

Production of recombinant collagen: state of the art and challenges

ISSN 2398-6182
 Received on 24th February 2017
 Revised on 30th March 2017
 Accepted on 24th April 2017
 doi: 10.1049/enb.2017.0003
 www.ietdl.org

Tianyi Wang¹, Jiewei Lew¹, Jayaraman Premkumar², Chueh Loo Poh², May Win Naing¹ ✉

¹Bio-Manufacturing Programme, Singapore Institute of Manufacturing Technology, Singapore

²Department of Biomedical Engineering, National University of Singapore, Singapore

✉ E-mail: winnaingm@SIMTech.a-star.edu.sg

Abstract: Collagen, which is often used in healthcare materials and biomedical research, is largely extracted from animal sources. Recombinant human collagen has the potential to be a promising alternative to animal collagen which has many shortcomings, including immunogenicity and lack of biocompatibility. Currently, recombinant human collagen has been expressed in both eukaryotic and prokaryotic hosts with varying degrees of success. One issue with recombinant collagen across all hosts is the inability to achieve full length collagen with native amounts of post-translational modifications, prompting much exciting research in this direction. There has also been much effort in improving yield and biomimicry of recombinant collagen. This review discusses collagen structure and current methods for extracting animal collagen, before introducing current research in synthesising recombinant human collagen in various hosts, and finally highlighting challenges in the field.

1 Introduction

Collagen is an important biomaterial that has many medical applications, ranging from fillers for cosmetic use, carriers in drug delivery systems [1], surgical suture [2], to scaffolds in tissue engineering [3, 4]. In particular, collagen exhibit great flexibility as a biomaterial which can be modified by scientists into collagen sponges for various biomedical applications, including neuronal regeneration [5], bone regeneration [6, 7] and wound healing [8]. Collagen can also form injectable hydrogels [9, 10] or serve as delivery vehicles [1, 11, 12] due to its innate ability to gel with temperature change, and is amenable to chemical modifications. Despite their versatility, collagens that are currently in use by the medical field are mostly animal derived, which may potentially cause immunogenic reactions and suffer from unwanted batch-to-batch variabilities [13]. Therefore, given the high demand for collagen, there is a need to develop novel methods to bio-manufacture large quantities of recombinant collagens in a consistent and efficient manner.

Synthetic biology is a promising strategy to address this need. There has been rapid development in synthetic biology, the engineering of biology, over the past decade. By applying engineering principles, synthetic biology aims to design and build biological-based parts, novel devices and systems as well as to redesign natural biological systems for a variety of biotechnological applications [14]. The emergence and development of synthetic biology has created an opportunity to engineer biological systems for large-scale production of useful but difficult to produce recombinant proteins such as human collagen. The purpose of this review is to discuss current methods of engineering recombinant collagen and highlight challenges to achieve large-scale production of high-quality recombinant collagen using a synthetic biology approach. The review first covers the understanding of collagen structure and its biosynthesis in the next section.

1.1 Collagen structure and post-translational modifications

Collagen forms the largest component of extra-cellular matrix and is also responsible for various signalling and structural roles in the

human body [15]. All human collagens contain repeating Gly-X-Y triplets, in which X and Y represent any amino acids but most commonly proline and hydroxyproline, respectively [16]. Changes in the type of amino acids present in X and Y positions have profound effects on the structure and function of collagen. Furthermore, procollagen undergoes several post-translational modifications before it matures into tropocollagen that can be crosslinked into collagen fibrils [16]. These post-translational modifications are essential in ensuring that the three single-stranded procollagen molecules associate with each other to form either heterotrimeric or homotrimeric tropocollagen [16, 17]. Only thermally stable tropocollagen can associate with each other to form collagen fibrils and finally collagen fibres. In particular, the hydroxylation of prolines to hydroxyproline is critical in ensuring the thermal stability of triple helical procollagen [18], without which many diseases, most prominently scurvy, occurs [16, 19, 20]. The hydroxylation of lysine to hydroxylysine is necessary for the condensation of collagen molecules to form a strong fibril that can withstand tension [19, 21, 22]. Hydroxylysine also serve as sites for glycosylation, without which, collagen fibrils would not form. Hydroxylation of lysines and proline is mediated by lysyl hydroxylase and prolyl 4-hydroxylase, respectively, and requires ascorbate as a co-factor, without which collagen peptides are unable to form fibres and will be degraded inside the cell [23–25]. Human collagens are ~42–54% hydroxylated [26]. As collagen structure and post-translational modifications have been extensively reviewed, the authors recommend interested readers to refer to [16, 23] for more comprehensive reviews in this topic.

Together, these subtle changes have led to the presence of 28 types of collagen, but the majority of collagens belong to types I, II and III [4, 19]. Type I collagen, the most common heterotrimeric collagen, is mainly present in our bones, skin, fibrous cartilage and tendon. These type I collagen fibrils have high tensile strength and are able to withstand more forces as compared with the same mass of steel, thus maintaining the integrity of our tissue architecture [4]. Type II collagen, a homotrimer, is present in large amounts within the hyaline cartilage and gives cartilage its shock absorbing properties [27]. Type III collagen, a homotrimer, is also present in human muscles, vasculature and skin, but in lesser amounts as compared with type I collagen [19].

1.2 Current methods for collagen extraction

As collagen denatures to gelatin at high temperatures, extraction procedures must be carried out at 4°C, either on ice or in the cold room. General procedures for collagen extraction involve first obtaining tissue specific to the type of collagen needed. These tissues are then cleaned by rinsing in water to remove all traces of dirt. Subsequently, fats and connective tissues are trimmed and hair removed. Next, the tissues are cut and placed into dilute sodium acetate solution to remove all soluble non-collagenous proteins and polysaccharides. Supernatant is discarded and residual sodium acetate rinsed out from the tissues. The sample is then freeze-dried and subsequently re-dissolved in acetic acid, which will solubilise only collagen. As the supernatant now contains a crude collagen solution, dialysis has to be performed for further purifications [28, 29]. Depending on the starting material, the resulting collagen solution obtained would be a mixture of different types of collagens. Should a specific type of collagen be desired, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) or chromatography may be performed to isolate this preferred collagen type. As the whole extraction process can take up to a month [30], optimisation is necessary to reduce time spent. To achieve a shorter extraction procedure, Pacak *et al.* [29] shortened the time needed for each step significantly. They also replaced the lengthy dialysis process with centrifugation in a 100 k molecular weight cut-off (MWCO) centrifugal filter tube, saving several days. Since collagen extraction protocols are based on collagen's solubility in various solvents, this protocol can be applied for all types of collagen.

To obtain different types of collagen, different starting materials must be used. For instance, if type I collagen is needed, then animal skins and tendons may be used, whereas if type II collagen is needed, animal hyaline cartilage would be preferred. As the starting material will determine the final types of collagen obtained, it is more difficult to obtain less-common collagen (e.g. type V or type VII collagen) due to the low presence of these collagens in tissues. Moreover, purified collagen is expensive due to tedious extraction processes, as well as intensive purification procedures [13, 29–31]. This makes collagen prohibitively expensive if clinically relevant amounts are needed. Furthermore, collagen extraction methods vary among different companies, resulting in collagen with significant differences in mechanical and chemical properties, thereby leading to unwanted batch-to-batch variations [13].

1.3 Shortcomings of extracted collagen

Due to collagen's relative abundance and importance in biological function, it has been commonly used to form biomimetic scaffolds to support cell growth for tissue engineering [32, 33]. Collagens used in tissue engineering or for in vitro toxicity testing are typically extracted from the tendons, skin and other connective tissues of unwanted animal by-products. Common animal sources of collagen include marine, bovine and porcine [34]. Unfortunately, bovine sources harbour the risk of carrying bovine spongiform encephalopathy while porcine sources may be unsuitable due to certain religious beliefs [34]. Furthermore, animal sources of collagen may have immunoglobulin E (IgE) mediated immunogenicity issues for a small proportion of people, leading to granuloma or localised inflammation [35, 36]. Animal derived collagen matrices may also result in poor predictions for in vitro toxicity testing. Although the use of human collagen matrices may be able to mitigate these problems, human collagen is rare, making it non-economical for common use. Finally, less-common collagen (e.g. type V or type VII collagen) are very expensive to obtain due to their low presence in animal tissues. Due to these limitations in extracted collagen, a promising alternative strategy would be to use synthetic biology to manufacture recombinant human collagen in a scalable and reproducible manner and this is discussed in the following sections.

2 Recombinant human collagen production using synthetic biology

2.1 Prokaryotic organisms

2.1.1 *Escherichia coli*: *Escherichia coli* (*E. coli*) is commonly used as a host in synthetic biology as it has been fully sequenced [37], and well-studied, allowing for convenient and predictable genetic manipulation. Moreover, *E. coli* exhibits fast growth kinetics, and can be grown in high-density cultures and has also demonstrated scalability in which it has been used to produce insulin on an industrial scale; *E. coli* is currently an industrial workhorse. While prokaryotic organisms produce their own native collagen like motifs containing the Gly-X-Y repeat sequence [38], this bacterial collagen differs from native human collagen in terms of amino acid type and distribution, and in turn structure. Strikingly, due to the absence of post-translational modification mechanisms, hydroxyproline is absent from prokaryotic collagen. Therefore, as a host for recombinant human collagen production, *E. coli* is limited by its lack of endogenous post-translational modifications such as hydroxylation. There has been a nascent effort to overcome this by Rutschmann *et al.* [39], using hydroxylases from mimivirus, but the collagen produced is still far from being similar to native human collagen in a multitude of parameters, including hydroxylation, thermal stability and structure. Expressing a fully functional prolyl hydroxylase in *E. coli* has also been very challenging [40]. Therefore, most work on *E. coli* produced collagen involves only small functional sections of human collagen, or is on human-like collagen [41–44]. Relatively little work on recombinant human collagen production has been performed in *E. coli* as compared with eukaryotic organisms, suggesting challenges, and hence opportunities for research in this area.

2.2 Eukaryotic organisms

2.2.1 *Yeast*: Yeast is often used in synthetic biology due to its ease of genetic manipulation and relatively low cost of maintenance. To this end, many have used yeast, most commonly *Saccharomyces cerevisiae* (*S. cerevisiae*) or *Pichia pastoris* (*P. pastoris*) as chassis for the production of recombinant human collagen. Since yeast lacks native prolyl and lysyl hydroxylases and does not perform post-translational hydroxylation, collagen produced by yeast will not be hydroxylated and will be thermally unstable at 37°C. To overcome this issue, human prolyl hydroxylases are commonly co-expressed with human collagen such that hydroxylase enzyme will be present to hydroxylate proline. Vuorela *et al.* further discovered that collagen expression increased the number of stable and active prolyl hydroxylases, resulting in the production of stable triple helical type III collagen with near native hydroxylation percentages in *P. pastoris* [45, 46]. However, hydroxylation of lysine was absent and yield (~15 mg/l) was low. Further studies by Myllyharju *et al.* [47] was able to show improvements in yield of triple helical collagen of up to 0.6 g/l by culturing *P. pastoris* in a bioreactor with a constant supply of oxygen.

Heterotrimeric collagens are inherently more difficult to express as compared with homotrimeric collagens because additional genes are needed. However, heterotrimers such as type I collagen have wider applications as compared with homotrimers such as type III collagen, making it important to develop means to synthesise heterotrimers. Nokelainen *et al.* [45] optimised their cloning strategies to express both subunits of prolyl hydroxylases as well as the polypeptides that make up type I collagen, procollagen1 α 1 and procollagen1 α 2, in *P. pastoris*. They successfully obtained type I collagen's characteristic 2:1 chain ratio of procollagen1 α 1 and procollagen1 α 2. Similar results have been observed by this group in *S. cerevisiae* [48].

Indeed, *S. cerevisiae* has also been used by many to produce recombinant collagen [48–50]. Chan *et al.* [50] engineered *S. cerevisiae* containing prolyl hydroxylases to produce type III



collagen, but were only able to obtain 0.5% proline hydroxylation, which is low compared with native hydroxylation levels in humans. Vaughan *et al.*, however, were able to achieve higher hydroxylation in human type III collagen by *S. cerevisiae* by using a different cloning strategy, though hydroxylation levels are still below that of native human collagen [51]. While most papers focused on achieving native levels of proline hydroxylation, Olsen *et al.* [49] reported that N- and C-propeptide regions are not needed for collagen to fold into a triple helical conformation, and their removal increased collagen yield, suggesting that yield can be improved through deliberate designs of the genetic construct.

2.2.2 Plants: Plant biotechnology was developed to deal with the sustainability of food supply due to the increasing world population, and also to tackle the effects of environmental changes on agricultural yield. The development of better genetic engineering tools coupled with a growing demand for medical materials have driven research in the direction of transgenic plants that produce biomaterials. Much research on plant-based recombinant human collagen has been performed by the Shoseyov Laboratory, with a focus on type I collagen being produced by tobacco plant [52, 53]. Although tobacco plants possess prolyl hydroxylase for hydroxylating proline residues, plant propyl hydroxylase is insufficient for hydroxylating human collagen to that of native human levels [54, 55], making it necessary to engineer plants to produce human prolyl hydroxylases instead [56]. In addition, Shoseyov Laboratory also expressed lysyl hydroxylase in tobacco plants for lysine hydroxylation. While *E. coli* transformation efficiency decreases with increasing plasmid sizes, agrobacterium mediated plant transformation allows much larger genetic sequences to be transfected into plant cells [57], making it possible to express full length procollagen1 α 1, procollagen1 α 2 along with both subunits of prolyl hydroxylase and lysyl hydroxylase within the same cell [52]. Signal peptides were added to direct the accumulation of collagen into subcellular compartments of the tobacco cell, such as plant vacuole [58]. Stein *et al.* from the Shoseyov Laboratory were able to demonstrate hydroxylation of both proline and lysine to levels similar to that of human type I collagen. Furthermore, they were able to demonstrate that type I collagen produced by tobacco was able to produce fibrils [52, 59]. In addition, the alignment of collagen fibrils is necessary to improve collagen's mechanical properties. To achieve aligned collagen fibrils akin to that of the native tendon, shear forces had to be imparted through extrusion during the liquid crystalline phase of collagen formation. This marked the first demonstration of fibril formation in full length recombinant human collagen [53]. Research findings from the Shoseyov Laboratory cumulated into a company – Collplant Ltd. Currently, Collplant is administering two human trials with plant derived recombinant human collagen for wound healing and tissue repair [60, 61].

Besides tobacco, researchers have also engineered maize to produce hydroxylated recombinant human procollagen1 α 1 peptide, with comparable hydroxyproline content as native human collagen1 α 1 [62]. While more work needs to be done before recombinant collagen produced in maize can reach the same level of technological maturity as that produced in tobacco, this study has nevertheless demonstrated the potential of maize as an alternative host for producing recombinant collagen.

As compared with other prokaryotic and eukaryotic cells, in which scaling up involves complex bioreactor set up and expensive food source, a great advantage of plant derived collagen is the scalability plant provides; scaling up is as simple as growing large amounts of plants. Alternatively, it is also possible to grow plants in the form of single cells in a bioreactor or as callus if desired.

2.2.3 Mammalian cells: Human collagen is naturally produced in mammals, thus mammalian cells have the advantage of having the necessary post-translational modifications for producing hydroxylated collagen, an important parameter which is not present in bacterial or yeast cells. Various types of mammalian cell lines have been used for the production of recombinant human

collagen, including human sarcoma cell line HT 1080 [63–65], Chinese hamster ovary cells [66], HeLa cells [67] and human embryonic kidney cells 293-EBNA [68, 69]. Papers reporting successful recombinant collagen expression using non-mammalian host systems largely focus on reporting methods to achieve native amounts of hydroxylation in collagen. This is usually done by creatively cloning hydroxylases into non-mammalian hosts and then performing tedious optimisation steps to ensure high biological activities of these enzymes [39, 45, 70]. Conversely, cloning strategies for the production of human collagen in mammalian cells do not involve prolyl or lysyl hydroxylases, as these enzymes are inherently present in mammalian cells as part of the post-translational modification pathway. In one of the earlier demonstrations of collagen synthesis, only full length α 1 chain of collagen V was expressed in 293-EBNA cells but yet the authors were able to obtain some amounts hydroxylation on both proline and lysine, yielding a thermally stable peptide, by adding ascorbate, a potent co-factor of hydroxylation [68], demonstrating that it is unnecessary to engineer in additional hydroxylases. Similarly, hydroxylation was observed in partial sequence of type XII collagen [67] and full length type VII collagen [71] only when ascorbate was present. In these instances, cloning of hydroxylases was also unnecessary. Furthermore, glycosylation of hydroxylated lysine, which is important in the formation of triple helical collagen, and largely unreported in non-mammalian hosts, also occurs natively in mammalian expression systems [64, 68]. Thus, the use of mammalian cells can simplify the extent of genetic manipulation needed. This means that smaller constructs are required, putting less stress on the cell thereby increasing cloning efficiency. Furthermore, glycosylation, which has been largely undetermined and unreported in non-mammalian expression systems, is being observed in collagens synthesised by mammalian cells.

Although mammalian cells have the requisite post-translational modification machinery as compared with other hosts, the yield of recombinant human collagen was significantly lower compared with that of yeast, with a maximum expression of 80 mg/l [43]. A common way to improve yield is to replace native promoters by stronger promoters. However, rapid production of large amounts of collagen meant that prolyl hydroxylases were unable to fully hydroxylate collagen [65], demonstrating that improving yield is a complex process requiring more than just a simple over-expression of proteins. Other roadblocks of large-scale production of recombinant human collagens in mammalian cells include costly culture, and the complexity of mammalian cell metabolism means a long bench-to-market time. Scaling up can also be challenging for adherent mammalian cell cultures [72].

2.2.4 Insect cells: The use of insect cells, such as Sf 9 or High Five cell lines for recombinant protein production has been increasing over the years. Transfection of insect cells is very well established and commonly performed via baculovirus. Insect cells can be cultured in high-density suspension cultures and therefore have the advantage of volumetric scalability in a bioreactor. Furthermore, expression of large amounts of recombinant proteins can be obtained at a lower cost in insect cells as compared with mammalian cells. In one of the earlier studies on collagen production in insect cells, Tomita *et al.* [73] made use of Sf 9's native prolyl hydroxylases in the production of recombinant human procollagen1 α 1 chains as well as heterotrimeric type I collagen [74]. However, results showed that while insect cells contain prolyl hydroxylases that can hydroxylate proline residues, this hydroxylation was insufficient. An interesting point to note is that despite the insufficient hydroxylation on proline residues, procollagen peptides were still able to associate to form heterotrimers, suggesting that the formation of trimeric collagen may still be possible with lower levels of prolyl hydroxylation [74]. In contrast, Li *et al.* [75] reported that inter-chain disulphate bond formation between collagen peptides was only present when type XXI minicollagen was co-expressed with prolyl hydroxylase in *Drosophila* S2 cells. These differences could be due to the different amounts of native prolyl hydroxylases in Sf 9 and



Drosophila S2 cells, and types of collagen expressed. Later studies on recombinant human production in insect cells recognised the need to clone prolyl hydroxylases into insect cells to ensure sufficient proline hydroxylation, and cloning of prolyl hydroxylases became more common [75–79].

Lysine hydroxylation, which is often left out in recombinant collagen papers, is present inherently in High Five cells. To demonstrate this point, Lamberg *et al.* found that hydroxylysine content in procollagen3 α 1 in the presence of human prolyl hydroxylase but absence of lysyl hydroxylase was 60% that of native mammalian type III collagen, suggesting that insect cells must have a considerable level of lysyl hydroxylase activity. Similarly, Nokelainen *et al.* reported increased hydroxylysine levels when recombinant collagen was co-expressed with lysyl hydroxylase [79]. As glycolysis occurs primarily on hydroxylysine, Nokelainen *et al.* also observed increased glycosylation along with increased lysine hydroxylation. However, it remains unknown if insect glycosylation post-translational modifications are sufficient for glycosylation to the levels similar to that in native mammalian collagen.

3 Challenges

While many groups have tried synthesising recombinant human collagens using different hosts, insufficient post-translational modifications and low yield remain as unresolved issues. Post-translational modifications, especially hydroxylation remains challenging in prokaryotic systems, which lack native post-translational modification mechanisms. Even in eukaryotic systems, native hydroxylation modules are unable to adequately hydroxylate collagen molecules, making it necessary to incorporate non-native hydroxylases when engineering hosts for collagen production. As the introduction of non-native sequences into the genome of the host takes a metabolic toll on the host, yield will be compromised. Thus, another important parameter in engineering collagen producing organisms is to ensure that yield is high, so as to facilitate future scaled up production.

3.1 Challenges in collagen hydroxylation

One challenge in synthesising recombinant human collagen is ensuring that the extensive post-translational modifications collagens undergo during biosynthesis happen in a non-native host. These post-translational modifications, including glycosylation and hydroxylation affect the integrity of collagen fibrils and improperly modified procollagen represents diseased states. As bacteria and yeast, which can produce recombinant proteins in a more cost efficient manner, do not have the necessary post-translational modification mechanisms, the production of affordable recombinant human collagen remains challenging. To produce hydroxyproline containing collagen using *E. coli*, Buechter *et al.* [80] cultured *E. coli* in media containing both proline and hydroxyproline and demonstrated that the control of hydroxyproline incorporation can be tuned by varying the ratio of proline to hydroxyproline in media. Although this is a relatively simple way to achieve collagen hydroxylation, hydroxyproline can occur on either the X or Y position in the Gly-X-Y sequence of collagen with equal probability. Since the positioning of hydroxyproline has effects on collagen stability, with hydroxyproline on the Y position conferring greater stability to collagen, this method is not without its shortcomings [81, 82]. Recently, Pinkas *et al.* [70] have demonstrated the possibility of post-translational hydroxylation of the proline residue in *E. coli*. A critical barrier preventing hydroxylation in *E. coli* is the lack of ascorbate production, an essential co-factor in the catalysis of hydroxylation. To overcome this, Pinkas *et al.* engineered a synthetic strain of *E. coli* by overexpressing the D-arabinono-1,4-lactone oxidase gene borrowed from *S. cerevisiae*. This engineered *E. coli* strain contained a biosynthetic shunt that can produce ascorbate-like molecules in which human prolyl hydroxylase was able to utilise to hydroxylate

a short synthetic collagen sequence, consisting of (Pro-Pro-Gly)_n. However, the controllability of hydroxylation levels, and the distribution of hydroxylated residues in native human collagen polypeptides, which has a much longer sequence, remains unknown. Furthermore, the method described by Pinkas *et al.* has yet been proven to be applicable for lysine hydroxylation. Another method of overcoming the lack of hydroxylation in prokaryotes like *E. coli* is to utilise both prolyl and lysyl hydroxylases borrowed from other organisms, such as that from mimivirus isolated from giant aquatic viruses [83, 84]. Rutschmann *et al.* have demonstrated that mimivirus hydroxylases enabled higher hydroxylation levels as compared with other animal hydroxylases in *E. coli* and the use of it also led to native percentages and distribution of hydroxylation in recombinant type III collagen [84]. However, only partial type III collagen sequence was used in this paper, thus the effectiveness of using mimivirus hydroxylases to hydroxylate full sequence collagen remains unexplored. Hydroxylating proline has been the focus of many papers, rather than hydroxylating lysine. This may be because thermal stability, controlled by proline hydroxylation, is a major measurable parameter in determining collagen quality. To the best of our knowledge, no group has been able to produce full length recombinant human collagen with native amounts of post-translational modifications in any host organism and this remains an essential research question.

3.2 Challenges in optimising yield

An important criterion in the production of recombinant human collagen for medical use is yield; increasing yield in a cost-effective manner is necessary to make end products affordable. Currently, yield of recombinant collagen using both prokaryotic and eukaryotic organisms are low, ranging from micrograms to milligrams per litre of culture depending on the length and type of collagen produced, and also the host organism [85]. For instance, type III collagen production by *P. pastoris* in a fermenter was up to 0.5 g/l [45] whereas collagen production using transgenic mammals was much lower at 50–200 μ g/ml [86]. One way to improve yield is to perform metabolic flux analysis using metabolic models. Metabolic flux analysis has shown great promise in pinpointing the rate limiting steps in a batch culture such that the step can be optimised for improved yield. In some instances, optimising carbon/nitrogen molar ratios in the media can improve collagen yield secreted by engineered *E. coli* [87], in other instances, the addition of another energy source such as NADPH improved yield [88]. Another way to improve yield is by optimising cloning strategies. There have been various literature describing ways on how to improve yield in bacterial [88–90], yeast [91] and mammalian expression systems [92, 93], and these methods are applicable for the production of recombinant collagen. Another approach is to apply metabolic engineering of the host to boost production of protein which has highly repetitive structures and rich in glycine, e.g. silk protein [94]. Advances in bioinformatics, such as more advanced algorithms for codon optimisation, will also aid in obtaining a higher yield.

4 Conclusions and future perspectives

Currently, bio-production of recombinant collagens still suffer from low yield, high cost and the products are unable to achieve the same post-translational modifications seen in native collagens, making current recombinant collagens both expensive and non-biological. Despite their potential, these unresolved issues make recombinant collagen less attractive as compared with animal derived collagens, thus animal derived collagens are still the gold standard. Alternatively, cell-free expression systems might be an option for deriving large-scale expression and purification of human collagen. Recently, cell-free systems derived from cultured mammalian cells are gaining momentum due to the successful demonstration of glycosylation and key post-translational modifications [95, 96].



This platform of cell-free protein synthesis system has also been rapidly transforming from a cost-effective micro-scale synthesis to macro-scale up production system [97, 98]. Nevertheless, recent advances in DNA manipulation technologies such as CRISPR, in addition to cheaper DNA synthesis, will reduce the timeline and cost of research and development, translating to lowered prices and better biomimicry for recombinant collagen. More importantly, scalability should be engineered into the design of synthetic collagen constructs so that bio-manufacturing techniques can be applied to make economical large-scale manufacturing of recombinant collagen a possibility.

5 References

- An, B., Lin, Y.-S., Brodsky, B.: 'Collagen interactions: drug design and delivery', *Adv. Drug Deliv. Rev.*, 2016, **97**, pp. 69–84
- Chattopadhyay, S., Raines, R.T., Glick, G.D.: 'Review collagen-based biomaterials for wound healing', *Biopolym.*, 2014, **101**, (8), pp. 821–833
- Glowacki, J., Mizuno, S.: 'Collagen scaffolds for tissue engineering', *Biopolym.*, 2008, **89**, (5), pp. 338–344
- Cen, L., Liu, W., Cui, L., et al.: 'Collagen tissue engineering: development of novel biomaterials and applications', *Pediatr. Res.*, 2008, **63**, (5), pp. 492–496
- Onuma-Ukegawa, M., Bhatt, K., Hirai, T., et al.: 'Bone marrow stromal cells combined with a honeycomb collagen sponge facilitate neurite elongation in vitro and neural restoration in the hemisectioned rat spinal cord', *Cell Transplant.*, 2015, **24**, (7), pp. 1283–1297
- Kobayashi, E., Nakahara, T., Inoue, M., et al.: 'Experimental study on in situ tissue engineering of the temporomandibular joint disc using autologous bone marrow and collagen sponge scaffold', *J. Hard Tissue Biol.*, 2015, **24**, (2), pp. 211–218
- Huang, R.-L., Chen, G., Wang, W., et al.: 'Synergy between Il-6 and soluble Il-6 receptor enhances bone morphogenetic protein-2/absorbable collagen sponge-induced bone regeneration via regulation of Bmpria distribution and degradation', *Biomaterials*, 2015, **67**, pp. 308–322
- Bai, W.-F., Xu, W.-C., Zhu, H.-X., et al.: 'Efficacy of 50 Hz electromagnetic fields on human epidermal stem cell transplantation seeded in collagen sponge scaffolds for wound healing in a murine model', *Bioelectromagnetics*, 2017, **38**, (3), pp. 204–212
- Taguchi, T., Xu, L., Kobayashi, H., et al.: 'Encapsulation of chondrocytes in injectable alkali-treated collagen gels prepared using poly(ethylene glycol)-based 4-armed star polymer', *Biomaterials*, 2005, **26**, (11), pp. 1247–1252
- Liu, Y., Gan, L., Carlsson, D.J., et al.: 'A simple, cross-linked collagen tissue substitute for corneal implantation', *Invest. Ophthalmol. Vis. Sci.*, 2006, **47**, (5), pp. 1869–1875
- Ruszcak, Z., Friess, W.: 'Collagen as a carrier for on-site delivery of antibacterial drugs', *Adv. Drug Deliv. Rev.*, 2003, **55**, (12), pp. 1679–1698
- Kouris, N.A., Squirrell, J.M., Jung, J.P., et al.: 'A non-denatured, non-crosslinked collagen matrix to deliver stem cells to the heart', *Regen. Med.*, 2011, **6**, (5), pp. 569–582
- Schmidt, M.M., Domelles, R.C.P., Mello, R.O., et al.: 'Collagen extraction process', *Int. Food Res. J.*, 2016, **23**, (3), pp. 913–922
- Clarke, L.: 'A synthetic biology roadmap for the UK', UK Synthetic Biology Roadmap Coordination Group, (Technology Strategy Board, 2012)
- Alberts, B.: 'Molecular biology of the cell' (Garland Science, 2002, 4th edn.)
- Shoulders, M.D., Raines, R.T.: 'Collagen structure and stability', *Annu. Rev. Biochem.*, 2009, **78**, (1), pp. 929–958
- Vitagliano, L., Berisio, R., Mazzarella, L., et al.: 'Structural bases of collagen stabilization induced by proline hydroxylation', *Biopolym.*, 2001, **58**, (5), pp. 459–464
- Berg, R.A., Prockop, D.J.: 'The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple-helix of collagen', *Biochem. Biophys. Res. Commun.*, 1973, **52**, (1), pp. 115–120
- Lodish, H.F., Darnell, J.E.: 'Molecular cell biology' (W.H. Freeman, 2000, 4th edn.)
- Leger, D.: 'Scurvy: reemergence of nutritional deficiencies', *Can. Fam. Phys.*, 2008, **54**, (10), pp. 1403–1406
- Takaluoma, K., Hyry, M., Lantto, J., et al.: 'Tissue-specific changes in the hydroxylysine content and cross-links of collagens and alterations in fibril morphology in lysyl hydroxylase 1 knock-out mice', *J. Biol. Chem.*, 2006, **282**, (9), pp. 6588–6596
- Schegg, B., Hulsmeier, A.J., Rutschmann, C., et al.: 'Core glycosylation of collagen is initiated by two (1-O)galactosyltransferases', *Mol. Cell Biol.*, 2008, **29**, (4), pp. 943–952
- Yamauchi, M., Sricholpech, M.: 'Lysine post-translational modifications of collagen', *Essays Biochem.*, 2012, **52**, (3), pp. 113–133
- Pihlajaniemi, T., Myllylä, R., Kivirikko, K.I.: 'Prolyl 4-hydroxylase and its role in collagen synthesis', *J. Hepatol.*, 1991, **13**, (Suppl 3), pp. S2–S7
- Peterkofsky, B.: 'Ascorbate requirement for hydroxylation and secretion of procollagen: relationship to inhibition of collagen synthesis in scurvy', *Am. J. Clin. Nutr.*, 1991, **54**, (6 Suppl), pp. 1135S–1140S
- Chan, S.W., Greaves, J., Da Silva, N.A., et al.: 'Assaying proline hydroxylation in recombinant collagen variants by liquid chromatography-mass spectrometry', *BMC Biotechnol.*, 2012, **12**, (1), p. 51
- Sophia Fox, A.J., Bedi, A., Rodeo, S.A.: 'The basic science of articular cartilage: structure, composition, and function', *Sports Health*, 2009, **1**, (6), pp. 461–468
- Gallop, P.M., Seifter, S.: 'Preparation and properties of soluble collagens', 1963, **6**, pp. 635–641
- Pacak, C.A., Powers, J.M., Cowan, D.B.: 'Ultrapurification of collagen type I for tissue engineering applications', *Tissue Eng. C. Methods*, 2011, **17**, (9), pp. 879–885
- Rajan, N., Habermehl, J., Coté, M.-F., et al.: 'Preparation of ready-to-use, storable and reconstituted type I collagen from rat tail tendon for tissue engineering applications', *Nat. Protoc.*, 2007, **1**, (6), pp. 2753–2758
- Cliche, S., Amiot, J., Avezard, C., et al.: 'Extraction and characterization of collagen with or without telopeptides from chicken skin', *Poult. Sci.*, 2003, **82**, (3), pp. 503–509
- Wallace, D.: 'Collagen gel systems for sustained delivery and tissue engineering', *Adv. Drug Deliv. Rev.*, 2003, **55**, (12), pp. 1631–1649
- Lee, K.Y., Mooney, D.J.: 'Hydrogels for tissue engineering', *Chem. Rev.*, 2001, **101**, (7), pp. 1869–1880
- Gómez-Guillén, M.C., Giménez, B., López-Caballero, M.E., et al.: 'Functional and bioactive properties of collagen and gelatin from alternative sources: a review', *Food Hydrocolloids*, 2011, **25**, (8), pp. 1813–1827
- Hori, H., Hattori, S., Inouye, S., et al.: 'Analysis of the major epitope of the alpha2 chain of bovine type I collagen in children with bovine gelatin allergy', *J. Allergy Clin. Immunol.*, 2002, **110**, (4), pp. 652–657
- Mullins, R.J., Richards, C., Walker, T.: 'Antigenicity. anaphylactic risk for surgeons', *Aust. N Z J. Ophthalmol.*, 1996, **24**, (3), pp. 257–260
- Blattner, F.R., Plunkett, G. III, Bloch, C.A., et al.: 'The complete genome sequence of Escherichia coli K-12', *Science*, 1997, **277**, (5331), pp. 1453–1462
- Rasmussen, M., Jacobsson, M., Björck, L.: 'Genome-based identification and analysis of collagen-related structural motifs in bacterial and viral proteins', *J. Biol. Chem.*, 2003, **278**, (34), pp. 32313–32316
- Rutschmann, C., Baumann, S., Cabalzar, J., et al.: 'Recombinant expression of hydroxylated human collagen in Escherichia coli', *Appl. Microbiol. Biotechnol.*, 2014, **98**, (10), pp. 4445–4455
- Werkmeister, J.A., Ramshaw, J.A.M.: 'Recombinant protein scaffolds for tissue engineering', *Biomed. Mater.*, 2012, **7**, (1), p. 012002
- Xue, W.J., Fan, D.D., Shang, L., et al.: 'Production of biomass and recombinant human-like collagen in Escherichia coli processes with different Co2 pulses', *Biotechnol. Lett.*, 2009, **31**, (2), pp. 221–226
- Tang, Y., Yang, X., Hang, B., et al.: 'Efficient production of hydroxylated human-like collagen via the co-expression of three key genes in Escherichia coli origami (De3)', *Appl. Biochem. Biotechnol.*, 2016, **178**, (7), pp. 1458–1470
- Ruggiero, F., Koch, M.: 'Making recombinant extracellular matrix proteins', *Methods*, 2008, **45**, (1), pp. 75–85
- Delacoux, F., Fichard, A., Geourjon, C., et al.: 'Molecular features of the collagen V Heparin binding site', *J. Biol. Chem.*, 1998, **273**, (24), pp. 15069–15076
- Nokelainen, M., Tu, H., Vuorela, A., et al.: 'High-level production of human type I collagen in the yeast Pichia pastoris', *Yeast*, 2001, **18**, (9), pp. 797–806
- Vuorela, A.: 'Assembly of human prolyl 4-hydroxylase and type Iii collagen in the yeast Pichia pastoris: formation of a stable enzyme tetramer requires coexpression with collagen and assembly of a stable collagen requires coexpression with prolyl 4-hydroxylase', *EMBO J.*, 1997, **16**, (22), pp. 6702–6712
- Myllyharju, J., Nokelainen, M., Vuorela, A., et al.: 'Expression of recombinant human type I-iii collagens in the yeast Pichia pastoris', *Biochem. Soc. Trans.*, 2000, **28**, (4), pp. 353–357
- Toman, P.D., Chisholm, G., McMullin, H., et al.: 'Production of recombinant human type I procollagen trimers using a four-gene expression system in the yeast Saccharomyces cerevisiae', *J. Biol. Chem.*, 2000, **275**, (30), pp. 23303–23309
- Olsen, D.R., Leigh, S.D., Chang, R., et al.: 'Production of human type I collagen in yeast reveals unexpected new insights into the molecular assembly of collagen trimers', *J. Biol. Chem.*, 2001, **276**, (26), pp. 24038–24043
- Chan, S.W.P., Hung, S.-P., Raman, S.K., et al.: 'Recombinant human collagen and biomimetic variants using a De novo gene optimized for modular assembly', *Biomacromolecules*, 2010, **11**, (6), pp. 1460–1469
- Vaughan, P.R., Galanis, M., Richards, K.M., et al.: 'Production of recombinant hydroxylated human type Iii collagen fragment in Saccharomyces cerevisiae', *DNA Cell Biol.*, 1998, **17**, (6), pp. 511–518
- Stein, H., Wilensky, M., Tsafir, Y., et al.: 'Production of bioactive, post-translationally modified, heterotrimeric, human recombinant type-I collagen in transgenic tobacco', *Biomacromolecules*, 2009, **10**, (9), pp. 2640–2645
- Yaari, A., Posen, Y., Shoseyov, O.: 'Liquid crystalline human recombinant collagen: the challenge and the opportunity', *Tissue Eng. A*, 2013, **19**, (13–14), pp. 1502–1506
- Tanaka, M., Sato, K., Uchida, T.: 'Plant prolyl hydroxylase recognizes poly (L-proline) li helix', *J. Biol. Chem.*, 1981, **256**, (22), pp. 11397–11400
- Ruggiero, F., Exposito, J.Y., Boumat, P., et al.: 'Triple helix assembly and processing of human collagen produced in transgenic tobacco plants', *FEBS Lett.*, 2000, **469**, (1), pp. 132–136
- Merle, C., Perret, S., Lacour, T., et al.: 'Hydroxylated human homotrimeric collagen I in agrobacterium tumefaciens-mediated transient expression and in transgenic tobacco plant', *FEBS Lett.*, 2002, **515**, (1–3), pp. 114–118
- Gelvin, S.B.: 'Agrobacterium-mediated plant transformation: the biology behind the 'gene-jockeying' tool', *Microbiol. Mol. Biol. Rev.*, 2003, **67**, (1), pp. 16–37, table of contents
- Shoseyov, O.S., Hanan: 'Collagen producing plants and methods of generating and using' (Collplant Ltd, 06 April, 2006)
- Brodsky, B., Kaplan, D.L.: 'Shining light on collagen: expressing collagen in plants', *Tissue Eng. A*, 2013, **19**, (13–14), pp. 1499–1501
- Rooney, C., Sethi, T.: 'Advances in molecular biology of lung disease: aiming for precision therapy in non-small cell lung cancer', *Chest*, 2015, **148**, (4), pp. 1063–1072



- 61 Miller-Fleming, L., Olin-Sandoval, V., Campbell, K., *et al.*: 'Remaining mysteries of molecular biology: the role of polyamines in the cell', *J. Mol. Biol.*, 2015, **427**, (21), pp. 3389–3406
- 62 Xu, X., Gan, Q., Clough, R.C., *et al.*: 'Hydroxylation of recombinant human collagen type I alpha 1 in transgenic maize co-expressed with a recombinant human prolyl 4-hydroxylase', *BMC Biotechnol.*, 2011, **11**, (1), p. 69
- 63 Geddis, A.E., Prockop, D.J.: 'Expression of human colla1 gene in stably transfected Ht1080 cells: the production of a thermostable homotrimer of type I collagen in a recombinant system', *Matrix*, 1993, **13**, (5), pp. 399–405
- 64 Fertala, A., Sieron, A.L., Ganguly, A., *et al.*: 'Synthesis of recombinant human procollagen Ii in a stably transfected tumour cell line (Ht1080)', *Biochem. J.*, 1994, **298**, (1), pp. 31–37
- 65 Frischholz, S.: 'Characterization of human type X procollagen and its Nc-1 domain expressed as recombinant proteins in Hek293 cells', *J. Biol. Chem.*, 1998, **273**, (8), pp. 4547–4555
- 66 Fukuda, K., Hori, H., Utani, A., *et al.*: 'Formation of recombinant triple-helical [alpha 1(Iv)]2 alpha 2(Iv) collagen molecules in cho cells', *Biochem. Biophys. Res. Commun.*, 1997, **231**, (1), pp. 178–182
- 67 Mazzorana, M., Gruffat, H., Sergeant, A., *et al.*: 'Mechanisms of collagen trimer formation. construction and expression of a recombinant minigene in hela cells reveals a direct effect of prolyl hydroxylation on chain assembly of type Xii collagen', *J. Biol. Chem.*, 1993, **268**, (5), pp. 3029–3032
- 68 Fichard, A., Tillet, E., Delacoux, F., *et al.*: 'Human recombinant alpha(V) collagen chain. Homotrimeric assembly and subsequent processing', *J. Biol. Chem.*, 1997, **272**, (48), pp. 30083–30087
- 69 Unsold, C., Pappano, W.N., Imamura, Y., *et al.*: 'Biosynthetic processing of the pro-alpha 1(V)2pro-alpha 2(V) collagen heterotrimer by bone morphogenetic protein-1 and furin-like proprotein convertases', *J. Biol. Chem.*, 2002, **277**, (7), pp. 5596–5602
- 70 Pinkas, D.M., Ding, S., Raines, R.T., *et al.*: 'Tunable, post-translational hydroxylation of collagen domains in Escherichia coli', *ACS Chem. Biol.*, 2011, **6**, (4), pp. 320–324
- 71 Chen, M., Costa, F.K., Lindvay, C.R., *et al.*: 'The recombinant expression of full-length type Vii collagen and characterization of molecular mechanisms underlying dystrophic epidermolysis bullosa', *J. Biol. Chem.*, 2002, **277**, (3), pp. 2118–2124
- 72 Datar, R.V., Cartwright, T., Rosen, C.G.: 'Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasminogen activator', *Biotechnol. N Y*, 1993, **11**, (3), pp. 349–357
- 73 Tomita, M., Ohkura, N., Ito, M., *et al.*: 'Biosynthesis of recombinant human pro-alpha 1(Iii) chains in a baculovirus expression system: production of disulphide-bonded and non-disulphide-bonded species containing full-length triple helices', *Biochem. J.*, 1995, **312**, (Pt 3), pp. 847–853
- 74 Tomita, M., Kitajima, T., Yoshizato, K.: 'Formation of recombinant human procollagen I heterotrimers in a baculovirus expression system', *J. Biochem.*, 1997, **121**, (6), pp. 1061–1069
- 75 Li, H.-C., Huang, C.-C., Chen, S.-F., *et al.*: 'Assembly of homotrimeric type Xxi minicollagen by coexpression of prolyl 4-hydroxylase in stably transfected drosophila melanogaster S2 cells', *Biochem. Biophys. Res. Commun.*, 2005, **336**, (2), pp. 375–385
- 76 Lamberg, A., Helaakoski, T., Myllyharju, J., *et al.*: 'Characterization of human type Iii collagen expressed in a baculovirus system: production of a protein with a stable triple helix requires coexpression with the two types of recombinant prolyl 4-hydroxylase subunit', *J. Biol. Chem.*, 1996, **271**, (20), pp. 11988–11995
- 77 Myllyharju, J., Lamberg, A., Notbohm, H., *et al.*: 'Expression of wild-type and modified pro chains of human type I procollagen in insect cells leads to the formation of stable [1(I)]2 2(I) collagen heterotrimers and [1(I)]3 homotrimers but not [2(I)]3 homotrimers', *J. Biol. Chem.*, 1997, **272**, (35), pp. 21824–21830
- 78 Pihlajamaa, T., Perala, M., Vuoristo, M.M., *et al.*: 'Characterization of recombinant human type Ix collagen. Association of alpha chains into homotrimeric and heterotrimeric molecules', *J. Biol. Chem.*, 1999, **274**, (32), pp. 22464–22468
- 79 Nokelainen, M., Helaakoski, T., Myllyharju, J., *et al.*: 'Expression and characterization of recombinant human type Ii collagens with low and high contents of hydroxylysine and its glycosylated forms', *Matrix Biol.*, 1998, **16**, (6), pp. 329–338
- 80 Buechter, D.D., Paoletta D.N., Leslie, B.S., *et al.*: 'Co-translational incorporation of trans-4-hydroxyproline into recombinant proteins in bacteria', *J. Biol. Chem.*, 2002, **278**, (1), pp. 645–650
- 81 Holmgren, S.K., Taylor, K.M., Bretscher, L.E., *et al.*: 'A hyperstable collagen mimic', *Nature*, 1998, **392**, (6677), pp. 666–667
- 82 Eberhardt, E.S., Panasik, N., Raines, R.T.: 'Inductive effects on the energetics of prolyl peptide bond isomerization: implications for collagen folding and stability', *J. Am. Chem. Soc.*, 1996, **118**, (49), pp. 12261–12266
- 83 Luther, K.B., Hulsmeier, A.J., Schegg, B., *et al.*: 'Mimivirus collagen is modified by bifunctional lysyl hydroxylase and glycosyltransferase enzyme', *J. Biol. Chem.*, 2011, **286**, (51), pp. 43701–43709
- 84 Rutschmann, C., Baumann, S., Cabalzar, J., *et al.*: 'Recombinant expression of hydroxylated human collagen in Escherichia coli', *Appl. Microbiol. Biotechnol.*, 2013, **98**, (10), pp. 4445–4455
- 85 Bulleid, N.J., John, D.C., Kadler, K.E.: 'Recombinant expression systems for the production of collagen', *Biochem. Soc. Trans.*, 2000, **28**, (4), pp. 350–353
- 86 John, D.C.A., Watson, R., Kind, A.J., *et al.*: 'Expression of an engineered form of recombinant procollagen in mouse milk', *Nat. Biotechnol.*, 1999, **17**, (4), pp. 385–389
- 87 Guo, J., Luo, Y.e., Fan, D., *et al.*: 'Medium optimization based on the metabolic-flux spectrum of recombinant Escherichia colifor high expression of human-like collagen Ii', *Biotechnol. Appl. Biochem.*, 2010, **57**, (2), pp. 55–62
- 88 Luo, Y.E., Fan, D.D., Shang, L.A., *et al.*: 'Analysis of metabolic flux in Escherichia coli expressing human-like collagen in fed-batch culture', *Biotechnol. Lett.*, 2007, **30**, (4), pp. 637–643
- 89 Bahreini, E., Aghaiypour, K., Abbasalipourkabir, R., *et al.*: 'An optimized protocol for overproduction of recombinant protein expression in Escherichia coli', *Prep. Biochem. Biotechnol.*, 2014, **44**, (5), pp. 510–528
- 90 Grabherr, R., Nilsson, E., Striedner, G., *et al.*: 'Stabilizing plasmid copy number to improve recombinant protein production', *Biotechnol. Bioeng.*, 2002, **77**, (2), pp. 142–147
- 91 Nordén, K., Agemark, M., Danielson, J.Å., *et al.*: 'Increasing gene dosage greatly enhances recombinant expression of aquaporins in Pichia pastoris', *BMC Biotechnol.*, 2011, **11**, (1), p. 47
- 92 Dalton, A.C., Barton, W.A.: 'Over-expression of secreted proteins from mammalian cell lines', *Protein Sci.*, 2014, **23**, (5), pp. 517–525
- 93 Xiao, S., Shiloach, J., Betenbaugh, M.J.: 'Engineering cells to improve protein expression', *Curr. Opin. Struct. Biol.*, 2014, **26**, pp. 32–38
- 94 Xia, X.X., Qian, Z.G., Ki, C.S., *et al.*: 'Native-sized recombinant spider silk protein produced in metabolically engineered Escherichia coli results in a strong fiber', *Proc. Natl. Acad. Sci.*, 2010, **107**, (32), pp. 14059–14063
- 95 Brodel, A.K., Sonnabend A Fau - Kubick, S., Kubick, S.: 'Cell-free protein expression based on extracts from cho cells', (1097-0290, (Electronic))
- 96 Brodel, A.K., Wustenhagen Da Fau - Kubick, S., Kubick, S.: 'Cell-free protein synthesis systems derived from cultured mammalian cells', (1940-6029 (Electronic))
- 97 Carlson, E.D., Gan, R., Hodgman, C.E., *et al.*: 'Cell-free protein synthesis: applications come of age', *Biotechnol. Adv.*, 2012, **30**, (5), pp. 1185–1194
- 98 Nguyen, T.A., Lieu, S.S., Chang, G.: 'An Escherichia coli-based cell-free system for large-scale production of functional mammalian membrane proteins suitable for X-ray crystallography', *J. Mol. Microbiol. Biotechnol.*, 2010, **18**, (2), pp. 85–91

