



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 17 March 2015

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 $\beta 1$ and $\alpha 2$ subunits of integrin, respectively, with cell adhesion on average reduced by 86% with anti- $\beta 1$ and 38% with anti- $\alpha 2$ 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.

© 2015 Elsevier Ltd. All rights reserved.

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).

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-X-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 DH5 α (Zymo Research) was used for plasmid maintenance and amplification. Haploid *S. cerevisiae* strain BY α 2 β 2 [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 α 2 (MAB1650Z), and anti-integrin β 1 (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 α 2 β 1 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-OG) with the native α 2 β 1 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 α 2 β 1 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-OG.

An integrin-binding site from collagen I (GFOGER [29]) was systematically introduced into the rCol-OG 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-OG 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-OG, 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-OG 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.

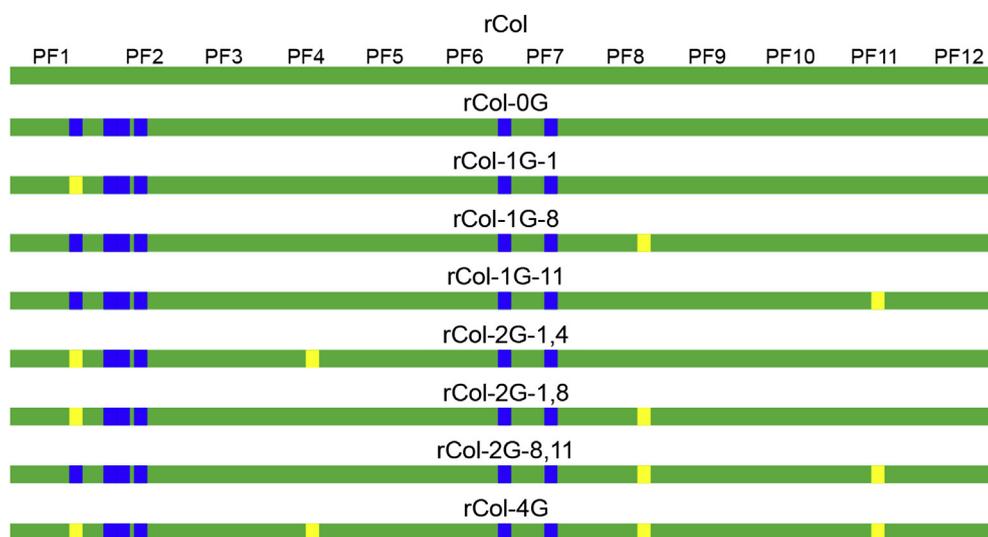


Fig. 1. Summary of recombinant collagen variants containing integrin-binding sites in the triple-helical domain. Variants were produced by removing all native integrin sites from human collagen III (rCol-0G) and systematically introducing non-native integrin-binding sites (GFOGER) in different frequencies and locations. Green regions denote amino acid sequences native to human collagen III, blue regions denote where native integrin-binding sites have been replaced with non-binding sites, and yellow regions denote non-native GFOGER sites from collagen I. “PF” denotes the primary fragment of the triple-helical region (one of 12 modules used in gene synthesis). The N- and C-propeptide regions are not pictured; they are expressed but removed after triple-helical assembly. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.2.3. Fabrication of genes encoding full-length collagen variants

Full-length genes were assembled based on our previously-reported methods [14,23], in which secondary fragments were assembled by PCR from PFs, and full-length genes for collagen variants were assembled from the secondary fragments. Details of the protocols are given in SI.

2.3. Expression and purification of recombinant collagen variants

Expression and purification were performed using previously-described protocols as a departure point [23,24], and details are given in SI. The full-length collagen genes were cloned in the CEN/ARS plasmid and transformed in *S. cerevisiae* strain BY α 2 β 2 in which we had integrated two copies each of α - and β -prolyl hydroxylase genes into the genome [24]. Expression of collagen was performed, yeast cells were harvested, and resulting collagen variants were purified.

2.4. Stability and structural analysis

Confirmation of triple helix formation and thermostability analysis were performed via circular dichroism (CD). Atomic force microscopy was used to image proteins and measure structural characteristics. These methods are based on previously published methods [23,24] and details are described in SI.

2.5. Cell adhesion on collagen substrates

2.5.1. Substrate deposition and cell adhesion imaging

HT-1080 cells were grown in DMEM (with 10% FBS, 100 units/ml penicillin, and 100 μ g/ml streptomycin) and subcultured every 3 days. Cells were trypsinized,

washed, and resuspended in DMEM (with 0.1% BSA) before seeding onto the protein-adsorbed surfaces. To prepare these surfaces, 20 μ g/ml native and recombinant collagens (in PBS) were incubated on 96-well non-tissue culture treated polystyrene plates (Corning) for 24 h at 4 $^{\circ}$ C. Our prior investigations showed that protein adsorption onto non-tissue culture treated wells was saturated when surfaces were incubated with ≥ 10 μ g/ml collagen [24]. Relative surface protein concentrations were quantified as described in SI.

Protein solutions were removed, and surfaces were incubated with DMEM with 0.5% BSA for 1 h. HT-1080 cells suspended in DMEM containing 0.1% BSA were then seeded at 3.15×10^4 cells/cm² on the protein-adsorbed surfaces and incubated for 4 h at 30 $^{\circ}$ C. The wells were washed with serum-free DMEM and imaged (Olympus IX-51). To quantify cell spreading, we measured cell areas. The boundaries for individual cells were manually defined using ImageJ (v 1.48) [53], and the areas from fifty adherent cells (N = 50) on each substrate were used to obtain statistical information.

2.5.2. Quantifying the number of cells adhered to surfaces

HT-1080 cells were incubated on the different protein-adsorbed surfaces for 4 h at 30 $^{\circ}$ C. To remove non-specifically bound cells, wells were filled with DMEM containing 0.1% BSA, and the plate was sealed with Microseal tape (Bio-Rad). The plates were inverted, centrifuged, and washed with PBS. Adhered cells were incubated with 4 μ M of calcein-AM (in PBS and 0.1% BSA) at 37 $^{\circ}$ C for 1 h. Solutions were removed, and cells were lysed with Cellytic M (Sigma). Fluorescence was measured at 494/517 nm (Ex/Em) (SpectraMax M2) and normalized to values for hCol I.

Table 1

Summary of recombinant collagen variants and native human collagen controls. Listed are the abbreviated name of the variant and the number and location of GFOGER integrin-binding sites. Primary fragment (PF) indicates the region in which the GFOGER sequence is located.

Name of collagen	Number of GFOGER sites	Description of collagen variant
hCol I	1	Human collagen I from native source; 1 native GFOGER site
hCol III	–	Human collagen III from native source (no GFOGER, but contains GROGER, GAOGER, GLOGEN, GLKGGEN, and GMOGER binding sites)
rCol	–	Human collagen III amino acid sequence (no GFOGER, but contains GROGER, GAOGER, GLOGEN, GLKGGEN, and GMOGER binding sites)
rCol-0G	0	Human collagen III with GROGER, GAOGER, GLOGEN, GLKGGEN, and GMOGER integrin-binding hexamers replaced with non-binding GSPGGK
rCol-1G-1	1	rCol-0G with GFOGER in PF1
rCol-1G-8	1	rCol-0G with GFOGER in PF8
rCol-1G-11	1	rCol-0G with GFOGER in PF11
rCol-2G-1,4	2	rCol-0G with GFOGERs in PF1 and PF4
rCol-2G-1,8	2	rCol-0G with GFOGERs in PF1 and PF8
rCol-2G-8,11	2	rCol-0G with GFOGERs in PF8 and PF11
rCol-4G	4	rCol-0G with GFOGERs in PF1, PF4, PF8, and PF11

2.5.3. Cellular adhesion modulation assays

To evaluate whether modulation of adhesion could be achieved through mixtures of binding and non-binding variants, native hCol I, rCol-1G-8, and rCol-4G were diluted with rCol-0G, yielding solutions containing 0.3–20 $\mu\text{g/ml}$ of the GFOGER-containing collagen, but with a 20 $\mu\text{g/ml}$ total protein concentration. As described above, protein mixtures were adsorbed and incubated with cells, and cell adhesion was quantified.

2.5.4. Integrin inhibition assays

We examined the specificity of the cell binding via $\alpha 2\beta 1$ integrins to the collagen substrates. HT-1080 cells were resuspended in DMEM supplemented with 0.1% BSA at 3.15×10^4 cells/cm². Anti- $\alpha 2$, anti- $\beta 1$, or non-specific IgG antibodies (at 10 $\mu\text{g/ml}$) were incubated with the cell suspension for 15 min at 37 °C. These cell/antibody solutions were incubated with collagen-adsorbed surfaces for 4 h at 30 °C. Cells were quantified with calcein-AM.

2.6. Quantification of focal adhesion complexes

To investigate the effect of varying integrin-binding sites within collagen substrates on the formation of focal adhesion complexes, we measured the relative amount of vinculin based on a previous protocol [33]. Wells of a non-tissue culture treated 6-well plate were coated with collagen-variant proteins. HT-1080 cells were seeded at 6.3×10^4 cells/cm² in DMEM supplemented with 10% FBS and incubated for 4 h at 30 °C. The solution was gently removed and replaced with DPBS, and bound integrins were crosslinked with 1 mM DTSSP in DPBS for 30 min. Unreacted DTSSP was quenched with 50 mM Tris–HCl in DPBS for 5 min. Uncrosslinked cellular components were extracted with 0.1% SDS in DPBS supplemented with 1 mM PMSF for 5 min. Wells were gently washed with DPBS, and crosslinked proteins were released by the addition of 50 mM DTT and 0.1% SDS in DPBS (without Ca²⁺ and Mg²⁺) at 37 °C for 30 min. Vinculin was detected through Western blot using monoclonal anti-vinculin antibodies. Band intensities were quantified using ImageJ and normalized to the hCol I band intensity for each experiment. Statistical analysis was performed on three independent experiments (N = 3).

2.7. Statistical analyses

For cell adhesion and focal adhesion quantification assays, statistical analyses were performed using at least three independent experiments. Each independent cell adhesion and antibody inhibition experiment contained at least three replicates (total N = 9); cell area analyses contained 50 cells (N = 50). Statistical significance was determined using the software program R [34]. We performed a one-way analysis of variance (ANOVA), followed by Tukey's HSD post-test comparison. Data is reported as mean \pm standard error of mean (S.E.M.) unless otherwise stated. Values were considered statistically significant when p-values were <0.05.

3. Results and discussion

3.1. Design strategy and synthesis of collagen genes encoding different combinations of non-native integrin-binding sites

To examine the feasibility of introducing non-native cell-responsive sites into a collagen scaffold and to study their respective bioactive properties, we designed a set of collagen variants containing the collagen I GFOGER $\alpha 2\beta 1$ integrin-binding sequence in different locations, frequencies, and combinations within a collagen III scaffold (summarized in Table 1, Fig. 1, and Supplementary Information [SI]). The gene synthesis for this set of collagen variants was performed using our modular human collagen III gene, described previously [14,24], coding for the biopolymer backbone.

Since collagen III naturally exhibits strong adhesion to cells, it was necessary first to design a full-length collagen variant in which native integrin interaction sites were removed to serve as a non-binding scaffold. Studies have identified the main $\alpha 2\beta 1$ integrin binding amino acid sequences in human collagen III to be the peptide hexamers GROGER, GAOGER, GLOGEN, GLKGEN, and GMOGER [27,28] (single amino acid abbreviations used). Based on the overlap between calculated optimized oligonucleotide melting temperatures (T_m) for hexameric sequences in gene synthesis [14], together with experimentally-determined triple-helical sequences that are both thermostable and exhibit minimal cell adhesion properties [28], we identified the hexameric sequence GSPGGK to replace each of the five native integrin-binding sites in collagen III.

This yielded the DNA encoding the protein sequences for the collagen variant “rCol-0G”. This gene assembled correctly as confirmed by DNA sequencing.

The integrin-binding sequence GFOGER was then introduced into the rCol-0G scaffold in primary fragments 1, 4, 8, and 11. GFOGER (Gly-Phe-Hyp-Gly-Glu-Arg) is found once within natural collagen I but never in collagen III, and it is specifically recognized by $\alpha 1\beta 1$, $\alpha 2\beta 1$, $\alpha 10\beta 1$, and $\alpha 11\beta 1$ integrins [26,29]. As with rCol-0G, all genes encoding these variants correctly assembled, demonstrating the modularity of the optimized collagen III gene and the ability to fabricate the genes encoding both single-GFOGER variants and those with a combination of GFOGER sites (Table 1 and Fig. 1). These full-length genes were then placed into the yeast expression system for expression, purification, and characterization of each collagen variant.

3.2. The collagen scaffold was tolerant of numerous changes while retaining triple-helical architecture

All collagen variants were successfully expressed and purified (Fig. S-1, SI). CD showed that all variants yielded spectra similar to data for rCol [24], and results confirmed their triple-helical structure (Fig. 2; Fig. S-2, SI). Fig. 2a shows that the variant with the maximum number of changes (rCol-4G) still remained triple-helical. The linear, triple-helical structure was confirmed by AFM imaging (Fig. 2c). Height measurements through the centers of single polymeric strands yielded 1.1 ± 0.3 nm, 0.9 ± 0.2 nm, and 0.9 ± 0.2 nm for rCol-0G, rCol-4G, and the hCol III control, respectively, consistent with previously-reported values for collagen assembled in a triple helix [24,35].

Thermostability scans were performed to determine apparent melting temperatures T_m (Fig. 2b, SI-3). Results for all variants are summarized in Tables S-1 (SI). Recombinant collagen III (with native human collagen III sequence; rCol) gave an average T_m of 35.8 ± 0.5 °C, and all recombinant GFOGER variants yielded comparable T_m values. We note that all of the recombinant proteins consistently yielded T_m values approximately 3.5 °C lower than human collagen from native tissue sources. This slightly lower stability has also been observed for fibrillar-type collagens expressed in other *S. cerevisiae* systems [13] and is likely due to the heterologous expression in yeast [14,24]. Yeast does not natively hydroxylate proline; however, hydroxyproline imparts stability to the collagen triple helix. Thus, we designed the strain to contain two copies of the α - and β -prolyl hydroxylase genes integrated into the yeast genome (BY $\alpha 2\beta 2$), with the collagen gene placed on a CEN/ARS plasmid (YCpMCOL) to facilitate variant manipulation [14,24]. The resulting proline hydroxylation level of our yeast system measured using amino acid analysis yielded $31.1 \pm 2.4\%$, which is lower than the $45 \pm 1.4\%$ measured for native hCol III. Ongoing investigations continue to optimize the yeast expression system further to increase hydroxylation levels and thermostability of the collagen trimer.

The high tolerance for mutations in the collagen scaffold shows the robustness of the modular platform and the scaffold's structural backbone. The trimeric assembly and thermostability results were not obvious *a priori*, as these variants contained significantly more amino acid changes than had been reported previously. For example, rCol-0G and rCol-4G contained 19 and 40 amino acid changes from the native human collagen III sequence, respectively. In contrast, the maximum number of designed mutations (of which we are aware) was five amino acids in a localized cluster [36], and even single amino acid changes in the Gly-position of fibrillar collagen can cause severe phenotypes [37]. Our results demonstrate that the collagen backbone can indeed serve as an underlying scaffold for modular assembly, as long as the Gly-X-Y sequence

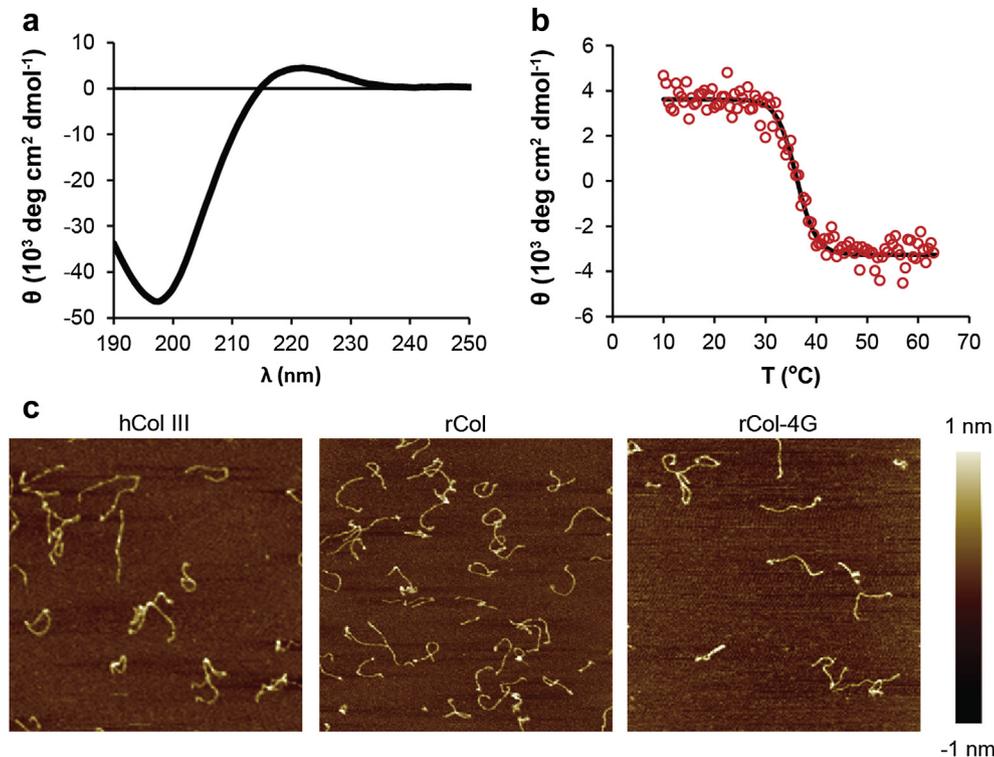


Fig. 2. Structural characterization of collagen variant with 4 GFOGER sites introduced into a non-binding collagen scaffold (rCol-4G) and controls. (a) Representative circular dichroism (CD) wavelength spectra for rCol-4G. (b) Representative thermostability scan measuring ellipticity at 221 nm vs. temperature for rCol-4G. (c) Representative AFM images for native human collagen III (hCol III), recombinant human collagen (rCol), and rCol-4G. Scanned areas are $1 \mu\text{m} \times 1 \mu\text{m}$.

pattern is retained and the sequence motifs that are substituted into the scaffold form stable trimers.

The retention of triple-helical structure after bioactive site incorporation is particularly important for the functional properties of the collagen-mimetic polymer, since integrin receptors for collagen simultaneously interact with multiple strands within intact triple-helical collagen [38]. This result also suggests that prior challenges to create extensively-mutated collagen *in vitro* were likely due to limitations of recombinant DNA technology, rather than the inherent triple-helical stability of collagen variants. Codon optimization for gene synthesis was necessary to minimize oligonucleotide mishybridizations [14]. It is unlikely that these triple-helical recombinant collagens will form higher-order structures, such as fibrils. Telopeptides are necessary for fibril assembly [39], and our purification protocol uses pepsin to remove endogenous yeast proteins, which likely also cleaves the telopeptide region. Alternative avenues to produce collagen variants with intact telopeptides are currently being explored. Although fibril assembly to obtain higher-order three-dimensional structures is an important aspect of collagen's use as a tissue engineering hydrogel scaffold, hydrogel formation can also be accomplished with other strategies, such as utilizing introduced non-native cysteines [23].

3.3. Cellular adhesion to collagen III was abrogated upon removal of integrin binding sites

Differences between the adhesion of HT-1080 cells to the variants of collagen were observed in cell microscopy images (Fig. 3a–l), and relative cell numbers were quantified by calcein-AM (Fig. 3m). As expected, cells adhered to native human collagen I and III controls from tissue sources (Fig. 3l, k), but did not adhere to bovine serum albumin (BSA) (Fig. 3a). The baseline

recombinant collagen (rCol), produced in yeast and comprising the same amino acid sequence as native human collagen III, bound as many cells as the native collagens ($p > 0.05$) (Fig. 3j). Accessible surface concentrations of collagens were confirmed to be at saturated and comparable levels between all samples, assuring any observed differences in cell binding are due to molecular-level differences in the collagen substrates rather than surface protein concentrations (Fig. S-4, SI).

Replacement of the five integrin-binding sites from collagen III [27,28] with the non-binding sequence GSPGGK [28] yielded the variant rCol-0G, on which no adhesion of HT-1080 cells was observed (Fig. 3b). Since this collagen variant remained triple-helical and showed no loss of stability relative to the recombinant baseline protein (Fig. S-2 and Table S-1, SI), the lack of cell adhesion must be specific to the loss of interactions between the integrin receptors on the cell surface and the hexameric peptide sequences within collagen. The variant rCol-0G, therefore, can serve as a non-binding collagen III scaffold, allowing for the evaluation of introduced integrin-binding sites.

3.4. Introduction of the collagen I integrin-binding site (GFOGER) at different frequencies and locations altered cellular adhesion

Crystal structures of a GFOGER-containing peptide bound by $\alpha 2\beta 1$ integrin show that integrin interacts with two of the three triple-helical collagen strands [38,40], confirming that the triple-helical tertiary structure of collagen is critical for recognition. We observed that the introduction of only a single-GFOGER motif into the ~1000-amino acid triple-helical region of the rCol-0G backbone recovered cellular adhesion (Fig. 3c–e). Interestingly, we observed a marked difference in adhesion, depending on the location of the GFOGER placement site. Single-GFOGER substitutions placed in

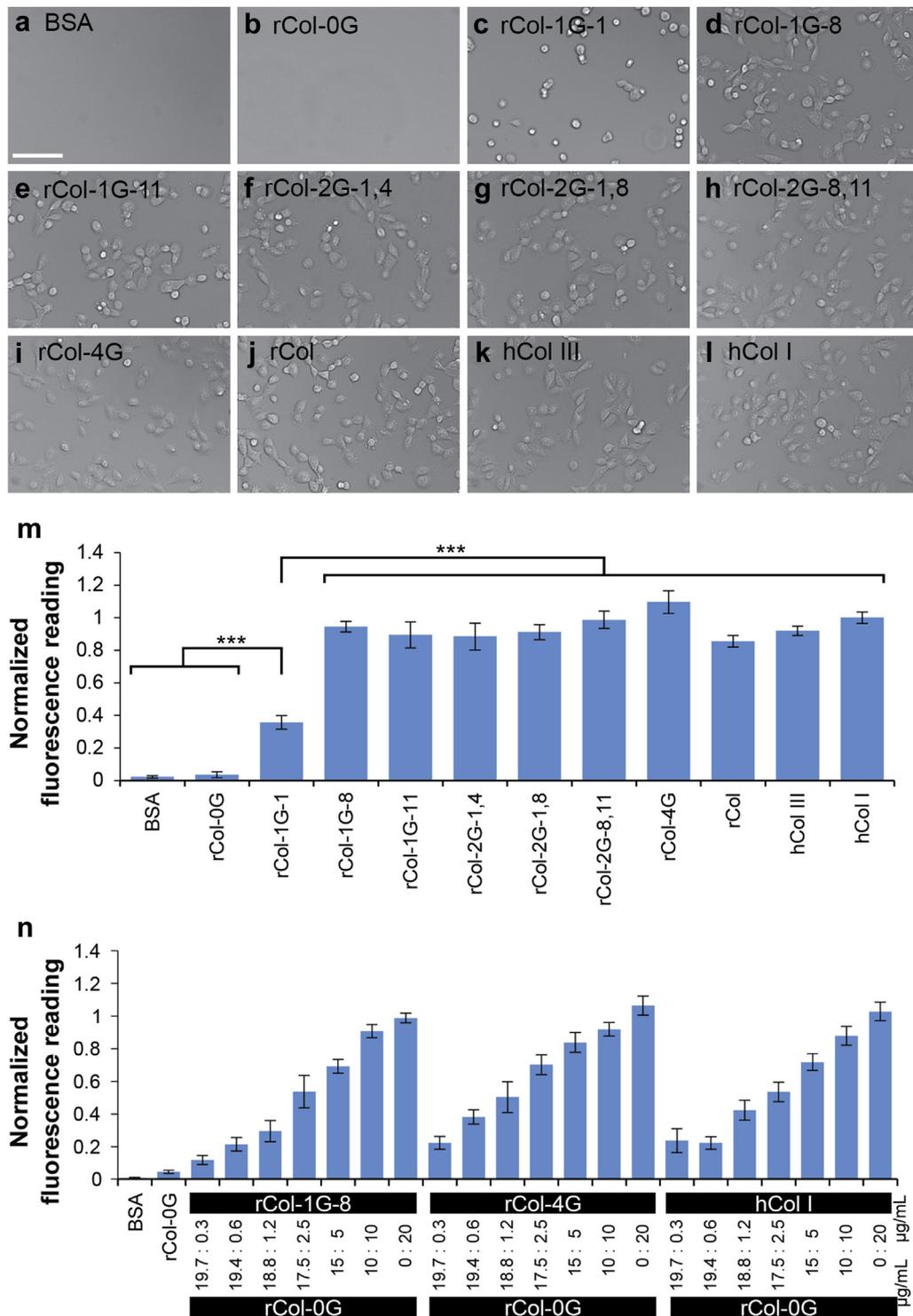


Fig. 3. Adhesion of HT-1080 cells to collagen and collagen-mimetic substrates. (a–l) Representative cell images for adhesion to collagen variants and controls. Scale bar = 100 µm. Substrates are: (a) bovine serum albumin (BSA); (b) rCol-0G; (c) rCol-1G-1; (d) rCol-1G-8; (e) rCol-1G-11; (f) rCol-2G-1,4; (g) rCol-2G-1,8; (h) rCol-2G-8,11; (i) rCol-4G; (j) recombinant human collagen III (rCol); (k) native human collagen III (hCol III); and (l) native human collagen I (hCol I). (m) Fluorescence measurements quantifying the relative number of cells adhered to collagen substrates using a calcein-AM assay. Values were normalized to average for hCol I. Cell adhesion on BSA and rCol-0G was significantly lower than on all other variants ($p < 0.001$). The adhesion on rCol-1G-1 was significantly different than all other groups ($p < 0.001$). Comparisons between substrates showing high adhesion (substrates shown in panels 3d through 3l) indicated no significant differences ($p > 0.05$). *** $p < 0.001$. Mean \pm S.E.M. (n) Fluorescence measurements quantifying the relative number of cells on surfaces coated with solutions of 0.3–20 µg/ml binding collagen (rCol-1G-8, rCol-4G, and hCol I) diluted in non-binding collagen (rCol-0G). Total protein concentrations (binding + non-binding collagens) were 20 µg/ml. Fluorescence values were normalized to average for 20 µg/ml hCol I. Mean \pm S.E.M.

primary fragment (PF) 8 or 11 (rCol-1G-8; rCol-1G-11) increased the number of adhered cells to values similar to those for native human collagen I and III (Fig. 3m). Differences between cell

spreading areas (per cell) for these two variants and native human collagen I and III were not significant ($p > 0.05$). Taken together, these results demonstrated similar cell numbers and quality of

attachment. It was surprising that a single hexamer sequence from collagen I can promote this degree of adhesion, considering that native collagens contain not only one but multiple integrin-binding sites. In contrast, cell adhesion to collagen with a single-GFOGER introduced into PF1 (rCol-1G-1) was weaker than on other GFOGER-containing collagen substrates (rCol, rCol-1G-8, rCol-1G-11, rCol-2G variants, rCol-4G) and native controls, with only approximately one-third the number of cells attached to the rCol-1G-1 substrate as compared to native controls (Fig. 3m). Furthermore, cells on rCol-1G-1 did not spread as well as on the other surfaces (Fig. 3c). The average surface area per cell on rCol-1G-1 was 50% and 54% of areas on hCol I and hCol III, respectively ($p < 0.001$).

The significant difference in binding between the rCol-1G mutation locations (in PF 1, 8, or 11) demonstrates the importance of the local collagen context for integrin-binding interactions. The results of the AFM analysis (Fig. 2d) support a possible mechanism for this difference. Although images showed that all recombinant collagens were linear and triple-helical, the apparent lengths for the imaged GFOGER mutants (224 ± 12 nm for rCol-0G, 210 ± 21 nm for rCol-1G-1, and 235 ± 27 nm for rCol-4G) were 10% shorter than for the controls (256 ± 10 nm for native hCol III and 251 ± 13 nm for rCol). Closer examination of individual GFOGER-variant trimers indicated that one end of the trimeric polymer often had a greater height or aggregate for a length ranging approximately 15–30 nm. This suggested that the decreased overall length could be due to one end of the GFOGER-collagens folding back on itself. If the folding propensity were in the N-terminus, then this could occlude GFOGER binding in PF1. Our data did not support the alternative scenario of proteolytic truncation; SDS-PAGE analysis showed all collagens in this investigation exhibit comparable molecular weights (Fig. S-1, SI). Furthermore, trimer misalignment as a mechanism for decreased adhesion in rCol-1G-1 was also unlikely, since the N- and C-terminal propeptide domains were expressed. Therefore, they were present to promote correct triple-helical assembly before their cleavage during purification.

3.5. Modulation of adhesion could be achieved by mixing different ratios of the binding and non-binding variants

Binding variants rCol-1G-8 and rCol-4G were diluted with rCol-0G, producing various concentrations of the GFOGER-containing proteins while keeping total protein amount at 20 $\mu\text{g/ml}$. We observed that the number of adhered cells could be tuned as rCol-1G-8 or rCol-4G was increased between 0.3 and 20 $\mu\text{g/ml}$ (Fig. 3n). Variant rCol-4G, containing four GFOGERS, did not significantly bind additional cells over the collagen with only one GFOGER (rCol-1G-8), hinting that the distance between GFOGER sites (intra-trimeric vs. inter-trimeric GFOGERS) may also be as important as the number of sites.

3.6. Inhibition of cell adhesion by anti- $\alpha 2$ and anti- $\beta 1$ integrin antibodies

We examined the subunit specificity of integrin-based cell adherence to GFOGER. Integrins that bind to collagens are $\alpha 1\beta 1$, $\alpha 2\beta 1$, $\alpha 10\beta 1$, and $\alpha 11\beta 1$ [26,29], with the most broadly expressed and well-studied being $\alpha 1\beta 1$ and $\alpha 2\beta 1$. These two integrins are involved in biological functions that include immunity, angiogenesis, and matrix remodeling [41]. In HT-1080 cells, the primary integrin that binds to triple-helical collagen (including the GFOGER sequence) is $\alpha 2\beta 1$ [17,42]. We therefore evaluated the specificity of cell binding to our GFOGER variants by inhibiting interactions with the $\alpha 2$ and $\beta 1$ subunits using antibodies.

Antibodies for the $\beta 1$ subunit strongly inhibited cell adhesion to all modular-GFOGER substrates (Fig. 4). All of the native human collagen controls and recombinant collagen variants containing GFOGER sites exhibited a significant decrease in cell adhesion when cells were first incubated with an antibody specific for the $\beta 1$ subunit, relative to media alone. Overall, cell adhesion was reduced by an average of 86% with anti- $\beta 1$ antibody for the variants that had demonstrated binding. This result is consistent with investigations showing that all collagen-binding integrins contain the $\beta 1$ subunit [26].

Incubation of cells with $\alpha 2$ -binding antibodies only partially inhibited cell adhesion to the GFOGER-containing substrates, with an average 38% decrease for the variants that showed binding with non-specific antibodies or media alone. The degree of inhibition varied with the underlying substrate. Incubation of cells with antibody concentrations higher than 10 $\mu\text{g/ml}$ yielded no additional decrease in cellular adhesion. In rCol-1G-1, which exhibited weaker binding compared to native collagens, the $\alpha 2$ antibody inhibited nearly all binding. However, for variants demonstrating stronger binding (e.g., rCol-2G-8,11; rCol-4G), cell binding was not significantly inhibited by anti- $\alpha 2$, suggesting that the increased number of GFOGER sites can sufficiently bind to non- $\alpha 2\beta 1$ collagen-binding integrins (e.g., $\alpha 1\beta 1$, $\alpha 10\beta 1$, $\alpha 11\beta 1$). This partial inhibition agrees with prior results for native collagen I [17] and is consistent with $\alpha 2\beta 1$ being the primary, but not sole, integrin receptor expressed by HT-1080 cells.

For studies utilizing short GFOGER-peptide substrates, however, the addition of $\alpha 2$ -binding antibodies to cells completely blocked all cell binding [17,43]. This apparent discrepancy suggests the importance of the full-length collagen in its native structure for integrin-mediated adhesion. This importance of the context of epitope presentation is also supported by prior work in which 30 times more copies of GFOGER were needed when the motif was surface-adsorbed and displayed in a peptide than when presented in full-length collagen I (by molar concentrations) [43]. While other recognition sites such as GLOGER and GASGER in collagen I have also been reported as binding to $\alpha 2\beta 1$ integrin [44], each site only occurs once in human collagen I [45] and at similar or weaker binding than GFOGER [27]. Therefore, their presence alone does not account for the significantly higher binding of full-length collagen I relative to a GFOGER-peptide, and suggests that the full structure of collagen in presenting the sequence plays a significant role in cell adhesion.

One possible explanation for the increased adhesion to full-length collagen relative to peptide may be that collagen provides numerous binding domains for other cell surface receptors, some of which support cooperative and synergistic interactions with integrin-mediated effects [46]. For example, several discoidin domain receptor (DDR) binding domains have been identified in human collagen III near the integrin-binding domains [46], and DDRs have been shown to enhance integrin-mediated cell adhesion [47]. Such effects are not recapitulated in systems that use only GFOGER peptides. Synergy between neighboring sequences has been reported in other protein systems. For example, cell binding to the RGD motif from fibronectin is significantly increased when presented concurrently with fibronectin's PHSRN peptide sequence in native spatial context [48,49]. Furthermore, different integrin-binding specificities are elicited for linear RGD and RGD-PHSRN peptides. Understanding the mechanisms of synergistic epitope effects can be important for regulating cell proliferation and differentiation [46,49,50].

Another factor which could potentially contribute to the observed increased binding to full-length collagen is the accessibility of the binding motifs. Immobilization of a GFOGER-peptide utilizing a PEO linker required approximately three times less

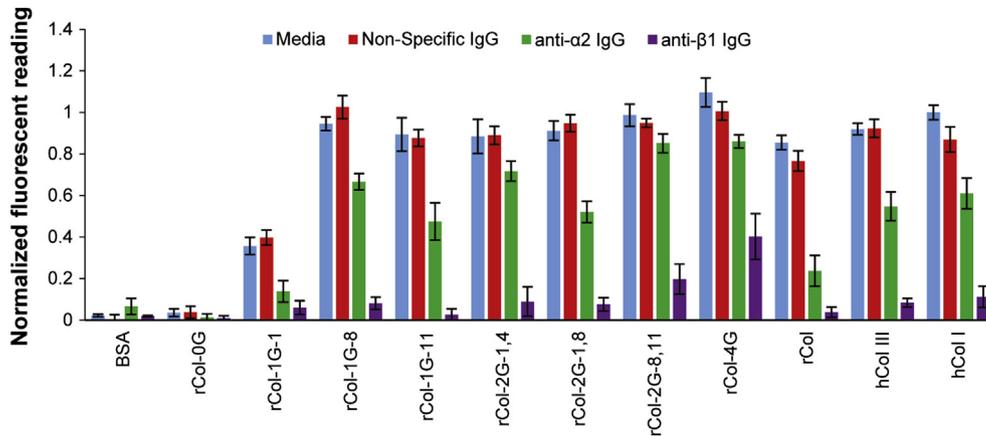


Fig. 4. Cell adhesion after inhibition of binding to $\alpha 2$ and $\beta 1$ integrin subunits. HT-1080 cells were incubated with no antibodies (media alone), non-specific IgG, anti- $\alpha 2$ IgG, and anti- $\beta 1$ IgG prior to incubation with different collagen substrates. Data quantifies the relative number of cells adhered using a calcein-AM assay. Fluorescence values were normalized to the average for hCol I with no antibodies. A significant decrease in adhesion due to anti- $\alpha 2$ antibody, relative to media-only conditions, was observed for hCol I, hCol III, rCol, rCol-1G-8, rCol-2G-1,8 ($p < 0.001$) and rCol-1G-1, rCol-1G-11 ($p < 0.01$). A significant decrease in adhesion due to anti- $\beta 1$ antibody, relative to media-only conditions, was observed for hCol I, hCol III, rCol, rCol-1G-1, rCol-1G-8, rCol-1G-11, rCol-2G-1,4, rCol-2G-1,8, rCol-2G-8,11, and rCol-4G ($p < 0.001$). Significant differences between $\alpha 2$ - and $\beta 1$ -subunit antibody inhibition for a given substrate was observed for hCol I, hCol III, rCol-1G-8, rCol-1G-11, rCol-2G-1,4, rCol-2G-1,8, rCol-2G-8,11 ($p < 0.001$) and rCol-4G ($p < 0.01$). Mean \pm S.E.M. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

peptide than peptide adsorbed to the surface (for the same number of cells adhered), suggesting that increased accessibility of GFOGER due to the linker could be important for adhesion [43]. This effect may occur in full-length collagen if only short regions of the collagen are required for adsorption to the surface. The total increased adhesion is likely a combination of multiple effects which includes both site accessibility and synergistic receptor interactions.

3.7. The number of focal adhesion complexes was modulated by GFOGER-collagen substrates

We examined the presence of vinculin, which is present in focal adhesion complexes, in cells seeded on varying collagen substrates

[51,52]. Our results demonstrated that the number of GFOGER sites on an external substrate can affect intracellular processes (Fig. 5). In Western blot analysis, the average vinculin band intensity for rCol-4G was greater than those for all other samples and controls. Cells seeded on rCol-4G exhibited approximately 56% higher levels of vinculin than on the native hCol I. Furthermore, cells incubated with the collagen variant containing no integrin-binding sites (rCol-0G) showed significantly less vinculin than with hCol I. Intensities were reduced by 96% from native human collagen I on average and close to that of the negative BSA control. The values for variant rCol-1G-1 averaged an intermediate intensity that was 12% of hCol I. These results are consistent with the cell adhesion data and show the potential to manipulate intracellular biochemical activity through integrin-binding sites in the underlying substrates [15].

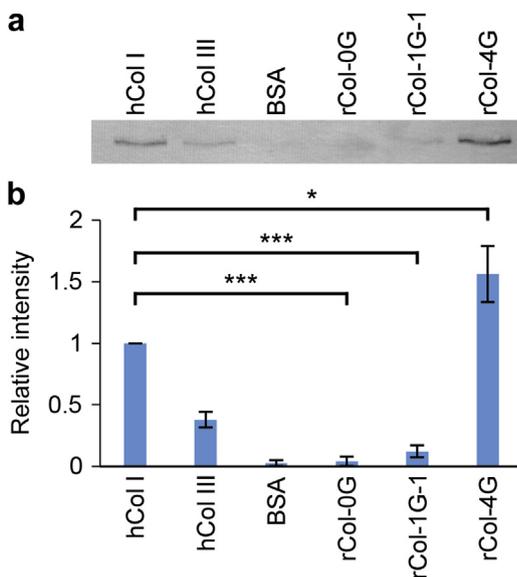


Fig. 5. Western blot analyses for vinculin in focal adhesion complexes. (a) Representative Western blot. (b) Band intensities, with values normalized to hCol I in each independent experiment. Intensity for rCol-4G was greater than those for all other samples and controls (with $p = 0.018$ compared to hCol I). * $p < 0.05$, *** $p < 0.001$. Mean \pm S.E.M.

4. Conclusions

In this study, we used a bottom-up, modular strategy that enabled systematic placement of integrin-binding sites in full-length collagen and allowed for the complete control of the number and location of these sites. Using the biopolymers created from this strategy, we found that the insertion of only a single-GFOGER site into a non-binding collagen variant was sufficient to restore full cellular adhesion. However, the location of the introduced site affected the degree of recovered adhesion and hints at the importance of epitope context. Cellular activity also could be modulated by mixing distinctly different collagen variants at tunable ratios. All variants were shown to be triple-helical, linear polymers, comparable to native human collagen. This investigation showed that one can re-design native collagen at the molecular level to tailor collagen-mimetic materials for manipulating cellular microenvironments and biochemical responses.

Acknowledgments

We gratefully acknowledge Dr. She-pin Hung for technical assistance in gene assembly of rCol-0G and rCol-1G-1, Cory Giertzak for help with the expression of collagen variants, and Dr. Adam Stieg (UCLA) for training and assistance with AFM images. CD was performed at the UCI Laser Spectroscopy Facility, and AFM was

performed at the Nano and Pico Characterization Facility at UCLA. This work was supported by the National Science Foundation (DMR-0706669, DMR-1006999, and CBET-1034566).

Appendix A. Supplementary data

Supplementary data related to this article can be found online at <http://dx.doi.org/10.1016/j.biomaterials.2015.02.074>.

References

- [1] Lutolf MP, Hubbell JA. Synthetic biomaterials as instructive extracellular microenvironments for morphogenesis in tissue engineering. *Nat Biotechnol* 2005;23:47–55.
- [2] Discher DE, Mooney DJ, Zandstra PW. Growth factors, matrices, and forces combine and control stem cells. *Science* 2009;324:1673–7.
- [3] von der Mark K, Park J, Bauer S, Schmuki P. Nanoscale engineering of biomimetic surfaces: cues from the extracellular matrix. *Cell Tissue Res* 2010;339:131–53.
- [4] Uemura M, Refaai MM, Shinoyama M, Hayashi H, Hashimoto N, Takahashi J. Matrigel supports survival and neuronal differentiation of grafted embryonic stem cell-derived neural precursor cells. *J Neurosci Res* 2010;88:542–51.
- [5] Lynn AK, Yannas IV, Bonfield W. Antigenicity and immunogenicity of collagen. *J Biomed Mater Res B Appl Biomater* 2004;71:343–54.
- [6] Lutolf MP, Weber FE, Schmoekel HG, Schense JC, Kohler T, Muller R, et al. Repair of bone defects using synthetic mimetics of collagenous extracellular matrices. *Nat Biotechnol* 2003;21:513–8.
- [7] Perlin L, MacNeil S, Rimmer S. Production and performance of biomaterials containing RGD peptides. *Soft Matter* 2008;4:2331–49.
- [8] Jia XQ, Kiick KL. Hybrid multicomponent hydrogels for tissue engineering. *Macromol Biosci* 2009;9:140–56.
- [9] DiMarco RL, Heilshorn SC. Multifunctional materials through modular protein engineering. *Adv Mater* 2012;24:3923–40.
- [10] Abraham LC, Zuena E, Perez-Ramirez B, Kaplan DL. Guide to collagen characterization for biomaterial studies. *J Biomed Mater Res B* 2008;87B:264–85.
- [11] Birk DE, Silver FH, Trelstad RL. Matrix assembly. In: Hay ED, editor. *Cell biology of extracellular matrix*. 2nd ed. New York: Plenum Press; 1991. p. 221–54.
- [12] Perret S, Eble JA, Siljander PR-M, Merle C, Farndale RW, Theisen M, et al. Prolyl hydroxylation of collagen type I is required for efficient binding to integrin $\alpha 1\beta 1$ and platelet glycoprotein VI but not to $\alpha 2\beta 1$. *J Biological Chem* 2003;278:29873–9.
- [13] Olsen DR, Leigh SD, Chang R, McMullin H, Ong W, Tai E, 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;276:24038–43.
- [14] Chan SWP, Hung SP, Raman SK, Hatfield GW, Lathrop RH, Da Silva NA, et al. Recombinant human collagen and biomimetic variants using a de novo gene optimized for modular assembly. *Biomacromolecules* 2010;11:1460–9.
- [15] Reyes CD, Garcia AJ. Alpha2beta1 integrin-specific collagen-mimetic surfaces supporting osteoblastic differentiation. *J Biomed Mater Res A* 2004;69:591–600.
- [16] Kotch FW, Raines RT. Self-assembly of synthetic collagen triple helices. *P Natl Acad Sci U. S. A* 2006;103:3028–33.
- [17] Yamazaki CM, Kadoya Y, Hozumi K, Okano-Kosugi H, Asada S, Kitagawa K, et al. A collagen-mimetic triple helical supramolecule that evokes integrin-dependent cell responses. *Biomaterials* 2010;31:1925–34.
- [18] Krishna OD, Jha AK, Jia X, Kiick KL. Integrin-mediated adhesion and proliferation of human MSCs elicited by a hydroxyproline-lacking, collagen-like peptide. *Biomaterials* 2011;32:6412–24.
- [19] Hernandez-Gordillo V, Chmielewski J. Mimicking the extracellular matrix with functionalized, metal-assembled collagen peptide scaffolds. *Biomaterials* 2014;35:7363–73.
- [20] Peng YY, Yoshizumi A, Danon SJ, Glattauer V, Prokopenko O, Mirochnitchenko O, et al. A Streptococcus pyogenes derived collagen-like protein as a non-cytotoxic and non-immunogenic cross-linkable biomaterial. *Biomaterials* 2010;31:2755–61.
- [21] An B, Kaplan DL, Brodsky B. Engineered recombinant bacterial collagen as an alternative collagen-based biomaterial for tissue engineering. *Front Chem* 2014;2.
- [22] Majsterek I, McAdams E, Adachi E, Dhume ST, Fertala A. Prospects and limitations of the rational engineering of fibrillar collagens. *Protein Sci* 2003;12:2063–72.
- [23] Que R, Mohraz A, Da Silva NA, Wang S-W. Expanding functionality of recombinant human collagen through engineered non-native cysteines. *Biomacromolecules* 2014;15:3540–9.
- [24] Jabaiah A, Wang X, Raman SK, Ragan R, Da Silva NA, Wang S-W. Nanoscale architecture and cellular adhesion of biomimetic collagen substrates. *J Biomater Appl* 2014;28:1354–65.
- [25] Hynes RO. Integrins - versatility, modulation, and signaling in cell-adhesion. *Cell* 1992;69:11–25.
- [26] Barczyk M, Carracedo S, Gullberg D. Integrins *Cell Tissue Res* 2010;339:269–80.
- [27] Kim JK, Xu Y, Xu X, Keene DR, Gurusiddappa S, Liang X, et al. A novel binding site in collagen type III for integrins alpha1beta1 and alpha2beta1. *J Biol Chem* 2005;280:32512–20.
- [28] Raynal N, Hamaia SW, Siljander PRM, Maddox B, Peachey AR, Fernandez R, et al. Use of synthetic peptides to locate novel integrin alpha(2)beta(1)-binding motifs in human collagen III. *J Biological Chem* 2006;281:3821–31.
- [29] Knight CG, Morton LF, Peachey AR, Tuckwell DS, Farndale RW, Barnes MJ. The collagen-binding A-domains of integrins alpha(1)beta(1) and alpha(2)beta(1) recognize the same specific amino acid sequence, GFOGER, in native (triple-helical) collagens. *J Biol Chem* 2000;275:35–40.
- [30] Persikov AV, Ramshaw JAM, Kirkpatrick A, Brodsky B. Amino acid propensities for the collagen Triple-Helix. *Biochemistry* 2000;39:14960–7.
- [31] Larsen LSZ, Wassman CD, Hatfield GW, Lathrop RH. Computationally optimized DNA assembly of synthetic genes. *Int J Bioinform Res Appl* 2008;4:324–36.
- [32] Liu H, Naismith J. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol* 2008;8:91.
- [33] Gallant N, Garcia A. Quantitative analyses of cell adhesion Strength. In: Coutts A, editor. *Adhesion protein protocols*. Humana Press; 2007. p. 83–95.
- [34] R Core Team. R: A Language and Environment for statistical computing. Vienna, Austria. 2013.
- [35] Bella J, Eaton M, Brodsky B, Berman HM. Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science* 1994;266:75–81.
- [36] Williams KE, Olsen DR. Matrix metalloproteinase-1 cleavage site recognition and binding in full-length human type III collagen. *Matrix Biol* 2009;28:373–9.
- [37] Kuivaniemi H, Tromp G, Prockop DJ. Mutations in collagen genes – causes of rare and some common diseases in humans. *Faseb J* 1991;5:2052–60.
- [38] Emsley J, Knight CG, Farndale RW, Barnes MJ. Structure of the integrin alpha2beta1-binding collagen peptide. *J Mol Biol* 2004;335:1019–28.
- [39] Kadler K, Holmes D, Trotter J, Chapman J. Collagen fibril formation. *Biochem J* 1996;316:1–11.
- [40] Emsley J, Knight CG, Farndale RW, Barnes MJ, Liddington RC. Structural basis of collagen recognition by integrin alpha2beta1. *Cell*. 2000;101:47–56.
- [41] Heino J. The collagen family members as cell adhesion proteins. *Bioessays* 2007;29:1001–10.
- [42] Pfaff M, Aumailley M, Specks U, Knolle J, Zerwes HG, Timpl R. Integrin and Arg Gly Asp dependence of cell-adhesion to the native and unfolded triple-helix of collagen type-VI. *Exp Cell Res* 1993;206:167–76.
- [43] Reyes CD, Garcia AJ. Engineering integrin-specific surfaces with a triple-helical collagen-mimetic peptide. *J Biomed Mater Res Part A* 2003;65A:511–23.
- [44] Siljander PR, Hamaia S, Peachey AR, Slatter DA, Smethurst PA, Ouwehand WH, et al. Integrin activation state determines selectivity for novel recognition sites in fibrillar collagens. *J Biol Chem* 2004;279:47763–72.
- [45] Tromp G, Kuivaniemi H, Stacey A, Shikata H, Baldwin CT, Jaenisch R, et al. Structure of a full-length cDNA clone for the prepro alpha 1(I) chain of human type I procollagen. *Biochem J* 1988;253:919–22.
- [46] Leitinger B. Transmembrane collagen receptors. *Annu Rev Cell Dev Biol* 2011;27:265–90.
- [47] Xu H, Bihan D, Chang F, Huang PH, Farndale RW, Leitinger B. Discoidin domain receptors promote $\alpha 1\beta 1$ - and $\alpha 2\beta 1$ -integrin mediated cell adhesion to collagen by enhancing integrin activation. *PLoS One* 2012;7. e52209.
- [48] Aota S, Nomizu M, Yamada KM. The short amino acid sequence Pro-His-Ser-Arg-Asn in human fibronectin enhances cell-adhesive function. *J Biol Chem* 1994;269:24756–61.
- [49] Petrie TA, Capadona JR, Reyes CD, Garcia AJ. Integrin specificity and enhanced cellular activities associated with surfaces presenting a recombinant fibronectin fragment compared to RGD supports. *Biomaterials* 2006;27:5459–70.
- [50] Cheng S-L, Lai C-F, Blystone SD, Avioli LV. Bone mineralization and osteoblast differentiation are negatively modulated by integrin $\alpha \beta 3$. *J Bone Mineral Res* 2001;16:277–88.
- [51] Humphries JD, Wang P, Streuli C, Geiger B, Humphries MJ, Ballestrem C. Vinculin controls focal adhesion formation by direct interactions with talin and actin. *J Cell Biol* 2007;179:1043–57.
- [52] Zamir E, Geiger B. Molecular complexity and dynamics of cell-matrix adhesions. *J Cell Sci* 2001;114:3583–90.
- [53] Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012;9:671–5.