



## King's Research Portal

DOI:

[10.1039/C6CP03403A](https://doi.org/10.1039/C6CP03403A)

*Document Version*

Peer reviewed version

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Domene, C., Jorgensen, C., & Abbasi, S. W. (2016). A perspective on structural and computational work on collagen. *Physical Chemistry Chemical Physics*, 18(36), 24802-24811. <https://doi.org/10.1039/C6CP03403A>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

## A perspective on structural and computational work on collagen

Carmen Domene,<sup>\*,a,b</sup> Christian Jorgensen,<sup>a</sup> Sumra Wajid Abbasia<sup>a</sup>*Received (in XXX, XXX) Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXXX 20XX*

DOI: 10.1039/b000000x

Collagen is the single most abundant protein in the extracellular matrix in the animal kingdom, with remarkable structural and functional diversity and regarded one of the most useful biomaterials. Etymologically, the term collagen comes from Greek *kola* 'glue' and *gen* 'giving birth to'. Thus, it is not surprising that the various collagens and the structures they form all serve the same purpose, to help tissues withstand stretching. Among the functions the various collagens are involved in are cell adhesion and migration, tissue repair, scaffolding and morphogenesis. Thus knowledge about the structure and properties of collagen, how they change depending on the nature of the local environment as well as the nature and specificity of collagen interactions with its partners is central to discerning the role of collagen in medical applications such as imaging, drug delivery, and tissue engineering, and in the design and construction of synthetic collagen-like materials for tools in biomaterial science and nanotechnology. The main focus of this perspective is to review the molecular and packing structures of collagen and the computer simulations work performed so far, to further highlight the significance of collagen.

### Introduction

Collagen is a key structural protein responsible for the constitutional integrity of vertebrates.<sup>1</sup> It is a representative of insoluble crystalline proteins and a model for its structure was first proposed by Wyckoff, Corey and Biscoe in 1935 when they could not associate categorically the diffraction pattern with a chain-like structure or an ordinary molecular crystal built up by the regular arrangement of very large molecules.<sup>2</sup> Subsequently, Astbury & Bell<sup>3</sup> proposed that the collagen molecule consisted of a single extended polypeptide chain in the *cis* conformation. They suggested that the data presented earlier by Wyckoff and co-workers was better fitted using a gliadin class molecule. The first suggestion that the collagen structure might be helical was proposed by Pauling & Corey in 1951.<sup>4</sup> The model was a three helical polypeptide chain with a one third content of proline or hydroxyproline. In their manuscript they affirmed that 'the structure of the molecule provided an immediate explanation of the principal mechanical property of collagen'. Independently, Cohen & Bear and Cowan, North & Randall in 1953 suggested the correct non-integer screw axis.<sup>5</sup> One of Ramachandran's major contributions to structural biology was the proposal, made along with Kartha, of a triple-helical structure for the fibrous protein collagen published in *Nature* in 1954.<sup>6</sup> At a time when the structure of only one globular protein had been determined, the model they proposed solved the controversy regarding the so-called short inter-atomic distances in previous collagen models, and re-examination of them led to the famous 'Ramachandran plot' which graphically illustrates the restrictions on polypeptide conformation.<sup>7</sup> In 1955, this structure was refined by Rich & Crick<sup>8</sup> and by North and coworkers<sup>9</sup> to the triple-helical structure accepted today, with modifications focused on the arrangement of the hydrogen bonds as Ramachandra's models presented some stereochemical limitations.

The molecular organization in the collagen fiber was also later determined and the axial repeat usually designated as the D period was established. D-periodic fibrils contain intermolecular covalent cross-links that lead to their high tensile strength and mechanical stability.<sup>10</sup>

In 1994, Berman and co-workers reported the first high-resolution crystal structure of triple-helical collagen-related peptides<sup>11</sup> and subsequently, several other high resolution crystal structures of oligopeptides related to collagen as well as synthetic mimics were determined.

The umbrella term 'collagen' covers all proteins that form a right-handed three polypeptide three-dimensional complex (Figure 1). This denomination is used for all the members belonging to the collagen family characterised by varying tissue distribution, size and function.<sup>1a, 12</sup> Collagen is conserved in the animal kingdom, and it is a key component of the majority of tissues accounting for about 20–30% of total body proteins.<sup>13</sup> Collagen is a primary component of bones, muscles, skin and tendons of vertebrates supporting delicate organs.<sup>14</sup> Collagen types are classified in several sub-families according to sequence homologies and to similarities in their structural organization and supramolecular such as fibrils, networks and filaments (Scheme 1). To date, at least 29 different types of collagen have been identified with the designation I-XXVIII,<sup>15</sup> along with additional proteins that have collagen-like domains including adiponectin, C1q macrophage receptor, acetyl cholinesterase, conglutinin, ectodysplasin collectin or ficolin among many others (Figure 2).<sup>1a</sup> Crucially, insight into collagen-protein interactions will facilitate the progress in new approaches in drug discovery, targeting and delivery.<sup>16</sup> Collagen types I, II and III account for the majority of collagen in the human body, and about one half of the total body collagen is in the skin, and about 70% of the material other than water present in dermis of skin and tendon is collagen.<sup>17</sup> In contrast, invertebrate collagen genes encode only for fibrillar and basement membrane collagen.<sup>18</sup> The presence of over a hundred

bacterial protein sequences containing the collagen characteristic domain in genomic databases has been reported, which suggests a whole new family of collagen-like proteins.<sup>19</sup>

Therefore, understanding how the relatively simple molecular building blocks of collagen self-assemble is imperative for artificial tissue development, growth, regeneration, and disease, cosmetics formulation, pharmacology or in its application in plastic surgery and medicine in general.<sup>17</sup>

### Structural characteristics of collagen

The molecular structure of collagen was assigned based on the evidence from studies of chemical composition, physicochemical analysis of solutions, X-ray crystallography and electron microscopy, and a variety of spectroscopy techniques.<sup>20</sup> Under physiological conditions, the basic collagen molecule is rod-shaped with a length and a width of about 3000 and 15 Å, respectively.<sup>21</sup> It consists of three polypeptide chains twined around one another with individual twists in the opposite directions and flanked at both termini by short non-helical telopeptides also known as amino-terminal collagen crosslinks.

The three chains are staggered by one residue which creates non-identical environments for the three chains. Generally, the triple helix is composed by two identical chains and a third chain that differs in its chemical composition. The main constituents of collagen are glycine, which is found at almost every third residue, proline that makes up ~28% of collagen, and two uncommon derivative amino acids, hydroxyproline and hydroxyllysine both formed by enzymatic post-translational modification. Stereoelectronic effects and preorganization play a key role in collagen stability.<sup>1d</sup> Each chain contains a repeating (Gly-Pro-Y) or (Gly-X-Hyp) triad, consisting of a glycine repeat motif followed by proline or 4-hydroxyproline and various other residues in the X and Y positions, respectively (Figure 1). The most common triplet in collagen is Pro-Hyp-Gly accounting for about 10% of the total sequence, and the peptide (Pro-Hyp-Gly)<sub>10</sub> forms a very stable triple helix. A high content of glycine and amino acid residues dictates the principal features of collagen.<sup>22</sup> The strands are held together primarily by inter-chain hydrogen bonds formed between individual chains with the involvement of hydroxyproline residues and facilitated by the high content of Gly residues.<sup>23</sup> A single collagen molecule is referred as tropocollagen. Collagen structures are either homotrimers or heterotrimers with either two or three dissimilar chains.<sup>24</sup> A group of tropocollagen molecules self-assemble to form fibrils having a distinct periodicity, and they also differ in their ability to form fibres and to organize the fibres into networks. Some covalent crosslinking within the triple helices can be identified as well as a variable amount of covalent crosslinking between tropocollagen molecules forming well organized aggregates.<sup>1c</sup> A group of tropocollagen may be broadly divided into fibrillar types and non-fibrillar collagens. The fibrillar collagen chains all have a perfect (Gly-X-Y)<sub>n</sub> repeating sequence. In contrast, all non-fibrillar collagens have sites where the repeating tripeptide pattern is interrupted.<sup>25</sup>

Collagen I, IV, V, VI, IX, XI are heterotrimers, and the remaining are homotrimers. Among the 29 different types of collagens characterised up to now, type I, II, III, V, and XI are fibril-forming collagens. Collagen fibrils are characterised by

remarkable strength and stability conferred by its self-aggregation and cross-linking; for instance collagen type I fibrils are stronger than steel. These fibrils assemble into well-structured supra-molecular linear aggregates of length >1µm with a characteristic supra-structure. Fibrillar collagens contain a relatively high content of charged residues (~15%–20%) and a small percentage of hydrophobic residues (~6%).<sup>16</sup> Fibers can form hydrogels, films or sponges. Several models have been also proposed for the arrangement of the tropocollagen within collagen fibrils. One of the initial ones was by Hodge and Petruska.<sup>26</sup> In this model, five tropocollagen molecules are staggered side-by-side with an offset of 67 nm between two neighbors as revealed by transmission electron microscopy. Subsequently, Schmitt et al. proposed that 234 amino acid residues is the period of the helix. In other words, the length of the tropocollagen molecule is about 4.4 times that of the native collagen period.<sup>27</sup> Not all collagens occur as periodic-structured fibrils though. However, the model cannot describe the spatial extension of the quarter staggered in two or three dimensions. This model is known as the Quarter staggered stacking model. Subsequently,<sup>27</sup> the Smith model was proposed where five tropocollagen molecules are arranged concentrically into a hollow filament, known as the microfibril.<sup>28</sup> Its limitation is that it cannot predict the organisation of fibrils with a diameter greater than 3.5 nm. Next, Hulmes and Miller proposed the Quasihexagonal packing model where periodic tropocollagen molecules were assigned the character of a molecular crystal and without microfibrillar sub-structures.<sup>29</sup> The Compressed microfibril model followed.<sup>30</sup> In this model, five-stranded microfibrils are compressed to place the molecules in fibril cross-section on a pseudo-hexagonal lattice, and in the longitudinal direction the molecules are supercoiled with a left handed twist. Using x-ray diffraction, Orgel et al. presented the Microfibrillar model which supported a microfibril structure composed by five staggered tropocollagen molecules arranged with a right-handed tilt, rather than just axially staggered.<sup>14f</sup> This model seems to be the one that best fits the native x-ray diffraction data and other experimental observations on the organization of the molecular segments in the overlap region of the fibril.

Type I collagen of fibril-forming collagens makes up over than 90% of organic mass of bones, as well as being the main collagen in a number connective tissues, including skin, tendons, cornea, ligaments, as well as vitreous body, brain, cartilage and hyaline tissues.<sup>1b</sup> Type II collagen makes up over 50% of all protein in cartilage and 85-90% of collagen of articular cartilage.<sup>1b</sup> Type III collagen is essential for Type I fibrillogenesis. Normal type I collagen is a heterotrimer triple-helical linear molecule consisting of two α-1 chains and one α-2 chain.<sup>31</sup> In contrast, type II and III form homotrimers, and can assemble into globular homotrimeric domains.

The fibril-associated collagens with interrupted triple helices (FACIT) subclass, comprise collagen types IX, XII, XIV, XVI, XIX, and XX; these structure are short non-helical domains.<sup>1b</sup> Type VI is a heterotrimeric collagen with small helical domains and stretched globular termini.<sup>32</sup> Stretching the α-3-chain of type VI, of double length when compared to the rest of chains, is due to the larger globular domain at N- and C-termini. These extended collagen domains are subject to extensive intracellular and extracellular post-translational modifications.<sup>33</sup>

Collagen types VII and X are formed by short chains, and type X, is homotrimeric exhibiting long C-terminal and short N-terminal domains. *In vivo* studies revealed it assembles into hexagonal networks.<sup>34</sup> Type VIII collagen is similar in structure to type X collagen, although it possesses a different spatial distribution which confers distinctive function.<sup>35</sup>

Collagen type IV is predominant in membranes, assimilating nitrogen atoms, the laminin proteins of the extracellular matrix, and other elements into a 2-dimensional supra-molecular aggregate. Type IV consists of a conformational flexible triple-helical structure with three domains. Until now, there are six recognized subunit chains ( $\alpha 1$  (IV)– $\alpha 6$  (IV)). Among these chains,  $\alpha 1$  (IV)  $2\alpha 2$  (IV) heterotrimers have been reported to be vital in forming important network in most embryonic as well as adult basement membranes.

Numerous experimental studies link the misregulation of collagen to a broad range of diseases. Collagen serves as binding sites for cytokines and multiple growth factor proteins. These cytokines and growth factors in turn regulate vital cell functions including survival, differentiation, motility and polarity.<sup>1e, 36</sup> One of the therapeutic potential for collagen is in the field of drug delivery, where the binding ability of collagen makes it a promising agent in the delivery of drugs, while the anchoring and network-generation ability of some collagen types has potential in tissue regeneration and repair.<sup>37</sup> Furthermore, experimental studies have widely pointed to deficiencies in type III collagen, as well as elastin, as linked to cardiac aneurysm formation.<sup>38</sup> Finally, collagen bio prostheses have been studied for roles in surgery.<sup>39</sup> Mutations in different regions can have different effects and defects at the molecular structure or collagen organization into mature fibers result in different diseases associated with connective tissues and even some types of osteoporosis and arthritis (1, 2). A number of excellent books and reviews on collagen describe the biochemical and biomedical aspects in detail,<sup>10, 40</sup> and some others focus on describing how understanding the key biochemical and physical properties of collagen lead to strategies to create, control and modify the structure and function of collagen-based biomaterials.<sup>16, 41</sup>

### Collagen-like peptide models and collagen-protein systems

The inability to crystallize collagen and its unsuitability for NMR characterisation rendered collagen model peptides an invaluable alternative to study native collagen structural features using biophysical techniques. Advances in peptide chemistry, in the 60's, set the stage for the synthesis and characterization of poly-tripeptides to be used as collagen models.<sup>25</sup> In these systems, the packing is reminiscent of collagen and usually, there is a high proportion of Pro-Hyp-Gly triplets. These studies have explored the effect of amino and imino acid content on the triple-helix structure, they have unravelled the basic principles of collagen self-assembly and into supramolecular structures and they have aided in clarifying the sequence dependent nature of biologically important features of native collagen and their alterations in diseased states as well as. Collagen mimetic peptide assemblies and functionalized collagen mimetic peptides have been also synthesised, and these have been reviewed elsewhere.<sup>42</sup> Bella and co-workers<sup>11c</sup> designed and crystallised at 1.9 Å resolution, a peptide to model the effect of interrupting the

repeating (Gly-X-Y) motive with a single Gly substitution to Ala at the center of a 30-amino acid peptide as these substitution had been identified in several diseases. The crystal structure published in 1994 provided structural information on the effect of a glycine substitution in a triple helix, an alteration which usually leads to pathological states in fibrillar collagens.

A first structural analysis of the polypeptide (Pro-Gly-Pro)<sub>n</sub> was reported by Yonath and Traub in 1969<sup>43</sup> followed by work on fiber diffraction by Scheraga<sup>44</sup> and Blout's groups.<sup>45</sup> In 1981, Okuyama et al<sup>11b</sup> presented crystallographic studies on the polypeptide model (Pro-Pro-Gly)<sub>10</sub> and reported 72 symmetry as opposed to the 103 triple helical symmetry in earlier natural collagen studies which animated a debate around the actual symmetry of natural collagen. In the 90's, new polypeptide structural models were synthesized and characterized, and continued to illustrate key features of collagen related systems. More recently, the first report of the full-length structure of the collagen-like polypeptide [(Pro-Pro-Gly)<sub>10</sub>]<sub>3</sub> at 1.3 Å was given by Berisio and co-workers in 2001. Model peptides have also been used to define the basic principles of collagen self-association to the supramolecular structures found in tissues.<sup>25, 46</sup> The interactions between the triple helical structure of collagen and proteins play important roles in collagen binding and degradation and for example in healing and repair of the body's tissues. Crystallography has yielded atomistic structures of a variety of collagen types, which has allowed for detailed studies of collagen complexes. Over the past decades, much information has been also gained about the interactions of collagen with cell surface receptors, extracellular matrix components and enzymes such as matrix metalloproteinase. Currently, around over 300 different crystal structures of collagen in complex with other proteins have been reported, and some of these selected examples will be described next to illustrate relevant pivotal interactions and their relation to function. In addition, the triple helix structural motif is found in a few non-collagenous proteins. The binding of a monoclonal antibody (MAb) to the triple-helical region of type III collagen was one of the first examples where a region of collagen which binds to another molecule was studied in some detail to clarify specific recognition and binding properties.<sup>47</sup> The molecular features involved in triple-helix interactions with another macromolecule were characterised, and the observation of unstable Gly-Gly-Y triplets adjacent to the recognition region was made, suggesting involvement of some flexibility or instability near the actual binding site. Human cysteine cathepsin is a protein crucial in pathophysiological and physiological cellular mechanisms, and is a key therapeutic agent for a range of diseases as it hydrolyses various extracellular matrix components among which are some types of collagens. Sage et al have described an inhibition mechanism of this protein mediated by glycosaminoglycan that involved *in vivo* modulation of its collagenase activity.<sup>48</sup> Another process of vital importance where protein-collagen interactions are crucial is the degradation of collagen to maintain the correct collagen homeostasis in tissues. In the collagenases, hemopexin C domain exosites bind native collagen, which is required for triple helicase activity during collagen cleavage. The active site of collagenolytic matrix metalloproteinases (MMPs) can only accommodate a single-chain of collagen. Thus the

collagen helix must be initially unwound by a triple helicase in order to expose the scissile bonds, and then cleavage of the chains occurs sequentially. The Overall lab has proposed models for the regulation of type I collagen levels upon stimulation of the activity of several matrix metalloproteinases. The collagen binding properties and the role of the ectodomain and the hemopexin C domain of the collagenolytic membrane type-1 matrix metalloproteinase (MT1-MMP) in collagenolysis were detailed characterised.<sup>49</sup> It was reported that collagen is a unique substrate for the proteases responsible for its cleavage, and these interactions recruit and regulate collagenolytic and gelatinolytic activities in a homeostatic manner.<sup>49</sup>

### Computer simulations and modelling of collagen systems

It has become possible to clarify at the atomic level the features and interactions of collagen in various forms and environments, primarily due to progress in the experimental determination of three-dimensional structures of collagen and collagen-protein complexes together with improved algorithms for computer simulations and the increasing speed and availability of supercomputers.

Molecular dynamics (MD) is a powerful computational technique that provides accurate descriptions of the structure and dynamics of biological systems, contributing to their understanding at an atomic level. In MD simulations the motion of interacting atoms is calculated by integrating Newton's equations of motion. The potential energy of the system and the forces, derived from the negative gradient of the potential with respect to displacements in a specified direction, are used to forecast the time evolution of the system in the form of a trajectory. Equilibrium quantities are then calculated using statistical mechanics by averaging over trajectories of sufficient length which would have sampled a representative ensemble of the state of the system. Specific MD procedures to study collagen tightly packed have been described in the literature.<sup>50</sup> These protocols are notably different from conventional MD simulations of proteins, which generally only treat individual protein molecules or complexes fully solvated.<sup>50</sup> These alternative approaches exploit ideas borrowed from modelling crystalline solids such as periodic boundaries to replicate the super molecular arrangement of collagen proteins within fibrils.<sup>50</sup> Numerous MD simulation studies applied to collagen have provided useful information to relate the diverse structural characteristics with their function.<sup>51,52</sup> A representative example of a simulation system of tropocollagen-peptide complex in solution is illustrated in figure 3.

Early computational studies in this area focused on microfibril and fibril packing issues such as longitudinal alignment and stabilization of fibrils.<sup>1c, 53</sup> Subsequently, collagen-like peptides were used in combined experimental and theoretical calculations to get insight into the structural features of collagen.<sup>54</sup> The structural importance of prolines and hydroxyproline for helix and fibril stability was established using computational models by Sheraga and coworkers.<sup>53c</sup> However, the first simulations in this area are the 0.5–1.0 ns molecular dynamics simulations of collagen-like peptides performed by Klein and Huang.<sup>55</sup> This was followed by MD simulations studying the role hydroxylated prolines play in stabilizing the collagen triple helix.<sup>56</sup>

Subsequent work focused on the modelling of telopeptides, which are crucial for the formation of enzymatic covalent crosslinks that form in collagens near their N- and C-ends, as these crosslinks provide structural integrity, strength, and stiffness to collagenous tissues.<sup>57</sup> One study reported conformational and packing studies of cross-linked structures of the fibril-forming Type I collagen N-telopeptide heterotrimer.<sup>58</sup> Due to the absence of high-resolution crystallographic structures of telopeptides, a triple-helical structure was built on the basis of crystallographic coordinates of a collagen-like sequence coordinates and then replaced with the actual bovine collagen residues. It was further found that if individual N-telopeptides were considered, their chain structures were essentially random, but when they were docked to their helix domain receptors, very ordered and specific conformations were created.<sup>58</sup> A second study addressed type I collagen C-telopeptide conformations using all three chains of the heterotrimer before and after it was docked to its receptor domain.<sup>59</sup> The computational models showed that the N- and C-telopeptide regions have different molecular packing and intrafibrillar crosslinking patterns that control the relative azimuthal orientations of molecules in the fibril.<sup>59</sup> In a later study, the deformation mechanisms of N- and C-crosslinks and the functional roles for the N- and C-telopeptide conformations were investigated via MD simulations.<sup>57</sup>

Other computational work focused on mutagenic disruptions of collagen functioning, and their associated pathologies.<sup>60</sup> In one study,<sup>61</sup> collagen-like molecules designed to mimic the site of mutations in collagen type I are used in combination with MD simulations to contrast general structural properties of the peptides with and without the mutation to examine the effect of the single-point mutation on the surrounding residues.

On the development front, a novel set of molecular mechanics parameters for hydroxyproline by Park *et al.* allowed for the reproduction of the correct pucker preference of the collagen backbone motif, which were tested in a set of simulations of collagen-like peptides. The role of hydroxylation in the stability of the collagen triple helix by adjusting to the right pucker conformation was reproduced.<sup>62</sup> Various other studies have been performed by several groups to investigate the relationship between interchain salt bridge formation and triple-helical stability using detailed molecular simulations with the aim to guide the design of collagen-like peptides that have specific interchain interactions.<sup>63</sup> To further clarify the stereospecificity of ion pairs, MD simulations were computed for triple-helical peptides containing reversed sequences, comparing EGK with KGE, for example. In combination with experimental studies, the results indicated that the reversal of charges lowered thermal stability, highlighting the importance of cross-chain ionic interactions for the stability of the collagen triple helix in solution.<sup>64</sup>

MD simulations together with experiment have investigated the pathways and molecular mechanisms for peptide assembly into triple-helical protomers, as well as their subsequent organization into structurally defined, linear assemblies.<sup>65</sup> These studies showed that collagen-mimetic fibrils and microfibers, which are very similar to those formed *in vivo*, could be obtained through the linear assembly of a small collagen-mimetic peptide driven through electrostatic interactions with precisely defined periodic

features with potential applications in material design. Experimental studies showed the positional preference of different amino acids to form a stable triple helical collagen motif, and the structural basis for the variations in the sequence and the positional propensity was systematically investigated with computational techniques.<sup>66</sup> Specifically, MD simulations were carried out on 39 collagen-like peptides showing that the propensity of the different amino acids to adopt collagen-like conformations depends primarily on their  $\phi$  and  $\psi$  angle preferences.<sup>66</sup>

Several experimental and modelling studies have been carried out to understand mechanical properties of bone, a biological nanocomposite that exhibits a highly optimized and complex multi-level hierarchical structure composed primarily of Type I collagen and hydroxyapatite.<sup>67</sup> Among the computational studies, molecular dynamics and steered molecular dynamics were employed to study directional dependence of deformation response of collagen with respect to the hydroxyapatite surface<sup>67</sup> and collagen interactions with rutile surfaces without hydroxylation.<sup>68</sup> The early process in the nucleation of hydroxyapatite at a collagen template was studied by immersing a triple helical collagen molecule in a stoichiometric solution of  $\text{Ca}^{2+}$ ,  $\text{PO}_4^{3-}$  and  $\text{OH}^-$  ions and compared with simulations of collagen interacting with surfaces of hydroxyapatite from the crystal.<sup>69</sup> In the context of drug delivery, medical diagnosis and molecular engineering, the interactions of collagen-like peptides with carbon nanotubes (CNTs) were also investigated with MD simulations.<sup>70</sup> A collagen-like peptide with a hydrophobic center and hydrophilic surfaces could be inserted spontaneously but slowly and the mechanism of the encapsulation process was characterised. Other examples of computational studies of the interaction of collagen with materials are those with gold nanoparticles,<sup>71</sup> graphene nanoribbons<sup>72</sup> as well as those to get insight into collagen self-assembly on substrates.<sup>73</sup>

There are numerous computational studies illustrating collagen-protein interactions, and few selected ones are highlighted below. To start, the binding of a C-terminal fragment of collagen XVIII, endostatin, to heparin and heparan sulfate was studied experimentally and further characterised by docking and molecular dynamics simulations.<sup>74</sup> Endostatin interacts with the heparan sulfate chains of the cell surface contributing to its biological activities. The aim of this study was to determine the affinity of these interactions, identify the structural features of heparin/heparan sulfate-endostatin complexes, as well as to investigate the effect of divalent cations on the interaction.<sup>74</sup>

Collagen type II is a specific target in the collagen-induced arthritis model. In a study using a homology model of an antigen-antibody complex, using 200 ns MD trajectory, the critical amino acids conferring Collagen type II epitope specificity to a variety of autoantibodies were investigated. The presence of a few anchoring residues in the antibody regions was shown to be probably sufficient to confer a moderately high affinity key for the recognition.<sup>75</sup>

A combined computational and experimental study illustrated the nature of the ligand-receptor interactions between single and triple-helical strands of collagen and integrin. Integrins are the main receptor proteins that cells use to both bind to and respond to the extracellular matrix, with these interactions regulating

many different cell functions, and thus a detailed understanding of the fine-tuning of collagen binding to integrin is essential as it might be a potential tool for therapeutic purposes.<sup>76</sup> Combined NMR and MD simulation methods addressed the question of why single-stranded collagen fragments are unable to establish a stable specific binding interaction with the integrin receptor, finding only weak complexes in solution. In another study,<sup>77</sup> a biomimetic design strategy of platelet adhesion inhibitors was proposed to develop potent inhibitors for the integrin  $\alpha 2\beta 1$ -collagen binding, using a combination of molecular docking, structure similarity analysis, MD simulations and experimental validation. These studies highlight the potential for combined *in silico* and *in vitro* for furthering our understanding of collagen-receptor interactions.

## Conclusions

Collagen is an example of a naturally occurring structural material composed by relatively simple chemical building blocks, the amino acids. The numerous types of collagen described in the literature have been reviewed in terms of their structure and interactions with other proteins and peptides. In addition, an overview of the computational studies carried out up-to-date has been presented, emphasising the potential of MD simulations to enhance our understanding of collagen. The importance of collagen has been highlighted through various of its applications, including the design of modern engineered biomaterials with precise formulated properties to serve a unique purpose, as well as constituting novel potent targets for the pharmaceutical industry for *de novo* drug design against a number of diseases as well as a drug delivery system owing to its excellent biocompatibility and safety. Collagen-linked pathologies have been reviewed, for which more than a thousand mutations are known to result in various diseases, including osteoporosis, arterial aneurysms, osteoarthritis, as well as rare diseases.<sup>1a</sup> Tropoelastin has emerged as key biomarker to identify patients that are at risk of a heart attack, and binding imaging agents are under development to target tropoelastin and to provide an imagable signal that will allow clinicians to identify the location and extend of vulnerable plaques.<sup>78</sup> The biomarker consists of a small peptide binding sequence coupled to a radioactive element that is detectable by MRI.<sup>78</sup> Consequently, further knowledge on collagen interactions in bulk materials and in protein environments, is central to understand its role in such medical applications, and in the design of tools in biomaterial science and nanotechnology.<sup>17</sup>

## Acknowledgements

C. Domene thanks the Biotechnology and Biological Sciences Research Council and the Engineering and Physical Sciences Research Council for financial support. C. Jorgensen thanks King's College London for a Graduate Teaching Assistant studentship. S.W. Abbasi thanks the High Commission of Pakistan for a travel fellowship award.

## Notes and references

<sup>a</sup>Department of Chemistry, King's College London, Britannia House, 7 Trinity Street, London SE1 1DB, UK Tel: 44 - (0)207848754; E-mail: carmen.domene@kcl.ac.uk

<sup>b</sup>Chemistry Research Laboratory, Mansfield Road, University of Oxford, Oxford OX1 3TA, UK

1. (a) Myllyharju, J.; Kivirikko, K. I., Collagens and collagen-related diseases. *Ann Med* **2001**, *33* (1), 7-21; (b) K. Gelsea, E. P., T. Aigner, a., Collagens—structure, function, and biosynthesis. *Advanced Drug Delivery Reviews* **2003**, *55* (12), 1531-1546; (c) Perumal, S.; Antipova, O.; Orgel, J. P. R. O., Collagen fibril architecture, domain organization, and triple-helical conformation govern its proteolysis. *Proceedings of the National Academy of Sciences* **2008**, *105* (8), 2824-2829; (d) Shoulders, M. D.; Raines, R. T., Collagen Structure and Stability. In *Annual Review of Biochemistry*, 2009; Vol. 78, pp 929-958; (e) Brown, J. C. T., R., The Collagen Superfamily. *Int Arch Allergy Imm* **1995**, *107* (4), 484-490.
2. Wyckoff, R. W.; Corey, R. B.; Biscoe, J., X-ray reflections of long spacing from tendon. *Science (New York, N.Y.)* **1935**, *82* (2121), 175-6.
3. Astbury, W.; Bell, F., The molecular structure of the fibers of the collagen group. *Nature* **1940**, *145*, 421-422.
4. Pauling, L.; Corey, R. B., The structure of fibrous proteins of the collagen-gelatin group. *P Natl Acad Sci USA* **1951**, *37* (5), 272-281.
5. (a) Cohen, C.; Bear, R. S., Helical polypeptide chain configuration in collagen. *Journal of the American Chemical Society* **1953**, *75* (11), 2783-2784; (b) Cowan, P. M.; North, A. T. C.; Randall, J. T., In *The Nature and Structure of Collagen*, Randall, J. T., Ed. Butterworths.: London, 1953; p 241.
6. (a) Ramachandran, G. N.; Kartha, G., Structure of collagen. *Nature* **1954**, *174* (4423), 269-270; (b) Ramachandran, G. N.; Kartha, G., Structure of collagen. *Nature* **1955**, *176* (4482), 593-595.
7. Bhattacharjee, A.; Bansal, M., Collagen structure: The Madras triple helix and the current scenario. *Iubmb Life* **2005**, *57* (3), 161-172.
8. Rich, A.; Crick, F. H. C., Molecular structure of collagen. *Journal of Molecular Biology* **1961**, *3* (5), 483-&.
9. Cowan, P. M.; McGavin, S.; North, A. C. T., Polypeptide chain configuration of collagen. *Nature* **1955**, *176* (4492), 1062-1064.
10. Barbara Brodsky, B.; Werkmeister, J.; Ramshaw, J., Collagens and Gelatins. Wiley-VCH, Ed. Vol. 108, pp 119-128.
11. (a) Kramer, R. Z.; Bella, J.; Brodsky, B.; Berman, H. M., The crystal and molecular structure of a collagen-like peptide with a biologically relevant sequence. *Journal of Molecular Biology* **2001**, *311* (1), 131-147; (b) Okuyama, K.; Okuyama, K.; Arnott, S.; Takayanagi, M.; Kakudo, M., Crystal and molecular-structure of a collagen-like polypeptide (PRO-PRO-GLY)<sub>10</sub>. *Journal of Molecular Biology* **1981**, *152* (2), 427-443; (c) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M., Crystal-structure and molecular structure of a collagen-like peptide at 1.9-Angstrom resolution. *Science* **1994**, *266* (5182), 75-81; (d) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A., Crystal structure of the collagen triple helix model (Pro-Pro-Gly)<sub>10</sub> (3). *Protein Science* **2002**, *11* (2), 262-270.
12. (a) MARK, K. V. D., Structure, biosynthesis and gene regulation of collagens in cartilage and bone, Dynamics of Bone and Cartilage Metabolism. In *Dynamics of Bone and Cartilage Metabolism: Principles and Clinical Applications* Markus Seibel, S. R., John Bilezikian, Ed. Academic Press: 2006; pp 3-26; (b) Ku'hn, K., The collagen family-variations in the molecular and supermolecular structure. *Rheumatology* **1986**, *10*, 29-69; (c) Mayne, R.; Brewton, R. G., New members of the collagen superfamily. *Current opinion in cell biology* **1993**, *5* (5), 883-90; (d) van der Rest, M.; Garrone, R., Collagen family of proteins. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **1991**, *5* (13), 2814-23; (e) Sato, K.; Yomogida, K.; Wada, T.; Yorihuzi, T.; Nishimune, Y.; Hosokawa, N.; Nagata, K., Type XXVI collagen, a new member of the collagen family, is specifically expressed in the testis and ovary. *J Biol Chem* **2002**, *277* (40), 37678-37684.
13. (a) Kadler, K., Extracellular matrix 1: Fibril-forming collagens. *Protein profile* **1995**, *2* (5), 491-619; (b) Harkness, R. D., Biological functions of collagen. *Biological Reviews of the Cambridge Philosophical Society* **1961**, *36* (4), 399-&.
14. (a) Vincent, J. F. V., *Structural biomaterials*. Rev. ed.; Princeton University Press: Princeton, N.J., 1990; p xii, 244 p; (b) Peter Fratzl, R. W., Nature's hierarchical materials. *Progress in Materials Science* **2007**, *52* (8), 1263-1334; (c) Nalla, R. K.; Kruzic, J. J.; Kinney, J. H.; Ritchie, R. O., Mechanistic aspects of fracture and R-curve behavior in human cortical bone. *Biomaterials* **2005**, *26* (2), 217-231; (d) Ritchie, R. O.; Kruzic, J. J.; Muhlstein, C. L.; Nalla, R. K., & Stach, E. A., Characteristic dimensions and the micro-mechanisms of fracture and fatigue in nano'andbio'materials. *International journal of fracture* **2004**, *128* (1-4), 1-15; (e) Hulmes, D. J. S.; Wess, T. J.; Prockop, D. J.; Fratzl, P., Radial Packing, Order, and Disorder in Collagen Fibrils. *Biophys J* **1995**, *68* (5), 1661-1670; (f) Orgel, J. P. R. O.; Irving, T. C.; Miller, A.; Wess, T. J., Microfibrillar structure of type I collagen in situ. *P Natl Acad Sci USA* **2006**, *103* (24), 9001-9005; (g) Freeman, J. W.; Silver, F. H., Elastic energy storage in unmineralized and mineralized extracellular matrices (ECMs): A comparison between molecular modeling and experimental measurements. *J Theor Biol* **2004**, *229* (3), 371-381.
15. Ricard-Blum, S., The Collagen Family. *Cold Spring Harbor Perspectives in Biology* **2010**, *3* (1), a004978-a004978.
16. An, B.; Lin, Y.-S.; Brodsky, B., Collagen interactions: Drug design and delivery. *Advanced Drug Delivery Reviews* **2016**, *97*, 69-84.
17. Lee, C. H.; Singla, A.; Lee, Y., Biomedical applications of collagen. *International Journal of Pharmaceutics* **2001**, *221* (1-2), 1-22.
18. Huxley-Jones, J.; Robertson, D. L.; Boot-Handford, R. P., On the origins of the extracellular matrix in vertebrates. *Matrix Biol* **2007**, *26* (1), 2-11.
19. Xu, C.; Yu, Z.; Inouye, M.; Brodsky, B.; Mirochnitchenko, O., Expanding the family of collagen proteins: Recombinant bacterial collagens of varying composition form triple-helices of similar stability. *Biomacromolecules* **2010**, *11* (2), 348.
20. (a) Gross, J., An intermolecular defect of collagen in experimental lathyrism. *Biochimica Et Biophysica Acta* **1963**, *71* (1), 250-&; (b) Ramachan.Gn; Sasisekh.V, Refinement of structure of collagen. *Biochimica Et Biophysica Acta* **1965**, *109* (1), 314-&.
21. Traub, W.; Piez, K. A., The chemistry and structure of collagen. *Advances in protein chemistry* **1971**, *25*, 243-352.
22. Piez, K. A., Molecular and aggregate structures of the collagens. In *Extracellular Matrix Biochemistry*, K.A. Piez, A. H. R. E., Ed. Elsevier: New York, 1984; pp pp 1-40.
23. Brodsky, B.; Persikov, A. V., Molecular structure of the collagen triple helix. *Adv Protein Chem* **2005**, *70*, 301+.
24. Ramachandran, G. N., Structure of Collagen. *Nature* **1956**, *177* (4511), 710-711.
25. Brodsky, B.; Thiagarajan, G.; Madhan, B.; Kar, K., Triple-helical peptides: An approach to collagen conformation, stability, and self-association. *Biopolymers* **2008**, *89* (5), 345-353.
26. Petruska, J. A.; Hodge, A. J., SUBUNIT MODEL FOR TROPICOLLAGEN MACROMOLECULE. *P Natl Acad Sci USA* **1964**, *51* (5), 871-&.
27. Schmitt, C.; CE, H.; Jakns, M., Electron microscopy investigations of the structure of collagens. *J Cell Comp Physiol* **1942**, *11*-33.
28. Smith, J. W., Molecular Pattern in Native Collagen. *Nature* **1968**, *219* (5150), 157-158.
29. Hulmes, D. J. S.; Miller, A., Quasi-hexagonal molecular packing in collagen fibrils. *Nature* **1979**, *282* (5741), 878-880.
30. Piez, K. A.; Trus, B. J., A new model for packing of type I molecules in the native fibril. *Biosci. Rep.* **1981**, *1*, 801-810.
31. Chang, S.-W.; Shefelbine, Sandra J.; Buehler, Markus J., Structural and Mechanical Differences between Collagen Homo- and Heterotrimers: Relevance for the Molecular Origin of Brittle Bone Disease. *Biophysical Journal* **2012**, *102* (3), 640-648.
32. (a) Weil, D.; Mattei, M. G.; Passage, E.; N'Guyen, V. C.; Pribula-Conway, D.; Mann, K.; Deutzmann, R.; Timpl, R.; Chu, M. L., Cloning and chromosomal localization of human genes encoding the three chains of type VI collagen. *American journal of human genetics* **1988**, *42* (3), 435-45; (b) Chu, M.-L.; Mann, K.; Deutzmann, R.; Pribula-Conway, D.; Hsu-Chen, C.-C.; Bernard, M. P.; Timpl, R., Characterization of three constituent chains of collagen type VI by peptide sequences and cDNA clones. *European Journal of Biochemistry* **1987**, *168* (2), 309-317.
33. (a) Aigner, T.; Hambach, L.; Soder, S.; Schlotzer-Schrehardt, U.; Poschl, E., The C5 domain of Col6A3 is cleaved off from the Col6 fibrils immediately after secretion. *Biochem Bioph Res Co* **2002**, *290* (2),

- 743-748; (b) Rupert Timpl, M.-L. C., Microfibrillar Collagen Type VI. In *Extracellular Matrix Assembly and Structure*, Peter D. Yurchenco, R. P. M. a. D. E. B., Ed. Academic Press: Orlando, 1994; pp 207–242.
34. T.M. Schmid, T. F. L., Type X collagen. In *Structure and Function of Collagen Types*, R. Mayne, R. E. B. E., Ed. Academic Press: Orlando, 1987; pp 195-222.
35. Yamaguchi, N.; Mayne, R.; Ninomiya, Y., The Alpha-1(Viii) Collagen Gene Is Homologous to the Alpha-1(X) Collagen Gene and Contains a Large Exon Encoding the Entire Triple Helical and Carboxyl-  
10 Terminal Non-Triple Helical Domains of the Alpha-1(Viii) Polypeptide. *J Biol Chem* **1991**, *266* (7), 4508-4513.
36. Ricard-Blum, B. D. a. M. v. d., *Unconventional Collagens Types VI, VII, VIII, IX, X, XIV, XVI and XIX*. Oxford University Press 2000.
- 15 37. (a) Schuppan, D.; Schmid, M.; Somasundaram, R.; Ackermann, R.; Ruehl, M.; Nakamura, T.; Riecken, E. O., Collagens in the liver extracellular matrix bind hepatocyte growth factor. *Gastroenterology* **1998**, *114* (1), 139-152; (b) Frenkel, S. R.; Toolan, B.; Menche, D.; Pitman, M. I.; Pachence, J. M., Chondrocyte transplantation  
20 using a collagen bilayer matrix for cartilage repair. *J Bone Joint Surg Br* **1997**, *79B* (5), 831-836; (c) Wakitani S, K. T., Hirooka A, Ochi T, Yoneda M, Yasui N, Owaki H, Ono K., Repair of rabbit articular surfaces with allograft chondrocytes embedded in collagen gel. *J Bone Joint Surg Br* **1989**, *71* (1), 74-80.
- 25 38. (a) Rizzo, R. J.; McCarthy, W. J.; Dixit, S. N.; Lilly, M. P.; Shively, V. P.; Flinn, W. R.; Yao, J. S., Collagen types and matrix protein content in human abdominal aortic aneurysms. *Journal of vascular surgery* **1989**, *10* (4), 365-373; (b) Dobrin, P.; Mrkvicka, R., Failure of elastin or collagen as possible critical connective tissue alterations  
30 underlying aneurysmal dilatation. *Vascular* **1994**, *2* (4), 484-488.
39. Samouillan, V.; Lamure, A.; Maurel, E.; Dandurand, J.; Lacabanne, C.; Ballarin, F.; Spina, M., Characterisation of elastin and collagen in aortic bioprostheses. *Medical and Biological Engineering and Computing* **2000**, *38* (2), 226-231.
- 35 40. Abou Neel, E. A.; Bozec, L.; Knowles, J. C.; Syed, O.; Mudera, V.; Day, R.; Hyun, J. K., Collagen — Emerging collagen based therapies hit the patient. *Advanced Drug Delivery Reviews* **2013**, *65* (4), 429-456.
41. Walters, B. D.; Stegemann, J. P., Strategies for directing the  
40 structure and function of three-dimensional collagen biomaterials across length scales. *Acta Biomaterialia* **2014**, *10* (4), 1488-1501.
42. Li, Y.; Yu, S. M., Targeting and mimicking collagens via triple helical peptide assembly. *Curr Opin Chem Biol* **2013**, *17* (6), 968-975.
- 45 43. Yonath, A.; Traub, W., Polymers of tripeptides as collagen models. *Journal of Molecular Biology* **1969**, *43* (3), 461-477.
44. (a) Miller, M. H.; Nemethy, G.; Scheraga, H. A., CALCULATION OF THE STRUCTURES OF COLLAGEN MODELS - ROLE OF INTERCHAIN INTERACTIONS IN DETERMINING THE  
50 TRIPLE-HELICAL COILED-COIL CONFORMATION .2. POLY(GLYCYL-PROLYL-HYDROXYPROLYL). *Macromolecules* **1980**, *13* (3), 470-478; (b) Miller, M. H.; Nemethy, G.; Scheraga, H. A., CALCULATION OF THE STRUCTURES OF COLLAGEN MODELS - ROLE OF INTERCHAIN INTERACTIONS IN DETERMINING THE  
55 TRIPLE-HELICAL COILED-COIL CONFORMATION .3. POLY(GLYCYLPROLYLALANYL). *Macromolecules* **1980**, *13* (4), 910-913; (c) Miller, M. H.; Scheraga, H. A., CALCULATION OF STRUCTURES OF COLLAGEN MODELS - ROLE OF INTERCHAIN INTERACTIONS IN DETERMINING TRIPLE-HELICAL COILED-  
60 COIL CONFORMATION .1. POLY(GLYCYL-PROLYL-PROLYL). *Journal of Polymer Science Part C-Polymer Symposium* **1976**, (54), 171-200; (d) Nemethy, G.; Miller, M. H.; Scheraga, H. A., CALCULATION OF THE STRUCTURES OF COLLAGEN MODELS - ROLE OF INTERCHAIN INTERACTIONS IN DETERMINING THE TRIPLE-  
65 HELICAL COILED-COIL CONFORMATION .4. POLY(GLYCYLALANYLPROLYL). *Macromolecules* **1980**, *13* (4), 914-919.
45. (a) Brown, F. R.; Blout, E. R.; Hopfinge, A., COLLAGEN-LIKE TRIPLE HELIX TO RANDOM-CHAIN TRANSITION -  
70 EXPERIMENT AND THEORY. *Journal of Molecular Biology* **1972**, *63* (1), 101-&; (b) Doyle, B. B.; Traub, W.; Lorenzi, G. P.; Blout, E. R., CONFORMATIONAL INVESTIGATIONS ON POLYPEPTIDE AND OLIGOPEPTIDES WITH REPEATING SEQUENCE L-ALANYL-L-  
145 PROLYLGLYCINE. *Biochemistry-U.S.* **1971**, *10* (16), 3052-&; (c) Doyle, B. B.; Bendit, E. G.; Blout, E. R., INFRARED SPECTROSCOPY OF COLLAGEN AND COLLAGEN-LIKE POLYPEPTIDES. *Biopolymers* **1975**, *14* (5), 937-957; (d) Doyle, B. B.; Traub, W.; Lorenzi, G. P.; Brown, F. R.; Blout, E. R., SYNTHESIS AND STRUCTURAL INVESTIGATION OF POLY(L-ALANYL-L-ALANYL-GLYCINE).  
80 *Journal of Molecular Biology* **1970**, *51* (1), 47-&; (e) Brown, F. R.; Blout, E. R.; Corato, A. D.; Lorenzi, G. P., SYNTHESIS AND STRUCTURAL STUDIES OF 2 COLLAGEN ANALOGS - POLY (L-PROLYL-L-SERYL-GLYCYL) AND POLY (L-PROLYL-L-ALANYL-GLYCYL). *Journal of Molecular Biology* **1972**, *63* (1), 85-&.
- 85 46. Simon-Lukasik, K. V.; Persikov, A. V.; Brodsky, B.; Ramshaw, J. A. M.; Laws, W. R.; Ross, J. B. A.; Ludescher, R. D., Fluorescence determination of tryptophan side-chain accessibility and dynamics in triple-helical collagen-like peptides. *Biophys J* **2003**, *84* (1), 501-508.
- 90 47. Shah, N. K.; Sharma, M.; Kirkpatrick, A.; Ramshaw, J. A. M.; Brodsky, B., Gly-Gly-containing triplets of low stability adjacent to a type III collagen epitope. *Biochemistry-U.S.* **1997**, *36* (19), 5878-5883.
48. Sage, J.; Mallevre, F.; Barbarin-Costes, F.; Samsonov, S. A.; Gehrcke, J. P.; Pisabarro, M. T.; Perrier, E.; Schnebert, S.; Roget, A.;  
95 Livache, T.; Nizard, C.; Lalmanach, G.; Lecaille, F., Binding of Chondroitin 4-Sulfate to Cathepsin S Regulates Its Enzymatic Activity. *Biochemistry-U.S.* **2013**, *52* (37), 6487-6498.
49. Tam, E. M.; Wu, Y. I.; Butler, G. S.; Stack, M. S.; Overall, C. M., Collagen binding properties of the membrane type-1 matrix  
100 metalloproteinase (MT1-MMP) hemopexin C domain. *J Biol Chem* **2002**, *277* (41), 39005-14.
50. Streeter, I.; de Leeuw, N. H., Atomistic Modeling of Collagen Proteins in Their Fibrillar Environment. *J Phys Chem B* **2010**, *114* (41), 13263-13270.
- 105 51. Torres, P. H. M.; Sousa, G. L. S. C.; Pascutti, P. G., Structural analysis of the N-terminal fragment of the antiangiogenic protein endostatin: A molecular dynamics study. *Proteins* **2011**, *79* (9), 2684-2692.
52. (a) Andersson, I. E.; Andersson, C. D.; Batsalova, T.;  
110 Dzhambazov, B.; Holmdahl, R.; Kihlberg, J.; Linusson, A., Design of Glycopeptides Used to Investigate Class II MHC Binding and T-Cell Responses Associated with Autoimmune Arthritis. *Plos One* **2011**, *6* (3); (b) Prior, S. H.; Byrne, T. S.; Tokmina-Roszyk, D.; Fields, G. B.; Van Doren, S. R., Path to Collagenolysis: COLLAGEN V TRIPLE-HELIX  
115 MODEL BOUND PRODUCTIVELY AND IN ENCOUNTERS BY MATRIX METALLOPROTEINASE-12. *J Biol Chem* **2016**, *291* (15), 7888-7901.
53. (a) Cunningham, L. W.; Davies, H. A.; Hammonds, R. G., ANALYSIS OF ASSOCIATION OF COLLAGEN BASED ON  
120 STRUCTURAL MODELS. *Biopolymers* **1976**, *15* (3), 483-502; (b) Jones, E. Y.; Miller, A., ANALYSIS OF STRUCTURAL DESIGN-FEATURES IN COLLAGEN. *Journal of Molecular Biology* **1991**, *218* (1), 209-219; (c) Fraser, R. D. B.; Macrae, T. P.; Miller, A.; Suzuki, E., MOLECULAR-CONFORMATION AND PACKING IN COLLAGEN  
125 FIBRILS. *Journal of Molecular Biology* **1983**, *167* (2), 497-521; (d) Lee, E.; Nemethy, G.; Scheraga, H. A.; Ananthanarayanan, V. S., BETA-BEND CONFORMATION OF CH3CO-PRO-PRO-GLY-PRO-NHCH3 - IMPLICATIONS FOR POSTTRANSLATIONAL PROLINE  
130 HYDROXYLATION IN COLLAGEN. *Biopolymers* **1984**, *23* (7), 1193-1206; (e) Nemethy, G.; Scheraga, H. A., STABILIZATION OF COLLAGEN FIBRILS BY HYDROXYPROLINE. *Biochemistry-U.S.* **1986**, *25* (11), 3184-3188.
54. (a) Brodsky, B.; Li, M. H.; Long, C. G.; Apigo, J.; Baum, J., NMR AND CD STUDIES OF TRIPLE-HELICAL PEPTIDES.  
135 *Biopolymers* **1992**, *32* (4), 447-451; (b) Long, C. G.; Braswell, E.; Zhu, D.; Apigo, J.; Baum, J.; Brodsky, B., CHARACTERIZATION OF COLLAGEN-LIKE PEPTIDES CONTAINING INTERRUPTIONS IN THE REPEATING GLY-X-Y SEQUENCE. *Biochemistry-U.S.* **1993**, *32* (43), 11688-11695; (c) Venugopal, M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu, D.; Brodsky, B., ELECTROSTATIC INTERACTIONS IN  
140 COLLAGEN-LIKE TRIPLE-HELICAL PEPTIDES. *Biochemistry-U.S.* **1994**, *33* (25), 7948-7956.
55. Klein, T. E.; Huang, C. C., Computational investigations of structural changes resulting from point mutations in a collagen-like  
145 peptide. *Biopolymers* **1999**, *49* (2), 167-183.



56. Mooney, S. D.; Kollman, P. A.; Klein, T. E., Conformational preferences of substituted prolines in the collagen triple helix. *Biopolymers* **2002**, *64* (2), 63-71.
57. Kwansa, A. L.; De Vita, R.; Freeman, J. W., Mechanical recruitment of N- and C-crosslinks in collagen type I. *Matrix Biol* **2014**, *34*, 161-169.
58. Malone, J. P.; George, A.; Veis, A., Type I collagen N-telopeptides adopt an ordered structure when docked to their helix receptor during fibrillogenesis. *Proteins-Structure Function and Genetics* **2004**, *54* (2), 206-215.
59. Malone, J. P.; Veis, A., Heterotrimeric type I collagen C-telopeptide conformation as docked to its helix receptor. *Biochemistry-U.S.* **2004**, *43* (49), 15358-15366.
60. Lee, K. H.; Holl, M. M. B., Free Energy Simulation to Investigate the Effect of Amino Acid Sequence Environment on the Severity of Osteogenesis Imperfecta by Glycine Mutations in Collagen. *Biopolymers* **2011**, *95* (6), 401-409.
61. Radmer, R. J.; Klein, T. E., Severity of Osteogenesis Imperfecta and Structure of a Collagen-like Peptide Modeling a Lethal Mutation Site. *Biochemistry-U.S.* **2004**, *43* (18), 5314-5323.
62. Park, S.; Radmer, R. J.; Klein, T. E.; Pande, V. S., A new set of molecular mechanics parameters for hydroxyproline and its use in molecular dynamics simulations of collagen-like peptides. *Journal of Computational Chemistry* **2005**, *26* (15), 1612-1616.
63. Gurry, T.; Nerenberg, P. S.; Stultz, C. M., The Contribution of Interchain Salt Bridges to Triple-Helical Stability in Collagen. *Biophys J* **2010**, *98* (11), 2634-2643.
64. Keshwani, N.; Banerjee, S.; Brodsky, B.; Makhatazde, G. I., The Role of Cross-Chain Ionic Interactions for the Stability of Collagen Model Peptides. *Biophys J* **2013**, *105* (7), 1681-1688.
65. Rele, S.; Song, Y.; Apkarian, R. P.; Qu, Z.; Conticello, V. P.; Chaikof, E. L., D-periodic collagen-mimetic microfibers. *Journal of the American Chemical Society* **2007**, *129* (47), 14780-7.
66. Raman, S. S.; Gopalakrishnan, R.; Wade, R. C.; Subramanian, V., Structural Basis for the Varying Propensities of Different Amino Acids To Adopt the Collagen Conformation. *J Phys Chem B* **2011**, *115* (11), 2593-2607.
67. Katti, D. R.; Pradhan, S. M.; Katti, K. S., Directional dependence of hydroxyapatite-collagen interactions on mechanics of collagen. *J Biomech* **2010**, *43* (9), 1723-1730.
68. Chen, M.; Zheng, T.; Wu, C.; Xing, C., Molecular dynamics simulations of collagen adsorption onto grooved rutile surface: The effects of groove width. *Colloids and Surfaces B: Biointerfaces* **2014**, *121*, 150-157.
69. Almora-Barrios, N.; De Leeuw, N. H., Molecular Dynamics Simulation of the Early Stages of Nucleation of Hydroxyapatite at a Collagen Template. *Cryst Growth Des* **2012**, *12* (2), 756-763.
70. (a) Kang, Y.; Wang, Q.; Liu, Y. C.; Wu, T.; Chen, Q.; Guan, W. J., Dynamic mechanism of collagen-like peptide encapsulated into carbon nanotubes. *J Phys Chem B* **2008**, *112* (15), 4801-4807; (b) Ravimohan Gopalakrishnan, K. B., Ettayapuram Ramaprasad Azhagiya Singam, Subramanian Sundaramana and Venkatesan Subramanian Adsorption of collagen onto single walled carbon nanotubes: a molecular dynamics investigation. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13046-13057.
71. (a) Gopalakrishnan, R.; Azhagiya Singam, E. R.; Vijaya Sundar, J.; Subramanian, V., Interaction of collagen like peptides with gold nanosurfaces: a molecular dynamics investigation. *Phys. Chem. Chem. Phys.* **2015**, *17* (7), 5172-5186; (b) Gopalakrishnan, R.; Azhagiya Singam, E. R.; Vijaya Sundar, J.; Subramanian, V., Interaction of collagen like peptides with gold nanosurfaces: a molecular dynamics investigation. *Physical Chemistry Chemical Physics* **2015**, *17* (7), 5172-5186.
72. Ebrahimi, S.; Montazeri, A., & Rafii-Tabar, H, Molecular dynamics study of the interfacial mechanical properties of the graphene-collagen biological nanocomposite. *Computational Materials Science* **2013**, *69*, 29-39.
73. Narayanan, B.; Gilmer, G. H.; Tao, J.; De Yoreo, J. J.; Ciobanu, C. V., Self-Assembly of Collagen on Flat Surfaces: The Interplay of Collagen-Collagen and Collagen-Substrate Interactions. *Langmuir* **2014**, *30* (5), 1343-1350.
74. Ricard-Blum, S.; Feraud, O.; Lortat-Jacob, H.; Rencurosi, A.; Fukai, N.; Dkhissi, F.; Vittet, D.; Imberty, A.; Olsen, B. R.; van der Rest, M., Characterization of endostatin binding to heparin and heparan sulfate by surface plasmon resonance and molecular modeling: role of divalent cations. *J Biol Chem* **2004**, *279* (4), 2927-36.
75. Boiers, U.; Lanig, H.; Schnert, B.; Holmdahl, R.; Burkhardt, H., Collagen type II is recognized by a pathogenic antibody through germline encoded structures. *Eur J Immunol* **2008**, *38* (10), 2784-2795.
76. Hans-Christian Siebert, M. B.-R., Thomas Eckert, Sabine Stötzl, Ulrike Kirch, Tammo Diercks, Martin J. Humphries, Martin Frank, Rainer Wechselberger, Emad Tajkhorshid, Steffen Oesser, Interaction of the  $\alpha 2A$  domain of integrin with small collagen fragments. *Protein & Cell* **2010**, *1* (4), 393-405.
77. Zhang, L.; Zhang, C.; Sun, Y., Biomimetic Design of Platelet Adhesion Inhibitors to Block Integrin  $\alpha 2\beta 1$ -Collagen Interactions: II. Inhibitor Library, Screening, and Experimental Validation. *Langmuir* **2014**, *30* (16), 4734-4742.
78. Passacquale, G.; Di Giosia, P.; Ferro, A., The role of inflammatory biomarkers in developing targeted cardiovascular therapies: lessons from the cardiovascular inflammation reduction trials. *Cardiovascular Research* **2016**, *109* (1), 9-23.