The Collagen Microfibril Model as a Tool for Leather Scientists

Eleanor M. Brown*

Fats, Oils and Animal Coproducts Research Unit, U.S. Department of Agriculture, Agricultural Research Service, Eastern Regional Research Center, 600 East Mermaid Lane, Wyndmoor, PA 19038

Abstract: Collagen, a structural protein of the extracellular matrix, gives strength and form to the skin, tendons, bones, cornea and teeth of mammals. The discovery by early humans that the skin of an animal, slaughtered for meat, could be preserved by exposing it to smoke or rubbing with fat, led to the production of leather. Through empirical methods, a number of tanning agents with a variety of properties were identified. The methods for production of leather evolved over several centuries as art and engineering with little understanding of the underlying science. Scientific advances of the twentieth century, including increasing use of collagen in medical device research, began to provide a basis for understanding the relationship between collagen structure and function in both biology and technology. The unique structure of fibrous collagens, a Gly-X-Y repeating sequence that results in a long triple helix that further associates into fibers, makes solution based studies of protein interactions challenging. Nearly twenty years ago, leather researchers at the Eastern Regional Research Center of the United States Department of Agriculture began the construction of a type I collagen microfibril model for the simulation of interactions of tanning chemistries with and between collagen triple helices. Insights gained, and possible new directions for research and development will be discussed.
Key words: collagen micofibril; molecular simulation; tanning mechanisms

1 Introduction

The shape and mechanical attributes of a vertebrate body are defined by its connective tissue, the cells of which are embedded in an extracellular matrix that is a complex mixture of proteins and carbohydrates and functions as a support for cellular materials. Collagen, a fibrillar protein of the extracellular matrix, is the most abundant protein in mammals, and has throughout history played important roles in physiology and technology. Collagen gives strength and form to the skin of mammals and serves as the substrate for the production of leather. Collagen also serves as the basis of biomaterials for other industries, including medical materials, food and adhesives. Protein structure is a determining factor in protein function, and an understanding of collagen structure is necessary for the development of collagen-based processes and products. The insolubility and noncrystalline nature of collagen presented a challenge to scientists who attempted to elucidate its structure either for physiological or technological purposes. In the 1940's and 50's, X-ray diffraction patterns and electron micrographs of collagen showing a highly organized supramolecular structure led observers to conclude that the structure was helical, but not α-helical.[1] Tropocollagen, the triple helical collagen monomer, stabilized by peptide and hydrogen bonds as well as steric interactions was isolated. The unique features of the amino acid composition, which contains 33% glycine, high levels of imino acids, and the post translationally modified hydroxyproline (Hpr) encouraged early model builders to visualize a $(Gly\text{-}Pro\text{-}Hpr)_n$ helix in ball and stick representation (Fig. 1).[2]

In the 1960's the microfibril, a structure consisting of four to six tropocollagen molecules staggered lengthwise to produce the alternating gap and overlap regions observed in electron micrographs, was proposed as the limiting unit for three-dimensional arrangement in collagen fibers.[3] By the 1970's, Heidemann[4] and others began synthesizing collagen-like peptides for solution and crystallization studies. Hulmes et al.[5] constructed two-dimensional models by aligning the amino acid sequence of collagen with electron micrographs to explain the stagger of the chains required to produce the D-space pattern as a reflection of the distribution of large hydrophobic residues. In vitro studies of collagen fiber assembly implicated the microfibril as the second stage in the hierarchical path to a super helical coiled-coil.[6] At each stage, the handedness of the super helix is reversed, leading to interdigitation that strengthens the

* Corresponding author: Phone: 215-233-6481. Email: ellie.brown@ars.usda.gov
§Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

whole and makes isolation of individual substructures larger than tropocollagen impractical.[7] By the 1980's, computer based molecular simulations had become feasible, and the Scheraga group at Cornell University explored the effects of various Gly-X-Y residues on the stability of the collagen triple helix.[8]

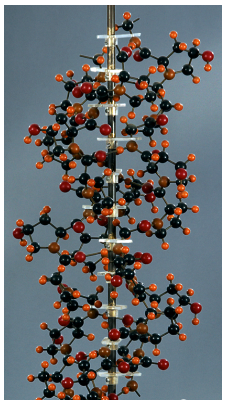

Fig. 1 This early ball and stick model of the collagen triple helix is located in the library at ERRC, ARS, USDA

2 Model building

The ERRC bovine collagen microfibril model is based on the experimental and theoretical data available for collagen structure in 1989, with modifications to accommodate more recent data. James Chen used computer algorithms in protein modeling software to construct a left-handed helix with the $(Gly\text{-}Pro\text{-}Hpr)_{12}$ sequence, and assemble three identical chains into an energy-minimized right-handed coiled-coil to produce a triple helical template. Five of these triple helices were then coiled into a left-handed Smith microfibril.[3,9] A sequence alignment template[3] was used to identify the amino acid side chains that would represent a 36 residue long slice of the microfibril. These side chains were mutated, computationally, into the positions originally occupied by Pro and Hpr side chains (Fig. 2).[10,11]

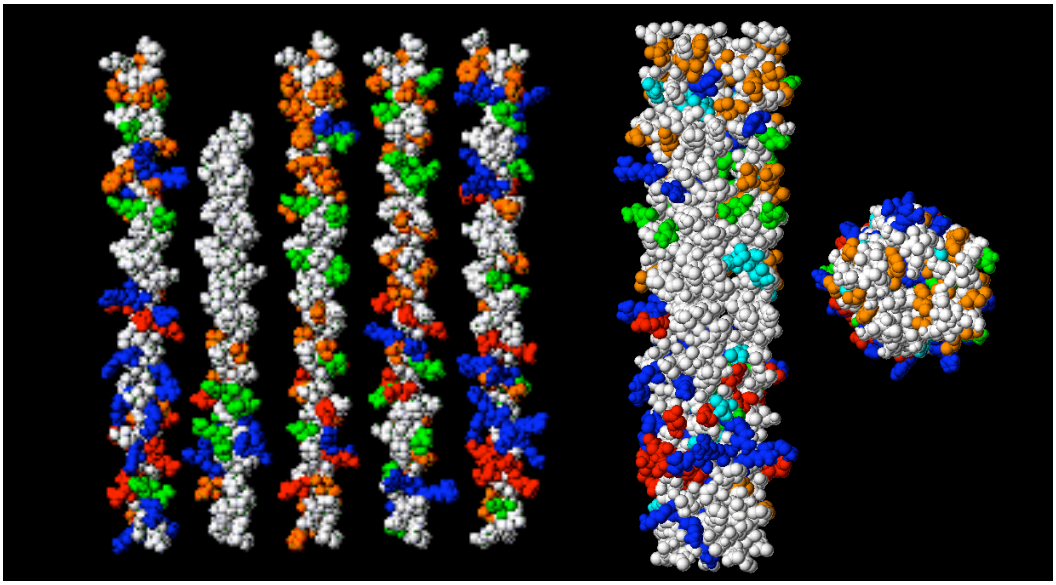

Fig. 2 Space-filling representations of a 36 residue long segment of a Type I collagen microfibril, shown from left to right as the individual energy minimized triple helical structures, where the carboxyl terminus is the top-end of the second structure, and the superhelical microfibril in both lateral and end on views. The color code is red for acidic, blue for basic, orange for hydrophobic, green for neutral polar residues and white for the characteristic collagen residues, (Gly, Pro, Hpr)

When a similar approach was attempted for the construction of the complete, 315-residue long, ERRC microfibril model, the result was many bad atomic contacts and general distortion of the backbone. Gregory King then developed a procedure for growing new side chains in place, whereby side chains were enlarged or diminished, one bond at a time. The entire molecule was then energy minimized between successive additions or subtractions. The resulting model, when colored to represent a negative microscopy stain, showed a banding pattern typical of those seen in collagen micrographs (Fig. 3). This 15 chain model encompasses a single gap region, 159 residues long with a 78-residue overlap region at either end. The model contains the entire helical portion of the tropocollagen sequence.[12]

Although the role of telopeptides in fibril stabilization and their importance as anchors for the natural crosslinks that form in collagen as the animal ages was recognized, the initial model ignored these nonhelical regions because of the limited data then available on their secondary structures. In 1999, based on spectrophotometric data for isolated telopeptides,[13,14] Phoebe Qi began the construction of models of the N-terminal[15] and C-terminal[16] telopeptides. These were fit into the gap region of the larger model and attached to the appropriate triple helicies. The telopeptide conformations have since been modified (Fig. 3b,c) in light of more recent studies of their interactions with the triple helical domain of collagen.[17,18]

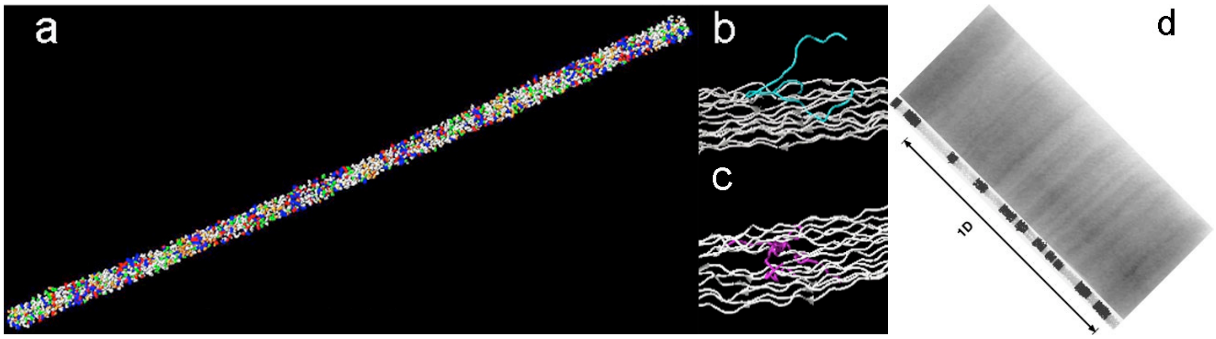

Fig. 3(a) ERRC collagen microfibril model in space-filling representation, colored as described for Figure 2, enlarged views of the (b) N-terminal and (c) C-terminal gap-overlap interfaces in ribbon representation with telopeptides in color, and (d) the microfibril shaded as with a microscopy stain, and aligned with a collagen micrograph

3 Utilization

The impetus for construction of this model was to provide a basis for the study of those interactions between tanning materials and collagen that might contribute to the stabilization of the fibril structure. The findings by Scholnick et al.[19] demonstrated that although dicarboxylic acids with varying chain lengths bound to collagen, those with C-7 to C-12 chains had a much greater effect on collagen thermal stability than did either shorter or longer chains, was explained by simulating the dicaboxylic acid molecules under experimental conditions to estimate the potential crosslinking span of different chains. Short chain dicarboxylic acids could crosslink within a triple helix, medium length chains could form inter helical crosslinks, and longer chain molecules could bind to the surface, but were unable to fit within the microfibril.

Although tanning may better be described in terms of protein modification than as simple crosslinking,[20-22] crosslinking or binding to collagen is generally accepted as one role of a tanning agent. Attempts at identifying binding sites have generally focused on the triple helical domains of collagen because these were the regions best described and understood. The primary chrome-tanning agent, "basic chrome sulfate," had been characterized as a mixture of Cr(III) complexes, where two or more chromium atoms connected by oxygen and/or hydroxyl bridges have varying numbers of associated sulfate ions.[23] However, our expertise, and the modeling software available, are based on protein structures, and could not easily accommodate complex inorganic structures. As a result, our rather primitive attempts at modeling chrome tanning were restricted to identifying potential binding sites with suitably spaced acidic residues (Fig. 4a).[24]

A greater awareness of the roles of collagen modification and fibril coating in tanning has recently increased interest in tanning agents other than chromium, and a better understanding of the role of telopeptides in fibril stabilization has led to a closer examination of the gap region. A successful tanning agent is one that interacts with the collagen matrix of the hide in a way that provides stability. Under the conditions of tanning, most tanning agents are in oligomeric form, and effective interactions with collagen are intermolecular. Thus, one of the requirements for tanning may be adequate open space within the fiber structure to accommodate a moderately sized oligomer without major distortion of the collagen ultrastructure. Glutaraldehyde tanning, the basis for many of the chrome-free tannages in use today was developed at ERRC half a century ago to produce washable sheepskin bed pads for hospitals. Glutaraldehyde molecules self associate to form oligomers that are large enough to form interhelical crosslinks with basic side chains on collagen (Fig. 4b). We have demonstrated tanning properties on powdered hide for genipin, an iridoid compound isolated from the fruit of *Gardenia jasmindides* Ellis, alone and in combination with aluminum.[25] Genipin dimerizes and acts essentially as a dialdehyde, much like glutaraldehyde, but with a better defined structure (Fig. 4c).

We chose catechin, a polyphenolic molecule often considered to be a model vegetable tannin to evaluate potential interactions of collagen with vegetable tannins.[26] This model was used to estimate the accessibility of potential interaction sites in the microfibril structure, and to suggest the most likely types of interactions between tannin and collagen after simulated tanning. As would be expected, the atomic density around potential binding sites was lower in the gap region than in the overlap region. Intra- and interhelical hydrogen bonding and hydrophobic interactions were observed with this model (Fig. 4d).

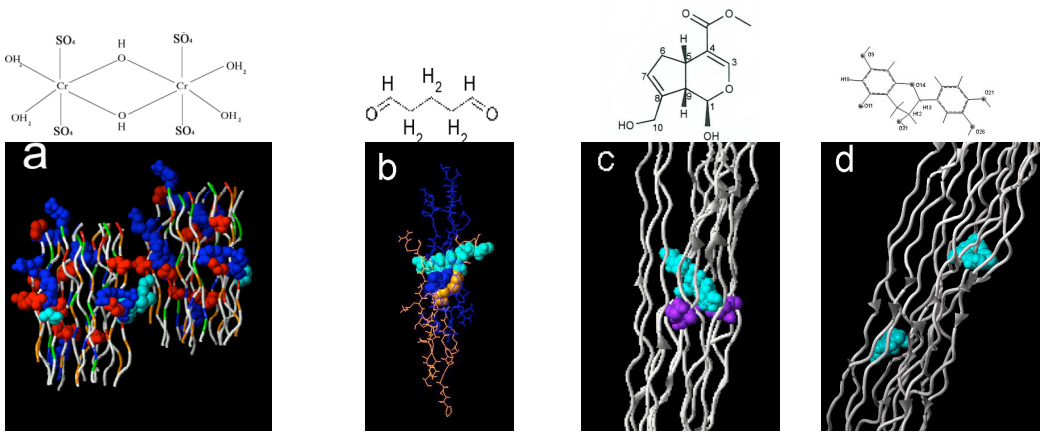

Fig. 4 Models of interaction sites for potential tanning agents (cyan) with collagen (a) binuclear Cr(III) complex forming a bridge between acid residues on different microfibrils (b) glutaraldehyde dimer crosslinking lysine residues on different triple helices (c) digenipin bridging lysine residues in the gap region of the microfibril and (d) catchin molecules in hydrogen bonding positions within the gap region of the microfibril. Structures of the tanning agent molecules are shown above those of the corresponding interaction sites

Computer-modeling studies are quite valuable in predicting or explaining experimental observations, and they may suggest which experiment among several is most likely to confirm or refute a hypothesis. Nevertheless, the experimental test of the hypothesis is essential. When working with a large insoluble substrate such as cattle hide, it is difficult to isolate the effects of individual processing steps on the molecular or supramolecular structure of collagen. To complement our computer modeling studies, we have developed laboratory scale model-tanning systems that employ pepsin-solubilized collagen, insoluble ball-milled collagen or powdered hide. Using conditions that approximated steps in a chrome tanning process, measurable effects on the conformation and conformational stability of soluble collagen under the conditions of chrome tanning could be seen. With a powdered hide model, the necessary conditions for tanning with genipin and natural tannins were explored.[25]

Covalent crosslinks, electrostatic interactions (salt bridges), hydrophobic interactions, hydrogen bonds, and water activity are among the contributing factors in a stabilization mechanism. Molecular modeling is an excellent approach to the study of the effects of these factors, individually and

collectively, on collagen structure. Limited resources have made it impractical for us to perform the kinds of studies with the microfibril model that would be most useful to an understanding of tanning mechanisms. Other researchers with different resources have accepted the challenge of tanning-relevant modeling. Madhan et al.[27,28] used a small collagen-like polypeptide and several vegetable tannin models to estimate binding energies as a function of tannin structure and peptide configuration. Triple helical models were used by Raman et al.[29] to show that n = 5 in the sequence $(Gly\text{-}Pro\text{-}Hpr)_n$ was a minimum requirement for a stable helix under molecular dynamics simulation, and later to evaluate the effect on helix stability of substituting acidic or basic residues for the Pro and Hpr.[30] A simulated tanning study by Fennen[31] looked at the bridging interactions of a binuclear chromium complex with two triple helices.

The studies most relevant to tanning have used our microfibril model as their starting point. Since 2004, a modeling group at the University of Pisa, Italy, has used the short segments of the ERRC microfibril model as a basis for studies of the structural and binding properties of supramolecular collagen. They discovered that large changes in the conformation of a 23-residue long microfibril segment resulted when molecular dynamics were simulated in pure water, and that these changes were minimized by the addition of even small numbers of formaldehyde or gallic acid molecules.[32-34] This group has developed more realistic tanning agent models, a flavonoid based vegetable tannin model[35] that is more complex than the gallic acid and catechin models typically used, and a chromium-flavonoid complex.[36] In separate studies, Siggel and Molnar[37] modeled the density and swelling behavior of a microfibril as a function of pH and the conditions of pickling with 12% NaCl. Bulo et al.[38] evaluated the effects of calcium cations and sulfate anions on the swelling of microfibrils at low pH, and also the interactions[39] of polymeric retanning agents with the collagen model.

4 Conclusions

The substrate for leather is the animal hide, although collagen is by far the major component, other hide substances including minor proteins and carbohydrates are present and may participate in the tanning process. The ERRC modeling group developed a collagen model suitable for studies of tanning mechanisms that has also served as a catalyst for the initiation of those studies by others. The feedback from other researchers has enabled us to validate and strengthen the model. We continue to use this model to explore collagen/tanning agent interactions, particularly in the less dense gap region of the microfibril. The requirement for experimental models as an adjunct to computer models is recognized, as is the necessity to be aware of and utilize appropriate models developed by scientists with different goals. The literature describing collagen structure/function relationships and modeling from a medical perspective is huge, and expanding. The use of collagen as a biomaterial is a rapidly growing field and one expected to develop knowledge valuable for the tanner. A start has been made towards a molecular level understanding of the interactions of the supermolecular structure of collagen with individual components of a chrome or vegetable tanning process. Much more will be needed to develop a comprehensive tanning mechanism. This is an auspicious moment: computing power has become relatively inexpensive, software more sophisticated, scientific literature available electronically, many modelers have been trained, and models that include solvents or processing chemicals have been published. The challenge is to expand the scope of leather modeling to take full advantage of microfibril models, to include the interactions with other hide substances or added chemicals, and to integrate relevant data from researchers in the medical and biomaterials science arenas in the development of a comprehensive mechanism of tanning.

Acknowledgements

The author appreciates the contributions of ERRC molecular modelers, fellow research scientists, leather professionals, and visiting scientists for guidance on tanning processes, the ERRC leaders, and the Research Liaison Committee of the American Leather Chemists Association who supported this work.

References

[1] G. N. Ramachandran; G. Kartha. Nature, 1954, 174(4423):269-270.
[2] G. N. Ramachandran. Structure of collagen at the molecular level. Treatise on Collagen, (Ramachandran, G. N., ed.) Chap. 3, Academic Press, 1967.
[3] J. W. Smith. Nature, 1968, 219:157-158.
[4] E. R. Heidemann; B. S. Harrap; H. D. Schiele. Biochemistry, 1973, 12:2958-2963.

[5] D. J. S. Hulmes; A. Miller; D. A. D. Parry. J. Mol. Biol, 1973, 79(1):137-148.
[6] G. C. Na; L. J. Butz; D. G. Bailey; R. J. Carroll. Biochemistry, 1986, 25:958-966.
[7] J. P. Orgel; T. C. Irving; A. Miller; T. J. Wess. Proc. Natl. Acad. Sci. USA, 2006, 103:9001-9005.
[8] M. Miller; G. Nemethy; H. A. Scheraga. Macromolecules, 1980, 13:910-913.
[9] J. M. Chen; C. E. Kung; S. H. Feairheller; E. M. Brown. J. Prot. Chem, 1991, 10:535-552.
[10] J. M. Chen; S. H. Feairheller; E. M. Brown. J. Am. Leather Chem. Assoc, 1991, 86:475-486.
[11] J. M. Chen; S. H. Feairheller; E. M. Brown. J. Am. Leather Chem. Assoc, 1991, 86:487-497.
[12] G. King; E. M. Brown; J. M. Chen. Protein Engineering, 1996, 9:43-49.
[13] A. George; J. P. Malone; A. Veis. Proc. Indian Acad. Sci.-Chem. Sci, 1999, 111:121-131.
[14] A. Otter; P. G. Scott; G. Kotovych. Biochemistry, 1988, 27:3560-3567.
[15] P. X. Qi; E, M. Brown. J. Am. Leather Chem. Assoc, 2002, 97:235-242.
[16] E. M. Brown. J. Am. Leather Chem. Assoc, 2004, 99:376-385.
[17] J. P. Malone; A. George; A. Veis. Proteins, 2004, 54:206-215.
[18] J. P. Malone; A. Veis. Biochemistry, 2004, 43:15358-15366.
[19] F. Scholnick; L. L. Liao; E. M. Brown; S. H. Feairheller. J. Am. Leather Chem. Assoc, 1992, 87:333-338.
[20] A. D. Covington; L. Song. Proceed. 27th Congress IULTCS, 2003, pp. 148-152.
[21] T. Ramasami. J. Am. Leather Chem. Assoc, 2001, 96: 290-304.
[22] E. M. Brown; M. M. Taylor. J. Am. Leather Chem. Assoc, 2003, 98:408-414.
[23] K. Takenouchi; K. Kondo; F. Nakamura. J. Soc. Leather Techn. Chem, 1981, 75:190-196.
[24] E. M. Brown; J. M. Chen; S.H. Feairheller. J. Am. Leather Chem. Assoc, 1993, 88:2-11.
[25] K. Ding; M. M. Taylor; E. M. Brown. J. Am. Leather Chem. Assoc. 2007, 102:164-170.
[26] E. M. Brown; P. X. Qi. J. Am. Leather Chem. Assoc, 2008, 103:290-297.
[27] B. Madhan; P. Thanikaivelan; V. Subramanian; J. Raghava Rao; B. Unni Nair; T. Ramasami. Chemical Physics Letters, 2001, 346:334-340.
[28] B. Madhan; V. Subramanian; J. Raghava Rao; B. Unni Nair; T. Ramasami. Int. J. Biol. Macromol, 2005, 37:47–53.
[29] S. S. Raman; R. Parthasarathi; V. Subramanian; T. Ramasami. J. Phys. Chem. B, 2008, 112:1533-1539.
[30] S. S. Raman; R. Parthasarathi; V. Subramanian; T. Ramasami. J. Mol. Struct.: THEOCHEM, 2008, 851:299-312.
[31] J. Fennen. J. Soc. Leather Tech. Chem, 1998, 82:5-10.
[32] S. Bronco; C. Cappelli; S. Monti. J. Phys. Chem. B, 2004, 108:10101-10112.
[33] S. Monti; S. Bronco; C. Cappelli. J. Phys. Chem. B, 2005, 109:11389-11398.
[34] S. Monti. J. Phys. Chem. C, 2007, 111:6086-6094.
[35] C. Cappelli; S. Bronco; S Monti. Chirality, 2005, 17:577-589.
[36] S. Bronco; C. Cappelli; S. Monti. J. Phys. Chem. B, 2006, 110:13227-13234.
[37] L. Siggel; F. Molnar. J. Am. Leather Chem. Assoc, 2006, 101:179-191.
[38] R. E. Bulo; L. Siggel; F. Molnar; H. Weiss. Macromol. Biosci, 2007, 7:234-240.
[39] L. Siggel; R. E. Bulo; F. Molnar; H. Weiss; T. Taeger. J. Am. Leather Chem. Assoc, 2007, 102:333-336.