Vol. 14, pp. 23–28 (2017)
doi: 10.2142/biophysico.14.0_23

Biophysics and Physicobiology
https://www.jstage.jst.go.jp/browse/biophysico/

Regular Article

Enhanced mRNA-protein fusion efficiency of a single-domain antibody by selection of mRNA display with additional random sequences in the terminal translated regions

Kazuki Takahashi[1]*, Masato Sunohara[1]*, Takuya Terai[1], Shigefumi Kumachi[1] and Naoto Nemoto[1]

[1]*Graduate School of Science and Engineering, Saitama University, Saitama 338-8570, Japan*

Received October 12, 2016; accepted January 8, 2017

In vitro display technologies such as mRNA and cDNA display are powerful tools to create and select functional peptides. However, in some cases, efficiency of mRNA-protein fusion is very low, which results in decreased library size and lower chance of successful selection. In this study, to improve mRNA-protein fusion efficiency, we prepared an mRNA display library of a protein with random N- and C-terminal coding regions consisting of 12 nucleotides (i.e. four amino acids), and performed an electrophoresis mobility shift assay (EMSA)-based selection of successfully formed mRNA display molecules. A single-domain antibody (Nanobody, or VHH) was used as a model protein, and as a result, a pair of sequences was identified that increased mRNA-protein fusion efficiency of this protein by approximately 20%. Interestingly, enhancement of the fusion efficiency induced by the identified sequences was protein-specific, and different results were obtained for other proteins including VHHs with different CDRs. The results suggested that conformation of mRNA as a whole, rather than the amino acid sequence of the translated peptide, is an important factor to determine mRNA-protein fusion efficiency.

* These two authors contributed equally to this work.
Corresponding author: Naoto Nemoto, Graduate School of Science and Engineering, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama City, Saitama 338-8570, Japan.
e-mail: nemoto@fms.saitama-u.ac.jp

Key words: Nanobody, variable domain of heavy chain of heavy chain antibody, electrophoresis mobility shift assay, in vitro translation

In vitro display technology has been a powerful tool to discover functional peptides and proteins by in vitro selection methods. Although cell-based display/selection systems such as phage display or yeast surface display are quite popular, in vitro display using a cell-free translation system has advantages in terms of library size, throughput, and feasibility of chemical modifications [1,2]. In particular, we have focused on mRNA [3–5] and cDNA [6] display technology where genotype (i.e. mRNA or cDNA) and phenotype (i.e. its encoding polypeptide) are linked by an antibiotic, puromycin. These technologies have been successfully used in the selection of peptide aptamers against some target molecules [7–9]. However, low efficiency of translation and mRNA-protein fusion for some proteins is problematic for practical use because such low efficiency results in decreased library size and lower chance of successful selection. To overcome this problem, we previously performed an EMSA (electrophoresis mobility shift assay) to find novel translation enhancer sequences at 5′-untranslated region (UTR) [10]. However, although optimization of 5′-UTR should improve *the initiation efficiency of translation*, it would not directly

◀ ***Significance*** ▶

mRNA display is a robust technology to identify functional proteins from large libraries by *in vitro* selection, which involves covalently fusing the mRNA to its coding protein via a puromycin-linker. However, low fusion efficiency for some proteins, in particular antibodies, remained to be solved. Here, we describe improvement of the mRNA-protein fusion efficiency of a single-domain antibody by an electrophoresis mobility shift assay-based selection of mRNA display molecules. We believe the work makes the *in vitro* selection of affinity reagents using mRNA display more user friendly and practical.

enhance the formation efficiency of mRNA-protein conjugates. At present, factors determining this efficiency is not well known, and the N- and C-terminal region of structural gene might affect overall efficiency of mRNA-protein fusion. In this study, to further increase the final yields of mRNA-protein fusions in mRNA display, we performed a selection of mRNA display using a library with additional random sequences at 5′- and 3′-translated regions. Specifically, we used a single-domain antibody (i.e., variable domain of heavy chain of heavy chain antibody; Nanobody, or VHH) [11,12], a small antibody derived from camelids, as a model protein for selection that needs improvement of mRNA-protein fusion efficiency, considering its wide range of applications and relatively small size.

Materials and Methods

General

Primers used in this study are shown in Supplementary Table S1. DNA sequences of the constructs are described in Supplementary Table S2 and schematically shown in Supplementary Figure S1. Detailed protocols are described in Supplementary Text S1.

DNA construction

The wild type VHH DNA sequences were shown in Supplementary Table S2. The B domain of protein A (BDA) [13] and Pou-specific DNA-binding domain of Oct1 (PDO) [14]-coding DNAs were prepared as described previously [15]. Insertion of the selected sequence to these proteins was performed by polymerase chain reaction (PCR). After purification of PCR products, second PCR was performed to add a T7 promoter. The full construct DNAs were subject to gel-purification. T_{12} and A_{12} sequences were inserted to 3′-terminus of VHH in a similar way.

Library construction

A VHH library was constructed by adding random nucleotides (N_{12}) to the N- and C-termini of the wild type VHH1 gene by PCR. After adding T7 promoter, the full construct DNA library was purified and checked by polyacrylamide gel electrophoresis (PAGE) and DNA direct sequencing (see Supplementary Fig. S1).

EMSA selection

The DNA library was in vitro transcribed into mRNA. The mRNA was ligated with a puromycin DNA linker by T4 ligase as described previously [15]. The ligation products were added to a reaction mixture containing rabbit reticulocyte lysate. The cell free translation mixture was incubated to generate mRNA display molecules. The mRNA display formation was analyzed by 8 M urea containing 4% stacking-6% separating SDS-PAGE and detected with a fluorescence image analyzer. The band containing mRNA display was excised from the gel, and mRNA display molecules were eluted from the gel. The mRNA display molecules were ethanol precipitated and reverse transcribed to synthesize cDNA display. Synthesized cDNA display molecules were amplified by PCR to add a T7 promoter for the next round. The selected DNAs at final round were cloned and their sequences were analyzed.

Evaluation of mRNA-protein fusion efficiency by SDS-PAGE

mRNA-protein fusion efficiency of each clone was evaluated by SDS-PAGE. Efficiency of mRNA-protein fusion with inserted A_{12} and T_{12} sequences at 3′-terminus of VHH1 gene was also confirmed. Each DNA was transcribed into mRNA and the mRNA was ligated with puromycin linker as described above. mRNA-protein fusion was formed by cell free translation. The mRNA display molecules were analyzed by SDS-PAGE and fluorescently detected using FITC attached to the linker moiety. The mRNA-protein fusion efficiency was calculated from band intensity. mRNA displays of BDA, PDO, and VHH2-4 inserted with selected sequence were also formed by the method described above, and their mRNA-protein fusion efficiency was evaluated.

Results

Sample preparation

Random sequences of 12 nucleotides were inserted into both ends of DNA coding a model VHH protein (named as VHH1 in Supplementary Fig. S1) with a His × 6 tag at C-terminus by PCR. For efficient transcription and translation, T7 promoter, tobacco mosaic virus Ω, and Kozak sequences were attached as 5′-UTR (see Supplementary Table S2 for the sequence) [8]. The created library (7×10^{12} molecules at the first round) was transcribed into mRNA, and ligated to our original puromycin linker incorporating fluorescein and biotin (SBP-linker) [15]. The ligation product was translated in vitro using rabbit reticulocyte lysate, and incubation under high salt concentration was performed to promote mRNA-protein fusion [8]. Thus-obtained sample was used for EMSA selection. Schematic representation of the overall selection cycle is shown in Figure 1.

Selection by EMSA

To iteratively concentrate sequences that had high efficiency of mRNA-protein fusion, four rounds of EMSA selection was performed according to our previous literature [10]. In particular, mRNA-protein conjugate was separated from non-conjugated mRNA by denaturing SDS-PAGE, and the band containing the conjugate was carefully excised. Extracted RNA was reverse-transcribed and the cDNA was used as a template for the next round after addition of UTR. Reduced time of translation and high-salt incubation was used as a selection pressure. After the fourth selection, mRNA-protein fusion efficiency of the initial and selected library was compared directly by SDS-PAGE (Fig. 2). The

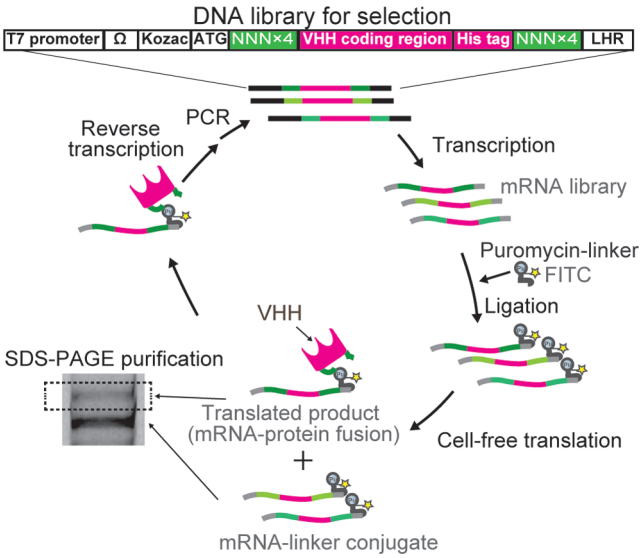

Figure 1 Overall schematic representation of this study. DNA library coding VHH with 12 nt random sequences at both termini were constructed by PCR. After transcription, ligation of puromycin linker was performed. Following cell-free translation, correctly formed mRNA-protein fusion, or mRNA display, was separated from non-fused mRNA by SDS-PAGE. The fusion was collected by dissecting gel, and reverse-transcribed cDNA was amplified by PCR, which was used as a template for the next round. LHR: linker hybridization region.

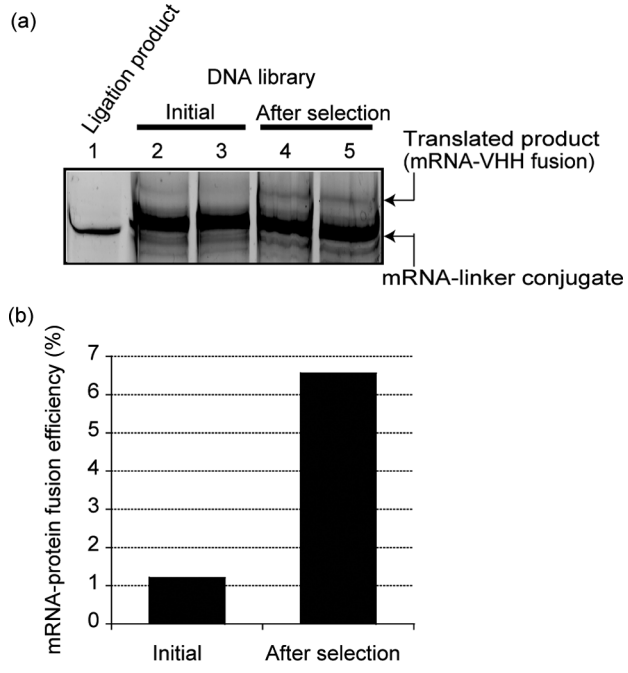

Figure 2 SDS-PAGE analysis of mRNA-protein fusion efficiency. (a) Initial library (lane 2, 3) and selected library (lane 4, 5) were translated under the same conditions and electrophoresed. mRNA-linker conjugate only was applied as a reference (lane 1). Gel was imaged using fluorescence of FITC, which was attached to the linker moiety (see Fig. 1). Samples were prepared duplicate and run on the same gel to confirm reproducibility. (b) Quantification of bands shown in (a). mRNA-protein fusion efficiencies were calculated and are shown as mean of two lanes.

band of the conjugate had increased intensity after the selection (6.6% of total mRNA) than before (1.2% of total mRNA), which indicated that the sequences featuring high mRNA-protein fusion efficiency were successfully enriched by this approach. Then, a portion of the selected library was cloned into *E. coli* and the colonies were randomly picked up for DNA sequencing.

Analysis of selected sequences

Twelve clones in total were successfully sequenced. As shown in Supplementary Table S3, no sequences were found duplicate or more, and we could not see clear homology among the identified DNA sequences or the peptides they coded, suggesting low convergence of selection at this stage. Even so, because the increase of average fusion efficiency of library was clear from Figure 2, mRNA-protein fusion efficiency of each clone was evaluated (Supplementary Fig. S2). Contrary to our expectation, however, only a few clones including #4 showed increase of efficiency compared with the original wt-VHH1 (i.e. protein without additional terminal sequences). This would be mainly because the resolution of EMSA selection is not high due to small difference of mobility. Indeed, when we prepared cDNA library from the dissected gel and analyzed by SDS-PAGE, majority of the sample was found to be non-conjugated cDNA (Supplementary Fig. S3). We note, however, that the population of cDNA-protein conjugate (15% of total cDNA) became much higher after four rounds of selection, and that the ratio of conjugate was increased by dissection (see Fig. 2 for comparison). This experiment also suggested that we might increase the library size of cDNA display for functional

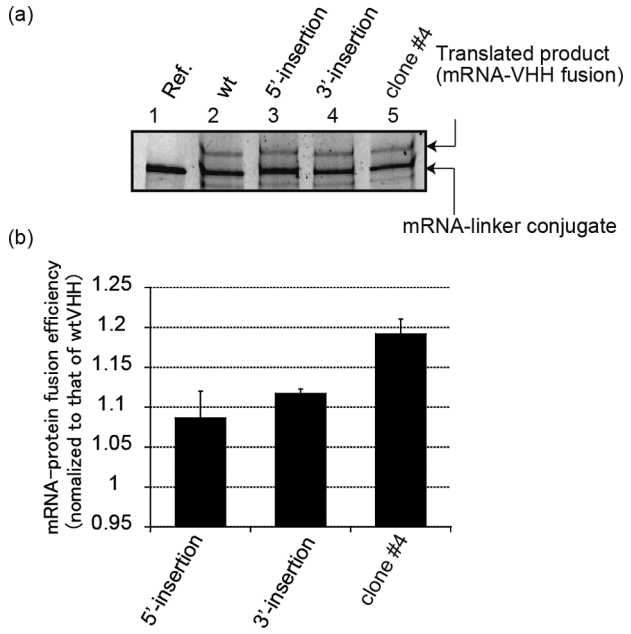

Figure 3 mRNA-protein fusion efficiency of mRNAs having both or only one of the identified 5′- and 3′-terminal sequences. (a) SDS-PAGE of translated samples. Lane 1, reference (non-translated mRNA of wt-VHH1). Lane 2, wt-VHH1 without any additional sequences. Lane 3, VHH1 with the additional 5′-sequence only. Lane 4, VHH1 with the additional 3′-sequence only. Lane 5, clone #4 (i.e. VHH1 with both additional sequences). (b) Quantification of bands shown in (a). mRNA-protein fusion efficiencies were normalized to that of wt-VHH1, and shown as mean±S.D. (n=3 independent experiments)

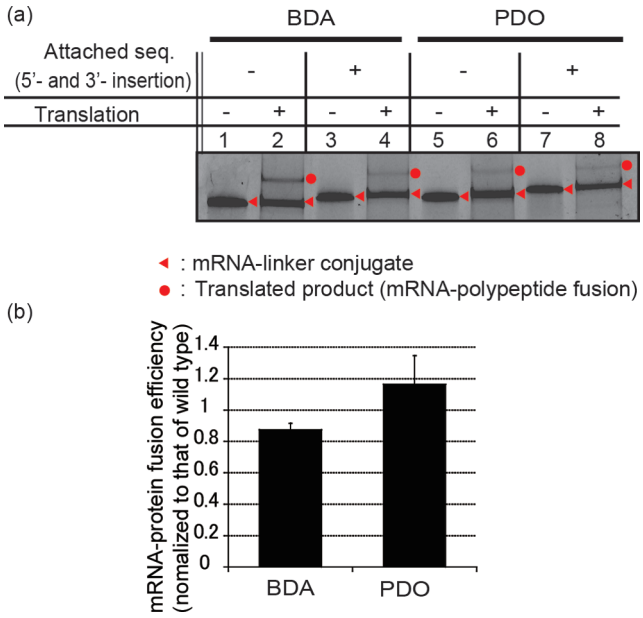

Figure 4 Effect of additional sequence of clone #4 on the mRNA-protein fusion efficiencies of BDA and PDO mRNAs. (a) SDS-PAGE of mRNAs before (odd lanes) and after (even lanes) translation reaction. Lane 1, 2: wt-BDA. Lane 3, 4: mRNA of BDA attached with 5′- and 3′-additional sequences. Lane 5, 6; wt-PDO. Lane 7, 8: mRNA of PDO attached with 5′- and 3′-additional sequences. (b) Quantification of bands shown in (a). mRNA-protein fusion efficiencies were normalized to those of wild types, and shown as mean±S.D. (n=3 independent experiments)

selection, if the selected sequences could be applied to other proteins as well.

Next, we focused on clone #4, which was the most promising clone identified in our selection (Supplementary Fig. S2). This clone had a DNA sequence of GTCGGTGTTTTT (translated into VGVF amino acids) at the 5′-terminal random region and CTAATTAGAAAT (translated into LIRN amino acids) at the 3′-terminal region. First, we examined which of the two regions contributed to the enhancement of mRNA-protein fusion efficiency, by translating mRNAs that had only one of the terminal sequences. As shown in Figure 3, introduction of one of the sequences alone to wt-VHH1 increased the mRNA-protein fusion efficiency by approximately 10%. This value was smaller than that of clone #4 (approximately 20% increase), indicating that both 5′- and 3′-terminal RNA (or amino acid) sequences played positive roles for the enhancement. We also attached T_{12} (translated to FFFF) or A_{12} (translated to KKKK) sequences to the 3′-terminus of wt-VHH1 mRNA, and compared their fusion efficiency with that of clone #4 (Supplementary Fig. S4). Only clone #4 showed significant enhancement, implying that the sequence of clone #4 had some special characteristics and that the achieved enhancement should not be explained by hydrophobicity or basicity of the selected C-terminal amino acids.

Introduction of the identified sequence to other proteins

We then asked if the selected sequence could enhance mRNA-protein fusion efficiency of other genes in general. As model proteins, the B domain of protein A (BDA) [13] and the Pou-specific DNA-binding domain of Oct-1 (PDO) [14] were chosen. The additional sequences of clone #4 was attached to the both termini of DNA coding these proteins (see Supplementary Table S2 for complete sequences). The DNAs were converted to mRNA-display as described above for VHH1, and the formation efficiencies of mRNA-protein conjugates were quantified by SDS-PAGE. As shown in Figure 4, significant increase of the efficiency was not observed for these proteins, implying that the effect of terminal sequences on mRNA-protein fusion efficiency is protein-dependent. To further test this idea, we attached T_{12} and A_{12} sequences to the 3′-terminus of BDA and PDO and checked their mRNA-protein fusion efficiencies (Supplementary Fig. S5). T_{12} did not change the efficiency, and A_{12} decreased mRNA-protein fusion efficiency of PDO only. These results also suggest the protein-dependent nature of the effect of terminal sequences on mRNA-protein fusion efficiency.

Introduction of the identified sequence to VHHs with different CDRs

The above results motivated us to further examine the

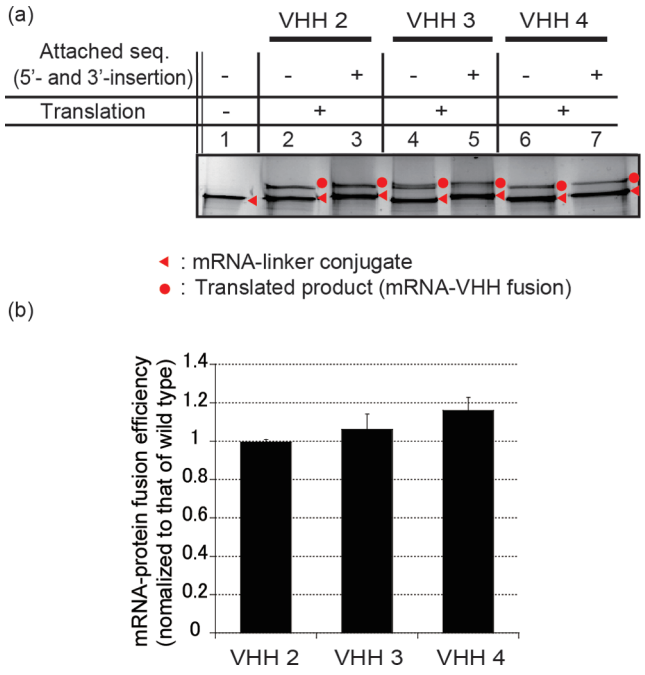

Figure 5 Effect of additional sequence of clone #4 on the mRNA-protein fusion efficiencies of other VHHs. (a) SDS-PAGE of translated mRNAs without (wild type, odd lanes) and with (even lanes) 5′- and 3′-additional sequences. Lane 1, reference (non-translated mRNA of wt-VHH1). Lane 2, 3: VHH2. Lane 4, 5: VHH3. Lane 6, 7: VHH3. (b) Quantification of bands shown in (a). mRNA-protein fusion efficiencies were normalized to those of wild types, and shown as mean ±S.D. (n=3 independent experiments)

scope of limitation of the identified sequence. So, we investigated whether the additional sequences of clone #4 had positive effects on mRNA-protein fusion efficiencies of other VHHs (i.e. antibodies with different complementarity determining regions (CDRs)). In particular, we randomly picked up three clones from a recombinant VHH library in our laboratory (named wt-VHH2-4), and inserted the identified sequences at the termini of coding regions (see Supplementary Table S2 for complete sequences). The DNAs were converted to mRNA display and the formation efficiencies of mRNA-protein conjugates were quantified as above (Fig. 5). To our surprise, the effect of the additional sequences was different among even VHHs, and significant increase of the efficiency was not observed for VHH2 or VHH3. The results suggested that the identified sequences herein could only enhance the fusion efficiency of a part of the VHH library, against our initial expectation.

Discussion

Control of translation efficiency is a long-standing question in biology, and several elegant theoretical models have been proposed mainly on the initiation process [16,17]. Also, selections of 5′-UTR sequences that enhanced translation efficiency were conducted earlier [10,18,19]. However, for mRNA/cDNA display technologies, not only translation initiation rate but the efficiency of puromycin incorporation are important. To our knowledge, the latter factor is rarely studied so far. In this work, by our original selection we identified a pair of sequence at the termini of coding region that increase this efficiency of a VHH, and demonstrated that the enhancing effect of the found sequences was highly protein-specific.

In the course of this study, Doi group addressed a similar problem [20]. They performed a selection to improve formation efficiency of mRNA-protein conjugates of single-chain variable fragment (scFv) using PURE system [21], and discussed that destabilization of mRNA secondary structure increased the efficiency of the formation of mRNA protein conjugates. Although the aim was the same as that of our paper, we think the following three points are different. First, the translation system and the translated protein are different. Second, in their work, a random sequence of 7 or 12 amino acid residues was inserted *before* FLAG tag at the C-terminus, unlike this study where the C-terminus of the translated region without FLAG tag was randomized. Our strategy should be more direct analysis of the effect of C-terminal sequences. Third, their selection method was dependent not only on the mRNA-protein fusion efficiency but also on the function (i.e. binding ability to the target protein) of displayed molecules. Though such strategy is meaningful to increase the yield of mRNA display that has a known function, we believe our EMSA selection should be more advantageous for creating a library of mRNA/cDNA display for subsequent functional selection.

Although the detailed mechanism of altered fusion efficiency by additional terminal sequences remains to be studied, the results in this work as well as those reported by Doi and others [20] suggested that mRNA as a whole is more important than the nature of translated polypeptides or protein scaffolds, at least in the case of antibodies. Perhaps, structures of mRNA (thermodynamic stability of secondary structures, interaction with ribosomes/tRNAs, and others) are a key factor that determines mRNA-protein fusion efficiency. Other factors may play roles, but further study is necessary. Also, it is currently unknown which part of the selected 12-nt sequences is important for the enhancement of efficiency. High-level in silico calculation or structural analysis of mRNA may provide clues to these remaining questions.

The fact that only the difference at CDR of VHH changed the enhancement effect of the added sequence was not what we expected, and it suggests that the identified sequence cannot be regarded as a universal and unbiased tool to increase the library size of VHH and other proteins in mRNA/cDNA display. However, we believe the obtained results provided fruitful insights into our basic understanding on the efficiency of mRNA-protein fusion as well as on the mRNA-ribosome complex.

Conclusion

In vitro display technologies are powerful tools to create and select functional peptides, but efficiency of mRNA-protein fusion is sometimes low, which needs to be addressed to improve the chance of successful selection. Also the factors that affect this efficiency is not well understood. Here, we inserted random sequences at both termini of the coding regions of a VHH, and performed our original EMSA-based selections. As results, we identified specific sequences that enhanced formation of mRNA-protein conjugate of the VHH. The degree of enhancement was only about 20%, but the difference from the original VHH was significant. Also, the 5′- and 3′-sequences seemed to have additive effects for the enhancement. Although the found sequences should not be generally used to increase the formation of mRNA-protein conjugates, this strategy itself will be applicable to other proteins. Moreover, the results gave us important suggestion that conformation of mRNA as a whole is a critical factor to determine mRNA-protein fusion efficiency.

Acknowledgements

This work was in part supported by KAKENHI (grant numbers 25008919,26350970 and 15H00797 to N. N., 15K05529 and 16J01710 to T. T.). T. T. was also supported by JSPS fellowship, and by The Uehara Memorial Foundation.

Conflicts of Interest

The authors declare that no competing interests exist.

Author Contributions

N. N. directed the project and wrote the paper. K. T. and M. S. performed the experiments and analyzed the data. S. K. and T. T. analyzed the data and co-wrote the paper with N. N.

References

[1] Josephson, K., Ricardo, A. & Szostak, J. W. mRNA display: from basic principles to macrocycle drug discovery. *Drug Discov. Today* **19**, 388–399 (2014).
[2] Valencia, C. A., Zou, J. & Liu, R. *In vitro* selection of proteins with desired characteristics using mRNA-display. *Methods* **60**, 55–69 (2013).
[3] Nemoto, N., Miyamoto-Sato, E., Husimi, Y. & Yanagawa, H. In vitro virus: bonding of mRNA bearing puromycin at the 3′-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett.* **414**, 405–408 (1997).
[4] Roberts, R. W. & Szostak, J. W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. USA* **94**, 12297–12302 (1997).
[5] Tabata, N., Sakuma, Y., Honda, Y., Doi, N., Takashima, H., Miyamoto-Sato, E., *et al.* Rapid antibody selection by mRNA display on a microfluidic chip. *Nucleic Acids Res.* **37**, e64 (2009).
[6] Yamaguchi, J., Naimuddin, M., Biyani, M., Sasaki, T., Machida, M., Kubo, T., *et al.* cDNA display: a novel screening method for functional disulfide-rich peptides by solid-phase synthesis and stabilization of mRNA–protein fusions. *Nucleic Acids Res.* **37**, e108 (2009).
[7] Ueno, S., Yoshida, S., Mondal, A., Nishina, K., Koyama, M., Sakata, I., *et al.* In vitro selection of a peptide antagonist of growth hormone secretagogue receptor using cDNA display. *Proc. Natl. Acad. Sci. USA* **109**, 11121–11126 (2012).
[8] Kumachi, S., Husimi, Y. & Nemoto, N. An RNA binding peptide consisting of four types of amino acid by in vitro selection using cDNA display. *ACS Omega* **1**, 52–57 (2016).
[9] Keefe, A. D. & Szostak, J. W. Functional proteins from a random-sequence library. *Nature* **410**, 715–718 (2001).
[10] Biyani, M., Biyani, M., Nemoto, N., Ichiki, T., Nishigaki, K. & Husimi, Y. Gel shift selection of translation enhancer sequences using messenger RNA display. *Anal. Biochem.* **409**, 105–111 (2011).
[11] Muyldermans, S. Nanobodies: Natural single-domain antibodies. *Annu. Rev. Biotechem.* **82**, 775–797 (2013).
[12] Harmsen, M. M. & De Haard, H. J. Properties, production, and applications of camelid single-domain antibody fragments. *Appl. Microbiol. Biotechnol.* **77**, 13–22 (2007).
[13] Gouda, H., Torigoe, H., Saito, A., Sato, M., Arata, Y. & Shimada, I. Three-dimensional solution structure of the B domain of staphylococcal protein A: comparisons of the solution and crystal structures. *Biochemistry* **31**, 9665–9672 (1992).
[14] Dekker, N., Cox, M., Boelens, R., Verrijzer, C. P., van der Vliet, P. C. & Kaptein, R. Solution structure of the POU-specific DNA-binding domain of Oct-1. *Nature* **362**, 852–855 (1993).
[15] Mochizuki, Y., Biyani, M., Tsuji-Ueno, S., Suzuki, M., Nishigaki, K., Husimi, Y., *et al.* One-pot preparation of mRNA/cDNA display by a novel and versatile puromycin-linker DNA. *ACS Comb. Sci.* **13**, 478–485 (2011).
[16] Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950 (2009).
[17] Borujeni, A. E. & Salis, H. M. Translation initiation is controlled by RNA folding kinetics via a ribosome drafting mechanism. *J. Am. Chem. Soc.* **138**, 7016–7023 (2016).
[18] Kamura, N., Sawasaki, T., Kasahara, Y., Takai, K. & Endo, Y. Selection of 5′-untranslated sequences that enhance initiation of translation in a cell-free protein synthesis system from wheat embryos. *Bioorg. Med. Chem. Lett.* **15**, 5402–5406 (2005).
[19] Mie, M., Shimizu, S., Takahashi, F. & Kobatake, E. Selection of mRNA 5′-untranslated region sequence with high translation efficiency through ribosome display. *Biochem. Biophys. Res. Commun.* **373**, 48–52 (2008).
[20] Nagumo, Y., Fujiwara, K., Horisawa, K., Yanagawa, H. & Doi, N. PURE mRNA display for in vitro selection of single-chain antibodies. *J. Biochem.* **159**, 519–526 (2016).
[21] Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., *et al.* Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**, 751–755 (2001).