Summary

1. COLLAGENS AT A GLANCE

Collagens are the most abundant proteins of human body and the principal components of mammalian connective tissues (≈ 30% of total protein mass). Nearly 30 different collagen types numbered with roman numerals have been described up to now: they provide tensile strength and flexibility to bones, dermis, tendons, ligaments and structural integrity to internal organs. Among them, collagen I is the most common, representing over the 90 percent of our total collagen[1] and the archetype of all fibrillar collagens, endowed with a predominant tissue assembly and maintenance role.

Collagens structural task is strictly related to the triple helical skeleton built up by three polypeptide chains, whose molecular hallmark is the multiple repetition of the triplets Gly-Xaa-Yaa, in which the amino acid residues in X- and Y-positions are frequently proline or hydroxyproline (hyp). The prototypal collagen I is generally a heterotrimer, comprising two proα1(I) and one proα2(I) chains.

Diversification of collagen types depends on molecular structure, supramolecular assemblies and slightly on chain composition, disclosing different roles and tissue distribution.[2]

1.1 Structural hints on molecular and supramolecular organization

Collagen triple helix is a widespread structural element, occurring in many other proteins. Each polypeptide chain consists of a left-handed polyproline-II-type helix, which is in contrast a relatively rare structural element in proteins.

INTRODUCTION

Individual chains supercoil along a common axis to form a right handed superhelix, as schematically represented in Figure 1.

The repetitive motif Gly-Xaa-Yaa is an essential requirement for triple helix stabilization. The presence of glycine residues buried in the core of the helix provides stability, since a larger side chain than H at the Cα would hamper the formation of an interchain H-bond between the backbone NH-group of glycine and the backbone CO of a residue in X-position, faced to the Y-position of the third neighboring chain.

Besides, the presence of the two imino acids proline and hydroxyproline, usually located at X- and Y-position, represents a major steric restriction, contributing to freeze Cα-N angle (Φ-angle). Rotation around Cα-CO (Ψ angle) is also restricted by unfavorable van der Waals interactions. Therefore, the conformation of the polyproline-II-helix is uniquely determined by Φ and Ψ dihedral angles, which define a narrow area in the Ramachandran plot (Φ vs Ψ).

A key aspect in the superhelix stability is also the orientation of Xaa and Yaa residues side chains, which point out of the helix and thus are freely accessible for binding interaction, favoring the cooperation with other molecules of the extra-cellular matrix (ECM).[3,4]

Additional effects supporting triple helical skeleton are either electrostatic interaction involving lysine residues or interchain H-bonds formed by the OH-groups of 4(R)-hydroxyproline; the latter contribute to increase the superhelix thermal stability.[5,6,7]

Considering the key aspects of triple helix stabilization, it is noteworthy that bonds within each chain in the unfolded state are up to 20% in *cis*-configuration and slow *cis-trans* isomerization, necessary for the formation of native triple helix, represents a kinetic unfavorable step often causing collagen chain misalignment.[3]

Native collagens comprehend two non-collagenous (NC) domains, namely the N- and C-propeptides, which are proteolitically removed only briefly before fibrils formation. Their presence is crucial for proper chain alignment, therefore playing a major role in superhelix nucleation.[3,8]

Similarly, in fibrillar collagens, the correct alignment of the triple helices is strictly related to the chain sequences and to the helices symmetry.

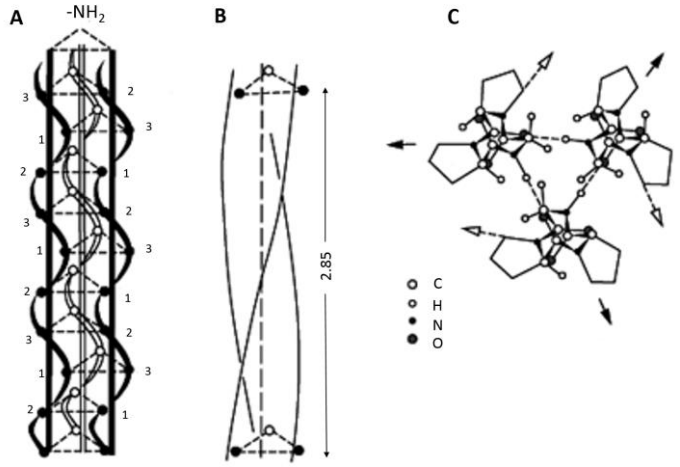


Figure 1. *Model of the collagen triple helix. The structure is shown for (Gly-Pro-Pro)$_n$ in which glycine is designated by 1, proline in X-position by 2 and proline in Y-position by 3; A, B: side views. C: top view in the direction of the helix axis.*[3]

In fibril-forming collagens (types I, II, III, V and XI) single chains are synthesized as precursors termed procollagens, flanked by NC domains at both N- and C-terminal ends. Therefore, unlike other collagen superfamily members, fibrillar collagens require the proteolytic processing by specific collagenases to cleave the extensions and achieve the mature and functional form. The propetides direct the triple helix formation in a zip-like manner from C- to N- terminus and their removal represents a crucial step, triggering spontaneous assembly into fibrils that range in diameter from 10 to 300 nm.[9,10]

The C-propeptide is a highly conserved domain, which is responsible for the correct α chain alignment and recognition via the CRS (chain recognition sequence)[11] and provides solubility to procollagens molecules, thus preventing improper aggregation. N-propetide, whose sequence is rather deeply diversified on collagen type, controls fibrils diameter and shape.[12] Collagen fibrils are mainly etherotypic, consisting of one or two quantitatively major collagens (I–III) as well as one quantitatively minor collagen (V or XI).[13]

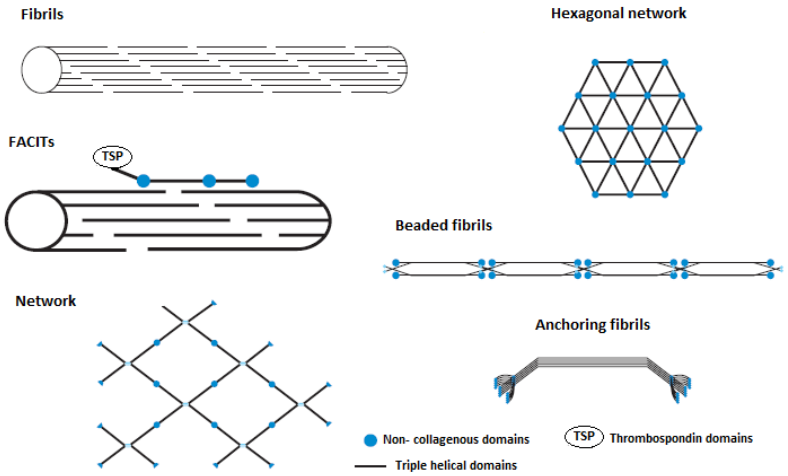


Figure 2. *Supramolecular assemblies formed by collagen.*[14]

Proteoglycans, non-fibrillar collagens and NC proteins are included in heterothypic fibrils and contribute to the fibrillogenesis process, controlling collagen crosslinking[15] and acting as nucleators for fibrillogenesis of collagen type I and II, in case of integrin and fibronectin.[16] Collagen fibrils are made of collagens II, XI, and IX or of collagens II and III in cartilage, [17] of collagens III and I in skin, and of collagens I and V in cornea.[18] The organization of each tissue depends not only on fibrils features, but also on suprafibrillar arrangements, mapping out a series or different areas: in the dermis of skin, fibrils take on a complex three-dimensional weave; in tendons, fibrils are parallel or anti-

parallel, subject to a planar undulation, while in bone osteons they are disposed forming concentric lamellae.[9]

In most tissues, notwithstanding collagen type I is the quantitatively predominant form, collagen V, which usually represent less than 5% of the total collagen content, inserts between strands of the microfibril, playing a fundamental part in fibrils assembly.[19]

Introducing non-fibrillar collagens, schematically represented in Figure 2, fibrils-associated collagen with interrupted triple helix (FACIT) are the most relevant among ECM components: they have small non-triple helical regions, leading to kinks that straighten under small strains, providing elasticity to tissues and organizing fibrillary network in the ECM. FACIT collagens are incorporated into fibrils selectively altering the surface properties and their interactions with other ECM elements. Small leucine-rich repeat proteoglycans (SLRPs), such as decorin and fibromodulin,[10,20,21] rather than forming covalent bonds, are essential for the surface interaction between fibrillar and FACIT collagens. Interestingly, collagen XII expression was found remarkably induced in response to wounding caused by excess of mechanical stress.[22]

The novel FACIT collagens comprehend five newly described collagen types (XVI, XIX, XX, XXI XXII), sharing structural features of the previously identified FACITs, although very little is known about their role in connective tissues.

Other non-fibrillar supramolecular structures are listed in Table 1.[23]

Classes	*Type*	*Genes encoding proteins typical of collagen composition*	*Description*
Fibrillar	I	COL1A1, COL1A2	Quarter-stagger packed fibrils that are primarily found in fibrous stromal matrices, such as skin, bone, tendons and ligaments
	II	COL2A1	
	III	COL3A1	
	V	COL5A1, COL5A2, COL5A3	
	XI	COL11A1	
	XXIV	COL24A1	
	XXVII	COL27A1	
FACIT	IX	COL9A1, COL9A2, COL9A3	Fibril-associated molecular bridges that are associated with type I (XII, XVI, XIX, XXI) and type II (IX, XVI, XIX) collagen fibrils
	XII	COL12A1	
	XIV	COL14A1	
	XVI	COL16A1	
	XIX	COL19A1	
	XX	COL20A1	
	XXI	COL21A1	
	XXII	COL22A1	
Basement membrane	IV	COL4A1, COL4A2, COL4A3, COL4A4, COL4A5, COL4A6	Network structure comprised of laminins and basement membrane proteins
Long chain	VII	COL7A13	Anchoring fibrils that are associated with the basement membrane
Filamentous	VI	COL6A1, COL6A2, COL6A3, COL6A5, COL6A6	Beaded microfibrils
Short chain	VIII	COL8A1	Hexagonal lattice structure
	X	COL10A1	These collagens both regulate, and are regulated by, hypertrophy in cartilage
Multiplexins	XV	COL15A1	

	XVII	COL18A1	Multiple triple helical domains with interruptions containing chondroitin sulphate and heparin sulphate glycosaminoglycans
MACIT (transmembrane domain)	XIII	COL13A1	Cell surface molecules with extracellular membrane-spanning and intracellular domains
	XVI	COL17A1	
	XXIII	COL23A1	

***Table 1** . Primary types of collagens. Collagens are here classified on the basis of domain structure homology and suprastructural assembly. MACIT, membrane-associated collagens with interrupted triple helices.[23]*

1.2 Collagen metabolism

1.2.1 Biosynthesis

The processes by which procollagen molecules are synthesized include unique steps that are not found in other globular proteins biosynthesis pathways.[24] The nucleation of the triple helix is coupled with a number of specific post-translational modifications, such as prolyl 4-hydroxylation at the Y-positions, prolyl 3-hydroxylation at the X-positions and lysyl hydroxylation to form hydroxylysine (Hyl) residues. Some of the Hyl residues are subsequently glycosylated to form galactosyl-Hyl and glucosyl-galactosyl-Hyl residues.

This delicate modification steps occur during polypeptides translocation in the in the lumen of the rough ER, where they are assisted by a set of ER resident molecular chaperones and enzymes, listed and briefly described in Table 2.

These chaperones implied in procollagen biosynthesis fall into two main classes: Bip/Grp78, calnexin, calreticulin, protein disulphide isomerase (PDI) and peptydil prolyl *cis-trans* isomerases (PPIases) are general ER-resident chaperones with broad binding specificity, and hence function with a variety of client proteins. Apart from the PPIases, they bind to the NC domains of

procollagens. The members of the other chaperones class, including prolyl 4-hydroxylase (PH-4) and heat shock protein 47 (hsp47), rather interact specifically with the collagenous domain and are involved in folding, cellular trafficking and quality control of procollagen molecules.

Chaperone	*Major binding-region*	*General function*
Bip/Grp78	C-propeptide	Hsp70 homolog in the ER; it may capture procollagen molecules that have inappropriately folded C-propeptide domains
Calnexin	C-propeptide (N-linked sugar)	ER-resident lectin with a transmembrane domain; recognizes the pro-α-chain in the ER when N-linked oligosaccharide is attached to its C-propeptide region
Calreticulin	C-propeptide (N-linked sugar)	Calnexin homolog without the transmembrane domain; it recognizes the proα-chain in the ER when N-linked oligosaccharide is attached to its C-propeptide region
PDI	C-propetide	Also functions as the β-subunit of P4-H; catalyzes the correct pairing of the disulphide bonds, coordinating the trimeric assembly of the proα-chains
P4-H	Gly-X-Y repeats (single chain)	Known as the principal modification enzyme for procollagen; hydroxylates of the Pro residues at Y-positions; retains in the ER abnormal procollagen
PPIase	Gly-X-Y repeats (single chain)	Some members may be involved in procollagen folding; catalyzes prolyl peptide bonds cis-trans interconversion
Hsp47	Gly-X-Y repeats (triple helical)	Collagen-specific heat-shock protein found in the ER, specifically recognizes procollagen triple helix, favoring the correct folding and preventing improper aggregation

Table 2. *Molecular chaperones and enzymes implied in collagen biosynthesis.*[24]

C-propeptide units association, stabilized by intermolecular disulphide bridges, initiates the folding of procollagen triple helix and must be accomplished with the correct pairing according on collagen type.

Since most cells produces different types of collagens, they synthesize homologous α-chains that do not possess the intrinsic ability to align with the correct partners; rather, 15 discontinuous residues within the C-propeptide portion are retained to assure the type-specific assembly of the individual procollagen chains.[11] C-propetide domains consist of a 250 amino acid sequence highly homologous to one other, containing eight conserved cysteine residues, oxidized into four cystine pairs during the trimerization process. The four cysteine residues located at the N-terminus of the C-propeptide contribute to this intermolecular crossing while the latter four form intramolecular disulfide bridges.[25] Additional interchain disulphide bridges are formed in case of collagen type III in the C-telopeptide region, a short NC sequence linking the C-propeptide to the triple helical domain.[24]

Following the three pro-α-chains trimerization at the C-propeptide domains, the assembly of the triple helix propagates toward the N-terminus.

The rate-limiting step of this process is the isomerization of the *cis*-peptide bonds connected to a Pro or Hyp residues to a *trans* configuration,[26] accelerated by peptydil prolyl *cis-trans* isomerases (PPIases) that also function as molecular chaperones.

Correctly folded and opportunely modified procollagen units are transferred from the ER to the Golgi apparatus; the specific molecular chaperone hsp47 listed in Table 2, plays a key role within this translocation step of procollagen molecules, as will be more extensively discussed in paragraph 3.1

In the *cis*-Golgi, procollagen molecules are stacked together and form lateral aggregates that may either protect newly synthesized procollagen from matrix

metalloproteases attack or afford them sufficient thermal stability at body temperature.[27]

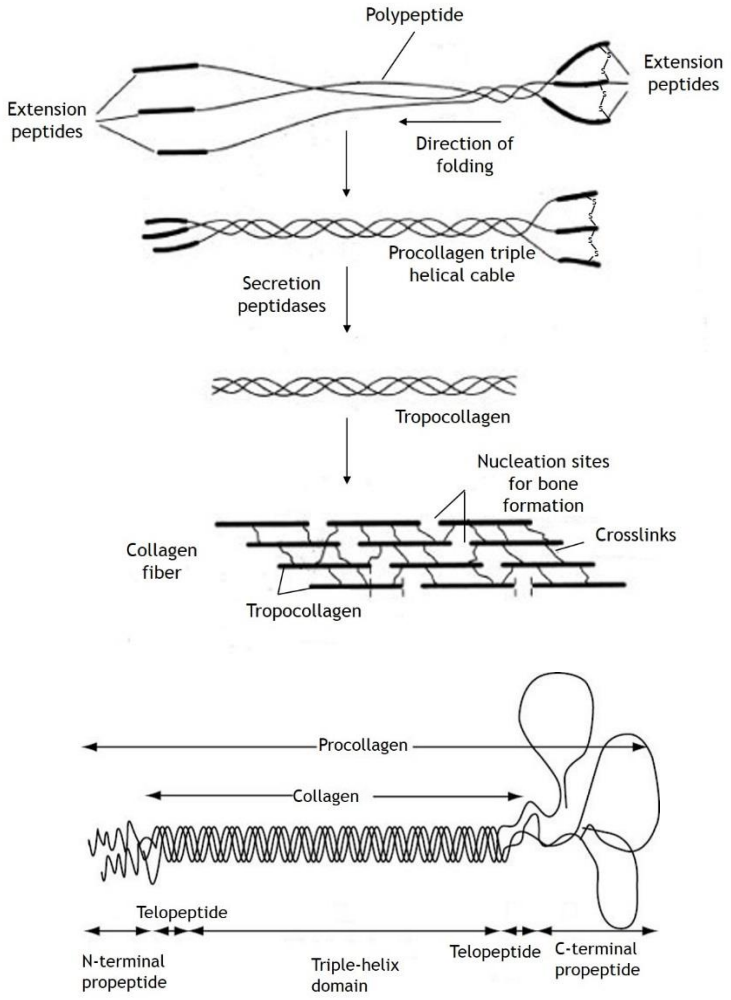


Figure 3. *Scheme for collagen biosynthesis: fibrillar collagens are synthesized in the form of procollagen with N- and C-propeptide extensions at both end of the α-chains.*[24]

Finally, these aggregates are incorporated into vesicles and are subsequently secreted into the extracellular space, where upon enzymatic cleavage of the propeptide extensions, the triple-helical collagen molecules self-assemble into

fibrillar supramolecules, thus generating the mature collagen molecules. The principal steps of collagen biosynthesis pattern are graphically summarized in Figure 3.

In conclusion, either procollagen specific chaperones or proteins endowed with a general chaperoning role take part to the complex events that occur in the ER lumen, contributing to the superhelix quality control mechanisms and ensuring that only properly folded collagen is secreted.

1.2.2 Degradation pathways

Extensive turnover of ECM components takes place during a variety of both physiological and pathological events implying tissue remodeling, such as development, tissue repair, degenerative connective tissue diseases, and cancer.[28] In these very diverse contexts, collagens under-go continuous renewal and their degradation is involved in physical expansion of tissues, liberation of latent growth factors embedded within the ECM, regulation of vascular development and cellular differentiation.[29,30]

To date, the main collagen turnover pathways comprehend pericellular digestion through K-cathepsin, specific for osteoclast-mediated bone resorption, extracellular cleavage by zinc-dipendent membrane-associated metalloproteases (MMPs) and endocytosis followed by lysosomal degradation. In healthy tissues, collagen is present in its fully hydroxylated form, implicated in insoluble highly cross-linked fibers and sheets of triple helical structures, resistant to attack from most proteases.[31] Nevertheless, in the context of tissue remodeling and renewal necessary in a variety of either abnormal or physiologic conditions, proteolytic enzymes such as MMP-1, MMP-2, MMP-8, MMP-13, MMP-14, MMP-15, and MMP-16 and the cysteine protease cathepsin K are capable to trigger collagen degradation, via the cleavage of intact fibers.[32,33]

Up to now, 23 MMPs or matrixines have been identified and most of them are barely expressed or even absent in healthy tissues, yet they are up-regulated in inflammation, tissue remodeling, wound repair and a variety of different situations. MMPs are secreted as activated enzymes or as latent pro-enzymes, subsequently activated by plasmin or other serine proteases[34,35] and their gene expressions is finely controlled by citokynes, matrikines and growth factors such as transforming growth factor β1 (TGF- β1).[35] MMPs generally consist of a prodomain, a catalytic domain, containing three histidines that bind zinc, a hinge region and a hemopexin domain.

Based on their substrate specificity, MMPs can be divided into 6 groups, among which, the so termed collagenases, MMP-1, -8, -13 and -14 target interstitial collagen I, II and III and are responsible of the very first step of fibers degradation, resulting in the obtainment of characteristic ¼ and ¾ fragments. Recent studies demonstrated that next to MMP-1, MMP-2 also plays a key role in collagenolysis.[36]

To achieve efficient collagenolytic activity the enzyme in charge must be able to (1) bind the collagen, (2) unwind the triple helix and (3) cleave the individual strands of the helix.

According to the MMP collagen cleavage model developed by Lauren-Fields et al., collagen fibers cleavage is initiated at the border between a tight triple helical region, rich in imino-residues and a loose triple helical region, rather displaying a poor content in imino acids.[37] Lauren-Fields and colleagues furthermore established that the ability of MMPs to cleave collagen depends on the cooperation of the catalytic domain, the C-terminal hemopexin, and the hinge region of the enzyme. The hemopexin domain is responsible for the binding of MMPs to the fibers and for the orientation and destabilization of the fibers. The catalytic domain unwinds the triple helix and cleaves each strand.

The precise role of the hinge region is still unknown, but removal of this region results in a loss of collagenolytic activity, thus proving its crucial role.[37]
Following the initial degradation step, denatured collagen, also called gelatin, is further degraded by the gelatinases (MMP-2, and -9).
Although most studies focused on degradation patterns mediated solely by extracellular proteases, during the last decades several research groups demonstrated that the complete collagen fibers digestion requires an interplay between the extracellular collagenases and the endocytic pathway.[38,39,40]
Thus, pericellular and extracellular degradation would act in concert with internalization of the obtained soluble partially digested collagen fragments, which are recognized by specific membrane receptors, belonging either to the integrin family (β1-integrins) or to the mannose receptor family, such as urokinase-type plasminogen activator receptor-associated protein (uPARAP).
On one hand, mature collagen structures exceed in orders of magnitudes the proper size for single cell endocytic apparatus, on the other hand MMPs are incapable of completely degrading collagen, but rather facilitate the formation of fragments with a manageable size that become appropriate targets for endocytic uptake by cells and subsequent lysosomal degradation in a cysteine cathepsin-dependent manner.[41]

1.3 Critical aspects of collagen homeostasis

Since collagen predominance within all connective tissues, disturbance in the delicate balance of collagen biosynthesis and degradation pathways may determine the transition from the physiologic to the disease state.
On one hand, prevalence of biosynthesis over degradation is connected to fibrosis states and severe scarring; on the other hand low-than-normal collagen levels, due i.e. to either MMPs up-regulation or inefficient

replacement with newly synthesized collagen fibers may cause a loss of elasticity at the skin level and abnormal fragility of bones. Indeed, in the early 80s, the dynamic state of collagen mass had become a clear well-defined concept, in contrast with the previous idea of collagen as an inert protein, synthesized and degraded at very slow rates in adult tissues.[42]

While introducing collagens metabolism, it is noteworthy to consider that their biosynthesis and degradation patterns consist in highly complex multi-step processes; hence, collagen mass in the body offers many different levels of regulation; some of them are highlighted in the following paragraphs.

1.3.1 MMPs

Extracellular matrix degradation pathway is finely controlled via MMPs activity modulation, in the context of transcriptional and post-transcriptional regulation, starting from gene expression all the way to zymogen activation and endogenous inhibition, with each level controlled by multiple factors.

Firstly most MMPs are regulated at a transcriptional level by several ECM functional molecules, including hormones, inflammatory cytokines and growth factors such as TGFβ, epidermal growth factor receptor (EGFR), tumor necrosis factor-α (TNF-α) and interleukin-1β (IL-1β), which finely control their expression.[43,44]. Interestingly the expression of genes encoding for MMP-1, MMP-2, MMP-9 and MMP-13 was found to be up-regulated by hyperglycemia,[45,46] a factor that may act synergistically with other ECM molecules, such as interleukin-6 (IL-6).[47] In contrast, in human endothelial cells, hypoxia/reoxygenation conditions were found to increase or decrease the expression level of MMP-2, depending on the duration of exposure to hypoxia.[48]

Remarkably, cell-cell and ECM-cell interaction may also influence MMPs expression;[49,50] recent works demonstrated than their regulation upon intervention of ECM functional molecules and external factors strictly depends on ECM type and cell type.[51,52]

Secondarily, since most MMPs are secreted as pro-enzymes, the activation of zymogen represents an additional key step for their regulation, which can be achieved either intracellularly or in the extracellular space.

Intracellular activation, for example, occurs in MMP-11 and MT1-MMP via proteolytic activation by cellular endoproteases,[53,54] *via* oxidative stress in the case of latent collagenases[55,56] and via phosphorylation in the case of MMP-2.[57,58]

At the extracellular environment, various classes of enzymes, such as fibrinolytic protease plasmin, serine protease tissue kallikrein, thrombin, trypsin and other MMPs, regulate the proteolytic activation of MMPs,[59,60] which can be achieved via different mechanisms. Among them the most known is the release of the active MMP after cleavage of the pro-domain shielding the catalytic moiety. Alternatively, activation may occur through a conformational change destabilizing the propeptide domain, thus resulting in loss of cysteine switch–zinc interaction. Of note, zymogen activation may be controlled by chemical factors such as hypoxia and nitrogen monoxide (NO): both were shown to be activators of proMMP-2 and pro-MMP-9, respectively.[48,61]

Collagen turnover within connective tissues is impacted by MMPs activation mechanisms, alongside those responsible for MMPs inhibition. Prominent among these inhibitors are the tissue inhibitors of metalloproteinases (TIMPs), secreted by various cells and able to recognize selectively individual MMPs, irreversibly chelating the catalytic zinc ion though the insertion of a conserved anchor into the active site of the target enzyme.[62]

In a similar manner to MMPs, TIMPs' expression is influenced by diverse ECM factors.[52,63,64]

TIMPs have a very short half-life in vivo, hence they represents an additional level of local MMPs extracellular activity control, thus contributing to the key events maintaining ECM homeostasis.

1.3.2 Serine Proteases

Several members belonging to the superfamilies of serine proteases and of their inhibitors (Serpins) are endowed with a crucial role in ECM components turnover, in particular in the context of inflammation, chronic wounds and more generally in tissues homeostasis.

Serine proteases, comprising urokinase-type plasminogen activator (u-PA) and neutrophil elastase, are released by neutrophils in conditions of tissue damage. The former is thought to be involved in tissue remodeling and cell migration processes; the latter has a broad substrate specificity and is responsible for the digestion of ECM components, such as elastin, collagen and fibronectin, allowing pro-inflammatory cytokines and leucocytes infiltration through the lesion area.[65]

Although serine proteases activity represents an essential part of physiologic tissue remodeling, response to invading pathogens and wound healing patterns, their unchecked activity can have devastating destructive effects on tissues, causing the collapse of the whole organism.[59,66]

Therefore, a set of inhibiting enzymes, consisting in the “antiproteases shield”, regulates proteases activity limiting their effect within the area interested by the injury and hampering the inflammation spread in adjacent undamaged regions. Examples include α1-antitrypsin, α2-macroglobulin, α1-antichymotrypsin and plasminogen activator inhibitor (PAI-1).

Evidences suggest that their role in both plasma and interstitial fluids can be overwhelmed in chronic wounds and more generally in cases of recruitment of large numbers of inflammatory cells, implying the subsequent release of proteases into involved tissues.[65]

1.3.3 Focus on SERPINs superfamily

SERPINs is a superfamily of genes encoding for Serine Protease Inhibitors, a group of more than 500 homologous proteins implied in diverse functions such as blood coagulation, fibrinolysis, inflammation, cell differentiation and apoptosis, with a medium weight comprised between 40 and 50 kDa.

In human plasma, they represent approximately 2% of the total protein content, of which 70% is α-1-antitrypsin (A1AT) or serpin A1, the archetype of the family.

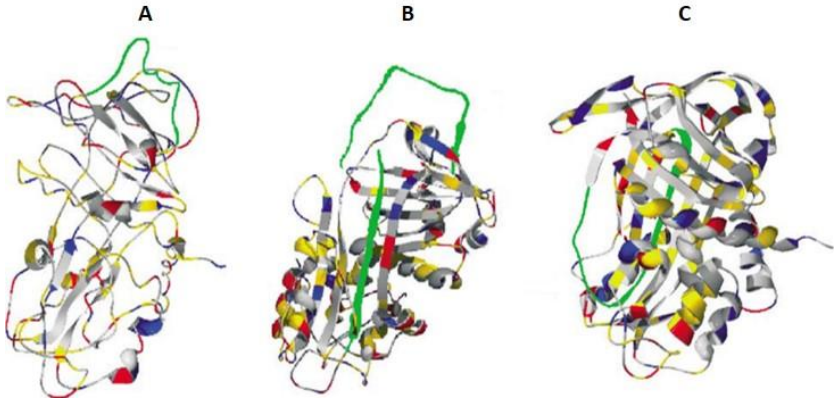


Figure 4. *Structural features of serpin proteins from various clades: (A) the uncleaved state of A1AT (adapted from H.K. Song, S.W. Suh, PDB n. 1KCT); (B) antithrombin III (adapted from R.W. Carrell, P.E. Stein, G. Fermi, M.R. Wardell, PDB n. 1ANT); (C) the inactive state of PAI-1 (adapted from T.J. Stout, H. Graham, D.I. Buckley, D.J. Matthews, PDB no. 1DVN). RCL is represented in green.*[67,68,69]

Serpins share secondary structure elements, since they are characterized by 3 β-sheets and 8 or 9 α-helices[70] and by the RCL (Reactive Centre Loop), a 20

amino acid motif near the C-terminus, comprising the scissile bond, which is cleaved by the target proteinase. Depending on RCL structure, the family members can be identified as inhibitory or non-inhibitory serpins.

The former are also referred to as “suicide substrate inhibitors”, considering that they irreversibly bind their target protease by kinetic trapping of the acyl enzyme intermediate of the normal substrate cleavage pathway.

Upon cleavage, the RCL moves to the opposite pole of the serpin, which undergo a distortion of the catalytic site, resulting in a ‘stressed to relaxed’ (S → R) conformational change (Figure 4).[71]

The inhibition pathway is capable of generating two reaction products: the covalent 1:1 serpin-protease complex, or the cleaved product of the reaction.

Serpins can be present in various conformations: native inhibitory with an exposed RCL, latent with a partially inserted RCL or non-inhibitory due to complex formation, cleavage, oxidation of reactive center residues or polymerization.[72]

1.4 Collagen related disorders: an overview

The following paragraphs are intended to roughly outline several diseases states related to collagen alterations, comprehending both acquired and genetic disorders, with a particular focus on those affecting collagen I.

Since its ubiquitous occurrence, clinical findings of pathologies involving collagen type I are very diverse and heterogeneous and can be considered as showcases that lead to understand matrix biology.

1.4.1 Ehlers-Danlos Syndrome

Ehlers-Danlos Syndrome (EDS) is the term employed to identify a very heterogeneous group of heritable pathologic conditions affecting soft connective tissues, classified in types and subtypes depending on degrees of soft connective tissues fragility, including the skin, ligaments, blood vessels and internal organs.[73]

In 50% of patients diagnosed with the classical form of EDS, mutations are reported in COL5A1 and COL5A1 genes, encoding for collagen type V α-chains.[74] Clinical signs are skin hyperextensibility, slow wound healing, easy bruising and joint hypermobility; these clinical findings are observed in both EDS types I and II, which form a continuum in classical EDS types, differing only in phenotypic severity.

The identification of mutations in genes encoding for collagen type V, resulting in classical EDS phenotype, highlighted an unexpected role of collagen type V, which appear significant in guiding unusual fibrils organization in the cornea and in nucleation of large fibrils in skin and other tissues as well.[19]

Vascular EDS (EDS type IV) is characterized by thin, translucent skin, acrogeria, easy bruising, and a peculiar facial appearance. Most importantly, affected individuals are at risk for arterial rupture, aneurysm or dissection, gastrointestinal perforation or rupture, and uterine rupture during pregnancy. This autosomal dominant condition is determined by mutations in gene COL3A1, encoding for the proα1(III) chain of collagen III, which is critical as a developmental scaffold in regulating arterial wall thickness and arterial diameter.[75] Most patients harbor a point mutation leading to the substitution of glycine within the triple helical region, some mutations result in an almost complete failure of fibroblasts to secrete type III procollagen, with an accumulation of the protein within the rough endoplasmic reticulum.

Whereas classical forms of EDS involve collagen type V, arthrochalasia types of EDS (types VIIA and B) are caused by the loss of exone 6 in mRNA encoding for α1 or α2 chain of procollagen I.[76] The principal clinical findings are joint hypermobility, bilateral hip dislocation, osteopenia and diverse skin involvement with slow wound healing and atrophic scars. This exon-skipping leads to the lack of N-telopeptide expression, a fragment linking N-propeptide to the triple helical domain of collagen. Since this portion contains key lysine residues critical for crosslinking and procollagen-I-N-proteinase cleavage site as well, its absence determines a deficient N-propeptide processing.

A subset of mutations in genes encoding for amino-terminal part of collagen type I polypeptides identify EDS subtypes, that may give rise to a superimposition with osteogenesis imperfecta (OI) phenotype, manifesting variable joint hypermobility and bone fragility.

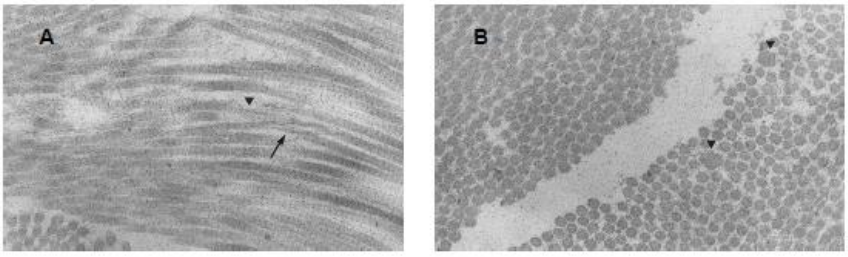


Figure 5. *In this patients a mutation Arg → Cys in the triple helical domain of proα1(I) chain determines untypical intermolecular disulphide bond formation and subsequent abnormal aggregation of collagen I molecules. (A) On the transverse section, the collagen fibrils show a distinct variability in diameter; (B) on the longitudinal section, the deposition of granulofilamentous material along the collagen fibrils is visible. In some foci, collagen fibrils have an unraveled and disorganized aspect (arrow).*[77]

Nuytinck et al. provided additional insights into genetic heterogeneity of EDS, describing a classical form affecting collagen type I, characterized by untypical propensity to arterial complications. They identified a C→T transition, which resulted in the substitution of a basic highly conserved arginine residue for an

uncharged cysteine residue at the Xaa position of the triplet G-Xaa-Yaa within the triple helical domain of proα1(I) chains.[77]

The cysteine residue projects on the outside of the triple helix, enabling the formation of intermolecular disulfide bridges between two adjacent collagen molecules. This results in the formation of molecular aggregates that cannot be secreted efficiently. Ultrastructural findings in patients affected are shown in Figure 5.

At other positions within the triple helical domain of proα1(I), substitutions of arginine by cysteine result in bone fragility with joint hypermobility.[78]

Finally, mutations leading to total absence of the pro-α2(I) chain, due to a homozygous or compound heterozygous COL1A2 mutation, can be associated with severe cardiac valvular anomalies necessitating surgery at young adult age.[79]

In other forms of EDS, collagens are not directly impacted; rather they are etiologically connected to defects in other proteins implied in ECM organization, such as proteoglycans and tenascin X, or mutations in genes encoding for enzymes involved in collagen biosynthesis and proper folding, i.e. prolyl *cis-trans* isomerase.[75]

No therapy is currently available for this disorder; main treatments are rather aimed to control and alleviate symptoms, such as anti-inflammatory drugs for joint pain and anti-hypertensive drugs, which are generally recommended to reduce pressure against the vessel walls in case of the vascular type EDS.

1.4.2 Osteogenesis Imperfecta

OI is the term used to address a heterogeneous group of disease states, whose common feature is the reduced amount of bone mass with frequent occurrence of fractures and deformities.[80]

The phenotype differs substantially, as shown in Figure 6, starting with patients that suffer only few fractures until puberty and those that die immediately after birth or after few days or weeks of life due to rib fractures and pulmonary hypoplasia.[81]

Other clinical findings comprehend disproportioned dwarfism, severe kyphoscoliosis with compression of lungs and internal organs,[82] deafness that may occur later in life and shortened life span. Blue sclera manifest in 50% of all patients and dentinogesis imperfecta (DI) with brittleness of teeth and increased risk of caries is prevalent only in some patients as well. Collagen defects at the skin levels are also reported, with abnormal mechanical properties.[83]

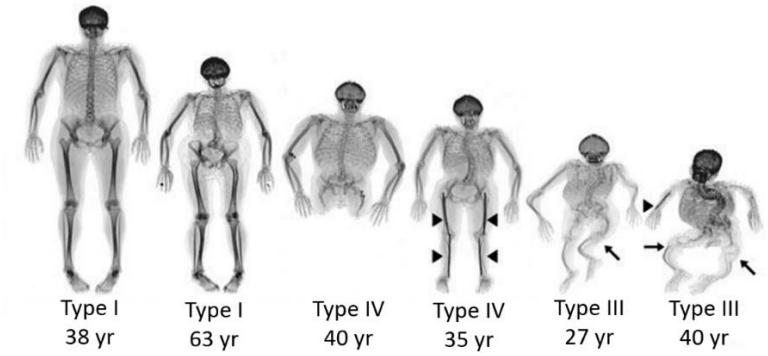


Figure 6. *Wide spectrum of skeletal manifestations in OI affected adults recruited for a clinical trial.*[84]

Such a various phenotype is not completely explained by the genetic background. Most patients (about 85%) carry heterozygous mutations in genes COL1A1 and COL2A2, encoding for the two α-chains of collagen type I, causing a dominant autosomal type of OI. These two genes may harbor two different types of mutations, causing distinct phenotypes.

OI type I is the mild form caused by nonsense mutations, which typically lead to haploinsufficiency, not influencing collagen architecture, yet causing lower-than-normal collagen levels.[85] In contrast, glycine substitutions result in the production of structurally impaired collagen and are accompanied by a more severe skeletal phenotype.

During the past decade, mutations in more than 10 additional genes have been identified; some recessive OI forms are caused by mutations of genes encoding for enzymes involved in post-translational modification of collagen, such as CRTAP, LEPRE1, PPIB and FKBP10.[86,87,88]

CRTAP, LEPRE1 and PPIB encode for three proteins implied in a cartilage associated complex responsible for the hydroxylation in 3-position of the ring of a proline residue sited at position 986 of the proα1(I) triple helix. Defects in prolyl 3-hydroxylation results in overmodifications of proα chains due to delayed assembly into trimers.

Instead, recessive mutations in FKBP10, which encodes the ER-resident cis-trans isomerase FKBP65, do not imply chain overmodifications of polypeptide chains.[89]

Recently, Christiansen et al. identified an autosomal missense mutation in SERPINH1, encoding the collagen chaperone-like protein hsp47, which results in a severe OI phenotype.[90] Their research underlined hsp47 critical role within procollagen molecules processing and its essential contribution in assisting the triple helix proper folding.

Besides surgical and orthopedic treatment, few opportunities of drug therapies are available for OI patients up to now.

In children and adolescents with forms from moderate to severe, bones are more likely to be influenced, since the skeleton is still developing. Therefore, elected therapies include antiresorptive drugs such as bisphosphonates, which

bind to hydroxyapatite crystals and limit physiologic bone resorption by osteoclasts, thus increasing bone mass and decreasing fracture rate.[91]

Instead, in adults affected by moderate forms of OI, osteoanabolic treatment based on teriparatide has been evaluated, in order to increase bone mineral density.

Recently, new promising antiresorptive drugs like denosumab became available. This antibody targets directly osteoclasts reducing their differentiation and activity.

1.4.3 Caffey's Disease

Infantile Cortical hyperostosis or Caffey's Disease is a pathologic condition affecting infants, manifested by cortical thickening of bones and swelling of contiguous soft tissues.[92]

When occurring during the first 5 months of life, it is a self-limiting benign syndrome, which usually undergoes spontaneous remission.

Figure 7. *Frontal radiograph of the whole skeleton of a fetus at 30 weeks gestation affected by lethal perinatal Caffey's diseases. Long bones are short and slightly angulated.*[93]

During the last decade, a lethal perinatal form was described,[93] and three cases from unrelated families were reported, displaying severe deformities of skull and short angulated long bones (Figure 7). Besides, abnormalities of the skin were also observed, with generalized edema and fragility.

In all cases, a missense mutation was found on gene COL1A1, encoding for α1-chain of collagen type I, predominant component of the organic part in bones and teeth. In particular, the three individuals were heterozygous for the identical mutation, a 3040C→T transition, resulting in the substitution of an arginine by a cysteine residue at position 836 (R 836 C), within the helical domain.

1.4.4 Osteoporosis

Osteoporosis is the most prevalent condition associated with aging, whose main clinical signs are low bone mass and micro-architectural deterioration that lead to increased risk of fractures.

Traditional classification of osteoporosis identifies two main categories: the primary and the secondary forms. The former is related to aging and decreased gonadal functions i.e. reduced levels of estrogens, whereas the latter is caused by other chronic illnesses, such as those leading to disuse and hence lack of mechanical stimulation.[94]

Loss of collagen mass and biochemical changes in collagen suprastructures organization are significantly involved in the pathogenesis of the disease; in particular, little changes were reported in the proportion of type III collagen, but in some cases there was a significant loss of the type VI. However, the major differences were observed in the post-translational modifications, namely, in the stabilizing cross-links and the hydroxylation of procollagen units.[95]

Of note, the hypotheses arose that the loss in bone density in affected individuals corresponds to fluctuations in collagen amounts at the skin level and as disserted previously, similar correlations were found in EDS, OI and Caffey's disease patients as well, suggesting a coherence in the mechanisms implied.[96]

Current therapeutic strategies include prevention methods based on diet integration with vitamin D supplement and physical activity, and pharmacologic treatment with bisphosphonates, recombinant PTH and selective estrogen receptor modulators.[97]

1.4.5 Osteoarthritis

Osteoarthritis (OA) is a degenerative joint disease progressively leading to the loss of joint functions causing chronic pain, disability and life quality impairment. In 50% of patients, cartilage affection is accompanied by histological changes of the synovium, which in most cases manifest at early stages of the disease with localized areas of synovitis.

OA predominantly occurs in industrialized nations and clearly offers a dramatic picture of a condition initiated by the imbalance in collagen homeostasis.

To date, there are no current interventions proven to restore cartilage or reduce the disease progress and the etiology is not known, though evidences suggest a multifactorial profile;[98] risk factors in initiating the injury process at the joints level comprehend age, gender, trauma, overuse, genetics and obesity.

The gradual thinning of the articular cartilage is the result of a decoupling of the normal degenerative and regenerative/repair processes, driven by cytokine cascades and the production of inflammatory mediators. Among them, IL-1β and TNF-α in turn decrease anabolic collagen synthesis and

increase catabolism (MMPs) along with the levels of other inflammatory mediators such as IL-8, IL-6, prostaglandin E2 and NO.[99]

Of note, NO contributes to articular cartilage damage and plays multiple roles, promoting cartilage degradation by inhibition of collagen and proteoglycan synthesis, MMP activation, and increased susceptibility to other oxidant injury.

In parallel to NO, reactive oxygen species (ROS) contribute to chondrocyte apoptosis, catabolic processes and matrix degradation, thus inducing the premature senescence of joints, as demonstrated in chondrocytes from OA patients by histological signs such as shortened telomeres, increased levels of β-galactosidase and decreased ATP production from mitochondrial dysfunction.[100,101]

1.4.6 Skin aging: pathophysiologic aspects

Skin aging represents a common condition characterized by alteration in collagen metabolism and by changing in both ECM topography and dermis histology, affecting connective tissue functions of scaffolding and supporting tissue integrity.[102]

As a matter of fact, both intrinsic and photo-induced skin aging outline conditions in which fibrillar collagens III and I are still functional but their amount per cell is remarkably reduced,[103,104,105] causing wrinkles formation and elasticity loss.

Clinical appearance of naturally aged skin and photo-aged skin differs substantially, since the former is smooth, pale and finely wrinkled, the latter is in contrast coarsely wrinkled and associated with dyspigmentation and telangiectasia, due to drastic modification of the dermis architecture.

In fact, in photodamaged skin, the superimposition of environmental factors, in particular UV exposure, results in emphasized collagen loss with fibers degeneration. Chronologically aged skin is generally thinner, the protection barrier is not compromised, yet alterations in fibers alignment are observed with unraveled collagen bundles and reduced space between fibers.[105]

On one hand collagen-degrading matrix metalloproteases (MMPs) are strictly involved in skin aging pathways, since their up-regulation is either acutely photo-induced[106,107] or gradually occurs in natural aging,[108,109] causing in both cases lower-than-normal collagen levels. On the other hand, failure to replace damaged collagen with newly synthesized material represents also a critical aspect in the pathophysiology of skin aging: in photo-aged skin, high molecular weight fragments obtained from increased collagen degradation impede efficient collagen biosynthesis, not providing the adequate basis to resident fibroblasts, necessary to support properly the mechanical tension required.[110,111]

Similarly, in chronologically aged skin, the loss of intact collagen fibers results in a lower mechanical stimulation, likely affecting procollagen biosynthesis.[105]

2. DESIGN OF COLLAGEN TURNOVER MODULATORS

Considering collagen type I predominance all over the body and the involvement in many diverse pathologic and physiopathologic conditions, the regulation of its biosynthesis and degradation pathways appears as a very appealing perspective in light of the complexity of these processes, both offering various levels of intervention.

Hence, the main goal of our research program is the search of newly designed molecules capable of influencing either collagen production or digestion. The novel lead compounds could allow us enlightening still unclear aspects in collagen turnover patterns and most importantly could represent the starting point for the development of more drug-like molecules for future insertion in new conception pharmaceutical formulations in innovative treatment of collagen related diseases.

In the search of new lead molecules, we focused on the pool of peptides, which are desirable candidate drugs, characterized by high safety and selectivity.

Indeed, peptide-based drugs generally display an excellent pharmacologic profile in humans, exhibiting low toxicity and high specificity to the designated target; in a physiologic setting, peptides tightly bind to a specific receptor, thus rarely interfering with other cell compartments.[112,113]

Besides, peptide as pharmaceutical leads are susceptible of a number of chemical modifications, such as cyclization, conjugation to different tags and insertion of non-natural amino acids, thus suggesting numerous opportunities to control either their efficacy to the designated clinical target or their pharmacokinetic properties, i.e. increasing the resistance to protease digestion.

AIM OF THE WORK

We organized our work in three main steps: design, synthesis and biological activity evaluation of collagen turnover modulators.
In particular, two parallel approaches were evaluated and developed:

PART A. The design of a Template Assembled Synthetic Protein (TASP)[114] mimicking the chaperone moiety of hsp47 or serpin H1, one of the principal collagen superhelix stabilizer (paragraph 1.2.1).

PART B. The synthesis and in vitro screening of short chain linear peptides from A1AT or serpin A1 C-terminal portion, endowed with putative enhancer effect on collagen production.

Both studies directly involve members of the Serpins superfamily (paragraph 1.3.3): hsp47 and serpin A1, yet they are related to collagen turnover in very different fashion.

3. PART A. Mimicking hsp47 chaperone function via TASP approach

3.1 Hsp47 role in collagen biosynthetic pathway

Hsp47 or serpin H1 is a 47 kDa heat shock protein belonging to the Serpins superfamily; it contains the retrieval signal RDEL sequence at its C-termini and thus resides in the ER lumen.[115,116]

As previously introduced in paragraph 1.2.1 while discussing about post-translational modification steps implied in collagen biosynthesis, hsp47 is not provided with a protease inhibitory activity but it is rather endowed with a crucial molecular chaperone role in procollagen molecules processing.

Briefly, serpin H1 can bind selectively procollagen in the ER lumen,[115] assisting triple helix proper folding and transport from the ER to the Golgi apparatus, where procollagen molecules are finally released in the extracellular space. The interaction is transient and the procollagen units dissociate from hsp47 upon or just before reaching the *cis*-Golgi.[117] Procollagen release may be caused by pH lowering while proceeding toward the *cis*-Golgi or by hsp47-procollagen complex dilution at the Golgi apparatus level, since this cell compartment harbors only a few free hsp47 molecules.[118,119]

Unlike other molecular chaperones directly involved in collagen biosynthesis steps, serpin H1 does not possess other known client proteins and, in striking contrast to other chaperones, it shows strong preference in recognizing the correct folded procollagen triple helix, while the interaction with the corresponding single-chain polypeptides is negligible.[120] Since hsp47 shows such a conformational preference, it may be considered the best candidate as the correctly folded triple helix stabilizer.[24]

Consistently with this role as the main folded superhelix stabilizer, hsp47 increases triple helix resistance toward proteases digestion[121,122] and inhibits procollagen molecules lateral aggregation[123]; in fact the addition of hsp47 prevents self-association of the triple helix form of type I collagen *in vitro* at pH values over than 6.[124]

3.2 Fields of interest and aim

In 2000 Nagai et al. succeeded in disrupting the hsp47 gene in mice: although heterozygotic mice (*hsp47*+/–) showed no evident phenotype, homozygotic mice (*hsp47*–/–) showed an embryonic lethal phenotype, thus underlining the essential role of glycoprotein serpin H1 within collagen biosynthetic pathway.[125]

In particular, *hsp47*-knockout mouse embryos were found deficient in the maturation of collagen types I and IV, and collagen triple helices formed in the absence of hsp47 showed increased susceptibility to protease digestion. Accordingly, there is supportive evidence that hsp47 constitutive expression is correlated to major collagen types expression: cells producing high levels of hsp47, produce high levels of collagen and *vice versa*; this positive correlation is regulated with transcriptional mechanisms.[126,127,128]

On one hand, hsp47 higher levels in fibrotic diseases correlates to collagen overproduction and abnormal accumulation,[129,130] accompanied by lowering or loss of functionality in related organs; on the other hand reduced functional levels of hsp47 were reported in severe recessive forms of OI.[131,132]

Of note, there are evidences that hsp47 expression also correlates with ageing. In vitro studies by Miyaishi et al. revealed an alteration in hsp47 expression in cultured dermal fibroblasts obtained from old persons compared with cells from young subjects: in particular, the lower levels of hsp47 are regulated by

transcriptional mechanism and may cause procollagen molecules retention in the ER, due to defects in transport to the Golgi apparatus.[133]

Interestingly, significantly elevated levels of autoantibodies directed to hsp47 were found in sera of patients affected by connective tissues autoimmune diseases, including rheumatoid arthritis, systemic lupus erythematosus, Sjögren's syndrome, and mixed connective tissue disease (MCTD).[134]

The critical role in assisting collagen maturation also implies a relevance within cartilage formation and endochondral ossification, as elucidated in a recent work by Masago et al., in which they reported that null mutant mice exhibited severe generalized chondrodysplasia and bone deformities with lower levels of type II and type XI collagen. At the level of tissue organization, they observed an accumulation of misaligned type I collagen molecules in the intervertebral discs and a substantial decrease in type II collagen fibers.[135]

Besides, the prevention of improper fibers aggregation could be an interesting goal in approaching the treatment of those pathologic conditions in which premature lateral aggregation occurs, such as some types of EDS, as elucidated in paragraph 1.4.1.[77]

Previous observations led us to identify hsp47 as a selective target to manipulate collagen production in the context of collagenopathies with altered collagen levels and organization, such as some types of EDS and OI (paragraphs 1.4.1, 1.4.2). Our main intent should be elucidating serpin H1 role in collagen related diseases, through the monitoring of collagen production either in normal or pathological situations.

In order to fulfill this purpose, we should avail a structural model able to mimic the chaperone function of the protein, detached from the rest of the molecule. The newly designed pharmacological tool should be employed primarily in in vitro assays in order to investigate the impact on collagen levels fluctuations,

and secondarily could be proposed as a putative starting point for further inquiry about innovative therapeutic approaches development.

3.3 Serpin H1 structural inquiry

Serpins are single chain proteins containing a conserved domain structure of 370-390 residues and a highly conserved tertiary structure made up by eight/nine α-helices and three β-sheets. Hsp47 exhibits the typical serpin fold, consisting of three β-sheets and nine α-helices.

There is no significant rearrangement upon collagen binding;[119] serpin H1 ability to bind specifically collagen and procollagen (types I to V) resides on a long deep cleft whose base is formed by a β-sheet termed and sides formed by helices termed hA and hG/hH, according to Huber & Carrell labelling system.[119,136,137]

The main interaction directly involves a strictly conserved aspartic residue (Asp385) within β-sheet B that binds through a salt bridge an arginine residue sited at the Y-position of the repetitive collagen triplet Gly-Yaa-Xaa. Procollagen binding is pH-dependent, thus emphasizing the electrostatic nature of hsp47 interaction with its own unique client: mutation of Asp385 to asparagine drastically weakens the interaction,[119] probably due to unfavorable charges redistribution at the interface.

Widmer et al. inquired on hsp47-collagen interaction, availing of a synthetic homotrimeric collagen model. They found that hsp47 recognizes all three collagen strands, through the establishment of hydrophobic interactions, especially involving Leu381 and Tyr383 and water-mediated contacts with Ser305 and Ala303 residues, which interact with threonine side chain in collagen models (Figure 8).[119]

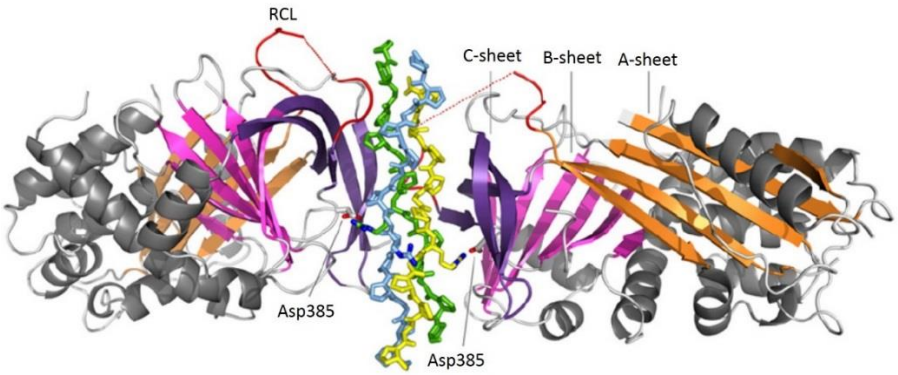


Figure 8. *Overall structure of hsp47 in complex with homotrimeric collagen model. Hsp47 is shown in cartoon representation, and the collagen model chains are depicted as sticks. α-Helices are drawn in gray. β-Sheet A is shown in orange, sheet B in magenta, and sheet C in violet. The segment corresponding to the RCL of inhibitory serpins is drawn in red. Asp385 is shown as sticks. The leading strand of the collagen triple helix is drawn in yellow, the middle strand in blue, and the trailing strand in green.*[119]

Basing on data reported in the literature[119,137] and aiming to design a structural model able to mimic the chaperone function of hsp47, we focused on the protein area mainly implied in procollagen binding and processing, comprising three strands of β-sheet B as the base of collagen binding cleft.

Since maintaining β-sheet B tridimensional structure is clearly a crucial aspect of procollagen binding and processing, it was necessary to pre-organize the secondary structure elements resembling their folding within the native protein. To this aim, the design of an artificial molecular scaffold was planned according to the Template Assemble Synthetic Protein (TASP) Approach.[114]

3.4 Protein design via the TASP approach

Since the theoretical introduction in late 80s,[114,138] protein design and synthesis has certainly been retained one of the most challenging goals in biomimetic chemistry, yet it could be considered an interesting perspective on the way of

introducing novel effective instruments to improve our ability to understand the main criteria of protein folding and tridimensional organization, providing specific sequence-to-folding information. Thus, the design of novel proteins with predetermined structures and properties can be a useful tool to test and expand our ability to understand natural proteins, their folding pathways, thermodynamic stability and catalytic properties.[139]

Besides, *de novo* design and chemical synthesis of proteins, unitedly to artificial structure, which mimic them, is a central strategy for accessing proteins with new functions, suited to many different applications, taking advantage of the chemical properties inherent to synthetic molecules.

Template Assembled Synthetic Proteins (TASPs), introduced by Mutter[114,140], combine structural features of natural proteins with synthetic elements. The concept of TASP approach is based on the intent to overcome the complexity of protein folding by inserting a topological 'built in' device properly designed to assist and direct the folding of covalently bound peptide blocks selected from the primary structure of the protein. The artificial template pre-organizes secondary structure elements, guiding the formation of a tertiary structure, similar to the native protein tridimensional outcome.

Protein-protein interactions, one of the predominant manner in which protein functions are controlled in nature, cause the burial of a protein large surface, previously exposed to the solvent. Since the mimicry of the entire protein surfaces involved in the interaction would be a challenging goal, the TASP design have been simplified by focusing on the so called "hot spots". These protein regions or groups of residues confer the majority of the free energy necessary for protein-protein association and are usually displayed on highly ordered secondary structural elements.[141]

Based on these concepts, the idea has emerged of projecting small molecules that mimic these ordered secondary structure domains and exhibit functionality in a similar fashion (Figure 9).[142]

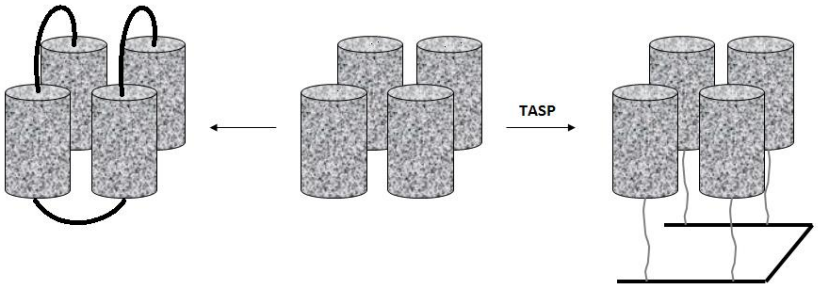


Figure 9. *Strategies for protein de novo design. The construction of tertiary structures from secondary-structure elements can occur by two different approaches, as shown here for a four-helix-bundle protein. In order to reach the designed tetrameric structure, which arises through self-association of the amphiphilic helix blocks, the helices may be linked by loops to yield a polypeptide chain with the linear connectivity found in natural proteins. In contrast, the same packing arrangement may be reached by grafting the individual helix blocks onto a carrier molecule (template).*[138]

Secondary structure elements in TASPs mainly comprise α-helix bundles, β-turn and β-sheets.

The synthesis of artificial α-helices have been pursued either in the past[143,144] or in recent works:[145] synthetic strategies employed to favor α-helices proper formation include peptoids, β-peptides and stapled peptides.

First attempts to build up four-helix bundle proteins were made by the groups of DeGrado and Richardson, following two complementary approaches: the former based on amino acids propensities, the latter intended to maximize the number of amino acids types used, in order to stress the similarity to the natural sequence (Figure 10).[143,146]

Their common intent was not to predict what 3-D structure a particular sequence would adopt, but rather identify a sequence that was compatible with a particular folding.[147]

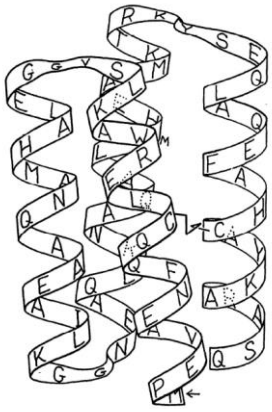


Figure 10. *Schematic illustration showing both the sequence and proposed three-dimensional structure of the designed four-helix bundle protein, Felix.*[146]

Recently, Malik et al. developed a three-helix bundle carboproteins on deoxyhexopyranosides: they speculated that the structural information received by the template could markedly facilitate protein folding.[145]

While the obtainment of α-helix bundles is based on the modular feature of each α-helix unit, the design of β-sheet proteins is a more challenging task.[148,149] In fact, a α-helix can exist in isolation, thus representing a building block for subsequent covalent attach on a synthetic template; in contrast, an isolated β-strand is not stable, rendering β-structure less modular and inherently more difficult to design.

In particular, this contrast stems from the fundamental difference in the hydrogen bonding network of the two types of secondary structures: in the α-helix backbone hydrogen bonding is intrasegmental, connecting the carbonyl group of residue "*i*" to the amide nitrogen of residue "*i + 4*". Thus, a α-helix motif can be self-contained, satisfying most of its hydrogen bonding needs without any help from a partner.

The β- strand nature is more gregarious, since hydrogen bonding involves groups located in neighboring chains, with several practical and theoretical

implications for protein design, such as the tendency to form aggregates and precipitate out of solution.

Coherently, in α-helices based TASPs, the interaction of each modular helix is based on hydrophobic contacts of non-polar faces, while artificial β-sheet design requires a balance between favorable and unfavorable contacts between β-strands, in order to prevent undesired aggregation and precipitation.[146]

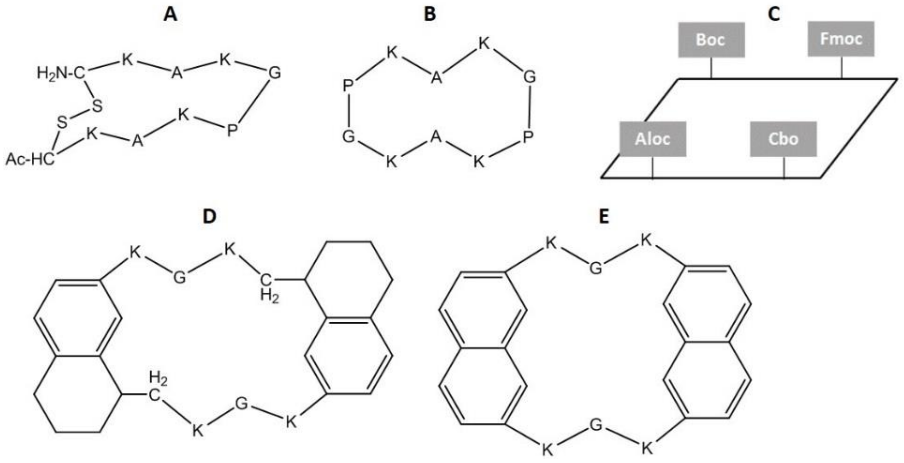


Figure 11. *Topological templates for use in TASP design: (A, B) cyclic peptides; (D,E) cyclic peptides containing turn-mimetics; orthogonal protection groups (VI) for the ε-amino-lysine groups.*[151]

In addition to α-helices and β-sheets, organized turn motifs are among the most well-studied secondary structural features of proteins. These turns generally involve three to five amino acids, and impose well-defined Φ and Ψ angles not otherwise available to helices, sheets or extended peptide sequences.[142]

Of note, during the last two decades, several scientific groups, including Nowick et al. succeded in the rational design and efficient synthesis of artificial β-turn and β-sheet motifs.[150]

One of the numerous implications of TASPs include the possibility to dispose on a synthetic construct functional groups sited at spatially defined position within the tertiary structure of a protein, thus mimicking the bioactive conformation.[151]

Based on similar concept, Howard et al. created a novel adaptor protein by fusing protein binding domains from different signaling networks.[152] These fusion proteins appear to become a powerful tool in molecular recognition studies and drug development, since they allow separating different activities of a signaling protein and investigate systematically each function.

Examples of molecular scaffold structures are reported in Figure 11.

3.4.1 Selection of peptide blocks from β-sheet B

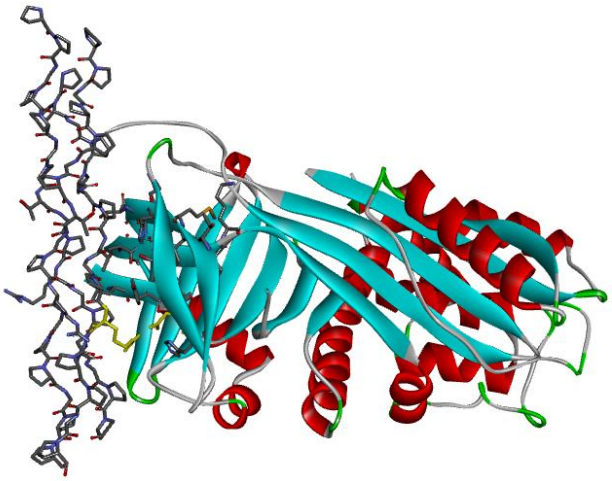


Figure 12. *Representation on DS 3.5 visualizer on the main area of interaction between serpin H1 and homotrimeric collagen model (ID on PDB: 3ZHA). Serpin H1 is visualized in cartoon representation, collagen model and β-sheet B main area of interaction are visualized in sticks. Asp-385 and recognized collagen arginine are colored in yellow.*[119]

Detrimental factors for artificial secondary structure folding comprise the small volume-to-surface ratio, which limits the number of hydrophobic interactions

and the lack of *long-range* interactions that significantly determine the final conformation of a polypeptide chain. Hence, during *de novo* design of TASPs, we should favor the arise of local interactions such as hydrogen-bonding, which may act as seeds of folding.[153] In fact, due to their ability to form early on in secondary structure elements, they can direct rapid and efficient chain folding. Coherently, the amphiphilicity of secondary structure building blocks represents the driving force for TASPs self-assembly,[154,155,156] creating an energetically favorable situation with a protein-like hydrophobic core and hydrophilic surface.

According to the TASP approach, amphiphilic peptide blocks selected from the functional portion of hsp47 should be covalently attached on an appropriately chosen template, which could serve as a built-in device for intramolecular folding.[157,158]

Referring to the Protein Data Bank (PDB) serpin H1 model crystalized in complex with synthetic homotrimeric collagen,[119] we explored the protein area at the interface of the interaction with the ligand (Figure 8) and selected three peptide blocks corresponding to three β-strands of sheet B,[136] mainly involved in procollagen binding, as reported in Table 3.

β-strand	*Protein portion*	*Orientation*	*Sequence*
1	Hsp47 (380-386)	*N-term → C-term*	KLFYADH
2	Hsp47 (302-308)	*N-term → C-term*	VAISLPK
3	Hsp47 (235-241)	*C-term ← N-term*	TMMHRTG

Table 3. *Peptide blocks selection*

3.4.2 Artificial β-sheet folding: main criteria

Between 1993 and 1997, several research groups conducted studies on β-sheet folding either in water or in non-polar solvents, systematically varying amino acids within small β-sheet containing proteins or small artificial β-sheets. They quantified the effect of each mutation upon thermodynamic stability of the proteins or using ^{1}H NMR and IR spectroscopy to measure the relative degrees of intramolecular hydrogen bonding within parallel β-sheet models.[159,160,161]

In 1997 Nowick et al. reported their findings on four different amino acids propensity to form parallel β-sheets (L, V, A, G), employing an artificial model which consists of two amino acid groups attached to a diamine backbone by way of two urea groups. The urea groups form a hydrogen-bonded turn structure,[162] which juxtaposes two amino acids and orients them to promote the formation of a parallel β-sheet. The trends in propensities to form β-sheet observed for the four amino acids (L > V > A > G) are largely similar to previous findings.[159,160]

Only one year later Kortemme et al. designed a small library of three-stranded models based on natural aminoacids, with Asn-Gly turns; the model which showed the higher degree of β-sheet nucleation in water, named *Betanova*, was hallmarked by hydrophobic side chained amino acids on the surface of the sheet, not hindered by β-branched aminoacids on the adjacent strand.[163]

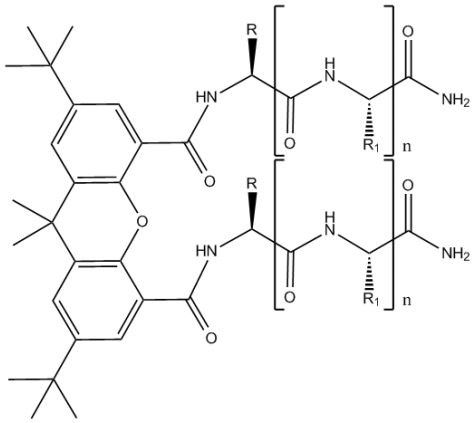
Following these criteria for artificial β-sheet folding, we modified in part the amino acidic sequences of native peptides. Since hydrophobic interactions represent the principal driving force for secondary structure nucleation in polar solvents, non-polar side chains sited at the surface of the sheet can favor the rising of H-bond establishment, in absence of β-branched amino acids in the adjacent strand. Hence, a leucine residue has opportunely replaced Ile304 in strand 2, adjacent to Tyr383 in strand 3.

Besides, the presence of a glycine residue can be unfavorable for β-sheet nucleation, due to the lack of one additional H-bond; thus Gly241 is opportunely replaced, as later explained.

3.4.3 Molecular modeling assisted choice of a proper scaffold

We designed two kinds of structural models: three-stranded models and simplified 2-stranded models (β-hairpins), both based on different templates.

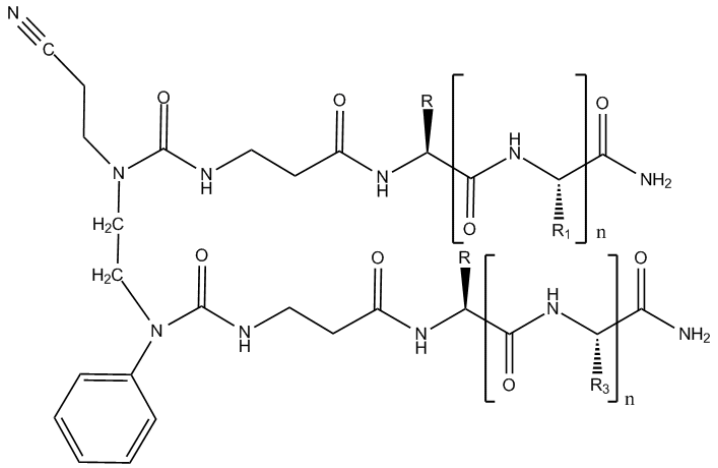
- 2-stranded models scaffolds



TASP 1

An artificial U-shaped xantene scaffold supports the β-hairpin 1,[164] anchoring strand 1 and strand 2 in a parallel fashion.

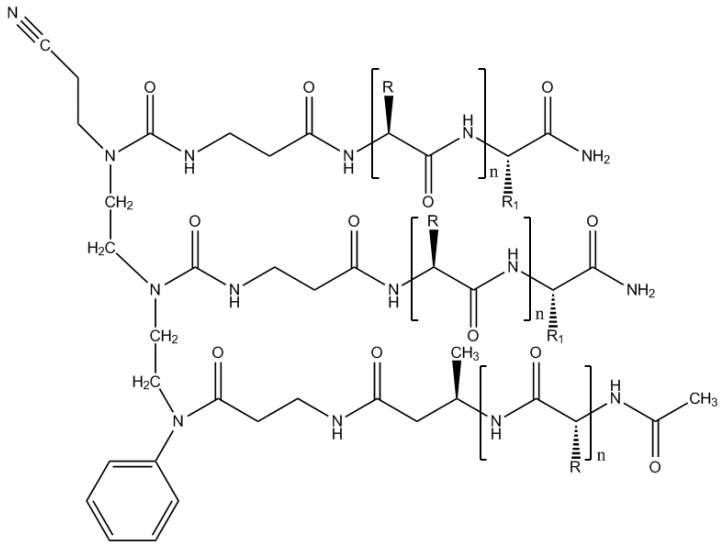
PART A



TASP 2

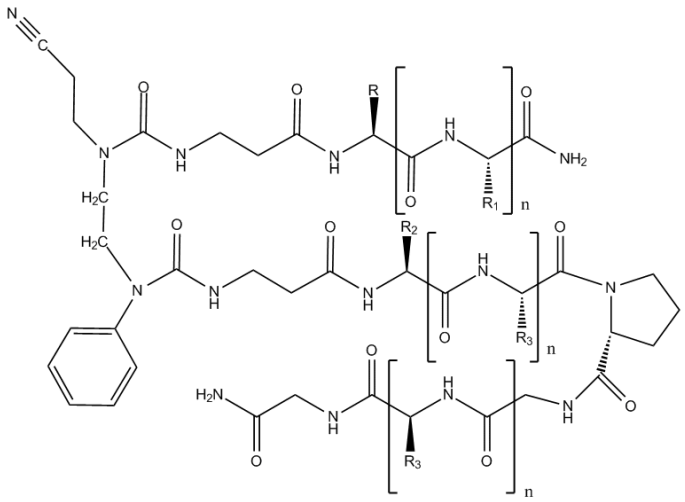
In β-hairpin 2, strand 1 and strand 2 are covalently attached to an oligourea scaffold proposed by Nowick et al.[161] through two β-alanine spacers; the phenyl group on the diurea molecular scaffold allows the two peptide blocks to be held in proximity.

- <u>3-stranded models scaffolds</u>



TASP 3

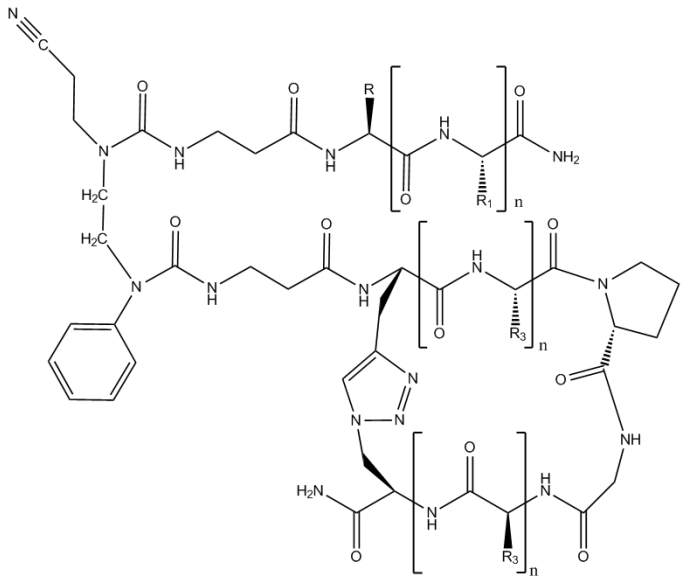
TASP 3 consists of an oligourea scaffold anchoring the three selected hsp47 peptide blocks, through β-alanine spacers: strand 1 and strand 2 are parallel, while the third strand is oriented in an antiparallel manner. In strand 3, we substituted Gly241 with a 3-aminobutyric acid unit, in order to increase hydrogen-bonding tendency, thus favoring β-sheet nucleation.



TASP 4

In designing TASP 4 we also employed an oligourea scaffold, anchoring strand 1 and strand 2 to the template through two β-alanine spacers, and linking the third strand to strand 2 by means an artificial DPro-Gly β-turn, proposed by Gellmann et al.[165,166,167]

We also designed two models identical to TASP 4, except for the insertion of different artificial β-turns: Asn-Gly (proposed by Kortemme at al. in *Betanova*[163]) and δ-linked ornitine, successfully inserted by Nowick et al. in β-sheet dimers that fold in water.[168]



TASP 5

Starting from TASP 4, we tried to decrease strand 3 mobility by anchoring it to strand 2 through a triazole bridge, that we could obtain synthetically, by the insertion of two clickable units. An artificial azido acid could replace Gly241 in strand 3, while Val302 could be substituted with a synthetic amino acid bearing an alkyne group in side chain; Cu(I) would be later employed as a catalyst within a click reaction.[169] The introduction of a triazole unit may produce several advantages, since the triazole can act as H-bond donor/acceptor, it is resistant to enzymatic degradation, hydrolysis and oxidation and early forms in mild conditions. Besides, the interatomic distances between Val302 and Gly241 C_α measured on the crystalized protein backbone are very similar to that found between substituents groups in 1-4 triazoles. (≈ 5 Å).[170,171]

With the support of molecular modeling advanced techniques, we preliminarily screened each designed TASP, by means Root Mean Square Deviation (RMSD) calculations and docking energies evaluations, performed using AMBER 9 (Assisted Model Building with Energy Refinement).[172]

3.4.4 RMSD calculations

Root mean square deviation (RMSD) calculations represent a useful tool to estimate the degree of spatial superimposition of the TASP candidate with the native protein tertiary structure.

In particular, RMSD values were calculated for each TASP, taking as a reference PDB serpin H1 crystallized in complex with homotrimeric collagen model.[119]

The atomic spatial coordinates of each TASP were compared to the correspondent in serpin H1 PDB model, previously selecting atoms either of the backbone or of the side chains from key amino acids for procollagen triple helix recognition and binding.

The results here reported for most promising TASP candidates are expressed in terms of potential energy and RMSD values vs time, during molecular dynamic simulations:

- **TASP 2**

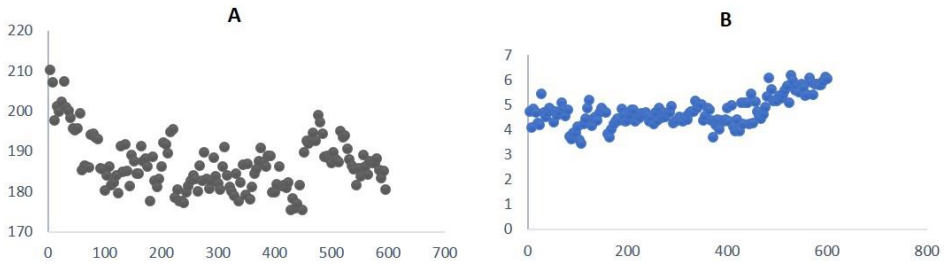


Figure 13. *(A) Potential energy values versus time (ps) during molecular dynamics simulations for TASP 2;) RMSD values versus time (ps) in molecular dynamics simulations for TASP 2; in RMSD calculations we referred to PDB file ID 3ZHA.*

- **TASP 3**

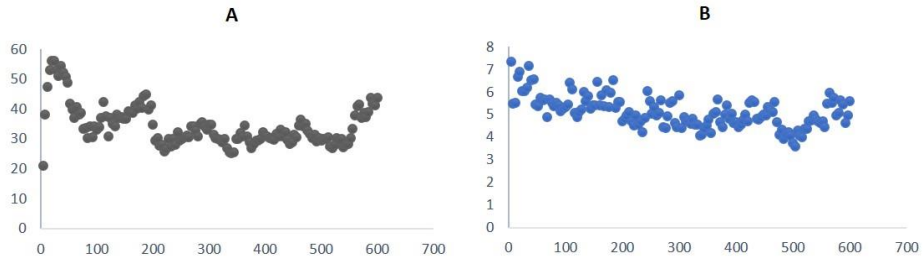


Figure 14. *(A) Potential energy values versus time (ps) during molecular dynamics simulations for TASP 3; b) RMSD values versus time (ps) during molecular dynamics simulations for TASP 3. Potential energy and RMSD fluctuations displays similar trends during molecular dynamics simulations: lower RMSD values correspond to higher stability conformations.*

- **TASP 5**

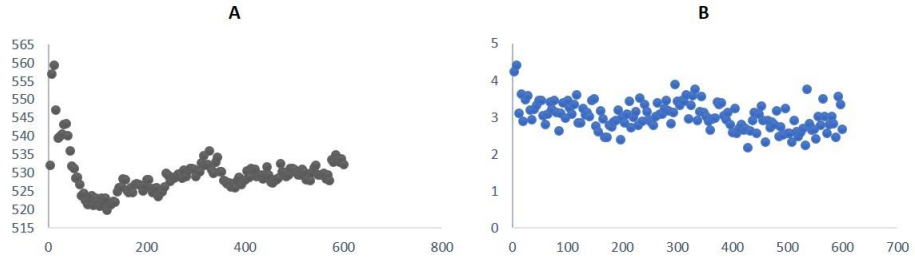


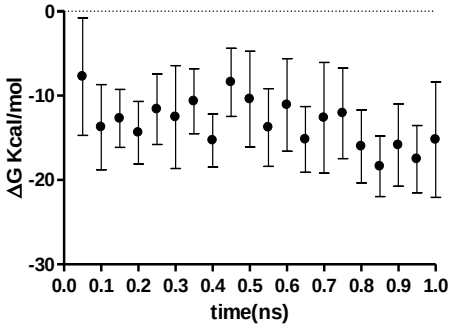
Figure 15. *(A) Potential energy values versus time (ps) during molecular dynamics simulations for TASP 5; b) RMSD values versus time (ps) during molecular dynamics simulations for TASP 5. The "clicked" unit anchoring strand 2 to strand 3 let us to increase the model rigidity, thus lowering RMSD values (in comparison to TASP 4, data not shown) even without any restraint applied.*

3.4.5 Docking energy evaluation

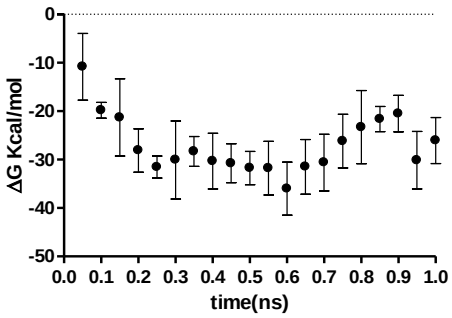
Free energy values were calculated in molecular dynamic simulations employing the module SANDER of AMBER 9. The aim was to compare the stability of each TASP candidate in complex with homotrimeric collagen model, thus establishing the tendency to bind and processing procollagen molecules

in a biological setting. Here we report the results in terms of docking energies for TASPs 2, 3, 5.

TASP 2 (A)



TASP 3 (B)



TASP 5 (C)

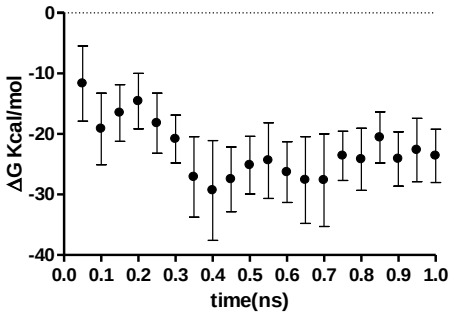


Figure 16. *(A, B, C) Docking energy versus time (ps) in molecular dynamics simulations for most promising TASPs. We can notice an increase in complex stability during the dynamics for TASP 3 and TASP 5.*

As a result of molecular modeling assisted screening of newly designed TASP candidates, we can finally deduce that a 3-stranded model (i.e. TASP 3 and TASP 5) is more desirable over a β-hairpin, due to the higher chance for the former to efficiently bind and process procollagen molecules.

Besides, an increase in strands number from two to three produces in general higher stability in antiparallel β-sheets.[167]

In particular we selected TASP 5 as the most promising model, which displays a higher degree of superimposition to crystallized serpin H1, probably due to the introduction of a clickable unit anchoring strands 2 and 3.

3.4.6 Synthesis protocol proposal for TASP 5

The synthesis of the selected TASP was programmed step by step and is currently on going. All the reactions performed and just scheduled are reported in this paragraph, along with the difficulties encountered.

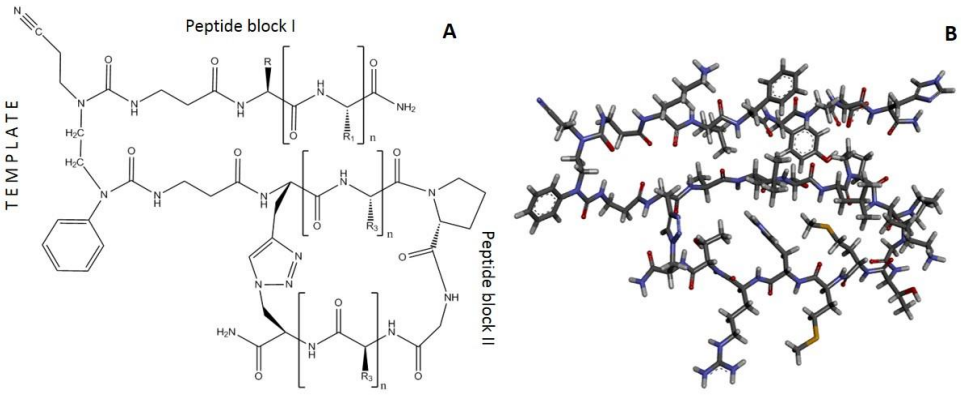


Figure 17. *(A) TASP 5 structure; Block I corresponds to serpin H1 (380-386), peptide block II comprises both fragments serpin H1 (302-308) and serpin H1 (235-241), namely strand 2 and 3. Two clickable units replaced Gly241 and Val302 in order to allow the formation of a triazole bridge between strands 2 and 3; (B) TASP 5 stick representation on DS 3.5 visualizer.*

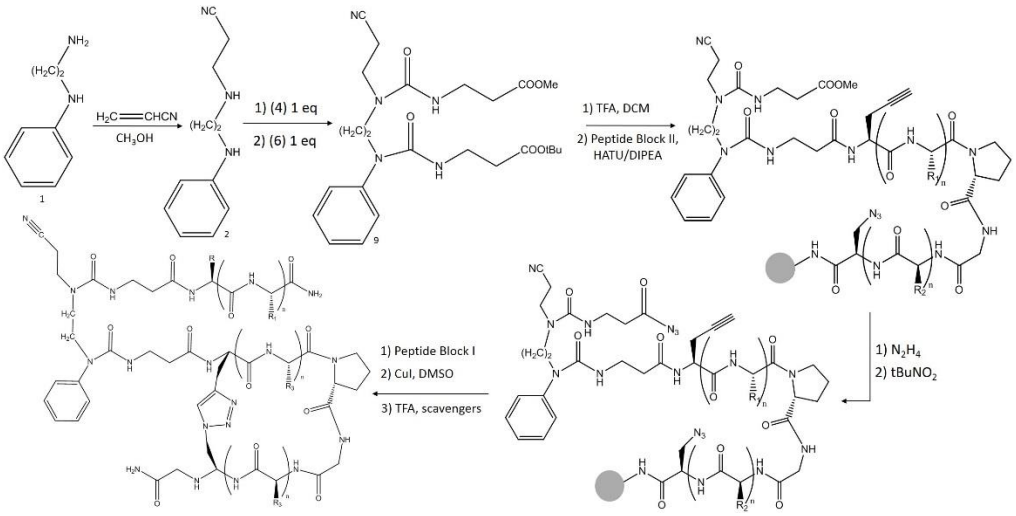


Figure 18. *Synthesis scheme scheduled for TASP 5.*

Azide aminoacid and alkyne aminoacid were purchased; such unnaturally modified aminoacids were inserted during Fmoc/*t*Bu SPPS (Fluorenylmethyloxycarbonyl/*tert*-Butyl solid phase peptide synthesis) of peptide block II (comprehending the artificial DPro-Gly β-turn[165]). Peptide blocks I and II were synthesized starting from a Rink-Amide resin and introducing in peptide block I a Lys380 residue, bearing on ε-amine the N-(1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl) (-Dde) protecting group. The diamine molecular scaffold 2 was obtained via Michael addition starting from commercial N-phenylethylenediamine;[162]

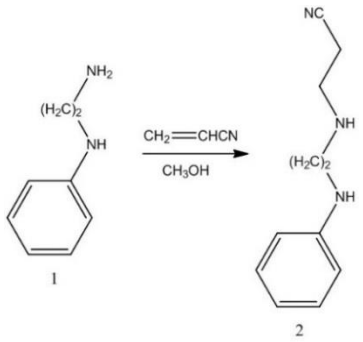


Figure 19. *Synthesis of the diamine scaffold.*

Synthesis of β-alanine spacers 4 and 6 (isocyanate aminoacids to be inserted in oligourea building block) started from two commercially available β-alanine residues, bearing orthogonal protecting groups on carboxyl functional group (H-β-Ala-OMe · HCl, H- β-Ala-OtBu · HCl). The following reactions reported in Figure 20 occur in DCM; triphosgene represents a valid alternative to highly toxic phosgene. The reactions were verified through ^{1}H NMR and the crude products were employed without further purification.[173] In particular, isocyanate 4 readily undergoes the following synthetic step, since it is very unstable, even when stored under nitrogen atmophere.

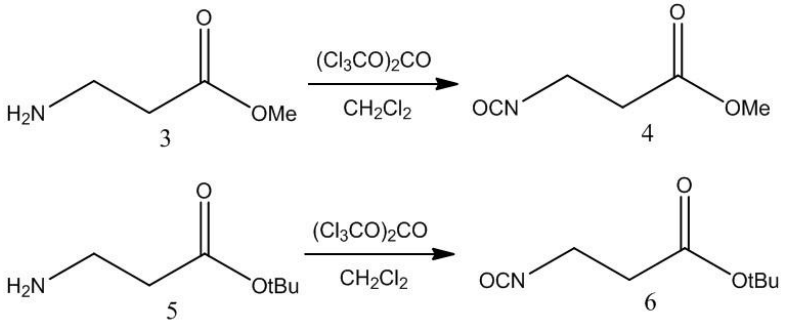


Figure 20. *Synthesis of isocyanate spacers*

The following steps should conduct to building block synthesis through sequential reactions of compound 2 with the spacers 4 and 6. Coherently with the data reported in the literature, [174] the 1:1 reaction of compound 2 with isocyanate 4 occurs in a regioselective fashion, since the aromatic amine group is remarkably less reactive than the other amine group, rather attached to an alkyl chain (Figure 21).

The desired regioisomer was then separated from the by-product 8 through flash chromatography on silica gel employing an EtOAc:Exane gradient.

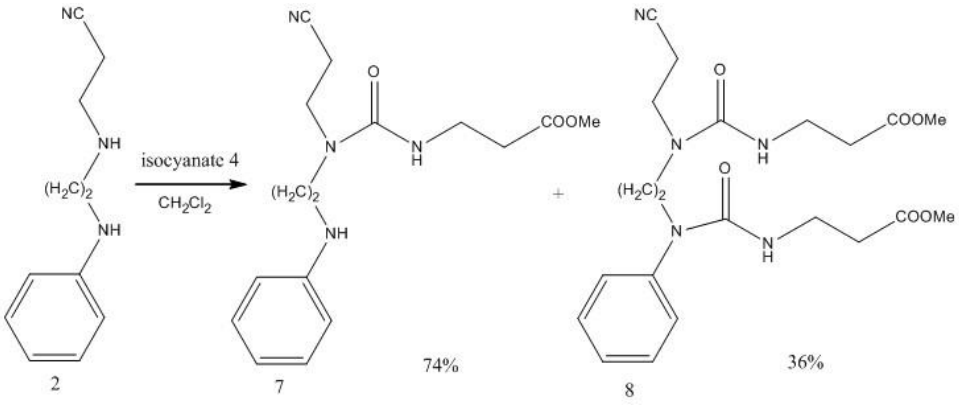


Figure 21. *Regioselective synthesis of urea 7. Compound 8 is obtained as a by-product.*

We attempted several times to perform the subsequent reaction of urea 7 with isocyanate 6 (Figure 22), as shown in Table 4, but none of the conditions tried so far allowed us to obtain an acceptable yield for compound 9.

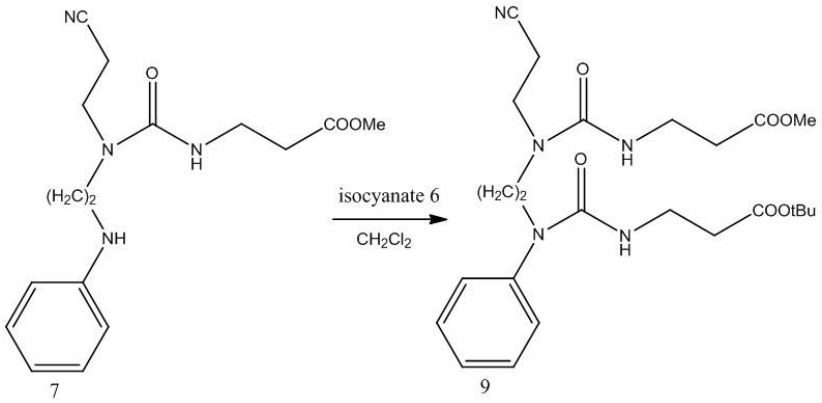


Figure 22. *Building block assembly.*

n	*T*	*t*	*solvent*	*7/6 ratio*	*yield*
1	r.t.	18 h	CH_3CN	1:1	\
2	r.t.	18 h	DCM	1:1.13	3%
3	r.t.	6 d	DCM	1:1	\
4	84°C	16 h	1,2 DCE	1:1.23	\

Table 4. *Different conditions tried for reaction of urea 7 with isocyanate 6.*

Next synthetic steps to achieve TASP 5 should be:

- *Tert*-butyl removal through diluted TFA treatment;
- Peptide block II anchorage to partially deprotected building block, when the former is still bound to the resin, while the latter is in solution (Figure 23);

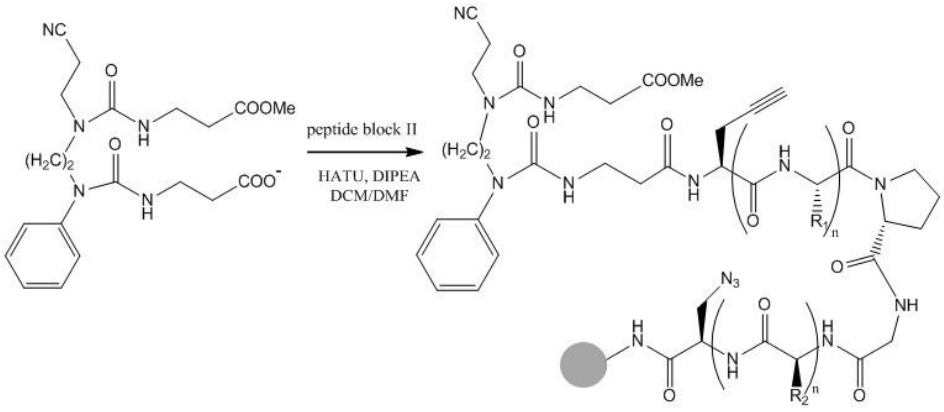


Figure 23. *Peptide block II anchorage.*

- Methyl ester protective group may be activated for subsequent coupling with peptide block I by resin treatment with DMF diluted hydrazine in presence of an excess of allyl alchol, in order to prevent possible reduction of allyl and azide groups upon hydrazine treatment; the resulting hydrazide should be later treated with *tert*-butyl nitrite in acidic conditions, finally obtaining the active acyl azide functional group (Figure 24);[175,176]

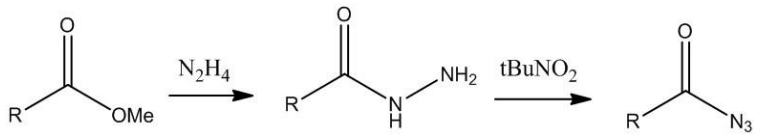


Figure 24. *Methyl ester → Acyl azide conversion*

- The next step should be the condensation via azide of the second spacer on the scaffold linked to the resin, with the block I in solution. For this purpose the block I (synthesized on resin) should be deprotected at the N-terminal amino group from the Fmoc group and then deprotected and cleaved from the resin with the lysine side chain still protected as Dde derivative. The Dde protective group could be removed after coupling by further treatment with diluted hydrazine;
- Cu(I) catalyzed cyclization through triazole bridge establishment should follow (Figure 25);[169]

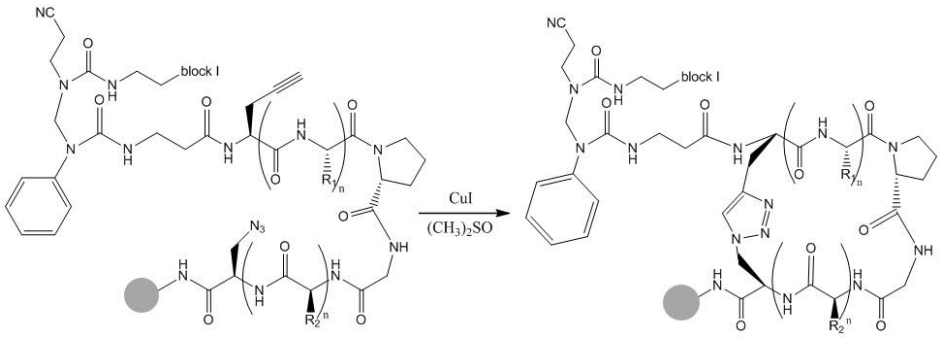


Figure 25. *Cu(I) catalysis for click reaction.*

- Final cleavage from resin and RP-HPLC purification.

3.4.7 *Future perspectives*

The difficulties found during the synthesis of TASP 5, in particular for the obtainment of compound 9, could be faced activating the aniline nitrogen as N-silylate derivative, adding isocyanate 6 to the reaction mixture, without purifying the reaction intermediate. N-silylation could be achieved employing neat TMSCN or alternatively N,O-Bis(trimethylsilyl)acetamide.

Once completed TASP 5 synthesis, several biological assays could be performed as already planned, aiming either to test TASP 5 binding to human collagen type I or the in vitro impact on collagen turnover.

Binding to collagen I may be evaluated by means Solid Phase ELISA (SP-ELISA), with the TASP construct coated on medium polarity polymeric plates and biotin-labeled collagen type I subsequently added to plate.

Surface Plasmon resonance (SPR) assays may be employed in parallel or assumed as a valid alternative to SP-ELISAs. In both assays, recombinant human SERPINH1 could serve as a positive control.

Regarding the evaluation of TASP 5 ability to affect collagen production both in normal and in pathologic conditions, an in vitro technique based on the establishment of Normal Human Dermal Fibroblasts (NHDFs) cultures has been recently set up. The aim was evaluating the activity of small linear peptides, selected and synthesized as putative collagen type I stimulating molecules (PART B, paragraphs 4.4.2, 4.4.3, 4.4.4).

In the context of in vitro assays on fibroblasts, internalization in liposomes could be useful for bigger constructs such as TASPs, in order to favor their entrance in treated fibroblasts.

4. PART B. Linear peptides from serpin A1 C-terminus: synthesis and biological screening

4.1 Role of serpin A1 in matrix remodeling patterns

Aiming to identify short chain peptides able to influence collagen production, we focused on serine protease inhibitor A1 (α1-antitrypsin, A1AT or serpin A1), a 52 kDa single chain glycoprotein, archetype of the Serpins superfamily (paragraph 1.3.3), deeply involved in tissue remodeling pathways.

The principal target of serpin A1 is neutrophil elastase, a protease released by phagocytic cells in the context of tissue damage to degrade ECM components, thus allowing infiltration of leukocyte and pro-inflammatory cytokines through the site. Serpin A1, along with other proteases inhibitors involved in the "anti-protease shield",[65] plays a major role in wound homeostasis, hampering the inflammation spread in adjacent undamaged areas.

Evidences suggest that serpin A1 might be a potential wound healing agent: it is absent or not functional in chronic wounds and in a case-report of complete A1AT deficiency severe skin abnormalities were reported.[177,178] Besides, serpin A1 implication in tissue repair processes is supported by a case study in which patients with steroid-resistant atopic dermatitis displayed a remarkable reduction of lesions if treated with topical serpin A1.[179]

Dabbagh et al., during their investigation in the balance of antiproteases and proteases at sites of tissue repair, reported an association of A1AT with fibroblasts proliferation; they also observed an enhancement of procollagen production in treated cells, independent from the mitogenic activity.[180]

4.2 Focus on serpin A1 C-terminal portion

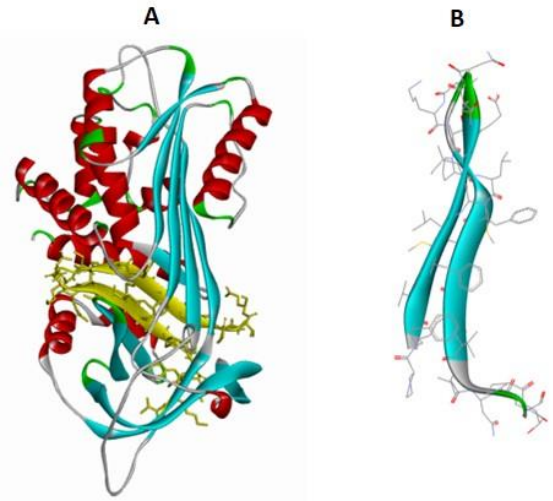


Figure 26. *(A) Serpin A1 crystal structure from PDB ID 3NE4. The C-terminal portion is showed in yellow; (B) Serpin A1 C.terminal β-hairpin domain selected from crystallized serpin A1, PDB ID 3NE4.*[181]

While inquiring about serpin A1 role in skin repair patterns, we focused on its C-terminal portion, in particular on the 36 amino acid C-terminal fragment (residues 382-418) released upon proteolytic cleavage during the formation of the complex with the substrate (Figure 26).

The A1-C36 peptide has been widely studied and found endowed with biological activities that are not dependent from the anti-protease function, located in the RCL, at a different region of the protein. A1-C36 is a specific transcriptional down-regulator of enzymes involved in bile acid synthesis in vitro and in vivo, forms amyloid fibrils present in atherosclerotic plaques leading to monocytes activation[182,183] and was attributed of chemotactic and pro-inflammatory activities.[184,185]

Congote and coworkers hypothesized that at least part of the observed biological activities resides in the C-terminal portion of A1-C36, consisting in the highly hydrophobic 26 amino acid sequence termed A1-C26,[186] containing

the β-sheet domain identified in the three dimensional serpin A1-trypsin complex.[187] In agreement with this hypothesis, they demonstrated that A1-C26 enhances collagen type I release in cultured NHDFs, as measured both in the culture media and in lysates (Figure 27).[188]

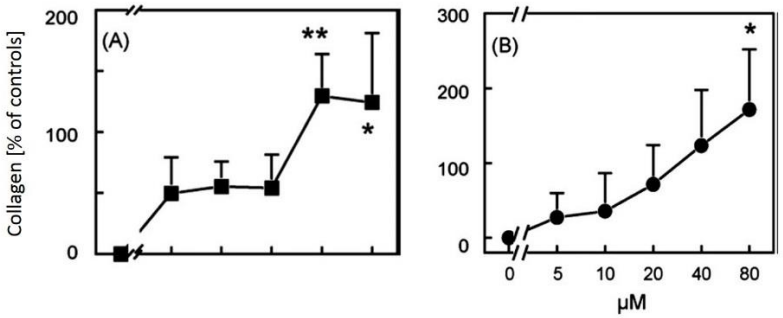


Figure 27. *Collagen I amounts measured in fibroblasts supernatants (A) and monolayers (B); results were calculated as the mean of 4-6 experiments ± S.E.M. and were expressed as percentage of stimulation over non-treated controls.*[188]

Based on these premises, we selected three overlapping peptides of A1-C26, aiming to identify shorter though still active sequences, thus providing insights on the mechanisms of action and, most importantly, trying to develop simpler molecules, suitable for biological screening and pharmaceutical applications. Therefore, the goal of this study was to identify the minimal protein fragment required for the effect on collagen production, easier to handle and to conjugate for further activity modulation.

4.3 Selection and synthesis of serpin A1 peptides

Our initial intent was selecting three 10-mer peptides of the entire highly hydrophobic sequence A1-C26, whose activity was reported in the literature and briefly described in the previous paragraph 4.2.

The overlap of each fragment to the adjacent allows preserving the integrity of functional motifs possibly located at the end of the selected sequences. In order to maintain the overlap of 2 residues, our idea was selecting two 10-mer fragments, the N-terminal and C-terminal peptides termed SA1-I and SA1-III, corresponding respectively to the protein portions serpinA1(393-402) and serpinA1(409-418), along with the 10-mer intermediate fragment SA1-II, serpinA1(401-410).

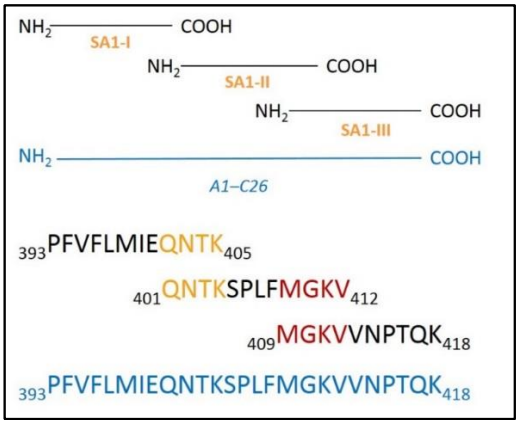


Figure 28. *Selection of overlapping fragments of A1-C26. Two 12-mer fragments and a 10-mer C-terminal peptide were selected, in order to maintain the overlap of four residues between the selected sequences.*

However, due to the high hydrophobicity of the N-terminal portion of A1-C26, two additional polar residues, namely a threonine and a lysine residue, were included at the C-terminal end of the peptide termed SA1-I, in order to prevent hydrophobic interactions, which may cause the raveling of the chain thus hiding the N-terminus and hampering the chain elongation. The intermediate fragment SA1-II was elongated as well, thus maintaining the overlap of 4 residues between the three peptides. Figure 28 reports the schematic representation of the selected serpin A1 overlapping fragments.

The synthesis of the three peptides, along with the entire sequence A1-C26, was scheduled, planning to obtain the N-termini acetylated and the C-termini ammidated for all peptides, thus avoiding the introduction of free charges not present in the native protein, which may putatively influence the biological activity profile. The C-terminal ends of SA1-III and A1-C26 were obtained in the ammidated forms as well to validate the comparison to the other N-acetylated and C-ammidated fragments in subsequent in vitro biological assays, achieving also increased resistance to protease digestion.

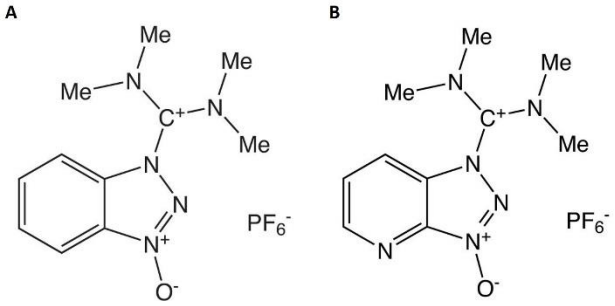


Figure 29. *Chemical structures of the main coupling reagents used for the synthesis of serpin A1 fragments: (A) HBTU; (B) HATU. HATU is mainly employed for difficult coupling reactions, since the presence of the additional aromatic nitrogen prevents epimerization due to the anchimeric assistance effect.*[189,190,191]

The synthesis of SA1-II and SA1-III, along with the entire sequence A1-C26 were performed on a manual batch synthesizer, according to standard Fmoc/*t*Bu SPPS strategy, generally employing HBTU/NMM as coupling reagents (Figure 29) and starting from a Rink-Amide resin 0,48 mmol/g, suitable to obtain the C-terminus ammidated (Figure 30, A).

Fmoc deprotection was achieved by treatment with a mild basic solution of pyperidine 20% in DMF.

α-N-terminal acetylation was carried out as reported in the Experimental Part B, paragraph 7.1.2.

PART B

Problematic coupling reactions, such as at the highly hydrophobic N-terminal portion of the entire sequence A1-C26 were carried on employing the more efficient coupling reagent HATU, and slightly changing the composition of the coupling solvents. In particular, the non-polar character of the mixture was increased, aiming to unravel the amino acid chain by favoring interactions of hydrophobic side chains with the solvent. Difficult coupling reactions were monitored via Kaiser test,[192] to reveal unreacted free amine groups; in case of positive outcome, the treatment with the coupling mixture was repeated at the same conditions, avoiding the deprotection cycles.

The synthesis of A1-C26 N-terminal fragment, termed SA1-I, required a lower loaded resin as well as the employment of HATU/NMM for each coupling reaction, in order to prevent epimerization and minimize hydrophobic interactions between the amino acid chains. In particular a Rink-Amide resin was functionalized with the C-terminal amino acid, using DCC as a coupling reagent and obtaining a final loading of 0,14 mmol/g. The effectiveness of each coupling reaction was monitored via Kaiser test and cleavages executed on a small portion of resin.

However, the solubility assays performed prior biological screening of the fragments (paragraph 4.4.1) showed that SA1-I was poorly soluble in the culture medium at all concentrations within the range of interest.

Therefore, the peptide was synthesized again, aiming to decrease the non-polar character and allow its dissolution and dilution in the culture media, necessary for subsequent in vitro biological activity evaluation. In particular, the peptide was N-conjugated to α-(9-Fluorenylmethyloxycarbonyl)amino-ω-carboxy poly(ethylene glycol) moiety (PEG9) and a serine residue corresponding to Ser405 in serpin A1 primary structure was added at the C-terminal end. The difficulties faced during the first effort to synthesize SA1-I

led to use a TentaGel™ resin with an elevated swelling capacity and HBTU/NMM as coupling reagents (Figure 30, B).

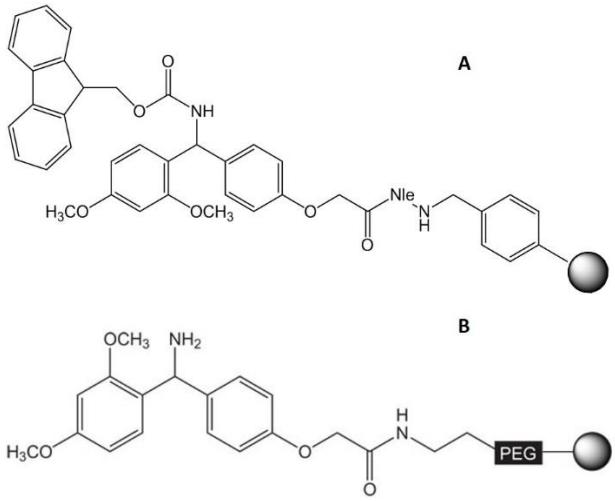


Figure 30. *Structures of the main resins that were employed for SPPS of serpin A1 fragments: (A) Rink-Amide; (B) Tentagel™ TGR R.*

After completion of each peptide sequence, cleavage from resin and contemporary deprotection of the amino acid side chains was carried out for 3 h at room temperature in a cleavage mixture consisting of Trifluoroacetic acid (TFA) supplemented with scavengers. The crude products were precipitated with diethyl ether, collected after centrifugation, dissolved in H_2O and lyophilized.

The products were purified by semi-preparative RP-HPLC to obtain a final purity ≥ 98% and characterized by RP-HPLC ESI-MS. Analytical data are reported in Table 5.

PART B

Peptide	*Protein fragment*	*Sequence*	*ESI-MS (m/z) Calc. (Found)*	*HPLC Rt[a,b]*
SA1-I	393-405	Ac-(PEG9)-PFVFLMIEQNTKS-NH_2	1738.99 (1741.85)	4.32[c]
SA1-II	401-412	Ac-QNTKSPLFMGKV-NH_2	1390.75 (1392.17)	3,32[d]
SA1-III	409-418	Ac-MGKVVNPTQK-NH_2	1142.63 (1143.74)	3.90[e]
A1-C26	393-418	Ac-PFVFLMIEQNTKSPLFMGKVVNPTQK-NH_2	1112.21 (1112.92)	4.04[f]

Table 5. *[a] Analytical HPLC on Alliance Chromatography (Waters) with a Phenomenex Kinetex C-18 column, 2.6 µm (100 × 3.0 mm), working at 0.6 mL/min. Solvent systems are A (0.1% TFA in H_2O) and B (0.1% TFA in CH_3CN). [b] Peptides purified on a Phenomenex Jupiter C-18 (250 × 4.6 mm) column using a Waters instrument (separation module 2695, detector diode array 2996) working at 4 ml/min. Solvent systems are A (0.1% TFA in H_2O) and B (0.1% TFA in CH_3CN). [c] Analytical gradient 30-80% in 5 min. [d] Analytical gradient 20-70% B in 5 min. [e] Analytical gradient 10-50% in 5 min. [f] Analytical gradient 40-90% in 5 min. Retention time (Rt) is expressed in minutes.*

4.4 Biological screening of serpin A1 peptides

4.4.1 Solubility tests on serpin A1 peptides

We performed solubility tests for serpin A1 peptides prior in vitro biological screening, since fibroblasts treatment implies the dissolution of the peptides in Dulbecco's Modified Eagle's Medium supplemented with 0.1% of fetal bovine serum (DMEM 0.1% FBS or DMEM 0.1) and subsequent sterile filtration via 0.22 µm pore size filters.

Therefore, we inquired about the solubility in culture media of serpin A1 fragments at the concentrations of interest, between 1 and 80 µM, namely the range within the control peptide A1-C26 displays activity, as reported in the literature.[188]

The peptides were dissolved/suspended in DMEM 0.1, adding Dimethyl sulfoxide (DMSO) in order to favor the dissolution at the maximum percentage still tolerated by the cells (1%). Each peptide solution at different dilution, was deposited in 6-multiwell (-mw) plates. The plates were incubated in 5% CO_2 humidified atmosphere and observed via optic microscope for 72 hours, corresponding to the entire duration of a standard treatment.

In this conditions, while testing the highest concentration of the range, i.e. 80 µM, the fragments termed SA1-I and A1-C26 resulted non-soluble in DMEM 0.1. Hence, A1-C26 was not employed as a positive control in the experiments, since it was not suitable to cover the entire concentration range of interest.

Investigating further SA1-I solubility, this peptide was found insoluble in DMEM 0.1, forming aggregates enlarging during the incubation time, even at the lowest concentration tested (1 µM), despite the N-conjugation with a PEG9 moiety. Therefore, SA1-I was not tested in vitro on NHDFs.

Besides, due to the poor solubility and the difficulties faced during the synthesis and the purification steps, we did not retain SA1-I an ideal candidate for an initial screening phase, looking for easy-to-handle molecules for future insertion in pharmaceutical formulations.

4.4.2 NHDFs establishment and treatments

Normal Human Dermal Fibroblasts (NHDFs) cultures were established from a primary culture obtained from neonatal foreskin dermis.

The cells were cultured in high-glucose (4,500 mg/l) DMEM supplemented with 10% FBS and penicillin-streptomycin at 37°C in 5% CO_2 humidified atmosphere. When at confluence, cultures were propagated by trypsinization and splitting. The experiments were conducted on young cycling cells with population doubling level (PDL) < 10 and were set up in order to measure either soluble

collagen or non-soluble collagen, referring in both cases to the archetypal collagen type I, characterized by ubiquitous occurrence.

Hence, the intention was to measure collagen in the soluble form, already secreted by fibroblasts and processed by collagenases, as well as collagen in its non-soluble form, i.e. intracellular collagen not yet released by the Golgi and collagen not suspended in cells media, rather implicated in the formation of supramolecular organized structures such as bundles and fibrils. The former was monitored in culture media via an immuno-assay, namely a sandwich ELISA set up in house; the latter was measured in lysates according to a western blotting (WB) technique.

The cells were seeded alternatively in 96-mw plates for subsequent measurement of soluble collagen type I, or in 6-mw plates for subsequent lysis and detection of non-soluble collagen type I. After trypsinization the number of cells per ml of suspension was estimated in Bürker counting chamber and 3000 cells/well were seeded in 96-mw plates, while 50000 cells/well were seeded in 6-mw plates.

After 24 hours, fibroblasts were treated with serpin A1 peptides along with L-ascorbic acid and TGF-β1 as positive controls. They are both substances that have been known for decades as collagen biosynthesis stimulators, both active at a transcriptional level.[193,194]

Collagen levels in samples from non-treated cells were referred to as controls. Due to the poor solubility exhibited by A1-C26 and SA1-I, the former was not used as a positive control, while the latter was not tested on NHDFs cultures, as more clearly explained in paragraph 4.4.1.

The peptides SA1-II and SA1-III were found readily soluble in DMEM 0.1 and were thus tested on NHDFs, treating the cells with increasing concentrations of both peptides, along with positive controls and collecting cell culture media and lysates after 72 hours incubation.

After media withdrawal, cell viability was measured in 96-mw through a ready-to-use MTS reagent to reveal possible changes due to a toxic or a mitogenic effect putatively elicited by peptides.

4.4.3 Immuno-enzymatic assays to detect collagen

A sandwich ELISA technique was set up in house, in order to monitor changes in soluble collagen type I levels in fibroblasts culture media.

The first attempt to set up an immune-assay suited for this purpose was based on the use of a couple of monoclonal antibodies recognizing two different epitopes of collagen type I and a detection method based on alkaline phosphatase (AP) converting the substrate p-Nitrophenyl Phosphate (pNPP) in the chromophore p-Nitrophenol (pNP), absorbing at λ 405 nm.

Since the sensitivity of this method was not adequate, a couple of polyclonal antibodies was preferred over monoclonal antibodies: a primary anti-collagen type I antibody for coating on medium polarity 96-mw plates and a secondary anti-collagen type I antibody biotin conjugated, with the same specificity, for subsequent detection of collagen I levels.

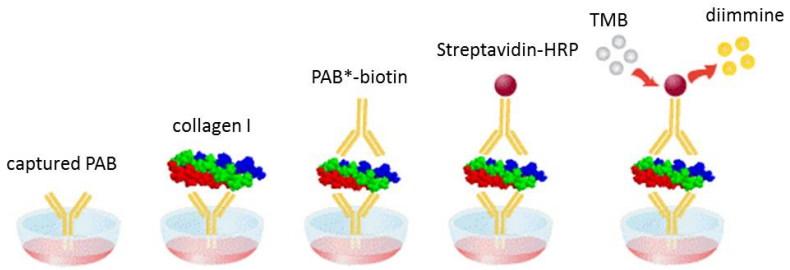


Figure 31. *Schematic representation of sandwich ELISA steps.*

The signal was further amplified, through the replacement of the AP-based detection system by horseradish peroxidase (HRP) streptavidin conjugated,

unitedly with the use 3,3',5,5'-tetramethylbenzidine (TMB) as a substrate. In fact, TMB is converted in a chromophore endowed with a much higher molar extinction coefficient if compared to pNP, and the signal is further intensified upon acidification with sulfuric acid. Figure 32 shows the calibration curves built up employing the two different sandwich ELISA methods to reveal increasing concentrations of standard human collagen type I.

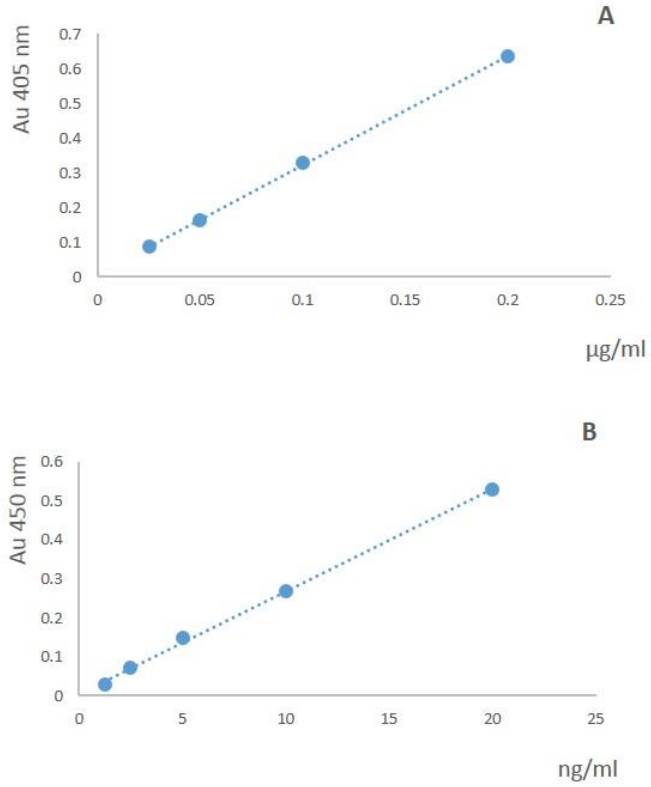


Figure 32. *Calibration curves built up with standard human collagen type I increasing concentrations, measured via sandwich ELISA technique: (A) monoclonal antibodies couple and AP-based detection method; (B) polyclonal antibody couple and HRP-based detection method. Sensitivity was amplified 10 times in case of method B.*

The immune-enzymatic method, improved in sensitivity as above described, allow detecting soluble collagen I in diluted culture media, up to 1 ng/ml, as shown in Figure 32.

The measurement of soluble collagen type I in fibroblast culture media was accomplished according to the sandwich ELISA technique previously described; along with each experiment a calibration curve was built up employing increasing concentrations of standard human collagen type I diluted in blocking buffer or in DMEM, aiming to check the linearity of the method.

4.4.4 Western blotting to measure non-soluble collagen

The measurement of collagen in fibroblast media do not furnish any information about neither intracellular nor fibrillar collagen amount, since this collagen forms are not released by the cells in the extra-cellular environment. Hence, the idea was conducting the monitor of collagen fluctuations within supernatants in parallel with the detection of non-soluble collagen within cell lysates.

Therefore, after opportune incubation with the treatments and the controls, cells in 6-mw plates were lysed by adding the Radio-ImmunoPrecipitation Assay buffer (RIPA) and collagen I was detected in recovered lysates via WB, using a radio-labeled anti-collagen type I antibody, as well as an anti-GAPDH antibody as an internal control to normalize values obtained for collagen type I bands.

4.4.5 Results

The basal amount of collagen found in the medium of untreated cells was 0.46 ± 0.04 Absorbance units (mean ± SE of 5 experiments). As reported in Figure 33 (A, B), both the positive controls TGF-β1 and ascorbic acid elicited a

significant increase in the amount of soluble collagen found in the culture media, compared to untreated cells. The two A1-C26 fragments, SA1-II and SA1-III both employed at the same concentration range, showed very different efficacies: while the effect of SA1-III was comparable to that of L-ascorbic acid, SA1-II did not significantly modify the amount of collagen in the culture medium. The corresponding percent increases compared to untreated cells were: 153% for TGF-β1, 85% for L-ascorbic acid, 13% for SA1-II and 91% for SA1-III 40 µM.

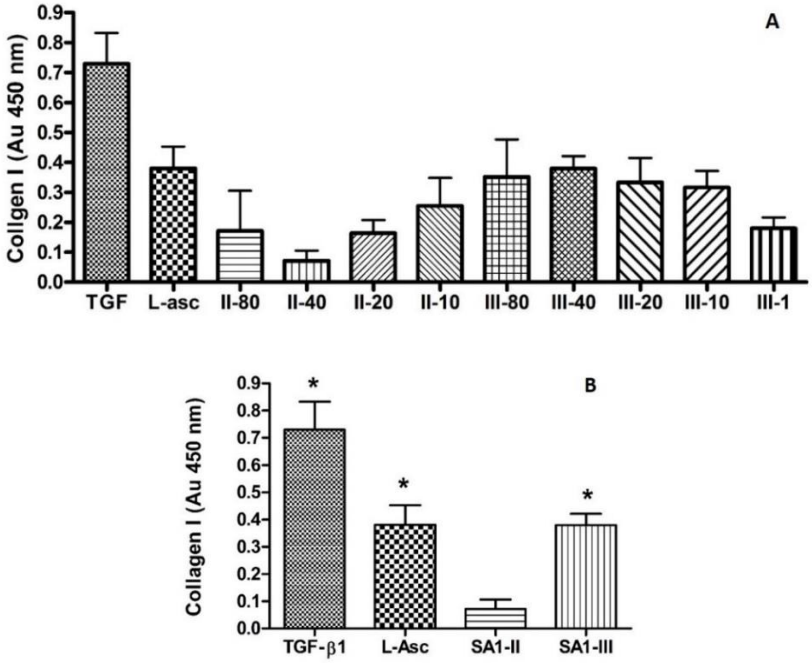


Figure 33. *Increase in collagen type I amount measured via sandwich ELISA in culture media collected from NHDFs submitted to different treatments. (A), from left to right: TGF-β1 10 ng/ml, L-ascorbic acid 10 µM, SA1-II 80, 40, 20, 10 µM, SA1-III 80, 40, 20, 10, 1 µM; (B), from left to right): TGF-β1 10 ng/ml, L-ascorbic acid 10 µM, SA1-II 40 µM, SA1-III 40 µM. Collagen I content is expressed in Absorbance units at λ 450 nm. Values are the mean ± SE of 4-5 different experiments and have been calculated subtracting to each value the basal level of collagen, measured in corresponding untreated controls. * p<0.05 statistically significant difference from untreated controls (one-way ANOVA)*

Figure 34 shows the dose-dependency of SA1-III effect, displaying a sigmoidal dose-response trend. The calculated EC_{50} resulted to be $9.75x10^{-7}$ M, corresponding to 33% increase in collagen compared to untreated cells.

Since SA1-III showed activity in terms of increasing soluble collagen I level as measured in the culture media, the biological activity of this fragment was further investigated via WB, conducted on lysates.

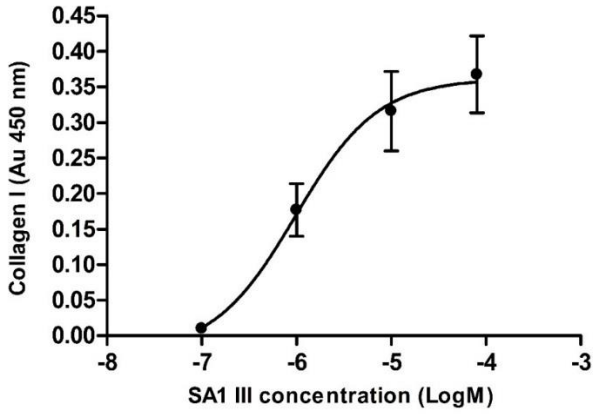


Figure 34. *(A, B); (C) Sigmoidal dose-response curve for SA1-III, serpin A1 (409-418). Collagen I content is expressed in Absorbance units at λ 450 nm. Values are the mean ± SE of 4-5 different experiments and have been calculated subtracting to each value the basal level of collagen, measured in corresponding untreated controls.*

Preliminary results were collected from 2 experiments in which non-soluble collagen I was measured via WB in samples recovered from fibroblasts treated with SA1-III 10 µM, along with positive controls L-ascorbic acid and TGF-β1. At a glance, the results even if preliminary, clearly confirm the trend exhibited in case of soluble collagen measured via ELISA in the media of SA1-III treated monolayers. In fact, collagen I levels were enhanced in treated cells if compared with untreated controls in both experiments.

PART B

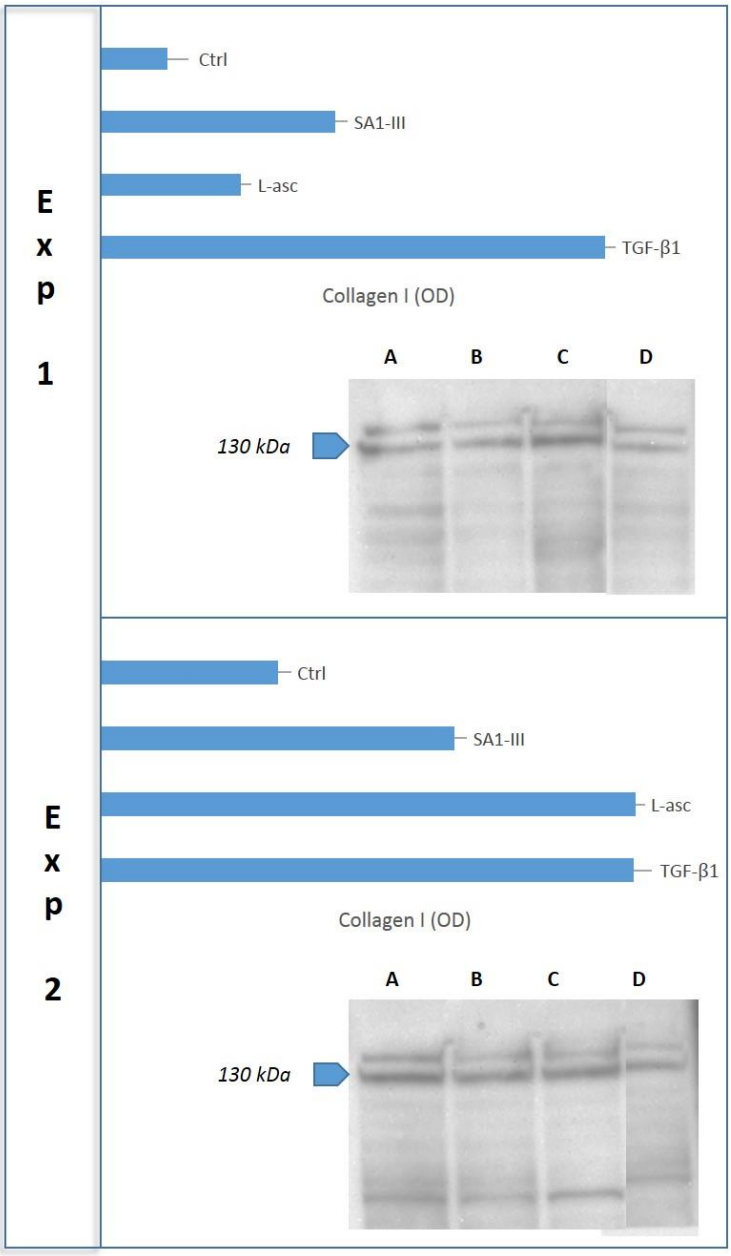


Figure 35. *Values in terms of band optical density (OD) obtained from WB performed on lysates of fibroblast monolayers: (A) TGF-β1 10 ng/ml, (B) L-ascorbic 10 µM, (C) SA1 III 10 µM, (D) untreated. GAPDH was employed as internal control to normalize values measured for collagen I. % change vs control in Exp. 1: TGF-β1 201, L-asc 203, SA1-III 100. % change vs control in Exp.2: TGF-β1 665, L-asc 111, SA1-III 255.*

4.5 Discussion and future perspectives

In the search for small molecules able to modulate collagen turnover, for either pharmaceutical or cosmeceutical purposes, peptides are an expanding category, generally exhibiting excellent safety and tolerability.[112,113] Accordingly, within the last two decades, several peptides targeting collagen production were described, mainly as active ingredients to be inserted in cosmeceutical formulas to improve skin firmness and texture.[195] Remarkable examples are the C-propeptide fragment KTTKS, whose conjugation to palmitic acid provides enhanced lipophilicity, and the tripeptide GHK, which is often employed in the form of copper complex.[196,197]

The main goal of this research was to investigate the involvement of serpin A1 C-terminal portion in collagen production stimulation, through the selection and synthesis of linear overlapping peptides of A1-C26, the A1AT hydrophobic C-terminal fragment whose in vitro impact on collagen turnover has been previously observed.[188]

As a first step, we selected and obtained through standard Fmoc/*t*Bu SPPS three subfragments of A1-C26, as listed in Table 5.

Serpin A1 fragments ability to affect collagen type I turnover was evaluated treating human NHDFs with each peptide diluted in culture media at increasing concentrations. Changes in the level of collagen type I were then monitored in collected culture media via a sandwich ELISA technique developed in house. Preliminary results regarding WB performed on lysates of monolayers treated with the active fragment SA1-III were also shown.

The peptide termed SA1-I comprises the highly hydrophobic sequence FVFLM, whose significance was extensively underlined by Joslin and colleagues. They identified high affinity receptors on human hepatoma HepG2 and human monocyte cell surface, able to bind the serpin-elastase complexes (SEC),

mediating their endocytosis and lysosomal degradation and inducing an increase in the synthesis of serpin A1. Besides, they demonstrated that the highly conserved pentapeptide FVFLM is the minimal domain sufficient for the sequence-specific recognition of the complexes by the SEC binding receptor.[198] Despite the premises regarding the hydrophobic sequence FVFLM belonging to the peptide termed SA1-I, this fragment was not tested, since it exhibited poor solubility in DMEM up to 1 µM, even if N-conjugated with a PEG9 moiety.

The fragments SA1-II and SA1-III contain the putative cleavage sites of proteolytic enzymes identified by Niemann and colleagues while studying the inhibitory activity of the C-terminal 44-residue peptide of serpin A1, termed SPAAT (short peptide from AAT).[199] In particular, SA1-III comprises a putative cleavage site of collagenase and both SA1-II and SA1-III contain cleavage sites of neutrophil elastase.

Fragment SA1-II was readily soluble but did not show any significant effect on collagen type I levels in culture medium, in a concentration range from 10 µM to 80 µM, thus acting as a negative control sequence.

Instead, fragment SA1-III, the C-terminal portion of A1-C26, elicited a significant increase in collagen type I amount as measured in media withdrawn from fibroblasts treated with different concentrations (range 0.1-100 µM), with an EC50 of 9.75x10-7 M. Dose-respose dependency for SA1-III is shown in Figure 34 (A). This effect was not due to changes in cell proliferation, as indicated by the MTS viability test. (data not shown).

Since the MTS assay allowed us to exclude a mitogenic effect from the mechanisms possibly involved in modifying collagen levels, the data reported in this work rather indicate SA1-III ability to alter the balance in collagen synthesis/degradation, inducing changes in collagen type I levels, as here observed in vitro. As putative mechanisms that may explain this effect, we

considered two main possible patterns: (1) stimulation of collagen type I biosynthesis and (2) decrease in collagen digestion upon proteases inhibition. Analyzing case (1), we did not find in the literature evidences supporting the hypothesis that serpin A1 fragments may behave as signal peptides in collagen biosynthesis pathway. Therefore, among the two hypothesis we focused in particular on mechanism (2).

The impact on collagen degradation may be attributed to a feedback-based induction of serpin A1 synthesis, which might in turn result in protease activity inhibition. Such an effect has been reported by Janchauskiene and colleagues in monocyte cultures treated with fibrillar A1-C36.[200] They observed that monocytes simultaneously treated with A1-C36 and LDL displayed a marked up-regulation of serpin A1 levels, as revealed via immunoprecipitation of serpin A1 and SDS-PAGE analysis followed by Western blot and autoradiography.[200]

Another possible explanation involves the presence within SA1-III sequence of cleavages sites specific for elastase and collagenase: peptides containing these sites may behave as competitive substrates for these enzymes, thus resulting in reduced digestion of the target proteins, such as collagen and elastin. In fact, Niemann et al. reported a putative cleavage site of collagenase, located within the SA1-III sequence (between Met-409 and Gly-410) and a cleavage site of neutrophil elastase, also comprised within the fragment SA1-III (between amino acids Val-412 and Val-413).[199] It is noteworthy that in a previous work the same authors inquired about SPAAT binding to collagen, laminin and other ECM proteins, supporting the view that the 44-residue peptide may provide in vivo protection of ECM components from inappropriate protease digestion.[201] We hypothesize that shorter fragments would act similarly in this context.

PART B

Targeting collagen turnover can be a desirable goal in all those pathologic and pathophysiologic conditions in which collagen maintains the regular triple helical structure but the amount per cell is deficient, thus affecting the integrity of tensile and support structures, such as bones and skin.

In such a scenario, we recognized in SA1-III a quite interesting molecule to proceed in further development and optimization, both in the field of rare collagen related diseases and more commonly in the expanding area of cosmeceutical products.[202]

In particular as extensively discussed in paragraph 1.4.6, both photo-induced and chronologically occurring skin-aging outline pathophysiologic conditions in which collagen amount and organization are clearly affected, yet its molecular triple-helical structure and functionality are not compromised.[103,104,105]

In the very diverse context of rare collagen-related diseases, we identified OI type I as a potential target disease; in fact, in subjects suffering from the most common form of OI, collagen type I architecture is not affected and yet biochemical tests on cultured skin OI fibroblasts show a lower-than-normal amount of type I collagen (paragraph 1.4.2).[85] A therapeutic approach based on affecting collagen turnover in OI type I patients should be more deeply evaluated, especially considering that treatments available up to now are exclusively based on surgery interventions and on the use of bisphosphonates to reduce fracture rate.[91]

In light of the results reported in this section, SA1-III can be retained a widely versatile lead compound, suitable for further studies devoted to the development of its pharmacological profile as a peptide therapeutic, involving the modulation of both its physicochemical properties to increase resistance to protease digestion, and its biological activity in view of the possible diverse applications.

Other perspectives of this work will be the definition of the implied mechanism/s of action, deepening the study of collagen degradation and working on shorter overlapping fragments selected from the active sequence SA1-III to furnish further insights on the patterns mediating the in vitro biological activity here reported.

Besides, the peptide SA1-I, which was not tested due to the poor solubility in DMEM, could be conjugated to different tags to improve the profile of this molecule, in the view of biological testing and most importantly considering a possible application.

5. MATERIALS

Peptide-synthesis-grade N,N-dimethylformamide (DMF) was from Scharlau (Barcelona, Spain). HPLC grade CH_3CN was purchased from Carlo Erba (Milano, Italy). Protected amino acids and resins were obtained from Iris Biotech AG (Marktredwitz, Germany). Coupling reagent N,N,N',N'-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate was purchased from Advanced Biotech Italia (Milano, Italy). Novasyn TGR R resin was from MerckMillipore (Darmstast, Germany).

Organic chemistry reactions were monitored by TLC, carried out on silica gel precoated plates 60 Å F_{254} from MerckMillipore (Darmstadt, Germany) and spots were revealed with Fluram®, fluorescamine, 4-phenyl-spyro[furan-2(3H),1'-isobenzofuran]-3,3'-dione), purchased by Fluka, Sigma-Aldrich Chemicals Co. (St. Louis, MO 63118 USA). Solvents were dried on molecular sieves purchased from Acros Organics, ThermoFisher scientific (Waltham, MA 02451 USA). ^{1}H and ^{13}C NMR spectra were recorded at 400 MHz on a Bruker instrument (Milano, Italia) in deuterated solutions and are reported in parts per million (ppm), with solvent resonance used as reference. Chemical shifts (δ) are reported in ppm relative to TMS and coupling constants (*J*) are reported in Hz. The multiplicity were marked as s =singlet, d = dublet, t = triplet, q = quartet, m = multiplet.

The peptides were synthesized on a manual batch synthesizer PLS 4 × 4 (Advanced ChemTech, Louisville, KY 40228 USA).

Semipreparative RP-HPLC was performed on a Phenomenex Jupiter C-18 (250 × 4.6 mm) column at 28 °C with a Waters separation module 2695 and diode-array detector 2996 (Waters, Milano, Italy), working at 4 ml/min with solvent system A (0.1% TFA in H_2O) and B (0.1% TFA in CH_3CN). Analytical HPLC system Alliance Chromatography with a Phenomenex Kinetex C-18 column 2.6 µm (100 × 3.0 mm) working at 0.6 ml/min, with the indicated linear gradients, coupled to a single quadrupole ESI-MS (Micromass ZQ) were purchased from Waters (Milano, Italy). The solvent systems used were A (0.1% TFA in H_2O) and B (0.1% TFA in CH_3CN). The peptides were lyophilized with an Edward Modulyo lyophilizer (5pascal, Milano, Italy).

Cell cultures (NHDFs) derived from the dermis of normal human neonatal foreskin (Clonetics, Lonza). DMEM was from GE Healthcare (Austria), FBS was purchased from BioWhittaker (Lonza), penicillin-streptomycin and L-ascorbic acid were obtained from Sigma-Aldrich Chemicals Co., St. Louis, MO 63118 USA. DMEM is filtered employing Nalgene Rapid flow filter (Thermo scientific) system. Cell proliferation measurement was performed with CellTiter 96® AQueous One Solution Cell Proliferation Assay kit by Promega (Madison, WI 53711 USA). Recombinant Human TGF-β1 was purchased from Peprotech (London, UK).

The BOLT® system for WB was purchased from Invitrogen, ThermoFisher scientific (Waltham, MA 02451 USA), the polyvinylidene difluoride (PVDF) membrane for protein transfer was from MerckMillipore (Darmstadt, Germany). Protein content was determined according to the Bio-Rad system (Hercules, CA 94547 USA). Antibodies for WB were purchased from Cell Signaling (Danvers, MA 01923 USA).

Chemiluminescence acquisition was achieved with the instrument ImageQuant Capture 350 from GE Healtcare (Little Chalfont, UK) and bands quantification was performed with the Quantity One software from Bio-Rad (Hercules, CA 94547 USA).

Centrifugation rates expressed in rpm are referred to a centrifuge Eppendorf model 5417R (Milano, Italy).

SP-ELISA assays were performed on 96-mw plates, Nunc Maxisorp (Sigma-Aldrich, Milano, Italy). Standard human collagen I, primary anti-collagen type I antibody and biotin conjugated anti-collagen type I antibody were purchased from MerckMillipore (Darmstadt, Germany). The antibodies used in immune-assays were polyclonal, specific for human collagen I, showing less than 10% cross reactivity with collagen types II, III, IV, V, and VI. Streptavidin-HRP conjugate was purchased from KPL (Gaithersburg, Maryland 20878 USA) and tetramethylbenzidine was employed as a ready-to-use substrate solution from eBioscience (San Diego, CA 92121 USA).

Absorbance values were measured on a Sunrise Tecan ELISA plate reader purchased by Tecan (Tecan Italia, Milano, Italy).

6. EXPERIMENTAL PART A

6.1 Molecular modeling assisted TASP design

The β-strands, along with the topological templates were designed employing the tLEAP routine in AMBER 9.[172]

Parameters regarding non-protein components, such as modified amino acids and scaffolds supporting the β-strands were inserted as elements defined by novel libraries properly created according to the xLEAP routine of AMBER and specifying the atoms implied in the connection with the peptide blocks.

6.2 RMSD values calculation

RMSD values were calculated for each candidate TASP, referring to the PDB serpin H1 model crystalized in complex with homotrimeric collagen. Key atoms in the crystalized complex and in each TASP, were manually selected either from the backbone or from amino acid side chains. Their IDs numbers were used to compare the relative position and calculate RMSD values.

6.3 Docking energy evaluation

Energy minimizations and molecular dynamic simulations (MD) were carried out by using the SANDER module of AMBER 9 with the FF99[203] force field.

Molecular dynamic simulations in implicit solvent were performed with GBSA. At this purpose, 500 steps of steepest-descent and 500 steps of conjugate-gradient minimization on the entire complex were performed with a generalized Born solvent model (igb=1), and no periodic boundary (ntb=0); the

surface area was computed and included in the solvation term, and a cutoff of 30 Å for non-bonded interactions was used.

The system was then heated from 0 to 300 K in 5 ps by holding the complex fixed with a harmonic constraint of a strength of 0.05 kcal/(mol $Å^2$).

After the minimization and heating, 1 ns dynamics simulations, with the complex constraint (strength of 0.001 kcal/(mol $Å^2$), were performed at a constant temperature of 300 K with SHAKE turned on for bonds involving hydrogens, allowing a time-step of 2 fs.

In this simulation, the binding free energies were calculated by using MM-PBSA method as implemented in the AMBER package. 200 snapshots were extracted from the MD simulation, and the binding free energies were calculated between each TASP candidate and the homotrimeric collagen model.[119] The average binding free energies were calculated every 10 snapshot getting 20 values of energy for MD trajectory.

In general, the binding free energies in condensed phase can be calculated according to the following equations.

$$\Delta G_{bind} = G_{complex} - (G_{guest} + G_{host})$$

$$G = E_{gas} + G_{solve} - TS$$

$$E_{gas} = E_{bond} + E_{angle} + E_{torsion} + E_{vdW} + E_{ele}$$

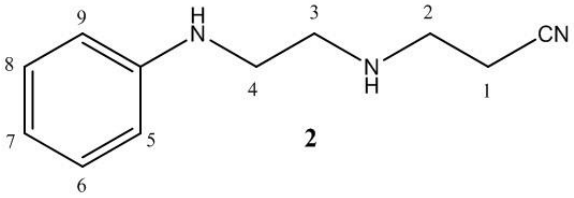
$$G_{solve} = G_{PB} + G_{SASA}$$

In the equations, $G_{complex}$, G_{guest} and G_{host} are the free energies of the complex, the molecule and the CD, respectively. E_{gas} is standard force field energy that consists of strain energies (E_{bond}, E_{angle} and $E_{torsion}$). The solvation free energy (G_{solve}) is further divided into a polar component (GPB) and a nonpolar one (GSASA). The polar component was calculated by using the PBSA program in AMBER 9.0[172], and the dielectric constant was set to 1 inside solute and 80 in solvent in this work.

The nonpolar component was determined by $\Delta G_{nonpol} = \gamma SASA + \beta$, in which SASA is the solvent-accessible surface area determined with MOLSURF[204]. In our calculations, the values for γ and β were set to 0.0072 kcal/mol Å^2 and 0 kcal/mol, respectively. The contribution of entropy (TS) to binding free energies via normal mode analyses are not evaluated as they usually have large error bars and require long simulation times.

6.4 Synthesis of the selected TASP

6.4.1 Synthesis of 3-((2-(phenylamino)ethyl)amino)propanenitrile



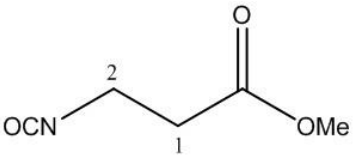
2

A 100-ml, two-necked round-bottomed flask, equipped with a magnetic stirring bar, a rubber septum and a nitrogen inlet adapter, was charged with a solution of diamine A (0.51 g, 3.75 mmol) in 50 ml of methanol, previously dried o.n. on activated molecular sieves (4 Å, 4 to 8 mesh).

Acrylonitrile (0.25 ml, 3.6 mmol) was added dropwise by syringe over one min. The resulting clear solution was stirred for 18 h at r.t. and concentrated by rotary evaporation, dissolved in 20 ml dichloromethane and concentrated again to afford 0.77 mg (yield 99%) of compound B as a yellow oil. The reaction was checked by TLC [AcOEt: n-hexane, 1:1] and spots revealed with fluorescamine in acetone and located with UV light 254 and 366 nm.

^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.18 (t, 2H, *J*=3 Hz, H-8, H-6), 6.72 (t, 2H, *J*=3.6 Hz, H-5, H-9), 6.65 (d, 1H, *J*=4.4 Hz, H-7), 3.23 (t, 2H, *J*=2.8, H-3, H-3'), 2.95 (t, 2H, *J*=3.2, H-1, H-1'), 2.92 (t, 2H, *J*=2.8 Hz, H-2, H-2'), 2.52 (t, 2H, *J*=3.3 Hz, H-3, H-3').

6.4.2 Synthesis of methyl 3-isocyanatopropanoate



4

A 25-ml, two-necked, round-bottomed flask, fitted with a rubber septum, a nitrogen inlet adapter and a magnetic stirring bar, was charged with β-alanine methyl ester hydrochloride (0.21 g, 1.5 mmol), 5 ml of CH_2Cl_2, previously distilled on $CaCl_2$, and DIEA (1.06 ml, 6.05 mmol). The resulting colorless solution was stirred at r.t. for 15 min.

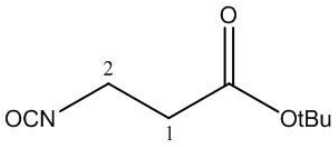
A solution of bis(trichloromethyl)carbonate (0.19 g, 0.64 mmol in 1 ml DCM) was added by syringe over 1 min and the resulting red solution was stirred 22 h at r.t.

The reaction was monitored after 1 h, 6 h and 22 h by TLC [MeOH: CH_2Cl_2, 1:1], spots were revealed with fluorescamine in acetone and located with UV light 254 and 366 nm.

The reaction mixture was extracted with 15 ml of cold 0.5 M aqueous HCl and 10 ml of crushed ice. The aqueous layer was then extracted with 5 ml of CH_2Cl_2 and the combined organic phases were extracted with a mixture of 15 ml of saturated aqueous NaCl solution and 10 ml of crushed ice, dried over $MgSO_4$, filtered and concentrated by rotary evaporation to afford 54 mg (yield 39%) of crude isocyanate as a dark red oil. During workup, the isocyanate was only exposed to water for a total time of 5-10 min.

^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.71 (s, 3H, H-Me), 3.48 (t, 2H, *J*=6 Hz, H-2, H-2'), 2.57 (t, 2H, *J*=6 Hz, H-1, H-1').

6.4.3 Synthesis of tert-butyl 3-isocyanatopropanoate



6

A 25-ml, two-necked, round-bottomed flask, fitted with a rubber septum, a nitrogen inlet adapter and a magnetic stirring bar, was charged with β-alanine methyl ester hydrochloride (0.27 g, 1.5 mmol), 5 ml of CH_2Cl_2, previously distilled on $CaCl_2$, and DIEA (1.06 ml, 6.05 mmol). The resulting colorless solution was stirred at r.t. for 15 min.

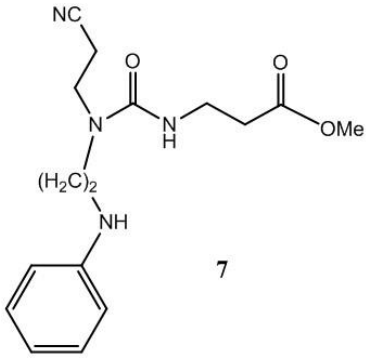
A solution of bis(trichloromethyl)carbonate (0.19 g, 0.64 mmol in 1 ml CH_2Cl_2) was added by syringe over 1 min and the resulting light yellow solution was stirred 5 h at r.t.

The reaction was monitored after 1 h and 5 h by TLC [MeOH: CH_2Cl_2, 1:2] and spots were revealed with fluorescamine in acetone and located with UV light 254 and 366 nm.

The reaction mixture was extracted with 15 ml of cold 0.5 M aqueous HCl and 10 ml of crushed ice. The aqueous layer was then extracted with 5 ml of CH_2Cl_2 and the combined organic phases were extracted with a mixture of 15 ml of saturated aqueous NaCl solution and 10 ml of crushed ice, dried over $MgSO_4$, filtered and concentrated by rotary evaporation and high vacuum to afford 183.5 mg (yield 71%) of the crude isocyanate as dark yellow crystals. During workup, the isocyanate was only exposed to water for a total time of 5-10 min.

^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.45 (t, 2H, *J*=4 Hz, H-2, H-2'), 2.48 (t, 2H, *J*=3 Hz, H-1, H-1'), 1.46 (s, 9H, H-*t*Bu).

6.4.4 Synthesis of urea 7



7

A 50-ml, two-necked, round-bottomed flask, fitted with a rubber septum, a nitrogen inlet adapter and a magnetic stirring bar, was charged with compound B (0.32 g, 1.69 mmol), 2 ml of CH_2Cl_2, previously distilled on $CaCl_2$. Isocyanate D was dissolved in 18 ml of CH_2Cl_2 under N_2 atmosphere and added dropwise under vigorous stirring. The resulting light yellow solution was stirred at r.t. 5h and monitored after 21h and 5h by TLC [AcOEt: n-hexane, 6:1].

The reaction mixture was concentrated by rotary evaporation and high vacuum. The resulting crude product was purified by flash chromatography on silica gel [AcOEt: n-hexane, 15:1] to afford 300 mg (yield 56%) of compound C. TLC [AcOEt: n-hexane, 15:1], Rf=0.45; ESI-MS: *m/z* calcd for $C_{16}H_{22}O_3N_4$ $(M + H)^+$=319.37, found 318.1; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.52 (d, 2H, *J*=3.9 Hz), 7.35 (t, 1H, *J*=4 Hz), 7.07 (t, 2H, *J*=4 Hz), 3.89 (t, 2H, *J*=3.6 Hz), 3.71-3.66 (m, 5H), 3.59 (t, 2H, *J*=3.2 Hz), 3.59 (t, 2H, *J*=3.2 Hz), 2.68 (t, 2H, *J*=3 Hz), 2.53 (t, 2H, *J*=3 Hz).

7. EXPERIMENTAL PART B

7.1 Peptide Synthesis

7.1.1 General procedure for manual SPPS

Peptides SA1-II and SA1-III were synthesized manually in reactors provided with filters, according to Fmoc/*t*Bu protection strategy, starting from a Fmoc-Rink Amide (AM)-PS resin (0.5 g, 0.48 mmol/g) and maintaining a stirring rate of 400 rpm. SA1-I was synthesized manually, starting from a Novasyn TGR R resin (0.3 g, 0.17 mmol/g,); coupling with PEG9 was performed at N-terminus previous α-N-acetylation.

Before solid phase synthesis cycles, the resins were swollen 40 minutes in DMF. The resins were treated with 1 ml of proper solution per 100 mg of resin and each amino acid cycle was characterized by the following four steps:

- Fmoc deprotection with a solution of 20% piperidine in DMF (1 × 5 min + 1 × 15 min);
- Washing cycles with DMF (3 × 2 min);
- Coupling reaction (1 × 40 min, 5 equiv amino acid and HBTU, 7 equiv DIPEA);
- Washing cycles with DMF (2 × 2 min) and DCM (3 × 1 min).

Critical coupling reactions were accomplished employing a mixture of DCM:DMF:NMP 1:1:1 as coupling solvent and their effectiveness was monitored via Kaiser Test,[192] treating a small sample of the resin with few drops of three previously prepared solutions:

Solution 1: Ninhydrin in ethanol (5g in 100 ml);

Solution 2: Phenol in ethanol (80 g in 20 ml);

Solution 3: Potassium cyanide in pyridine (2 ml of aq. Solution 0,001 M in 98 ml);

The colorimetric reaction was developed after 5 minutes incubation in a sand bath at 100°C. In case of the appearance of a blue coloration, attributed to the presence of free amine group ≥ 5%, another coupling reaction was performed after a rapid swelling in DMF.

7.1.2 α-N-Acetylation

Serpin A1 peptides were acetylated at the α-N-terminus, after removal of the Fmoc protecting group on the α-amino function of the last residue anchored to the resin. After 20 min of swelling in DCM, the resin was stirred in a solution of Ac_2O (20 equiv) and NMM (20 equiv) in DCM at r.t., and the solution was refreshed after 1 h (2 × 1 h). The resin was then washed with DCM (3 × 2 min) and dried.

7.1.3 Cleavage from resin

The peptides were cleaved from the resin, with contemporary deprotection of the amino acid side chains, by treatment with a cleavage mixture consisting of TFA/H_2O/TlS (95:2.5:2.5) or alternatively TFA/H_2O/TIS/EDT (94:2.5: 1:2.5). The resin was treated with the reaction mixture at r.t., stirring at 400 rpm. After 3 h the resin was filtered off, washed with 1 ml TFA and the solution was concentrated by flushing with N_2 until clouding. The peptides were precipitated from cold Et_2O and centrifuged. The precipitated peptides were then washed with Et_2O under stirring at 0°C (3 × 15 min), centrifuged and lyophilized.
Intermediate cleavage procedures were performed on a small sample of resin, aiming to control the synthesis progress, according to the same protocol

described for final cleavage of the complete amino acid sequence, avoiding the washing steps.

7.1.4 Purification and characterization

The peptides were purified through semipreparative RP-HPLC and subsequently characterized through RP-HPLC ESI-MS. The crude products were injected in the C18 column after dissolution in the initial phase of the chosen gradient and filtration to remove particulates. The methods employed for purification of each peptide are shown in Table 6. Analytical methods for serpin A1 peptides purification

Purification, along with analytical methods and observed and calculated masses via ESI-MS are summarized in Table 5.

Peptide	*Protein Fragment*	*Analytical Method*
SA1-I	Serpin A1 (393-405)	30-80% CH_3CN in 30 min
SA1-II	Serpin A1 (401-412)	20-70 % CH_3CN in 30 min
SA1-III	Serpin A1 (409-418)	10-50% CH_3CN in 30 min
A1-C26	Serpin A1 (393-418)	40-80% CH_3CN in 30 min

Table 6. *Analytical methods for serpin A1 peptides purification*

7.2 In vitro screening of serpin A1 peptides

7.2.1 Human fibroblast cultures maintenance and propagation

Primary Normal Human Dermal Fibroblasts (NHDFs) derived from the dermis of human neonatal foreskin were used. NHDF were cultured in DMEM, supplemented with 10% FBS, L-glutamine 1% and penicillin/streptomycin to a final concentration of 100 U/ml.

The culture media is filtered prior to use to guarantee sterility. Cells were maintained at 37°C in 5% CO_2 humidified atmosphere. At confluence, cultures were propagated by trypsinization, and the attained PDL was calculated according to the equation: PDL = 3.32 × log N/No (where N and No are the recovered and seeded cell numbers, respectively). The experiments were conducted with young cycling cells (PDL < 10), seeded into 96-mw plates and 6-mw plates. Cells seeding into mw plates from a standard 25 cm^2 culture flask was accomplished according to the following detailed procedure:

- Removal of culture medium;
- Washing cycles with PBS to remove residual DMEM (2 × 2 ml);
- Incubation at 37°C in 5% CO_2 humidified atmosphere with trypsin 0.05%-EDTA 0.02% in PBS (1 ml × 3-7 min);
- Addition of DMEM to block trypsin activity (2 ml);
- Transfer into tube and centrifugation (250 × g, 5 min);
- Supernatant removal and suspension in 1-2 ml of DMEM, favoring the mechanical disaggregation of the pellet;
- Cells count in Bürker counting chamber, performed on 20 µl of withdrawn suspension;
- Pre-conditioning of the wells with DMEM;
- Cell seeding, equally distributing the suspension into the wells (50000 cells per well in 6-mw plates, 3000 cells per well in 96-mw plates);
- Incubation until treatment at 37°C in 5% CO_2 humidified atmosphere.

7.2.2 NHDFs treatments

24 hours after seeding, DMEM was removed and monolayers were treated with each peptide diluted at the desired concentration in DMEM supplemented with a lower FBS concentration (DMEM 0.1% FBS), to avoid possible interference while subsequent collagen type I measurement. Peptide concentration range was 1-100 µM. Non-treated cells were referred to as controls and were added just of low serum DMEM, while L-ascorbic acid and TGF-β1 were employed as positive controls at a concentration of 10 µM and 10 ng/ml (≈ 0.4 nM) respectively.

The fragment SA1-I was found insoluble in DMEM, at the concentration employed for cells treatments, even in case of conjugation with a PEG9 moiety. Hence, a solubility test at 1-80 µM in DMEM was performed prior filtration, using 1% DMSO to facilitate dissolution. Aggregates were clearly visible via optic microscope since deposition in the wells, and became visibly larger after 72 h incubation.

Peptides SA1-II, SA1-III and L-ascorbic acid solutions were filtered prior cells treatment with 0.22 µm filters, in order to guarantee sterility of the solutions employed; TGF-β1 solution was rather prepared in sterile conditions, therefore the filtration step was skipped.

After 72 hours of incubation with the peptides in study and the controls, culture media were recovered and centrifuged at 250 × g for 5 min. The supernatants were then transferred into new microtubes and stored at -20°C until use.

7.2.3 MTS viability test

MTS assay is performed in order to detect changes in cell viability, possibly due to a cytotoxic or mitogen effect. The method is based on the reduction of MTS

tetrazolium compound by viable cells to generate a colored formazan product that is soluble in cell culture media. This conversion is carried out by NADPH-dependent dehydrogenase enzymes in metabolically active cells. The formazan dye produced by viable cells can be quantified by measuring the absorbance at 490-500 nm.

After media withdrawal, the cell proliferation assay was performed on monolayers into 96-mw plates, according to the following steps:

- Washing cycles with PBS (2 × 200 µl per well);
- Addition of phenol-red-free culture medium with 5% FBS (100 µl per well);
- Addition of a ready to use MTS solution (20 µl per well)
- 90 min incubation at 37°C in 5% CO_2 humidified atmosphere;
- Measurement of absorbance at 490 nm.

Two empty additional wells underwent the same procedure and were referred to as controls.

7.2.4 Cell lysates preparation

Cell lysis and recovery of nuclear, membrane and cytoplasmic proteins are accomplished treating the monolayers with Radio-ImmunoPrecipitation Assay (RIPA) buffer, composed as follows:

- Tris-HCl (pH 7.6) 25mM;
- NaCl 150 mM;
- NP-40 1% (surfactant);
- Sodium deoxycholate 1% (surfactant);
- SDS (sodium dodecyl sulfate) 0.1% (surfactant).

Phosphatase inhibitor and protease inhibitor were added both at 1% concentration to RIPA buffer, immediately prior to use.

Fibroblast monolayers in 6-mw were lysed after 72 h of incubation with treatments and controls, according to the following procedure:

- Culture media removal from wells;
- Washing cycles with PBS (3 × 2 ml);
- Addition of RIPA buffer to plates;
- The buffer was rinsed around all the well surface and transferred into micro-tubes;
- Sonication of the lysates (1 min);
- Centrifugation at 4°C (14000 rpm on a Eppendorf 5417R centrifuge, 1 min);
- Transfer of the supernatants into new micro-tubes;
- Storage at -20°C until use.

7.2.5 Total protein content determination in cell lysates

The total protein content in fibroblast lysates was evaluated employing the Bio-rad kit (DC Protein Assay), according to the Lowry method, in which proteins in the sample (mainly tryptophan and tyrosine, but also cystine, cysteine and histidine) react with an alkaline solution of copper tartrate and Folin reagent.
The assay was performed in 96-mw, adding in sequence:

- 5 µl of each sample;
- 25 µl reagent A (copper tartrate);
- 200 µl reagent B (Folin reagent).

Other five wells were prepared with increasing concentrations of an albumin solution in MQ water (0.125, 0.25, 0.5, 1 e infine 2 mg/ml), while an additional well was prepared with MQ water and considered as a reference control. After incubation of the plate (37°C, 30 min), absorbance values were read at 590 nm and the protein content of each sample was determined basing on the

calibration curve built up with values obtained from wells treated with albumin solutions at known concentrations, subtracting the control.

7.2.6 Protein concentration procedure

Protein concentration was performed in case of too diluted samples (protein concentration < 1 mg/ml), using 3K Amicon filters, according to the following procedure:

- Saturation of the filters with 1% nonfat milk powder in MQ water (overnight, 500 µl);
- Removal of the milk solution, avoiding direct contact with filters;
- Addition of 500 µl of MQ water and centrifugation (14000 rpm, 5 min), twice;
- Filter inversion and centrifugation (1000 rpm, 2 min);
- Filters transfer on new tubes and sample loading (≈ 100 µl);
- Centrifugation (14000 rpm, 15 min);
- Second centrifugation cycle may vary, depending on the final volume requested; the resulting volume of the sample was periodically monitored during centrifugation;
- Filters inversion on new tubes and centrifugation (1000 rpm, 2 min).
- Storage of the concentrated samples at -20°C until novel protein content measurement and subsequent employment in WB, as described in section 7.2.7.

7.2.7 General procedure for western blotting

Sample preparation. After determination of the protein content (≈ 15-30 µg) samples were prepared adding to lysates sample buffer 4X, containing SDS as denaturing agent, β-mercaptoethanol 10X (reducing agent) and MQ water to a

final volume of 25-30 µl. Once prepared as described, the samples underwent denaturation procedure at 70°C for 10 minutes; they were brought to r.t. and centrifuged at 14000 rpm for 1 minute.

Gel electrophoresis. Proteins in the samples were separated with Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis (SDS-PAGE) on acrylamide gel, in running buffer MOPS, used for proteins with weights from 200 Da to 14 kDa and supplemented with antioxidant agent (500 µl were added for 200 ml of running buffer MOPS).

Pre-cast gels were briefly washed with deionized water and, after removal of the comb, they were rinsed with running buffer filling the wells to remove residual acrylamide; running buffer was also employed to fill the internal and external chamber of the electrophoresis apparatus.

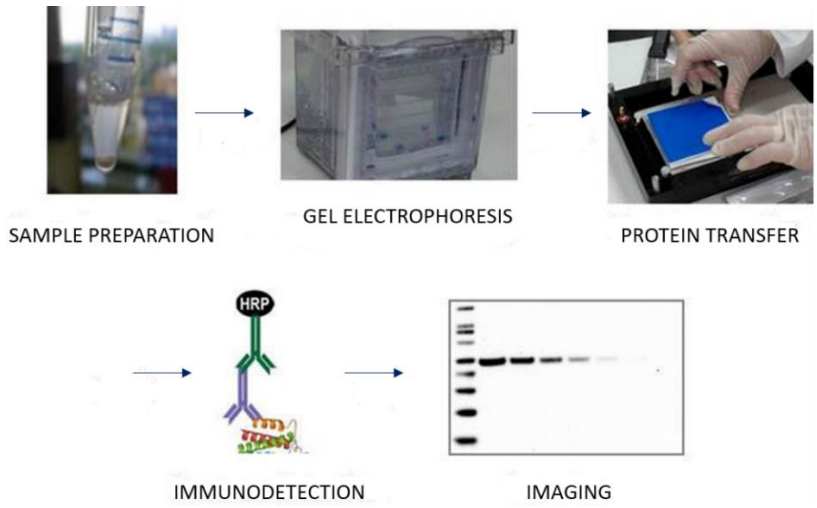


Figure 36. *Schematic representation of WB procedure.*

The samples were loaded, along with 2.5 µl of the two markers *See Blue standard* and *Magic Mark,* both useful to track the molecular weights corresponding to the visualized bands in the visible-range and via

chemiluminescence respectively. Protein separation was performed at 165 V for 45 minutes, with 40-125 mA current.

Protein transfer. Protein transfer on PVDF membrane was achieved in transfer buffer, obtained mixing prior to use (1) 25 ml of a ready-to-use solution (BOLT 20X) with (2) 50 ml methanol, (3) 500 µl antioxidant and (4) deionized water to reach a final volume of 500 ml.

The sandwich was assembled according to the order shown in Figure 37. Before the assembly, all components were rinsed in transfer buffer and the membrane was submerged in methanol for 10 minutes and washed 15 seconds in transfer buffer prior to use. Protein transfer was achieved in the same chamber employed for electrophoresis, filled with transfer buffer, setting up a voltage of 20 V for 1 hour, with a starting current intensity of 380 mA, progressively decreasing.

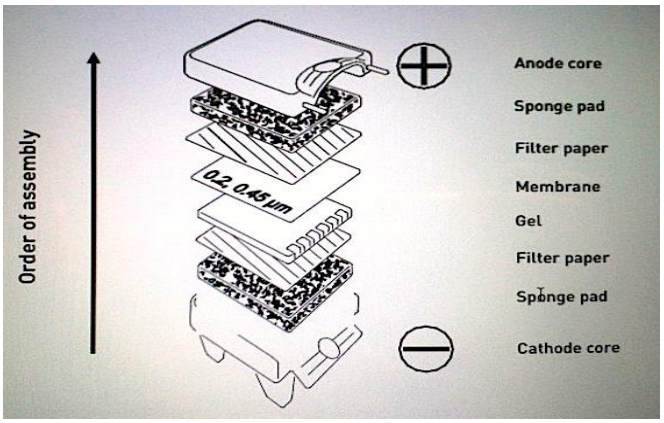


Figure 37. *Assembly of the sandwich for protein transfer on the PVDF membrane.*

The protein transfer on the membrane was observed by treating the membrane few seconds with Ponceau S stain (0.1% solution in 1% acetic acid), and then rinsing with deionized water. The occurrence of the blot was also

verified by the appearance of the *See Blue standard* bands profile on the membrane, which subsequently assisted the cut of the membrane basing on the molecular weight of the proteins of interest.

Immunodetection. Prior incubation with primary antibody, non-specific binding sites on the membrane were blocked by 1 hour treatment at r.t. with powder milk 6% in PBS-Tween. After the removal of the blocking solution, the membrane was incubated o.n. at 4°C under mild stirring, with primary anti-collagen type I antibody, 1:1000 and anti-GAPDH antibody 1:3000 in PBS-Tween 0.05%.

After incubation, the primary antibodies were removed and the membrane was washed 3 times with PBS-Tween 0.1%. Once executed the three washing cycles (10 ml × 5 min), the membrane was transferred into the blot holder of the *Snap i.d.® 2.0 Protein Detection System* apparatus (Figure 38) to perform the last washing (10 ml × 1 min). The washing buffer was filtered off and the secondary antibody was added (10 ml, 1:4000 in PBS-Tween 0.05%).

Figure 38. *Snap i.d.® 2.0 Protein Detection System apparatus; this system is directly connected with a vacuum filtration pump, hence optimizing the antibody-antigen interaction and reducing the washing time.*

After 20 minutes incubation at r.t., the antibody solution was filtered off and 4 washing cycles were executed (10 ml × 1 min). The membrane was then removed from the SNAP system and exposed for 2 minutes to luminol and H_2O_2

(500 + 500 µl respectively); the residual solution was then removed and chemiluminescence was read with *ImageQuant Capture 350*, subsequently employing the *Quantity One®* software for band quantification. Bands corresponding to GAPDH (MW 37 kDa) were considered as internal control reference for the determination of collagen type I levels.

Stripping and storage of the membranes. After acquisition, the membrane was washed once with PBS-Tween 1% to remove residual luminol and H_2O_2. The primary antibody was rather removed by treatment with stripping buffer under vigorous stirring at r.t. for 15 minutes. The membrane was washed with PBS-Tween 1% (4 × 5 min) and non-specific binding sites were blocked with 6 % milk powder solution in MQ water. The membrane were stored 3-4 days in PBS-Tween 1 % at 4°C; alternatively, they were dried, disposed on filter paper and stored at -20°C.

7.2.8 Immunoassays

Soluble collagen in culture media was measured through SP-ELISA. All buffers were brought to r.t. prior to use and filtered through 0.22 µm filters to remove suspended particles.

Activated polystyrene 96-well ELISA plates were coated with the primary anti-collagen type I antibody 100 µl/well 1:100 in pure carbonate buffer 0.05 M (pH 9.6). After overnight incubation at 4°C, the antibody solution was removed and plates were washed 3 times with 300 µl/well of washing buffer, consisting of PBS supplemented with 0.05% of Tween 20. Non-specific binding sites were blocked 1 h 40 minutes at r.t. with blocking buffer, consisting of 10% FBS in PBS (200 µl/well).

After 4 washes with PBS-Tween 0.05%, plates were incubated with 100 µl/well of cell-conditioned culture media diluted 1:20 in DMEM 0.1% FBS. Wells incubated with unconditioned DMEM 0.1 % FBS were referred to as controls.
In order to check the linearity oh the method, a calibration curve was obtained in parallel within each experiment, employing increasing concentrations of standard human collagen type I diluted in blocking buffer.
After 2 h 40 minutes incubation at r.t., plates were washed 5 times with washing buffer and treated with 100 µl/well of biotin-conjugated anti-collagen type I antibody 1:4500 in blocking buffer. After 1 h incubation at r.t., the antibody solution was removed, plates were washed 5 times and the streptavidin-HRP conjugate was added (100 µl/well of a 1:2000 solution in blocking buffer). After 30 minutes incubation at r.t., the conjugate solution was removed and plates were washed before adding 100 µl/well of a TMB ready-to-use solution. Absorbance values were read 8 minutes after the addition of the substrate solution, upon acidification with 2 N H_2SO_4 in water 50 µl/well. Collagen I content was expressed in Absorbance units at λ 450 nm.
The effect of the different treatments was shown as net increase over the basal level, calculated subtracting to each value the corresponding untreated control value.

7.2.9 Statistics

Values were expressed as mean ± SE of 4-5 different experiments. The statistical significance of the differences was analyzed by one-way ANOVA followed by Dunnett multiple comparison test.

8. GLOSSARY

A1AT: α-1-antitrypsin;

AP: Alkaline Phosphatase;

DCC: N,N'-Dicyclohexylcarbodiimide;

DCM: Dichloromethane;

Dde: N-(1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl);

DI: Dentinogenesis Imperfecta;

DIEA: N,N-Diisopropylethylamine;

DMEM: Dulbecco's Modified Eagle's Medium;

DMF: N,N-dimethylformamide

DMSO: Dimethyl sulfoxide;

ECM: Extra-Cellular Matrix;

EDS: Ehlers-Danlos Syndrome;

EDT: Ethanedithiol;

EGFR: Epidermal Growth Factor Receptor;

ELISA: Enzyme-Linked Immunosorbent Assay;

ER: Endoplasmic Reticulum;

ESI-MS: Electrospray Ionization-Mass Spectrometry;

FACIT: Fibrils-Associated Collagen with Interrupted Triple helix;

FBS: Fetal Bovine Serum;

Fmoc: Fluorenylmethyloxycarbonyl;

GAPDH: Glyceraldehyde-3-Phosphate Dehydrogenase;

HATU: 2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate;

HBTU: N,N,N',N'-tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate;

HRP: Horseradish Peroxidase;

Hyl: Hydroxylysine;

Hyp: Hydroxyproline;

HSP: Heat Shock Protein;

IL: Interleukin;

MACIT: Membrane-Associated Collagens with Interrupted Triple helix;

MCTD: Mixed Connective Tissue Disease;

MD: Molecular Dynamics;

MMPs: Matrix Metalloproteinases;

MTS: (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium);

NC: Non-Collagenous;

NHDFs: Normal Human Dermal Fibrosbasts;

NMM: N-methylmorpholine;

NMP: N-methylpyrrolidone;

n-MW: n-Multiwell;

NO: Nitrogen monoxide;

OA: Osteoarthritis;

OD: Optical Density;

OI: Osteogenesis Imperfecta;

PAI-1: Plasminogen Activator Inhibitor 1;

PBS: Phosphate Buffered Saline;

PDB: Protein Data Bank;

PDI: Protein Disulphide Isomerase;

PDL: Population Doubling Level;

PEG9: α-(9-Fluorenylmethyloxycarbonyl)amino-ω-carboxy poly(ethylene glycol);

PH-4: Prolyl 4-Hydroxylase;

pNP: p-Nitrophenol;

pNPP: p-Nitrophenyl Phosphate;

PPIases; Peptydil Prolyl *cis-trans* Isomerases;

PVDF: Polyvinylidene Difluoride;

RIPA: Radio-ImmunoPrecipitation Assay;

RCL: Reactive Centre Loop;

ABBREVIATIONS

RMSD: Root Mean Square Deviation

ROS: Reactive Oxygen Species;

SDS-PAGE: Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis

SEM: Standard error of measurement;

SERPINs: Serine Protease Inhibitors;

SLRPs: Small Leucine-rich Repeat Proteoglycans;

SPAAT: Short Peptide from α1-antitrypsin;

SP-ELISA: Solid Phase Enzyme-Linked Immunosorbent Assay;

SPPS: Solid Phase Peptide Synthesis;

SPR: Surface Plasmon Resonance;

TASP: Template Assembled Synthetic Protein;

t-Bu: *tert*-Butyl;

TFA: Trifluoroacetic acid;

TGF-β1: Transforming Growth Factor β1;

TIMPs: Tissue Inhibitors of Metalloproteinases;

TIS: Triisopropylsilane;

TLC: Thin Layer Chromatography;

TMS: Tetramethylsilane;

TNF-α: Tumor Necrosis Factor-α;

TMB: 3,3',5,5'-tetramethylbenzidine;

u-PA: urokinase-type Plasminogen Activator;

uPARAP: urokinase-type Plasminogen Activator Rreceptor-Associated Protein;

WB: Western Blotting.

9. REFERENCES

[1] R. F. Diegelmann, *Wounds*, **2001**, 5: 13;

[2] J. Brinckmann, *Top Curr Chem*, **2005**, 247: 1-6;

[3] J. Engel, H. P. Bächinger, *Top Curr Chem*, **2005**, 247: 7-33;

[4] S. Ricard-Blum, F. Ruggiero, M. van der Rest, *Top Curr Chem*, **2005**, 247: 35-84;

[5] J. Bella, M. Eaton, B. Brodsky, H. M. Berman, *Science,* **1994***,* 266: 75–81;

[6] B. Brodsky, A. V. Persikov, *Adv. Protein Chem.,* **2005***,* 70: 301–339;

[7] A. V. Persikov, J. A. M. Ramshaw, A. Kirkpatrick, B. Brodsky, *Biochemistry,* **2005***,* 44: 1414–1422;

[8] J. Engel, D. J. Prockop, *Ann. Rev. Biophys Biophys Chem*, **1991**, 20: 137-152;

[9] D. J. S. Hulmes, *J Struct Biol*, **2002**, 137: 2-10;

[10] K. E. Kadler, D. F. Holmes, J. A. Trotter, J. A. Chapman, *Biochem J*, **1996**, 316: 1-11;

[11] J. F. Lees, M. Tasab, N. J. Bulleid,, *EMBO J.* **1997**, 16: 908–916;

[12] T.F Linsenmayer, E. Gibney, F. Igoe F, M. K. Gordon, J. M. Fitch, L. I. Fessler, D. E. Birk, *J Cell Biol*, **1993**, 121: 1181–1189;

[13] J. Exposito , U. Valcourt, C. Cluzel, C. Lethias, *Int J Mol Sci,* **2010**, 11: 407-426;

[14] S. Ricard-Blum, *Cold Spring Harb Perspect Biol,* **2011***,* 3: a004978;

[15] S. Kalamajski, A. Oldber, *Matrix Biol*, **2010**, 29: 248–253;

[16] K. E. Kadler, A. Hill, E. G. Canty-Laird, *Curr Opin Cell Biol*, **2008**, 20: 495–501;

[17] J. J. Wu, M. A. Weis, L. S. Kim, D. R. Eyre, *J Biol Chem*, **2010**, 285: 18537–18544;

[18] P. Bruckner, *Cell Tissue Res*, **2010**, 339: 7–18;

[19] R. J. Wenstrup, J. B. Florer, E. W. Brunskill, S. M. Bell, I. Chervoneva, D. E. Birk, *J Biol Chem,* **2004***,* 279: 53331–53337;

[20] B. Font, E. Aubert-Foucher, D. Goldschmidt, D. Eichenberger, M. van der Rest, *J Biol Chem*, **1993**, 268: 25015-205018;

[21] T. Ehnis, W. Dieterich, M. Bauer, H. Kresse, D. Schuppan, *J Biol Chem*, **1997**, 272: 20414-20419;

[22] M. Fluck, M. N. Giraud, V. Tunc, M. Chiquet, *Biochim Biophys Acta*, **2003**, 1593: 239-248;

[23] J. K. Mouw, G. Ou, V. M. Weaver, *Nat Rev Mol Cell Biol,* **2014**, 15: 771-785;

[24] T. Koide, K. Nagata, *Top. Curr. Chem.*, **2005**, 247: 85-114

[25] K. Zararullah, E. M. Brown, H. Kuivaniemi, G. Tromp, A. L. Sieron, A. Fertala, D. J. Prockop, *Matrix Biol*, **1997**, 16: 201-209;

[26] A. V. Buevich, Q. H. Dai, X. Liu, B. Brodsky, J. Baum, *Biochemistry*, **2000**, 32: 4299-4308;

[27] E. Leikina, M. V. Mertts, N. Kuznetsova, S. Leikin, *Proc Natl Acad Sci USA*, **2002**, 99: 1314-1318;

[28] R. A. Wagenaar-Miller, L. H. Engelholm, J. Gavard, S. S. Yamada, J. S. Gutkind, N. Behrendt, T. H. Bugge, K. Holmbeck, *Mol Cell Biol*, **2007**, 27: 6309-6322;

[29] K. B. Hotary, E. D. Allen, P. C. Brooks, N. S. Datta, M. W. Long, S. J. Weiss, *Cell*, **2003**, 114: 33-45;

[30] J. D**.** Mott, Z. Werb, *Curr Opin Cell Biol*, **2004**, 16: 558–564;

[31] V. Ottani, D. Martini, M. Franchi, A. Ruggeri, M. Raspanti, *Micron*, **2002**, 33: 587-596;

[32] P. Saftig, E. Hunziker, O. Wehmeyer, S. Jones, A. Boyde, W. Rommerskirch, J. D. Moritz, P. Schu, P. K. von Figura, *Proc. Natl. Acad. Sci. U.S.A.,* **1998**, 95: 13453-13458;

[33] J. Shi, M. Y. Son, S. Yamada, L. Szabova, S. Kahan, K. Chrysovergis, L. Wolf, A. Surmak, A., K. Holmbeck, *Dev. Biol.,* **2008**, 313: 196-209;

[34] P.A.M. Snoek van Beurden, J.W. Von den Hoff, *BioTechniques*, **2005**, 38: 73–83;

[35] Y.Y. Li, C.F. McTiernan, A.M. Feldman, *Cardiovascular Research*, **2000**, 46: 214–224;

[36] R.T. Aimes, J.P. Quigley, *J Biol Chem*, **1995**, 270: 5872-5876;

[37] J.L. Lauer-Fields, D. Juska, G.B. Fields, *Biopolymer*, **2002**, 66: 19-32;

[38] L. Kjøller, L. H. Engelholm, M. Høyer-Hansen, K. Danø, T. H. Bugge, N. Behrendt, *Exp Cell Res,* **2004**, 293, 106-116;

[39] D. H. Madsen, L. H. Engelholm, S. Ingvarsen, T. Hillig, R. A. Wagenaar-Miller, L. Kjøller, H. Gårdsvoll, G. Høyer-Hansen, K. Holmbeck, T. H. Bugge, N. Behrendt, *J Biol Chem,* **2007**, 282: 27037-27045;

[40] D. H. Madsen, S. Ingvarsen, H. J. Jürgensen, M. C. Melander, L. Kjøller, A. Moyer, C. Honoré, C. A. Madsen, P. Garred, S. Burgdorf, T. H. Bugge, N. Behrendt, L. H. Engelholm, *J Biol Chem*, **2011**, 286: 26996-27010;

[41] T.H. Bugge, N. Behrendt, in *Biology of Extracellular Matrix 2* (Eds.: W. C. Parks, R. P. Mecham), Springer-Verlag Berlin Heidelberg, **2011,** pp. 10-20;

[42] L. J. Geoffrey, *Am J Physiol*, **1987**, 252: C1-C9;

[43] Z. S. Galis, M. Muszynski, G. Sukhova, E. Simon-Morrissey, P. Libby, N. Y. Ann, *Acad Sci*, **1995**, 748: 501-507;

[44] V. H. Rao, R. K. Singh, D. C. Delimont, R. H. Finnell, J. A. Bridge, J. R. Neff, B. P. Garvin, D. L. Pickering, W. G. Sanger, B. A. Buehler, G. B. Schaefer, *Int J Oncol*, **1999**, 14: 291-300;

[45] S. Uemura, H. Matsushita, W. Li, A. J. Glassford, T. Asagami, K. H. Lee, D. G. Harrison, P. S. Tsao, *Circ Res*, **2001**, 88: 1291-1298;

[46] D. C. Radisky, D. D. Levy, L. E. Littlepage, H. Liu, C. M. Nelson, J. E. Fata, D. Leake D, E. L. Godden, D. G. Albertson, M. A. Nieto, Z. Werb, M. J. Bissell, *Nature*, **2005**, 436: 123-127;

[47] Y. Li, D. J. Samuvel, K. P. Sundararaj, M. F. Lopes-Virella, Y. Huang, *J Cell Biochem*, **2010**, 110: 248–59;

[48] Y. Ben-Yosef, N. Lahat, S. Shapiro, H. Bitterman, A. Miller, *Circ Res*, **2002**, 90: 784-91;

[49] A. M. Romanic, J. A. Madri, *J Cell Biol*, **1994**, 125: 1165-78;

[50] Y. Hasebe, K. Egawa, M. Shibanuma, K. Nose, *Cancer Sci*, **2007**, 98: 58-67;

[51] I. Sehgal, T. C. Thompson, *Mol Biol Cell*, **1999**, 10: 407-16;

[52] L. R. Gomes, L. F. Terra, R. A. Wailemann, L. Labriola, M. C. Sogayar, *BMC Cancer*, **2012**, 12: 26;

[53] D. Pei, S. J. Weiss, *Nature*, **1995**, 375: 244-7;

[54] P. Stawowy, H. Meyborg, D. Stibenz, N. Borges Pereira, N. Stawowy, M. Roser, U. Thanabalasingam, J. P. Veinot, M. Chrétien, N. G. Seidah, E. Fleck, K. Graf, *Circulation*, **2005**, 111: 2820-2827;

[55] X. Fu, S. Y. Kassim, W. C. Parks, J. W. Heinecke, *J Biol Chem*, **2001**, 276: 41279-41287;

[56] S. J. Weiss, G. Peppin, X. Ortiz, C. Ragsdale, S. T. Test, *Science*, **1985**, 227: 747-749;

[57] M. Sariahmetoglu, B. D. Crawford, H. Leon, J. Sawicka, L. Li, B. J. Ballermann, C. Holmes, L. G. Berthiaume, A. Holt, G. Sawicki, R. Schulz, FASEB J, **2007**, 21: 2486-2495;

[58] A. L. Jacob-Ferreira, M. Y. Kondo, P. K. Baral, M. N. James, A. Holt, X. Fan X, R. Schulz, *PLoS One*, **2013**, 8: e71794;

[59] A. Page-McCaw, A. J. Ewald, Z. Werb, *Nat Rev Mol Cell Biol*, **2007**, 8: 221-233;

[60] F. J. Descamps, E. Martens, F. Ballaux, K. Geboes, G. Opdenakker, *J Pathol*, **2004**, 204: 555-561;

[61] Z. Gu, M. Kaul, B. Yan, S. J. Kridel, J. Cui, A. Strongin, J. W. Smith, R. C. Liddington, S. A. Lipton, *Science*, **2002**, 297: 1186-1190;

[62] F. X. Gomis-Rüth, K. Maskos, M. Betz, A. Bergner, R. Huber, K. Suzuki, N. Yoshida, H. Nagase, K. Brew, G. P. Bourenkov, H. Bartunik, W. Bode, *Nature*, **1997**, 389: 77-81;

[63] C. E. Campbell, A. M. Flenniken, D. Skup, B. R. Williams, *J Biol Chem*, **1991**, 266: 7199-206;

[64] P. Silacci, J. M. Dayer, A. Desgeorges, R. Peter, C. Manueddu, P. A. Guerne, *J Biol Chem*, **1998**, 273: 13625-9;

[65] D. R. Yager, B. C. Nwomeh, *Wound Repair Regen*, **1999**, 7: 433-441;

[66] K. J. Aitken, D. J. Bagli DJ, *Nat Rev Urol*, **2009**, 6: 596-611;

[67] R. W. Carrell, P. E. Stein, G. Fermi, M. R. Wardell, *Structure*, **1994**, 2: 257-270;

[68] H. K. Song, K. N. Lee, K. S. Kwon, M. H. Yu, S. W. Suh, *FEBS Lett*, **1995**, 377: 150-154;

[69] T. J. Stout, H. Graham, D. I. Buckley, D. J. Matthews, *Biochemistry*, **2000**, 39: 8460-8469;

[70] T. Nakashima, S. C. Pak, G. A. Silverman, P. M. Spring, M. J. Frederick, G. L. Clayman, *Biochim Biophys Acta*, **2000**, 1492: 441-446;

[71] J. A. Huntington, R. J. Read, R. W. Carrell, *Nature*, **2000**, *407*: 923-926;

[72] D. van Gent, P. Sharp, K. Morgan, N. Kalsheker, *Int J Biochem Cell Biol*, **2003**, 35: 1536-1547;

[73] F. Malfait, A. De Paepe, Brenner's Encyclopedia of Genetics, 2nd Edition, Volume 2, 2013, pp. 456-459;

[74] S. Symoens, D. Syx, F. Malfait, B. Callewaert, J. De Backer, O. Vanakker, P. Coucke, A. De Paepe, *Hum Mutat*, **2012** 33: 1485-1493;

[75] P. H. Byers, M. L. Murray, *Matrix Biology*, **2014**, 33: 10-15;

[76] P. H. Byers, M. Duvic, M. Atkinson, M. Robinow, L. T. Smith, S. M. Krane, M. T. Greally, M. Ludman, R. Matalon, S. Pauker, D. Quanbeck, U. Schwarze, *Am J Med Genet*, **1997**, 72: 94-105;

[77] L. Nuytinck, M. Freund, L. Lagae, G. E. Pierard, T. Hermanns-Le, A. De Paepe, *Am J Hum Genet*, **2000**, 66: 1398–1402;

[78] W. A. Cabral, E. Makareeva, A. D. Letocha, N. Scribanu, A. Fertala, A. Steplewski, D. R. Keene, A. V. Persikov, S. Leikin, J. C. Marini, *Hum Mutat*, **2007**, 28: 396-405;

[79] U. Schwarze, R. Hata, V. A. McKusick, H. Shinkai, H. E. Hoyme, R. E. Pyeritz, P. H. Byers, *Am J Hum Genet*, **2004**, 74: 917-930;

[80] F. Rauch, F. H. Glorieux, *Lancet*, **2004**, 363: 1377-1385;

[81] B. F. Pontz, H. Stoss, J. Spranger, *Ann N Y Acad Sci*, **1988**, 543: 30-39;

[82] A. Karbowski, M. Schwitalle, A. Eckardt, *Z Orthop Ihre Grenzgeb*, **1999**, 137: 219-22;

[83] B. Hansen, G. B. Jemec, *Arch Dermatol*, **2002**, 138: 909-911;

[84] P.P. Tsitsopoulos, I. Anagnostopoulos, V. Tsitouras, I. Venizelos, P. Tsitsopoulos, *Cent Eur J Med*, **2008**, 3: 517-520;

[85] M. C. Willing, C. J. Puchno, M. Atkinson, P. H. Byers, *A J Hum Genet,* **1992**, 51: 508-515;

[86] R. Morello, T. K. Bertin, Y. Chen, J. Hicks, L. Tonachini, M. Monticone, P. Castagnola, F. Rauch, F. H. Glorieux, J. Vranka, H. P. Bächinger, J. M. Pace, U. Schwarze, P. H. Byers, M. Weis, R. J. Fernandes, D. R. Eyre, Z. Yao, B. F. Boyce, B. Lee, *Cell*, **2006**, 127: 291-304;

[87] D. Baldridge, U. Schwarze, R. Morello, J. Lennington, T. K. Bertin, J. M. Pace, M. G. Pepin, M. Weis, D. R. Eyre, J. Walsh, D. Lambert, A. Green, H. Robinson, M. Michelson, G. Houge, C. Lindman, J. Martin, J. Ward, E. Lemyre, J. J. Mitchell, D. Krakow, D. L. Rimoin, D. H. Cohn, P. H. Byers, B. Lee, *Hum Mutat*, **2008**, 29: 1435-1442;

[88] J. C. Marini, A. R. Blissett, *J Clin Endocrinol Metab*, **2013**, 98: 3095-30103;

[89] Y. Alanay, H. Avaygan, N. Camacho, G.E. Utine, K. Boduroglu, D. Aktas, M. Alikasifoglu, E. Tuncbilek, D. Orhan, F. T. Bakar, B. Zabel, A. Superti-Furga, L. Bruckner-Tuderman, C. J. Curry, S. Pyott, P. H. Byers, D. R. Eyre, D. Baldridge, B. Lee, A. E. Merrill, E. C. Davis, D. H. Cohn, N. Akarsu, D. Krakow, *Am J Hum Genet*. **2010**, 86: 551-559;

[90] H. E. Christiansen, U. Schwarze, S. M. Pyott, A. AlSwaid, M. Al Balwi, S. Alrasheed, M. G. Pepin, M. A. Weis, D. R. Eyre, P. H. Byers, *Am J Hum Genet*, **2010**, 86: 389-398;

[91] H. Hoyer-Kuhn, C. Netzer, O. Semler, *Wien Med Wochenschr*, **2015**, DOI 10.1007/s10354-015-0361-x;

[92] J. Caffey, *J Pediatr*, **1946**, 29: 541–559;

[93] A. Kamoun-Goldrat, J. Martinovic, J. Saada, P. Sonigo-Cohen, F. Razavi, A. Munnich, M. Le Merrer, *Am Journal Med Genet Part A*, **2008**, 146A: 1820–1824;

[94] R. Y. Lau, X. Guo, *J Osteoporosis*, **2011**, doi:10.4061/2011/293808;

[95] A. J. Baileya, S. F. Wottonb, T. J. Simsa, P. W. Thompsonb, *Connect Tissue Res*, **1993**, 29: 119-132;

[96] S. Shuster, *Medical Hypotheses*, **2005**, 65, 426-432;

[97] R. Bernabei, A. M. Martone, E. Ortolani, F. Landi, E. Marzetti, *Clin Cases Miner Bone Metab*, **2014**, 11: 201-207;

[98] S. K. Das, A. Farooqi, *Best Pract Res Cl Rh*, **2008**, 22: 657-675;

[99] S. Krasnokutsky, M. Attur, G. Palmer, J. Samuels, S. B. Abramson, *Osteoarthr Cartil*, **2008**, 16: S1-S3;

[100] J. A. Martin, T. D. Brown, A. D. Heiner, J. D. Buckwalter, *Clin Orthop Relat Res*, **2004**, 427: S96-S103;

[101] K. Kuhn, D. D. D'Lima, S. Hashimoto, M. Lotz, *Osteoarthr Cartil*, **2004**, 12: 1-16.

[102] J. H. Chung, J. Y. Seo, H. R. Choi, M. K. Lee, C. S. Youn, G. Rhie, K. H. Cho, K. H. Kim, K. C. Park, H. C. Eun, *J Invest Dermatol*, **2001**, 117: 1218-1224;

[103] R. M. Lavker, *BA Gilchrest. Cambridge, MA, Blackwell Science*, **1995**, 123-135;

[104] J. Varani, D. Spearman, P. Perone, S. E. G. Fligiel, S. C. Datta, Z. Q. Wang, Y. Shao, S. Kang, G. J. Fisher, J.J. Voorhees, *Am J Pathol*, **2001**, 158, 931-942;

[105] J. Varani, M. K. Dame, L. Rittie, S. E.G. Fligiel, S. Kang, G. J. Fisher, J. J. Voorhees, *Am J Pathol*, **2006**, 168: 1861-1868;

[106] G. J. Fisher, S. C. Datta, H. S. Talwar, Z. Q. Wang, J. Varani, S. Kang, J. J. Voorhees, *Nature*, **1996**, 379: 335-338;

[107] G. J. Fisher, Z. Q. Wang, S. C. Datta, J. Varani, S. Kang, J. J. Voorhees, *New Eng J Med,* **1997**, 337: 1419-1428;

[108] A. J. Millis, J. Sottile, M. Hoyle, D. M. Mann, V. Diemer., *Exp Gerontol,* **1989**, 24: 559 -575;

[109] R. Ricciarelli, P. Mini, N. Ozer J. M. Zingg, A. Azzi, *Free Radic Biol Med,* **1999**, 27: 729-737;

[110] J. Varani, P. Perone, S. E. G. Fligiel, G. J. Fisher, J. J. Voorhees, *J Invest Dermatol,* **2002**, 119: 122-129;

[111] J. Varani, L. Schuger, M. K. Dame, C. Leonard, S. E. G. Fligiel, S. Kang, G. J. Fisher, J. J. Voorhees, *J Invest Dermatol*, **2004**, 122: 1471-1479;

[112] L. Otvos Jr, J. D. Wade, *Front Chem,* **2014**, 2: 1-4;

[113] K. Fosgerau, T. Hoffmann, *Drug Discov Today,* **2015**, 20: 122-127;

[114] M. Mutter, *Angew Chem*, **1985**, 97: 639-654;

[115] A. Nakai, M. Satoh, K. Hirayoshi, K. Nagata, *J. Cell Biol.*, **1992**, 117: 903;

[116] S. Saga, K. Nagata, W. T. Chen, K. M. Yamada, *J. Cell Biol.*, **1987**, 105: 517;

[117] M. Satoh, K. Hirayoshi, S. Yokota, N. Hosokawa, K. Nagata, *J. Cell Biol.*, **1996**, 133: 469;

[118] T. Natsume, T. Koide, S. Yokota, K. Hirayoshi, K. Nagata, *J. Biol. Chem.*, **1994**, 269: 31224;

[119] C. Widmer, J. M. Gebauer, E. Brunstein, S. Rosenbaum, F. Zaucke, C. Drögemüller, T. Leeb, U. Baumann, *PNAS*, **2012**, 109: 13243-13247;

[120] T. Koide, S. Asada, Y. Takahara, Y. Nishikawa, K. Nagata, K. Kitagawa, *J. Biol. Chem.*, **2006**, 281: 3432-3438;

[121] N. Jain, A. Brickenden, E. H. Ball, B. D. Sanwal, *Arch. Biochem. Biophys.*, 1994, 314: 23;

[122] E. H. Ball, N. Jain, B. D. Sanwal, *Adv. Exp. Med. Biol.*, 1997, 425:239;

[123] Y. Ishida, H. Kubota, A. Yamamoto, A. Kitamura, A. P. Bächinger, K. Nagata, *Mol. Biol. Cell*, **2006**, 17: 2346-2355;

[124] C. A. Thomson, V. S. Ananthanarayanan, *J.Biochem.*, **2000**, 349(3): 877;

[125] N. Nagai, M. Hosokawa, S. Itohara, E. Adachi, T. Matsushita, N. Hosokawa, K. Nagata, *J. Cell Biol.*, **2000**, 150: 1499;

[126] K. Nagata, *TIBS*, **1996**, 21: 22-26;

[127] K. Nagata, *Matrix Biol*, **1998**, 16:379;

[128] K. Nagata, *Sem. Cell Dev. Biol.*, **2003**, 14:275;

[129] K. Nagata, *TIBS*, **1996**, 21: 22-26;

[130] T. Taguchi, M. S. Razzaque, *Trends Mol Med*, **2007**, 13: 46-53;

[131] H. E. Christiansen, U. Schwarze, S. M. Pyott, A. AlSwaid, M. Al Balwi, S. Alrasheed, M. G. Pepin, M. A. Weis, D. R. Eyre, P. H. Byers, *Am. J. Hum. Gen.*, **2010**, 86: 389-398;

[132] U. Lindert, M. A. Weis, J. Rai, F. Seeliger, I. Hausser, T. Leeb, D. Eyre, M. Rohrbach, C. Giunta, *J Biol Chem*, **2015**, doi: 10.1074/jbc.M115.661025;

[133] O. Miyaishi, Y. Ito, K. Kozaki, T. Sato, H. Takechi, K. Nagata, S. Saga, *Mech Ageing Dev*, **1995**, 77: 231-226;

[134] S. Yokota, H. Kubota, Y. Matsuoka, M. Naitoh, D. Hirata, S. Minota, H. Takahashi, N. Fujii, K. Nagata, *Biochem Biophys Res Commun*, **2003**, 303: 413-418;

[135] Y. Masago, A. Hosoya, K. Kawasaki, S. Kawano, A. Nasu, J. Toguchida, K. Fujita, H. Nakamura, G. Kondoh, K. Nagata, *J Cell Sci*, **2011**, 125: 1118-1128;

[136] R. Huber, R.W. Carrell, *Biochemistry*, **1989**, 28: 8951-8966;

[137] J.W. Davids, T. S. El-Thaher, A. Nakai, K. Nagata, A. D. Miller, *Bioorg. Chem.*, **1995**, 23: 227-238;

[138] M. Mutter, S. Vuilleumier, *Angew Chem Int Ed Engl*, **1989**, 28: 535-554;

[139] M. H. Hecht, *Proc Nati Acad Sci USA*, **1994**, 91: 8729-8730;

[140] M. Mutter, *Proc 10th Am Peptide Symposium*, **1987**, ESCOM, Leiden, pp. 349-353;

[141] B. C. Cunningham, J. A. Wells, *Science*, **1989**, 224: 1081-1085;

[142] N. T. Ross, W. P. Katt, A. D. Hamilton, *Phil Trans R Soc A*, **2010**, 368: 989-1008;

[143] L. Regan, W. F. DeGrado, *Science*, **1988**, 241: 976-978;

[144] S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, M. H. Hecht, *Science*, **1993**, 262: 1680-1685;

[145] L. Malik, J. Nygaard, J. N. Cristensen, C. S. Madsen, H. I. Rösner, B. B. Kragelund, R. Hoiberg-Nielsen, W. W. Streicher, L. Arleth, P. W. Thulstrup, K. J. Jensen, *ChemBioChem*, **2015**, doi: 10.1002/cbic.201500285;

[146] M. H. Hecht, J. S. Richardson, D. C. Richardson, R. C. Ogden, *Science,* **1990**, 249: 884-891;

[147] L. Regan, D. Caballero, M. R. Hinrichsen, A. Virrueta, D. M. Williams, C. S. O'Hern, *Biopolymers*, **2015**, doi: 10.1002/bip.22639;

[148] T. P. Quinn, N. B. Tweedy, R. W. Williams, J. S. Richardson, D. C. Richardson, *Proc. Nat. Acad. Sci. USA*, **1994**, 91: 8747-8751;

[149] Y. Yan, B. W. Erickson, *Protein Sci*, **1994**, 3: 1069-1073;

[150] J. S. Nowick, *Acc Chem Res,* **2008**, 41: 1319-1330;

[151] G. Tuchscherer, M. Mutter, *J Biotechnol*, **1995**, 41: 197-210;

[152] P. L. Howard, M. C. Chia, S. Del Rizzo, F. -F Liu, T. Pawson, *Proc Natl Acad Sci USA,* **2003**, 100: 11267-11272;

[153] R. L. Baldwin, *Trends Biochem. Sci.* (Pers. Ed.) 11, **1986**, 6-9;

[154] E. T. Kaiser, F. J. Kezdy, *Science*, **1984**, 223: 249-255;

[155] M. Mutter, *Angew. Chem.*, **1985**, 97: 639-654;

[156] E. T. Kaiser, *Trends Biochem. Sci*, **1987**, 12: 305-309;

[157] M. Mutter, P. Dumy, P. Garrouste, C. Lehman, M. Mathieu, C. Peggion, S. Peluso, A. Razaname, G. Tuchscherer, Angew. *Chem Int Ed Engl*, **1996**, 35: 13-14;

[158] G. Tuscherer, M. Mutter, *J. Biotech.*, **1995**, 41: 197-210;

[159] C. A. Kim, J. M. Berg, *Nature*, **1993**, 362: 267-270;

[160] C. K. Smith, J. M. Withka, L. Regan, *Biochemistry*, **1994**, 33: 5510-5517.

[161] J. S. Nowick, S. Insaf, *J. Am. Chem. Soc*., **1997,** 119: 10903-10908;

[162] J. S. Nowick, N. A. Powell, E. J. Martinez, E. M. Smith, G. J. Noronha, *Org. Chem.* **1992**, 57, 3763-3765;

[163] T. Kortemme, M. Ramírez-Alvarado, L. Serrano, *Science*, **1998**: 253-256;

[164] J. S. Nowick, P. Ballester, F. Ebmeyer, J. Rebek, *J. Am. Chem. Soc.*, **1990**, 112: 8902-8906;

[165] S. H. Gellman, *Curr Opin Chem Biol,* **1998**, 2: 717-725;

[166] J. F. Espinosa, S. H. Gellman, *Angew. Chem.* **2000**, 39: 2330-2333;

[167] F. A. Syud, H. E. Stanger, H. S. Mortell, J. F. Espinosa, J. D. Fisk, C. G. Fry, S. H. Gellman, *J. Mol. Biol.,* **2003**, 326: 553–568;

[168] J. S. Nowick, J. H. Tsai, *J. Am. Chem. Soc.*, **1999**, 121: 8409-8410;

[169] H. C. Kolb, K. B. Sharpless, *Drug Discovery Today*, **2003**, 8: 1128-1137;

[170] V. D. Bock, R. Perciaccante, T. Jansen, H. Hiemstra, J. H. van Maarseveen, *Org. Lett.,* **2006,** 8: 919-922;

[171] X. Li, *Chem Asian J.*, **2011**, 6: 2606-2616;

[172] D. A. Case, T. A. Darden, T. A.; T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R., K. M. Merz, D. A. Perlman, M. Crowley, R. C. Walker, W. Zhang, W.; B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D. H. Mathews, C. Schafmeister, W. S. Ross, P. A. Kollman, *AMBER 9*, **2006**, University of California, San Francisco, CA, USA;

[173] J.S. Nowick, D.H. Holmes, G. Noronha, E.M. Smith, T.M. Nguyen, S. Huang, *J. Org. Chem.*, **1996**, 61: 3929-3934;

[174] J.S. Nowick, E.M. Smith, G. Noronha, *J. Org. Chem.*, **1995**, 60: 7386-7387;

[175] S.H. Nakagawa, E.T. Kaiser, *J Org Chem*, **1983**, 48: 678;

[176] E.T. Kaiser, H. Mihara, G.A. Laforet, J.W. Kelly, L. Walters, M.A. Findeis, T. Sasaki, *Science*, **1989**, 243: 187;

[177] C.N. Rao, D.A. Landin, Y.Y. Liu, K. Chilukuri, Z.Z. Hou, D.T. Woodley, *J Invest Dermatol,* **1995**, 105: 572-578;

[178] M. Cathomas, A. Schüller, D. Candinas, R. Inglin, *Int Wound J,* **2015**, 12, 601-604;

[179] A.M. Wachter, J. Lezdey, *Ann Allergy,* **1992**, 69: 407-414;

[180] K. Dabbagh, G.J. Laurent, A. Shock, P. Leoni, J. Papakrivopoulou, R.C. Chambers, *J Cell Physiol,* **2001**, 186: 73-81;

[181]A. O. Patschull, L. Segu, M. P. Nyon, D. A. Lomas, I. Nobeli, T. E. Barrett, B. Gooptu, *Acta Crystallogr*, **2011**, 67: 1492-1497;

[182] M.C. Gerbod-Giannone, A. Del Castillo-Olivares, S. Janciauskiene, G. Gil, P.B. Hylemon, *J Biol Chem,* **2002**, 277: 42973-42980;

[183] W. Dichtl, F. Moraga, M.P.S. Ares, M. Crisby, J. Nilsson, S. Lindgren, S. Janciauskiene, *Mol Cell Biol Res Commun,* **2000**, 4: 50-61;

[184] G. Joslin, R.J. Fallon, J. Bullock, S.P. Adams, D.H. Perlmutter, *J Biol Chem,* **1991**, 266: 11282-11288;

[185] S. Janciauskiene, H.T. Wright, S. Lindgren, *Atherosclerosis*, **1999**, 147: 263-275;

[186] L.F. Congote, N. Temmel, *FEBS Lett,* **2004**, 576: 343-347;

[187] J.A. Huntington, R.J. Read, R.W. Carrell, *Nature,* **2000**, 407: 923-926;

[188] L.F. Congote, N. Temmel, G. Sadvakassova, M.C. Dobocan, *Peptides*, **2008**, 29: 39-46;

[189] L. A. Carpino, *J Am Chem Soc*, **1993**, 115: 4397-4398;

[190] L. A. Carpino, A. J. El-Faham, *Org Chem*, **1995**, 60: 3561-3564;

[191] L. A. Carpino, F. J. Ferrer, *Org Lett*, **2001,** 3: 2793-2795;

[192] E. Kaiser, R.L. Colescott, C.D. Bossinger, P.I. Cook, Anal Biochem, **1970**, 34: 595-598;

[193] B. L. Lyons, R. I. Schwarz, *Nucleic Acids Res*, **1984**, 12: 2569-2579;

[194] R.A. Ignotz, T. Endo, J. Massagu, *J Biol Chem*, **1987**, 262: 6443-6446;

[195] M.P. Lupo, A.L. Cole, *Dermatol Ther,* **2007***,* 20: 343-349;

[196] K. Katayama, J. Armendariz-Borunda, R. Raghow, A.H. Kang, *Biochemistry*, **1991**, 30: 7097-7104;

[197] F.X. Maquart, L. Pickart, M. Laurent, P. Gillery, J. Monboisse, J. Borel, *FEBS Lett,* **1988**, 238: 343-346;

[198] G. Joslin, R.J. Fallon, J. Bullock, S.P. Adams, D.H. Perlmutter, *J Biol Chem*, **1991**, 266: 11282-11288;

[199] M.A. Niemann, J.E. Baggott, E.J. Miller, *Biochim Biophys Acta*, **1997**, 1340: 123-130;

[200] S. Janciauskiene, H.T. Wright, S. Lindgren, *Atherosclerosis*, **1999**, 147: 263-275;

[201] M.A. Niemann, J.E. Baggott, E.J. Miller, *J Cell Biochem*, **1997**, 6: 346-357;

[202] C. M. Choi, D. S. Berson, *Semin Cutan Med Surg,* **2009**, 6: 163-168;

[203] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R.Luo and T. Lee., *J. Comput. Chem.,*

2003*,* 24: 1999-2012;

[204] M.L. Connolly, *J. Appl. Cryst.* **1983,** 6: 548-558.