A homozygous pathogenic missense variant broadens the phenotypic and mutational spectrum of CREB3L1-related osteogenesis imperfecta

Guillemyn B, Kayserili H, Demuynck L, Sips P, De Paepe A, Syx D, Coucke PJ, Malfait F, Symoens S

Human Molecular Genetics, 28 (11), 1801-1809, 2019.

A homozygous pathogenic missense variant broadens the phenotypic and mutational spectrum of *CREB3L1*-related osteogenesis imperfecta

Brecht Guillemyn[1], Hülya Kayserili[2], Lynn Demuynck[1], Patrick Sips[1], Anne De Paepe[1], Delfien Syx[1], Paul J. Coucke[1], Fransiska Malfait[1], Sofie Symoens[1,*]

[1]Center for Medical Genetics Ghent, Ghent University Hospital, Department of Biomolecular Medicine, Ghent 9000, Belgium
[2]KOÇ University School of Medicine (KUSoM) Medical Genetics Department, Topkapi Zeytinburnu, 34010 Istanbul, Turkey

*To whom correspondence should be addressed: Department of Biomolecular Medicine, Center for Medical Genetics Ghent, Ghent University Hospital, Corneel Heymanslaan 10, Medical Research Building 1, 9000 Ghent, Belgium. Tel: 0032/9 332 02 33; Email: Sofie.Symoens@UGent.be

Abstract

The cyclic AMP responsive element binding protein 3-like 1 (*CREB3L1*) gene codes for the endoplasmic reticulum stress transducer old astrocyte specifically induced substance (OASIS), which has an important role in osteoblast differentiation during bone development. Deficiency of OASIS is linked to a severe form of autosomal recessive osteogenesis imperfecta (OI), but only few patients have been reported. We identified the first homozygous pathogenic missense variant (p.(Ala304Val)) in a patient with lethal OI, which is located within the highly conserved basic leucine zipper domain, four amino acids upstream of the DNA binding domain. *In vitro* structural modeling and luciferase assays demonstrate that this missense variant affects a critical residue in this functional domain, thereby decreasing the type I collagen transcriptional binding ability. In addition, overexpression of the mutant OASIS protein leads to decreased transcription of the *SEC23A* and *SEC24D* genes, which code for components of the coat protein complex type II (COPII), and aberrant OASIS signaling also results in decreased protein levels of SEC24D. Our findings therefore provide additional proof of the potential involvement of the COPII secretory complex in the context of bone-associated disease.

Introduction

Osteogenesis imperfecta (OI) is a clinically and genetically heterogeneous group of heritable bone dysplasias, with the severity of symptoms ranging from perinatal lethality to generalized osteopenia (1). This brittle bone disease affects one in 15,000-20,000 births and is characterized by typical clinical manifestations such as bone fragility, skeletal deformities, low bone mass and short stature. Extraskeletal features, including blue sclerae, dentinogenesis imperfecta, adult-onset hearing loss, joint hypermobility, restrictive pulmonary disease, cardiovascular abnormalities and easy bruising, contribute to the multisystemic disorder (1-3). The predominant autosomal dominant (AD) forms are caused by mutations in either *COL1A1* (MIM 120150) or *COL1A2* (MIM 120160), encoding the α1- and α2-chains of type I procollagen respectively. Another rare AD OI subtype is associated with mutations in interferon–induced transmembrane protein 5 (*IFITM5*, MIM 614757), which is involved in bone mineralization. In approximately 10% of OI cases, the disease has an autosomal recessive (AR) inheritance. Several genes have been associated with these AR forms of OI, and they are classified according to their mechanism and pathophysiology: collagen post-translational modification (*CRTAP*, MIM 605497; *P3H1*, MIM 610339; *PPIB*, MIM 123841), collagen processing and crosslinking (*SERPINH1*, MIM 600943; *FKBP10*, MIM 607063; *PLOD2*, MIM 601865; *BMP1*, MIM 112264), bone mineralization (*SERPINF1*, MIM 172860) and osteoblast differentiation/function (*SP7*, MIM 606633; *TMEM38B*, MIM 611236; *WNT1*, MIM 164820; *CREB3L1*, MIM 616215; *SPARC*, MIM 182120; *MBTPS2*, MIM 300294; *TAPT1*, MIM 616897) (1, 2, 4-18).

The *CREB3L1* gene (cAMP Responsive Element Binding Protein 3 Like 1) encodes the endoplasmic reticulum (ER)-stress transducer 'old astrocyte specifically induced substance' (OASIS), a basic leucine zipper (bZIP) transcription factor which belongs to the well-conserved family of the cyclic AMP responsive element binding protein/activating transcription factor (CREB/ATF) genes. OASIS is processed by regulated intramembrane proteolysis (RIP) in response to ER stress, and is highly expressed in osteoblasts (19, 20). OASIS$^{-/-}$ mice exhibit severe osteopenia and spontaneous fractures, resulting from a decrease in type I collagen in the bone matrix and a decline in the activity of osteoblasts. More recently, *Col1a1* was identified as a target of OASIS, and Murakami *et al.* demonstrated with murine studies that OASIS activates the transcription of *Col1a1* through an unfolded protein response element (UPRE)-

like sequence in the *Col1a1* promoter region, thereby revealing its critical role in bone formation (19-21).

Hitherto, only 3 reports have associated homozygous *CREB3L1* defects to an AR form of OI (a whole gene deletion, the in-frame deletion (c.934_936delAAG, p.(Lys312del)) and the nonsense variant (c.1284C>A, p.(Tyr428*))), which is currently classified as OI type XVI (2, 15, 22, 23).

Here, we present a Turkish family, in which molecular analysis of the proband revealed a previously unreported homozygous missense variant (c.911C>T, p.(Ala304Val)).

We applied structural modeling to study the effects of this missense variant on the OASIS protein. We then performed further *in vitro* studies to investigate the functional consequences regarding regulation of type I collagen and COPII component gene expression.

Results

Clinical phenotype

We report a consanguineous Turkish family of second cousins, who had a medically terminated pregnancy at 19 weeks of gestation due to skeletal changes highly suggestive for severe OI. Antenatal ultrasound findings of the female fetus (IV-3, Fig. 1) included short tubular bones, multiple rib fractures with beaded appearance, and a narrow thorax circumference of 81mm (2.5-5th percentile). Postnatal findings revealed a birth length, weight and occipitofrontal circumference of 16cm, 210g and 15.3cm, respectively, and the presence of soft calvaria, microretrognathia and short and bowed extremities.

The parents (III-7 and III-8, Fig. 1) did not show any overt clinical signs of OI and had no history of fractures. Bone densitometry revealed Z-scores of the left femur of -1.2 for the mother and -2.1 for the father, while Z-scores for the lumbar L1 to L4 vertebrae were -2.5 for the mother and -3.7 for the father, markedly lower values than expected for their age. Two prior pregnancies of the couple were terminated with early first trimester abortions of unknown cause (IV-1 and IV-2, Fig. 1). The parents have two healthy sons (IV-4, 3.5 years old; IV-5, 2 years old, Fig. 1) and interestingly, three individuals in the family had a history of fractures (III-9, paternal uncle, 5 fractures of the ankle and elbow after mild trauma; III-10, paternal uncle, one fracture; paternal grandmother (not included in the pedigree), 2 fractures of the wrist). Clinical assessment of other family members was not possible.

Molecular studies and structural modeling reveal a critical residue in the nuclear localization sequence of *CREB3L1*

Panel sequencing of all known OI-associated genes identified a homozygous missense variant c.911C>T p.(Ala304Val) in exon 7 of the *CREB3L1* gene, which was confirmed by direct Sanger sequencing. Both parents (III-7 and III-8, Fig. 1A) and two healthy siblings (IV-4 and IV-5, Fig. 1A) were found to be heterozygous carriers of the missense variant, and sequencing for the other family members (with or without fractures) was not possible as no blood samples were available. Based on the criteria of absence of the variant from the queried population databases, *in silico* prediction tools pointing to a possible pathogenic allele (PolyPhen-2, Probably Damaging (0.992); SIFT, Deleterious (0.03); Align CVGD, C0 (GV=58.02); DEOGEN2, overall deleterious score of 51.5%), and a clinical phenotype highly suggestive for

the *CREB3L1* related form of OI, we initially classified this variant as a variant of unknown significance (VUS, class 3) (24).

The affected alanine (Ala) residue is located within a highly conserved bipartite nuclear localization sequence (NLS) (RVRRKIKNKIS**A**QESRRKKKEY) within the bZIP domain, and only 4 amino acids (AA) upstream of the DNA binding domain (RRKKKEY), in which the earlier reported in-frame deletion p.(Lys312del) is located (RRKK**K**EY), (Fig. 2, Fig. 3A) (22). Based on these observations, we hypothesized that the pathogenic mechanism underlying p.(Ala304Val) is similar to what has been suggested for the p.(Lys312del) variant, *i.e.* that mutant OASIS cannot reach the nucleus and/or bind to its target DNA sequence (22). Therefore, we decided to include the p.(Lys312del) as a positive control in our further experiments.

Three dimensional structural modeling of the full-length WT, p.(Ala304Val), and p.(Lys312del) protein sequences of OASIS, by means of Iterative Threading ASSembly Refinement (I-TASSER) algorithms, showed that both mutations result in conformational changes within the NLS (Fig. 3). Substitution of an Ala to Val residue has been associated with a decreased ability to adopt an α-helical conformation (25), which is seen in all models for p.(Ala304Val), while being less pronounced for p.(Lys312del) (Fig. 3) (26-28).

Based on systematic AA replacement analysis in budding yeast, Kosugi *et al.* created a platform which enables researchers to study the functional contribution of AAs at each position of a NLS class (29). Use of these NLS mapper algorithms results in a predicted bipartite NLS loss for both p.(Ala304Val) and the positive control p.(Lys312del) (Supplemental Fig. S2) (22, 23, 29), suggesting that both mutant proteins might not be able to translocate to the nucleus.

To further investigate this hypothesis, we also modeled the p.(Ala304Val) variant on a homology model of the CREB bZIP-CRE complex (30). The results demonstrated that the AA at position 304 is facing inward, pointing towards the CRE binding site of the DNA helix. Importantly, a size difference is noted at this position when comparing the WT (Ala, 67 Angstrom cube) to the mutant protein (Val, 105 Angstrom cube) (Supplemental Figure S3) (31). No change in polarity was noted, and the variant did not disrupt H-bridges of adjacent AAs (31, 32).

Mutant OASIS affects expression of type I collagen and COPII vesicle proteins

Since *CREB3L1* is expressed at very low levels in fibroblasts, we reasoned that this is a suboptimal cell type to study the function of OASIS, and therefore chose the same overexpression system as used by Keller *et al.* to study the pathogenic nature of the identified variant (20-22). The expression constructs generated for this study are referred to as 'Empty'

(empty vector), 'WT' (WT OASIS), 'A304V' (variant reported in this study) and 'K312del' (variant previously reported, (22)). Transfection of these expression constructs in HEK293 cells resulted in the expression of stable OASIS proteins (Fig. 4G). Biochemical assays using a luciferase reporter were performed in order to validate the direct impact of the p.(Ala304Val) variant on the regulation of the expression of the downstream target genes of OASIS, using type I collagen expression as a representative example. Significantly decreased luciferase activity was observed for the A304V and K312del constructs compared to WT, indicating that the respective variants lead to a reduced transcriptional activation of the *Col1a1* promoter.

This was observed not only for the full-length *Col1a1* promoter, but also for the *Col1a1* promoter with a mutant UPRE sequence, which was previously used to demonstrate that OASIS directly binds to this specific *Col1a1* promoter sequence (Fig. 4A and Fig. 4B) (19-21). In addition, luciferase activity levels of cells transfected with both A304V and K312del constructs were comparable to transfection of the Empty vector, indicating that both mutant proteins have no detectable residual DNA binding and/or gene expression activation ability to regulate the *Col1a1* gene. Since *SEC23A* and *SEC24D*, both members of the COPII secretory pathway, were previously shown to be targets of CREB3L2 and CREB3L1, respectively, we performed overexpression studies in HEK293 cells in order to investigate the effects of the OASIS variants on the expression of these genes (22, 33). Measurements of *SEC23A* and *SEC24D* mRNA levels showed that overexpression of the A304V and K312del variants significantly decreased transcription of both COPII inner coat components (Fig. 4C-D). At the protein level however, only the level of SEC24D was significantly decreased after overexpression of either A304V or K312del (Fig. 4E-F and 4H-I), which is in line with the earlier report on the effects of the p.(Lys312del) in-frame deletion (22).

Taken together, these functional data enabled us to reclassify the p.(Ala304Val) VUS as a pathogenic variant (causal mutation, class 5) and to offer appropriate genetic counselling to the affected family.

Discussion

This is the first report linking a pathogenic missense variant to the *CREB3L1*-related AR form of OI, and the 4th case in total implicating this gene. In 2013, we described the association of *CREB3L1* to a lethal AR form of OI in a family with two affected relatives carrying a homozygous whole gene deletion (15). More recently, a pathogenic in-frame deletion (c.934_936delAAG, p.(Lys312del)) was reported, in which a qualitative alteration of the protein affected both the DNA binding capacity of OASIS and the COPII coat secretory pathway (22). Phenotypically, this p.(Lys312del) pathogenic variant led to intrauterine fetal demise in the homozygous proband, whereas heterozygous carriers presented with mild signs of OI (history of fractures, osteopenia and blue sclerae) (22). Previous medically terminated pregnancies in this family displayed similar severe and lethal signs of OI, but no molecular studies were conducted for these cases (22). Recently, a third homozygous nonsense pathogenic variant (c.1284C>A, p.(Tyr428*)) was reported by Lindahl *et al*., for which studies on osteoblasts and fibroblasts revealed a decrease in *COL1A1* transcription only in osteoblasts, thereby strengthening the role of OASIS as a tissue-specific transcription factor (23). Furthermore, the authors demonstrated that deficiency of OASIS affects transcription of several bone-associated genes (*COL1A1*, *COL1A2*, *ALPL*, *IBSP* and *OPN*), reduces glycosaminoglycan levels in bone extracellular matrix and has negative effects on osteoblasts (23). In contrast to the two earlier reports, the child presented by Lindahl *et al.*, survived infancy. The boy's healthy parents and four siblings did not display signs of (mild) OI (23).

Molecular studies of the identified homozygous c.911C>T, p.(Ala304Val) pathogenic variant reported in this study revealed that the mutated AA is a critical residue in the NLS, positioned in the same functional domain as the earlier described in-frame deletion c.934_936delAAG, p.(Lys312del) (22). The lethal phenotype observed in the proband with this homozygous pathogenic missense variant and the milder OI signs in heterozygous carriers are similar to these described by Keller *et al*. for the p.(Lys312del) variant (22). Our structural modeling data show impaired NLS protein conformation and predicted bipartite NLS loss for both p.(Ala304Val) and p.(Lys312del) mutants. In support of this observation, it is known that Ala residues can be involved in substrate recognition or specificity and that Lys residues are quite common in protein-active or -binding sites (25), which is further illustrated by structural modeling on the CREB bZIP-CRE complex (30). Luciferase assays further demonstrated that overexpression of both mutant A304V and K312del proteins have a similar negative impact on

activation of type I collagen transcription. Together, these findings suggest that both p.(Ala304Val) and p.(Lys312del) have similar working mechanisms; they both form stable mutant proteins, which subsequently might accumulate in the cytosol. Although we cannot fully exclude the possibility that residual mutant proteins are translocated to the nucleus, we hypothesize that these mutant proteins have compromised promotor binding ability.

By performing overexpression studies, we also demonstrated that both the p.(Ala304Val) and the p.(Lys312del) mutants do not have the same ability as WT OASIS to increase protein levels of SEC24D, and therefore confirmed SEC24D as a target of OASIS which is potentially relevant to the pathogenesis of OI (22). SEC24D is part of the well-conserved COPII coat secretory pathway, and bi-allelic pathogenic variants in SEC24D result in a syndromic form of OI, resembling Cole-Carpenter syndrome (MIM 616294), with skull ossification defects and fractures (22, 34, 35). COPII vesicles are processed when proteins, destined for downstream intracellular compartments, are sorted and packaged at discrete sites on the ER membrane, also called 'ER exit sites' (36). COPII structures consist of an inner (SEC23/24) and outer (SEC13/31) coat, providing stability. Their formation is aided by SAR1, SEC12 and SEC16, before transporting its cargo proteins to the ER-Golgi intermediate compartment and Golgi apparatus (37). Loading of large COPII vesicles, needed for packaging of procollagens, is enabled by the auxiliary proteins cTAGE5 (cutaneous T-cell lymphoma-associated 5) and TANGO1 (transport and Golgi organization 1), with the latter recruiting Sedlin (another helper protein) in a later stage. Recent studies have shown that monoubiquitylation of SEC31A helps to regulate COPII size, that glycosylation of both SEC24 and SEC23 is important for organization and regulation of COPII vesicles, and that phosphorylation of SEC23 and SEC24 confers directionality on COPII vesicles from ER to Golgi (34, 38). By now, it is well established that the secretion of procollagen requires an optimal working of the COPII secretory pathway. Perturbation of COPII components, and of global regulators of COPII expression, such as the transcription factors CREB3L2 and CREB3L1, which promote transcription of SEC23A and SEC24D, respectively, all result in defects in procollagen secretion and extracellular matrix assembly (34). Saito *et al*. demonstrated that *Creb3l2*$^{-/-}$ chondrocytes accumulated type II collagen and other cartilage matrix proteins in the ER lumen (22, 23, 33). Keller *et al*. first proposed that mutations in OASIS can lead to OI due to disruption of the important role this protein plays in the secretion of type I collagen and other bone matrix proteins from osteoblasts during osteogenesis (22, 23, 33). This hypothesis was first confirmed in a recent study by Lindahl *et al*. and is now strongly supported by the new evidence provided in this report (22, 23, 33).

In conclusion, this report of the first homozygous pathogenic missense variant broadens the mutational and phenotypic spectrum of *CREB3L1*-related OI and provides additional proof of the lack of an optimal working COPII secretory complex as a potentially critical factor in the context of bone-associated disease.

Materials and Methods

Ethical considerations

Written and signed informed consent was obtained from the parents of the patient participating in this study. Genomic DNA (gDNA) from the proband, siblings or parents was isolated from whole blood according to the standard procedures.

Molecular studies

We used conventional Sanger sequencing and next generation panel sequencing (MiSeq platform – Illumina) for molecular screening of the *COL1A1*, *COL1A2*, *CRTAP*, *LEPRE1*, *PPIB*, *CREB3L1*, *WNT1*, *PLS3*, *BMP1*, *FKBP10*, *IFITM5*, *PLOD2*, *SERPINF1*, *SERPINH1*, *SP7* and *TMEM38B* genes. For NGS, single bases (up to 20 bases intronic of all coding exons) were covered with a minimal of 30x. Confirmational Sanger sequencing and segregational analysis was performed using the BigDye Terminator Cycle Sequencing Kit (Life Technologies, Carlsbad, Ca, USA) and run on a ABI 3730XL DNA Analyzer (Life technologies).

Nucleotide numbering reflects cDNA numbering, with +1 corresponding to the A nucleotide of the ATG translation initiation codon in the reference sequence of *CREB3L1* (NM_052854.2). AA residues are numbered from the first methionine residue of the protein reference sequence (NP_005421.1). Variant nomenclature follows the Human Genome Variation Society (HGVS) guidelines (http://www.hgvs.org/mutnomen), and variant classification was done by using the Alamut Visual software (version 2.10) and according to the American College of Medical Genetics (ACMG) standards and guidelines (Genome Aggregation Database, http://gnomad.broadinstitute.org) (24, 39). In addition, DEOGEN2 was used to check the mutation effect prediction on protein level (overall score and amino acid similarity) (31), and the variant was checked and submitted to the OI Variant Database (http://www.le.ac.uk/ge/collagen/).

Structural modeling of the variant

By means of the I-TASSER server, which is an integrated platform for automated protein structure and function prediction based on the sequence-to-structure-to-function paradigm, 5 different (monomeric) three dimensional structural protein models were generated of the full length WT, p.(Ala304Val), and p.(Lys312del) protein sequences (26-28). We retained 3 models

in which a clear (functional) helical bZIP domain could be distinguished within the expected AA positions 292 and 353 (data for the other 2 models is not shown) (Figure 3).

The homology model of the CREB bZIP-CRE complex (PDB: 1DH3 – Mus musculus – generated in the expression system of *Escherichia coli*) was used as a template (30). The UCSF Chimera software package (version 1.13, build 41780) was used to visualize, study the localization, and model the effect of the specific protein variant (Dunbrack rotamers and FindHBond function), respectively (32, 40) (Supplemental Figure S3).

Expression vectors

Starting from a human WT cDNA OASIS construct (generated on a pCMV-3Tag-2 backbone, 'WT'), we used the QuikChange II Site-Directed Mutagenesis Kit to generate mutant constructs for our identified variant (c.911C>T, p.(Ala304Val), 'A304V') and the earlier reported in-frame deletion (c.934_936delAAG, p.(Lys312del), 'K312del') (22, 41). The primers for site-directed mutagenesis were designed using the QuikChange Primer Design tool (Agilent) and were purchased as HPLC-purified primers (primer sequences are listed in the Supplementary Table S1) (Integrated DNA Technologies). A pCMV-3Tag-2 empty vector (cat240196, Agilent) was purchased to use as a transfection control in our experiments ('Empty'). Final constructs were sequenced, and a control- digestion was performed to confirm correct vector structure (data not shown).

Luciferase reporter assay

For the luciferase experiments, 20,000 HEK293 cells were seeded in clear bottom 96 well plates (CLS3603-48EA, Sigma-Aldrich) in triplicate at day 1 and transiently co-transfected at day 3 using FuGene HD transfection reagent (E2311, Promega). Per reaction, 40ng of 'Empty', 'WT', 'A304V' or 'K312del' was combined with 40ng of reporter constructs for *Col1a1* ('*Col1a1* prom' contains the 2.3-kb *Col1a1* promoter and UPRE (TGACGTGG)-like sequence (CGACGTGG), '*Col1a1* prom mUPRE' contains the 2.3-kb *Col1a1* promoter and mutant UPRE-like sequence (CGAaGgGG), 10ng of Renilla luciferase expression construct and 0.3µl FuGene HD transfection reagent (21). Twenty-four hours post transfection, cells were lysed according to the manufacturers guidelines (Dual-Glo Luciferase Assay System, Promega) and luciferase activity was measured using a GloMax-Multi Detection System (E7031, Promega). Data were normalized to Renilla luciferase and log10 transformed. Graphs display data-points normalized to WT values.

In vitro overexpression studies

In brief, 200,000 HEK293 cells were seeded in 6-well plates in triplicate at day 1 and transiently transfected at day 2 using FuGene HD transfection reagent (E2311, Promega) at a 3:1 ratio (3µl reagent: 1µg plasmid) per well and incubated for 48 hours before harvesting. These cells were subsequently processed for quantitative reverse-transcription PCR (RT-qPCR) or immunoblotting.

Quantitative reverse transcription PCR

Total RNA was extracted from transfected HEK293 cells using the RNeasy Kit (QIAGEN). Starting from 2µg of RNA, cDNA was subsequently synthesized with the iScript cDNA Synthesis Kit (Bio-Rad Laboratories). Primer sequences are listed in the Supplementary Table S1) (Integrated DNA Technologies). RT-qPCR reactions were prepared with the addition of RealTime ready DNA Probes Master mix and ResoLight Dye (Roche) and were run in duplicate on a Roche LightCycler 480 System. Data were analyzed with qbase+ software (version 3.0, Biogazelle) (42), and expression was normalized to the housekeeping genes *HPRT1*, *RPL13A* and *YWHAZ*. Graphs display data-points normalized to WT values.

Immunoblotting

For immunoblotting of OASIS, SEC24D and SEC23A, protein lysates were prepared from transfected HEK293 cells using a 0.05M Tris-HCL buffer (pH 8.0, 0.15M NaCl, 5.0mM EDTA, 1% NP-40 and protease inhibitor cocktail) at 4°C and subjected to SDS-PAGE under reduced condition (6.25% 1M dithiothreitol) (NP0335BOX, Life Technologies Europe). Proteins were transferred to a nitrocellulose membrane with the iBlot 2 Dry Blotting System (Thermo Fisher Scientific). Membranes were blocked in 5% dry milk (OASIS and SEC24D) or 2% ECL Blocking Agent (GE Healthcare) (SEC23A and β -tubulin), incubated overnight with primary antibodies against OASIS (1/1000; ab33051; Abcam), SEC24D (1/1500, ab191566, Abcam), SEC23A (1/500, ab179811, Abcam) or β -tubulin (1/1500, ab6046, Abcam) and subsequently incubated with horseradish peroxidase-conjugated secondary antibody (1/1500, 7074S, Bioké BV). Membranes were scanned with an Amersham Imager 680 System (GE Healthcare), quantitation was achieved using ImageJ and normalized to the amount of b-tubulin. Graphs display data-points normalized to WT values.

Statistical analysis

Statistical analysis was performed using GraphPad Prism 7.04 software. The RT-qPCR, immunoblotting and luciferase reporter assay results are expressed as mean ± standard error of the mean (SEM) from three independent experiments, and statistical significance was determined by performing one-way ANOVA followed by Sidak's test for multiple comparisons (see figure legends).

Funding

This work was supported by Research Foundation Flanders [12Q5917N to D.S., 1842318N to F.M.]; Ghent University [08/01M01108 to A.D.P.]; and the European Union's Horizon 2020 research and innovation programme Marie Skłodowska-Curie [794365 to P.S.].

Acknowledgements

The *Col1a1* reporter and renilla constructs were kindly provided by Professor Kazunori Imaizumi (University of Miyazaki, Japan), the WT cDNA OASIS construct was a kind gift of Professor Allen Volchuk (Toronto General Research Institute, Canada).

Conflict of Interest statement. None declared.

References

1. Forlino,A. and Marini,J.C. (2016) Osteogenesis imperfecta. *Lancet*, **387**, 1657–1671.

2. Marini,J.C., Forlino,A., Bächinger,H.P., Bishop,N.J., Byers,P.H., De Paepe,A., Fassier,F., Fratzl-Zelman,N., Kozloff,K.M., Krakow,D., *et al.* (2017) Osteogenesis imperfecta. *Nature Publishing Group*, **3**, 1–19.

3. Kang,H., Aryal A C,S. and Marini,J.C. (2017) Osteogenesis imperfecta: new genes reveal novel mechanisms in bone dysplasia. *Translational Research*, **181**, 27–48.

4. Semler,O., Garbes,L., Keupp,K., Swan,D., Zimmermann,K., Becker,J., Iden,S., Wirth,B., Eysel,P., Koerber,F., *et al.* (2012) A mutation in the 5'-UTR of IFITM5 creates an in-frame start codon and causes autosomal-dominant osteogenesis imperfecta type V with hyperplastic callus. *Am. J. Hum. Genet.*, **91**, 349–357.

5. Van Dijk,F.S., Nesbitt,I.M., Zwikstra,E.H., Nikkels,P.G.J., Piersma,S.R., Fratantoni,S.A., Jimenez,C.R., Huizer,M., Morsman,A.C., Cobben,J.M., *et al.* (2009) PPIB mutations cause severe osteogenesis imperfecta. *Am. J. Hum. Genet.*, **85**, 521–527.

6. Becker,J., Semler,O., Gilissen,C., Li,Y., Bolz,H.J., Giunta,C., Bergmann,C., Rohrbach,M., Koerber,F., Zimmermann,K., *et al.* (2011) Exome sequencing identifies truncating mutations in human SERPINF1 in autosomal-recessive osteogenesis imperfecta. *Am. J. Hum. Genet.*, **88**, 362–371.

7. Morello,R., Bertin,T.K., Chen,Y., Hicks,J., Tonachini,L., Monticone,M., Castagnola,P., Rauch,F., Glorieux,F.H., Vranka,J., *et al.* (2006) CRTAP is required for prolyl 3-hydroxylation and mutations cause recessive osteogenesis imperfecta. *Cell*, **127**, 291–304.

8. Marini,J.C., Cabral,W.A., Barnes,A.M. and Chang,W. (2007) Components of the collagen prolyl 3-hydroxylation complex are crucial for normal bone development. *Cell Cycle*, **6**, 1675–1681.

9. Christiansen,H.E., Schwarze,U., Pyott,S.M., AlSwaid,A., Balwi,Al,M., Alrasheed,S., Pepin,M.G., Weis,M.A., Eyre,D.R. and Byers,P.H. (2010) Homozygosity for a missense

mutation in SERPINH1, which encodes the collagen chaperone protein HSP47, results in severe recessive osteogenesis imperfecta. *Am. J. Hum. Genet.*, **86**, 389–398.

10. Steinlein,O.K., Aichinger,E., Trucks,H. and Sander,T. (2011) Mutations in FKBP10 can cause a severe form of isolated Osteogenesis imperfecta. *BMC Medical Genetics 2018 19:1*, **12**, 152.

11. Ha-Vinh,R., Alanay,Y., Bank,R.A., Campos-Xavier,A.B., Zankl,A., Superti-Furga,A. and Bonafé,L. (2004) Phenotypic and molecular characterization of Bruck syndrome (osteogenesis imperfecta with contractures of the large joints) caused by a recessive mutation in PLOD2. *Am. J. Med. Genet. A*, **131**, 115–120.

12. Syx,D., Guillemyn,B., Symoens,S., Sousa,A.B., Medeira,A., Whiteford,M., Hermanns-Lê,T., Coucke,P.J., De Paepe,A. and Malfait,F. (2015) Defective Proteolytic Processing of Fibrillar Procollagens and Prodecorin Due to Biallelic BMP1 Mutations Results in a Severe, Progressive Form of Osteogenesis Imperfecta. *J. Bone Miner. Res.*, **30**, 1445–1456.

13. Volodarsky,M., Markus,B., Cohen,I., Staretz-Chacham,O., Flusser,H., Landau,D., Shelef,I., Langer,Y. and Birk,O.S. (2013) A deletion mutation in TMEM38B associated with autosomal recessive osteogenesis imperfecta. *Hum. Mutat.*, **34**, 582–586.

14. Keupp,K., Beleggia,F., Kayserili,H., Barnes,A.M., Steiner,M., Semler,O., Fischer,B., Yigit,G., Janda,C.Y., Becker,J., *et al.* (2013) Mutations in WNT1 Cause Different Forms of Bone Fragility. *The American Journal of Human Genetics*, **92**, 565–574.

15. Symoens,S., Malfait,F., D'hondt,S., Callewaert,B., Dheedene,A., Steyaert,W., Bächinger,H.P., De Paepe,A., Kayserili,H. and Coucke,P.J. (2013) Deficiency for the ER-stress transducer OASIS causes severe recessive osteogenesis imperfecta in humans. *Orphanet J Rare Dis*, **8**, 154.

16. Mendoza-Londono,R., Fahiminiya,S., Majewski,J., Care4Rare Canada Consortium, Tétreault,M., Nadaf,J., Kannu,P., Sochett,E., Howard,A., Stimec,J., *et al.* (2015) Recessive osteogenesis imperfecta caused by missense mutations in SPARC. *Am. J. Hum. Genet.*, **96**, 979–985.

17. Lindert,U., Cabral,W.A., Ausavarat,S., Tongkobpetch,S., Ludin,K., Barnes,A.M., Yeetong,P., Weis,M., Krabichler,B., Srichomthong,C., *et al.* (2016) MBTPS2 mutations

cause defective regulated intramembrane proteolysis in X-linked osteogenesis imperfecta. *Nat Commun*, **7**, 11920.

18. Schwarze,U., Cundy,T., Pyott,S.M., Christiansen,H.E., Hegde,M.R., Bank,R.A., Pals,G., Ankala,A., Conneely,K., Seaver,L., *et al.* (2013) Mutations in FKBP10, which result in Bruck syndrome and recessive forms of osteogenesis imperfecta, inhibit the hydroxylation of telopeptide lysines in bone collagen. *Hum Mol Genet*, **22**, 1–17.

19. Honma,Y., Kanazawa,K., Mori,T., Tanno,Y., Tojo,M., Kiyosawa,H., Takeda,J., Nikaido,T., Tsukamoto,T., Yokoya,S., *et al.* (1999) Identification of a novel gene, OASIS, which encodes for a putative CREB/ATF family transcription factor in the long-term cultured astrocytes and gliotic tissue. *Brain Res. Mol. Brain Res.*, **69**, 93–103.

20. Kondo,S., Murakami,T., Tatsumi,K., Ogata,M., Kanemoto,S., Otori,K., Iseki,K., Wanaka,A. and Imaizumi,K. (2005) OASIS, a CREB/ATF-family member, modulates UPR signalling in astrocytes. *Nat. Cell Biol.*, **7**, 186–194.

21. Murakami,T., Saito,A., Hino,S.-I., Kondo,S., Kanemoto,S., Chihara,K., Sekiya,H., Tsumagari,K., Ochiai,K., Yoshinaga,K., *et al.* (2009) Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation. *Nat. Cell Biol.*, **11**, 1205–1211.

22. Keller,R.B., Tran,T.T., Pyott,S.M., Pepin,M.G., Savarirayan,R., McGillivray,G., Nickerson,D.A., Bamshad,M.J. and Byers,P.H. (2018) Monoallelic and biallelic CREB3L1 variant causes mild and severe osteogenesis imperfecta, respectively. *Genet Med*, **20**, 411–419.

23. Lindahl,K., Åström,E., Dragomir,A., Symoens,S., Coucke,P., Larsson,S., Paschalis,E., Roschger,P., Gamsjaeger,S., Klaushofer,K., *et al.* (2018) Homozygosity for CREB3L1 premature stop codon in first case of recessive osteogenesis imperfecta associated with OASIS-deficiency to survive infancy. *Bone*, **114**, 268–277.

24. Richards,S., Aziz,N., Bale,S., Bick,D., Das,S., Gastier-Foster,J., Grody,W.W., Hegde,M., Lyon,E., Spector,E., *et al.* (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*, **17**, 405–423.

25. Barnes,M.R. (2003) Bioinformatics for Geneticists.

26. Roy,A., Kucukural,A. and Zhang,Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*, **5**, 725–738.

27. Zhang,Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **9**, 40.

28. Yang,J., Yan,R., Roy,A., Xu,D., Poisson,J. and Zhang,Y. (2015) The I-TASSER Suite: protein structure and function prediction. *Nat. Methods*, **12**, 7–8.

29. Kosugi,S., Hasebe,M., Tomita,M. and Yanagawa,H. (2009) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 10171–10176.

30. Schumacher,M.A., Goodman,R.H. and Brennan,R.G. (2000) The Structure of a CREB bZIP·Somatostatin CRE Complex Reveals the Basis for Selective Dimerization and Divalent Cation-enhanced DNA Binding. *J. Biol. Chem.*, **275**, 35242–35247.

31. Raimondi,D., Tanyalcin,I., Ferté,J., Gazzo,A., Orlando,G., Lenaerts,T., Rooman,M. and Vranken,W. (2017) DEOGEN2: prediction and interactive visualization of single amino acid variant deleteriousness in human proteins. *Nucleic Acids Res.*, **45**, W201–W206.

32. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, **25**, 1605–1612.

33. Saito,A., Hino,S.-I., Murakami,T., Kanemoto,S., Kondo,S., Saitoh,M., Nishimura,R., Yoneda,T., Furuichi,T., Ikegawa,S., *et al.* (2009) Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis. *Nat. Cell Biol.*, **11**, 1197–1204.

34. McCaughey,J. and Stephens,D.J. (2018) COPII-dependent ER export in animal cells: adaptation and control for diverse cargo. *Histochem. Cell Biol.*, **150**, 119–131.

35. Garbes,L., Kim,K., Rieß,A., Hoyer-Kuhn,H., Beleggia,F., Bevot,A., Kim,M.J., Huh,Y.H., Kweon,H.-S., Savarirayan,R., *et al.* (2015) Mutations in SEC24D, encoding a component

of the COPII machinery, cause a syndromic form of osteogenesis imperfecta. *Am. J. Hum. Genet.*, **96**, 432–439.

36. Hughes,H., Budnik,A., Schmidt,K., Palmer,K.J., Mantell,J., Noakes,C., Johnson,A., Carter,D.A., Verkade,P., Watson,P., *et al.* (2009) Organisation of human ER-exit sites: requirements for the localisation of Sec16 to transitional ER. *Journal of Cell Science*, **122**, 2924–2934.

37. Unlu,G., Levic,D.S., Melville,D.B. and Knapik,E.W. (2014) Trafficking mechanisms of extracellular matrix macromolecules: insights from vertebrate development and human diseases. *Int. J. Biochem. Cell Biol.*, **47**, 57–67.

38. Jin,L., Pahuja,K.B., Wickliffe,K.E., Gorur,A., Baumgärtel,C., Schekman,R. and Rape,M. (2012) Ubiquitin-dependent regulation of COPII coat size and function. *Nature*, **482**, 495–500.

39. Wildeman,M., van Ophuizen,E., Dunnen,den,J.T. and Taschner,P.E.M. (2008) Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. *Hum. Mutat.*, **29**, 6–13.

40. Dunbrack,R.L. (2002) Rotamer libraries in the 21st century. *Curr. Opin. Struct. Biol.*, **12**, 431–440.

41. Vellanki,R.N., Zhang,L., Guney,M.A., Rocheleau,J.V., Gannon,M. and Volchuk,A. (2010) OASIS/CREB3L1 induces expression of genes involved in extracellular matrix production but not classical endoplasmic reticulum stress response genes in pancreatic beta-cells. *Endocrinology*, **151**, 4146–4157.

42. Hellemans,J., Mortier,G., De Paepe,A., Speleman,F. and Vandesompele,J. (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.*, **8**, R19.

Legends to Figures

Figure 1 Pedigree and clinical findings. (A): Pedigree of the Turkish *CREB3L1* OI family. The proband is indicated with an arrow, asterisks denote family members available for molecular testing. (B): Postmortem examination of fetus IV:3 showed bowed extremities with bilateral angulation of the forearms due to fractures, bilateral femoral and tibial bowing. (C): Anterior-posterior and (D) lateral radiographs of fetus IV:3 revealed thin, wavy ribs and multiple fractures of tubular bones resembling accordion-like femora and humeri.

Figure 2 Protein structure and function of OASIS. OASIS is a 519 AAs long protein containing an N-terminal cytoplasmic part, which holds the conserved bZIP domain (AAs 292-353), and a transmembrane domain (TM), which anchors it in the rough endoplasmic reticulum membrane. When a (bone) cell is stressed or depleted, the full-length OASIS is transported to the Golgi, where it is cleaved through RIP at the S1P and S2P sites. The released N-terminal active factor of OASIS is translocated to the nucleus, where it activates the transcription of its target genes (*e.g. COL1A1*, *VEGFA*, *SEC24D*). The NLS is shown, the RRKKKEY DNA binding domain is depicted in bold, and the AAs at positions 304 (c.911C>T, p.(Ala304Val)), 312 (c.934_936delAAG, p.(Lys312del)) and 428 (c.1284C>A, p.(Tyr428*)) are highlighted (2, 15, 19, 20, 22, 23).

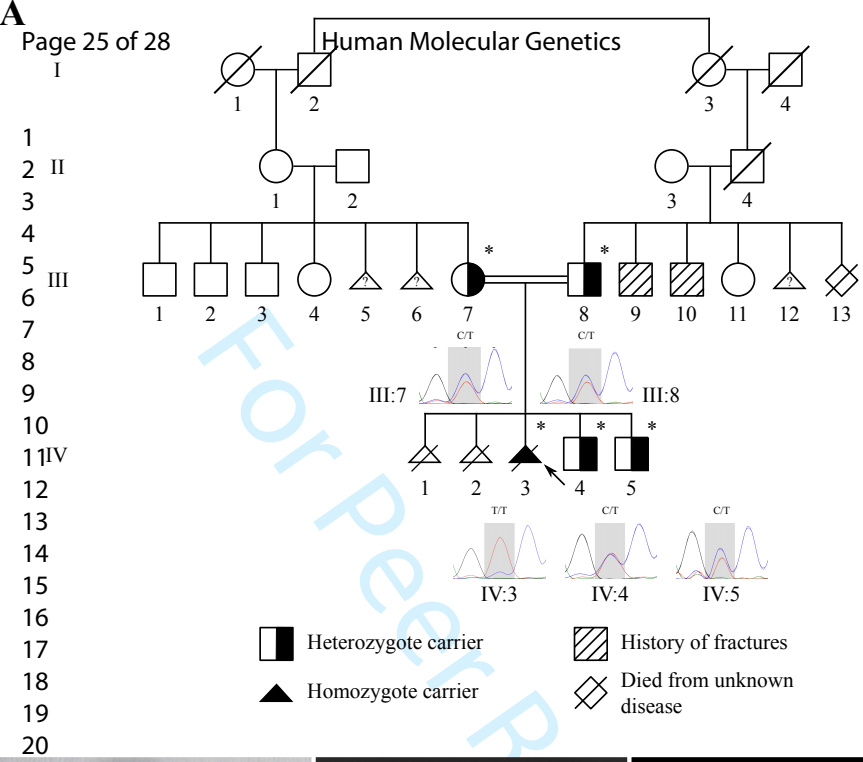
Figure 3 Structural modeling of human WT and mutant OASIS. Structural modeling of the full length WT, p.(Ala304Val), and p.(Lys312del) OASIS protein sequences, by means of I-TASSER algorithms. Conformational changes within the functional helical bZIP domain containing the NLS (blue) are marked with arrows, highlighting the effects of the p.(Ala304Val) and p.(Lys312del) variants.

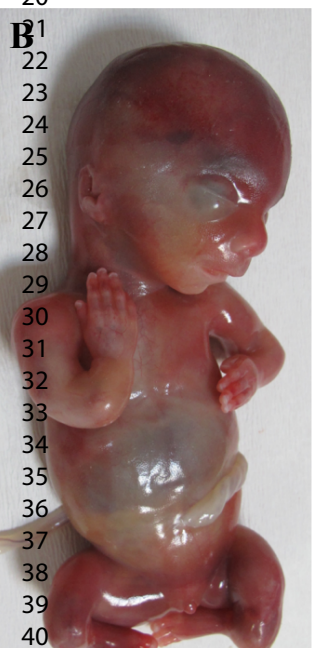
Figure 4 Effect of overexpression of the Ala304Val variant on *Col1a1* promoter activity, and on mRNA/protein levels of the COPII components SEC23A/SEC24D, respectively. (A) and (B): Luciferase assays provide strong evidence that A304V has a negative effect on OASIS-induced transcriptional activation of the *Col1a1* gene, similar to K312del. (C) and (D): RT-qPCR shows that both A304V and K312del have a negative effect on the expression of *SEC23A* and *SEC24D*, when compared to WT overexpressed *CREB3L1*. (E), (F), (G), (H) and (I): Immunoblotting shows that A304V is a stably expressed mutant protein that does not appear

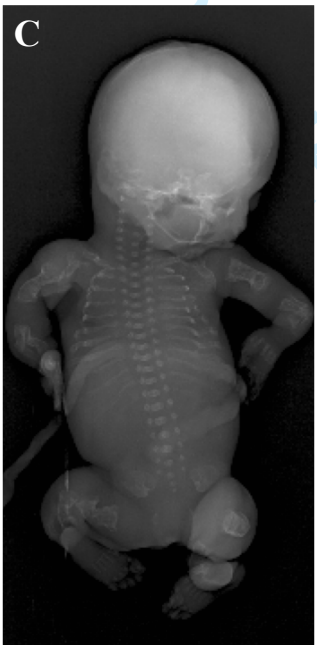
to be truncated when compared to WT protein. Both A304V and K312del prevent OASIS-induced increase of SEC24D protein levels, while SEC23A protein levels are not affected by OASIS. Values shown are the mean of three independent experiments; Empty, empty vector-transfected control; WT, wild type-transfected OASIS. Error bars, SEM. (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).

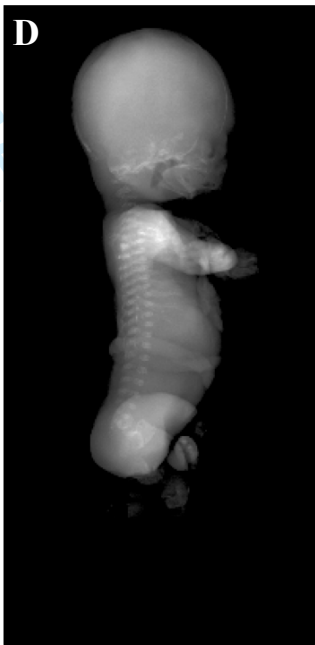
Abbreviations

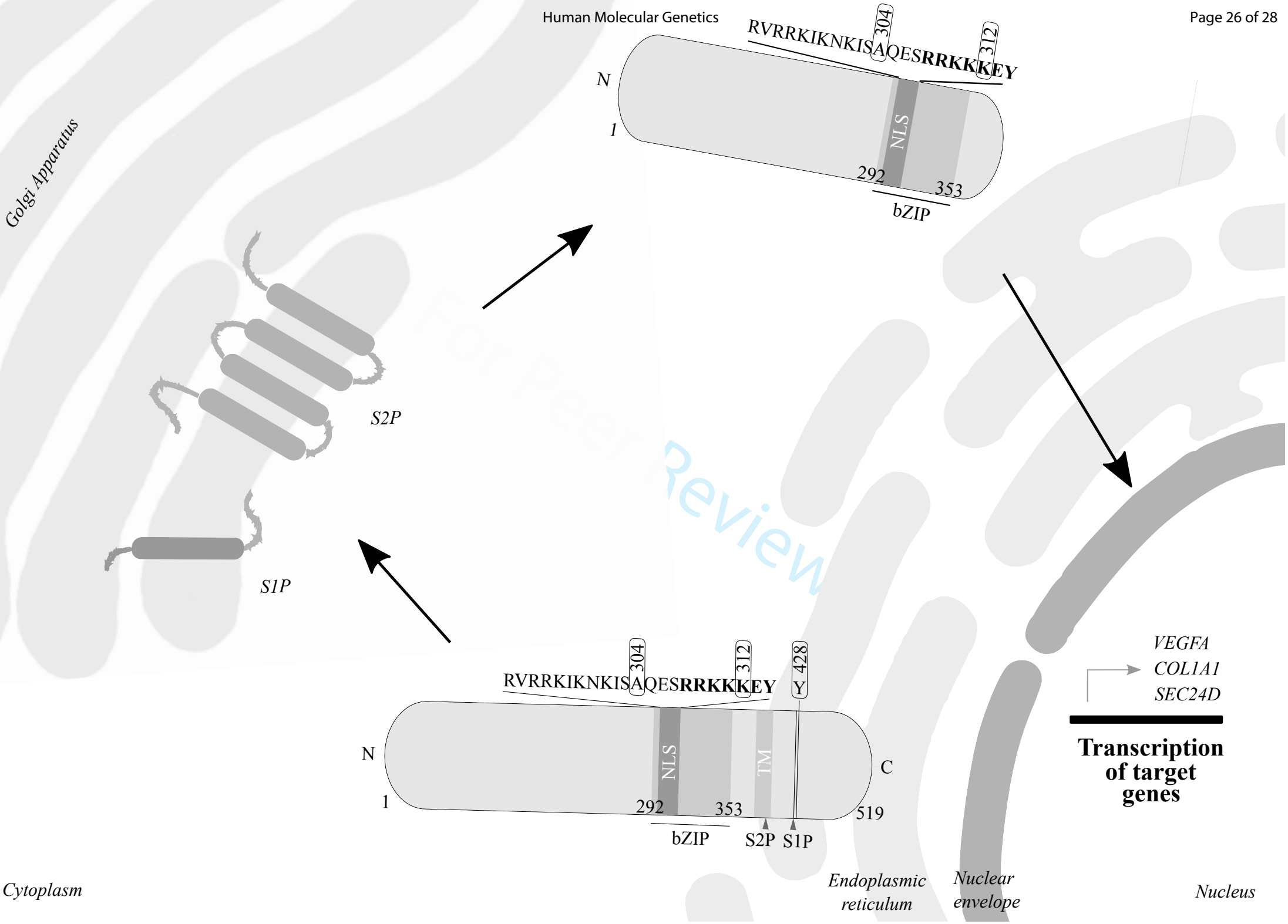
CREB3L1	cyclic AMP responsive element binding protein 3-like 1
OASIS	old astrocyte specifically induced substance
OI	osteogenesis imperfecta
COPII	coat protein complex type II
AD	autosomal dominant
AR	autosomal recessive
ER	endoplasmic reticulum
bZIP	basic leucine zipper
RIP	regulated intramembrane proteolysis
UPRE	unfolded protein response element
VUS	variant of unknown significance
NLS	nuclear localization sequence
AA	amino acid
I-TASSER	Iterative Threading ASSembly Refinement
cTAGE5	cutaneous T-cell lymphoma-associated 5
TANGO1	transport and Golgi organization 1
ACMG	American College of Medical Genetics


A
I
II
III
IV
1
2
3
4
5
6
7
8
9
10
11
12
13
C/T
III:7
III:8
T/T
IV:3
IV:4
IV:5
Heterozygote carrier
Homozygote carrier
History of fractures
Died from unknown disease

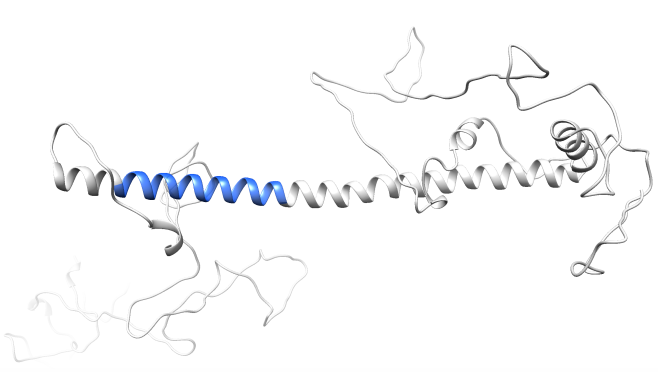

B


C

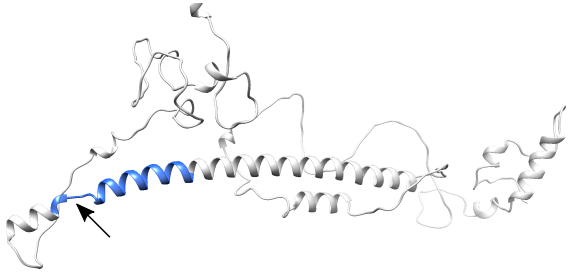

D


Golgi Apparatus
RVRRKIKNKISAQESRRKKKEY
304
312
N
1
NLS
292
353
bZIP
S2P
S1P
RVRRKIKNKISAQESRRKKKEY
304
312
428
Y
N
C
1
NLS
TM
292
353
519
bZIP
S2P
S1P
VEGFA
COL1A1
SEC24D
Transcription of target genes
Cytoplasm
Endoplasmic reticulum
Nuclear envelope
Nucleus

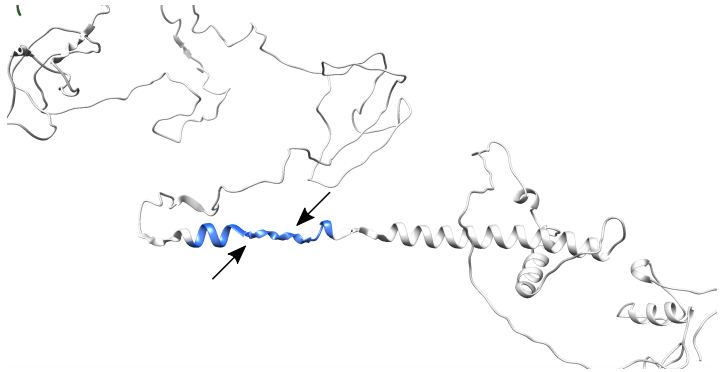
WT OASIS - MODEL1



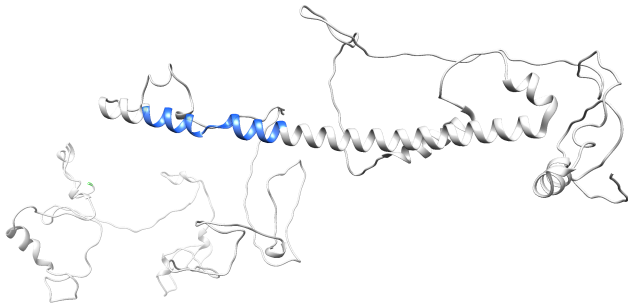
p.(Ala304Val) - MODEL1



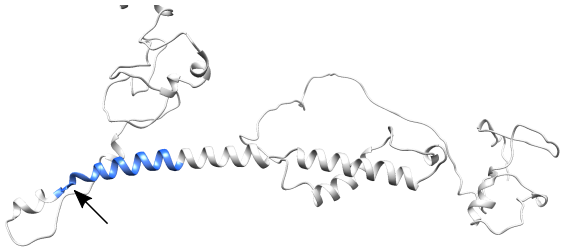
p.(Lys312del) - MODEL1



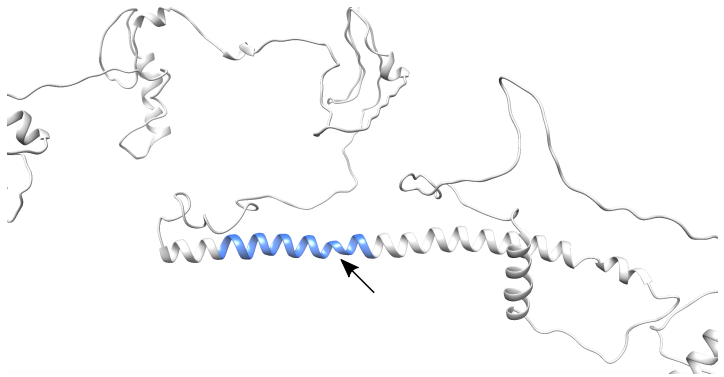
WT OASIS - MODEL2



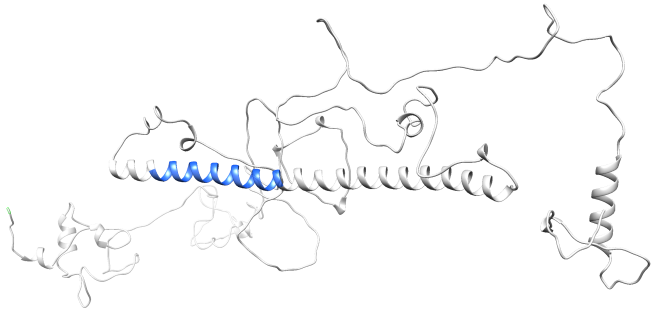
p.(Ala304Val) - MODEL2



p.(Lys312del) - MODEL2



WT OASIS - MODEL3



p.(Ala304Val) - MODEL3

p.(Lys312del) - MODEL3