ADVANCED SCIENCE Open Access
www.advancedscience.com

A Novel Gene Synthesis Platform for Designing Functional Protein Polymers

Toshimasa Homma, Rie Yamamoto, Lily Zuin Ping Ang, Alaa Fehaid, and Mitsuhiro Ebara*

Recombinant protein polymers with repeat sequences of specific amino acids can be regarded as sustainable functional materials that can be designed using genetic engineering. However, synthesizing genes encoding these proteins is significantly time-consuming and labor-intensive owing to the difficulty of using common gene synthesis tools, such as restriction enzymes and PCR primers. To overcome these obstacles, a novel method is proposed herein: seamless cloning of rolling-circle amplicons (SCRCA). This method involves one-pot preparation of repetitive-sequence genes with overlapping ends for cloning, facilitating the easy construction of the desired recombinants. SCRCA is used to synthesize 10 genes encoding hydrophilic resilin-like and hydrophobic elastin-like repeat units that induce liquid-liquid phase separation. SCRCA shows higher transformation efficiency and better workability than conventional methods, and the time and budget required for SCRCA are comparable to those required for non-repetitive-sequence gene synthesis. Additionally, SCRCA facilitates the construction of a repeat unit library at a low cost. The library shows considerably higher diversity than that of the current state-of-the-art method. By combining this library construction with the directed evolution concept, an elastin-like protein polymer with the desired functions can be rapidly developed. SCRCA can greatly accelerate research on protein polymers.

1. Introduction

Diverse repetitive amino acid sequences exist in nature, each possessing distinctive functions.[1–5] Recombinant protein polymers inspired by these characteristics have attracted attention as sustainable and biocompatible materials synthesized using environment-friendly processes.[6,7] The applications of these polymers are expanding to high-toughness fibers,[8,9] ultra-high elastic materials,[10,11] cell culture substrates,[12,13] bioseparation,[14,15] drug delivery,[16,17] proton-conducting membranes,[18] and smart cells.[19] The ability to freely design repeat units and chain lengths via genetic recombination technology represents a major advantage during the development of new protein-polymer materials.[3,20–22] To create new functional protein polymers, copolymers, and block copolymers consisting of two repeat units with different properties have been rationally designed.[23–26] Moreover, polymer libraries of different lengths and repetitive sequences have also been constructed for identifying mutants with novel functions through high-throughput screening.[27–29]

However, the synthesis of repetitive-sequence genes encoding protein polymers requires time-consuming methods due to the low number of unique parts whose sequences can be recognized using restriction enzymes or PCR primers.[30,31] Although a codon-scrambling algorithm has been developed to improve gene complexity,[32] screening for useful sequences is costly because it requires many custom oligo DNAs. Rolling-circle amplification (RCA) is promising because it can simultaneously synthesize multiple repetitive-sequence genes with different repeat numbers.[33,34] However, the previously reported RCA methods have several shortcomings, including the difficulty of isolating gene

T. Homma, R. Yamamoto, L. Z. P. Ang, A. Fehaid
Division of Chemical Engineering and Biotechnology
National Institute of Technology
Ichinoseki College, Takanashi
Hagisho, Ichinoseki, Iwate 021-8511, Japan
E-mail: hommatsh@ichinoseki.ac.jp

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/advs.202410903

© 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/advs.202410903

T. Homma, R. Yamamoto, A. Fehaid, M. Ebara
Research Center for Macromolecules and Biomaterials
National Institute for Materials Science (NIMS)
1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan

A. Fehaid
Forensic Medicine and Toxicology Department
Faculty of Veterinary Medicine
Mansoura University
Dakahlia, Mansoura 35516, Egypt

M. Ebara
Graduate School of Pure and Applied Sciences
University of Tsukuba
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan

M. Ebara
Graduate School of Advanced Engineering
Tokyo University of Science
6-3-1 Katsushika-ku, Shinjuku, Tokyo 125-8585, Japan

fragments at a desired length from a library, inability to control the direction of insertion,[33] and requirement for a specific nucleotide sequence,[34] leaving room for improvement in operability and versatility.

We propose a new gene synthesis method, called seamless cloning of rolling-circle amplicons (SCRCA). It consists of four steps: preparation of ssDNA ring templates, RCA reaction, selection of DNA fragment size, and seamless cloning (**Figure 1a**). Seamless cloning simplifies the introduction of DNA fragments into a plasmid vector at the desired insertion position and in a specific direction, although overlapping sequences are required at both ends of the DNA fragments.[35] In SCRCA, these overlapping sequences are added at each end of the gene when the repeat sequence is synthesized, improving the success rate and significantly reducing the time and labor required for gene synthesis. Moreover, through innovative strategies, this method can be used to construct copolymer and block copolymer genes, as well as polymer gene libraries with different repeat units (Figure 1b–d).

In this study, we first synthesized repetitive-sequence genes encoding repeat units of a resilin-like polymer (RLP) and an elastin-like polymer (ELP), the major protein-polymer backbone. Cost analysis revealed that the SCRCA method is more economical than the other methods in terms of repetitive gene synthesis and can be achieved with the same time and cost as the synthesis of non-repetitive sequence genes via a common method. In fact, copolymer sets designed by SCRCA can easily provide sequence and functional insights that can be beneficial for rational polymer designing. Next, we proved that repeat unit libraries with random mutations can be constructed at a low cost via SCRCA. By performing the directed evolutionary experiment using the SCRCA-constructed libraries, we successfully developed a multi-responsive ELP in several months, a task that would have typically required several years. To the best of our knowledge, there are no published reports describing the application of directed evolution for developing protein polymers. Our results confirm that SCRCA is an important platform for the advancement of protein–polymer research.

2. Results and Discussion

2.1. Development of the Gene Amplification Process

For the SCRCA, we devised a new RCA reaction using forward and reverse primers with overlapping sequences at their 5′ ends. The Bst DNA polymerase large fragment was selected as the strand displacement enzyme for RCA because it lacks 5′→3′ exonuclease activity, and therefore, the overlapping sequence of the primer would not be digested.[36,37] Further, its high optimum temperature (60–70 °C)[38,39] is appropriate for amplifying GC-rich sequences, which are often present in major repetitive polypeptides. A cyclized ssDNA with a sequence encoding the desired repeat unit was used as an RCA template. During isothermal DNA amplification using these materials, DNA elongation and strand displacement reactions are expected to occur continuously (**Figure 2a**). Consequently, a mixture of genes with repetitive sequences with different repeat numbers and overlapping sequences at both ends of the repetitive sequences can be prepared.

To verify whether the reaction occurred as expected, isothermal amplification was performed using three ssDNA rings, namely R, E, and RE, that have nucleotide sequences encoding the hydrophilic RLP repeat unit[40] $[(GRGDSPYS)_4]_n$, hydrophobic ELP repeat unit[41] $[(VGVPG)_6]_n$, and RLP-ELP tandem repeat unit $[GRGD\text{-}(SPYSGRGD)_3\text{-}(GVPGVGVPGV)_6\text{-}SPYS]_n$, respectively. RLP and ELP induce phase separation below an upper and above a lower critical solution temperature (UCST and LCST), respectively. We considered that the advanced design of these polymers could contribute to the elucidation and application of the phase separation phenomenon.[42,43] Therefore, we adopted these polymers in our experimental models. Forward and reverse primers with 15-base overlapping sequences at their 5′ ends were used in reactions. All reactions generated a mixture of amplified products of different lengths (Figure 2b–d). The positions of the bands matched the corresponding theoretical values (repeat number × ring size + overlapping sequence bases), indicating that the repetitive-sequence genes were successfully prepared (Figure 2e–g). As amplification products were obtained even when the large RE ring was used, SCRCA may also be suitable for gene synthesis of polymers with long repeat units and copolymers.

Gene length can be selected by performing agarose gel electrophoresis, fractionating the sequences of different lengths, and cutting the desired ones from the gel. However, the conventional RCA products exhibit smear-like bands due to the presence of incomplete double-stranded DNA, suggesting the presence of several sequences other than the targeted gene length.[29,33] In contrast, the RCA product of the method described herein has clearly separated bands. This makes it possible to easily isolate the gene with the target length (Figure 2h). By increasing the run time of agarose gel electrophoresis, the length of the genes isolated could be increased (Figure 2i,j).

2.2. Repetitive-Sequence Gene Construction

To evaluate the probability of constructing a transformant with the desired repetitive-sequence gene, *E. coli* transformation was performed using the In-Fusion seamless cloning system.[44] Homopolymer and copolymer genes *R6*, *E6*, *(RE)3*, *R12*, *E12*, and *(RE)6* were prepared via RCA reactions using R, E, and RE rings (the repetitive-sequence genes are written in italics according to their repeat unit and number; for example, *R6* represents a gene encoding a polypeptide with six repeats of the R ring sequence). Purified DNA fragments were subjected to fusion with the linear vector designed for expression and sequence confirmation (Figure S1, Supporting Information). Block copolymer genes were synthesized via the simultaneous seamless cloning of two DNA fragments (Figure 1b). Here, four block copolymer genes, *R3E3*, *R8E4*, *R6E6*, and *R4E8*, were constructed in parallel by combining *R3*, *R4*, *R6*, *R8*, *E3*, *E4*, *E6*, and *E8* blocks isolated from two RCA reactions using R and E as templates accordingly.

Colony PCR results showed that most transformants had the desired gene length (Figure S2, Supporting Information). Sanger DNA sequencing revealed that 81.4% of the transformants contained the desired gene when the gene length was ≈550 bp (*R6*, *E6*, *(RE)3*, and *R3E3*); even for genes with lengths >1000 bp (*R12*, *E12*, *(RE)6*, *R8E4*, *R6E6*, and *R4E8*), 41.5% of all the

© 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

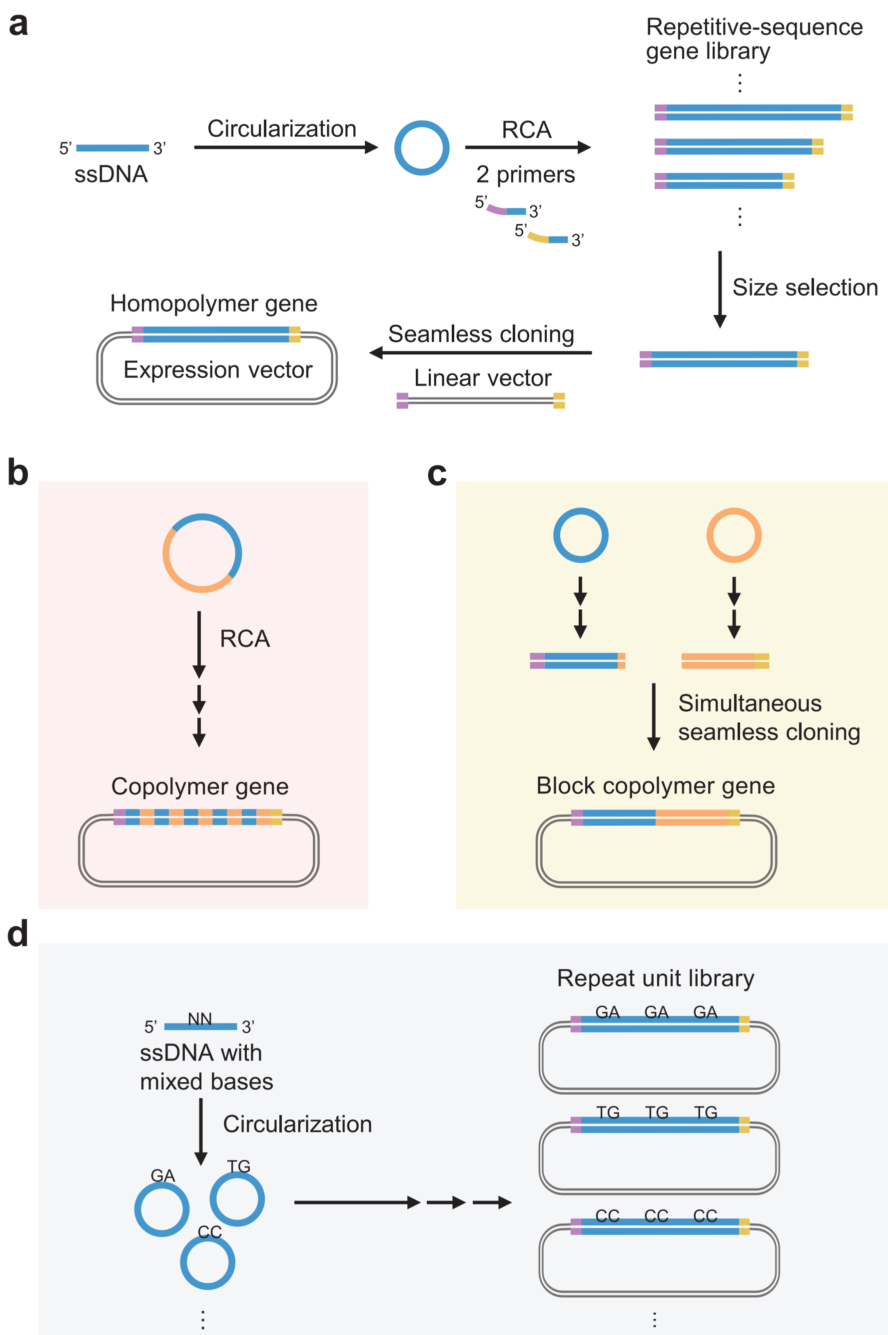


Figure 1. Seamless cloning of rolling-circle amplicons (SCRCA). a) Synthesis of a homopolymer gene using the SCRCA method: 1) ssDNA oligo with a repeat unit sequence is circularized; 2) A repetitive-sequence gene library with different repeat numbers is prepared via RCA (blue line: repetitive sequence; purple and yellow lines: overlapping sequences); 3) Amplified products are separated based on their repeat numbers via agarose gel electrophoresis, and the gene of the desired length is excised; 4) This gene is inserted into an expression vector via seamless cloning. Copolymer and block copolymer genes composed of two different repeat units are synthesized via RCA using a large-ring template b) and simultaneous seamless cloning c), respectively. d) A repeat unit library is constructed using an ssDNA oligo containing mixed bases.

transformants exhibited the desired gene (**Figure 3**). Of the 32 sequences analyzed, there was one large deletion resulting from cloning failure and two deletions resulting from amplification of incomplete templates (Figures S3–S7, Supporting Information). Base substitution or deletion errors were less than 0.02% of >29,000 bases, compared to those of the conventional RCA method.[33]

2.3. Cost Analysis

To evaluate the cost-effectiveness of SCRCA, we calculated the cost of synthesizing the homopolymer gene *E12* alone and three block polymer genes *R8E4*, *R6E6*, and *R4E8* together. Assuming that all conventional methods can be used for synthesis, gene synthesis roadmaps were prepared based on data from previously

© 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

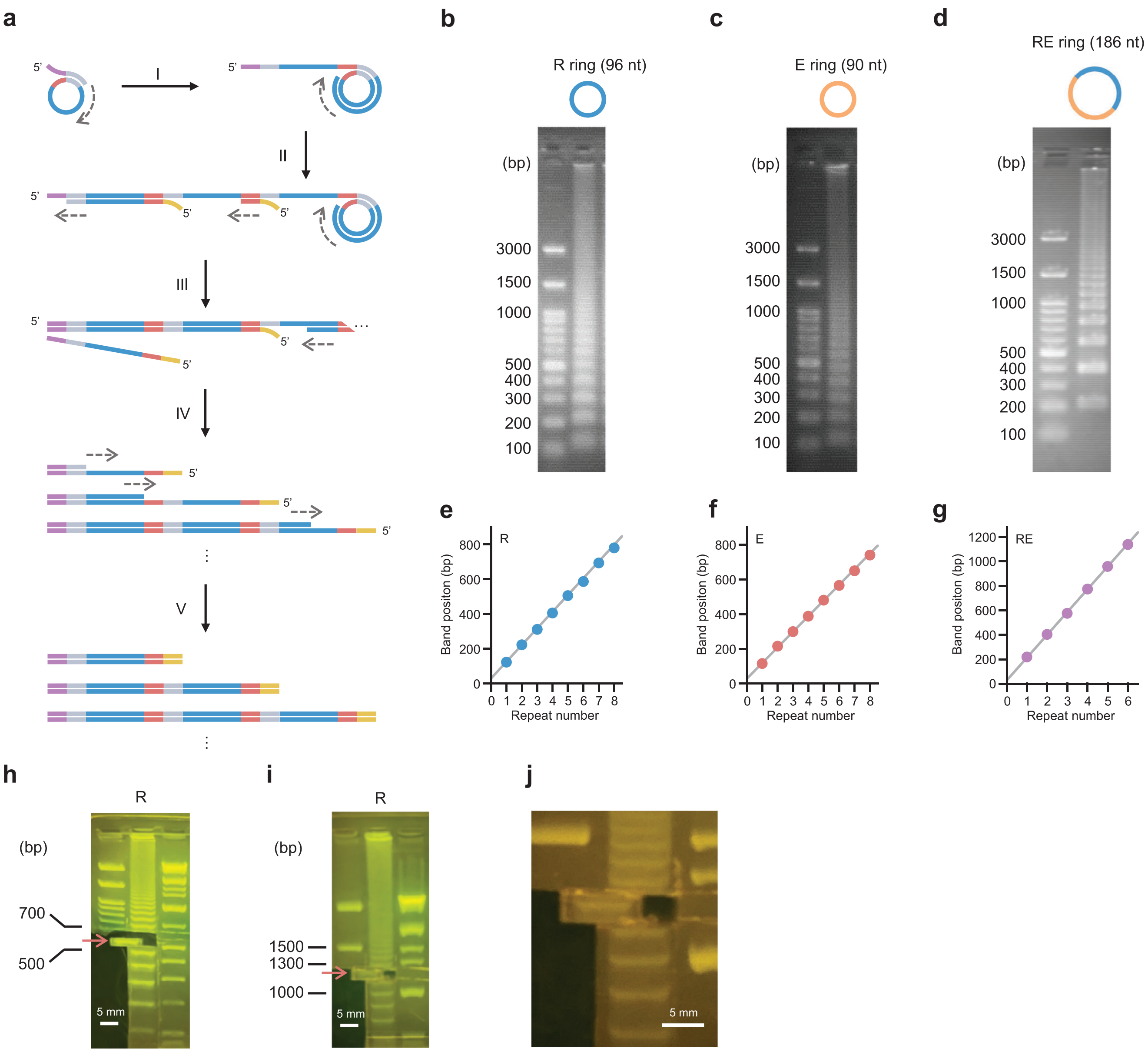


Figure 2. Isothermal rolling-circle amplification (RCA) for seamless cloning of rolling-circle amplicons (SCRCA). a) Schematic of the isothermal amplification. Blue, gray, and red lines: repetitive sequences; purple and yellow lines: overlapping sequences. The 3′ ends of the forward and reverse primers are homologous to the gray and red lines, respectively. I) A DNA chain, in which the complementary sequence of the ssDNA ring is repeated, is elongated from the reverse primer. II) Forward primer attaches to the elongated DNA chain, forming a dsDNA section. III) The nascent chain is dissociated by a subsequent elongation reaction. IV) Primers with complementary sequences attach to the dissociated nascent chain to form a dsDNA fragment. V) These reactions continue to occur to construct a library of repetitive genes with different repeat numbers. The 5′ and 3′ ends of the genes in the library have overlapping sequences derived from the 5′ ends of the forward and reverse primers, respectively. (b–d) Amplicons obtained by RCA using the b) R, c) E, and d) RE rings were separated using 1.5% agarose gel electrophoresis. e–g) Band positions of the amplicons obtained by RCA using the (e) R, (f) E, and (g) RE rings, and the linear function (gray line) of their theoretical lengths (repeat number × ring size + overlapping sequence bases) of these amplicons showed excellent correlation ($R > 0.999$). Each value represents the mean of three replicates, and the standard errors were less than 10 bp for all means. h–j) Two DNA fragments equivalent to the 6 (h) and 12 repeats (i) of R ring sequences were cut from the gels electrophoresed for 70 and 110 min, respectively. Panel (j) is an enlarged view of the cut gel, showing that the gel contained the 12-repeat DNA fragment.

published papers[31–33] (Figures S8 and S9, Supporting Information). Costs for oligo DNA were determined based on the information provided on the websites of a contract oligo DNA synthesis company (Tables S1 and S2, Supporting Information). For reagent costs, we calculated the required costs based on the quantities and concentrations listed in the Materials and Methods section of each paper (Tables S3 and S4, Supporting Information). For example, a reaction solution of overlap-extension rolling circle amplification (OERCA) contains 50 µL of 25 mM dNTPs and requires two reactions per gene synthesis; based on this, we

 © 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

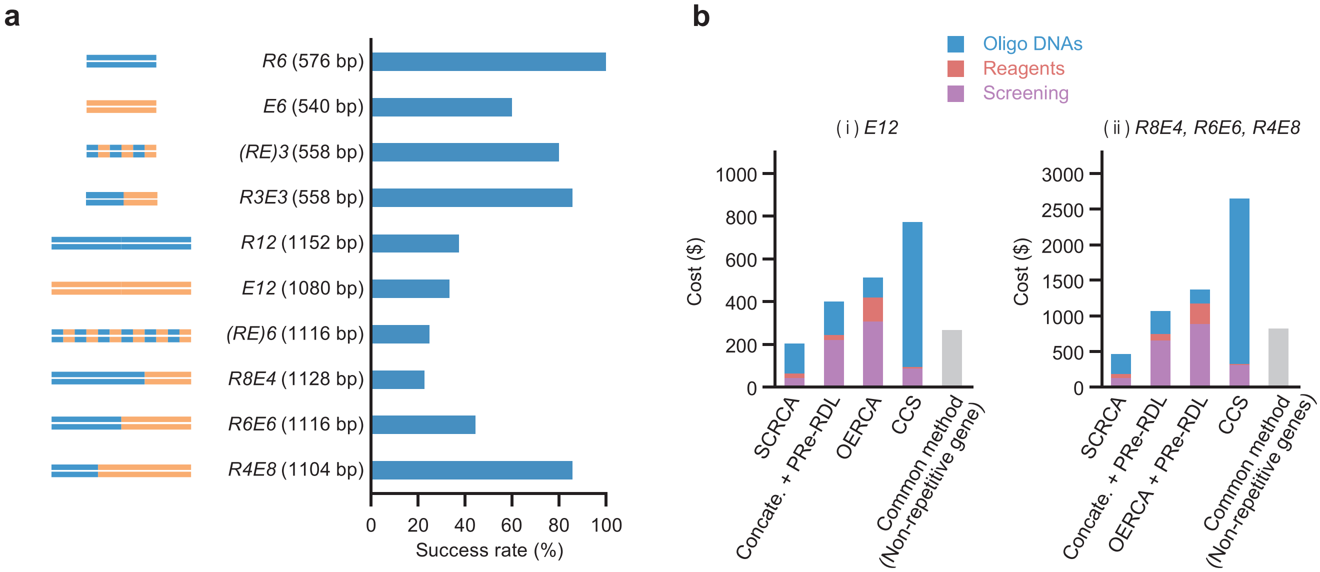


Figure 3. Success rates and costs of various repetitive-sequence gene syntheses. a) The success rates of gene syntheses were calculated from the positivity ratios obtained via colony PCR and Sanger DNA sequencing. b) Comparison of synthesis cost of SCRCA with that of conventional methods. Assuming that synthesis can be performed according to the roadmap, we calculated the cost of synthesizing i) a homopolymer gene (*E12*) alone and ii) block polymer genes (*R8E4*, *R6E6*, and *R4E8*) simultaneously. For the synthesis of the *E12* gene, the combination of concatemerization and recursive directional ligation by plasmid reconstruction (PRe-RDL),[31] OERCA,[33] and combinatorial codon scrambling (CCS)[32] were examined. For the synthesis of three block polymer genes, PRe-RDL combined with concatemerization, PRe-RDL combined with OERCA, and CCS were examined. The costs are expressed separately for oligo DNA, reaction reagents, and screening. For comparison, costs were calculated for the outsourced synthesis of non-repeat sequence genes of the same length.

estimated the cost of dNTPs to be \$16.7 (although 70% G/C dNTPs were actually used, to simplify the calculation, it was assumed that equal amounts of mixed dNTPs were used).[33] In contrast, a reaction solution of SCRCA comprises 50 µL of 1.5 mM dNTPs and requires only one reaction per gene synthesis; therefore, the cost of dNTPs is only \$1.0.

As the process of acquiring the desired constructs is critical in genetic recombination experiments, we also calculated the cost of screening (Tables S6 and S7, Supporting Information). The number of colonies required for screening was determined based on the detailed comparative results of Amiram et al.[33] For example, they reported that using the OERCA method required examining ≈200 colonies of which, only 2 colonies had >1 kb long genes. They first used colony PCR to ensure that the direction of insertion was correct. Although the authors did not reveal the percentage of colonies with the correct insertion orientation, we can safely assume it to be less than 50% based on the results of this colony PCR (Figure 3b in the reference) (This is a reasonable result since they cloned the gene at the blunt end). For 40% of the 200 colonies (80 colonies), we performed an additional colony PCR to determine the length. Candidate transformants were selected from these 40% colonies, and their sequences were confirmed by outsourced DNA sequencing. Assuming that the transformants carrying the *E12* gene were isolated by these operations, the number of colonies needed to be examined would be 280. In contrast, the method combining concatemerization and PRe-RDL was estimated to require 100 colonies to obtain the *E4* gene from the concatemerization product and 8 colonies each to construct the *E8* and *E12* genes from the obtained *E4* gene, thereby requiring a reagent cost of 116 colonies for the entire process (Figure S8, Supporting Information). Considering the results of the aforementioned previous studies[33] and principles,[31] these numbers of colonies would be appropriate. In contrast, with SCRCA, transformants with the desired sequence can be obtained from ≈8 colonies, greatly reducing the number of colonies to be tested (Figure S2, Supporting Information). The costs of the other methods were calculated in the same manner.

The SCRCA cost is estimated to be half of the conventional method cost and is equivalent to the cost of outsourced synthesis of non-repetitive genes (Figure 3b; Tables S1–S7, Supporting Information). Concatemerization and OERCA, like SCRCA, can simultaneously synthesize a variety of long and short repetitive genes,[31,33] but the screening cost is high due to the large number of colonies with genes of non-target length or those inserted in the reverse direction. In other words, the cost-effectiveness of SCRCA can be attributed to precise control of gene length and a high cloning success rate. Notably, only two weeks were required from ordering the oligo DNA to preparing the 10 repetitive-sequence genes. Since the outsourced synthesis of a gene of similar length takes 8–13 business days (Economy Gene Synthesis Service of Eurofins Genomics, Inc.), the workability of SCRCA is also comparable to that of common synthesis methods for non-repeat sequence genes.

2.4. Utilization of Copolymer Set Prepared by SCRCA

Copolymerization or block copolymerization using two repeat units with different functions is a powerful method to develop new functional protein polymers;[25,26,45–47] however, to create copolymers with the desired functionality, the effects of copolymerization efficacy and composition ratio must be investigated

 © 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

21983844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

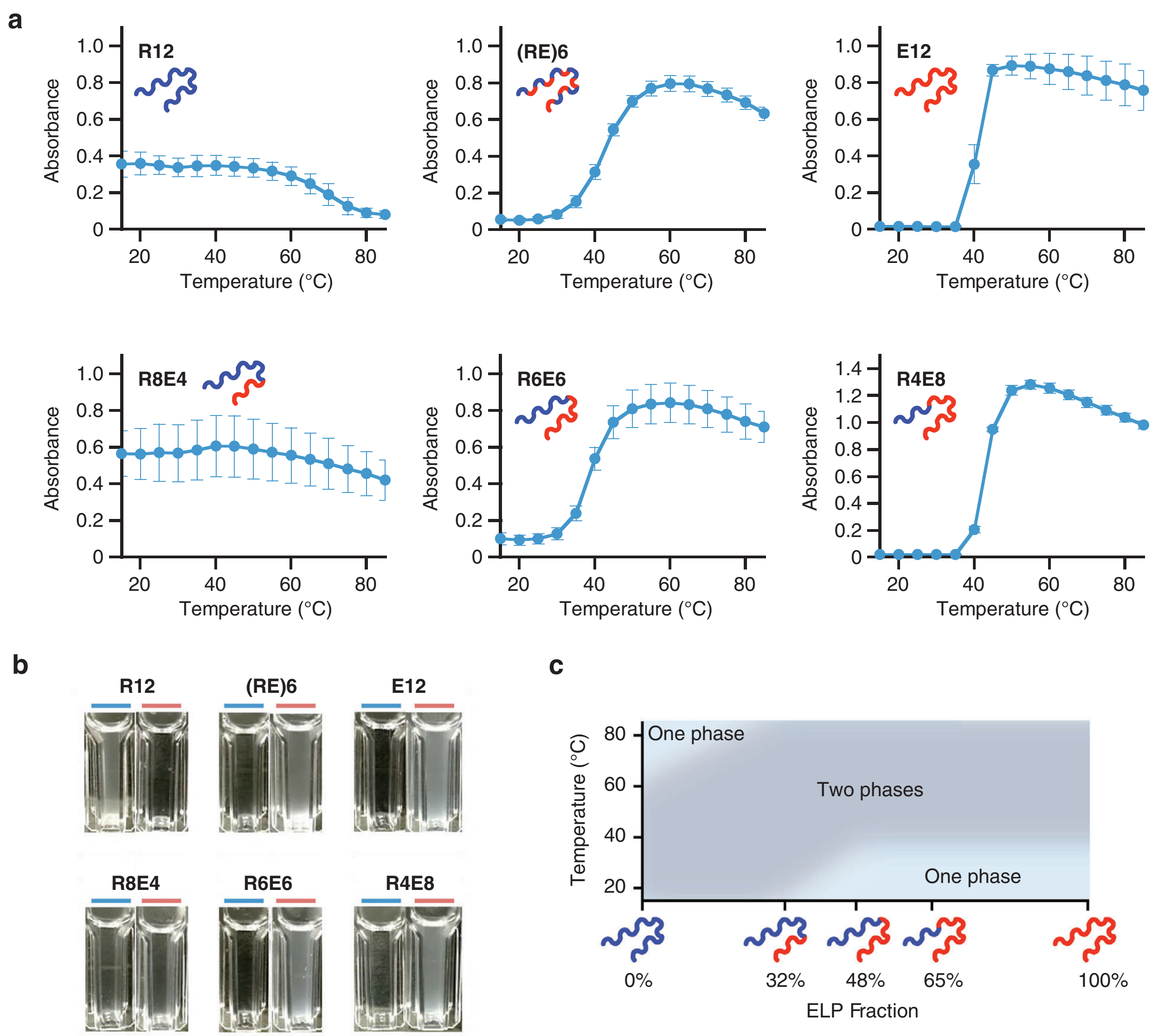


Figure 4. Temperature responses of the polymers prepared by seamless cloning of rolling-circle amplicons. a) Phase separation behaviors of polymers R12, E12, (RE)6, R8E4, R6E6, and R4E8 with the increase in temperature. The samples were prepared in PBS (pH 7.4) containing 2 µM polypeptide and 8 mM residual urea. Each value represents the mean of three replicates, with error bars showing the standard errors. b) Appearance of the polymer solution during absorbance measurements. Left: at 25 °C, right: at 80 °C. c) Effect of the ELP fraction on temperature-sensitive phase separation. The ELP fraction indicates the ratio of the ELP portion to the total repetitive sequence.

in advance. Copolymer sets that can be constructed using SCRCA can accelerate such investigations. To test this idea, we purified the constructed ELP-RLP copolymer set (Figure S10 (Supporting Information), the polymers corresponding to genes are written in Roman type) and investigated the effect of the block ratio on temperature-sensitive phase separation (UCST and LCST). As shown in **Figure** 4a,b, *E12, R4E8,* and *R6E6* showed cloudiness upon heating, R8E4 showed cloudiness over a wide temperature range of 15–85 °C, whereas R12 lost cloudiness upon heating. Microscopic observation confirmed that this cloudiness was attributed to phase separation (Figure S11, Supporting Information). These results indicate that to exhibit the phase separation phenomenon over a wide temperature range, the ELP fraction should be adjusted to ≈30% (Figure 4c). Considering that the function of a polymer is affected by its concentration and the surrounding environment,[40,48–52] it will be highly beneficial to prepare copolymer sets and evaluate their function in the actual environment where they will be used. Notably, for the conventional PRe-RDL, extra constructs, and multiple transformations are required when preparing copolymer sets,[26,45,53–55] whereas SCRCA does not require extra constructs and facilitates the preparation of copolymer sets in a single transformation (Figure S9, Supporting Information). As the simultaneous cloning of various block genes facilitates the construction of block polymers of different lengths and compositions at once, SCRCA can be an ideal gene synthesis method to comprehensively investigate the effects of sequence and length on the function of block polymers.

Interestingly, copolymer (RE)6 and block copolymer R6E6, which have the same amino acid composition, showed similar

 © 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

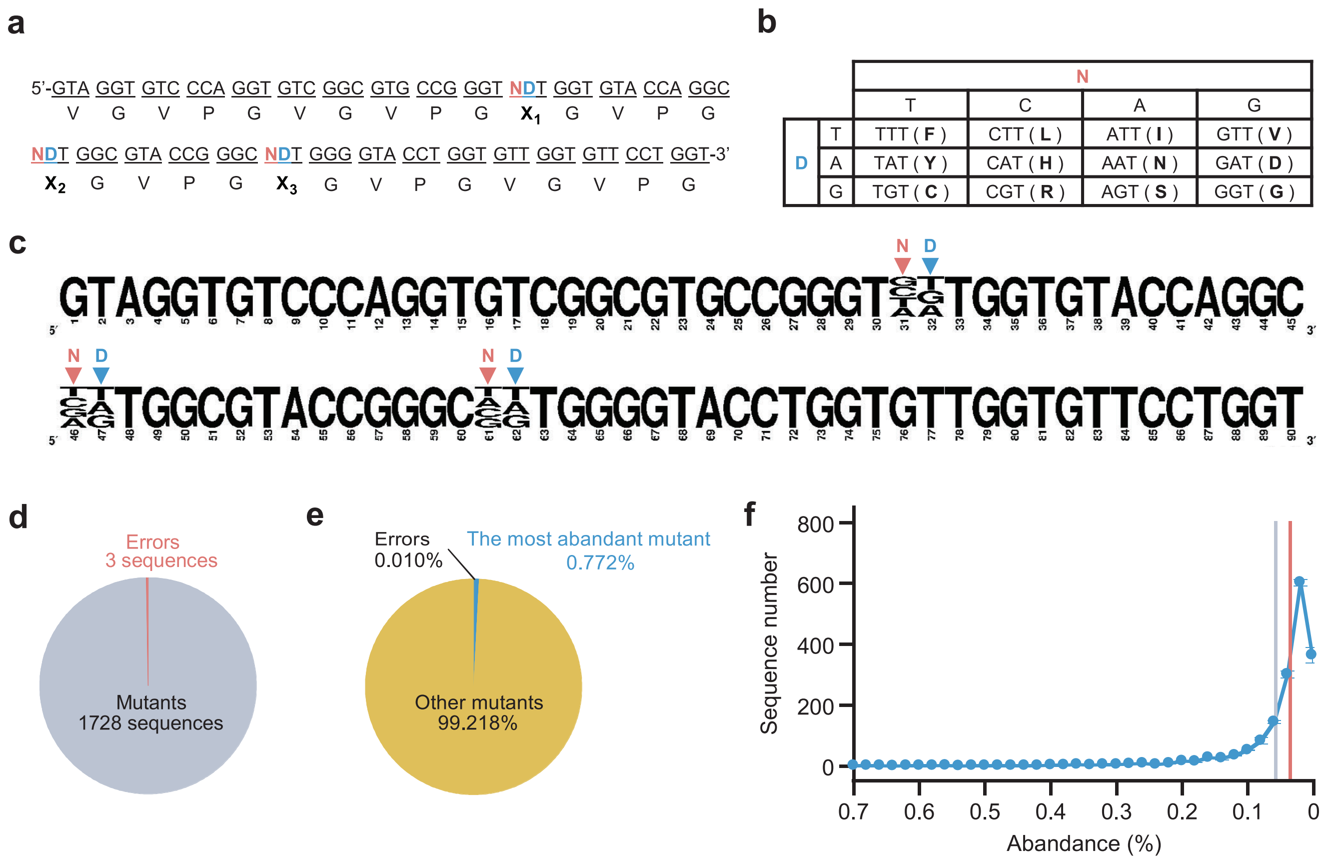


Figure 5. Construction of the repetitive-sequence gene library. a) Nucleotide sequence of the E-3NDT ring template library. b) At each NDT codon site, one of 12 types of amino acids is encoded by the combination of bases. c) Alignment result of the nucleotide repeat units found in the library constructed by RCA using the E-3NDT ring template library. The height of each nucleotide corresponds to its frequency. N and D represent the mixed base introduction position in E-3NDT ssDNA. d) Ratios of mutants and errors among all sequences. Each value represents the mean of two replicates. The corresponding standard deviations are 0.7 and 1.4, respectively. e) Abundance ratios of the most abundant mutant, other mutants, and error sequences. Each value represents the mean of two replicates. The corresponding standard deviations are 0.083%, 0.088%, and 0.005%, respectively. f) The frequency distribution of mutant sequences. Gray and red lines indicate theoretical and experimental medians, respectively. Values represent the mean of two measurements, and error bars represent standard errors.

temperature responses. When the ELP fraction was greater than 48%, the ELP moiety determined the temperature responsiveness of the polymer. As the protein-induced phase separation in cells is sequence-dependent,[56,57] the results of the present study may be specific to ELPs that undergo phase separation at high temperatures. Indeed, block and pseudorandom polymers of ELP mutants reportedly exhibit similar LCSTs.[58] Due to the small number of reported cases, future validation experiments using SCRCA on a diverse range of sequences are warranted.

2.5. Constructing the Repeat Unit Library

Directed evolution is a promising technique to develop highly functional enzymes and antibodies.[59] In this technique, the following four steps are repeated: 1) Construction of a gene library by mutagenesis; 2) Expression of genetic information in the form of protein; 3) Selection of the most suitable mutant using a simple evaluation method; 4) Designing next-generation libraries using the genes of the selected mutants as templates. To the best of our knowledge, directed evolution has not been previously used to develop protein polymers. This may be due to the fact that mutagenesis methods suitable for low-complexity proteins such as protein polymers have not yet been explored. We considered that a repetitive-sequence gene library with repeat units containing mutations would be useful for the directed evolution of protein polymers and investigated an approach for the preparation of a ring-template mixture using ssDNA containing mixed bases (Figure 1d). ELP was selected as the model sequence; the V sites of VGVPGVGVPGVGVPGVGVPGVGVPGVGVPG can be replaced with a guest amino acid, which changes the temperature responsiveness of the ELPs[60] (**Figure 5a**). We designed an E-3NDT ssDNA oligo in which the three codons (X_1, X_2, and X_3) encoding the V sites were replaced with NDT codons.[61] This codon encodes one of 12 diverse amino acids (Figure 5b).

To confirm the construction of the library, RCA products, which were presumed to have five repeats of the E-3NDT units, were analyzed using a next-generation sequencer. The nucleotide repeat units present at each end of the gene (the first at the 5′-end and the fifth at the 3′-end) could be read accurately. However, the sequence information at the second to fourth nucleotide repeat units was unreliable probably owing to poor

© 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

cluster formation due to the repetitive sequences. The counts of each sequence were almost the same in the first (at the 5′-end) and fifth (at the 3′-end) repeat units (R = 0.991), indicating that single nucleotide repeat units were repeated five times in a majority of the genes. Multiple alignments of the observed units confirmed that the E unit backbone was maintained (Figure 5c). In contrast, high complexity was observed at sites where mixed bases were introduced. Although this library contained all the 1728 expected sequences (all combinations of the mixed bases), it contained few error sequences (Figure 5d). Notably, even the most abundant sequences represented only <0.9% of the total number of sequences, and the median abundance ratio of all variants was 0.035% ± 0.01%, close to the ideal value (100%/1728 sequences = 0.058%) (Figure 5e,f). These results indicate that this library possessed excellent diversity. The minor deviation from the ideal value is believed to be due to a bias in the base types during the introduction of mixed bases into the oligo DNA.

Genes having six repeats of the E-3NDT unit were introduced into the *E. coli* expression system mentioned above. As E-3NDT is based on the E unit, this can be termed an E6 mutant library. We examined 42 transformed colonies and selected 28 colonies with the correct gene lengths (Figure S12, Supporting Information). Sanger sequencing revealed that the majority of the selected transformants carried E6 mutant genes with different amino acids (**Figure 6**a,b; the mutants are denoted by "*m*" and the colony number). Interestingly, block copolymer mutants with two or more different repeat units were obtained in addition to E6 mutants with the same unit repeated six times (as expected) (Figure 6b). We ascribe this to seamless cloning between the amplicons with ssDNA regions (Figure 6c). This phenomenon, which may also have occurred when constructing the genes comprising only one repeat unit, might facilitate the synthesis of long repetitive-sequence genes via SCRCA. Considering that polypeptide aggregation is affected by mean hydropathy and net charge,[62] the presence of block copolymer genes diversified the ELP library (Figure 6d).

Recently, a library construction technique applicable to low-complexity sequences, such as ELP, has been reported; however, this technique presents challenges, such as transformation efficiency of less than 25% and low library diversity (the occupancy of the most abundant sequences is more than 30%).[29] In contrast, libraries constructed using SCRCA have a transformation efficiency of ≈70% and excellent diversity. In addition, the cost associated with SCRCA is only one-tenth that of the previously reported method because the introduction of mixed bases is free. These advantages considerably reduce the hurdles in applying evolutionary molecular engineering to the development of functional protein polymers.

2.6. Development of the Desired Functional Protein Polymer by Directed Evolution

Finally, we tested whether the repeat unit library could be used to rapidly develop the desired protein polymers. As a development target, we selected a multi-response ELP that is soluble at low temperatures (4 °C, pH 7.4), insoluble near body temperature (37 °C, pH 7.4), and soluble in the environment around cancer cells (37 °C, pH 6.5). Such protein polymers are useful vehicles for drug delivery to cancer cells.[24,63] To develop polymers with these complex properties, molecules with LCSTs in the range of 4–37 °C at pH 7.4 and >37 °C at pH 6.5 must be designed (**Figure 7**a). The rational design of such polymers from the basic ELP sequence requires several years as factors, such as amino acid substitution, chain length, and influence of the operating environment, require careful examination to determine their effect on polymer responsiveness to temperature and pH changes.[64–67]

In the directed evolution experiment, we aimed to develop the desired polymer by optimizing the V̲ positions of E6, which has an LCST higher than body temperature and does not respond to pH changes (Figure 7b). A problem encountered during screening is that the temperature responsiveness of ELP is easily affected by its concentration and the presence of contaminants.[64] Therefore, prior to screening, we predicted the polymer functions from the amino acid sequences and then selected some mutants to reduce the number of samples. By evaluating the temperature responsiveness of the selected mutants after purification, the improvement in the function can be compared. We also conducted at least two rounds of experiments to confirm that the libraries constructed via SCRCA are applicable to the directed evolution.

The screening of the first directed evolution round involved searching for protein polymers that exhibited temperature and pH responsiveness in the E6 mutant library. Cysteine-containing mutants (m9, m33, and m40) were excluded owing to their potential to form irreversible aggregates. Block copolymers were also excluded because of the complexity of the next-generation library design. Consequently, three mutants, m8, m10, and m19, whose net charge changed significantly in response to a small change in pH, were selected and purified (Figure 7c; Figure S13, Supporting Information). The mutants m8 and m10 exhibited both temperature and pH responsiveness (Figure 7d). However, they were water-soluble near body temperature.

The first-round results (Figures 6b and 7c,d) revealed that the substitution of histidine results in a response to the small pH change. Lowering LCST while maintaining pH responsiveness would require the introduction of functional groups that interact with histidine, such as tyrosine, phenylalanine, and histidine.[68] Therefore, we designed the m10-derived sequence m10-3YWY, which has YWY codons encoding one phenylalanine, tyrosine, leucine, and histidine in the remaining three V̲ sites (positions X_4, X_5, and X_6) (**Figure 8**a,b). Among the 37 transformants examined using colony PCR and Sanger sequencing (Figure S14, Supporting Information), 27 mutants possessed the m10 backbone with different amino acids at positions X_4, X_5, and X_6 (Figure 8c; Figure S15, Supporting Information). This indicates that a template design can be used to construct a protein polymer library that inherits the characteristics of the useful sequences identified in the previous round.

To reduce the library size, we selected seven unique mutants based on their amino acid composition (Figure 8d) and purified them (Figure S16, Supporting Information). We measured turbidity in low-temperature (4 °C, pH 7.4), biological (37 °C, pH 7.4), and slightly acidic (37 °C, pH 6.5) environments on a microplate scale. Based on the result (Figure 8d), we selected m76, which agglutinated only at 37 °C, pH 7.4, as a strong candidate; its temperature and pH responses were close to the target properties (Figure 8e,f). This result indicates that the SCRCA-assisted di-

 © 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

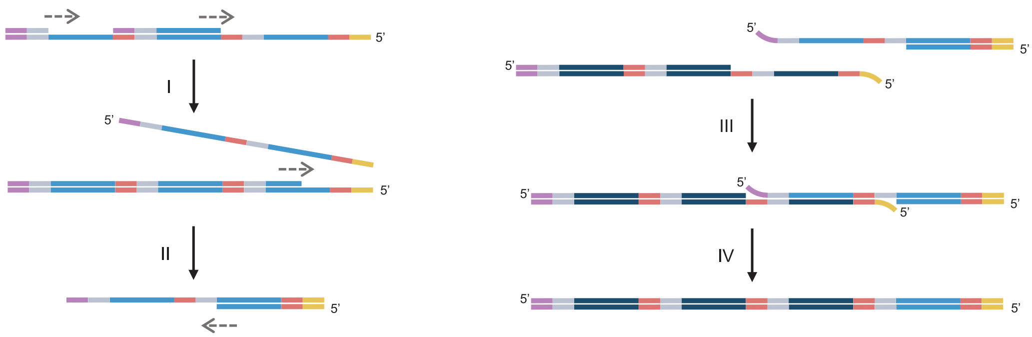
a

b

E6 mutants	1st			2nd			3rd			4th			5th			6th		
	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3
m1	Y	I	L	Y	I	L	Y	I	L	Y	I	L	Y	I	L	Y	I	L
m4	G	S	S	G	S	S	G	S	S	G	S	S	G	S	S	G	S	S
m5	V	I	L	V	I	L	V	I	L	V	I	L	V	I	L	V	I	L
m7	Y	G	N	Y	G	N	Y	G	N	G	H	H	Y	G	N	Y	G	N
m8	H	Y	S	H	Y	S	H	Y	S	H	Y	S	H	Y	S	H	Y	S
m9	C	Y	H	C	Y	H	C	Y	H	C	Y	H	C	Y	H	C	Y	H
m10	H	H	I	H	H	I	H	H	I	H	H	I	H	H	I	H	H	I
m12	L	V	I	L	V	I	L	V	I	L	V	I	L	V	I	L	V	I
m16	V	G	F	V	G	F	V	G	F	V	G	F	C	C	G	C	C	G
m17	V	G	I	V	G	I	V	L	G	R	G	V	F	C	G	R	G	D
m19	R	H	V	R	H	V	R	H	V	R	H	V	R	H	V	R	H	V
m20	F	V	F	F	V	F	F	V	F	F	V	F	F	V	F	F	V	F
m21	I	F	I	I	F	I	I	F	I	I	F	I	I	F	I	I	F	I
m22	F	G	S	F	G	S	F	G	S	F	G	S	F	G	S	F	G	S

E6 mutants	1st			2nd			3rd			4th			5th			6th		
	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3
m23	H	S	N	H	S	N	H	S	N	H	S	N	H	S	N	H	S	H
m24	F	F	L	F	F	L	F	F	L	F	F	L	F	F	L	F	F	L
m26	G	V	F	G	V	F	G	V	F	G	V	F	G	V	F	G	V	F
m29	D	I	I	D	I	I	D	I	I	D	I	I	D	I	I	D	I	I
m30	V	F	L	V	F	L	V	F	I	V	F	I	V	F	I	V	F	I
m31	V	G	L	V	Y	L	V	Y	L	V	G	F	V	G	F	V	G	F
m32	D	L	Y	D	L	Y	D	L	Y	D	L	Y	D	L	Y	D	L	Y
m33	I	C	R	I	C	R	I	C	R	I	C	R	I	C	R	I	C	R
m34	F	F	F	F	F	F	L	C	G	L	C	G	L	C	G	L	C	G
m37	V	R	Y	V	R	Y	V	R	Y	V	R	Y	C	D	F	V	G	D
m39	V	F	L	V	F	L	V	F	L	V	F	L	V	F	L	V	F	L
m40	C	C	N	C	C	N	C	C	N	C	C	N	C	C	N	C	C	N
m42	C	S	H	C	S	H	C	S	H	C	S	H	C	S	H	C	S	C

c



d

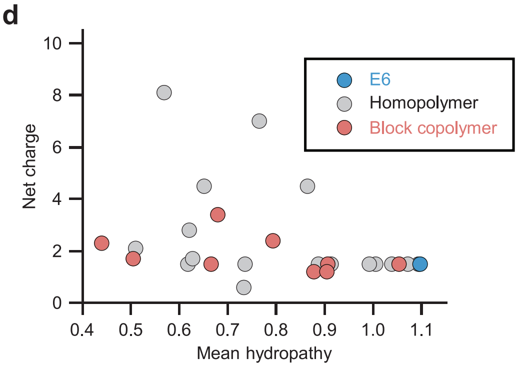


Figure 6. Composition of the E6 mutant library. a) Alignment result of the repetitive amino-acid-sequence parts (between 3 and 183 aa) of 27 E6 mutants fabricated by SCRCA using the E-3NDT ring template. The height of each amino acid corresponds to its frequency. The X_1, X_2, and X_3 positions represent NDT codons in the E-3NDT ring template. b) For 27 E6 mutants, the amino acids at X_1, X_2, and X_3 are summarized. The mutants highlighted in red have a block copolymer sequence consisting of two or more repeat units. c) Principle of block copolymer formation during library construction. During isothermal amplification (described in Figure 2a), dsDNAs I) with an ssDNA region at the 3′ end and II) with an ssDNA region at the 5′ end are generated, respectively (blue, gray, and red lines: repetitive sequences; purple and yellow lines: overlapping sequences). III) They are annealed, and IV) the nick is repaired during the In-Fusion reaction to generate a block copolymer gene. The dark blue line indicates a repeat unit that is partly different from that indicated by the blue line. d) Mean hydropathy and net charges are plotted for wild-type E6 (blue), 18 mutants with a homopolymer gene (gray), and 9 mutants with a block copolymer gene (red).

© 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

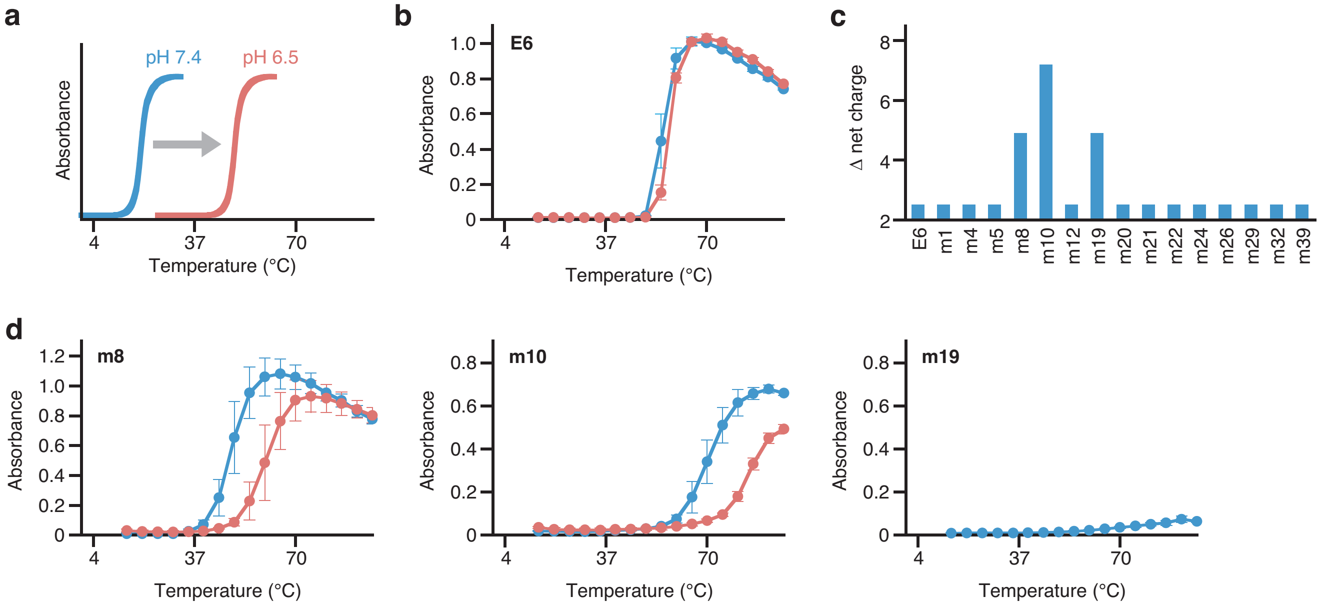


Figure 7. The first round of the directed evolution. a) The development target must have a transition temperature between 4 and 37 °C in a biological environment (pH 7.4) and >37 °C in a slightly acidic environment (pH 6.5). b) Temperature and pH responses of E6. The sample was prepared in phosphate-buffered saline (PBS) solutions containing 5 µM E6 (blue line: pH 7.4; red line: pH 6.5). Each value represents the mean of three replicates, with error bars showing the standard errors. c) Net charges at pH 7.4 and pH 6.5 are calculated based on the respective amino acid sequences. d) Temperature and pH responses of the three selected E6 mutants, i.e., m8, m10, and m19. The samples were prepared in PBS solutions containing 5 µM polymer (blue line: pH 7.4, red line: pH 6.5). Each value represents the mean of three replicates, with error bars showing the standard errors. The temperature response of m19 at pH 6.5 was not evaluated because this mutant did not show a lower critical solution temperature below 95 °C.

rected evolution strategy could facilitate the development of a desired protein polymer with a few transformation operations and within several months.

3. Conclusion

We developed a repetitive-gene synthesis method that combines RCA and seamless cloning and demonstrated that various genes encoding protein polymers can be easily synthesized. The SCRCA method has a high success rate of gene synthesis and excellent workability, and block copolymer genes, which normally require multiple cloning steps, can be constructed in a single cloning operation. As the operation days and cost of SCRCA are comparable to those of standard gene synthesis, SCRCA is expected to be a fundamental technology that will accelerate the rational design of functional protein polymers.

We also demonstrated that protein polymer libraries with different repetitive units can be easily constructed by introducing mixed bases into the template ssDNA. The library construction using SCRCA achieved superior transformation efficiency and diversity in relation to that using the state-of-the-art method. Furthermore, by combining this library construction technique with the directed evolution concept, we proved that the development of a highly functional protein polymer can be easily achieved without the need for diligent research. This development strategy may be suitable for further functionalization of conventional protein polymers and for the search for polymers with unknown functions. Considering its suitability for the directed evolution of low-complexity sequences, this strategy may also be used to understand diseases involving intrinsically disordered proteins[69,70] and to develop bioproduction processes using intracellular phase separation.[71] Future research will focus on efficiently screening desired protein materials and transformants. As proteins and polypeptides have a wide range of applications, a screening method that is suitable for development purposes must be proposed. Further studies will advance SCRCA to a technology that contributes substantially to multiple fields.

4. Experimental Section

Synthesis of the ssDNA Ring: DNA was synthesized by Eurofins Genomics, Inc. (Tokyo, Japan). The ssDNA ring template was designed based on the repeat-unit sequence and synthesized as 5′-phosphorylated ssDNA. The forward and reverse primers were designed based on the beginning and end sequences of the repeat unit, respectively. By designing the reverse primer to anneal to both ends of the 5′-phosphorylated ssDNA, the reverse primer was also used as a split oligo for ssDNA cyclization. The combinations of the ssDNA and primers used in this study are summarized in Table S8 (Supporting Information).

A mixture of 2 µL 50 µM 5′-phosphorylated ssDNA, 4 µL 50 µM reverse primer, and 29 µL milli-Q water was heated at 95 °C for 2 min. After cooling at 4 °C, 1 µL T4 DNA ligase (400 000 cohesive end units/mL, New England Biolabs) and 4 µL 10× T4 DNA ligase buffer were added, followed by incubation at 20 °C overnight. The product was purified using a GenElute PCR Clean-up Kit (Sigma-Aldrich, St Louis, MO, USA). The purity and concentration were determined using a Nanodrop 2000 device (Thermo Fisher Scientific, Waltham, MA, USA).

To prepare the RE ring template, a split DNA (ACCAGGAACACCATCACCACGACC) was designed based on both ends of the 5′-phosphorylated ssDNAs of the R and E rings. A mixture of 2 µL of each of the 50 µM 5′-phosphorylated ssDNAs, 4 µL of the 50 µM split DNA, 4 µL of the 50 µM reverse primer, and 23 µL milli-Q water was prepared. Cyclization was performed as described above.

 © 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

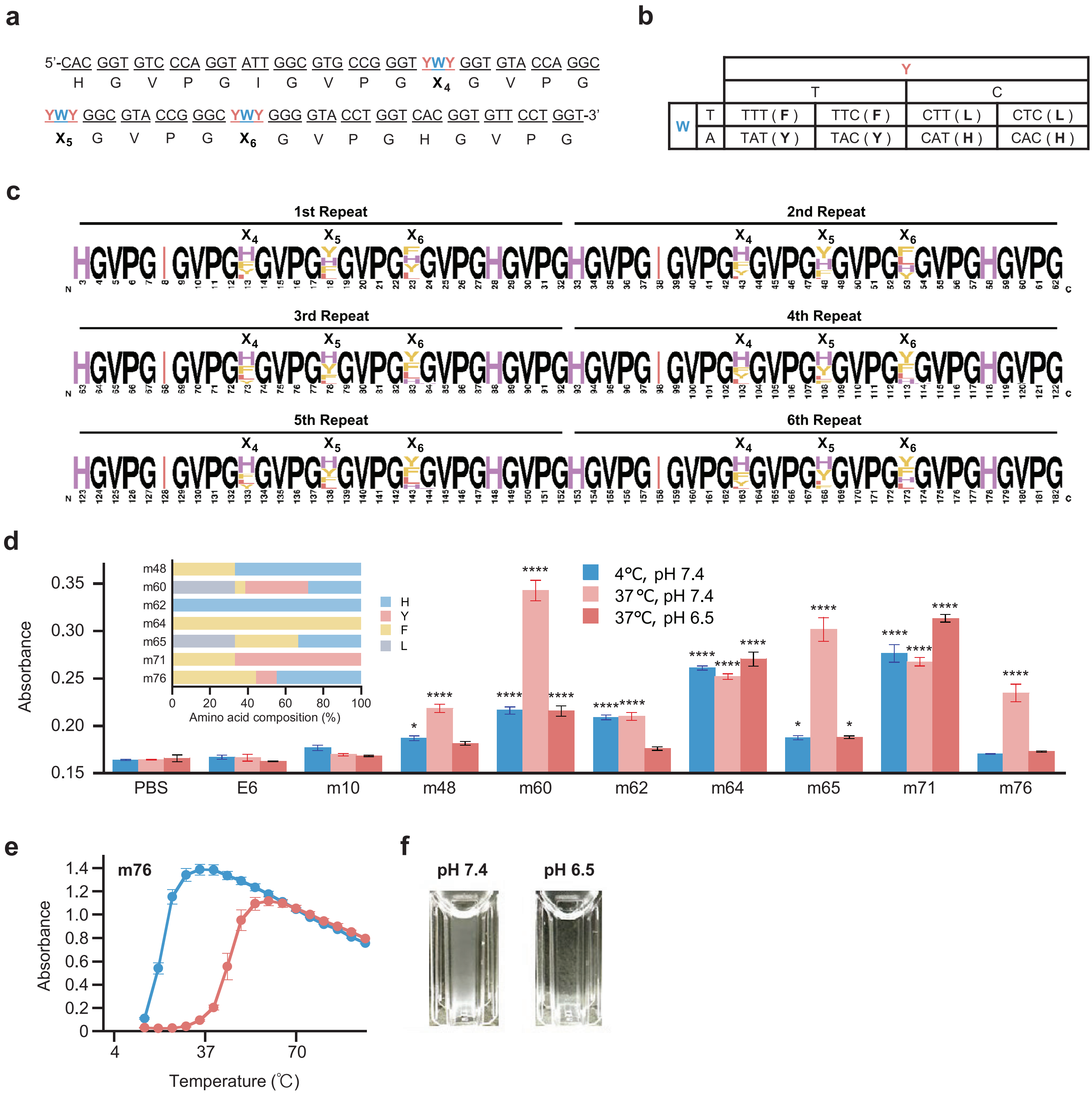


Figure 8. The second round of the directed evolution. a) Nucleotide sequence of the m10-3YWY ring template. b) At each YWY codon site, one of four types of amino acids is encoded by the combination of bases. c) Alignment result of repetitive amino-acid-sequence parts (in the range of 3–183 aa) of 27 m10 mutants prepared by SCRCA using the m10-3YWY ring template. The height of each amino acid corresponds to its frequency. The X_4, X_5, and X_6 positions represent YWY codons in the m10-3YWY ring template. d) Results of simple turbidity measurements. Seven m10 mutants, i.e., m48, m60, m62, m64, m65, m71, and m76, were selected based on their unique amino acid compositions (insert). The samples were prepared in PBS (pH 7.4) containing 5 µM polymer and <270 mM residual urea. Turbidity was measured at 4 and 37 °C, and then, the pH was adjusted to 6.5 by adding lactic acid. Each value represents the mean of three replicates, with error bars showing the standard errors. Statistical differences were determined using one-way ANOVA with Dunnett's multiple comparisons post-test. Differences in turbidity (at pH 7.4 and 4 °C) relative to the PBS blank: no asterisk $p > 0.1$; *$p < 0.1$; ****$p < 0.0001$. e) Temperature and pH responses of m76. The samples were prepared in PBS solutions containing 5 µM polymer and 80 mM residual urea (blue line: pH 7.4, red line: pH 6.5). f) Appearance of m76 polymer solution. Left: 37 °C, pH 7.4; right: 37 °C, pH 6.5.

© 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

RCA: A reaction solution of 100 μL containing 0.32 units μL^{-1} Bst DNA polymerase large fragment (New England Biolabs), 0.5 μM forward primer, 0.5 μM reverse primer, 0.2 ng μL^{-1} ssDNA ring template, 1.5 mM dNTP mix, 1× Thermopol buffer, and 2 mM magnesium sulfate was prepared. Isothermal amplification was performed at 60 °C for 12 h. A 1.5% TBE agarose gel was used to confirm the amplification products. The agarose gels were stained with ethidium bromide after electrophoresis and photographed using the Dolphin-Doc imaging system (Wealtec Corp., Sparks, NE, USA). Band positions were calculated using CS Analyzer 3.0 (ATTO Corporation, Tokyo, Japan).

Cloning and Transformation: For size selection of the DNA fragments, 2.5% TBE agarose gels containing 0.01% (v/v) SYBR Safe DNA were used. After electrophoresis, the band positions were checked using a blue-light illuminator, and the band corresponding to the desired length was cut using a cutter. The DNA fragments were purified using the FastGene Gel/PCR Extraction Kit (Nippon Genetics, Tokyo, Japan). The purity and concentration were determined using the Nanodrop 2000 device.

Expression vectors were constructed using an In-Fusion cloning kit (Takara Bio Inc., Shiga, Japan) according to the manufacturer's instructions. A linear vector was constructed via inverse PCR using the pET22b vector and two primers (TGGCCGACTCATCATCACCACCACCAC and TTCATATGTATATCTCCTTCTTAAAGTTAAAC). The sequences of these primers were designed to add the amino acid sequences MK to the N-terminus and WPTHHHHHH to the C-terminus of protein polymers. *Escherichia coli* BLR (DE3)-competent cells prepared in the laboratory were used for the transformation.

Colony PCR, Sanger DNA Sequencing, and Mass Spectrometry: The T7-promoter primer (TAATACGACTCACTATAGG) and T7-terminator primer (GCTAGTTATTGCTCAGCGG) were used for colony PCR and Sanger DNA sequencing. Colony PCR was performed according to the instructions for the Go Taq Green Mix (Promega, Madison, WI, USA), except that the annealing temperature was set to 50 °C, and the elongation time was set to 2 min for short genes (≈550 bp) and 3 min for long genes (≈1100 bp).

Sanger DNA sequencing was performed by Fasmac Co. Ltd (Atsugi, Japan). BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to analyze the sequence results. For genes with repetitive sequences of >1000 bp, the results of decoding in the 5′→3′ and 3′→5′ directions were combined (duplicated portions were confirmed to be >60 bp). The positivity rate (%) was calculated as [number of positive results from colony PCR / number of tests for colony PCR] × [number of positive results from Sanger DNA sequencing/number of tests for DNA sequencing].

For protein polymers whose repeat counts could not be confirmed by Sanger sequencing alone [R12, E12, and (RE)6], their masses were measured using matrix-assisted laser desorption ionization time-of-flight mass spectrometry. The measurements were performed by Apro Science (Tokushima, Japan).

Preparation of Block Copolymer Genes: For simultaneous seamless cloning, RLP-block genes were prepared via RCA with a forward primer (GATATACATATGAAAGGCGCGGTGACTCTCC) and a reverse primer (GACACCTACTGAGTAAGGTGAATCACCACGACC). These genes have an overlapping sequence at their 3′-end that recognizes the 5′-end of ELP-block genes. The ELP-block genes were prepared using RCA with a forward primer (TACTCAGTAGGTGTCCCAGGTGTCGG) and a reverse primer (ATGATGAGTCGGCCAACCAGGAACACCAACACCAGGTAC). The desired band of the block gene was cut out from the gel and purified. Two block genes and a linear vector were linked via an In-Fusion reaction.[44]

Cost Calculation: SCRCA synthesis costs were compared with those for PRe-RDL,[31] OERCA,[33] CCS,[32] and a common method of non-repeat gene synthesis. The costs were calculated separately for oligo DNA synthesis, reaction reagents, and screening. The costs for oligo DNA and non-repeat gene synthesis were determined based on the data provided on the websites of Eurofins Genomics Inc. (https://eurofinsgenomics.jp/jp/product/oligo-dna/standard-oligo-overview.aspx and https://eurofinsgenomics.jp/jp/service/gsy/overview.aspx). Reagent and screening costs were calculated by multiplying the cost per sample by the number of samples. Reagent prices were checked online in April 2024 and calculated at a currency conversion rate of ¥150 to $1. OERCA and concatemerization yield several mutants with non-purposeful repeat numbers, necessitating the screening of a large number of colonies. Therefore, based on the number of tests in the literature,[33] we determined that 280 and 100 colonies per transformation were required to evaluate OERCA and concatemerization, respectively.

Expression and Purification: Transformants were inoculated into a 200 mL ZYP-5052 medium and incubated at 25 °C for 2–3 days. After harvesting by centrifugation (9000×g, 5 min), the bacteria were stored at −30 °C. Protein polymers R12, E12, (RE6), R8E4, R6E6, and R4E8 were purified using His-tag affinity resin (His GraviTrap or Ni Sepharose 6Fast Flow, GE Healthcare, Chicago, IL, USA). Initially, 10 mL binding buffer (20 mM phosphate buffer (pH 7.4) containing 4 M urea, 0.5 м NaCl, and 40 mM imidazole) was added to the defrosted bacteria, and the bacteria were then sonicated on ice. After insoluble matter was removed via centrifugation, the supernatant was incubated at 37 °C for 10 min. Centrifugation was again performed to remove the generated insoluble matter, and the supernatant was passed through a 0.2 μm filter for sterilization. Next, affinity purification was performed according to the manufacturer's instructions, except that 0.04 and 4 м urea were added to the binding and elution buffers, respectively. The eluate was concentrated using an Amicon Ultra 10 kDa column (Millipore, Burlington, MA, USA). Protein polymers E6, m8, m10, and m19 were purified via the inverse transition cycling (ITC) method.[15] The purified polypeptides were dissolved in milli-Q water or PBS. Protein polymers m48, m60, m62, m64, m65, m71, and m76 were purified using the His-tag affinity resin and then further purified using the ITC method. The purified products were dissolved in 20 mM phosphate buffer (pH 7.4) containing 4 M urea to maintain their high concentrations and stored at −20 °C.

Verification of Protein Polymer Purity: Purified protein polymer concentrations were calculated by measuring the absorbance at 280 nm. The molar absorption coefficient was determined using the method reported by Pace et al.[72] To check for contamination by other proteins, sodium dodecyl sulfate-polyacrylamide gel electrophoresis was performed, and the gels were stained with Coomassie brilliant blue.

Evaluation of the Temperature Response: Absorbance was measured at 350 nm using a V-650 spectrophotometer (Jasco, Tokyo, Japan) connected to a temperature controller. PBS at pH 7.4 was used as the blank. To facilitate the adjustment of the polymer concentration, the measurement solution was prepared by diluting a concentrated solution. Some concentrated solutions contained 4 м urea to prevent phase separation at room temperature. The residual urea concentration in the measurement solution is presented in the figure legends. The temperature was increased by 1 °C per min. PBS at pH 6.5 was prepared by adding 22 μL of 50-fold diluted lactic acid to 1000 μL of the measurement solution.

Microscopic Observations: Images of phase separation were captured at 400 × magnification using an Olympus CKX53 microscope and Olympus CellSens software (Standard Version 4.1.1, Tokyo, Japan). A sample solution of 100 μL was placed in a transparent flat-bottomed 96-well plate, imaged at 25 °C, heated to 45 °C on a hot plate, and imaged again. For comparison, the same imaging was performed using PBS without the polymer.

Next-generation Sequencing Analysis: MiSeq Illumina sequencing (Illumina, San Diego, CA, USA) and data analysis were performed by Hokkaido System Science Co., Ltd (Hokkaido, Japan). The sequencing samples were synthesized via RCA using a set of primers (TCGTCG-GCAGCGTCAGATGTGTATAAGAGACAGTGTAGGTGTCCCAGGTGTCGG and GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACCAGGAACAC-CAACACCAGGTAC) with linker sequences at their 5′-ends. For size selection, agarose gel electrophoresis was conducted, and fragments corresponding to five repeats were excised from the gel. After purification, the samples were submitted to the company. After sequencing, the company removed the adapter sequence using Cutadapt[73] and trimmed the low QV region using Trimmomatic.[74] Cutadapt and Trimmomatic parameters are shown in the Supporting Information. The analysis revealed that the middle portions of the gene sequences were less reliable; therefore, only the sequences that were 90 bp from the 5′-ends and 90 bp from the 3′-ends were included in the analysis. However, since sequences with low counts lack reliability, only those with six or more counts were

© 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

analyzed. All sequences in a library were aligned in parallel using Weblogo (weblogo.berkeley.edu/logo.cgi).

Mean Hydropathy and Net Charge: The mean hydropathy and net charge were calculated using Expasy (https://web.expasy.org/protscale/) and PROTEIN CALCULATOR 3.4 (https://protcalc.sourceforge.net/), respectively. Calculations were performed using default settings.

Simple Turbidity Test: Polymer solutions of 150 µL (5 µM polypeptide in PBS, pH 7.4) were prepared in the cells of 96-well microplates (clear bottom, half area). As the measurement solutions were prepared by diluting the polymer concentrates with 4 M urea, the solutions contained 40–270 mM urea, which should have little effect on the transition temperature of ELPs.[75] The solutions were incubated at 4 °C for 30 min, and absorbance was measured at 350 nm. The solutions were further incubated at 37 °C for 30 min, and absorbance was measured again. Then, 3.3 µL of 50-fold diluted lactic acid was added to reduce the pH to 6.5. The solution was again incubated at 37 °C for 30 min, and absorbance was measured at 350 nm. PBS was used as a blank. The experiments were performed on three cells under identical conditions.

Statistical Analysis: The correlation coefficients were calculated, and statistical tests were performed using KaleidaGraph Version 4.5.3 (Synergy Software, Eden Prairie, MN, USA). One-way ANOVA with Dunnett's test was used for multiple comparisons. Differences with $p < 0.1$ were considered statistically significant for improving screening accuracy.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This study was partly supported by a Grant-in-Aid for Early-Career Scientists (No. 22K14549) and a Grant-in-Aid for Transformative Research Areas (No. 24H01378) from the Japan Society for the Promotion Science (JSPS), a project (No. JPNP20004) subsidized by the New Energy and Industrial Technology Development Organization (NEDO), the project "Research project for sericultural bio-industry" (No. JP22680575) commissioned by the Ministry of Agriculture, Forestry and Fisheries (MAFF), and the Research Collaboration Program (No. AM2100) of the National Institute for Materials Science (NIMS).

Conflict of Interest

The National Institute of Technology has filed a patent application on the repetitive-sequence gene library construction (PCT/JP2023/014202). The National Institute of Technology and NIMS have filed another patent application on the gene synthesis method. All other authors declare no competing interests.

Data Availability Statement

The data sets in this study are available within the source data.

Keywords

directed evolution, gene synthesis, protein polymers, repeat sequences, rolling-circle amplification

Received: October 30, 2024
Revised: January 15, 2025
Published online:

[1] A. Miserez, J. Yu, P. Mohammadi, *Chem. Rev.* **2023**, *123*, 2049.
[2] X. Zhang, J. Li, C. Ma, H. Zhang, K. Liu, *Acc. Chem. Res.* **2023**, *56*, 2664.
[3] M. P. Chang, W. Huang, D. J. Mai, *J. Polym. Sci.* **2021**, *59*, 2644.
[4] D. Chatterjee, F. J. Lewis, H. J. Sutton, J. A. Kaczmarski, X. Gao, Y. Cai, H. A. McNamara, C. J. Jackson, I. A. Cockburn, *Cell Rep.* **2021**, *35*, 108996.
[5] J. Forbes, A. Bissoyi, L. Eickhoff, N. Reicher, T. Hansen, C. G. Bon, V. K. Walker, T. Koop, Y. Rudich, I. Braslavsky, P. L. Davies, *Nat. Commun.* **2022**, *13*, 5019.
[6] T. Collins, J. Azevedo-Silva, A. da Costa, F. Branca, R. Machado, M. Casal, *Microb. Cell Fact.* **2013**, *12*, 21.
[7] T. Homma, S. Terui, F. Yokoyama, S. Okino, S. Ohta, C. Kato, N. Haraguchi, I. Fujisawa, S. Itsuno, L. Z. P. Ang, *Biotechnol. Bioeng.* **2023**, *120*, 194.
[8] D. Qin, M. Wang, W. Cheng, J. Chen, F. Wang, J. Sun, C. Ma, Y. Zhang, H. Zhang, H. Li, K. Liu, J. Li, *Angew. Chem., Int. Ed.* **2024**, *63*, e202400595.
[9] J. Li, B. Jiang, X. Chang, H. Yu, Y. Han, F. Zhang, *Nat. Commun.* **2023**, *14*, 2127.
[10] X. Hu, X. X. Xia, S. C. Huang, Z. G. Qian, *Biomacromolecules* **2019**, *20*, 3283.
[11] S. Lv, D. M. Dudek, Y. Cao, M. M. Balamurali, J. Gosline, H. Li, *Nature* **2010**, *465*, 69.
[12] D. Kam, A. Olender, A. Rudich, Y. Kan-Tor, A. Buxboim, O. Shoseyov, S. Magdassi, *Adv. Funct. Mater.* **2023**, *33*, 2210993.
[13] D. Ding, P. A. Guerette, J. Fu, L. Zhang, S. A. Irvine, A. Miserez, *Adv. Mater.* **2015**, *27*, 3953.
[14] M. Kim, W. G. Chen, J. W. Kang, M. J. Glassman, K. Ribbeck, B. D. Olsen, *Adv. Mater.* **2015**, *27*, 4207.
[15] S. R. MacEwan, W. Hassouneh, A. Chilkoti, *J. Vis. Exp.* **2014**, e51583.
[16] A. G. Goncalves, E. J. Hartzell, M. O. Sullivan, W. Chen, *Adv. Drug Delivery Rev.* **2022**, *191*, 114570.
[17] R. S. C. Su, R. J. Galas, C. Y. Lin, J. C. Liu, *Macromol. Biosci.* **2019**, *19*, 1900122.
[18] C. Ma, J. Dong, M. Viviani, I. Tulini, N. Pontillo, S. Maity, Y. Zhou, W. H. Roos, K. Liu, A. Herrmann, G. Portale, *Sci. Adv.* **2020**, *6*, eabc0810.
[19] Y. Dai, M. Farag, D. Lee, X. Zeng, K. Kim, H. Son, X. Guo, J. Su, N. Peterson, J. Mohammed, M. Ney, D. M. Shapiro, R. V. Pappu, A. Chilkoti, L. You, *Nat. Chem. Biol.* **2023**, *19*, 518.
[20] Z. G. Qian, F. Pan, X. X. Xia, *Curr. Opin. Biotechnol.* **2020**, *65*, 197.
[21] R. Price, A. Poursaid, H. Ghandehari, *J. Control Rel.* **2014**, *190*, 304.
[22] A. D. Malay, T. Suzuki, T. Katashima, N. Kono, K. Arakawa, K. Numata, *Sci. Adv.* **2020**, *6*, eabb6030.
[23] E. G. Roberts, N. G. Rim, W. Huang, A. Tarakanova, J. Yeo, M. J. Buehler, D. L. Kaplan, J. Y. Wong, *Macromol. Biosci.* **2018**, *18*, 1800265.
[24] M. Abdelghani, J. Shao, D. H. T. Le, H. Wu, J. C. M. van Hest, *Macromol. Biosci.* **2021**, *21*, 2100081.
[25] S. C. Huang, Z. G. Qian, A. H. Dan, X. Hu, M. L. Zhou, X. X. Xia, *ACS Biomater. Sci. Eng.* **2017**, *3*, 1576.
[26] A. Basheer, S. Shahid, M. J. Kang, J. H. Lee, J. S. Lee, D. W. Lim, *ACS Appl. Mater. Interfaces* **2021**, *13*, 24385.
[27] Q. Wang, X. Xia, W. Huang, Y. Lin, Q. Xu, D. L. Kaplan, *Adv. Funct. Mater.* **2014**, *24*, 4303.
[28] H. K. Binz, P. Amstutz, A. Kohl, M. T. Stumpp, C. Briand, P. Forrer, M. G. Grütter, A. Plückthun, *Nat. Biotechnol.* **2004**, *22*, 575.
[29] N. C. Tang, J. C. Su, Y. Shmidov, G. Kelly, S. Deshpande, P. Sirohi, N. Peterson, A. Chilkoti, *Nat. Commun.* **2024**, *15*, 3727.
[30] F. Teulé, A. R. Cooper, W. A. Furin, D. Bittencourt, E. L. Rech, A. Brooks, R. V Lewis, *Nat. Protoc.* **2009**, *4*, 341.
[31] J. R. McDaniel, J. A. MacKay, F. G. Quiroz, A. Chilkoti, *Biomacromolecules* **2010**, *11*, 944.
[32] N. C. Tang, A. Chilkoti, *Nat. Mater.* **2016**, *15*, 419.

© 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[33] M. Amiram, F. G. Quiroz, D. J. Callahan, A. Chilkoti, *Nat. Mater.* **2011**, *10*, 141.
[34] H. Jung, A. Pena-Francesch, A. Saadat, A. Sebastian, D. H. Kim, R. F. Hamilton, I. Albert, B. D. Allen, M. C. Demirel, *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 6478.
[35] P. H. N. Celie, A. H. A. Parret, A. Perrakis, *Curr. Opin. Struct. Biol.* **2016**, *38*, 145.
[36] J. M. Aliotta, J. J. Pelletier, J. L. Ware, L. S. Moran, J. S. Benner, H. Kong, *Genet. Anal. Biomol. Eng.* **1996**, *12*, 185.
[37] X. Li, Y. Huang, Y. Guan, M. Zhao, Y. Li, *Anal. Chim. Acta.* **2007**, *584*, 12.
[38] K. B. Ignatov, E. V. Barsova, A. F. Fradkov, K. A. Blagodatskikh, T. V. Kramarova, V. M. Kramarov, *BioTechniques* **2014**, *57*, 81.
[39] T. Yoshimura, T. Suzuki, S. Mineki, S. Ohuchi, *PLoS One* **2015**, *10*, e0136532.
[40] M. Dzuricky, B. A. Rogers, A. Shahid, P. S. Cremer, A. Chilkoti, *Nat. Chem.* **2020**, *12*, 814.
[41] J. R. Simon, N. J. Carroll, M. Rubinstein, A. Chilkoti, G. P. López, *Nat. Chem.* **2017**, *9*, 509.
[42] S. Rekhi, C. G. Garcia, M. Barai, A. Rizuan, B. S. Schuster, K. L. Kiick, J. Mittal, *Nat. Chem.* **2024**, *16*, 1113.
[43] S. E. Reichheld, L. D. Muiznieks, F. W. Keeley, S. Sharpe, *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E4408.
[44] B. Zhu, G. Cai, E. O. Hall, G. J. Freeman, *BioTechniques* **2007**, *43*, 354.
[45] P. Weber, M. Dzuricky, J. Min, I. Jenkins, A. Chilkoti, *Biomacromolecules* **2021**, *22*, 4347.
[46] X. X. Xia, Q. Xu, X. Hu, G. Qin, D. L. Kaplan, *Biomacromolecules* **2011**, *12*, 3844.
[47] B. Dai, C. J. Sargent, X. Gui, C. Liu, F. Zhang, *Biomacromolecules* **2019**, *20*, 2015.
[48] W. Hassouneh, M. L. Nunalee, M. C. Shelton, A. Chilkoti, *Biomacromolecules* **2013**, *14*, 2347.
[49] R. Balu, N. K. Dutta, N. R. Choudhury, C. M. Elvin, R. E. Lyons, R. Knott, A. J. Hill, *Acta Biomater.* **2014**, *10*, 4768.
[50] L. Li, T. Luo, K. L. Kiick, *Macromol. Rapid Commun.* **2015**, *36*, 90.
[51] M. R. Banki, L. Feng, D. W. Wood, *Nat. Methods* **2005**, *2*, 659.
[52] D. W. Lim, K. Trabbic-Carlson, J. A. MacKay, A. Chilkoti, *Biomacromolecules* **2007**, *8*, 1417.
[53] I. Weitzhandler, M. Dzuricky, I. Hoffmann, F. Garcia Quiroz, M. Gradzielski, A. Chilkoti, *Biomacromolecules* **2017**, *18*, 2419.
[54] S. Roberts, T. S. Harmon, J. L. Schaal, V. Miao, K. Li, A. Hunt, Y. Wen, T. G. Oas, J. H. Collier, R. V. Pappu, A. Chilkoti, *Nat. Mater.* **2018**, *17*, 1154.
[55] M. W. T. Werten, G. Eggink, M. A. Cohen Stuart, F. A. de Wolf, *Biotechnol. Adv.* **2019**, *37*, 642.
[56] E. W. Martin, A. S. Holehouse, I. Peran, M. Farag, J. J. Incicco, A. Bremer, C. R. Grace, A. Soranno, R. V. Pappu, T. Mittag, *Science* **2020**, *367*, 694.
[57] K. M. Ruff, Y. H. Choi, D. Cox, A. R. Ormsby, Y. Myung, D. B. Ascher, S. E. Radford, R. V. Pappu, D. M. Hatters, *Mol. Cell* **2022**, *82*, 3193.
[58] A. Chilkoti, M. R. Dreher, D. E. Meyer, *Adv. Drug Delivery Rev.* **2002**, *54*, 1093.
[59] Y. Wang, P. Xue, M. Cao, T. Yu, S. T. Lane, H. Zhao, *Chem. Rev.* **2021**, *121*, 12384.
[60] D. W. Urry, *Angew. Chemie Int. Ed. Eng.* **1993**, *32*, 819.
[61] M. T. Reetz, D. Kahakeaw, R. Lohmer, *ChemBioChem* **2008**, *9*, 1797.
[62] S. Roberts, M. Dzuricky, A. Chilkoti, *FEBS Lett.* **2015**, *589*, 2477.
[63] D. J. Callahan, W. Liu, X. Li, M. R. Dreher, W. Hassouneh, M. Kim, P. Marszalek, A. Chilkoti, *Nano Lett.* **2012**, *12*, 2165.
[64] D. E. Meyer, A. Chilkoti, *Biomacromolecules* **2004**, *5*, 846.
[65] Y. Cho, Y. Zhang, T. Christensen, L. B. Sagle, A. Chilkoti, P. S. Cremer, *J. Phys. Chem. B.* **2008**, *112*, 13765.
[66] J. R. McDaniel, D. C. Radford, A. Chilkoti, *Biomacromolecules* **2013**, *14*, 2866.
[67] J. A. MacKay, D. J. Callahan, K. N. FitzGerald, A. Chilkoti, *Biomacromolecules* **2010**, *11*, 2873.
[68] S.-M. Liao, Q.-S. Du, J.-Z. Meng, Z.-W. Pang, R.-B. Huang, *Chem. Cent. J.* **2013**, *7*, 44.
[69] B. Tsang, I. Pritišanac, S. W. Scherer, A. M. Moses, J. D. Forman-Kay, *Cell* **2020**, *183*, 1742.
[70] J. Wang, J. M. Choi, A. S. Holehouse, H. O. Lee, X. Zhang, M. Jahnel, S. Maharana, R. Lemaitre, A. Pozniakovsky, D. Drechsel, I. Poser, R. V. Pappu, S. Alberti, A. A. Hyman, *Cell* **2018**, *174*, 688.
[71] S. Lim, D. S. Clark, *Trends Biotechnol.* **2024**, *42*, 496.
[72] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, *Protein Sci.* **1995**, *4*, 2411.
[73] M. Martin, *EMBnet.journal* **2011**, *17*, 10.
[74] A. M. Bolger, M. Lohse, B. Usadel, *Bioinformatics* **2014**, *30*, 2114.
[75] F. G. Quiroz, N. K. Li, S. Roberts, P. Weber, M. Dzuricky, I. Weitzhandler, Y. G. Yingling, A. Chilkoti, *Sci. Adv.* **2019**, *5*, eaax5177.

© 2025 The Author(s). Advanced Science published by Wiley-VCH GmbH

21983844, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/advs.202410903 by National Institute For, Wiley Online Library on [02/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License