Biomaterials 53 (2015) 309–317

Contents lists available at ScienceDirect

Biomaterials

journal homepage: www.elsevier.com/locate/biomaterials

Tuning cellular response by modular design of bioactive domains in collagen

Richard A. Que [a], Sam Wei Polly Chan [b], Abeer M. Jabaiah [b], Richard H. Lathrop [c], Nancy A. Da Silva [a, b, *], Szu-Wen Wang [a, b, *]

[a] Department of Biomedical Engineering, University of California, Irvine, CA 92697-2715, USA
[b] Department of Chemical Engineering and Materials Science, University of California, Irvine, CA 92697-2575, USA
[c] Department of Computer Science, University of California, Irvine, CA 92697-3435, USA

ARTICLE INFO

Article history:
Received 28 October 2014
Received in revised form
9 February 2015
Accepted 15 February 2015
Available online

Keywords:
Collagen
Biomimetic material
Extracellular matrix (ECM)
Cell adhesion
GFOGER
Integrin

ABSTRACT

Collagen's ability to direct cellular behavior suggests that redesigning it at the molecular level could enable manipulation of cells residing in an engineered microenvironment. However, the fabrication of full-length collagen mimics of specified sequence *de novo* has been elusive, and applications still rely on material from native tissues. Using a bottom-up strategy, we synthesized modular genes and expressed recombinant human collagen variants in *Saccharomyces cerevisiae*. The resulting biopolymers contained prescribed cell-interaction sites that can direct and tune cellular responses, with retention of the important triple-helical self-assembled structure. Removal of the native integrin-binding sites GROGER, GAOGER, GLOGEN, GLKGEN, and GMOGER in human collagen III yielded collagen that did not support adhesion of mammalian cells. Introduction of GFOGER sequences to this scaffold at specified locations and densities resulted in varying degrees of cellular attachment. The recruitment of focal adhesion complexes on the different collagens ranged from a 96% reduction to a 56% increase over native collagen I. Adhesion to the GFOGER-containing variants was entirely dependent and partially dependent on the β1 and α2 subunits of integrin, respectively, with cell adhesion on average reduced by 86% with anti-β1 and 38% with anti-α2 integrin antibody incubation. Results support the importance of local context in collagen-cell interactions. The investigation demonstrates the flexibility of this approach to introduce targeted changes throughout the collagen polymer for producing fully-prescribed variants with tailored properties.

1. Introduction

The importance of the microenvironment in cellular development, self-renewal, and differentiation has propelled efforts to create biomimetic materials that can direct cellular response for therapeutic applications and regenerative medicine. Characteristics of the local extracellular matrix (ECM), including nanoscale architecture, mechanical properties, cell-interaction site presentation, and degradation rates, have been shown to regulate cellular behavior [1–3]. However, limitations of using naturally-derived ECM components have included the inability to decouple such properties for independent assessment [2], and the use of animal sources, which give materials that can be poorly-defined and immunogenic in clinical applications [4,5]. To address these limitations, strategies to generate cell-responsive synthetic materials have included the chimeric integration of bioactive ECM sites into synthetic peptides or polymers (e.g., polyethylene glycol, hyaluronic acid) [6–8] and the use of recombinant systems for molecular-level control (e.g., elastin) [9].

Given that >90% of natural ECM tissue and approximately 30% of all mammalian proteins comprise the protein collagen [10,11], it follows that collagen is extensively used in regenerative medicine investigations as a scaffold on which to grow cells. However, current tissue engineering efforts rely on "off-the-shelf" native collagen obtained from animal sources, such as bovine collagen or Matrigel, so there is little or no control over molecular-level parameters, such as protein sequence, cell–collagen interactions, and residual bioactive impurities [8]. Recombinant collagen has been

* Corresponding authors. Department of Chemical Engineering and Materials Science, University of California, Irvine, 916 Engineering Tower, Irvine, CA 92697-2575, USA. Fax: +1 949 824 2541.
E-mail addresses: ndasilva@uci.edu (N.A. Da Silva), wangsw@uci.edu (S.-W. Wang).

http://dx.doi.org/10.1016/j.biomaterials.2015.02.074
0142-9612/

previously expressed by other groups [12,13] and is commercially available; however, these recombinant systems generate collagen with native sequence and do not demonstrate the ability to design multiple, relatively significant changes at specified locations throughout the biopolymer. This is because a synthetic bottom-up approach has been elusive, due to the difficulties in generating synthetic genes encoding the $(glycine\text{-}X\text{-}Y)_N$ backbone of collagen and the stringent need to post-translationally hydroxylate prolines in the Y-position for stability [14]. Consequently, synthesis of designer collagen-mimetic material has primarily focused on peptides [15–19], unhydroxylated collagen-like polymers from bacteria [20,21], or tandem repeating collagen domains [22].

Our research group has developed a platform for creating recombinant human collagen III in which we can specifically tailor the identity, location, and frequency of functional sites within the biopolymer [14,23]. Collagen III was selected as the molecular scaffold because it is a homotrimer; experimentally, this requires only one gene to be synthesized with no need to separate the heterotrimeric populations of product. Synthetic genes encoding the collagen-mimetic biopolymers are fabricated from oligonucleotides in which DNA sequences are optimized for gene assembly and yeast expression using a biocomputational strategy. This optimization is necessary to overcome mishybridization propensities due to the repetitive glycine-X-Y amino acid sequences and G-C rich sequences [14]. Twelve gene modules (primary fragments, PF) were designed which span the entire triple-helical region of collagen III, and the baseline DNA modules encode for the human amino acid protein sequence. The protein is synthesized in a recombinant *Saccharomyces cerevisiae* yeast system that has been genetically altered to express human α- and β-prolyl hydroxylase, which impart stability to the collagen triple helix by post-translational hydroxylation [24]. In this investigation, we demonstrate the modularity of this platform and its ability to create defined, non-native variants of human collagen III in the context of integrin-binding sites.

Integrins are the primary receptors that mediate cell adhesion and mechanical interactions with the extracellular matrix and are important in processes such as adhesion, migration, and differentiation [25,26]. Genetic mutations in these receptors can result in pathological states such as tumor growth and metastasis, muscular dystrophy, and thrombosis [26]. While biomaterials engineering has focused on manipulating the fibronectin-based sequence RGD [7], only approximately one-third of all integrins bind to this sequence [26]. There is thus untapped potential to engage alternative integrins for biomaterials design, thereby expanding the potential to modulate alternative types of cells and their corresponding processes.

In this investigation, multiple native integrin-binding sites were removed from the natural amino acid sequence of collagen III [27,28], and a site from collagen I but nonexistent in collagen III (GFOGER) [29] was introduced at various locations and frequencies. Prior examples of biomaterials engineering with GFOGER have been limited to short peptides, which are able to support cellular adhesion and promote osteoblast-specific gene expression [15,17], but do not present the peptides in native context. Our bottom-up, modular strategy enables the re-design of full-length collagen at the molecular level to tailor collagen for tuning cellular microenvironments and response. This alternative strategy could expand the level of control over cellular behavior in therapeutics and regenerative medicine.

2. Materials and methods

2.1. Materials

Escherichia coli strain DH5α (Zymo Research) was used for plasmid maintenance and amplification. Haploid *S. cerevisiae* strain BYα2β2 [24] was used to express the collagen baseline and all mutants. TOPO vector used to hold PCR generated gene fragments, fetal bovine serum (FBS), penicillin, and streptomycin were purchased from Life Technologies. Restriction enzymes, DNA ligase, DNase, and RNase were purchased from New England Biolabs. KOD Hot Start DNA Polymerase, human collagen III, bovine collagen I, anti-integrin α2 (MAB1650Z), and anti-integrin β1 (MAB2253Z) were purchased from EMD Millipore. PfuUltra II DNA polymerase was purchased from Agilent Technologies. DNA fragments were purified using Zymoclean Gel DNA Recovery Kit (Zymo Research) after gel electrophoresis. Phenylmethylsulfonyl fluoride (PMSF) and BCA assay were purchased from Pierce. Pepsin and phosphate buffer saline (PBS) were purchased from MP Biomedicals. Human collagen I, bovine collagen III, and calcein-AM were purchased from BD Bioscience. Streptavidin-alkaline phosphatase was purchased from GE Healthcare. All *E. coli* and yeast growth media were purchased from Difco. Dulbecco's modified Eagle's medium (DMEM) and monoclonal anti-vinculin antibodies were purchased from Sigma. Unless otherwise noted, all other chemicals were purchased from Fisher Scientific.

2.2. Construction of genes encoding collagen-mimetic mutants

2.2.1. Integrin-binding site removal and insertion

Our goal was to develop collagen substrates with varying frequencies, location, and combinations of α2β1 binding sites and to examine their structure, stability, and biological effect on interacting cells. We used our modular human collagen gene, which has been described elsewhere [14], as the underlying backbone upon which desired variants were made. The design of full-length human collagen III includes twelve gene modules (or "primary fragments", PF), which span and encode for the entire triple-helical region (See Fig. 1 for "Baseline" collagen, rCol) and additionally includes the N- and C-terminal propeptides and telopeptides. These ends were included to promote correct triple-helical formation, and presence of the triple helix initiating C-propeptide was confirmed by Western blot after protein expression.

Mutants of human collagen were designed as described in Fig. 1 and Table 1. One of these variants included a collagen (rCol-0G) with the native α2β1 binding sites removed (GROGER, GAOGER, GLOGEN, GLKGEN, and GMOGER hexamer motifs [27,28]) and replaced with a non-binding sequence. To identify a replacement non-binding hexameric sequence, we calculated the thermodynamic melting temperatures (T_m) for oligonucleotides encoding an exhaustive list of hexamer peptides to identify sequences that would minimally interfere with PCR-based gene synthesis. The top-ten hexamer candidates were then cross-checked with their respective Toolkit III peptides [28] to determine the sequences likely to yield a well-defined, thermostable triple helix with low cellular binding. This analysis yielded the non-binding hexamer GSPGGK, which replaced the native α2β1 binding sites. Although the melting temperature of short peptides is lower when glycines are introduced into the X-position of Gly-X-Y [30], work by Raynal et al. reported that the region with the native GSPGGK site had a T_m which was approximately the average of all the 27-mer peptides spanning the entire triple-helical region of collagen III [28]. Therefore, we did not expect the introduction of this sequence to be disruptive to the overall melting temperature of rCol-0G.

An integrin-binding site from collagen I (GFOGER [29]) was systematically introduced into the rCol-0G scaffold. To preserve the immediate local context and stability of the GFOGER sequence, we included the three native flanking amino acids on both sides of the collagen I sequence. Therefore, to create the GFOGER variants, the amino acid sequence GER-GFOGER-GVQ was substituted for twelve amino acids in each of the respective replacement sites within the rCol-0G scaffold. Three factors went into the selection of the new locations for introducing the non-native GFOGER sequences. The first criterion was to choose primary fragments (PF) that yielded relatively uniform spacing between the GFOGERs. Secondly, the proximity to the center of the selected PFs was preferred, since mutagenesis in locations close to the fragment edges would overlap with the adjacent fragment and require additional oligonucleotide and PCR fragment synthesis. Finally, we favored regions that minimized the number of amino acid changes between the native and GFOGER sequences. For example, the GFOGER of PF1 was introduced into the same location as the native GROGER sequence.

2.2.2. Mutagenesis of primary fragments in collagen variants

The construction of the genes encoding the collagen-mimetic variants was performed according to protocols described in Chan et al. [14] and in Supplementary Information (SI). After defining the amino acid sequences of the different variants (Table 1), the DNA sequences were optimized by a computational algorithm to favor correct hybridizations between synthetic oligonucleotides for assembly by PCR, to disrupt incorrect hybridizations, and for expression in *S. cerevisiae* [14,31].

For rCol-0G, the PFs containing the non-binding hexamer GSPGGK (replacing the native hexameric integrin-binding sites) were assembled by PCR from oligonucleotides using previously-described conditions [14]. To introduce the non-native GFOGER sequences into PFs, we used previously-described strategies of assembly by PCR using oligonucleotides for encoding GFOGER into PF1 [14] and site-directed mutagenesis using PFs 4, 8, and 11 for rCol-0G as the template DNA [32]. Details of the protocol and a list of the oligonucleotides used for mutagenesis are given in SI. The correct final sequences for PFs were confirmed by DNA sequencing.