Abstract

**Background:** Osteogenesis imperfecta is an inherited disorder particularly of a human connective tissue. It is a worldwide extensive disorder regardless of age, gender or ethnic group. At present the disease includes nine clinically different types. Typical clinical features are brittle bones, high frequency of fractures and bone deformities. The other observed signs are blue sclera, dentinogenesis imperfect and otosclerosis. The first four types of the disease arise from mutations in collagen type I genes, composed from COL1A1 and COL1A2 chains. A result of these mutations is the production of shortened or structurally defective protein. Individuals affected by OI forms V to IX have mutations in proteins encoded by following genes: CRTAP, LEPRE1, PPIB, FKBP10. Collagenous types of the illness exhibit a broad range of severity depending on type and mutation localization in the structure of the collagen type I.

**Objectives and Methods:** The aim of this study is the description of the clinical forms of the disease, identifying mutations and polymorphisms of genes of the collagen type I by a molecular genetic analysis of genomic DNA of Czech OI patients.

**Results:** Currently in the Czech population there are described mutations and polymorphisms only of MLBR2 region, namely exons 31, 33 and 36 and introns 32 and 39, of the COL1A1 gene of 25 OI patients.

**Conclusion:** It is important to perform a further molecular genetic analysis of both collagen type I genes for the detection of the widest possible mutational spectrum for determination of possible genotype-phenotype relationship of affected individuals.

**Keywords**

Osteogenesis imperfecta, collagen type I, COL1A1, COL1A2, MLBR, mutations

---

1 Introduction

Osteogenesis imperfecta, type I-IV, is an inherited disorder of the connective tissue formation, especially of bones, joints and skin. Clinical features of this disease are bone fragility, high frequency of fractures, bone deformity, joint hypermobility, subnormal or short stature, dentinogenesis imperfecta (DI), bluish/greyish hue of sclera, hearing lose in adulthood, vascular, neurological and pulmonary complications and some other [30]. Presentation of these characters vary according to the type of a disease, but also within the same type of an illness. The incidence of non-lethal forms of the disease, so called OI tarda involving types I, III and IV, is reported in the range of 1: 25 000 to 1: 40 000 live births. The incidence of lethal type of this disorder - OI type II (known as OI congenita),
Based on radiographic features we distinguish type IIA or grey sclera and extremly large and soft cranium [13]. Significantly short. Infants have a triangular face, blue sclera which occurs already in the perinatal period. The extremities are severely deformed, multiple fractures of stillborns and perinatal mortality (up to 80% infants). Their bones are severely deformed, multiple fractures is very common for this type [18]. Another usual sign is blue sclera (intensity does not change with age) [30]. Affected individuals can have Otosclerosis occurs in some cases. Based on presence of DI, OI type IV is divided into two types: type IVA (absence of DI) and IVB (presence of DI). OI type IIB results from CRTAP gene mutations [16]. OI type III is the most severe form of OI. The first fractures occur in uterus and at birth. Patients have sub-normal stature with short extremities compared to the body, deformed short and long bones. Other distinctive signs include a triangular face, DI, blue sclera which usually turn to white with age, barrel shaped chest, weak muscles and severe scoliosis [16]. Radiographic findings of infants demonstrate undermineralized calvarium with Wormian bones, of adult detect osteopenia and popcorn-like calcification, especially metaphyseal and epiphyseal. This calcification disrupts the growth plate and reduces growth of long bones, especially femurs. Metaphyses of long bones are broad, diaphysis are thin. Osteopenia and joint hypermobility often lead to kyphoscoliosis. Basilar impression can occur in some cases. The patients require a wheelchair and crutches. Genetic transmission is autosomal dominant. This OI type is caused by mutations in collagen type I genes [13].

2.4 OI Type IV

OI type IV is the highly heterogeneous form of this disease. There is considerable intra- and interfamilial heterogeneity. Individuals can be mildly to severely affected. Their stature is variable short. The first fractures can occur at birth. Bones are mildly to severely deformed with popcorn-like calcification that is less common as in the type III. Patients have white sclera, although bluish and grayish color is also described.

Osteogenesis Imperfecta Classification

Based on clinical signs the first OI classification from David Sillence (created in 1979) distinguished four types of the disease (I-IV). In the past, related to the development methods of analysis, such as molecular-genetic techniques and histological findings, new forms were identified in the IV group of OI - OI type V-IX [13]. The disease exhibits a wide spectrum of clinical and radiological signs and varies in severity from mild to perinatal lethal forms. Inheritance is mainly autosomal dominant (AD), but there are some types with the autosomal recessive (AR) type of the inheritance [16].

2.1 OI Type I

This autosomal dominant type is the mildest form of OI. Patients do not have deformed bones and achieve normal or smaller growth. Fracture frequency is constant during childhood, decreases after puberty and then increases after menopause of women and after the sixth decade of men. Mild scoliosis resulting from vertebral fractures is very common for this type [18]. Another usual but not characteristic sign is blue sclera (intensity does not change with age) [30]. Affected individuals can have dentinogenesis imperfecta, mild joint hypermobility, tendency to bruising and suffer from partial or total hearing loss [20]. Based on the presence of DI we distinguish OI type IA (absence DI) and OI type IB (presence DI). Patients diagnosed with the type IB may have mild bowing of long bones of limbs. OI type I is a result of COL1A1 or COL1A2 genes mutations [13] [15].

2.2 OI Type II

OI type II is the lethal type of OI with high frequency of stillbirths and perinatal mortality (up to 80% infants die during the first week of life). Survival of the perinatal period is rare [1]. These individuals often die of a lung failure. Their bones are severely deformed, multiple fractures occur already in the perinatal period. The extremities are significantly short. Infants have a triangular face, blue or grey sclera and extremely large and soft cranium [15]. Based on radiographic features we distinguish type IIA (characterized by short and deformed long bones of lower limbs, short, deformed and continuously expanded ribs, dark blue sclera, and macrocephaly), IIB (which is similar to the type A, but individuals have small head circumference, shallow orbit and white or bluish sclerae) and IIC (this type differs by the presence of deformities and low bone density especially of ribs and long bones of limbs) [1] [20]. OI types IIA and IIC are inherited by the AD manner and are caused by COL1A1 or COL1A2 genes changes [11]. OI type IIB results from CRTAP gene mutations. It is the autosomal recessive form of OI type II [1].
unknown. All of these types (V-IX) are inherited in an autosomal recessive manner [13].

3 Collagen Type I

Type I collagen is the most abundant protein of extracellular matrix in connective tissue, primarily in bones. It is a heterotrimer which is composed of two alpha1(I) chains, encoded by COL1A1 gene which is located at 17q21.3-q22 position of chromosome 17, and of one alpha2(I) chain that is encoded by COL1A2 gene on chromosome 7 at locus 7q21.3-q22.1. COL1A1 gene is consisted of 51 exons, the coding sequence of COL1A2 consists of 52 exons. Even so, genetic information of the alpha chains is the same size in both of these two genes because the amino acids 568-603 encode exon 33 of COL1A1 gene, in the COL1A2 gene there are the same amino acid residues encoded by exons 33 and 34 [7]. Structurally we distinguish three areas of collagen genes: promoter, the 5 part of the gene which includes the signal sequence for binding the RNA polymerase and which contains binding sites of transcription factors, coding sequence carrying the genetic information of the alpha chain, and terminator, the 3domain where the polyT sequence and the termination codons, the termination signals for DNA transcription, are situated (Fig. 1).

The first form of alpha chain, prepro-alpha chain, is produced by fibroblasts, osteoblasts or odontoblasts [9]. There are three domains within the structure of prepro-alpha chain: N-terminal propetide, encoded by exons 1-5 and part of exon 6, helical domain encompassing exons 6-49 and C terminal propeptide encoded by exons 50 and 51 and part of exon 49 [7]. The N-propeptide structure further consists of signal peptide, von Willebrand factor binding site and Col 2 binding site of cell-specific proteins (Fig. 2). Pro-alpha chains arise in the endoplasmic reticulum by splitting the signal peptide and then they are assembled into procollagen molecules. The folding process proceeds from the C– to the N–terminus [5]. The subsequent extracellular N– and C–propeptides separation (by N– and C–peptidase) creates the final molecule collagen type I, which is subject to other posttranslational modifications such as glycation or hydroxylation of amino acid residues [21] [23]. The final collagen monomer is terminated by N– and C–telopeptides [7].

The triple-helical region of alpha chains is composed of 338 repeating Gly-X-Y sequences, in which Gly is glycine, X is frequently proline and Y is often hydroxyproline. It follows that the amino acid glycine is crucial for correct folding of collagen monomers due to inter-chains links production [28]. The main function of the proline and hydroxyproline is to stabilize the elongated nature of the alpha chains and to increase denaturation temperature of the protein [3] [24].

Collagen monomers are assembled into the collagen microfibrils and these into collagen fibrils. The basic repeat structure of collagen fibrils is called D-period. This period contains whole sequence of the monomer. It is 67nm long and is composed of one overlap and one gap zone (Fig. 3) [8].

There are several crucial areas of collagen type I genes promoters that influence DNA transcription. We rank to them transcription factors binding sites, activating proteins binding sites (such as YY1, c-Krox, IF1, IF2, AP1 and so on) stimulating or suppressing the transcription and CpG rich sequences of promoter (and of exon 1 and intron 1) whose methylation is prevented binding of transcription factors [9].

The first form of alpha chain, prepro-alpha chain, is produced by fibroblasts, osteoblasts or odontoblasts [9]. There are three domains within the structure of prepro-alpha chain: N-terminal propetide, encoded by exons 1-5 and part of exon 6, helical domain encompassing exons 6-49 and C terminal propeptide encoded by exons 50 and 51 and part of exon 49 [7]. The N-propeptide structure further consists of signal peptide, von Willebrand factor binding site and Col 2 binding site of cell-specific proteins (Fig. 2). Pro-alpha chains arise in the endoplasmic reticulum by splitting the signal peptide and then they are assembled into procollagen molecules. The folding process proceeds from the C– to the N–terminus [5]. The subsequent extracellular N– and C–propeptides separation (by N– and C–peptidase) creates the final molecule collagen type I, which is subject to other posttranslational modifications such as glycation or hydroxylation of amino acid residues [21] [23]. The final collagen monomer is terminated by N– and C–telopeptides [7].

The triple-helical region of alpha chains is composed of 338 repeating Gly-X-Y sequences, in which Gly is glycine, X is frequently proline and Y is often hydroxyproline. It follows that the amino acid glycine is crucial for correct folding of collagen monomers due to inter-chains links production [28]. The main function of the proline and hydroxyproline is to stabilize the elongated nature of the alpha chains and to increase denaturation temperature of the protein [3] [24].

Collagen monomers are assembled into the collagen microfibrils and these into collagen fibrils. The basic repeat structure of collagen fibrils is called D-period. This period contains whole sequence of the monomer. It is 67nm long and is composed of one overlap and one gap zone (Fig. 3) [8].

There are several crucial areas of collagen type I genes promoters that influence DNA transcription. We rank to them transcription factors binding sites, activating proteins binding sites (such as YY1, c-Krox, IF1, IF2, AP1 and so on) stimulating or suppressing the transcription and CpG rich sequences of promoter (and of exon 1 and intron 1) whose methylation is prevented binding of transcription factors [9].

The first form of alpha chain, prepro-alpha chain, is produced by fibroblasts, osteoblasts or odontoblasts [9]. There are three domains within the structure of prepro-alpha chain: N-terminal propetide, encoded by exons 1-5 and part of exon 6, helical domain encompassing exons 6-49 and C terminal propeptide encoded by exons 50 and 51 and part of exon 49 [7]. The N-propeptide structure further consists of signal peptide, von Willebrand factor binding site and Col 2 binding site of cell-specific proteins (Fig. 2). Pro-alpha chains arise in the endoplasmic reticulum by splitting the signal peptide and then they are assembled into procollagen molecules. The folding process proceeds from the C– to the N–terminus [5]. The subsequent extracellular N– and C–propeptides separation (by N– and C–peptidase) creates the final molecule collagen type I, which is subject to other posttranslational modifications such as glycation or hydroxylation of amino acid residues [21] [23]. The final collagen monomer is terminated by N– and C–telopeptides [7].

The triple-helical region of alpha chains is composed of 338 repeating Gly-X-Y sequences, in which Gly is glycine, X is frequently proline and Y is often hydroxyproline. It follows that the amino acid glycine is crucial for correct folding of collagen monomers due to inter-chains links production [28]. The main function of the proline and hydroxyproline is to stabilize the elongated nature of the alpha chains and to increase denaturation temperature of the protein [3] [24].

Collagen monomers are assembled into the collagen microfibrils and these into collagen fibrils. The basic repeat structure of collagen fibrils is called D-period. This period contains whole sequence of the monomer. It is 67nm long and is composed of one overlap and one gap zone (Fig. 3) [8].

3.1 Multi Ligand Binding Regions (MLBR) of the Collagen Type I Protein

There are several ligand binding sites situated at the level of collagen monomer. There are three hot spots on the alpha1 and alpha2 chains defined by codons 81200
(MLBR1), 682830 (MLBR2) and 9211040 (MLBR3) [26]. These sequences bind integrines that bind extracellular matrix molecules [25], keratan sulphate proteoglycans and dermatan sulphate proteoglycans, regulators of the fibrogenesis and formation of the inter-fibrils interactions and protectors of the fibrils against proteolytic damage [25].

COMP (Cartilage oligomeric matrix protein), fibronectin ensuring collagen type I molecules adhesion [26]. Hsp47 (heat-shock protein 47) that serves to thermal stabilization of the helix during procollagen synthesis, helps to folding and assembly of procollagen molecules and participates in the transport of structurally unaffected molecules from the endoplasmatic reticulum [26], belong also to ligands of collagen of the type I. Finally, the other important extracellular matrix proteins interacting with the collagen type I molecules are phosphophoryne inducing dentin matrix mineralization [8], osteonectin that binds extracellular matrix proteins, regulates production and storage of some extracellular matrix molecules or inhibits cell cycle [2, 11]. von Willebrand factor, protein affecting platelet function [22], and some more.

Generally, binding of extracellular matrix molecules to molecules of collagen of the type I increases strenght and elasticity of bone tissue [26].

4 Molecular Basis and Genotype-Phenotype Correlation of Osteogenesis Imperfecta, Type I-IV

Types I-IV of osteogenesis imperfecta are the result of collagen type I genes (COLIA1 and COLIA2) mutations. Essentially, mutations of these two genes manifest in two ways: 1) synthesis of a decreased number of alpha chains, 2) production of structurally defective protein.

Production of a decreased amount of collagen fibrils is associated with nondeforming OI type I. This type of OI disease results from null mutations, a single nucleotide substitutions, that lead to the STOP codons formation (presence of STOP codons terminates DNA transcription). Decreased expression of the protein also may result from genetic changes in splice-sites of pre-mRNA, provided that these one result in intron retention in mRNA or in STOP codon formation [16].

Deforming types of osteogenesis imperfecta, types II, III and IV, arise based on mutations affecting the structure of collagen. These changes are in 80% missense mutation (glycine substitution), the remaining 20% are frameshift mutations including deletion/insertion of one and more nucleotides (number not divisible by three) and splice-site mutations resulting in exon skipping or in new splice-sites production [13].

In terms of genotype-phenotype relationship, several links were defined. Severity of the disorder increases from N to C terminus of alpha chains. This relationship applies to MLBR regions. Specifically, MLBR 1 mutations result in mild to severe forms of OI while clinical pictures of MLBR 2 and 3 changes are primarily OI II and III types. Furthermore, there are eight lethal regions of the alpha2 chain (all of them are located in proteoglycans binding sites of the collagen fibrils) [26]. Regardless, lethal mutations are located rather in the alpha1 chain (about 35.6% glycine substitutions cause lethal OI) than in the alpha2 chain (only 19% are lethal) [13].

5 Current Knowledge of Osteogenesis Imperfecta Treatment

Treatment of patients diagnosed by OI includes using medicaments, surgery, orthotic treatment and rehabilitation. Recently, in the medical treatment the most widely used medicament are bisphosphonates. Effect of the bisphosphonates is reduction of the bone turnover with subsequent increase of the bone density but not the improving of the structure of the collagen type I molecules [13].

Surgical treatment is mainly performed to correct deformities and to reduce the bone brittleness as the result of bad bowing and to improve the physical condition of the individual. It includes osteotomy, intramedullary fixation due nails, wires, pins and other or spinal fusion with Harrington rod [20, 29]. Surgical intervention is also one of the possible solutions of otosclerosis in which the patients undergo stapedectomy (surgical removal of the stapes). The non-invasive therapy involves the use of orthotic trunk and limbs orthoses to correct scoliosis and mild limb deformations, such as genua valga or vara. Individuals suffering from hearing loss use a cochlear implant to improve its hearing. Patients are also advised to perform light physical activity (swimming, walking in water, Nordic walking) to strengthen the weakened muscles [17].

Recently gene and cell therapies are a current issue in the treatment of this disease. The essence of cell therapy is the transplantation of bone marrow matched donors. Normal osteoblasts formed by marrow donors have the ability to replace the mutant osteoblasts. OI patients who have undergone a cell therapy show an increase in a bone mineral content and an increase in body height. The aim of a gene therapy is to prevent expression of mutated alleles. This is achieved by binding of complementary antisense DNA/RNA fragments or hammerhead ribozymes to abnormal pre-mRNA. The intention is to prevent translation of this pre-mRNA and its subsequent degradation. A gene therapy using these mechanisms results in the conversion of severe types of OI in mild forms of the disease. Another approach in a gene therapy can be modification of mesenchymal stem cells of OI patients in vitro and consequent returning of these cells to the individuals. Using a gene therapy to treat OI currently has one obstacle. That is the small number of known mutations of genes of collagen type I. Thus, the creation of fragments, the use of
rRNA, or modification of stem cells of affected individuals is currently very difficult because of the high mutation spectrum of OI [3][4].

6 Conclusion

Osteogenesis imperfecta is a heterogeneous disorder with a wide spectrum of clinical characters and a large genetic diversity. Determination of genotype-phenotype relationship is a permanent problem because the same mutation may present different phenotypes among unrelated individuals also in members of one family with the same form of this disorder. The reason is the wide spectrum of clinical signs of identical mutations. Currently, 10% of all mutations that alter the glycine codon is described. In the future it is important to detect a lot of collagen of the type I mutations using the molecular genetic analysis to determination of the genotype-phenotype correlation in patients diagnosed by OI types I-IV. For this purpose the method of laser microdissection of affected tissue can be also used. This method can detect a specific mutations affecting bone formation.

The analysis of genes of the collagen type I should be aimed primarily at multi ligand binding regions (MLBR1-3), because changes in these sequences may prevent the creation of intra- or extramolecular bonds important for the quality of a bone structure, and at COL1A1 and COL1A2 regions of participating genes, important for initiation and process of the transcription.

Recently, the molecular genetic analysis (comprising polymerase chain reaction (PCR) and double-sided sequencing) focused on the part of the COL1A1 gene, containing MLBR2 region, in 25 Czech patients affected by OI type I-IV had carried out. We observed mutations in DNA samples of seven Czech patients. Four of these patients are affected by OI type IA, one patient suffering from OI type III and two patients were diagnosed with OI type IVB. All determined changes in our sample collection are single nucleotide mutations that result in either amino acid substitution, STOP codon production or the mutations do not alter the reading frame. Mutations of coding sequence were observed in exons 31, 33 and 36. Of these mutations only Gly523Cys, Gly526Cys and Arg519STOP were described in literature. Two non-coding sequence modifications were observed in introns 32 and 39. Both of these two intronic mutations were detected in two patients affected with OI type IA and in one patient diagnosed with OI type IVB.

Currently, we collect a biological material (venous blood and bone grafts) of Czech OI patients for the molecular genetic analysis of other important regions of the COL1A1 gene and for subsequent COL1A2 gene analysis. In the future it is important to perform the molecular genetic analysis of complete sequences of both collagen type I genes and subsequently to compare the clinical manifestations of the disease in patients having the same form of OI and the same change in DNA. This is crucial for the correct diagnosis of the type of the disease and for provision of timely treatment of affected patients to limitation potential health problems associated with osteogenesis imperfecta.

Acknowledgments

Acknowledgements belong to Doc. MUDr. I. Mářík, CSc. and to MUDr. Olga Hudáková, Ph.D. for the provision of useful informations regarding the clinical description of individual forms of the disease. The work was supported by the CBI project No. IM06014 of Ministry of Education, Youth and Sport CR and by the project SVV-2011-262514 of Charles University in Prague.

References


