2. BACKGROUND

2.1. Elastic fibers

2.1.1. Composition, structure and distribution

The elastic properties of many tissues are owing to the presence of highly branched elastic fibers in the extracellular space. The amount, distribution and structural organization of the elastic fibers vary in different organs. They are numerous in tissues, which undergo periodic or frequent stretching, such as aorta, lung, and elastic ligaments (Pasquali-Ronchetti and Baccarani-Contri, 1997). In normal human skin, the elastic connective tissue proteins represent a relatively small fraction of the total dermal proteins, collagen being the predominant component (Grant and Prockop, 1972; Uitto and Lichtenstein, 1976). They are organized into distinct morphologies in different tissues: small, rope-like networks in lung, skin and ligament; thin concentric sheets in blood vessels; and large three-dimensional honeycomb structures in elastic cartilage (Pasquali-Ronchetti and Baccarani-Contrri, 1997).

Structural studies on the elastic fibers revealed a complex structure consisting of 2 distinct components (Ross, et al., 1977). The major component, which has an amorphous appearance, represents the elastin protein. The amorphous elastin is surrounded by distinct fibrillar structures, which are rich in acidic glycoproteins and are organized into 10 to 12 nm fibrils of beaded appearance (Cleary, et al., 1981; Cotta-Pereira, et al., 1976). This component represents the elastic fiber microfibrillar proteins called microfibrils. The relative concentrations of elastin and the microfibrillar component vary during embryonic development; the newly developed fibers in young embryos are composed exclusively of the microfibrillar component, but during development the proportion of elastin increases progressively and in a fully developed fiber more than 90% is elastin (Ross, et al., 1977). Microfibrils themselves are made up of at least five distinct proteins (Gibson, et al., 1989), including two forms of the 350 kDa glycoprotein, fibrillin (Sakai, et al., 1986; Zhang, et al., 1994) and two microfibril-associated glycoproteins (MAGPs), MAGP-1 (Gibson, et al., 1991) and MAGP-2 (Gibson, et al., 1996). Several other components, such as lysyl oxidase (Kagan, et al., 1986), proteoglycans (Baccarani-Contrri, et al., 1990), osteopontin (Pasquali-Ronchetti and Baccarani-Contrri, 1997), emilin (Bressan, et al., 1993), and fibulin-1 (Roark, et al., 1995) are also reportedly present in the elastic fiber.
2.1.2. **Biosynthesis of elastic fibers**

Elastic fiber formation requires the coordination of a number of important processes including the intracellular transcription and translation of tropoelastin, intracellular further processing of the protein, secretion of the protein into the extracellular space, delivery of the tropoelastin monomers to sites of elastogenesis, alignment of the monomer units with previously deposited tropoelastin through associating microfibrillar proteins, and lastly, the conversion to the insoluble elastin polymer through the cross-linking action of lysyl oxidase (Mithieux and Weiss, 2005).

2.1.2.1. **Tropoelastin synthesis**

*Structure of the elastin gene*

In the last decades, a number of investigators have characterized the structure of the gene coding for tropoelastin, the precursor of elastin, from several species including humans, cows, and rats (Bashir, *et al.*, 1989; Indik, *et al.*, 1987; Olliver, *et al.*, 1987; Pierce, *et al.*, 1990). These studies have revealed the presence of only a single elastin gene in the mammalian genome that, in most species, is composed of 36 exons distributed throughout approximately 40,000 bp of genomic DNA, has an intron:exon ratio of almost 19:1, and has an unusually high frequency of repetitive DNA sequences within many introns, particularly at the 3’-end. Some of these investigations have also demonstrated that hydrophobic and cross-linking domains of tropoelastin are encoded by separate, alternating exons, reflecting the protein structure.

Some variation exists amongst species with respect to the presence of exon homologues. The human gene, which is localized to chromosome 7 (Fazio, *et al.*, 1991), has 34 exons, compared with 36 in the bovine and rat elastin genes (Bashir, *et al.*, 1989; Yeh, *et al.*, 1989). Exons 34 and 35 of the latter species are absent from the human gene, while the human gene contains the unusual hydrophilic-encoding exon 26A, not described in any other species (Bashir, *et al.*, 1989; Indik, *et al.*, 1987).

*Tropoelastin mRNA expression*

The expression of tropoelastin mRNA and elastic fiber synthesis is highest in early development and occurs primarily within a limited period during development (Parks, *et al.*, 1988). The changes in elastin synthesis appear to be a consequence of a change in amount of
elastin mRNA, and a strong correlation exists between mRNA levels and tropoelastin synthesis (Burnett, et al., 1982; Pollock, et al., 1990; Sephel, et al., 1987). Many potential control regions exist in the elastin gene including untranslated regions, intronic regions and promoter regions. Negative regulatory elements have been found in the first intronic region in the bovine gene in a region that is highly similar to the human gene (Manohar and Anwar, 1994). The promoter region of the human gene has been analyzed and found to contain both up-regulatory and down-regulatory elements (Rosenbloom, et al., 1991). Up to eight different transcription start sites have been identified (Bashir, et al., 1989; Rosenbloom, et al., 1993; Rosenbloom, et al., 1991), demonstrating that the elastin gene is under a complex control mechanism.

Age-dependence of the human elastin promoter coupled to the chloramphenicol acetyltransferase reporter gene has been demonstrated in mice in vivo (Hsu-Wong, et al., 1994). In chick aorta cells, the decrease in elastin synthesis that occurs with age has been shown to result partly from the destabilization of mRNA (Johnson, et al., 1995). Growth factors and hormones such as transforming growth factor β1 (Kahari, et al., 1992; McGowan, et al., 1997), insulin-like growth factor I (Wolfe, et al., 1993), vitamin D (Pierce, et al., 1992) and interleukin-1β (Mauviel, et al., 1993) have all been shown to affect tropoelastin synthesis at either the promoter level or post-transcriptionally by affecting the stability of tropoelastin mRNA. In addition, there is evidence that tropoelastin may be under negative feedback autoregulation whereby accumulation of tropoelastin in the ECM space may inhibit the further production of tropoelastin mRNA (Foster and Curtiss, 1990).

**Alternative splicing of the elastin gene**

A significant variation in nucleotide sequence and size of both mRNA and cDNA of tropoelastin has been observed within a species (Baule and Foster, 1988; Fazio, et al., 1988a; Pierce, et al., 1990). These variable cDNAs were shown to be the result of alternative splicing of tropoelastin mRNA. Two types of alternative splicing have been demonstrated; a complete excision of an entire exon, as seen with human exons 22, 23 and 32 or the excision of a portion of an exon, as is the case for human exon 26A (Indik, et al., 1987). At least 11 human tropoelastin splice variants have been identified with some exons shown to be subject to alternative splicing: exons 22, 23, 24, 26A, 32 and 33 (Boyd, et al., 1991; Fazio, et al.,
Exon 23 is a cross-linking domain while exons 22, 24, 32 and 33 are hydrophobic domains (Vrhovski and Weiss, 1998).

Translation and secretion

The intracellular biosynthesis of elastin polypeptides follows the same basic principles as the formation of other mammalian proteins. Elastin is synthesized by smooth muscle cells, endothelial and microvascular cells, chondrocytes and fibroblasts, and the translation of the mRNA molecules coding for elastin polypeptides takes place in the rough endoplasmic reticulum (RER) of these cells (Uitto, 1979; Vrhovski and Weiss, 1998). Polypeptide chains are released into the lumen of the RER with the release of the signal peptide (Saunders and Grant, 1984). Tropoelastin is secreted to the plasma membrane via secretory vesicles (Saunders and Grant, 1985). Inhibition of secretion of tropoelastin experimentally accumulated in the RER and Golgi apparatus results in intracellular degradation of the accumulated protein by cysteine proteases (Davis and Mecham, 1996). Before secretion, intracellular tropoelastin is, therefore, likely to be chaperoned by a 67-kDa elastin binding protein (galactolectin), which prevents intracellular self-aggregation and premature enzymatic degradation of the protein (Hinek and Rabinovitch, 1994). This association lasts until the complex is excreted into the extracellular space where the companion chaperone interacts with galactosugars of the microfibrils, dramatically decreasing its own affinity for the tropoelastin molecule, which then be released locally. The 67 kDa chaperone is recyclable and is a constitutive part of the elastin cellular receptor (Debelle and Tamburro, 1999). The elastin-binding protein binds tropoelastin primarily at the VGVAPG peptide site of human tropoelastin (Mecham, et al., 1989).

2.1.2.2. Elastic fiber assembly

Elastogenesis is a highly complex process, which consists of all events leading to the construction of a functional elastin within the elastic fiber. It starts inside the cell by specifically targeting and transporting the tropoelastin molecule to sites of fiber formation on the cell surface (Debelle and Tamburro, 1999). Before any elastin is deposited, microfibrils are secreted into the extracellular space close to the cell surface. The microfibrillar component works as a scaffold for the deposition of elastin and serve to align the tropoelastin molecules into the correct orientation for subsequent cross-linking (Mecham, 1991). The
molecular basis of how tropoelastin binds microfibrils during elastic fiber formation remains unclear, and it is the concern of current research in this field.

The formation of a specific transglutamase crosslink between fibrillin-1 and tropoelastin may act to covalently stabilize the newly deposited tropoelastin on the microfibrils (Rock, et al., 2004). This in turn is facilitated by the binding of MAGP-1 with tropoelastin (Brown-Augsburger, et al., 1996; Clarke and Weiss, 2004). When this binding is missing, the fiber integrity is severely altered, as noted in lamb ductus arteriosus, where a truncated tropoelastin is present and does not form fibers (Hinek and Rabinovitch, 1993). Calcium-dependent multiple binding sites were observed between tropoelastin and MAGP-1. The interaction of the N-terminal part of the MAGP-1 with the highly conserved C-terminal end of tropoelastin was particularly emphasized.

The only two Cys residues confined to the C-terminus of tropoelastin have been demonstrated to form an intrachain disulfide bond and molecular modeling has predicted the formation of a hairpin loop giving rise to a positively charged pocket with the C-terminal RKRK sequence at the end (Brown, et al., 1992). It was speculated that this pocket could provide a non-covalent binding site for the highly acidic microfibrils. It has been confirmed experimentally that tropoelastin can indeed bind to the microfibrillar protein MAGP-1 (Bashir, et al., 1994; Brown-Augsburger, et al., 1994); the binding being localized to the C-terminus of tropoelastin and the N-terminal half of MAGP-1 (Brown-Augsburger, et al., 1996; Clarke and Weiss, 2004). Both proteins appear to require an intact secondary structure, which has been cited as evidence that binding is not simply ionic (Brown-Augsburger, et al., 1994). Transglutaminase-mediated cross-linking of MAGP-1 to tropoelastin might then covalently lock this association (Brown-Augsburger, et al., 1994). Blocking MAGP-1 or the C-terminus of tropoelastin with antibodies reduces elastin accumulation in the ECM (Brown-Augsburger, et al., 1994), indicating that the interaction between tropoelastin and MAGP-1 is an important step in elastic fiber formation.

Fibrillin-1, a large cysteine-rich multidomain glycoprotein, polymerizes in the extracellular space in a head-to-tail manner as parallel bundles of six to eight molecules to form microfibrils that provide a force-bearing structural framework for fibrillogenesis (Gibson, et al., 1996; Kielty, et al., 2002). Mutations in fibrillin-1 result in Marfan syndrome, a heritable disease associated with severe aortic, ocular, and skeletal defects due to defective elastic

In a study aimed at screening the entire fibrillin-1 molecule for sites of interaction with tropoelastin and the major microfibril-associated molecule MAGP-1, Rock et al. (Rock, et al., 2004) have recently discovered high affinity interactions between tropoelastin and two overlapping fibrillin-1 fragments encoded by central fibrillin-1 sequences (exons 18-25 and 24-30) indicating that this region of fibrillin-1 is the major tropoelastin binding site. A further interaction of moderate affinity between tropoelastin and a proline-rich region of fibrillin-1 fragment encoded by exons 9-17 has been identified. The principal MAGP-1 binding site is suggested to be on fibrillin-1 sequence encoded by exons 1-8. These authors proposed the following model of elastic fiber formation (fig. 2.1).

I. Microfibrils are first formed by linear and lateral fibrillin-1 interactions.

II. MAGP-1 then associates onto microfibril bead surfaces via an interaction with N-terminal fibrillin-1 domains.

III. Tropoelastin is then deposited on an interbead region adjacent to the beads through strong interactions with the fibrillin-1 central sequence and subsequently becomes cross-linked to fibrillin-1. Tropoelastin and MAGP-1 may then interact on microfibrils.

IV. Further deposition of tropoelastin to microfibril-bound tropoelastin and MAGP-1 followed by lysyl oxidase cross-linking

V. Formation of mature elastic fiber
2.1.2.3. Coacervation and cross-linking of tropoelastin

Cross-linking of tropoelastin molecules is preceded by the association and alignment of the molecules bringing together each cluster of four positively charged lysine side chains necessary for cross-linking. Coacervation is thought to be the molecular mechanism through which aligning and concentrating can occur (Urry, 1978). Inappropriate coacervation may be detrimental to fiber formation and it appears that many different molecules may influence the coacervation of tropoelastin. The interaction of glycosaminoglycans with tropoelastin (Fornieri, et al., 1987), the binding of lipids by tropoelastin (Vrhovski, et al., 1997), and overhydroxylation of Pro residues (Urry, et al., 1979) are suggested to reduce the ability of tropoelastin to coacervate. Cross-linking of tropoelastin is greatly reduced at low temperatures when tropoelastin is not coacervated (Narayanan, et al., 1978). It was suggested that at low temperature, water forms a clathrate-like structure around the hydrophobic regions of tropoelastin, keeping the protein unfolded (Mithieux and Weiss, 2005). With increasing temperature, the ordered clathrate water is disrupted, rendering the hydrophobic domains,
Background

such as the oligopeptide repetitive sequences GVGVP, GGVP, and GVGVAP, free to fold and interact with other hydrophobic segments.

The initial step of the cross-linking reaction, which renders tropoelastin insoluble, is an oxidative deamination of lysine residues by Cu$^{2+}$ dependent lysyl oxidase to produce allysine, also known as α-amino adipic δ-semialdehyde. All subsequent reactions are spontaneous and involve the condensation of closely positioned lysine and allysine residues to produce cross-links such as allysine aldol, lysinonorleucine, merodesmosine, and tetrafunctional cross-links unique to elastin, such as desmosine and isodesmosine (Noblesse, et al., 2004; Reiser, et al., 1992; Vrhovski and Weiss, 1998). Desmosine and isodesmosine are thought to result from two different pathways (fig. 2.2).

Figure 2.2: Structure and route of formation of elastin cross-links (Vrhovski and Weiss, 1998).

Other than understanding the chemical mechanism of cross-link formation, little is known about how tropoelastin monomers interact one another to form the functional mature elastin. This is largely due to the highly insoluble nature of elastin, making it difficult to analyze.
Two types of cross-linking domains exist in tropoelastin: those rich in alanine (KA) and those rich in proline (KP). Within the KA domains, lysine residues are typically found in clusters of two or three amino acids, separated by two or three alanine residues. Molecular modeling and structural analysis show these regions to be α-helical structure, which has the effect of positioning two lysine sidechains on the same side of the helix (Brown-Augsburger, et al., 1995; Gray, et al., 1973). This conformation has been suggested to be critical to the formation of desmosines. Modeling studies have shown that allysine aldol and dehydrolysinonorleucine will fit on the α-helix without distortion with one interesting restriction; allysine aldol can only be accommodated when the precursor lysines are separated by three alanines while dehydrolysinonorleucine can form between lysines separated by either two or three residues (Gray, et al., 1973). In human tropoelastin, the KA domains are encoded by exons 6, 15, 17, 19, 21, 23, 25, 27, 29, and 31. The KP domains are encoded by exons 4, 8, 10, and 12. In these KP domains, the lysine pairs are separated by one or more proline residues and are flanked by prolines and bulky hydrophobic amino acids (Brown-Augsburger, et al., 1995). Desmosines and isodesmosines have not been found in association with KP domains. This could be attributed to the presence of multiple proline residues, which impose steric constraints that would not allow the formation of α-helical structure. There is little information describing how many tropoelastin molecules could be involved in the crosslink formation. However, one elastin cross-linking domain has been identified that joins three tropoelastin domains identified as 10, 19, and 25 (Brown-Augsburger, et al., 1995). The only two KA domains that contain three lysine residues are encoded by exons 19 and 25. Two lysines from each exon form a desmosine crosslink in an antiparallel arrangement of the exons, while the remaining lysine on each chain forms a lysinonorleucine crosslink with two lysine residues present on the KP domain encoded by exon 10.

2.2. Elastic fibers in the skin

The skin is a complex organ composed of two main tissues: epidermis and dermis. The epidermis, a stratified and differentiated epithelium, is responsible for the barrier function of the skin. The dermis is a connective tissue in which cells are embedded in their own and the abundant ECM, which is mainly composed of collagen and many macromolecules such as elastin. Elastic fibers are responsible for skin elasticity (Gibson, et al., 1989), and collagen fibers for its tensile strength. The progressive disorganization of elastic fibers and their
replacement by collagen fibers are major manifestations of skin aging. In addition, burns or deep wounds result in scarring with loss of elasticity.

The dermal elastic system consists of a meshwork of loose and wide branched fibers, which are larger and more represented in the reticular dermis, and which split repeatedly, becoming smaller as they ascend into the middle papillary dermis, where they form an arcade projected towards the dermal-epidermal junction. In this region, elastic fibers lose their elastin content and become almost exclusively composed of microfibrillar glycoproteins (Pasquali-Ronchetti and Baccarani-Contrì, 1997).

2.2.1. Skin aging

Aged human skin appears loose and sagging with reduced resilience and elasticity compared to young skin. Aging of the skin is thought to consist of two processes taking place simultaneously. The first process is intrinsic, chronologic aging and similar perhaps to aging of other tissues. The second process is photoaging, an environmentally induced remodeling of the dermis that arises as a result of repeated exposure of skin to sunlight. It is always rather difficult to distinguish the real age-related modifications from those induced by exogenous factors, including ultraviolet light and minor traumatic injuries (Pasquali-Ronchetti and Baccarani-Contrì, 1997). In both cases, however, the age-related alterations affect all dermal components (Fornieri, et al., 1989). Varani et al. have shown decreased procollagen gene expression and increased expression of genes encoding several matrix metalloproteinases as the result of intrinsic aging and photoaging (Varani, et al., 1998).

Elastin is produced early in life, and then appears to be stable and to have a turnover, which approaches the life span (Shapiro, et al., 1991). In a transgenic mouse line, the activity of the elastin promoter was found to increase during postnatal development and, in skin, to reach a peak at 3 months of age, then to decrease (Hsu-Wong, et al., 1994). Therefore, the age-related modifications extensively described in the elastic fibers of all organs and tissues may be largely interpreted as result of progressive degradation of a protein polymer that has been produced early in life.

Scanning electron microscopy showed that, in humans, the elastic meshwork grows largely undistorted during postnatal growth, where fibers seem to enlarge according to the growth of
tissue. Later, in adults and old subjects, elastic fibers gradually become tortuous, frayed and porous (Imayama and Braverman, 1989). Moreover, it is well documented that solar radiation induces a series of progressive structural and biochemical changes in the elastic fibers of exposed skin (Engel, *et al.*, ; Uitto, 1986), such as overproduction and basophilic degeneration up to destruction of the elastin meshwork in dermis (Bouissou, *et al.*, 1988; Brien, 1985; Frances and Robert, 1984) as well as in superficial vessels (O'Brien and Regan, 1991).

The physiological changes that occur on elastic fibers during aging and development were precisely described by Pasquali-Ronchetti and Baccarani-Contri (Pasquali-Ronchetti and Baccarani-Contri, 1997). In fetal skin, the elastin appears around the seventh month of gestation as thin strands of amorphous material associated with bundles of about 10-20 wide microfibrils. At birth, the elastic fibers in the reticular dermis are very small, and comprise a core of amorphous elastin surrounded by a discontinuous coat of microfibrils. Up to 10 years, the elastic fiber diameter tends to grow to its final size by further deposition of amorphous elastin, whereas the number of surrounding microfibrils decreases. The amorphous elastin seems to represent the major component of the fiber, whereas microfibrils are rare. Starting of the fourth decade of life, electron-dense materials accumulate within skin elastic fibers in an age-dependent manner. In very old subjects, these materials seem to have disappeared, leaving behind holes, which give to the fiber a sieve appearance. A concomitant phenomenon is the appearance of isolated elastic fibers, which do not exhibit the typical amorphous ultrastructural organization but seem to consist of interwoven filaments (Robert, *et al.*, 1988).

A few studies designed to distinguish actinic damage from chronological aging of skin elastic fibers demonstrated that the alterations were similar, but were most pronounced in sun-exposed skin (Bouissou, *et al.*, 1988; Braverman and Fonferko, 1982). The age-associated alterations consisted of swelling of fibers, granular degradation of elastin, with the appearance of lacunae within the elastic fibers, filled by osmiophilic materials (Braverman and Fonferko, 1982). This latter phenomenon has been described as being mostly associated with actinic damage (Bouissou, *et al.*, 1988). Rather interestingly, in the dermis of sun-exposed skin, inflammatory cells (macrophages and mast cells) were often observed. Therefore, although this has not yet been proved, lysosomal enzymes or mast cell products might contribute to the severe elastin damage in sun-exposed skin. On the other hand, elastin seems to accumulate in the superficial photoaged skin; in fact, the morphologic changes in
sun-damaged skin have been shown to be associated with accumulation of material having the staining properties of elastin, known as solar elastosis (Gilchrest, 1989).

The cellular mechanisms leading to solar elastosis are not understood and, indeed, controversial findings concerning the synthesis of elastic fibers during solar elastosis have been reported. Several reports have demonstrated that elastic fibers deposited during solar elastosis consist of the same components as normal elastic fibers. In response to UVA and/or UVB radiation, keratinocytes secrete many mediators that could stimulate fibroblast synthetic activity, and some of them, e.g., tumor necrosis factor b, interleukin-1b, and interleukin-10, have been shown to increase the promoter activity of the elastin gene, steady-state mRNA levels, and elastin accumulation (Kahari, et al., 1992; Mauviel, et al., 1993; Reitamo, et al., 1994). Whereas Bernstein et al (Bernstein, et al., 1994) have noted increased elastin mRNA levels in sun-damaged skin, others (Schwartz, et al., 1995; Werth, et al., 1997) have reported no difference in steady-state levels of elastin mRNA during solar elastosis.

Although ROS are widely accepted as crucial mediators of UV phototoxicity, the exact mechanism of their generation in irradiated skin is poorly understood and seems to be dependent on the presence of non-DNA chromophores acting as photosensitizers (Sander, et al., 2002). Photosensitization occurs as a consequence of initial formation of excited states of chromophores and their subsequent interaction with substrate molecules (type I photoreaction) or molecular oxygen (type II photoreaction) by energy and/or electron transfer (Foote, 1991). Various chromophores (e.g., riboflavin, porphyrins, nicotinamide adenine dinucleotide phosphate, ECM proteins, etc.) contained in human skin have been proposed as endogenous UV sensitizers of photo-oxidative stress (Dalle Carbonare and Pathak, 1992; Kipp and Young, 1999; Scharffetter-Kochanek, et al., 1997; Wondrak, et al., 2003). The skin-damaging effects of UVA appear to result from type II photoreactions in which UVA or near-UV radiation in the presence of certain photosensitising chromophores leads to the formation of reactive oxygen species (Dalle Carbonare and Pathak, 1992).

Upon UV photo-excitation human skin reveals a strong tissue auto-fluorescence indicative of the presence of various fluorescent UV chromophores, which in the dermis are mainly attributed to ECM protein bound heterocyclic cross-links of enzymatic and nonenzymatic origin (Tian, et al., 2001). Fluorescence excitation measurements on a 6-week-old hairless mouse skin in vivo revealed three maxima at 295, 340, and 360 nm. It was suggested that the
dominant excitation maximum at 295 nm originating in the epidermis is attributable to tryptophan, and the remaining two excitation maxima predominantly originating in the dermis are attributed to the presence of ECM protein fluorophores (Kollias, et al., 1998). Investigations have demonstrated that certain cross-link chromophores called advanced glycation end products (AGE) accumulate on dermal ECM proteins in human skin (Dyer, et al., 1993; Jeanmaire, et al., 2001) and may play an active part as mediators of skin phototoxicity acting as potent sensitisers of photo-oxidative stress (Masaki, et al., 1999; Wondrak, et al., 2002a; Wondrak, et al., 2002b). However, the observation that a significant amount of ROS is formed upon solar simulated light radiation of ECM proteins (collagen and elastin) not modified with AGE (Wondrak, et al., 2003) made the importance of the AGE doubtful.

2.2.2. Pathology

In a series of pathological conditions, either inherited or acquired, dermal elastic fibers can be severely affected. Such conditions include pseudoxanthoma elasticum, cutis laxa, and Marfan syndrome.

*Pseudoxanthoma elasticum*

Pseudoxanthoma elasticum is a relatively rare disorder affecting tissue rich in elastic fibers. Although most cases of the disease appear to be inherited, the precise mode of transmission through generations in many families is difficult to establish due to delayed age of onset and variable expression of the pseudoxanthoma elasticum phenotype (Christiano and Uitto, 1994). It is characterized by progressing mineralization and fragmentation of the elastic fibers (Pasquali-Ronchetti, et al., 1986) and accumulation in the mid dermis (Christiano, et al., 1992). The skin on the involved areas becomes thickened and leathery, forming yellowish plaques (Uitto, 1979). Typical skin lesions may contain relatively small, 1-3 mm yellowish papules that give the affected area a “plucked chicken skin” and the skin later on becomes lax, redundant and inelastic (Uitto, 1979). Although the elastic fibers of the affected tissue are clearly abnormal, the underlying mutations and etiologic mechanisms remain largely unknown. Findings indicated, however, that mineralization of the elastic fibers is caused by secretion of a series of glycoproteins with high affinity for calcium ions by dermal fibroblasts and their deposition within the fibers (Pasquali-Ronchetti and Baccarani-Contri, 1997).
Cutis laxa

Cutis laxa is an elastin related disorder, which induces the loss of elastin and elastic fibers in the cutaneous and other connective tissue compartments. It can occur in both genetic and acquired forms and exhibits a considerable heterogeneity in its clinical manifestations (Uitto, 1979). For example, in its most severe, perinatal form, the elastic fibers are almost undetectable in the skin and internal organs, leading to the early death of the patient. However, other cutis laxa phenotypes only lead to a mild wrinkling of the skin.

Cutis laxa is characterized by extreme loose and redundant skin, which frequently forms sagging folds on the face (Uitto and Lichtenstein, 1976). The major histopathologic feature of cutis laxa is the degeneration of elastic fibers in the dermis or other connective tissues due to very scarce or almost absent elastin. The disorder can be caused by different molecular defects: impaired synthesis of tropoelastin due to a variety of genetic defects affecting one of the steps of the synthetic pathway of this protein, increased degradation by abnormal activity of elastin-degrading enzymes (Fornieri, et al., 1994; Giro, et al., 1985), or decreased cross-linking of elastin due to decreased activity of lysine oxidase (Byers, et al., 1980). The latter is suggested to be associated with lysyl oxidase deficiency due to impaired copper metabolism (Pasquali-Ronchetti, et al., 1994). More recently, another genetic form of cutis laxa was successfully correlated to the elastin gene (Zhang, et al., 1999). In this case, a single base deletion in exon 30 leads to the missense translation of the C-terminal part of the molecule. As a consequence, the encoded tropoelastin lacks the C-terminal region.

Marfan syndrome

Marfan syndrome is a heritable connective tissue disorder manifested by defects in the dermal, skeletal, ocular, and cardiovascular systems (Godfrey, 1994). It is inherited as an autosomal dominant trait, with an incidence of nearly one per 10,000 people without gender or ethnic predilection. Life expectancy is significantly reduced, and more than 85% of patients die of cardiovascular complications. Defects in fibrillin are now known to cause the manifestations of this disease (Dietz, et al., 1994). Immunofluorescence studies of skin sections and dermal fibroblast cultures were the first to show this association.
2.3. Biochemical methods for elastin analysis

Although its role in the elastic property of elastic fibers is an essential element in the physiological function of connective tissues, elastin is a relatively less studied protein as compared to other human tissue proteins. The difficulty of performing biochemical studies explains the limited amount of work in this field. The heterogeneity of the elastic tissue composed of elastin and microfibrillar components, the small percentage of elastin in the skin, the difficulty in extracting and isolating elastin, the extreme insolubility of elastin neither in water nor in organic solvents are some of the problems encountered. Elastin is often defined as the residual protein obtained when all other components of connective tissue have been eliminated by drastic biochemical methods. Increasing interest in the biochemistry of elastin is derived from the realization that marked changes in the structure of the protein can be encountered in various acquired diseases, such as arteriosclerosis and pulmonary emphysema, and also during normal aging of the animal tissues (Sandberg, 1976; Urry, 1978). Aberrations in elastic fibers are also encountered in inherited diseases affecting the skin (Hashimoto and Kanzaki, 1975; Pasquali-Ronchetti, et al., 1981).

A limited number of studies were performed in order to measure the amount of elastin in skin and other connective tissues. Such biochemical analyses are based on the assumption that elastin is a very resistant protein which can be isolated from tissues by drastic chemical treatment or the amount of which may be evaluated by the desmosine content of the whole tissue. The first approach can be satisfactorily applied to “normal” elastin, as it is known that it resists to backbone cleavage by conditions by which all other biological constituents are removed. However, this approach cannot be applied to aged or pathological elastin, which could be already partially degraded and probably susceptible to chemical treatments. Moreover, the quantification of desmosines by radioimmunoassay (Uitto, et al., 1983) or by high performance liquid chromatography (HPLC) may underestimate immature elastin or pathological elastin in which, for some reason, lysyl oxidase activity is reduced (Fornieri, et al., 1987; Pasquali-Ronchetti, et al., 1994). The limitations of elastin analysis mentioned above call for alternative analytical methods devoid of such drawbacks. Qualitative and quantitative characterization of elastin on the basis of enzymatically-produced peptides was considered in this study.
2.4. LC-MS of peptides and proteins

The identification and measurement of peptides and proteins particularly at the nanoscale level has become a critical area of research. Most traditional approaches for the analysis of proteins have involved separation by one- or two-dimensional poly-acrylamide gel electrophoresis (1D or 2D-PAGE). Even though this technique is an excellent quantitative tool, it is often inadequate for the separation of complex protein mixtures. In addition, it is very difficult to automate, is labor intensive, and requires a very experienced analyst to read the gel fingerprints. For these reasons, liquid chromatography (LC), coupled with mass spectrometry (MS) has become the technique of choice. It can either be applied to the analysis of the actual 2D spots to improve the resolution of the separation or be used for the direct analysis of the protein mixture. LC-MS is an analytical technique that couples high-resolution chromatographic separation with sensitive and specific mass spectrometric detection. This includes HPLC-MS, capillary electrophoresis-MS (Banks, 1997; von Brocke, et al., 2001) and capillary electrochromatography-MS (Verheij, et al., 1991).

MS is a chemical analysis technique that is used to measure the mass of unknown molecules by ionizing, separating and detecting ions according to their mass-to-charge ratios \((m/z)\). The utilization of mass spectrometry to investigate biological processes dates back to the late 1930's and early 1940's with the use of stable isotopes and isotope ratio mass spectrometers (Ratner, et al., 1943; Schoenheimer, et al., 1937). The last twenty years has seen a dramatic increase in the capabilities of MS. At the beginning of this period, the invention of fast atom bombardment (FAB) (Barber, et al., 1981), plasma desorption (PD) and thermospray (TSP) that permitted the production of gas phase ions from charged and polar compounds without prior chemical derivatization enabled easier analysis of involatile and thermally unstable molecules, especially those of biological interest. These new ionization techniques made high mass macromolecules amenable to mass spectrometric analysis and they opened the room for the development of other ionization techniques, such as matrix-assisted laser desorption/ionization (MALDI) and ESI, applicable for the analysis of even bigger molecules, such as polypeptides and proteins (Lim and Lord, 2002). The remarkable advances in analyzer technology and electronics accompanied this development, so that some of today's most powerful mass spectrometers are relatively small and easy to use.
ESI has had a tremendous commercial impact over the last few years on the use of mass spectrometry in biological research (Smyth, 2003). It was the first method to extend the useful mass range of instruments to well over 50,000 amu. Although introduced in its present form in 1984 (Yamashita and Fenn, 1984), a major discovery took place almost unnoticed in 1968, when Malcolm Dole and coworkers were able to bring macromolecules into the gas-phase at atmospheric pressure (Dole, et al., 1968). Building on Dole et al.’s ideas, Fenn and co-workers developed electrospray as a true interface for mass spectrometry (Yamashita and Fenn, 1984). Since then, a wide range of biomolecules has been investigated by ESI.

Figure 2.3 shows the steps involved in the formation of ions by ESI. Briefly, liquid sample introduced from the LC or syringe pump then enters the ESI needle, to which a high voltage is applied. The needle sprays the sample solution into very fine droplets that retain electrical charges at their surface. As the droplets fly through a region of dry gas at atmospheric pressure, solvent from the droplets evaporates, thereby concentrating the number of charges on a smaller area. The solvent evaporates completely, leaving a multiply charged analyte molecules. The mist of charged molecules then hits a heated capillary skimmer that leads to the ion optics and mass analyzer.

A significant advance in the ESI technique was the introduction of nano-ESI, which utilizes a very low solvent flow rate and consequently a very small sample volume. The signal-to-noise ratio of the technique is enhanced compared with standard ESI and a small aliquot of the sample (e.g., 1 µL of 10^{-6} M) will spray for ca. 30 minutes, thus enabling a peptide mass map and several MS/MS sequence tag analyses to be performed on single minimal sample introduction without chromatographic separation (Karas, et al., 2000). Moreover, advantages, such as a lower susceptibility to salt contamination and a more uniform response to chemically very different analytes such as peptides and oligosaccharides (Karas, et al., 2000), and high sensitivities down to the low attomol range (Valaskovic, et al., 1995) have been reported.
A major breakthrough in the analysis of biomolecules came in 1988 with the introduction of MALDI (Karas and Hillenkamp, 1988; Tanaka, et al., 1988), a technique that is capable of detecting molecules over 300,000 amu. The technique involves embedding the analyte in a solid matrix, which absorbs laser radiation, typically in the UV region. Upon drying the matrix-analyte mixture, the analyte molecules are intercalated into the matrix crystals. A number of matrices including $\alpha$-cyano-4-hydroxy cinnamic acid, 2,5-dihydroxybenzoic acid, and sinapinic acid are commonly used in peptide and protein analysis.

An important development to improve the mass resolving power of MALDI-MS is the implementation of delayed extraction (DE) (Bahr, et al., 1997; Brown and Lennon, 1995; Whittal and Li, 1995), a method based on the idea of “time lag energy focusing” originally described in 1955 by Wiley and McLaren (Wiley and McLaren, 1955). In conventional MALDI instruments, the ions generated by the laser beam near the surface of the sample probe are extracted by a dc potential. In DE mode, a short time delay is inserted between the laser ionization and ion extraction events. Within the delay time, ions will spread out into the
Background

extraction gap owing to their initial velocities, thus spreading out in space and acquiring different total kinetic energies, when accelerated by the delayed extraction. Application of the appropriate pulse voltage provides the energy correction necessary to simultaneously detect all ions of the same \( m/z \) regardless of their initial energy. The initially less energetic ions closer to the voltage source receive more energy (from the pulsed potential) than the initially more energetic ions further from the source at the time the pulse is applied. An energy/spatial correction is thus provided such that all ions of the same mass/charge reach the detector plane simultaneously.

2.5. Peptide and protein sequencing

2.5.1. Edman degradation

To sequence a protein some years back, a substantial amount had to be purified and a technique known as Edman degradation had to be used. This method, which was developed by Pehr Edman (Edman, 1950), relies on the identification of amino acids that have been chemically cleaved in a stepwise fashion from the amino terminus of the protein and requires much expertise. Phenylisothiocyanate reacts with the uncharged terminal amino group of the peptide to form a phenylthiocarbamyl derivative. Under mildly acidic conditions, a cyclic derivative of the terminal amino acid is liberated leaving the intact peptide shortened by one amino acid. The hydrolysis reaction results in a rearrangement of the released N-terminal residue to a phenylthiohydantoin derivative, which can be identified by chromatographic procedure. Often no sufficiently long or unambiguous peptide sequence could be assigned and the method fails completely if the protein is acetylated at its amino terminus or is otherwise blocked to the Edman reaction, which requires a free amino terminus.

2.5.2. Tandem mass spectrometry

During the 1990s, tandem MS displaced Edman degradation significantly, because it is much more sensitive and can fragment the peptides in seconds instead of hours or days (Wilm, et al., 1996). Furthermore, MS does not require proteins or peptides to be purified to homogeneity, it has no problem of identifying blocked or otherwise modified proteins, and the sample amount necessary for analysis is usually less than 1 pmol to obtain a high quality mass spectrum (Mendes, et al., 2004). In the last few years, further breathtaking
technological advances have established MS not only as the definitive tool to study the primary structure of proteins, but also as a central technology for the field of proteomics.

A desirable tandem mass spectrometer in peptide sequencing should provide high mass accuracy for both precursor and product ions, high resolution, simple selection of precursor ions, precise tuning of the collision energy, and a much simplified calibration procedure. The fragmentation ions produced by tandem MS can be separated into two classes. If, after the peptide fragments along the amino acid backbone, the charge stays on the N-terminal fragment, then the ion is designated an "a", "b", or "c" ion depending on which particular bond is cleaved in the peptide chain (fig. 2.4). Conversely, if, after the peptide fragments, the charge remains on the C-terminal fragment, the ions are labeled "x", "y", or "z" ions.

Figure 2.4: Peptide fragment ions in MS/MS

Peptide sequencing by tandem mass spectrometry begins with the acquiring of fragment spectra of the peptides in a mixture of enzymatically digested protein using different approaches, such as LC-MS/MS, nano-ESI-MS/MS, and MALDI-PSD-TOF MS. Deducing peptide sequences from the raw MS/MS data is slow and tedious when it is done manually. Instead, the most popular approach is to convert the resulting mass spectra into mass list and to subject it to peptide sequencing either by matching with protein databases (Creasy and Cottrell, 2002; Mann and Wilm, 1994; Perkins, et al., 1999; Yates, et al., 1995) or by de novo sequencing (Ma, et al., 2003; Shevchenko, et al., 1997; Taylor and Johnson, 2001). Even more accurate results could be obtained by coupling both methods when the protein of interest exists in databases.
**LC-MS/MS**

In this method, the peptides that are generated by protein digestion are injected onto an HPLC column that is directly coupled to a mass spectrometer. The peptides are eluted from the column using a solvent gradient of increasing organic content, so that the peptide species elute in order of their hydrophobicity. As the mass spectrometer can distinguish the peptides by their masses, there is no need to separate them into non-overlapping chromatographic peaks and usually many peptides arrive at the end of the column at any given time. The signal intensity in the mass spectrum is directly proportional to the analyte concentration, so the peptides are eluted in as small a volume as possible.

Having determined the $m/z$ values and the intensities of all the peaks in the spectrum, the mass spectrometer then proceeds to obtain fragment ion spectra. In this process, a particular peptide ion is isolated, energy is imparted by collisions with an inert gas (such as nitrogen molecules, or argon or helium atoms), and this energy causes the peptide to break apart. A mass spectrum of the resulting fragments - the tandem MS (also called MS/MS or MS$^2$) spectrum - is then generated. In MS terminology, the species that is fragmented is called the ‘precursor ion’ and the ions in the tandem-MS spectrum are called ‘product ions’.

**Nano-ESI MS/MS**

For many applications, where on-line coupling with HPLC is not needed, the high flow rates typically invoked by syringe pumps are neither required nor applicable, if low sample volumes are to be investigated. Dilution of sample to obtain a sufficiently large sample volume usually does not solve the problem, since satisfactory results are obtained in a concentration range between $10^{-5}$ and $10^{-7}$ mol/L, depending on the kind of analyte and the purity of the sample. Only in special cases, lower concentrations have been analyzed successfully. Hence, when only low volumes and amounts of sample are available, the logical consequence is to reduce the flow rate as low as possible, and this is one of the ideas underlying the concept of nano-ESI (Wilm and Mann, 1996). In the mid-1990s, Wilm and Mann (Wilm and Mann, 1994) introduced nanospray to the MS community, and Valaskovic et al. pioneered the development of nanospray ionization emitters (Valaskovic, et al., 1995).

In nano-ESI glass capillaries are used as spray capillaries which are drawn out at one end either by a mechanical or a laser puller to give orifices of only 1-10 µm in diameter. For sufficient conductivity the capillaries are coated with conductive material, e.g. gold, or a thin
metal wire is inserted into the capillary. The capillaries are loaded from the back with only 1-5 µL of sample solution. The spray needle is adjusted under stereomicroscopic or video camera control directly in front of the counter electrode orifice. The distance between needle and orifice is only 0.5 to 2 mm. No liquid feed system is used; the dispersion of the liquid and the generated flow rates of about 20-50 nL/min are exclusively due to the forces of the electrical fields when voltages between 500 and 1000 V are applied. Some backpressure to the spray capillary by a gas-tight syringe is usually applied and is helpful in initiating the flow. Spray capillaries are only used once, which gives the additional advantage that contamination and memory effects are avoided (Karas, et al., 2000).

Nano-ESI provides many advantages over conventional ESI for the analysis of macromolecules, including decreased sample consumption, lower flow rate, and increased sensitivity (Valaskovic, et al., 1996). The high molecule-ion intensity is a prerequisite for structure elucidation with MS/MS experiments. Some problems associated with nano-ESI are that the handling of the glass capillaries (sample loading, opening procedure of the capillary) requires some training, as well as the spatial adjustment. Moreover, nano-ESI needles vary in quality and reliability and one has to be aware of absorption effects to the glass walls.

**MALDI-PSD-TOF MS**

PSD analysis is an extension of MALDI MS that allows one to observe and identify structurally informative fragment ions from decay taking place in the field free region after leaving the ion source. After leaving the ion source, all ions have the same nominal kinetic energy, most of them are still unfragmented precursor molecular ions and they have already acquired internal energy by various mechanisms (gas-phase collisions, laser irradiation, thermal mechanisms, etc.). During their flight through the field-free drift region they have a long time available for post-source decay into product ions. These product ions still have basically the same velocity as their precursor ions but now they have a much lower kinetic energy owing to their lower mass. The kinetic energy of the product ions is a measure of their mass. In linear instruments, PSD ions are detected at the same time as their precursors and, therefore, cannot be mass analyzed. The ion reflector, in classical time-of-flight instruments used as a device for flight time compensation of initial energy distributions, is used here as an energy analyzer and thus as a mass analyzer for PSD ions. Owing to their mass-dependent kinetic energies, PSD ions are reflected at different positions within the reflector and thus have mass-dependent total flight times through the instrument (Spengler, 1997).
Conversion of MS/MS raw data to mass list

In tandem MS identification of peptides, the fragment ion spectra derived from peptides should either be matched with theoretical spectra calculated from known DNA or protein sequences, or be used for automated de novo sequencing using special software tools. Algorithms used for this purpose usually have their own input formats and are not able to read the proprietary binary file formats of the mass spectrometer manufacturers. They usually demand conversion of the MS/MS raw data to peak list. There are some software programs, such as Mascot Distiller, MassLynx with its ad-on program MaxEnt, and wiff2dta (Boehm, et al., 2004) that are capable of creating high quality peak lists from each mass spectrum. Although they use different approaches to detect and pick a signal as a true peak, all of these algorithms have common aims, such as removal of noises to subsequently reduce data and selection of peaks of the monoisotopic mass. In this work, Mascot Distiller and MassLynx were used to create peak lists for the MS/MS spectra produced.

2.5.3. De novo sequencing

Theoretically, each peptide fragment in a series differs from its neighbor by one amino acid (fig. 2.4). It is, therefore, possible to determine the amino-acid sequence by considering the mass difference between neighboring peaks in a series. Determining the sequence in this way is called de novo sequencing. A number of algorithms and software packages have been reported for the deduction of protein sequences from MS/MS data (Chen, et al., 2001; Ma, et al., 2003; Taylor and Johnson, 1997; Taylor and Johnson, 2001). A powerful software, PEAKS, which extracts amino acid sequence information from MS/MS data with or without the use of databases was used in this work (Ma, et al., 2003).

De novo sequencing may not be possible for the majority of spectra obtained because of the following reasons (Peng and Gygi, 2001).

1. More than one ion series is usually present (i.e. sequencing the peptide from the N- and C-termini simultaneously),
2. Ion series are rarely complete,
3. Fragment ions are present in varying abundances, often with associated losses of water and/or ammonia, and
4. Some amino acids have very similar or identical molecular masses (e.g. isoleucine and leucine, glutamine and lysine, etc.).
For these reasons, database-searching algorithms have become an indispensable tool for the interpretation of tandem mass spectra.

2.5.4. Database searching

A decade ago, it has been realized that the peptide-sequencing problem could be converted to a database-matching problem, which would be much simpler to solve. The reason database searching is easier than de novo sequencing is that only an infinitesimal fraction of the possible peptide amino-acid sequences actually occur in nature. A peptide-fragmentation spectrum might, therefore, not contain sufficient information to unambiguously derive the complete amino-acid sequence, but it might still have sufficient information to match it uniquely to a peptide sequence in the database on the basis of the observed and expected fragment ions. A limitation of database searching compared to de novo sequencing is that large-scale proteomic experiments should only be carried out using organisms that have had their genome sequenced, so that all the possible peptides are known.

There are several different algorithms that are used to search sequence databases with tandem-MS-spectra data, and they have names such as PeptideSearch (Mann and Wilm, 1994), Sequest (Eng, et al., 1994), and Mascot (Perkins, et al., 1999).

PeptideSearch makes use of the fact that fragmentation spectra usually contain at least a small series of easily interpretable sequence (Mann and Wilm, 1994). This series constitutes an amino acid tag. The lowest mass in the series contains information about the distance, in mass units, to one terminus of the peptide, and the highest mass contains information about the distance to the other peptide terminus. Together, the peptide-sequence tag consists of three parts (the amino-terminal mass, a short amino-acid sequence and the carboxy-terminal mass). This construct can be matched against sequences in the database and, if desired, the peptide that is identified can be made to comply with the cleavage event of the proteolytic enzyme used.

The software algorithm Sequest matches a peptide sequence with a tandem mass spectrum using the following steps (Eng, et al., 1994):
1. Peptides with molecular masses matching that of the peptide ion sequenced in the tandem mass spectra are extracted from a protein database;

2. Each peptide is given a preliminary score by examining the number of predicted fragment ions from the database peptide that match the acquired fragment ions in the tandem mass spectrum; and

3. The top 500 best matching peptides undergo a more rigorous ion-matching algorithm that generates a cross-correlation score. A list of peptides with good correlation is returned to the user with the top-scoring peptide being considered the best candidate.

Mascot search engine calculates the theoretically predicted fragments for all the peptides in the database and matches them to the experimental fragments in a top-down fashion, starting with the most intense b- and y-ions (Perkins, et al., 1999). It is called probability-based matching, since it calculates the probability that the number of fragment matches is random. The negative logarithm of this number multiplied by 10 is the identification score.