

THREE-DIMENSIONAL-ENERGY MINIMIZED MODELS FOR CALF SKIN TYPE I COLLAGEN TRIPLE HELIX AND MICROFIBRIL: II. THE 'SMITH' MICROFIBRIL

by

JAMES M. CHEN, STEPHEN H. FEAIRHELLER AND ELEANOR M. BROWN

U.S. Department of Agriculture, ARS

Eastern Regional Research Center

600 EAST MERMAID LANE

PHILADELPHIA, PENNSYLVANIA 19118

Abstract

Molecular modeling methods were applied to the development of a three-dimensional structure for a Type I collagen microfibril model. The model was based on the energy minimized 'Smith' microfibril model containing the consensus polypeptide sequence (Gly-Pro-Hyp)₁₂. The amino acid sequences of the alpha 1 and 2 chains of calf skin Type I collagen were substituted into the 'collagen-like' (Gly-Pro-Hyp)₁₂ 'Smith' microfibril model. The model was then energy minimized using molecular mechanics and evaluated structurally using graphics. Specific interaction sites which may contribute to the stability of collagen packing can be identified in this model. Computer graphics was used to display the Type I 'Smith' microfibril model to determine specific domains containing sites which would facilitate the evaluation of potential crosslinking agents. The Type I model can be used to suggest domains in collagen that contain possible sites for crosslinking reactions with chromium and other agents; what constraints on size, shape and molecular characteristics are likely to describe the optimum crosslinking agents; and the application of which analytical and spectroscopic techniques to real systems will give the most rigorous test of the computer models.

Introduction

In the first paper of this series⁽¹⁾ we presented a description of the three-dimensional model for Type I collagen triple helices. In this study, a three-dimensional microfibril model for Type I collagen, based on the energy minimized 'Smith' microfibril model containing the consensus polypeptide sequence (Gly-Pro-Hyp)₁₂⁽²⁾, is presented. This initial (Gly-Pro-Hyp)₁₂ 'Smith' microfibril model was based on the model for the packing of collagen triple helices first proposed by Smith⁽³⁾. Specific regions in the linear amino acid sequences of the alpha 1 and 2 polypeptide chains were substituted into the three-dimensional (Gly-Pro-Hyp)₁₂ model. This Type I 'Smith' microfibril model was then energy minimized in order that all the possible stabilizing interactions are formed. The goals of this study are to construct a three-dimensional model for the 'Smith' microfibril structure of Type I collagen and to identify the structure-function relationships pertaining to how tanning ligands (i.e., Cr(III)⁽⁴⁾) modify collagen interactions in the process of leather manufacturing⁽¹⁾. Inter- and intra-polypeptide interactions between sidechains and/or backbone groups will be described. In the microfibril models, clusters or dense regions containing specific functional groups will be described. Identification of important sites for ligand

interactions will also be discussed. In addition, procedures for further developments in order to test and improve the Type I 'Smith' microfibril model will be discussed.

COLLAGEN PACKING MODELS

Present collagen packing models are based on electron microscopy and X-ray diffraction analysis⁽⁵⁻¹²⁾. The two dimensional alignment of adjacent collagen molecules was determined by electron microscopy of negative and positive stained transverse collagen tissue samples⁽¹⁰⁾. In the negative stained samples, tungstate or uranyl salts tend to deposit within the less tightly packed regions of the fibrils⁽¹⁰⁾. Electron microscopy of a negatively stained sample shows alternating dark and light bands, which are perpendicular to the length of the fibrils, formed along the fibrillar long axis. The accepted model explains this banding pattern as an array of adjacently aligned molecules staggered along their helical long axis by 1 D intervals, where 4.4D intervals constitute the entire length of a single collagen molecule^(1-3,10,12). Evidence for the relative length of a D interval is derived from data for positively stained collagen samples where heavy metal ions bind to charged amino acid sidechains. Positively stained samples also show distinct banding patterns that are thinner and more numerous, corresponding to the distribution of charged amino acids within collagen packing. Computer generated alignments of the collagen polypeptide sequence when compared to the banding pattern of a positive stained collagen sample gives a D interval length of 234 amino acid residues along a single polypeptide chain⁽¹⁰⁾. Electron microscopy studies indicate that the end-to-end alignment of collagen molecules is separated by a "gap" region corresponding to a length of 0.6D^(1-3,10,12). Smith⁽³⁾ incorporated these results into a three-dimensional packing model for collagen molecules. In this model, five collagen triple helices are packed in a circular array while each adjacent molecule is staggered laterally by 1 D spacing⁽³⁾. X-ray diffraction data were used to propose alternative models such as the hexagonal and 'quasi'-hexagonal packing model^(13,14). In both models, collagen molecules are hexagonally closed packed, but in the quasi hexagonal model, each respective molecule is tilted by approximately 5 degrees, with respect to the fibrillar long axis, to better account for x-ray diffraction patterns of collagen tendon samples⁽¹⁵⁾. Although these methods provide important structural information concerning collagen interactions, the results directly depend on the method of tissue preparation and on the tissue type. For example, both x-ray diffraction and electron microscopy techniques require the tissue sample to be modified (preparation techniques include stretching, clamping, fixation and modification of the solution which contains the dissected fiber samples) in order to obtain high resolution data for analysis⁽¹⁰⁾. Hence, the native structure of collagen interaction (packing) may be altered in these modified tissue samples.

CONSTRUCTION OF THE TYPE I MICROFIBRIL MODEL

Construction of the three-dimensional structure for the Type I microfibril was based on the energy minimized structure of the (Gly-Pro-Hyp)₁₂ 'Smith' microfibril⁽²⁾. Substitution of the region of the alpha 1 and 2 sequences shown in Fig. 1 was made into the (Gly-Pro-Hyp)₁₂ microfibril model⁽¹⁾. The specific microfibril region in Fig. 1 was chosen for study because it contains regions with many charged sidechains which may be involved in the

FIG. 1. — The microfibril region based on the packing model proposed by Smith⁽³⁾ which will be incorporated into the energy minimized microfibril model of (Gly-Pro-Hyp)₁₂. All references of microfibril regions made in the text will refer to the alignment based on the tripeptide numbers for COLLAGEN 1. Tripeptide 338 in COLLAGEN 2 pertains to the true carboxyl terminus of a single triple helical Type I molecule and this region contains part of the gap region (labeled 0). Each polypeptide chain has a length of 36 amino acid residues or 12 tripeptides. The triple helices are polarized in the same direction with respect to COLLAGEN 2.

SEQUENCES USED IN THE TYPE I 'SMITH' MICROFIBRIL MODEL

COLLAGEN 1			COLLAGEN 2			COLLAGEN 3			COLLAGEN 4			COLLAGEN 5							
A1	A2	A1																	
17	GKN	GKA	GKN	329	Gp	GPA	Gp	251	GPA	GFV	GPA	173	GPR	GSR	GPR	95	GAA	GPS	GAA
18	GDD	GED	GDD	330	GPR	GIR	GPR	252	GEK	GEp	GEK	174	GAN	GPS	GAN	96	GEE	GEE	GEE
19	GEA	GHp	GEA	331	GRT	GSQ	GRT	253	GAp	GPS	GAp	175	GAp	GPp	GAp	97	GKR	GKR	GKR
20	GKP	GKP	GKP	332	GDA	GSQ	GDA	254	GAD	GEp	GAD	176	GND	GPD	GND	98	GAR	GST	GAR
21	GRp	GRp	GRp	333	GPA	GPA	GPA	255	GPA	GTA	GPA	177	GAK	GNK	GAK	99	GEp	GEI	GEp
22	GER	GER	GER	334	Gp	Gp	Gp	256	GAp	Gp	GAp	178	GDA	GEp	GDA	100	GPS	GPA	GPS
23	GPp	GVP	GPp	335	Gp	Gp	GPp	257	GTP	GTT	GTP	179	GAp	GVV	GAp	101	GLp	GPp	GLp
24	GPQ	GPQ	GPQ	336	Gp	Gp	GPp	258	GPQ	GPQ	GPQ	180	GAp	GAp	GAp	102	GPp	GPp	GPp
25	GAR	GAR	GAR	337	Gp	Gp	GPp	259	GIA	GLL	GIA	181	GSQ	GTA	GSQ	103	GER	GLR	GER
26	GLp	GFP	GLp	338	GPP	GPP	GPP	260	GQR	GAp	GQR	182	GAp	GPA	GAp	104	GGp	GNp	GGp
27	GTA	GTP	GTA	0				261	GVV	GFL	GVV	183	GLQ	GPS	GLQ	105	GSR	GSR	GSR
28	GLp	GLp	GLp	0				262	GLp	GLp	GLp	184	GMP	GIp	GMP	106	GFp	GLp	GFp

binding of tanning ligands. After incorporation of the Type I amino acid sequences into the (Gly-Pro-Hyp)₁₂ microfibril model⁽²⁾, the structure was energy minimized. Energy minimization procedures were presented elsewhere^(1,2).

Results and Discussion

ENERGETIC EVALUATION OF THE TYPE I COLLAGEN 'SMITH' MICROFIBRIL MODEL

Table 1A presents the total potential energy (Etots) computed for the 'sum' of the five collagen triple helices, as described previously⁽¹⁾. Etotp (Table 1B) is the computed potential energy of the Type I 'Smith' microfibril model. Etotd (Table 1C) is the potential energy difference between the Type I 'Smith' microfibril packed (Etotp) and not packed (Etots). The Etotd value is negative for formation of the 'Smith' microfibril model (Table 1C, Etotd = -1263.911 kcal/mole), with respect to the sum of the potential energies of the five separate Type I collagen triple helices. This indicates that the packed microfibril model is more stable than the separate collagen molecules. Stabilizing interactions were formed in the 'Smith' microfibril for Type I collagen.

TABLE I

Computed Total Potential Energies^a For The Type I Collagen Models

A. Sum of potential energies of collagens 1-5 ^b not packed (etots) ^a	B. Computed potential energy for type I 'Smith' microfibril model (etotp) ^a	C. Potential energy difference between the type I 'Smith' microfibril model and the five separate type I triple helices (etotd) ^a
-4740.187	-6004.098	-1263.911

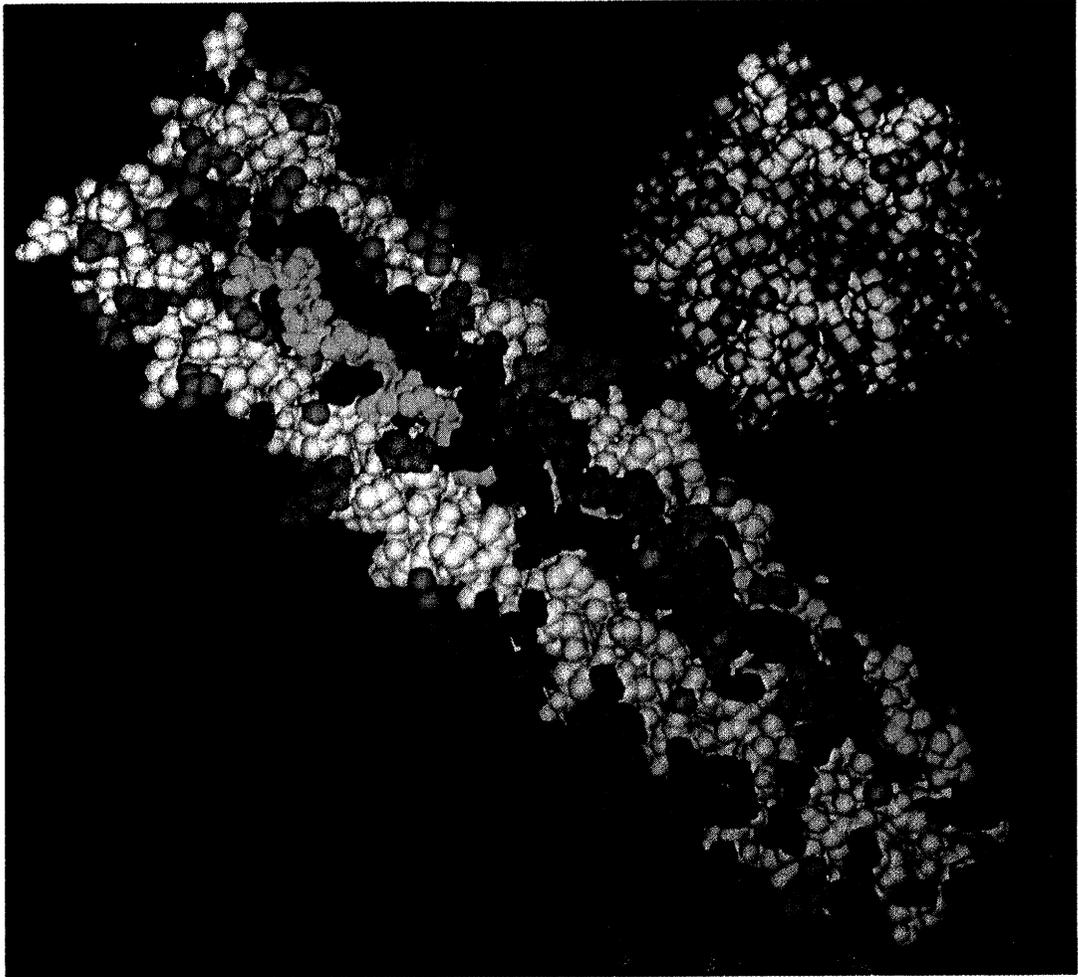
^aEnergies were computed in kcal/mole using the AMBER force fields with united-atoms^(1,2) according to the potential energy function in equation 1 as described in Methods^(1,2). Etots is the sum of the total potential energy of the five type I collagen triple helices as described in the a previous paper⁽¹⁾. Etotp is the computed potential energy for the type I 'Smith' microfibril model as described in the text. (E_{totd}) is computed by taking the difference: (E_{totp})-(E_{tot}). Although the following energy values are not shown in this table, the individual Energy terms are defined as follows: E_{bs} is the sum of energies arising from bond stretching or compression beyond the optimum bond length; E_{ab} is the sum of energies for angles which are distorted from their optimum values; E_{tor} is the sum of the torsional energies which arise from rotations about each respective dihedral angle; E_{op} is the sum of energies for the bending of planar atoms out of the plane; E_{vdw} and E_c are the sum of energies due to non-bonded van der Waals and electrostatic interactions, respectively; E_{14vdw} and E_{14e} are the sum of the van der Waals and electrostatic interaction energies, respectively for atoms connected by three bonds and E_{hb} is the sum of energies due to hydrogen bond interactions. The non-bonded E_{vdw} and E_c interaction terms were given a cutoff distance of 8 angstroms in which all the interacting atom pairs separated by a distance greater than 8 angstroms are not accounted for in the energy calculations. Solvent effects of water were not included explicitly, but were accounted for implicitly in a dielectric function (R_{ij} + 1) where R_{ij} is the distance between two separated atoms⁽³⁾.

^bSee Figure 1.

ANALYSIS OF THE TYPE I 'SMITH' MICROFIBRIL MODEL FOR COLLAGEN PACKING

Figure 2 shows the lateral and cross-sectional view of the energy minimized 'Smith' microfibril for Type I calf skin collagen. In the microfibril model, each polypeptide chain of one triple helix is shown in a different color (red, yellow and purple) to emphasize the right handed twist of the individual chains. This triple helix, with respect to the microfibril structure, has a left handed twist. Figure 2 also shows specific amino acid residues colored based on their chemical properties. For example, nonpolar residues (methionine, valine, isoleucine, leucine, tyrosine, phenylalanine) are orange, polar and neutral residues (serine, threonine, glutamine and asparagine) are green, charged basic amino acid sidechains (lysine, arginine and histidine) are blue, and charged acidic sidechains (glutamic acids and aspartic acids) are red. Red was also used in one of the polypeptide chains in Fig. 2, but the general features are distinguishable. In both the lateral and cross sectional views for the Type I microfibril model, it is clear that the functional sidechains are found clustered into groups. For example, in the lateral view of the microfibril model (Fig. 2), it is clear that the nonpolar sidechains (orange) are seen mainly in the upper portion of the model and the charged sidechains (red and blue) are seen in the lower region. This clustering effect may indicate what type of interactions are important for fibril stability⁽²⁾. The cross sectional view of the Type I microfibril model (upper-right figure, Fig. 2) also shows similar features as above for the non-polar and charged sidechains positioned around the circumference of the structure.

FIG. 2. — A color figure of the energy minimized 'Smith' microfibril model for Type I calf skin collagen constructed according to the sequence alignment in Fig. 2. The Type I microfibril is shown in a lateral and cross-sectional view. In the microfibril model, collagen 2 (see Fig. 2, tripeptide sequences 329-338) is colored (red, purple and yellow) such that each polypeptide chain is distinctly seen to have a right-handed twisting around the triple helix. The entire triple helix has a left handed twist around the microfibril structure. In this figure, specific amino acid residues are colored differently to distinguish different chemical properties. For example, nonpolar residues (methionine, valine, isoleucine, leucine, tyrosine, phenylalanine) are colored in orange, neutral/polar (serine, threonine, glutamine and asparagine) residues are colored in green, charged basic amino acid sidechains (lysine, arginine and histidine) are colored in blue and charged acidic sidechains (glutamic acids and aspartic acids) are colored in red.



STRUCTURAL ANALYSIS OF THE 'SMITH' MICROFIBRIL MODEL FOR THE PACKING OF FIVE TYPE I COLLAGEN TRIPLE HELICES

The Type I microfibril contains five collagen triple helices packed along their helical long axis. From a cross-sectional view (Fig. 2), the 'Smith' model packs the five collagens in a circular array. Each triple helix has a left-handed twist around the microfibril structure and the tilt angle of each molecule is approximately 10 degrees, with respect to the microfibrillar long axis⁽²⁾. The collagen polypeptide chain, triple helix and microfibril have a left, right and left handed twist, respectively. This process of alternating the twist handedness at each higher level of fibril formation probably maximizes fibrillar strength⁽¹⁶⁾.

ANALYSIS OF SIDECHAIN INTERACTIONS IN THE TYPE I 'SMITH' MICROFIBRIL MODEL

Previous studies revealed that non-bonded van der Waals interactions contributed the greatest to the formation of the (Gly-Pro-Hyp)₁₂ microfibril⁽²⁾. It was also shown that electrostatic interactions probably stabilize the packed structure and contribute to the specificity of inter-helical collagen interactions⁽²⁾.

The energy minimized Type I microfibril is based on the alignment of the alpha 1 and 2 polypeptide sequences of five collagen molecules as shown in Fig. 1⁽¹⁾. In studies involving the linear amino acid sequences of Type I collagen, amino acid residues with similar properties tend to form clusters along the collagen polypeptide chains. For instance, clusters of charged sidechains were distinguished from clusters formed by non-polar sidechains. These charged regions correlated with the dark bands seen under an electron microscope for samples of positively stained transverse sections of collagen⁽¹⁰⁾. Since the Type I model is based on this mode of collagen packing, it would be desirable to determine if the clusters of charged and uncharged sidechains can effectively interact, three-dimensionally, to stabilize the present microfibril model.

In the interior of the microfibril model, sidechains with positive charges are usually paired with or surrounded by neighboring sidechains with negative charges. Non-polar sidechains are usually found in clusters packed with adjacent non-polar groups. Unlike analysis involving linear polypeptide sequences, which can only suggest possible sidechain contacts, the present three-dimensional microfibril model allows for the observation, identification and study of specific interactions which may occur in the native state.

STABILIZING INTERACTIONS OF CHARGED SIDECHAINS WITHIN THE MICROFIBRIL MODEL: ARGININE AND LYSINE

In a previous study⁽¹⁾ describing the collagen triple helical models, it was shown that arginines in the 'X' position of Gly-X-Y formed intra-polypeptide interactions to stabilize the polypeptide chain, whereas 'Y' position arginines did not form this interaction and are probably involved in inter-helical packing interactions. The microfibril model shows that arginines in the 'Y' positions form stabilizing interactions with surrounding charged and polar groups. Examination of the three-dimensional model shows that arginine sidechains in the interior are usually paired with acidic groups such as glutamic acids. Specific and energetically favorable interacting complexes can be observed. Although the structure-function relationship of tanning agents for arginine sidechains in the interior are usually paired with acidic groups such as glutamic acids. Specific and energetically favorable interacting complexes can be observed. Although the structure-function relationship of tanning agents for arginine sidechains is not discussed here, future studies using other computational methods, such as molecular dynamics^(17,18), will be applied to determine possible ligand interactions within these domains.

In the interior of the representative Type I microfibril model, specific interactions due to lysine sidechains are less well defined. The sidechains of lysine are very flexible (i.e., four main sidechain torsional angles in $-(\text{CH}_2)_4-\text{NH}_3^+$) and probably interact less specifically for a given region than arginines. Hence, lysine sidechains are better targets for potential chemical modifying agents than arginines which are more likely to be bound.

In the exterior surface of the 'Smith' microfibril model, many arginine and lysine sidechain positions are similar to those described for the triple helical structures⁽¹⁾. In cases where arginines are near acidic groups, charged interactions do form but are probably non-specific since these exterior regions will interact with adjacent collagen structures during fibrillar growth.

STABILIZING INTERACTIONS INVOLVING HYDROXYPROLINE CLUSTERS WITHIN THE INTERIOR OF THE 'SMITH' MICROFIBRIL

Figure 3 shows a space filling model (in the 'relaxed' stereo mode) of a cluster formed by adjacent hydroxyprolines. The top stereo-figure is a lateral view and the bottom stereo-figure is a cross sectional view. Here, hydroxyprolines (purple with the $-\text{OH}$ group in red/blue) from adjacent collagen molecules are clustered in the interior of the microfibril. The hydroxyl group of each hydroxyproline is polarized toward the acidic sidechain of a glutamic acid (sidechain is red/white, see upper stereo-figure). This glutamic acid is also stabilized by an adjacent serine (green with the $-\text{OH}$ group in red/blue, see upper stereo-figure). These neighboring hydroxyproline clusters within the microfibrillar interior are sites for potential chemical ligands. Since each hydroxyproline $-\text{OH}$ is from a different triple helix, very effective stabilization of the fibrillar structure is possible if a specific ligand could be designed to form multiple crosslinks at each cluster site.

LIMING IN THE TANNING PROCESS MAY INCREASE THE NUMBER OF POTENTIAL BINDING SITES FOR CHROMIUM (III)

Liming is a stage in the leather manufacturing process where the skin is bathed within a saturated solution of calcium hydroxide. This step removes hair, epidermis and other undesired non-collagenous materials. Liming also produces a swelling effect in the skin which results in more chromium intake⁽¹⁹⁾. Liming may hydrolyze glutamine and asparagine sidechains to glutamic and aspartic acid sidechains, respectively⁽²⁰⁾. Since Cr(III) is thought to complex with and crosslink acidic sidechain groups⁽⁴⁾, liming may actually increase the number of Cr(III) binding sites within the collagen packing. Figure 4 is a color figure of the Type I 'Smith' microfibril model where the structure on the left represents a pre-limed model and the structure on the right is the limed model. Acidic sidechains are colored red and the glutamine and asparagine sidechains are colored green. It is clear that upon hydrolysis of the glutamine and asparagine sidechains by calcium hydroxide, the limed structure shows a significant increase in the number of potential acidic sites.

As to the swelling effect of the skin tissue due to liming, it is possible that the increased number of acidic functional groups perturbs the native packing of Type I collagen. Our modeling studies indicate that electrostatic interactions within collagens are important for the specificity of packing between the molecules⁽²⁾. Therefore, if there were an increase in the proportion of specific charged sidechains, the packing arrangement of native collagens would be altered and this could be a cause of the swelling effect observed during liming. Furthermore, liming probably disrupts the original interactive states of the glutamines and asparagines. For example, the possible stabilizing interactions between

STRUCTURAL ANALYSIS OF THE 'SMITH' MICROFIBRIL MODEL FOR THE PACKING OF FIVE TYPE I COLLAGEN TRIPLE HELICES

The Type I microfibril contains five collagen triple helices packed along their helical long axis. From a cross-sectional view (Fig. 2), the 'Smith' model packs the five collagens in a circular array. Each triple helix has a left-handed twist around the microfibril structure and the tilt angle of each molecule is approximately 10 degrees, with respect to the microfibrillar long axis⁽²⁾. The collagen polypeptide chain, triple helix and microfibril have a left, right and left handed twist, respectively. This process of alternating the twist handedness at each higher level of fibril formation probably maximizes fibrillar strength⁽¹⁶⁾.

ANALYSIS OF SIDECHAIN INTERACTIONS IN THE TYPE I 'SMITH' MICROFIBRIL MODEL

Previous studies revealed that non-bonded van der Waals interactions contributed the greatest to the formation of the (Gly-Pro-Hyp)₁₂ microfibril⁽²⁾. It was also shown that electrostatic interactions probably stabilize the packed structure and contribute to the specificity of inter-helical collagen interactions⁽²⁾.

The energy minimized Type I microfibril is based on the alignment of the alpha 1 and 2 polypeptide sequences of five collagen molecules as shown in Fig. 1⁽¹⁾. In studies involving the linear amino acid sequences of Type I collagen, amino acid residues with similar properties tend to form clusters along the collagen polypeptide chains. For instance, clusters of charged sidechains were distinguished from clusters formed by non-polar sidechains. These charged regions correlated with the dark bands seen under an electron microscope for samples of positively stained transverse sections of collagen⁽¹⁰⁾. Since the Type I model is based on this mode of collagen packing, it would be desirable to determine if the clusters of charged and uncharged sidechains can effectively interact, three-dimensionally, to stabilize the present microfibril model.

In the interior of the microfibril model, sidechains with positive charges are usually paired with or surrounded by neighboring sidechains with negative charges. Non-polar sidechains are usually found in clusters packed with adjacent non-polar groups. Unlike analysis involving linear polypeptide sequences, which can only suggest possible sidechain contacts, the present three-dimensional microfibril model allows for the observation, identification and study of specific interactions which may occur in the native state.

STABILIZING INTERACTIONS OF CHARGED SIDECHAINS WITHIN THE MICROFIBRIL MODEL: ARGININE AND LYSINE

In a previous study⁽¹⁾ describing the collagen triple helical models, it was shown that arginines in the 'X' position of Gly-X-Y formed intra-polypeptide interactions to stabilize the polypeptide chain, whereas 'Y' position arginines did not form this interaction and are probably involved in inter-helical packing interactions. The microfibril model shows that arginines in the 'Y' positions form stabilizing interactions with surrounding charged and polar groups. Examination of the three-dimensional model shows that arginine sidechains in the interior are usually paired with acidic groups such as glutamic acids. Specific and energetically favorable interacting complexes can be observed. Although the structure-function relationship of tanning agents for arginine sidechains in the interior are usually paired with acidic groups such as glutamic acids. Specific and energetically favorable interacting complexes can be observed. Although the structure-function relationship of tanning agents for arginine sidechains is not discussed here, future studies using other computational methods, such as molecular dynamics^(17,18), will be applied to determine possible ligand interactions within these domains.

FIG. 3. — A space-filling model in the relaxed stereo-viewing mode, showing an interesting sidechain interaction complex within the microfibrillar interior. Here, hydroxyprolines (purple with —OH group in red/blue) from adjacent collagen molecules are in the interior of the microfibril model. The hydroxyl group of each hydroxyproline is polarized toward the acidic sidechain of a glutamic acid residue (sidechain in red/white, see top stereo-figure). This glutamic acid is also stabilized by an adjacent serine residue (green with —OH in red/blue, see upper stereo-figure).

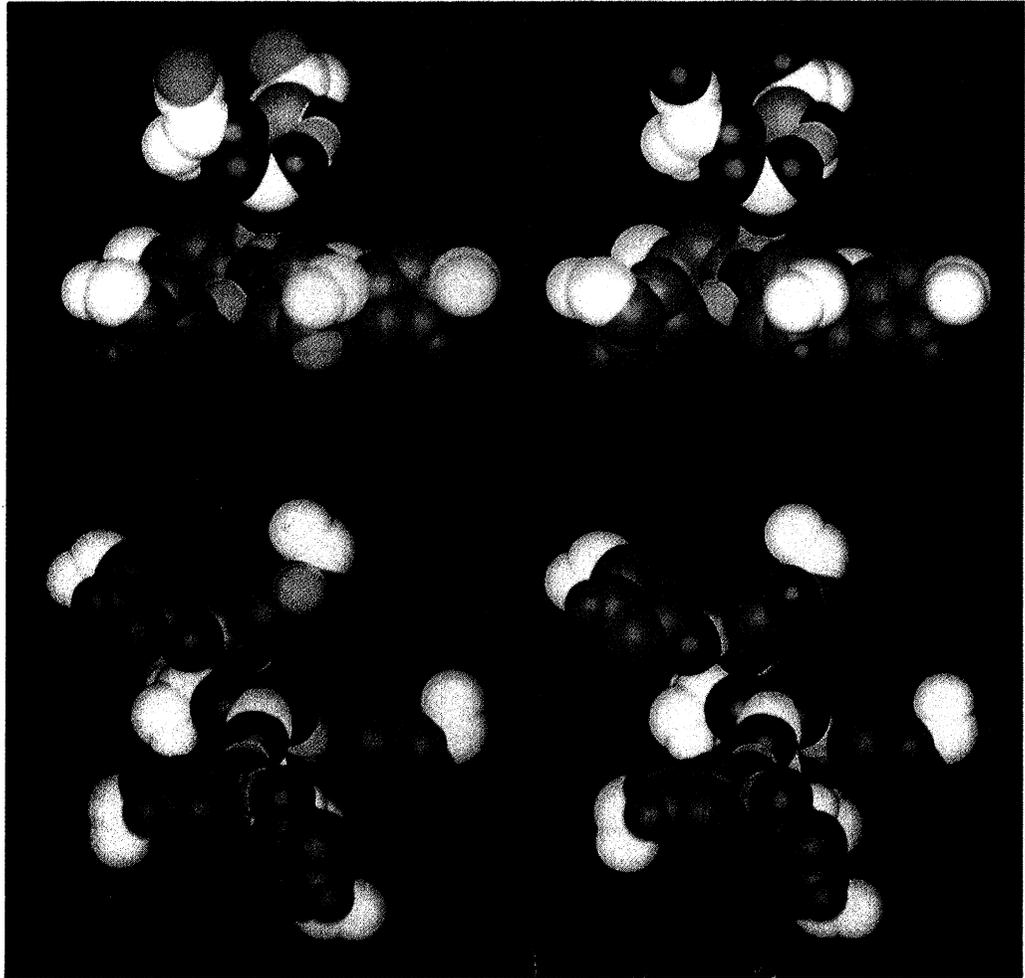
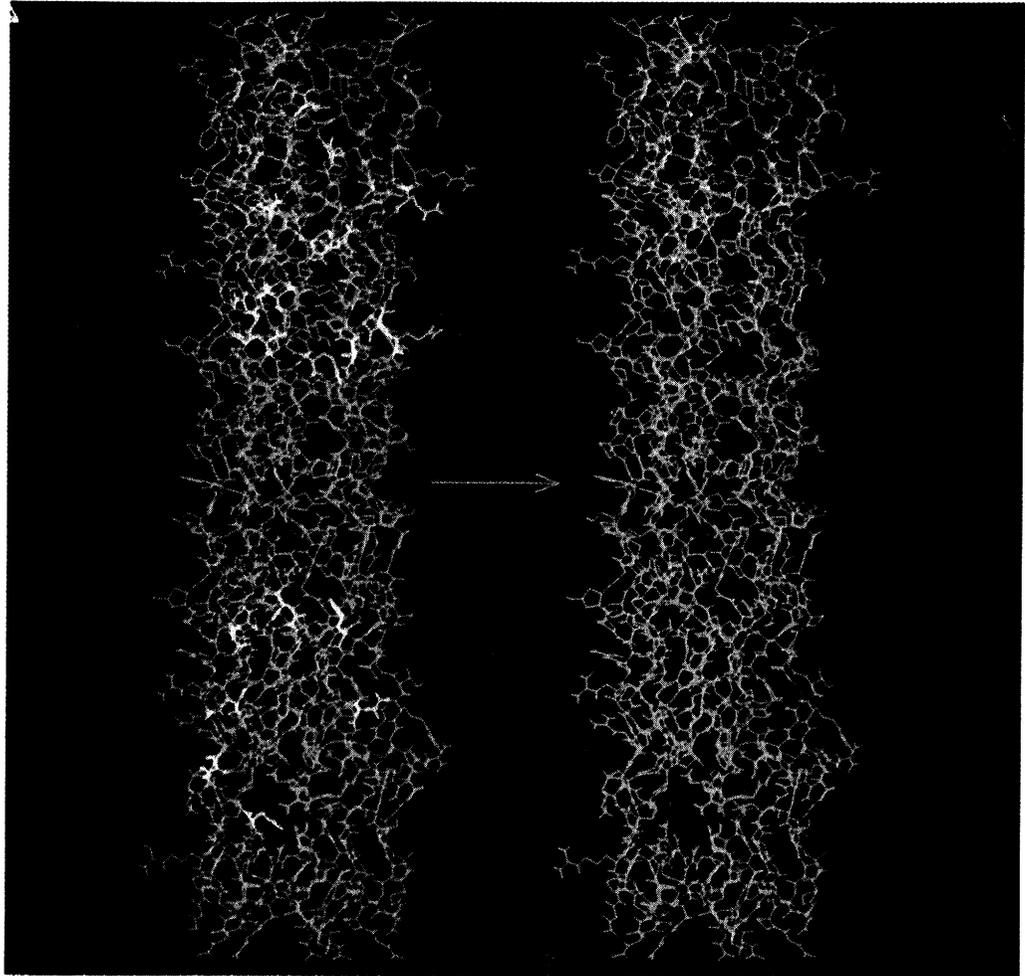


FIG. 4. — A color figure of the Type I 'Smith' microfibril model where the structure on the left represents a pre-limed model and the structure on the right is the limed model. The acid sidechains are colored red, and the glutamine and asparagine sidechains are colored green. It is clear that upon oxidation by the calcium hydroxide, the limed structure shows a significant increase in the number of potential acidic sites.



aspartic acid and asparagine sidechains, described in a previous study⁽¹⁾, would be disrupted. The region containing the Asp/Asn interactions is now a destabilized region containing closely packed negative charged sidechains of the type Asp/Asp* (i.e., the asterisk defines the Asn --> Asp modification). Within collagen, these negatively charged regions would weaken and/or alter the native packing state. The overall increased hydrophilicity and weakened native packing forces between the collagen molecules and microfibrillar units would permit an increased hydration state. The limed skin would permit better Cr(III) distribution or intake which results in maximizing chromium-collagen interactions.

Conclusion

This study described the use of the (Gly-Pro-Hyp)₁₂ 'Smith' microfibril model⁽²⁾ to determine possible Type I collagen interactions. The energy minimized structure of the Type I 'Smith' microfibril model was relatively stable when packed similarly to the starting (Gly-Pro-Hyp)₁₂ microfibril model. Non-bonded van der Waals and electrostatic interactions showed similar contributions within this initial model. However, these values may show further stabilization as the structure is refined through both interactive and energy minimization procedures (work in progress). This refinement step is necessary since the initial Type I model is based on the energy minimized packing of the (Gly-Pro-Hyp)₁₂ microfibril model. Since the linear amino acid sequences of Type I collagen contain varying tripeptide sequences, regions with a high density of large/long sidechains will require a modification of the packing found for the microfibril model of (Gly-Pro-Hyp)₁₂⁽²⁾. Nevertheless, the initial energy minimized Type I model, overall, shows favorable stability when compared to the five unpacked Type I collagen triple helices (Table 1A-C). The microfibril model allows sidechains to form stabilizing interactions. Further study of the initial Type I microfibril model will involve the identification of all possible stabilizing sidechain-sidechain and sidechain-backbone interactions (i.e., such as hydrogen bonds, charge-charge and nonpolar interactions). This Type I model will be modified in order to maximize these potential interactions. However, the model described here provides a reasonable three-dimensional view of possible collagen interactions.

The design of chemical ligands which interact specifically with proteins requires a detailed knowledge of the three-dimensional structure of the proteins. Although a method such as 'Quantitative Structure-Activity relationships', (QSAR), can be used to determine the binding affinity characteristics of a set of active compounds, this method also requires a large set of binding data for known ligands which may or may not interact specifically at an active site⁽²¹⁾. In the case of a protein environment, which is complex and consists of multiple binding sites, QSAR analysis is probably less effective.

The specific interactions of tanning agents in modifying calf skins are not precisely defined. Since there is no information concerning the actual three-dimensional structure of Type I collagen, knowledge of the interactions of these tanning ligands is based mainly on experimental studies. Although these studies are important to identify possible reactive conditions or groups (i.e., such as acidic or basic amino acid sidechains) and may provide information useful in the modification of known tanning solutions, they provide little information however, leading to the design of a specific ligand. Unlike many globular proteins whose functional properties are contained within a single structure or small polymeric complexes of the single globular protein, the functional properties of Type I skin are based on the specific packing of collagen molecules. The geometric parameters for possible ligands are probably dependent on the successful identification of channels formed by the packed collagen molecules or microfibrillar units. Possible active sites for

specific cross-linking interactions are based on the identification of specific reactive environments within the Type I collagen packing. The successful design of a ligand must satisfy both the geometric and reactive constraints of Type I tissue. Molecular modeling is an effective tool when applied to understanding the structure-function relationships of collagen packing.

Acknowledgment

The authors would like to thank Drs. Arthur J. Miller and Chun E. Kung for helpful suggestions and review of this manuscript.

Note Added in Proof

The three-dimensional structures for the collagen structures as described in ref. 2 have been deposited with the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973, from which copies are available.

References

1. Chen, J.M., Fairheller, S.H., and Brown, E.M., *J. Am. Leather Chemists' Assoc.*, (1991).
2. Chen, J.M., Kung, C.E., Fairheller, S.H., and Brown, E.M., *J. Protein Chem.*, **10**(5), 535-552, (1991).
3. Smith, J.W., *Nature*, **219**, 157-158, (1968).
4. Bienkiewicz, K.J., In "Physical Chemistry of Leather Making", R.E. Krieger Publishing Co., Florida, pp. 324-352, (1983).
5. Brodsky, B., Eikenberry, E.F., Belbruno, K., and Sterling, K., *Biopolymers*, **21**, 935-951, (1982).
6. Chapman, J.A., and Hulmes, D.J.S., In "Ultrastructure of the Connective Tissue Matrix", Ruggeri, A., Motta, P.M., (eds.), Martinus Nijhoff Publishers, Boston, pp. 1-33, (1984).
7. Chapman, J.A., In "Connective Tissue Matrix: Topics in Molecular and Structural Biology", vol. 5, Hukins, D.W.L. (ed.) Verlag Chemie, Weinheim, pp. 89-132, (1984).
8. Chew, M.W.K., and Squire, J.M., *Int. J. Biol. Macromol.*, **8**, 27-36, (1986).
9. Eikenberry, E.F., and Brodsky, B., *J. Mol. Biol.*, **144**, 397-404, (1980).
10. Piez, K.A., In "Extracellular Matrix Biochemistry", Piez, K.A., Reddi, A.H. (eds.) Elsevier, New York, pp. 1-39, (1984).
11. Ramachandran, G.N., *J. Am. Leather Chemists' Assoc.*, **63**, 160-191, (1968).
12. Hodge, A.J., and Petruska, A.J., In "Aspects of Protein Structure", Ramachandran, G.N. (ed.) Academic Press, London, pp. 289-300, (1963).
13. Miller, A., and Wray, J.S., *Nature*, **230**, 437-439, (1971).
14. Hulmes, D.J.S., and Miller, A., *Nature*, **282**, 878-880, (1979).
15. Miller, A., and Tocchetti, D., *Int. J. Biol. Macromol.*, **3**, 9-18, (1981).
16. Piez, K.A., and Trus, B.L., *J. Mol. Biol.*, **110**, 701-704, (1990).
17. McCammon, J.A., and Karplus, M., *Ann. Rev. Phys. Chem.*, **31**, 29-45, (1980).
18. Gao, J., Kuczera, K., Tidor, B., and Karplus, M., *Science*, **244**, 1069-1072, (1989).
19. See ref., **4**, pp. 275-299.
20. Beek, J., Jr. and Sookne, A.M., *J. Am. Leather Chemists' Assoc.*, **34**, 641-643, (1939).
21. Gupta, S.P., *Chem. Rev.*, **89**, 1765-1800, (1989).