

## Shining Light on Collagen: Expressing Collagen in Plants

Barbara Brodsky, PhD, and David L. Kaplan, PhD

Collagens are a remarkable group of proteins that are critical from a physiological perspective due to their diverse and versatile functions *in vivo*. However, collagens are challenging to generate *ex vivo* for biomaterials or regenerative medicine due to their complex processing and assembly into functional materials. Therefore, collagen availability remains a major unmet need for biomaterials, as relatively limited supplies of the protein in pure form are available mainly through harvesting bovine tissues. This animal source, subsequent to purification, remains associated with significant safety concerns due to the potential carryover of animal-derived diseases. Other more limited sources of animal collagens are also commercially available, as well as collagens generated in heterologous hosts; however, the challenge to these sources remains both economic and structural. The need for new safe sources of collagens remains high, with a significant potential impact in areas of medicine when considering the opportunity to mimic native collagen features. The articles in this issue of the journal focus on plant-derived collagens to address some of these needs. Progress toward plant production of collagens, the ability to self-assemble these recombinant proteins into higher-order structures, and the utility of these materials in various medical applications suggest an important path forward for the field.

**C**OLLAGEN PROVIDES THE fundamental basis for many biological functions in the human body. The relatively simple repetitive  $(\text{Gly-X}_{aa}\text{-Y}_{aa})_n$  sequence motif can be used as a building block for self-assembly to create structures in which mechanical function and chemical signaling are orchestrated to control physiological function. The remarkable diversity generated by collagen structures includes soft to stiff tissues; signaling cells to respond to homeostatic conditions or to more acute conditions that require rapid remodeling; and regeneration and repair due to tissue damage. The molecular and fibril structure of collagen are well defined, but the code that relates the collagen amino-acid sequence and structural features to specific biological responses is only beginning to be elucidated. The “Toolkit” of overlapping collagen peptides has allowed definition of the collagen sequences that are required for binding to receptors, other matrix proteins, and matrix metalloproteinases (MMPs).<sup>1</sup> Such information has created a firm starting point for visualizing the molecular interactions that set off signaling cascades. Tissue engineering seeks to mimic the structural and informational features of collagen systems. Basic information about collagen-specific signaling is important in experimental design of regenerative medicine, but the results from tissue engineering will also improve our understanding of how collagen directs function. A question such as “Will a biomedical material prepared from collagen type I versus type II improve

tissue regeneration of bone or cartilage?” may only get answered by a combination of basic science and tissue engineering strategies.

Bovine type I collagen provides the majority of current collagen protein used by the biomedical community. Due to the challenges presented with current sourcing of collagens and the lack of ability to recapitulate the important structural hierarchy, chemical modifications of reprocessed collagens are often pursued to help stabilize the material when used for biomaterials in tissue repairs or related medical needs. While such methods work well, including dehydrothermal treatments, carbodiimide cross-linking chemistry, glutaraldehyde cross-linking and related approaches, the chemistry utilized invariably changes the normal signaling achieved by collagens due to changes in structure, epitope displays, and different peptide fragments released on degradation of the materials. Thus, improved sources of recombinant human collagens that would self-assemble into the correct structural hierarchy and avoid the need for post-processing cross-linking chemistry would transform the field. Both basic science and tissue engineering approaches would benefit from the availability of new collagen materials in which collagen sequences could be varied to make basic structure–function correlations and to provide a means of optimizing control of material properties, signaling, and degradation. Alternative sources of collagens that are being pursued include synthetic

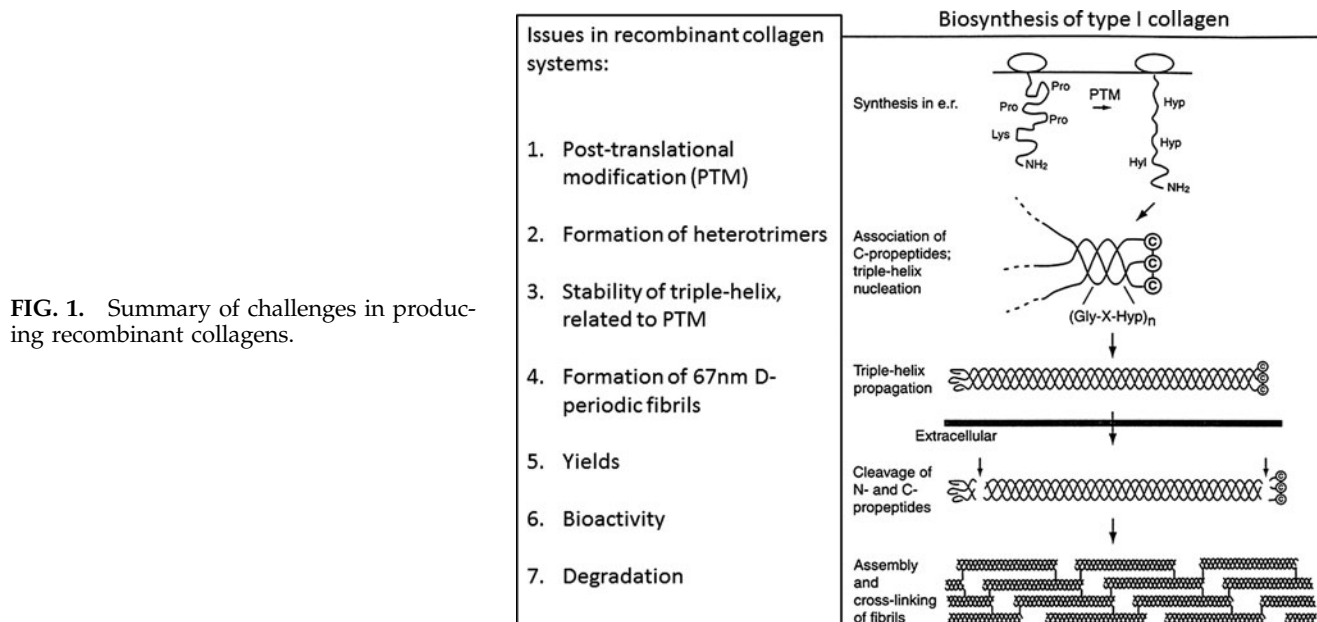
peptides and recombinant collagens. Synthetic peptides have been successfully used for fundamental studies and to characterize some elements of higher-order structures.<sup>2-4</sup> However, costs for peptide synthesis and the lack of robust mechanical properties due to the low-molecular-weight building blocks limits the utility of these systems for other than fundamental inquiries into structure function, including cell signaling. There is much interest in suitable host systems for scale up and expression of purified recombinant human collagens that are biologically active. Heterologous expression of human collagens has been demonstrated in mammalian cell lines, baculovirus, and a number of prokaryotes.<sup>5</sup> Such recombinant collagens could circumvent concerns with allergenicity, contamination, and quality control in current bovine type I collagen sources.

The transgenic plant approach described in this issue represents a useful addition to other recombinant collagens that are being pursued in a number of laboratories (see Ruggiero and Koch<sup>5</sup> for review), including the longstanding yeast-derived recombinant human collagens from the pioneering work of Kivirikko.<sup>6</sup> Significant advances have been made in other plant systems as well as tobacco, including barley and maize, where the specific proline hydroxylation sites were determined.<sup>7</sup> Although most molecular biology efforts have focused on expression of human collagen, a complementary approach has been to express collagen-like proteins from bacteria in *Escherichia coli*, as these proteins have a stability that is similar to human type I collagen in the absence of post-translational modifications.<sup>8,9</sup> The high yield and ability to include biologically active sequences, along with favorable biocompatibility properties, makes these collagen-like proteins attractive candidates for biomaterial applications.<sup>10</sup> Recently, the ability to include post-translational modifications in such bacterial systems has also been developed.<sup>11</sup>

Collagens have a number of unique features that should be addressed in any recombinant system (Fig. 1). The post-translational modification of Pro to hydroxyproline (Hyp) is required for triple-helix stability. Mammalian cell lines

produce hydroxylated collagen, but cannot be scaled up sufficiently to meet the needs of tissue engineering. The hydroxylation problem has largely been met by incorporating the two genes for prolyl hydroxylase within the system and by optimizing hydroxylation. For instance, Merle *et al.*<sup>12</sup> showed that incorporation of the prolyl hydroxylase genes within tobacco increased the stability of the recombinant human collagen produced. Human collagens also require hydroxylation of specific Lys residues, which is necessary for cross-linking in the fibrils. The Shosayev lab has incorporated the gene for an enzyme that is capable of hydroxylating and glycosylating Lys, as well as the prolyl hydroxylase genes,<sup>13</sup> in an attempt to better model the human system. Type I collagen is a heterotrimer, composed of two  $\alpha 1$  chains and one  $\alpha 2$  chain. Although initial studies were on homotrimers, both chains have been inserted in yeast, tobacco, and other systems.<sup>14,15</sup> In the well-characterized yeast system, it appears that both homotrimers and heterotrimers were formed when both chains were expressed in the system.<sup>13</sup> One challenge for recombinant collagens is their ability to form periodic fibrils that mimic type I collagen fibrils, as the fibril structure is considered important for mechanical properties, degradation, and perhaps for cell signaling. Recombinant type I and type III collagens in yeast yielded fibrils with a 67nm axial periodicity, which is very promising for tissue engineering applications. Recombinant human collagen expressed in tobacco formed fibrils, but they did not show the characteristic axial periodicity.<sup>13</sup>

The readily available extracted collagen materials have already been successfully applied in a wide range of drug delivery and biomaterials applications.<sup>16</sup> The utility of recombinant collagens, such as those derived from plants, will have to be demonstrated in their effectiveness as scaffolds or for wound healing; so, the articles in this volume are a welcome contribution in this direction. New ways to bioengineer and understand these systems, as well as to derive new commercial materials, will help the fields of biomaterials and regenerative medicine in many ways. The examples presented



**FIG. 1.** Summary of challenges in producing recombinant collagens.

in this issue provide a glimpse into what is possible. With the advances in recombinant systems and the increasing range of research tools available in bioinformatics, genetics, and biophysics, opportunities to pursue new sources of collagens are being realized. The way in which the chemistry, structure, and morphology of the collagen determines human tissue structure and function is becoming better defined, and this information will form the basis for developing new commercial materials in the biomaterial and regenerative medicine fields. A vision for the cropping of plants to use sunlight as the energy source to generate viable human medical materials is very appealing, and the development of successful large-scale production systems for collagen material may usher in a new era of exciting applications.

## Conclusions

Collagen-omics represents the underlying instruction set for extracellular matrix (ECM) structure and function. Virtually all physiological functions in our body, from the molecular to macroscopic level, originate from the underlying collagen chemistry encoded by the genetic templates in all our cells. Our ability to dissect the rules by which collagen chemistry controls cell function, tissue remodeling, mechanical properties, and inflammation should propel the field into a new realm of major impact in medical materials. To achieve such an impact, we should overcome the current constraints with available collagen materials. The articles in this issue offer directions toward this goal, where plant-derived collagens offer a route to new supplies of human collagens. With broadened approaches such as those described, we can envision options to generate the multitude of different collagen types in high yields, and with high purity, and without bioburdens, materials that will self-assemble into the required structural complexity to mimic what happens *in vivo*, and that will have utility in biomedical products. This is the beginning, but seeing the tools to achieve these goals emerge, we suggest not only a major impact in the medical community, but also major insights into collagen-omics to inform the next generation of fundamental scientists and engineers in the areas of matrix biology, diseases associated with collagen mutations and the ECM, and material science and engineering in general.

## Acknowledgment

The authors thank the National Institutes of Health [R01 EB011620, U01 EB014976] for support of studies on collagen.

## Disclosure Statement

No competing financial interests exist.

## References

- Farndale, R.W., Lisman, T., Bihan, D., Hamaia, S., Smerling, C.S., Pugh, N., Konitsiotis, A., Leiting, B., de Groot, P.G., Jarvis, G.E., and Raynal, N. Cell-collagen interactions: the use of peptide Toolkits to investigate collagen-receptor interactions. *Biochem Soc Trans* **36**, 241, 2008.
- Brodsky, B., and Persikov, A.V. Molecular structure of the collagen triple helix. *Adv Protein Chem* **70**, 301, 2005.
- Kar, K., Amin, P., Bryan, M.A., Persikov, A.V., Mohs, A., Wang, Y.H., and Brodsky, B. Self-association of collagen triple helical peptides into higher order structures. *J Biol Chem* **281**, 33283, 2006.
- Kotch, F.W., and Raines, R.T. Self-assembly of synthetic collagen triple helices. *Proc Natl Acad Sci U S A* **103**, 3028, 2006.
- Ruggiero, F., and Koch, M. Making recombinant extracellular matrix proteins. *Methods* **45**, 75, 2008.
- Nokelainen, M., Tu, H., Vuorela, A., Notbohm, H., Kivirikko, K.I., and Myllyharju, J. High-level production of human type I collagen in the yeast *Pichia pastoris*. *Yeast* **18**, 797, 2001.
- Xu, X., Gan, Q., Clough, R.C., Pappu, K.M., Howard, J.A., Baez, J.A., and Wang, K. Hydroxylation of recombinant human collagen type I alpha 1 in transgenic maize co-expressed with a recombinant human prolyl 4-hydroxylase. *BMC Biotechnol* **11**, 69, 2011.
- Xu, Y., Keene, D.R., Bujnicki, J.M., Höök, M., and Lukomski, S. Streptococcal Scl1 and Scl2 proteins form collagen-like triple helices. *J Biol Chem* **277**, 27312, 2002.
- Mohs, A., Silva, T., Yoshida, T., Amin, R., Lukomski, S., Inouye, M., and Brodsky, B. Mechanism of stabilization of a bacterial collagen triple helix in the absence of hydroxyproline. *J Biol Chem* **282**, 29757, 2007.
- Peng, Y.Y., Yoshizumi, A., Danon, S.J., Glattauer, V., Prokopenko, O., Mirochnitchenko, O., Yu, Z., Inouye, M., Werkmeister, J.A., Brodsky, B., and Ramshaw, J.A. A *Streptococcus pyogenes* derived collagen-like protein as a non-cytotoxic and non-immunogenic cross-linkable biomaterial. *Biomaterials* **31**, 2755, 2010.
- Pinkas, D.M., Ding, S., Raines, R.T., and Barron, A.E. Tunable, post-translational hydroxylation of collagen domains in *Escherichia coli*. *ACS Chem Biol* **6**, 320, 2011.
- Merle, C., Perret, S., Lacour, T., Jonval, V., Hudaverdian, S., Garrone, R., Ruggiero, F., and Theisen, M. Hydroxylated human homotrimeric collagen I in *Agrobacterium tumefaciens*-mediated transient expression and in transgenic tobacco plant. *FEBS Lett* **515**, 114, 2002.
- Stein, H., Wilensky, M., Tsafir, Y., Rosenthal, M., Amir, R., Avraham, T., Ofir, K., Dgany, O., Yayon, A., and Shoseyov, O. Production of bioactive, post-translationally modified, heterotrimeric, human recombinant type-I collagen in transgenic tobacco. *Biomacromolecules* **10**, 2640, 2009.
- Baez, J., Olsen, D., and Polarek, J.W. Recombinant microbial systems for the production of human collagen and gelatin. *Appl Microbiol Biotechnol* **69**, 245, 2005.
- Olsen, D.R., Leigh, S.D., Chang, R., McMullin, H., Ong, W., Tai, E., Chisholm, G., Birk, D.E., Berg, R.A., Hitzeman, R.A., and Toman, P.D. Production of human type I collagen in yeast reveals unexpected new insights into the molecular assembly of collagen trimers. *J Biol Chem* **276**, 24038, 2001.
- Browne, S., Zeugolis, D.I., and Pandit, A. 2013 Jan 21. [Epub ahead of Print]. *Tissue Eng Part A*.

Address correspondence to:

David L. Kaplan, PhD

Department of Biomedical Engineering

Tufts University

4 Colby Street

Medford, MA 02155

E-mail: david.kaplan@tufts.edu

Received: February 25, 2013

Accepted: February 26, 2013

Online Publication Date: April 30, 2013