Received: 28 September 2021 | Revised: 11 April 2022 | Accepted: 13 May 2022

DOI: 10.1002/mgg3.1996

Molecular Genetics & Genomic Medicine Open Access WILEY

CLINICAL REPORT

The first glycine-to-tryptophan substitution in the *COL1A1* gene identified in a patient with progressively-deforming Osteogenesis imperfecta

Kinga Sałacińska[1] | **Izabela Michałus**[2] | **Iwona Pinkier**[1] | **Lena Rutkowska**[1] | **Danuta Chlebna-Sokół**[3] | **Elżbieta Jakubowska-Pietkiewicz**[4] | **Łukasz Kępczyński**[1] | **Dominik Salachna**[1] | **Agnieszka Gach**[1]

[1]Department of Genetics, Polish Mother's Memorial Hospital Research Institute, Lodz, Poland

[2]Department of Endocrinology and Metabolic Diseases, Polish Mother's Memorial Hospital Research Institute, Lodz, Poland

[3]Department of Bone Metabolic Diseases, University Centre of Paediatric, Medical University of Lodz, Lodz, Poland

[4]Department of Paediatric Propedeutics and Bone Metabolic Diseases, Medical University of Lodz, Lodz, Poland

Correspondence
Kinga Sałacińska, Department of Genetics, Polish Mother's Memorial Hospital Research Institute, Lodz, Poland.
Email: kinga.salacinska@iczmp.edu.pl

Funding information
Narodowe Centrum Nauki, Grant/Award Number: 2014/13/B/NZ5/03102; Polish Mother's Memorial Hospital Research Institute, Grant/Award Number: 2016/IV/MN-57

Abstract

Background: Osteogenesis imperfecta (OI) is a genetic disorder of connective tissue with variable phenotype and heterogeneous genetic background. Majority of reported mutations are glycine substitutions, whose clinical outcome ranges from mild to perinatal lethal. The phenotype appears to be influenced by the properties of amino acid side chain and the degree of structural aberration of collagen molecules. Since the genotype–phenotype correlation remains unclear, the severity of mutation is mostly predicted according to previously-reported cases. Although the number of OI variants is constantly expanding, no glycine-to-tryptophan substitutions have been reported in *COL1A1* gene.

Methods: A sample from a 15-year-old girl presenting with progressively-deforming OI type III was tested using an NGS custom gene panel. Multiple bioinformatic and interpretation tools, including mutation databases and conservation analysis, were used for variant classification. The presence of the mutation was verified by Sanger sequencing.

Results: A novel heterozygous mutation c.733G>T was identified in the *COL1A1* gene (p.Gly245Trp).

Conclusions: The discovery of this novel glycine-to-tryptophan substitution located in the *COL1A1* gene broadens the spectrum of mutations underlying this rare disease and provides useful information on the clinical outcome of such substitutions.

KEYWORDS

collagen type I, genotype–phenotype correlation, Next Generation Sequencing, osteogenesis imperfecta

1 | INTRODUCTION

Osteogenesis imperfecta (OI) is a rare genetic connective tissue disorder. As the disease affects a number of tissues, the patients present a variety of skeletal and extra-skeletal features, the most characteristic being: susceptibility to fractures resulting from mild trauma, growth deficiency, long-bone and chest deformations, triangular-shaped face, blue sclera and dental abnormalities (Dentinogenesis imperfecta). The classification of OI was first established by Sillence et al. in 1979, who divided the disease into four subgroups based on clinical severity: Type I, with a mild non-deforming phenotype, Type II, associated with perinatal lethality, Type III, with a severe nonlethal progressively-deforming outcome, and Type IV, with moderate severity, i.e., intermediate between types I and III (Forlino & Marini, 2016; Steiner & Basel, 2005).

The disease is characterized by a heterogeneous genetic background with multiple loci, resulting in the OI phenotype. About 85%–90% of cases are associated with dominantly-inherited pathogenic variants in the *COL1A1* (OMIM 120150) and COL1A2 (OMIM 120160) genes, most of which result from glycine substitutions known to cause qualitative changes in collagen type I structure. It is essential that glycine, which is the smallest amino acid, is present in every third position of the triple helical domain to maintain the proper structural conformation of pro-alpha1(I) and pro-alpha2(I) chains. Any substitution for glycine causes a qualitative alteration of the chains that are incorporated into heterotrimers, resulting in impaired helix formation, thus influencing the stability of the resulting collagen molecule and its interactions with other extracellular molecules. The clinical outcomes of mutations associated with amino acid substitutions range from mild to perinatal lethal depending on the substituent (Bodian et al., 2008; Forlino et al., 2011).

Potentially pathogenic mutations may occur in any of the 338 Gly-Xaa-Yaa triplets following the change of a single nucleotide in the codon for glycine (GGG, GGA, GGC, GGT). Depending on the location, the change can result in the glycine substitution for one of eight possible amino acids, i.e., alanine, arginine, aspartic acid, cysteine, glutamic acid, serine, tryptophan, or valine (Dalgleish, 2014). Previous studies on patients with OI have reported examples of seven such substitutions in the *COL1A1* gene, involving all but the tryptophan residue, this being a large, non-polar aromatic amino acid containing an indole side chain.

There are only ten potential loci in the *COL1A1* gene that may allow for glycine (GGG)-to-tryptophan (TGG) substitution. In the *COL1A2* gene, three distinct glycine-to-tryptophan mutations have been noted in the OIVD (Osteogenesis Imperfecta Variant Database) to date: p.Gly319Trp, p.Gly367Trp and p.Gly814Trp (Dalgleish & Kheng Teh, 2020; Hrušková et al., 2015; Nuytinck et al., 2000).

Our present study reports the first glycine-to-tryptophan substitution in the *COL1A1* gene, identified in a patient with a progressively-deforming OI phenotype. This description of a novel mutation broadens our understanding of the heterogeneous genetic background of this rare disease.

2 | MATERIALS AND METHODS

2.1 | Subjects

A 15-year-old girl of Caucasian origin was admitted to hospital with a suspected severe form of OI. At the age of 11, molecular screening for pathogenic mutation was performed, as the patient had been included in a research project on genetic background of OI. The study was approved by the local Ethics Committee (No. 108/2015). The legal guardian gave informed consent for the patient to take part in the study.

2.2 | Methods

The pathogenic variant was detected by Next Generation Sequencing (NGS), using a custom-designed gene panel for resequencing exons and flanking non-coding sequences of known and candidate genes related to OI (Design Studio software, Illumina, USA) (Sałacińska et al., 2021). Genomic DNA was isolated automatically from biobanked peripheral blood (MagCore Genomic DNA Whole Blood Kit, RBC Bioscience, Taiwan). A library composed of 710 amplicons, with a mean length of 175 bp and 99% coverage, was prepared using TruSeq Custom Amplicon reagents, according to the manufacturer's protocol. The library was sequenced on a MiniSeq platform (Illumina, USA).

The obtained data was prioritized and variant filtering was performed using Illumina VariantStudio software (human genome reference GRCh37 version). Further analysis was performed using the Varsome freely-available variant interpretation tool (https://varsome.com/), supplemented by information available in public databases (dbSNP, ClinVar, HGMD, OIVD) and literature. Variant analysis was performed according to the ACMG recommendations. Class 3, 4, and 5 variants were validated by Sanger sequencing (3500 Genetic Analyzer, ThermoFisher Scientific, USA). Clustal Omega (1.2.4) and Jalview (2.11.0) were used to generate alignments between

multiple sequences and analyze the conservation of novel variants across species.

3 | RESULTS

3.1 | Patient profile

The patient is a 15-year-old girl, a second child of non-consanguineous parents, delivered by cesarean section in week 40 with an Apgar score of 9/10. She was suspected of OI in the second week of life after first postnatal fracture of right femur. Diagnosis was confirmed by Babygram examination and the presence of characteristic clinical features including a triangular shaped face and soft skull bones. She had sustained about 20 postnatal fractures of the upper and lower limbs, the most recent fracture occurred at the age of 13. Weight and body height are both below the third centile. Patient presents skeletal deformations such as bowed and shortened lower limbs, a barrel-shaped chest and reduced height of the lower thoracic to lumbar *vertebrae* (Th11-L3; vertebral compression fractures). Further lower limb deformations were reduced by surgical interventions (2011—corrective osteotomies of both tibia bones; 2012—telescopic rods in both femurs) (Figure 1). The girl suffers from contractures of knee and elbow joints. Extra-skeletal features characteristic for OI, such as blue sclera, prominent forehead and Dentinogenesis imperfecta are observed. Densitometry results from 2016 revealed decreased bone mineral density, indicating osteoporosis (total body BMD 0.408 g/cm^2, *Z*-score −4.3; spine L1-L4 BMD 0.525 g/cm^2, Z-score −2.7). Treatment with bisphosphonates started at week 5 of life. Within 10 years, she had received 31 cycles of pamidronate. She currently uses a wheelchair or walks with additional support. The family history is negative, and the patient has a healthy brother. Based on the pathogenic character of the variant, we suspect that it is of de novo origin in the patient; however, no DNA sample was available from the parents.

3.2 | Characteristics of the novel variant

A novel single nucleotide variant, c.733G>T, that results in glycine-to-tryptophan substitution (p.Gly245Trp) was identified in exon 10 of the *COL1A1* gene (NM_000088.3) at position 17:48274558. The variant reached a coverage of 277 reads and 114 alternative reads. The variant meets the criteria of ACMG for a likely pathogenic variant, as supported by a considerable body of evidence. The variant was absent in control samples from population databases (Exome Sequencing Project, 1000 Genomes Project, Exome Aggregation Consortium), nor was it found in the gnomAD exomes or gnomAD genomes, reaching good coverage of the locus (31.0–94.4).

Identified glycine-to-tryptophan missense mutation represents a common mechanism of glycine substitutions to other amino acids characteristic for OI pathogenesis. For the *COL1A1* gene, a low rate of benign missense

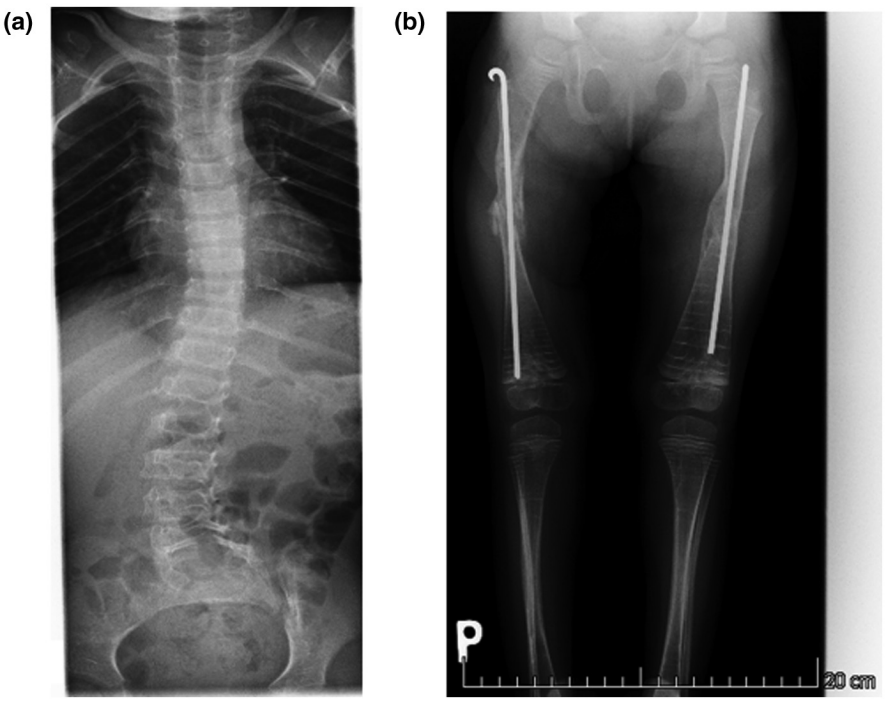

FIGURE 1 (a) X-ray of the spine AP projection (b) X-ray of lower limbs at the age of 4 years showing Fassier-Duval telescopic rods in femurs.

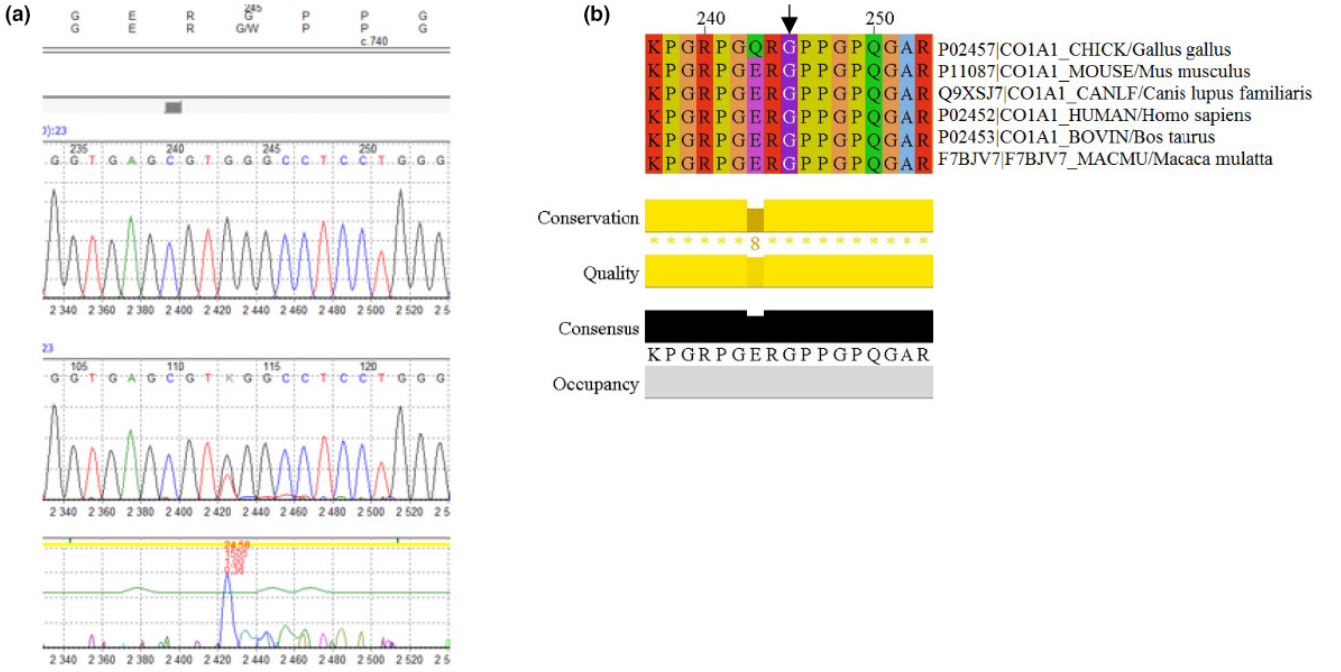

FIGURE 2 (a) Sanger sequencing chromatogram of the identified variant. (b) Multiple sequence alignment analysis of the novel variant across several species in Jalview 2.11.0, Clustal omega 1.2.4. The conserved columns with the highest score of 11 are indicated by "*".

mutations has been reported: 411 out of 451 non-VUS missense variants, and 772 out of 1217 clinically-reported variants are pathogenic. Multiple lines of computational evidence indicate a deleterious effect on the gene or gene product: a pathogenic verdict was obtained, based on nine pathogenic predictions compared to no benign predictions and one uncertain prediction. The variant was submitted to ClinVar by the Polish Mother's Memorial Hospital Research Institute (accession number SCV000994669.1) and added to the dbSNP database (rs1598299275).

Multiple Sequence Alignment analysis found the novel variant site to have the highest conservation score across several species (Figure 2). Neither the c.733G>T (p.Gly245Trp) mutation, nor any variant at c.733 or p.245 position in the *COL1A1*, has been previously reported in any human mutation database including the dedicated OIVD. No 3D structural model for molecular modeling is available in the Protein Data Bank. The possibility that the performed NGS procedure demonstrated technical bias was excluded, as none other of the 47 patients with a diagnosis of OI sequenced simultaneously with the proband had the c.733G>T variant.

4 | DISCUSSION

The substitution of glycine by other amino acid is well known to play a role in the pathogenesis of OI. Over the years, several models have been developed for relating missense mutations to particular OI type. The substituting reside, its position in the chain and the gene in which mutation occurs are all suspected to influence the final clinical outcome (Bodian et al., 2008, 2009; Marini et al., 2007). Genotype-to-phenotype correlation studies have shown that the type of amino acid substituting for glycine remains an important variable, as recurrence at a single glycine locus by different residues has been found to often result in distinct OI types (Bodian et al., 2009).

The mortality associated with missense mutations ranges from 0% to 64%, depending on the type of substituent amino acid, and this value differs between *COL1A1* and *COL1A2* (Sałacińska et al., 2021). This phenomenon is possibly related to the spacial constraints of the triple helix and the identity of the residue replacing glycine, which in turn depends on properties of its side chain: amino acids with a branched nonpolar (Val) or charged side chain (Arg, Asp, Glu) are associated with a more severe phenotype than small polar (Cys, Ser) or nonpolar (Ala) residues (Basel & Steiner, 2009; Bodian et al., 2008, 2009). This indicates that larger amino acids cause greater disruption of the triple helix and hence, a more severe phenotype (Basel & Steiner, 2009).

The presented mutation is the first record of a glycine-to-tryptophan substitution in the *COL1A1* gene. The variant, p.Gly245Trp, was identified in a patient presenting with a progressively-deforming OI phenotype. As tryptophan belongs to a group of amino acids with a branched

hydrophobic side chain, similarly to valine, the severe phenotype resulting from this novel mutation is consistent with the proposed correlation between residue identity on collagen structure and clinical outcome (Forlino & Marini, 2016). Although the lethality of amino acids varies significantly between collagen type I genes, the three glycine-to-tryptophan substitution reported in *COL1A2* resulted in progressively-deforming Type III, similarly to the novel variant reported herein (Bodian et al., 2009; Dalgleish & Kheng Teh, 2020; Hrušková et al., 2015; Nuytinck et al., 2000).

However, several unsolved problems still remain concerning the genotype–phenotype correlation of OI. The unclear influence of the type of substituent amino acid on the clinical manifestation, the impact of modifying factors, and the limited success of prediction models, all highlight the complexity of the OI pathomechanism (Bodian et al., 2009; Forlino et al., 2011). Although it is often hard to predict the phenotype associated with a given substitution, healthcare professionals can still make a rough prognosis based on certain principles and previously-reported cases. Therefore, our finding provides valuable information on the clinical outcome of glycine-to-tryptophan substitutions in the *COL1A1* gene. By broadening our knowledge of the spectrum of mutations underlying OI, we move toward a fuller understanding of the genotype–phenotype correlation of this heterogeneous disease.

AUTHOR CONTRIBUTIONS

Agnieszka Gach and Kinga Sałacińska: conceptualization. Kinga Sałacińska: formal analysis, writing-original draft preparation, visualization, and project. administration. Kinga Sałacińska, Iwona Pinkier, Lena Rutkowska, and Dominik Salachna: investigation. Izabela Michałus, Danuta Chlebna-Sokół, Elżbieta Jakubowska-Pietkiewicz, Łukasz Kępczyński: resources. Agnieszka Gach: writing-review, editing and supervision. Kinga Sałacińska and Agnieszka Gach: funding acquisition. All authors agreed to be accountable for the content of the work.

ACKNOWLEDGMENTS

We thank the patient for participation in this study.

CONFLICT OF INTEREST

The authors declare that there are no competing interests associated with the manuscript.

ETHICS STATEMENT

The study was approved by the local Ethics Committee (No. 108/2015). The legal guardian gave informed consent for the patient to take part in the study.

ORCID

Kinga Sałacińska https://orcid.org/0000-0002-1802-8795

REFERENCES

Basel, D., & Steiner, R. D. (2009). Osteogenesis imperfecta: Recent findings shed new light on this once well-understood condition. *Genetics in Medicine*, *11*(6), 375–385. https://doi.org/10.1097/GIM.0b013e3181a1ff7b

Bodian, D. L., Madhan, B., Brodsky, B., & Klein, T. E. (2008). Predicting the clinical lethality of Osteogenesis imperfecta from collagen glycine mutations. *Biochemistry*, *47*(19), 5424–5432. https://doi.org/10.1021/bi800026k

Bodian, D. L., Chan, T. F., Poon, A., Schwarze, U., Yang, K., Byers, P. H., Kwok, P. Y., & Klein, T. E. (2009). Mutation and polymorphism spectrum in Osteogenesis imperfecta type II: Implications for genotype-phenotype relationships. *Human Molecular Genetics*, *18*(3), 463–471. https://doi.org/10.1093/hmg/ddn374

Dalgleish III, R. (2014). Clinical and molecular genetics of Osteogenesis imperfecta, Chapter 10: Osteogenesis imperfecta genotypes and genotype–phenotype relationships. In J. Shapiro, P. H. Byers, F. H. Glorieux, & P. D. Sponseller (Eds.), *Osteogenesis imperfecta—A translational approach to brittle bone disease* (pp. 103–110). Elsevier Inc. https://doi.org/10.1016/B978-0-12-397165-4.00010-1

Dalgleish, R., & Kheng Teh, W. (2020). Collagen type I alpha 1 chain (COL1A1); collagen type I alpha 2 chain (COL1A2). In *Global variome shared LOVD*. Leiden Open Variation Database. LOVD v.3.0 Build 26c. https://databases.lovd.nl/shared/genes/COL1A1, https://databases.lovd.nl/shared/genes/COL1A2

Forlino, A., Cabral, W. A., Barnes, A. M., & Marini, J. C. (2011). New perspectives on Osteogenesis imperfecta. *Nature Reviews Endocrinology*, *7*(9), 540–557. https://doi.org/10.1038/nrendo.2011.81

Forlino, A., & Marini, J. C. (2016). Osteogenesis imperfecta. *Lancet*, *387*(10028), 1657–1671. https://doi.org/10.1016/S0140-6736(15)00728-X

Hrušková, L., Mařík, I., Mazurová, S., Martásek, P., & Mazura, I. (2015). COL1A2 gene analysis in a Czech Osteogenesis imperfecta patient: A candidate novel mutation in a patient affected by Osteogenesis imperfecta type 3. *Advances in Genomics and Genetics*, *5*, 275–281. https://doi.org/10.2147/AGG.S58766

Marini, J. C., Forlino, A., Cabral, W. A., Barnes, A. M., San Antonio, J. D., Milgrom, S., Hyland, J. C., Körkkö, J., Prockop, D. J., De Paepe, A., Coucke, P., Symoens, S., Glorieux, F. H., Roughley, P. J., Lund, A. M., Kuurila-Svahn, K., Hartikka, H., Cohn, D. H., Krakow, D., … Byers, P. H. (2007). Consortium for Osteogenesis imperfecta mutations in the helical domain of type I collagen: Regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. *Human Mutation*, *28*(3), 209–221. https://doi.org/10.1002/humu.20429

Nuytinck, L., Tükel, T., Kayserili, H., Apak, M. Y., & De Paepe, A. (2000). Glycine to tryptophan substitution in type I collagen in a patient with OI type III: A unique collagen mutation. *Journal of Medical Genetics*, *37*(5), 371–375. https://doi.org/10.1136/jmg.37.5.371

Sałacińska, K., Pinkier, I., Rutkowska, L., Chlebna-Sokół, D., Jakubowska-Pietkiewicz, E., Michałus, I., Kępczyński, Ł., Salachna, D., Jamsheer, A., Bukowska-Olech, E., Jaszczuk, I., Jakubowski, L., & Gach, A. (2021). Novel mutations within collagen Alpha1(I) and Alpha2(I) ligand-binding sites, broadening the spectrum of Osteogenesis imperfecta—Current insights into collagen type I lethal regions. *Frontiers in Genetics*, *12*, 692978. https://doi.org/10.3389/fgene.2021.692978

Steiner, R. D., & Basel, D. (2005). COL1A1/2 Osteogenesis imperfecta. In M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. W. Gripp, G. M. Mirzaa, & A. Amemiya (Eds.), *GeneReviews®* (pp. 1993–2022). University of Washington.

How to cite this article: Sałacińska, K., Michałus, I., Pinkier, I., Rutkowska, L., Chlebna-Sokół, D., Jakubowska-Pietkiewicz, E., Kępczyński, Ł., Salachna, D., & Gach, A. (2022). The first glycine-to-tryptophan substitution in the *COL1A1* gene identified in a patient with progressively-deforming Osteogenesis imperfecta. *Molecular Genetics & Genomic Medicine*, *10*, e1996. https://doi.org/10.1002/mgg3.1996