Collagen and Collagen Disorders

Açelya YALOVAÇ *, N. Nuray ULUSU**°

**Collagen and Collagen Disorders**

**Summary**
Collagen has been studied extensively by a large number of research laboratories since the beginning of the 20th century. Collagens are the major fibrous glycoproteins present in the extracellular matrix and in connective tissue such as tendons, cartilage, the organic matrix of bone, and the cornea, and they maintain the strength of these tissues. Collagen has a triple-helical structure. This molecule consists of repeating unusual amino acids 35% glycine, 11% alanine, 21% proline, and hydroxyproline. The importance of collagens was clearly defined after the inherited collagen disorder studies. Mutations that alter folding of the triple helix result in identifiable genetic disorders, such as: osteogenesis imperfecta, Ehlers-Danlos syndromes, Alport syndrome, Bethlem myopathy, some subtypes of epidermolysis bullosa, Knobloch syndrome, some osteoporoses, arterial aneurysms, osteoarthrosis, and intervertebral disc diseases. The collagen family will be one of the most important research topics in future years because of its many important functions.

**Key Words:** Collagen, osteogenesis imperfecta, Ehlers-Danlos syndrome, Alport syndrome, Bethlem myopathy

**INTRODUCTION**
Enzymopathies affect humanity and our health in a multitude of ways. These disorders may cause life-threatening pathophysiological problems, structuropathies and various morbidities and mortalities. Some of these enzymopathies are quite common, such as glucose-6-phosphate deficiency. This enzyme defect is the most common enzymopathy in humans (1). On the other hand, there are various other enzymopathies including 6-phosphogluconate dehydrogenase deficiency (2). Collagen gene mutations and collagen modifying enzyme defects are rarely encountered. In this review, we discuss some recent data about collagen diseases.

Collagens, which account for one-third of the body’s proteins, are the biggest protein family of connective tissue and the extracellular matrix. The collagen family has various physiological functions. This superfamily includes 28 different types. These groups can also be classified into subtypes according to their structural and functional properties (3). The primary structure of the collagen molecule was identified as (Gly-X-Y)(n). The collagen structure is frequently characterized by the Gly-Pro-Hyp tripeptide (4). Glycine makes up one-third of all the amino acid residues, and the repetitive structure consists of many amino acids (5). Although this structure can theoretically
include more than 400 different compositions, when they are analyzed, only a limited number of differences can be determined (4). The triple-helix structure is important for some cellular functions such as adhesion and extracellular matrix activation and for enzymatic functions such as hydroxylation of collagen’s Lys and Pro residues (6). Even though its catabolism by matrix metalloproteinases is not fully understood, it is accepted that the triple-helix structure is required for the catabolism of both the collagen molecule and the cell surface receptors of macrophages (6,7). Collagen is usually found in fibril form in the extracellular matrix. The fibril-forming collagens are 300 nm in length and 67 nm in diameter (8). Each collagen chain consists of approximately 1000 amino acid residues. Fibril-forming collagens are types I, II, III, V, and XI, and these types of collagens are most common in the body (9). Type I collagen is the most abundant protein in humans (10). The collagen biosynthesis requires eight specific post-translational enzymes (11).

The critical roles of collagens are identified with the diseases that are caused by the mutations on the genes coding them, among which are osteogenesis imperfecta, chondroplasias, some subtypes of Ehlers-Danlos syndrome, Alport syndrome, Bethlem myopathy, some subtypes of epidermolysis bullosa, Knobloch syndrome, some osteoporoses, arterial aneurysms, osteoarthrosis, and intervertebral disc diseases. In this review, we intend to present the most prevalent diseases caused by several mutations.

**Osteogenesis Imperfecta**

Osteogenesis imperfecta, also known as “brittle bone disease”, is the most common (estimated prevalence is 1/10000 and 1/20000) autosomal dominant disease of the connective tissue (12,13).

The disease is characterized by extremely fragile bones, reduced bone mass, blue sclera, dentinogenesis imperfecta, hearing loss, and scoliosis (14). The bone fragility in this disease is caused by reduced bone mass, degenerated organization of bone tissue and altered bone geometry in size and shape (15).

The first classification of osteogenesis imperfecta was made by Sillence in 1979, and it was divided into four groups (types I-IV) due to the defects in the collagen gene. However, today this disease is divided into seven groups with the discovery of types V-VII (14,16).

Most of the cases are related with the mutations in two genes that encode the proalpha1 or proalpha2 polypeptide chains of type I collagen. A small proportion of these diseases are from a mutation in a cartilage protein or the expression of 3-prolyl-hydroxylase (17). Type I collagen is known to be the main structural protein of bone and skin. However, type I collagen gene mutations cause a bone disease. It has been shown previously that mutant collagen expression is different in bone and in skin (13). Clinical manifestation of this disease ranges from normal to lethal (18). If the mutations affect the amount of type I collagen, clinical manifestation of the disease will result in mild forms. Mutations in COL1A1 or COL1A2 genes result in more severe and lethal forms. Glycine has a critical role in the formation of collagen triple-helix. The most common mutations of this disease are seen in the substitution of glycine with a bigger amino acid (13). The mutations seen in types V, VI and VII are recently defined types (15). Type VII osteogenesis imperfecta arises from the mutations in “cartilage-associated protein” gene (19).

Osteogenesis imperfecta has many phenotypic differences and can be seen in every age, and diagnosis of the disease is difficult (16). Prenatal diagnosis can be made using clinical, radiographic, biochemical or genetic methods; the remainder are diagnosed after birth (20). Today, only symptomatic treatment is available - nonsurgical treatment consists of physical therapy and rehabilitation; surgical treatment and drugs used for increasing bone strength are the other treatment options (21). Treatments with bisphosphonates have led to significant improvements in adolescents and increased patient quality of life (12,22). A potent inhibitor of bone resorption drug and growth hormone have also been used in recent years (21,23). Animal studies and laboratory studies consisting of somatic gene therapy that aims at the inactivation of the mutations are still being evaluated (24).

**Marfan Syndrome**

Marfan syndrome is an autosomal dominant, multisystemic disease of the connective tissue (25). It does not have a
particular gender, racial or geographic tendency (26). This syndrome affects approximately 1 in 5000 to 1 in 10,000 individuals (27). It results from the mutations in the FBN1 gene, which encodes the fibrillin-1 (28). More than 500 FBN1 mutations have been determined for this syndrome (29). For clinical diagnosis, 86% of the cases can be diagnosed clearly by using Ghent nosology, but for certain diagnosis, genetic test should be performed (30). Its characteristics are aortic dilatation that progresses with aortic valve failure, mitral valve prolapsus and failure, lens dislocation and myopia, long and slim body and long limbs due to excessive bone growth, arachnodactyly, chest deformations, defects in skeletal system, skin, nervous system and lungs, and sometimes scoliosis (31-33).

One of the reasons for sudden cardiac death in athletes is the rupture of the aortic root due to Marfan syndrome (34). Furthermore, it can lead to various surgical pathologies such as hernia, ileus and abdominal vein aneurysms. These pathologies can sometimes be life-threatening (35). Individuals with Marfan syndrome are often tall and agile and they can unwittingly join physical activities and sports (27). Symptoms related to muscle and skeleton are very common (30).

**Ehlers-Danlos Syndrome**

The estimated prevalence of Ehlers-Danlos syndrome ranges from 1/10,000 to 1/25,000 [36]. However, this disease can be seen on all continents and can affect all races and both genders (37). At least 10 subtypes are classified based upon genetic, biochemical and clinical features (38). The syndrome is characterized by fragile and tender skin, easy bruising and atrophic injury. In approximately half of the cases, mutations in the COL5A1 and COL5A2 genes encoding the alpha1 and alpha2 chains of type V are defined (39).

Type IV Ehlers-Danlos syndrome accounts for approximately 5 to 10% of cases, and this disease is caused by 19 different mutations in the COL3A1 gene coding for type III procollagen (36,40). Characteristic features of the disease are easy bruising, serious vein, uterus and digestion complications, acrogeria, and translucent skin with highly visible subcutaneous vessels. The vascular type can result in fatal bleedings (36,41). Type VI is a result of the homozygote and heterozygote mutations in the lysyl hydroxylase gene (38). The expression of the disorder is changeable in different tissues. In skin, a complete lack of hydroxylsine was observed, but in some tissues such as lung, tendon, kidney or bone it is less expressed (42). The skin of type VI patients are hyperelastic but bruise easily and heal difficultly and slowly (43). Type VII is related with the defect in the process of transformation of procollagen to collagen. Inadequate procollagen amino protease activity is seen (44). On the other hand, many types of this disorder have been described, some of which are rarely seen, such as types V, VII and X (45). Types I and III result from decreasing lysyl hydroxylase activity (36). Very little is known about the genetic or biochemical defects responsible for the other subtypes of Ehlers-Danlos syndrome, but they are still being researched (38,46).

**Alport Syndrome**

Alport syndrome is caused by deletion mutations in COL4A5 and COL4A6 genes or COL4A3 and COL4A4 genes encoding the alpha3(IV) and alpha4(IV) chains. Eighty percent of Alport syndrome cases are X-linked and its prevalence is 1/5000 (47,48). This syndrome affects some of the basal membranes and is characterized by renal failure, loss of hearing and lens abnormalities (49). The first pathological sign in this disease is the loss of type IV collagen network formed by alpha3, alpha4 and alpha5(IV) chains (47). The damage of collagen IV due to mutations causes dysfunction of bound epithelium and results in organ damage (50). Diagnosis of this disease is predominantly based on histological studies (47).

**Epidermolysis Bullosa**

Epidermolysis bullosa is a heterogeneous group of heritable disorders. It is characterized by blistering of the skin and mucous membranes as a result of a small injury (51,52). It has been demonstrated that 10 genes are responsible for the different forms of this disease (51).

Epidermolysis bullosa has been classified under three anatomic groups: epidermolytic, junctional, and dermolytic types. These groups can be divided into subtypes based on histological and clinical criteria (53). Genetic abnormalities have been examined in different subtypes: in dominant simplex type, mutations in the K5 or K14 genes cause...
basal cell damage and blistering, while in recessive simplex types, muscular dystrophy results from abnormal plectin and skin fragility syndrome from PKP1 mutations. The junctional form results from the laminin 5, type XVII collagen and alpha-6-beta-4 integrin gene mutations. The dystrophic type is related with mutations in the type VII gene (54). The cause of squamous cell carcinoma development in epidermolysis bullosa patients is still unknown (55).

Current treatment of the disease includes nursing of the skin and other organs, infection care for chronic wounds, nutritional support and surgical treatment if needed (56).

**Knobloch Syndrome**

Knobloch syndrome is an autosomal recessive disorder characterized by high myopia, vitreoretinal degeneration, occipital bone damage, and congenital encephalocele (57). Genetical analysis of families with Knobloch syndrome is heterogeneous (58). In these families, pathological mutations in the COL18A1 gene at 21q22.3 have been identified (57). These mutations cause the loss of one or all collagen COL18A1 isoforms or endostatin (58).

**Bethlem Myopathy**

Bethlem myopathy is an autosomal dominant inherited relatively mild disease that is characterized by muscle weakness and distal joint contractures. It is supposed that type VII collagen plays a role in connecting cells and extracellular matrix and that mutations in genes encoding collagen VI result in Bethlem myopathy (59,60). While it is reported that the first attack in this disease is seen in early childhood, two-thirds of patients are over 50 years of age because of the slow progression of the disease (61).

**CONCLUSION**

Collagen, the most abundant protein in mammals, is the main component of the extracellular matrix and as a result is the main component of connective tissue. It is unavoidable that the undesirable effects of molecular changes in the structure of this fibrous protein may affect many systems of the human body, from the central nervous system to the musculoskeletal and cardiovascular systems. Collagen disorders can affect different systems such as the cardiovascular or ocular system, making these syndromes very important, even though the collagen disorders are not so common in our country. In recent years, despite the increasing information obtained about the structure and synthesis of collagen, the genetic and molecular bases of the collagen disorders, which are still accepted as untreatable or incurable clinical syndromes, remain undetermined. In addition to having the opportunity to understand numerous molecular disorders, progressive biochemical tests, molecular findings and also the technological resources will facilitate an explanation of the systematic and physiopathologic processes. Molecular studies will help to explain the different syndromes and clinical entities that are classified in the same groups. In the future, it will be possible to treat these patients and improve the quality of their daily life,

**REFERENCES**


43. Rauma T, Kumpumäki S, Anderson R, Davidson BL,


