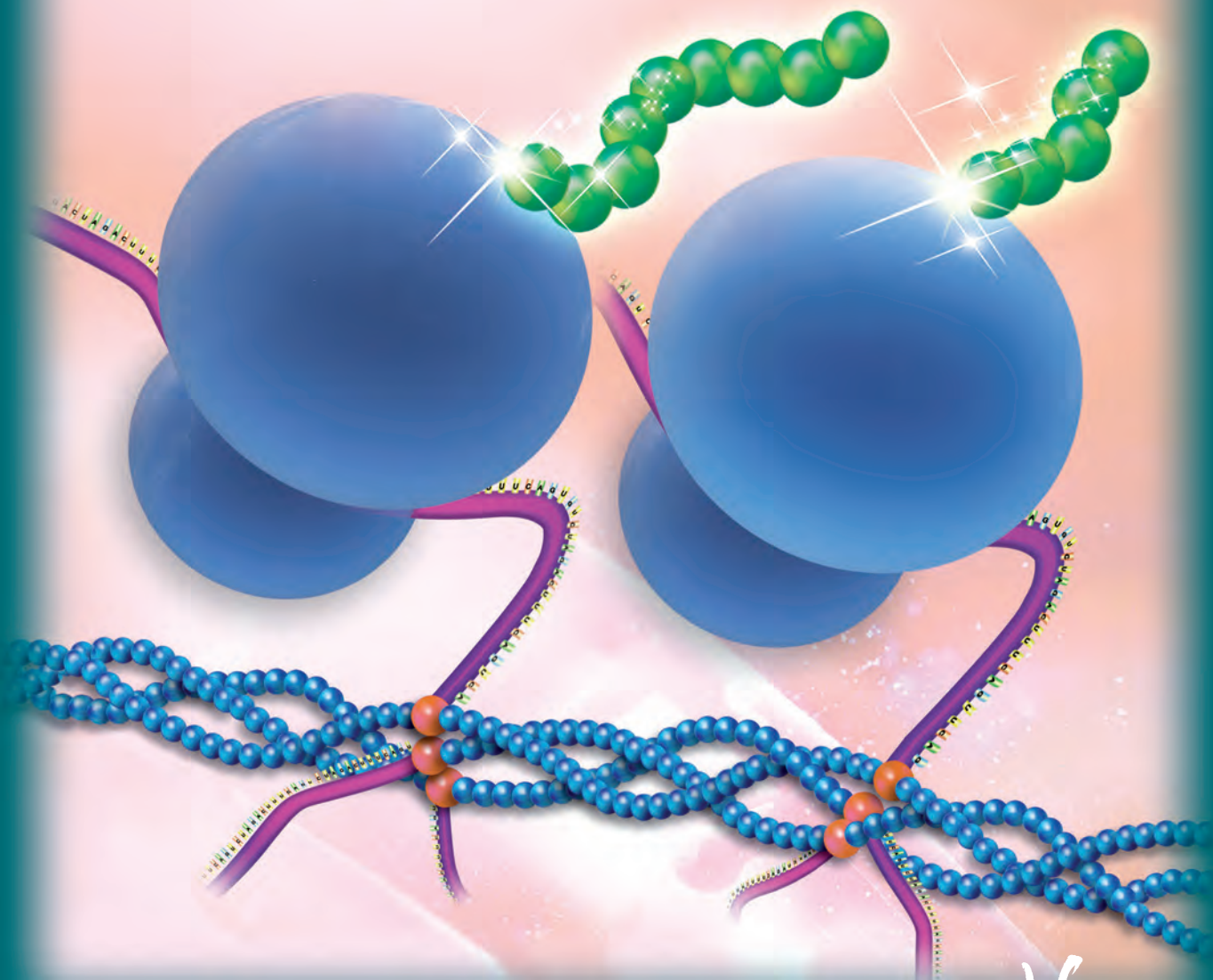


A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEM **BIO** CHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY



10/2015

Chemistry & *Life* Sciences

A Journal of

**Minireview:** The Utilities of Chemical Reactions and Molecular Tools for O-GlcNAc Proteomic Studies (E. J. Kim)**Highlight:** Recent Efforts towards a Systems-Level Hypothesis for an Early RNA World (G. Springsteen)www.chembiochem.org

WILEY-VCH

A Recombinant Collagen–mRNA Platform for Controllable Protein Synthesis

Liping Sun,^[a] Yunjing Xiong,^[a] Anat Bashan,^[b] Ella Zimmerman,^[b] Shirley Shulman Daube,^[b] Yoav Peleg,^[b] Shira Albeck,^[b] Tamar Unger,^[b] Hagith Yonath,^[c, d] Miri Krupkin,^[b] Donna Matzov,^[b] and Ada Yonath^{*[b]}

We have developed a collagen–mRNA platform for controllable protein production that is intended to be less prone to the problems associated with commonly used mRNA therapy as well as with collagen skin-healing procedures. A collagen mimic was constructed according to a recombinant method and was used as scaffold for translating mRNA chains into proteins. Cysteines were genetically inserted into the collagen chain at positions allowing efficient ribosome translation activity while minimizing mRNA misfolding and degradation. Enhanced green fluorescence protein (eGFP) mRNA bound to collagen was successfully translated by cell-free *Escherichia coli* ribosomes. This system enabled an accurate control of specific protein synthesis by monitoring expression time and level. Luciferase–mRNA was also translated on collagen scaffold by eukaryotic cell extracts. Thus we have demonstrated the feasibility of controllable protein synthesis on collagen scaffolds by ribosomal machinery.

Local protein deficiency is a main cause of medical problems ranging from local injury, wounds or surgery that can lead to death to disability or genetic disorders such as cystic fibrosis (CF).^[1] Direct injection of the relevant proteins into the tissue might require repeated injections that could cause severe side effects, including initial burst release.^[2] Other strategies for tissue regeneration include cell transplantation and gene therapy. A prominent problem associated with autologous cells im-

plants is the difficulty in harvesting a sufficient amount of cells. Moreover, allogeneic and xenogeneic cells possess risks of viral infection and can be immunogenic. Similarly, gene therapy with DNA plasmids or viral vectors has potential disadvantages, including poor control of dosage and duration, low gene-transfer efficiency, endogenous virus recombination, oncogenic effects, and unexpected immune responses.^[3] Thus, an efficient local protein synthesis platform for wound areas, deficient tissues or damaged organs should help regain proper functionality as well as promote tissue repair.

Ribosomes are the universal molecular machines that perform protein biosynthesis according to the genetic code provided to them by messenger RNA (mRNA). So far, in vivo on-site-induced protein biosynthesis by ribosomes has hardly been attempted. Nanoparticles encapsulating luciferase-encoding plasmid and *Escherichia coli* extract are capable of autonomous synthesis of protein in vitro and in vivo.^[4] mRNA vectors have also been used as templates for protein synthesis rather than DNA plasmid due to RNA's low toxicity, low immunogenicity, and high expression efficacy in vivo. It has been shown that modified RNA encoding human vascular endothelial growth factor-A (VEGF-A) led to marked improvement in heart function in a mouse myocardial infarction model.^[5]

Collagen, which is widely used for tissue engineering, is the main structural protein of various connective tissues.^[6] All collagen molecules are made up of three polypeptide strands, twisted together into a right-handed triple helix. A distinctive feature of collagen is the repeated GXY sequence, in which G represents glycine and X and Y may be any amino acid. The glycines are located in the interior of the helix, and the amino acids at the Y positions are located at the helix surface.^[7] Hence their side chains may be chemically modified without perturbing the stable helical structure (Figure S1 in the Supporting Information). Natural collagen can suffer from contamination by infectious agents, heterogeneity, potential immunogenicity, loss of structural integrity, and product standardization.^[8] An efficient recombinant collagen should be free of most of these complications, as this is less immunogenic and offers excellent homogeneity. More importantly, collagen gene sequence can be designed according to specific requirements.^[9]

The use of immobilized mRNA has been reported previously. Solid surfaces, that is, chip surfaces coated with neutravidin were used for smFRET studies^[10] and streptavidin-coated beads were useful for other applications.^[11] However, such systems are less suitable for in situ tissue repair, which requires soft car-

[a] Dr. L. Sun, Y. Xiong
Department of Biomaterials, College of Materials, Xiamen University
422, Siming South Road, Xiamen 361005 (China)

[b] Dr. A. Bashan, Dr. E. Zimmerman, Dr. S. Shulman Daube, Dr. Y. Peleg,
Dr. S. Albeck, Dr. T. Unger, M. Krupkin, D. Matzov, Prof. A. Yonath
Weizmann Institute of Science
234 Herzl Street, Rehovot 7610001 (Israel)
E-mail: ada.yonath@weizmann.ac.il

[c] H. Yonath
Sheba Medical Center
1 Sheba Street, Tel Hashomer 52621 (Israel)

[d] H. Yonath
Sackler School of Medicine, Tel Aviv University
10 Levanon Street, Tel Aviv 69978 (Israel)

Supporting information for this article is available on the WWW under
<http://dx.doi.org/10.1002/cbic.201500205>.

© 2015 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA.
This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

riers. Hence, we focus on mRNA immobilization on natural biological scaffolds, such as collagen.

Benefiting from knowledge of the accurate ribosomal machinery^[12] alongside collagen's chemical and structural properties,^[7] we have developed a new, controllable, recombinant collagen–mRNA system for local protein production. We designed and cloned a collagen gene into which cysteine codons were inserted at positions allowing efficient mRNA binding and ribosome translation activity (Figure S2). As natural collagen contains no cysteine in the GXY repeats region, we inserted two glycine-proline-cysteine (GPC) triplets into our recombinant collagen to provide a functional sulfhydryl group for further reaction with amine-modified mRNA. A bacteriophage T4 fibrin foldon domain at the C terminus serves as a nucleation site to facilitate the correct folding of the collagen triple helix.^[13]

Enhanced green fluorescence protein (eGFP) was used as a reporter protein to investigate the feasibility of controllable protein synthesis on collagen scaffold by the ribosomal machinery. mRNA of eGFP was crosslinked to the collagen strand, and eGFP was translated when ribosomes were provided by *E. coli* extract. Thus, here we demonstrate the construction of a new recombinant collagen–mRNA platform for controllable continuous specific protein synthesis.

Our construct contains collagen bound to an N-terminal maltose binding protein (MBP) domain, a hexahistidine tag, a TEV cleavage site, a Flag tag, and a foldon domain at the C terminus (Figure 1). In principle, our construct could be of a large range of lengths (repeats of the collagen portion), according to need.

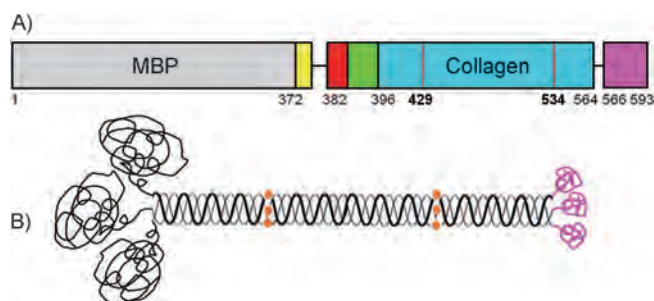


Figure 1. A) Schematic representation of the recombinant MBP–collagen. The domains are not drawn to scale; numbers show amino acid positions, with the cysteine residues marked in bold. Yellow, red, green, and purple represent the His₆-tag, TEV cleavage site, Flag-tag, and foldon domain, respectively. B) Triple-helix representation of MBP–collagen. Left: N-terminal folded MBP domains are (black), right C-terminal foldon domains (purple). Orange spheres represent cysteines in collagen strands.

The tunable distance between the inserted cysteines can be designed to maximize the efficiency of the platform, which is designed such that 1) the SH– insertions should minimize mRNA aggregation and/or misfolding and 2) the platform's multiple sizes are fully under control. This distance, which controls the usability of the platforms (as it directly influences the possible interactions between the bound mRNA chains), depends on the structure of the recombinant modified collagen

scaffold. Hence we verified that it maintains the triple helicity of natural collagen (see below).

An additional parameter is the ribosome size, which is different in prokaryotes and eukaryotes. In this study it was set to 28.8 nm (105 amino acids of the collagen chain, after verifying the formation of triple helices), which is sufficient for *E. coli* ribosome binding; as its longest dimension is about 25 nm. The chimeric protein MBP–collagen (Figure S2) was expressed in *E. coli* BL21(DE3) cells and was found to fold as a trimer with a calculated molecular weight (MW) of 184.5 kDa (Figure 2A,

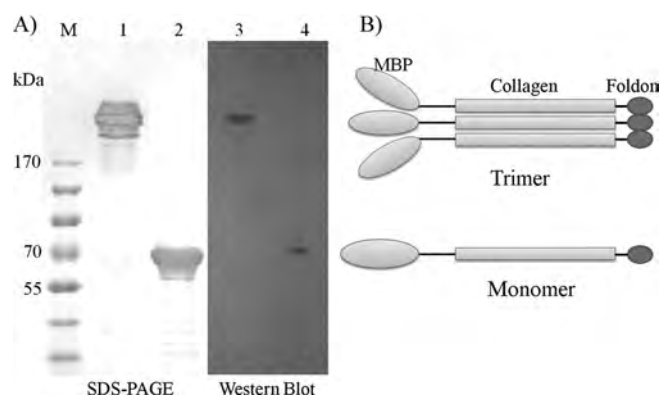


Figure 2. A) SDS-PAGE and Western blot of collagen trimer (unheated sample, lanes 1 and 3) and monomer (heated sample, lanes 2 and 4). B) Schematic representation of the collagen trimer and monomer structures. The calculated molecular weight of the collagen monomer is 61.5 kDa.

lane 1). After 5 min of heating at 90 °C under reducing conditions, a monomer band appeared (lane 2). These observations were further supported by western-blot analysis with anti His-tag antibody. Lanes 3 and 4 represent trimeric and monomeric MBP–collagen, respectively. These data demonstrated that the collagen trimer was converted to unfolded monomer under reducing condition. Schematic structures of the trimer and monomer are shown in Figure 2B.

The thermostability of MBP–collagen was further studied by circular dichroism (CD) spectroscopy. The typical peaks for a collagen triple helix, namely a negative peak at 196 nm and a positive peak at 224 nm,^[14] were observed (Figure 3A). When a sample was heated from 20 to 90 °C, the heights of these peaks decreased gradually, thereby indicating denaturation of the collagen. The melting temperature (T_m) was around 55 °C (Figure 3B). Given the T_m of human type I collagen (<36 °C), the recombinant MBP–collagen is highly stable. This thermal stability is attributed to the foldon domain and the multiple GPP repeats in MBP–collagen.^[10] We also investigated the thermostability of MBP–collagen by SDS-PAGE (Figure 3C). MBP–collagen samples were incubated at different temperatures (30–90 °C) for 5 min and analyzed by 10% SDS-PAGE. We found that melting took place at 50 °C, in accordance with the CD result.

We designed a protein synthesis platform in which mRNA was covalently bound to a collagen scaffold. In this system, the 3'-end of the desired mRNA is modified with an amine and

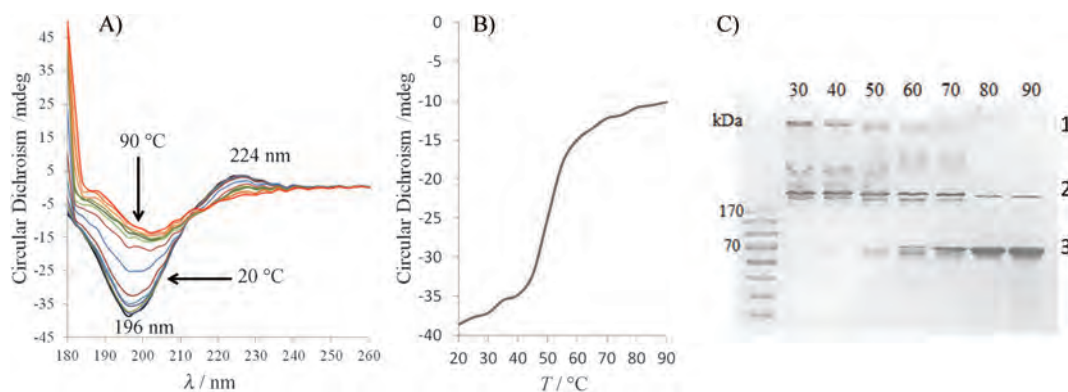


Figure 3. A) CD spectra of MBP-collagen taken at 20–90 °C. B) Melting curve of MBP-collagen at 196 nm. C) SDS-PAGE of MBP-collagen at 30–90 °C. The positions of the loading well, trimer, and monomer are marked 1, 2 and 3, respectively.

crosslinked to engineered cysteines on MBP-collagen by using the hetero-bifunctional crosslinker sulfo-SMCC.^[15] Crosslinking the mRNA through its 3'-end rather than the 5' might increase the accessibility of the 5'-end of the mRNA to ribosomes.

A titration experiment was performed to determine the optimal molar ratio of collagen to RNA. Accordingly, crosslinking was performed at a 20:1 collagen:mRNA ratio in order to ensure that no free RNA remained in the solution (Figure 4A).

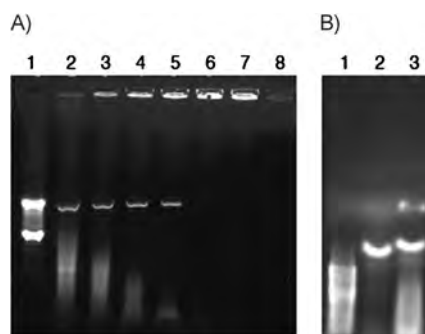


Figure 4. Titration experiment aimed at resolving RNA-collagen complexes by gel electrophoresis. A) Unmodified RNA (lane 1) and collagen (lane 8) were incubated at a final molar ratio of collagen to RNA of 1:1, 2:1, 4:1, 8:1, 16:1 and 32:1 (lanes 2–7, respectively) and resolved on non-denaturing 1% agarose gel in 0.5% × TBE buffer. B) Resolution of RNA-collagen complexes by SDS denaturing agarose gel. Complexes formed between unmodified RNA and collagen (lane 3), or modified RNA and crosslinked collagen (lane 2) were resolved and compared to unmodified RNA only (lane 1). Bands were imaged after ethidium bromide staining.

The products of the two-step crosslinking reaction were resolved by SDS denaturing agarose gel electrophoresis (Figure 4B). Nevertheless, it is conceivable that part of the mRNA was nonspecifically bound to the collagen. Indeed, we found that under denaturing conditions the collagen-RNA complexes migrate into the gel, providing a comparison between the crosslinking and nonspecific collagen-RNA complexes (Figure 4B lanes 2 and 3, respectively). The product of the crosslinked reaction migrated slightly faster than the nonspecific complex. In addition, the mRNA was better protected from degradation in the crosslinked reaction, consequently practical-

ly no smeared band was observed in the crosslinked RNA-collagen complex (Figure 4B, lane 2).

In order to test whether the mRNA that is part of a collagen-RNA complex could undergo translation by ribosomes, complexes formed in the crosslinking reaction at a 20:1 collagen:RNA ratio were incubated with *E. coli* ribosome extract S12 as the source of ribosomal machinery, energy and ions required for the protein translation.^[16] The 20:1 ratio ensured that very little free RNA was present in the reaction (Figure 4A). We measured the kinetics of eGFP at 37 °C. The relative fluorescence units (RFU) of eGFP increased gradually over time (Figure 5A), reaching maximal fluorescence intensity after

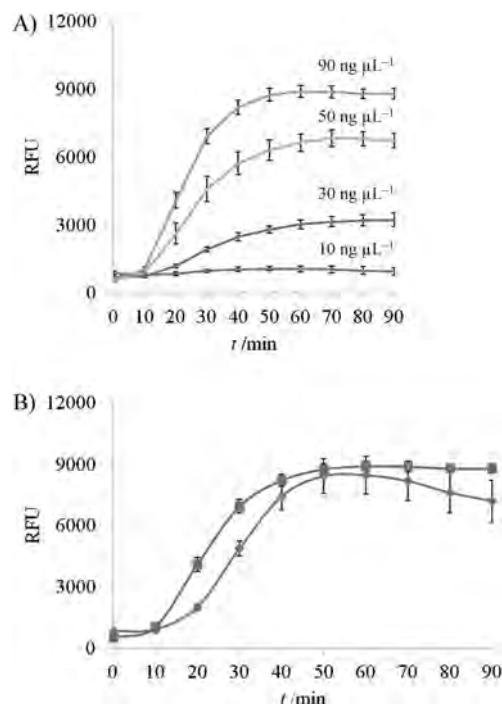


Figure 5. A) Time- and concentration-dependent eGFP biosynthesis on a collagen scaffold ([RNA] = 10–90 ng μL⁻¹). B) Comparison of eGFP expression of RNA (♦) and RNA-collagen (■). All results are representative of three repeated experiments. Relative fluorescence units (RFU) represent the mean ± standard deviations.

50 min of incubation; eGFP expression level was also found to be dependent on RNA concentration. The RNA–collagen complex showed a similar kinetic profile to eGFP expression from non-crosslinked RNA (Figure 5B). These results indicated that eGFP was successfully synthesized while associated with the collagen scaffold.

To expand the application of the collagen–mRNA platform in eukaryotic systems, we investigated luciferase synthesis on a collagen scaffold with eukaryotic translation extracts (rabbit reticulocyte lysate) instead of *E. coli* ribosome extract. The luciferase mRNA–collagen system showed a significantly higher luminescence intensity (Table S2) than the negative control sample, but a lower one than a pure mRNA system; this might be due to mRNA degradation.

So far mRNA treatment is in its embryonic phase because of issues concerning its stability and its tendency to interact with other mRNA chains, self-fold, or aggregate. Our design should minimize these problems because we can select the location of the mRNA on the collagen platform in a fashion that limits its mobility. One way is to position neighboring mRNA chains at intervals compatible with the size of the ribosome, so that the translating ribosome can occupy the space between them (Figure S3). Nevertheless, insertion of mRNA into cells could lead to side effects such as activating the mammalian innate immune response system; this might be solved by incorporating modified nucleosides in mRNA.^[17]

The main advantage of our platform is the controllability of its size, shape, and composition. Thus, the sequences of the mRNA chains can be preselected as well as readily exchanged or modified during treatment according to the medical purpose. Other advantages stem from the properties of the recombinant collagen in the proposed platform. First, its sequence can be designed so as not to include any of the collagen antigenic determinants. Second, we have full control of its size—from very large to relatively small—so that it will be endocytosed by protein-deficient cells. In addition, the concentration of synthesized protein can be controlled by the amount of the cell-free system in local surface applications, such as wounds or by the concentration of the added mRNA platform. Moreover, multiple mRNA chains encoding different proteins can be linked to a single collagen scaffold, or to an assembly of scaffolds, thus enabling a combination therapy of multiple protein and peptide drugs. Consequently, the construct can be used to synthesize diverse functional proteins (e.g., growth factors that can promote tissue repair). As this platform enables accurate control of the amounts of specific proteins by selecting the required mRNA sequence, expression time and expression level, it could also provide a platform for personalized medicine.

Experimental Section

The collagen gene was optimized for *E. coli* codon usage and cloned into pUC57 plasmid. Collagen was subcloned into a pET-MBP-TevH plasmid by transfer PCR,^[18] and the sequence was verified by DNA sequencing. The constructed plasmid was transformed into chemically competent *E. coli* BL21(DE3) cells. The cells were

grown at 37 °C in lysogeny broth. Protein expression was induced by isopropyl- β -thiogalactopyranoside (IPTG). MBP–collagen was purified by nickel affinity column chromatography and gel filtration column chromatography on a fast protein liquid chromatography system. SDS-PAGE and western blot were performed to confirm the collagen expression.

EGFP or the luciferase gene was amplified by PCR. PCR products were transcribed to mRNA by T7 RNA polymerase. EGFP RNA was modified with 3'-NH₂-ATP by using poly-A polymerase. Amine-modified RNA was conjugated to sulfhydryl-containing MBP–collagen by sulfo-SMCC. RNA, collagen and RNA–collagen complex were analyzed by 1% agarose gel electrophoresis with ethidium bromide staining. EGFP and luciferase syntheses on collagen scaffolds were performed by using an *E. coli* cell-free translation assay and rabbit reticulocyte lysate system, respectively.

S12 cell-free extract was prepared from *E. coli* strain BL21(DE3) according to a previous report.^[16] An aqueous solution containing amino acid mix, ATP, *E. coli* tRNA mixture, S12 extract, and other components was prepared. Collagen–mRNA or eGFP–mRNA (250 ng μ L⁻¹) was added to this solution at 37 °C, thereby commencing the translation reaction. Fluorescence data were collected on a Synergy HT Microplate Reader (BioTek Instruments; λ_{ex} = 485 nm, λ_{em} = 528 nm). Collagen–luciferase mRNA complex or luciferase–mRNA was translated in rabbit reticulocyte lysate system at 30 °C for 1.5 h. Luciferase activity was analyzed by using an Infinite 200 Pro microplate reader (Tecan, Männedorf, Switzerland).

Acknowledgements

This work was supported by the Kimmelman Center for macromolecular assemblies, the Weizmann Institute and the National Natural Science Foundation of China (81171453). Cloning, expression and purification of recombinant collagen were performed at the Israel Structural Proteomics Center (ISPC), which is supported by the Dana & Yossie Hollander Center for Structural Proteomics.

Keywords: collagen · mRNA · protein synthesis · ribosomes

- [1] D. C. Gadsby, P. Vergani, L. Csanády, *Nature* **2006**, *440*, 477.
- [2] a) R. M. Boehler, J. G. Graham, L. D. Shea, *BioTechniques* **2011**, *51*, 239; b) Y. Ikada, *J. R. Soc. Interface* **2006**, *3*, 589.
- [3] a) K. Lee, E. A. Silva, D. J. Mooney, *J. R. Soc. Interface* **2011**, *8*, 153–170; b) T. Niidome, L. Huang, *Gene Ther.* **2002**, *9*, 1647.
- [4] A. Schroeder, M. S. Goldberg, C. Kastrop, C. G. Levins, R. Langer, D. G. Anderson, *NanoLetters* **2012**, *12*, 2685.
- [5] L. Zangi, K. O. Lui, A. Gise, Q. Ma, W. Ebina, L. M. Ptaszek, D. Später, H. Xu, M. Tabebordbar, R. Gorbatov, B. Sena, M. Nahrendorf, D. M. Briscoe, R. A. Li, A. J. Wagers, D. J. Rossi, W. T. Pu, K. R. Chien, *Nat. Biotechnol.* **2013**, *31*, 898.
- [6] a) J. Glowacki, S. Mizuno, *Biopolymers* **2008**, *89*, 338; b) B. Weng, X. Liu, R. Shepherd, G. G. Wallace, *Synth. Met.* **2012**, *162*, 1375.
- [7] a) W. Traub, A. Yonath, D. M. Segal, *Nature* **1969**, *221*, 914; b) A. Yonath, W. Traub, *J. Mol. Biol.* **1969**, *43*, 461.
- [8] a) M. D. Shoulders, R. T. Raines, *Annu. Rev. Biochem.* **2009**, *78*, 929; b) L. Cen, W. Liu, L. Cui, W. Zhang, Y. Cao, *Pediatr. Res.* **2008**, *63*, 492.
- [9] a) Y. Liu, M. Griffith, M. A. Watsky, J. V. Forrester, L. Kuffova, D. Grant, K. Merrett, D. J. Carlsson, *Biomacromolecules* **2006**, *7*, 1819; b) D. Olsen, C. Yang, M. Bodo, R. Chang, S. Leigh, J. Baez, D. Carmichael, M. Perala, E. R. Hamalainen, M. Jarvinen, J. Polarek, *Adv. Drug Delivery Rev.* **2003**, *55*, 1547.
- [10] A. Tsai, G. Kornberg, M. Johansson, J. Chen, J. D. Puglisi, *Cell Rep.* **2014**, *7*, 1521.
- [11] M. Biyani, Y. Husimi, N. Nemoto, *Nucleic Acids Res.* **2006**, *34*, e140.

- [12] A. Bashan, I. Agmon, R. Zarivach, F. Schluenzen, J. Harms, R. Berisio, H. Bartels, F. Franceschi, T. Auerbach, H. A. S. Hansen, E. Kossoy, M. Kessler, A. Yonath, *Mol. Cell* **2003**, *11*, 91.
- [13] S. Frank, R. A. Kammerer, D. Mechling, T. Schulthess, R. Landwehr, J. Bann, Y. Guo, A. Lustig, H. P. Bächinger, J. Engel, *J. Mol. Biol.* **2001**, *308*, 1081.
- [14] K. A. Piez, M. R. Sherman, *Biochemistry* **1970**, *9*, 4129.
- [15] a) G. T. Hermanson in *Bioconjugate Techniques*, 2nd ed., Academic Press, San Diego, **2008**, pp. 283–286; b) E. Hemaprabha, *J. Pharm. Sci. Innov.* **2012**, *1*, 22; c) B. A. Tannous, N. H. L. Chiu, T. K. Christopoulos, *Anal. Chim. Acta* **2002**, *459*, 169; d) B. A. R. Williams, J. C. Chaput, *Curr. Protoc. Nucleic Acid Chem.* **2010**, unit 4.41.
- [16] T. W. Kim, J. W. Keum, I. S. Oh, C. Y. Choi, C. G. Park, D. M. Kim, *J. Biotechnol.* **2006**, *126*, 554.
- [17] K. Karikó, H. Muramatsu, F. A. Welsh, J. Ludwig, H. Kato, S. Akira, D. Weissman, *Mol. Ther.* **2008**, *16*, 1833.
- [18] a) Y. Peleg, T. Unger in *Methods in Molecular Biology, Vol. 426: Application of High-Throughput Methodologies to the Expression of Recombinant Proteins in E. coli* (Eds.: B. Kobe, M. Guss, T. Huber), Humana, Totowa, **2008**, pp. 197–208; b) A. Erijman, A. Dantes, R. Bernheim, J. M. Shifman, Y. Peleg, *J. Struct. Biol.* **2011**, *175*, 171.

Manuscript received: April 21, 2015

Accepted article published: April 30, 2015

Final article published: May 26, 2015
