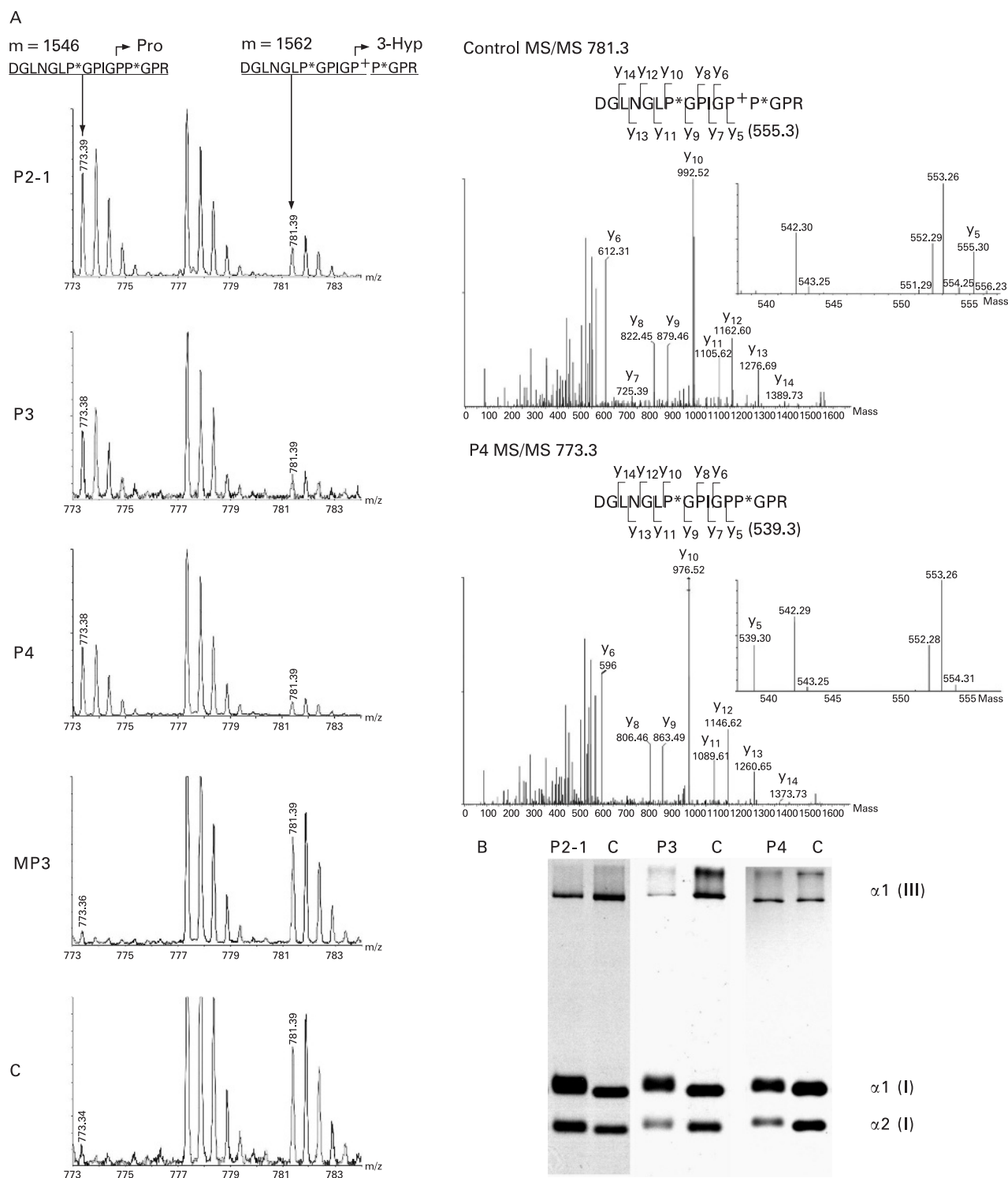
Figure 4 Tandem mass spectrometry and SDS-urea-PAGE of secreted type I collagen. (A) Left side: mass spectra of α1(I) collagen tryptic peptide ions, containing hydroxylated Pro986 (doubly charged precursor at nominal m/z of 781 Th) or non-hydroxylated Pro986 (doubly charged precursor at nominal m/z of 773 Th), shown for patients P2-1, P3 and P4, a heterozygous nonsense mutation carrier (MP3) and a control. Since both peptide forms elute differently from RP columns, the mass spectra shown were merged over an elution window of about 1 min. Note that an additional peak at m/z of about 777 Th is also present but the corresponding peptide was not identified by MS/MS analysis. The 3-hydroxylation at α1(I)Pro986 is low compared to the normal level in the control and MP3. Right side: MS/MS spectra of the 781 (upper panel) and the 773 Th peptide ions, containing the α1(I)Pro986 residue (lower panel), shown for a control and P4, respectively. Following manual interpretation, the y-type of fragment ions are localising the 16 dalton mass difference between these peptides to the C terminal five residues. Further manual interpretation leads to identification of the 781 Th ion as collagen α1(I) 975-DGLNGLP*GPIGP+P*GPR-990 (P+ and P* denoting 3- and 4-hydroxylated proline, respectively), whereas the 773 Th ion corresponds to collagen α1(I) 975-DGLNGLP*GPIGPP*GPR-990, thus with Pro986 in the non-hydroxylated form. For the sake of clarity, only the sequence tag that was readable using y-type of fragment ions is shown. (B) SDS-urea-PAGE analysis of type I collagen in the fibroblast culture medium. Delay of electrophoretic migration of type I collagen α-chains in P2-1, P3 and P4 compared to that secreted from control cell strains indicates overmodification of all these α-chains.

contrary to what is observed in severe AD OI, where usually both normal and aberrantly migrating type I collagen chains are seen.

DISCUSSION

In four unrelated probands with severe OI without *COL1A1* or *COL1A2* mutation, four novel mutations in the *LEPRE1* gene were detected. The frameshift and nonsense mutations in P1 and P2–1/2 respectively and the splice site mutation in P4 are homozygous and associated with parental consanguinity. P3, the longest survivor among all OI patients with *LEPRE1* mutations reported so far, is compound heterozygous for the same splice site mutation as in P4 and a novel nonsense mutation.[7 10]

Our observations confirm the conclusion by others[10] that, in the perinatal period, it is very difficult to distinguish radiographically severe OI due to *LEPRE1* mutations from severe OI caused by *COL1A1* or *COL1A2* mutations. A similar conclusion has recently been drawn regarding the *CRTAP* mutations.[10] However, with increasing age some distinctive clinical and radiologic features become apparent. Long term survivors with OI due to *LEPRE1* mutations present a particularly severe phenotype, which significantly impairs quality of life, as illustrated in P3, who has required daily use of morphine analogues for several years. Clinically, extreme short stature and shortening of the four extremities, with rhizomelia of both upper and lower limbs, is striking in both P3 and P4. In both patients the forehead with frontal bossing is remarkably tall and narrow. The OFC in P3 is only relatively macrocephalic. Osteoporosis is particularly severe, with DEXA Z-scores that are generally lower than those seen in AD OI.[7] In the radiographic follow-up in P3, three consecutive stages can be discerned. In infancy and early childhood, lack of long bone and vertebral remodelling and growth results mainly in rhizomelic long bones, diaphyseal widening and bowing as well as widening and deformation of all epi- and metaphyses. The second stage is apparent from mid-childhood. Continuing long bone deformation is associated with peculiar agglomeration of the bony trabeculae, especially prominent in the wide metaphyseal ends in the knees. It results in round popcorn shaped dense and cyst-like translucent areas organised in a honeycomb-like pattern. The first and second radiographic stages were observed also in the 8-year-old P4. Beyond the age of 15 years, the third stage is noticeable in P3 and is characterised by a complete disappearance of the honeycomb-type trabecular agglomeration, extreme osteoporosis and additional widening of the rhizomelic diaphyses, but progressive narrowing and bowing of the mesomelic diaphyses. The knee joint spaces have been reduced to a mere line between the femora and tibiae. We have not observed similar features in several long term surviving OI type III patients with documented *COL1A1* or *COL1A2* mutation.

On the one hand, the lack of prolyl-3 hydroxylation most probably has its severe phenotypic effects because of abnormal posttranslational modification of type I collagen molecules, which were also shown to have a moderately slower but increased secretion than in control cells.[7] On the other hand P3H1 deficiency also affects the structural integrity and maintenance of type II collagen in human cartilage. This combined adverse effect had already been documented in Crtap -/- mice, which show decreased prolyl 3-hydroxylation in cartilage, as well as abnormalities in the growth columns of proliferating chondrocytes of the growth plate.[8] Together, these findings suggest that recessive OI due to defects in the

prolyl 3-hydoxylation complex is in fact an osteochondrodysplasia (OCD), characterised by both defects in cartilage and bone. The contribution to the disease phenotype of deficient prolyl 3-hydroxylation of type IV and V collagen, which contain multiple prolyl 3-hydroxylation sites, remains unclear, although overmodification of type V collagen, which is present in association with type I collagen in skin and bone, has been shown in patients with P3H1 deficiency.[7]

This report brings the number of already characterised mutations in *LEPRE1* to 17, 16 of which would be expected to affect all three *LEPRE1* mRNA splice forms detected in this study, and most of which would be expected to result in severely decreased levels of total *LEPRE1* mRNA.[7 10] It is of interest that the splice site mutation identified in P3 and P4 affects only one of the three *LEPRE1* mRNA splice forms (*LEPRE1a*). Western blotting and IF analysis of fibroblast cultures both revealed the absence of the corresponding protein splice form P3H1a (736AA) in P2-1, P3 and P4. Moreover, mass spectrometry showed severe reduction of α1(I)Pro986 3-hydroxylation in these probands, resulting in overmodification of type I (pro)collagen chains in skin fibroblasts on SDS-urea-PAGE. This overmodification results from a longer exposure of the collagen chains to lysyl and prolyl 4-hydroxylation due to delayed collagen type I triple helix formation in P3H1 deficient cells.[7] The fact that the reduction of prolyl 3-hydroxylation in P3 and P4, both carrying the splice site mutation, was comparable to that in P2-1, homozygous for a nonsense mutation, supports the notion that α1(I)-Pro986 hydroxylation is restricted to P3H1a. Some residual 3-hydroxylation activity was detectable in the patients, despite the absence of detectable P3H1 protein on western blot and IF analysis. This could indicate compensation for P3H1 loss-of-function—for example, by other members of the P3H family members, such as P3H2 or P3H3.[19] Tiainen *et al*[20] showed that recombinant P3H2 hydroxylates a synthetic peptide corresponding to the single prolyl 3-hydroxylation site in the α1(I) collagen chain. P3H1a is the only protein splice form that contains the "KDEL" ER retention signal. The retention of P3H1a in the ER, where it co-localises with type I collagen and exerts its prolyl 3-hydroxylation function, was demonstrated by means of IF analysis. Hence, only P3H1a would catalyse the prolyl 3-hydroxylation, known to occur in the ER, and therefore P3H1a probably represents the true prolyl 3-hydroxylase-1.

The non-redundant function of P3H1a and P3H1b/c suggests distinct roles of the latter splice forms. The in silico protein structure analysis predicts that the splice forms P3H1b/c are directed towards the secretory pathway with a putative function in the extracellular matrix. In addition, the presence of N-glycosylation sites, GAG attachment sites and an RGD cell attachment site supports the idea of their role as extracellular matrix cell attached proteoglycan. Wassenhove-McCarthy *et al*[21] isolated P3H1 as the chondroitin sulfate proteoglycan, leprecan, from a rat L-2 cell line and localised the proteoglycan to basement membranes. Irrespective of a role as matrix proteoglycan for P3H1b/c, we have been unable to detect any corresponding protein in the medium and cellular fractions of fibroblast cell cultures, using a polyclonal antibody raised against the full length human P3H1 protein. This paradox may be explained by the possibly low abundance of these proteins in fibroblast cultures, in which the medium fraction differs substantially from the natural extracellular environment of skin. Preferably, bone or skin extracellular matrix extracts should be used to study proteoglycans secreted into the extracellular matrix. On the other hand, the alternative splicing

may have an on-off switch role, regulating the amount of biologically active protein (P3H1a) with only poor or no translation of splice forms P3H1b/c, a mechanism which has been shown earlier for other genes.[22] It is of interest that so far no *LEPRE1b/c* transcripts have been detected in closely related species such as *Mus musculus* or *Rattus norvegicus* and conservation of intron 14 is also low between the latter species and *Homo sapiens*.

Electrophoresis of type I (pro)collagen chains in fibroblasts can be helpful in the detection of *LEPRE1* mutations, since in this instance only over-modified type I collagen chains are detected, in contrast to what is observed in severe AD OI, where both normal and aberrantly migrating type I collagen chains are usually seen. The latter distinction, however, is less clear for collagen studies performed in fetal cells, in which collagen chains show a variable delay in electrophoretic mobility. In our experience, *COL1A1/COL1A2* mutations are identified in ~93% of OI patients with a documented abnormal electrophoretic pattern for type I collagen. So far, the proportion of *LEPRE1* mutations in this cohort is <3%. Hence, mutations in *LEPRE1/CRTAP* account for only a small proportion of patients with severe/lethal OI. Other, yet to be identified, genes may be involved in the pathogenesis of OI. Candidates may have a role either in post-translational modification of type I collagen and/or intracellular trafficking of collagen components.

In conclusion, we present four novel *LEPRE1* mutations in four unrelated probands with a severe congenital OCD with important growth deficiency, many fractures and extreme skeletal undermineralisation that, at birth, cannot be distinguished from OI in patients with *COL1A1/COL1A2* mutations. Clinical follow-up of the longest survivor of AR OI due to *LEPRE1* mutations provides evidence for the evolution towards a particularly severe OCD with distinct clinical and radiographic features. Obviously the prospective study of the natural course in long term surviving patients with severe OI with known causal mutations will establish more firmly whether or not specific OI phenotypes may be assigned to some or all of the different genotypes. The finding of the newly identified splice site mutation that specifically affects the P3H1 splice form responsible for $\alpha1(I)Pro986$ 3-hydroxylation, kindles interest in the role and function of the other splice forms of P3H1 (*LEPRE1* transcripts) that cannot rescue this hydroxylation activity.

REFERENCES

1. **Rauch F,** Glorieux FH. Osteogenesis imperfecta. *Lancet* 2004;**363**:1377–85.
2. **Bruckner P,** Eikenberry EF. Formation of the triple helix of type I procollagen in cellulo. Temperature-dependent kinetics support a model based on cis in equilibrium trans isomerization of peptide bonds. *Eur J Biochem* 1984;**140**:391–5.
3. **Myllyharju J,** Kivirikko KI. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet* 2004;**20**:33–43.
4. **Vranka JA,** Sakai LY, Bachinger HP. Prolyl 3-hydroxylase 1, enzyme characterization and identification of a novel family of enzymes. *J Biol Chem* 2004;**279**:23615–21.
5. **Sillence DO,** Senn A, Danks DM. Genetic heterogeneity in osteogenesis imperfecta. *J Med Genet* 1979;**16**:101–16.
6. **Ward LM,** Rauch F, Travers R, Chabot G, Azouz EM, Lalic L, Roughley PJ, Glorieux FH. Osteogenesis imperfecta type VII: an autosomal recessive form of brittle bone disease. *Bone* 2002;**31**:12–8.
7. **Cabral WA,** Chang W, Barnes AM, Weis M, Scott MA, Leikin S, Makareeva E, Kuznetsova NV, Rosenbaum KN, Tifft CJ, Bulas DI, Kozma C, Smith PA, Eyre DR, Marini JC. Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. *Nat Genet* 2007;**39**:359–65.
8. **Morello R,** Bertin TK, Chen Y, Hicks J, Tonachini L, Monticone M, Castagnola P, Rauch F, Glorieux FH, Vranka J, Bachinger HP, Pace JM, Schwarze U, Byers PH, Weis M, Fernandes RJ, Eyre DR, Yao Z, Boyce BF, Lee B. CRTAP is required for prolyl 3-hydroxylation and mutations cause recessive osteogenesis imperfecta. *Cell* 2006;**127**:291–304.
9. **Barnes AM,** Chang W, Morello R, Cabral WA, Weis M, Eyre DR, Leikin S, Makareeva E, Kuznetsova N, Uveges TE, Ashok A, Flor AW, Mulvihill JJ, Wilson PL, Sundaram UT, Lee B, Marini JC. Deficiency of cartilage-associated protein in recessive lethal osteogenesis imperfecta. *N Engl J Med* 2006;**355**:2757–64.
10. **Baldridge D,** Schwarze U, Morello R, Lennington J, Bertin TK, Pace JM, Pepin MG, Weis M, Eyre DR, Walsh J, Lambert D, Green A, Robinson H, Michelson M, Houge G, Lindman C, Martin J, Ward J, Lemyre E, Mitchell JJ, Krakow D, Rimoin DL, Cohn DH, Byers PH, Lee B. CRTAP and LEPRE1 mutations in recessive osteogenesis imperfecta. *Hum Mutat* 2008;**29**:1435–42.
11. **Pattyn F,** Robbrecht P, De Paepe A, Speleman F, Vandesompele J. RTPrimerDB: the real-time PCR primer and probe database, major update 2006. *Nucleic Acids Res* 2006;**34**(Database issue):D684–8.
12. **Hellemans J,** Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* 2007;**8**(2):R19.
13. **Gogiel T,** Jaworski S. Proteoglycans of human umbilical cord arteries. *Acta Biochim Pol* 2000;**47**:1081–91.
14. **Nuytinck L,** Dalgleish R, Spotila L, Renard JP, Van Regemorter N, De Paepe A. Substitution of glycine-661 by serine in the alpha1(I) and alpha2(I) chains of type I collagen results in different clinical and biochemical phenotypes. *Hum Genet* 1996;**97**:324–9.
15. **Laskey RA,** Mills AD. Quantitative film detection of 3H and 14C in polyacrylamide gels by fluorography. *Eur J Biochem* 1975;**56**:335–41.
16. **Pattyn E,** Verhee A, Uyttendaele I, Piessevaux J, Timmerman E, Gevaert K, Vandekerckhove J, Peelman F, Tavernier J. HyperISGylation of Old World monkey ISG15 in human cells. *PLoS ONE* 2008;**3**:e2427.
17. **Ghesquiere B,** Van Damme J, Martens L, Vandekerckhove J, Gevaert K. Proteome-wide characterization of N-glycosylation events by diagonal chromatography. *J Proteome Res* 2006;**5**:2438–47.
18. **Kaul SC,** Sugihara T, Yoshida A, Nomura H, Wadhwa R. Gros1, a potential growth suppressor on chromosome 1: its identity to basement membrane-associated proteoglycan, leprecan. *Oncogene* 2000;**19**:3576–83.
19. **Jarnum S,** Kjellman C, Darabi A, Nilsson I, Edvardsen K, Aman P. LEPREL1, a novel ER and Golgi resident member of the Leprecan family. *Biochem Biophys Res Commun* 2004;**317**:342–51.
20. **Tiainen P,** Pasanen A, Sormunen R, Myllyharju J. Characterization of recombinant human prolyl 3-hydroxylase isoenzyme 2, an enzyme modifying the basement membrane collagen IV. *J Biol Chem* 2008;**283**:19432–39.
21. **Wassenhove-McCarthy DJ,** McCarthy KJ. Molecular characterization of a novel basement membrane-associated proteoglycan, leprecan. *J Biol Chem* 1999;**274**:25004–17.
22. **Yang TP,** Agellon LB, Walsh A, Breslow JL, Tall AR. Alternative splicing of the human cholesteryl ester transfer protein gene in transgenic mice. Exon exclusion modulates gene expression in response to dietary or developmental change. *J Biol Chem* 1996;**271**:12603–9.