Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 9860–9864, October 1992
Biophysics

Helical model of nucleation and propagation to account for the growth of type I collagen fibrils from symmetrical pointed tips: A special example of self-assembly of rod-like monomers

DEBORAH SILVER*, JOHN MILLER*, ROBERT HARRISON†, AND DARWIN J. PROCKOP‡§

*Department of Electrical and Computing Engineering, Rutgers University, Piscataway, NJ 08855; and †Department of Pharmacology, Jefferson Cancer Institute, and ‡Department of Biochemistry and Molecular Biology, Jefferson Institute of Molecular Medicine, Jefferson Medical College, Thomas Jefferson University, Philadelphia, PA 19107

Contributed by Darwin J. Prockop, July 10, 1992

ABSTRACT **A model was developed to account for the recent observations indicating that type I collagen fibrils assembled *in vivo* grow from symmetrical pointed tips. The essential features of the model are (*i*) a distinctive structural nucleus forms at each end of a growing fibril and growth of the fibril then proceeds by propagation of the two structural nuclei, (*ii*) the two structural nuclei have similar spiral or helical conformations, and (*iii*) assembly of each structural nucleus requires two kinds of specific binding steps defined as 3.4 *D*-period and 0.4 *D*-period overlaps, but propagation of the nucleus requires only the 3.4 *D*-period binding step.**

The fibrils of type I collagen found in skin, tendons, bone, and other connective tissues of vertebrates are among the largest and most abundant protein polymers in living organisms (see ref. 1). The packing of the monomer into fibrils is closely related to the repeating clusters of charged and hydrophobic side chains of amino acids that are located on the surface of the molecule and that divide it into 4.4 segments or *D* periods (1–12). In the longitudinal direction, the monomers are packed head-to-tail with a gap of ≈0.6 *D* period and, therefore, with a repeat of 5 *D* periods. The continuity of the fibril is maintained by many of the monomers being staggered by 1, 2, 3, or 4 *D* periods relative to nearest neighbors so as to generate gap and overlap regions. The high degree of order in the longitudinal packing of collagen fibrils has been agreed upon for several decades. However, there has been continuing disagreement about the lateral packing of the monomers (1–12). One current view is that the monomers are laterally packed in a tilted quasi-hexagonal lattice (4, 12). A related view is that the fibril consists of "compressed" microfibrils, each of which consists of monomers helically coiled into a rope-like pentameric structure (3, 5). However, still another current view is that the lateral packing of collagen in many fibrils is either liquid-like (13) or a biological equivalent of a liquid crystal (6, 11). Also of interest is that some observations (1, 14, 15) on the reassembly of fibrils from soluble collagen extracted from tissues suggested that the first structures formed are monomers bound by 0.4 *D*-period overlaps (4 *D* staggers). Others (1, 15, 16) suggested that the initial stages involve assembly of long and thin filaments similar to microfibrils. Still other observations (13) indicated that the first fibrils assembled have pointed tips similar to the pointed tips subsequently seen in serial sections of chicken tendons (17).

Recently, we have developed a system for studying assembly of type I collagen into fibrils by enzymic cleavage of a purified soluble precursor of collagen under physiological conditions (18–24). The fibrils formed at 37°C are as thin as type I collagen fibrils seen in tissues (20), but the fibrils formed at 29°C–34°C are thicker. Because the thicker fibrils settle from solution, it was possible to use dark-field light microscopy to follow growth of the fibrils from intermediate stages (22). The first fibrils detected had a blunt end and a pointed end or tip. Initial growth of the fibril was exclusively from the pointed end or α-tip. Later, β-tips appeared on the blunt ends and the fibrils then grew in both directions. Scanning transmission electron microscopy analysis (24) revealed that both the α-tips and β-tips were near-paraboloidal in shape. Also, all the monomers were oriented with N termini directed toward the tips (22, 24).

METHODS

Computer simulations of fibril models were carried out with a Silicon Graphics workstation. Each collagen monomer was represented as a cylinder with a length defined as 4.4 *D* periods. Models for fibril assembly were simulated by addition of one monomer at a time according to rules for binding through unidirectional 3.4 *D*-period overlaps or 0.4 *D*-period overlaps as described in *Results*. Interactive programs were written so that the models could be rotated in three dimensions after each monomer addition and so that different growth patterns could be explored in two or three dimensions.

RESULTS

Criteria for the Model. The major criteria for developing a model were that it account for the following observations made on fibrils assembled *de novo* at 37°C (Fig. 1): (*i*) in cross-section, the α-tips were circular down to a tip of <25 nm (22, 24); (*ii*) axially, the α-tips were near-paraboloidal, as indicated by a decrease in mass toward the tip that was an average of 17 molecular *D* segments per *D* period and that was linear over 100 or more *D* periods (24); (*iii*) the α-tips on both short and long fibrils were essentially the same (24), an observation implying that the contour remained constant as the fibrils grew; (*iv*) the β-tips were also near-paraboloidal but were more irregular with slopes that varied from 50 to 200 molecular *D* segments per *D* period (24); (*v*) after the β-tips appeared, both α-tips and β-tips grew simultaneously on the same fibril (24); (*vi*) all monomers were oriented so that the N termini were directed toward the tips (22, 24); and (*vii*) the shafts of all the fibrils formed under the same conditions had about the same diameters (21, 23).

Initial Considerations of Alternative Models. Initial considerations indicated that the observations could not be explained in any simple manner by several previous suggestions about fibril growth (Table 1). For example, assembly driven primarily by surface tension and liquid-like forces (see ref. 11) cannot readily explain the persistent differences in con-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

§To whom reprint requests should be addressed.

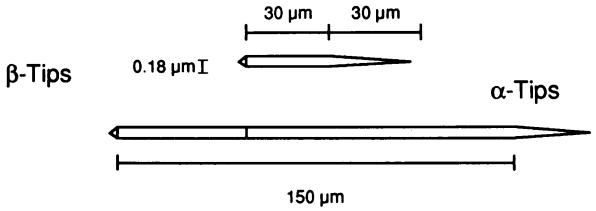

FIG. 1. Schematic for growth of fibrils from pointed tips (22, 24). Vertical line indicates the site of apparent change in polarity of monomers in the fibril.

tour between α-tips and β-tips as the two kinds of tips grow simultaneously (Fig. 1). Lateral aggregation or coalescence of long microfibrils (1, 15–17) is excluded by the microscopic observations on growth of the intermediate fibrils (22, 24). Several complex assumptions are required to explain growth of the tips of fibrils formed by laterally packed microfibrils (Table 1).

The Helical Model of Nucleation and Propagation. Models were developed here on three assumptions: (*i*) the fibril is probably assembled by a limited number of specific binding steps, (*ii*) the growth probably occurs by formation of a structural nucleus for each tip, and then propagation of the two nuclei, and (*iii*) the structural nuclei are probably spiral or helical. After a series of trial experiments, the simplest model developed was one in which monomers were bound unidirectionally in 3.4 *D*-period overlaps to form strands, and the strands were coiled in the same direction to form concentric helical layers. The strands that formed the center of the tip were coiled into a pentameric helical core similar to a Smith-type microfibril (Fig. 2 *Top*). Additional strands were then coiled around the pentameric core (Fig. 2 *Middle* and *Bottom*). The strands in each layer were interlocked by assuming that each monomer in a strand had a 0.4 *D*-period overlap on one side, with a monomer in a trailing strand, and another 0.4 *D*-period overlap on the opposite side, with a monomer in a leading strand. A computer simulation of the model demonstrated that a fibril with a paraboloidal tip was readily assembled from 12,000 cylinders (Figs. 3 and 4).

Stepwise Assembly of the Model. In assembly of the model, the first step was assumed to be a 3.4 *D*-period overlap that was repeated unidirectionally (Fig. 5, frames 1–5). As the fifth monomer was added, the strand coiled to form a pentameric helical core or a microfibril (frame 5). The pentameric core was then propagated longitudinally by monomers, each of which was bound by both a 3.4 and a 0.4 *D*-period overlap (frame 6). To simplify analysis, the propagation of each strand in Fig. 5 was arbitrarily stopped after it acquired six monomers. The second strand in the model was then initiated through a 0.4 *D*-period overlap binding step (frame 7). After initiation of the strand, it was propagated by a series of repetitive 3.4 *D*-period binding steps. All subsequent strands were formed in a similar manner. Therefore, initial assembly of the model (Fig. 5) involved three general kinds of binding steps: (*i*) 3.4 *D*-period overlaps, (*ii*) 0.4 *D*-period overlap, and (*iii*) less specific layer-to-layer binding steps. To explain the uniform diameters of the shafts (22–24), it was assumed that initiation of new strands via the 0.4 *D*-period overlap became less favorable as the diameter increased and the surface became more planar, leading to the formation of a structural nucleus (Fig. 3, frame 4). Thereafter, the nucleus was propagated through nearby equivalent 3.4 *D*-period overlaps (Fig. 3, frames 5–8).

Table 1. Alternative models

Model	Assumptions necessary to explain observations on tip growth
Random additions of monomers or oligomers followed by liquid-like reassembly	Mechanism (?) to explain difference in geometry between simultaneously growing α-tips and β-tips
Microfibrillar subunits	(*i*) Mechanism (?) to explain difference in geometry between simultaneously growing α-tips and β-tips, (*ii*) mechanism (?) to explain smooth near-paraboloidal contour of the tips,* (*iii*) vectoral insertion of every fifth monomer or of pentamers along the tips†

*Change in slope through a cross-section of an α-tip was 1 molecular diameter (≈1.4 nm) per 4–8.6 *D* periods or per ≈1 monomer length (5 *D* periods) with an SD of only about ±8% (see ref. 24).
†If microfibrils form the tip and grow by addition of monomers, every fifth monomer must be inserted vectorally between the growing microfibril, and previously assembled microfibrils along the slope of the tip. If microfibrils grow by addition of pentamer subunits, each pentamer must be inserted vectorally.

Three Arbitrary Choices in the Model. In assembling the model, one arbitrary choice made was that the first binding step (Fig. 5, frame 1) was a 3.4 *D*-period overlap. Essentially the same results were obtained if the first binding step was a 0.4 *D*-period overlap (data not shown). A second arbitrary choice was that the strands of monomers were coiled into concentric layers or hollow cylinders. A model in which the strands were coiled into an equable spiral with strands continuous from one layer to the next also seemed acceptable but was not examined in detail. A third arbitrary choice was made to resolve the 2π molecular problem (10), whereby geometric considerations dictate that successive concentric layers differ in circumference by 2π molecular diameters, while maintenance of the *D* periodicity requires increments in multiples of 5. Galloway (10) resolved this problem in his helical model by assuming that the density of packing decreased from the outside of the fibril inward. In the model described here, our solution was to introduce occasional head-to-tail gaps that were larger than the 0.6 *D* period between some of the strands in a layer (Fig. 2 *Middle* and *Bottom*), so that each layer in cross-section contained an integral multiple of 5 molecules (Fig. 4). These occasional

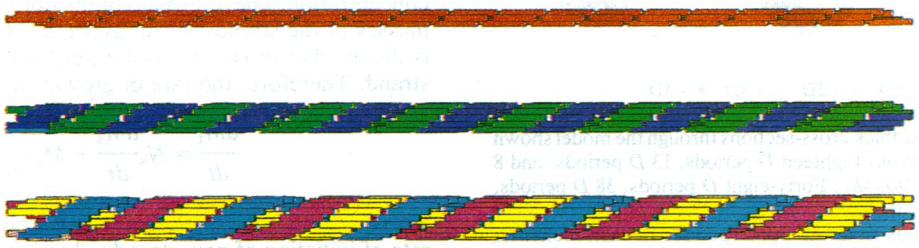

FIG. 2. Computer simulations of the model developed here for fibril assembly. (*Top*) An inner pentamer core. (*Middle*) Two strands wrapped around an inner pentamer core to form a helical layer. (*Bottom*) Three additional strands wrapped around the structure in *Middle* to form a second helical layer around the pentameter core. A gap of >0.6 *D* periods was introduced between two of the strands.

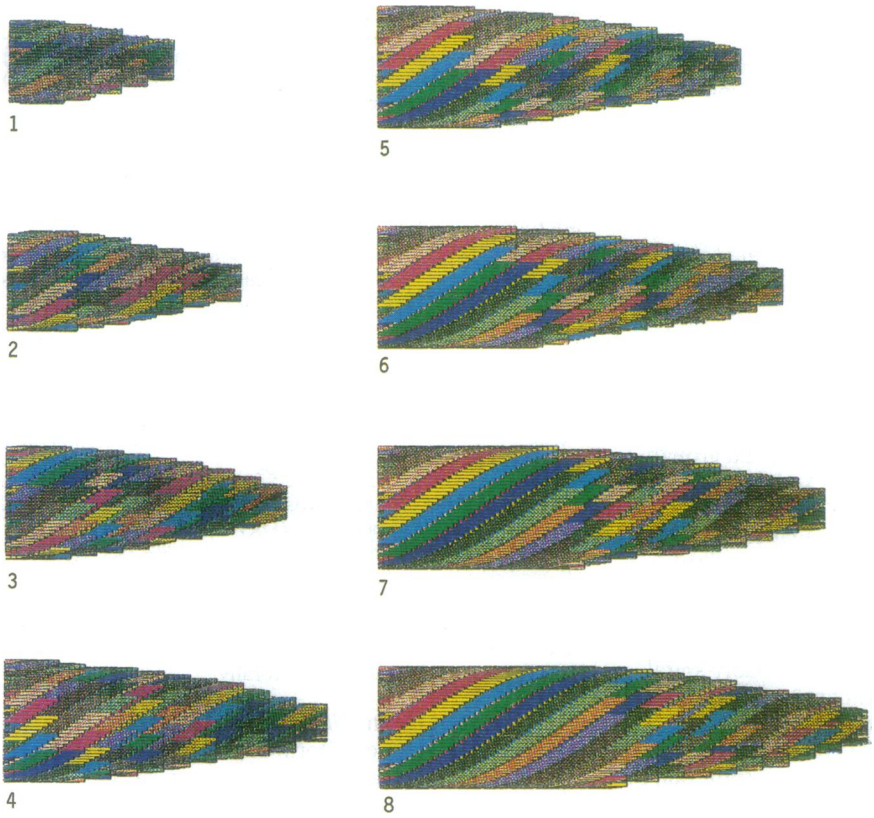

FIG. 3. Computer-simulated growth of the model. The first frame contains 1500 monomers. An additional 1500 molecules were added in each frame so that the structure in the eighth frame contains 12,000 molecules, has a tip of 10 molecular lengths (50 *D* periods), and a shaft with a diameter of 30 molecules (equivalent to ≈50 nm). The slope of the tip is 17 molecules per *D* period. Each monomer in a strand has the same color, with the intensity of the outer strands less than the intensity of the inner strands.

gaps were not propagated between layers and were not apparent in all cross-sections of the model (Fig. 4).

Mechanisms for Generating Different Nuclei and Growth Patterns. Several mechanisms were considered for generating different paraboloidal nuclei for the α-tips and β-tips. One was a stutter mechanism that postulated there was a finite probability of nonproductive additions of monomer during the binding steps required to initiate a new strand or to propagate a strand. Therefore, there was a temporary delay or stutter in growth of the strand. Since a strand in one layer of the model cannot propagate beyond a strand immediately beneath it, the propagation of multiple strands in successive layers would be slowed. Computer simulation indicated that the slope of the tip was proportional to the probability of an event that temporarily delayed initiation or growth of a strand. Figs. 3 and 4 present a tip in which growth of one strand in 17 was temporarily delayed to generate a tip with an incremental mass slope of 17 molecular *D* segments per *D* period.

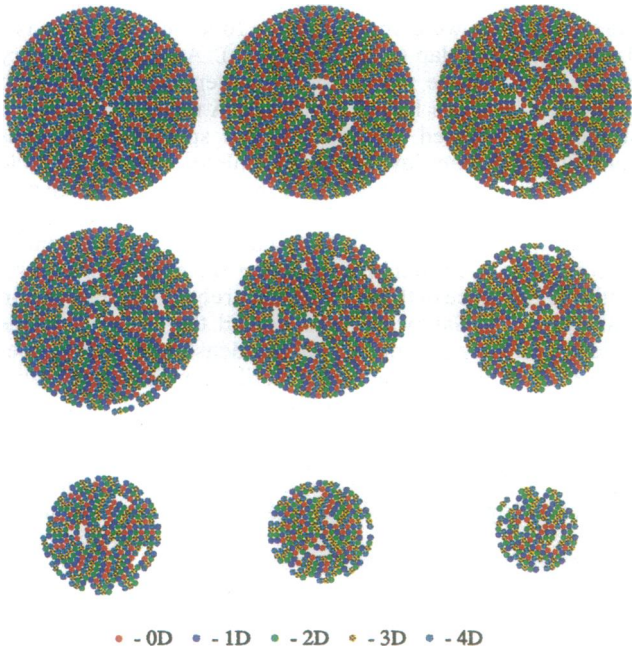

FIG. 4. One *D*-period-thick cross-sections through the model shown in Fig. 3 (frame 8). (*Bottom*) Eighteen *D* periods, 13 *D* periods, and 8 *D* periods from the tip. (*Middle*) Forty-eight *D* periods, 38 *D* periods, and 28 *D* periods from the tip. (*Top*) Eighty-eight *D* periods, 73 *D* periods, and 58 *D* periods from the tip. *D* periods in each monomer were colored as indicated. Light blue for the 4 *D* periods included both the last 0.4 *D* period of the monomer and the 0.6 *D*-period gap. Open areas reflect gaps of >0.6 *D* period introduced to resolve the 2π problem (see text).

A second mechanism for accounting for the differences between the α- and β-tips was based on the observation that β-tips arose from the blunt ends of the first fibrils detected by dark-field light microscopy (22). Irregularities on the surfaces of the blunt ends may well have caused assembly of nuclei that were more irregular and had a steeper slope than the nuclei for the α-tips (24).

A third mechanism for generating different α-tips and β-tips was a kinetic one that assumed that the 0.4 *D*-period overlap to initiate new strands and the 3.4 *D*-period overlap to extend existing strands had a different dependence on monomer concentration. The total mass of the fibril (M_t) is the sum of the masses of the strands, or on average: $M_t = N_s M_s$, where N_s is the number of strands in the fibril and M_s is the mass of a strand. Therefore, the rate of growth of mass (dM_t/dt) is

$$\frac{dM_t}{dt} = N_s\frac{dM_s}{dt} + M_s\frac{dN_s}{dt}, \qquad [1]$$

where dM_s/dt is the rate of growth of strands and dN_s/dt is the rate of initiation of new strands. Assuming a constant rate of growth of strands dependent on monomer concentration yields $dM_s/dt = k_1[m]$, where k_1 is a rate constant for strand growth through 3.4 *D*-period overlaps and [m] is the monomer con-

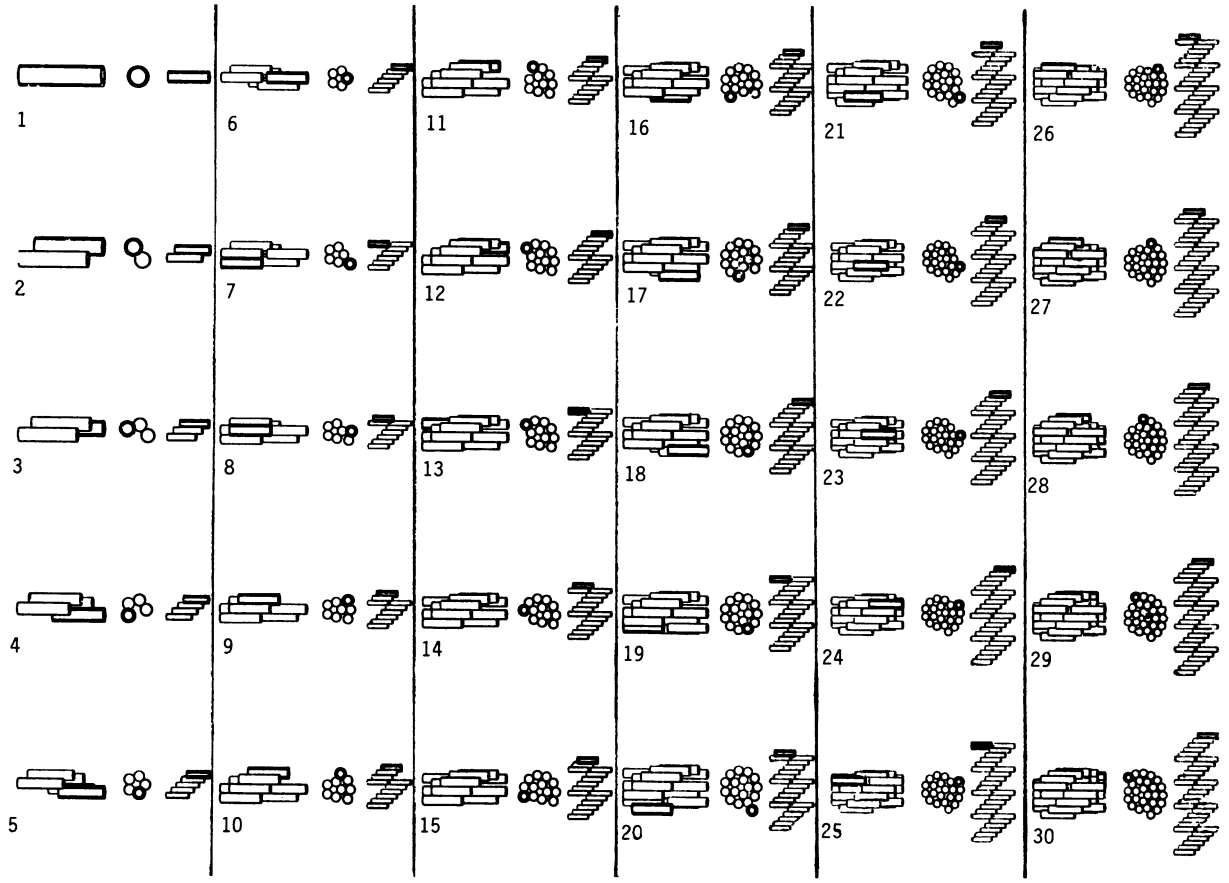

FIG. 5. Individual steps in assembly of the model. Three views of the structure formed through 30 stepwise additions of monomer are shown: lateral view, axial view (in projection), and a two-dimensional schematic to illustrate the 3.4 *D*-period and 0.4 *D*-period binding steps. The last monomer added is outlined more heavily than previously added monomers. Schematic does not reflect the layers of the structure. To simplify the presentation, growth of the pentamer core was arbitrarily stopped after it acquired 6 monomers. Therefore, the slope of the tip is infinite (see text).

centration. To account for the observation that all fibrils have about the same maximal diameters, lateral growth can be limited by using a standard equation for limited processes,

$$\frac{dN_s}{dt} = k_2[\mathrm{m}](N_{\mathrm{smax}} - N_s), \qquad [2]$$

where k_2 is the rate constant associated with the initiation of new strands through 0.4 *D*-period overlaps and the boundary conditions are $N_s = 0$ at $t = 0$, and $N_s = N_{\mathrm{smax}}$ at $t = \infty$. Eq. 2 then solves to $N_s = N_{\mathrm{smax}}(1 - e^{-k_2[\mathrm{m}]t})$ and its first derivative is $dN_s/dt = k_2[\mathrm{m}]N_{\mathrm{smax}}e^{-k_2[\mathrm{m}]t}$.

With the appropriate substitutions, Eq. 1 becomes

$$\frac{dM_t}{dt} = k_1[\mathrm{m}]N_{\mathrm{smax}}(1 - e^{-k_2[\mathrm{m}]t}) + k_2[\mathrm{m}]N_{\mathrm{smax}}\, e^{-k_2[\mathrm{m}]t}\, M_s. \qquad [3]$$

When monomer concentration is high, as occurs early in fibril assembly in the *in vitro* system (see refs. 18–20 and 24), the exponential terms in Eq. 3 are very close to zero and the growth of the fibril is primarily by elongation (k_1) to give highly tapered tips. When monomer concentration is low, as occurs later in the *in vitro* system, the exponential terms are near unity and growth occurs mainly by strand initiation (k_2). Therefore, the tips are more blunt.

DISCUSSION

Cylindrical lattice and spiral lattice models were originally suggested by Ramachandran (2) as a structure for 20-nm collagen fibrils. More recently, Galloway (10) suggested a helical model for large collagen fibrils. All three of the models were based on x-ray diffraction data, but none has been examined in detail since their original presentations. The helical model presented here was developed to account for the geometric constraints imposed on the process of collagen self-assembly that arose from recent observations (22, 24) on the growth of collagen fibrils *in vitro* from intermediate-sized fibrils that were ≈10 μm long to fibrils of several mm (Fig. 1). The essential features of the model are (*i*) a distinctive structural nucleus forms at each end of a growing fibril and growth of the fibril then proceeds by propagation of the two structural nuclei, (*ii*) the two structural nuclei have similar spiral or helical conformations, and (*iii*) assembly of each structural nucleus requires two kinds of specific binding steps defined as 3.4 *D*-period and 0.4 *D*-period overlaps, but propagation of the nucleus requires only the 3.4 *D*-period binding step.

The extensive previous studies on collagen fibrils do not provide a definitive conclusion about the structure, and some of the observations are contradictory (1–16, 25, 26). The model presented here is consistent with much of the available data, but there are some discrepancies. Crystalline domains observed by x-ray diffraction (4, 12) and electron microscopy (9) of native rat tail tendon are absent from the model, although the model contains some of the concentric organization suggested by the electron microscopic data (9). The model also shows some features that are reminiscent of the patterns recently observed in mineralizing turkey tendon (26), where rectangular plates of apatite crystals are aligned in the gap regions in parallel arrays across the full width of the

fibril. Such alignment of gap regions can be seen in some parts of the model (Fig. 4), although, because of the concentric organization, the pattern does not extend across the full fibril width. It may be that some molecular rearrangements occur during maturation or mineralization of collagen fibrils.

The model incorporates the principle of nucleation and propagation because extensive studies have established the principle for assembly of collagen fibrils (see refs. 1, 15, 16 and 19). The assumption that the structural nuclei have a spiral or helical conformation was based on the difficulty of explaining growth of fibrils from near-paraboloidal tips by alternative mechanisms (Table 1). Also, growth of collagen fibrils from a spiral or helical nucleus is similar to the process of screw dislocation and spiral growth frequently seen with crystals of small molecules (27). Many crystals in nature are generated by screw dislocation and are frequently observed to grow from spiral-growing surfaces that are flatter but similar to the helical tips proposed here (27).

The authors are grateful for extensive and important suggestions from Drs. Eric Eikenberry, Jürgen Engel, Hans Hofmann, David J. S. Hulmes, Karl E. Kadler, Elton Katz, and Eric Raab. The animations were performed on instruments at the Laboratory for Visiometries and Modeling (Center for Computer Aids for Industrial Productivity; Rutgers University). [The programs are available from D.S. and J.M. (1992), Simulated Rod Growth, CAIP Center Technology Report, Rutgers University. A videotape is available for a fee.] The work was supported in part by grants from the National Institutes of Health (AR 38188), the March of Dimes/Birth Defects Foundation, and the Lucille P. Markey Charitable Trust. D.J.P. is grateful for support from the von Humboldt Foundation that enabled him to carry out a major part of the work in the laboratory of Dr. Klaus Kühn (Max-Planck Institut für Biochemie, Martinsried bei München).

1. Piez, K. A. (1984) in *Extracellular Matrix Biochemistry*, eds. Piez, K. A. & Reddi, A. H. (Elsevier, New York), pp. 1–40.
2. Ramachandran, G. N. (1967) in *Treatise on Collagen: Chemistry of Collagen*, ed. Ramachandran, G. N. (Academic, New York), Vol. 1, pp. 148–150.
3. Smith, J. W. (1968) *Nature (London)* **219**, 157–158.
4. Hulmes, D. J. S. & Miller, A. (1979) *Nature (London)* **282**, 878–880.
5. Piez, K. A. & Trus, B. L. (1981) *Biosci. Rep.* **1**, 801–810.
6. Hulmes, D. J. S., Jesior, J.-C., Miller, A., Berthet-Colominas, C. & Wolff, C. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3567–3571.
7. Brodsky, B. & Eikenberry, E. F. (1982) *Methods Enzymol.* **82**, 127–173.
8. Woodhead-Galloway, J. (1984) in *Connective Tissue Matrix*, ed. Hukins, D. W. L. (Verlag Chemie, Basel), pp. 133–160.
9. Hulmes, D. J. S., Holmes, D. F. & Cummings, C. (1985) *J. Mol. Biol.* **184**, 473–477.
10. Galloway, J. (1985) in *Biology of Invertebrate and Lower Vertebrate Collagens*, eds. Bairoti, A. & Garrone, R. (Plenum, New York), pp. 73–82.
11. Chapman, J. (1989) *Biopolymers* **28**, 1367–1382.
12. Jones, E. Y. & Miller, A. (1991) *J. Mol. Biol.* **218**, 209–219.
13. Holmes, D. F. & Chapman, J. A. (1979) *Biochem. Biophys. Res. Commun.* **87**, 993–999.
14. Ward, N. P., Hulmes, D. J. S. & Chapman, J. A. (1986) *J. Mol. Biol.* **190**, 107–112.
15. Veis, A. & Payne, K. (1988) in *Collagen: Biochemistry*, ed. Nimni, M. E. (CRC, Boca Raton, FL), Vol. 1, pp. 113–138.
16. Gelman, R. A. & Piez, K. A. (1980) *J. Biol. Chem.* **155**, 8098–8102.
17. Birk, D. E., Zycband, E. I., Winkelmann, D. E. & Trelstad, R. L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4549–4553.
18. Miyahara, M., Njieha, F. K. & Prockop, D. J. (1982) *J. Biol. Chem.* **257**, 8442–8448.
19. Kadler, K. E., Hojima, Y. & Prockop, D. J. (1987) *J. Biol. Chem.* **262**, 15696–15701.
20. Kadler, K. E., Hulmes, D. J. S., Hojima, Y. & Prockop, D. J. (1990) *Ann. N.Y. Acad. Sci.* **580**, 214–224.
21. Hulmes, D. J. S., Kadler, K. E., Mould, A. P., Hojima, Y., Holmes, D. F., Cummings, C., Chapman, J. & Prockop, D. J. (1989) *J. Mol. Biol.* **210**, 337–345.
22. Kadler, K. E., Hojima, Y. & Prockop, D. J. (1990) *Biochem. J.* **268**, 339–343.
23. Romanic, A. M., Adachi, E., Kadler, K. E., Hojima, Y. & Prockop, D. J. (1991) *J. Biol. Chem.* **266**, 12703–12709.
24. Holmes, D. F., Chapman, J. A., Prockop, D. J. & Kadler, K. E. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9855–9859.
25. Farber, S., Garg, A. K., Birk, D. E. & Silver, F. H. (1986) *Int. J. Biol. Macromol.* **8**, 37–42.
26. Traub, W., Arad, T. & Weiner, S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9822–9826.
27. Boistelli, R. & Astier, J. P. (1988) *J. Cryst. Growth* **90**, 14–30.