COMMUNICATION

cDNA TRAP display for rapid and stable *in vitro* selection of antibody-like proteins

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Taishi Kondo,[a] Minori Eguchi,[a] Seita Kito,[a] Tomoshige Fujino,[a] Gosuke Hayashi [ab] and Hiroshi Murakami*[ac]

We developed a cDNA TRAP display for the rapid selection of antibody-like proteins in various conditions. By modifying the original puromycin linker in the TRAP display, a monobody was covalently attached to the cDNA. As a proof-of-concept, we demonstrated a rapid model selection of an anti-EGFR1 monobody in a solution containing ribonuclease.

In vitro selection technologies have been used to obtain various functional peptides[1-4] and antibody-like proteins (ALPs)[5-12]. mRNA display, which is one of the most popular methods used for *in vitro* selection[13-15], utilizes puromycin to connect the translated polypeptide and the coding mRNA. Although this method can use a huge library with a diversity of more than 10^{13}, it requires complicated multi-step procedures for the preparation of the ALP–puromycin linker (PuL) –mRNA complexes ("–" and "/" represent covalent and non-covalent linkages, respectively). Therefore, skillful techniques and a long experimental time are required to obtain functional peptides and ALPs.

To facilitate the *in vitro* selection process, we previously developed the Transcription–translation coupled with Association of Puromycin linker (TRAP) display[16] (Fig. 1a), and performed a selection of macrocyclic peptides[17] against the vascular endothelial growth factor receptor 2. Furthermore, we improved the TRAP display for the *in vitro* selection of ALPs, and rapidly obtained various monobodies and nanobodies against multiple targets, including the epidermal growth factor receptor 1, the human epithelial growth factor receptor 2, and the receptor-binding domain (RBD) of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike protein[18]. Moreover, the monobodies for which sequences were obtained within only 4 days after the arrival of the RBD had sub-nanomolar dissociation constants and were able to capture and neutralize the SARS-CoV-2.

This method is useful for the rapid selection of ALPs against various targets in an ordinal condition, but it also has a drawback because of the non-covalent linkage between ALPs (Fig. 1b). For example, this method cannot be used in the

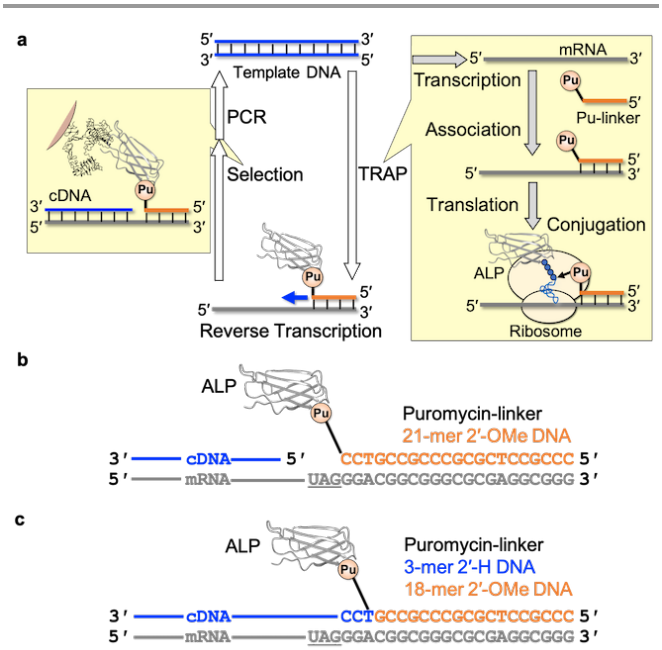
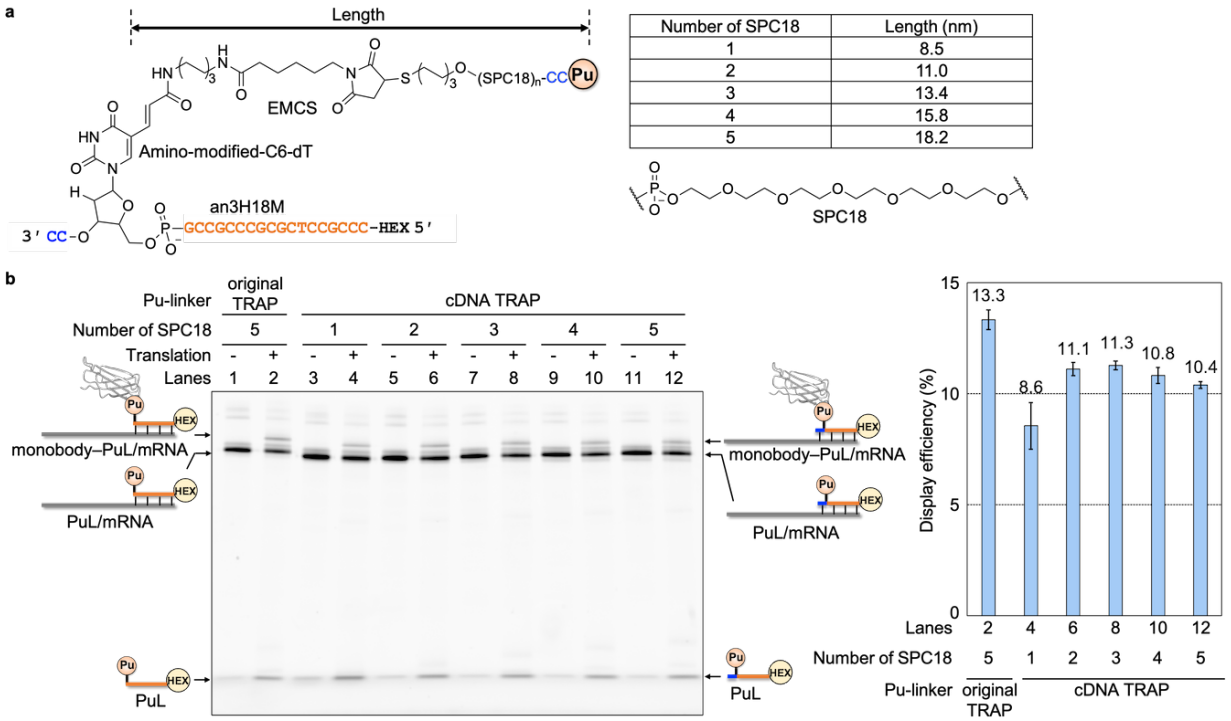


Fig. 1 TRAP display and puromycin linkers. (a) Schematic representation of the TRAP display. The addition of DNA to the TRAP reaction mixture results in the induction of four continuous reactions (transcription, association, translation, and conjugation), and produces an ALP–PuL/mRNA complex. After the addition of the RT primer, cDNA is synthesized by reverse transcription, and the resulting ALP–PuL/mRNA/cDNA complex is used for the selection of ALPs that bind to a target. (b) Design of a PuL in the original TRAP display. In the original TRAP display, the 21-mer of nucleic acids in the PuL was 2'-OMe DNA, with the PEG-spacer attached at the 3' end. (c) Design of a PuL in the cDNA TRAP display. The 2'-H of the three bases at the 3' end (CCT; coloured in blue) was unmodified, and the PEG spacer was attached at the thymine base at the thread position from the 3' end. "–" and "/" represent covalent and noncovalent linkages respectively. Abbreviations: RT, reverse transcription; PCR, polymerase chain reaction; PuL, Pu-linker; ALP, antibody-like protein

[a] *Graduate School of Engineering, Nagoya University, Nagoya, 464–8603, Japan. E-mail: murah@chembio.nagoya-u.ac.jp*
[b] *Japan Science and Technology Agency (JST), PRESTO, Kawaguchi, 332-0012, Japan.*
[c] *Institutes of Innovation for Future Society, Nagoya University, 464–8603, Japan.*
† Footnotes relating to the title and/or authors should appear here.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x



Number of SPC18	Length (nm)
1	8.5
2	11.0
3	13.4
4	15.8
5	18.2

Fig. 2 Optimization of the spacer length of a PuL for the cDNA TRAP display. (a) Structure of PuLs carrying an an3H18M ("H" and "M" represent 2′-H DNA and 2′-OMe DNA, respectively) oligonucleotide with 1–5 units of SPC18. Unmodified DNAs with 2′-H and modified DNAs with 2′-OMe are colored in blue and orange, respectively. The 5′ end of a PuL was labelled with HEX, to visualize the PuL, mRNA/PuL, and the mRNA/PuL monobody in a gel. (b) Urea SDS–PAGE analysis of the display efficiency of the wild-type monobody on the mRNA. The PuLs carrying a variable number of SPC18 units were annealed with the mRNA and added to a cell-free translation system (lanes 3–12). The PuL used in the original TRAP display was also used as a control (lanes 1 and 2). Error bars, standard deviation of each experiment (in triplicate). Abbreviations: SPC18, hexa-ethyleneglycol spacer; HEX, hexachlorofluorescein. Other abbreviations are as mentioned above.

presence of an RNase-containing solution, such as a cell-extraction, because of the absence of a bridge molecule between the cDNA and the ALPs once the mRNA is degraded. This method is also difficult to apply to a k_{off} selection including extremely long washing steps, or under high temperature, because they might result in the exchange of PuL on the mRNA and cause disruption of genotype–phenotype correlations.

Here, we reported development of a cDNA TRAP display in which the ALP and the cDNA were connected through a covalent linkage (Fig. 1c) by applying a stratagem reported previously, the so-called cDNA display[19-23].

In the original TRAP display, the 21-mer nucleic acid part of the PuL, which anneals with the mRNA, was 2′-OMe DNA (an21M), to prevent the synthesis of the complementary RNA via promoter-independent transcription by T7-RNA polymerase[24, 25]. For the cDNA TRAP display, the 3′ end of the an21M must be converted to 2′-H DNA to enable reverse transcription from the 3′ end. To optimize the quantity of 2′-H DNA at the 3′ end of the an21, we first tested the an3H18M-pri (in which "H" and "M" represent 2′-H DNA and 2′-OMe DNA, respectively) oligonucleotide and the an5H16M-pri oligonucleotide as primers for the reverse transcription (Fig. S1, ESI). We found that both oligonucleotides worked as efficient primers, at a level similar to that of the original reverse transcription primer.

Because the an3H18M-pri sequence contains 16 pyrimidine bases, it tends to form a triplex with the corresponding mRNA. Therefore, we optimized the an3H18M-pri sequence to prevent triplex formation. We found that the an3H18M-2 and 3 sequences were optimized (Fig. S2, ESI) and thus selected the an3H18M-3 sequence for further study.

We removed release factor 1 from a reconstituted *Escherichia coli* cell-free translation system[26-28], to generate a UAG codon as a blank codon in the TRAP display, thus causing the ribosome to pause at the UAG codon placed before the an3H18M/mRNA duplex (Fig. 1c); the puromycin would then attack the peptidyl-tRNA at the P-site of the ribosome to conjugate the ALP and the PuL. To maximize the conjugation efficiency, next we optimized the length of the linker between an3H18M and the puromycin. According to previous research[21], we incorporated amino-modified-C6-dT at the third position from the 3′ end of the an3H18M and attached *N*-(6-maleimidocaproyloxy)-succinimide (EMCS), followed by HS-$(SPC18)_n$-CC(puromycin) (SPC18 represents the hexa-ethyleneglycol spacer; n = 1–5; Figs 2a, S3, and S4, ESI), to the modified base. The resulting PuLs were annealed with the wild-type monobody (the 10th human fibronectin type-III domain) mRNA, and the resulting PuL/mRNA complexes were added to a translation reaction mixture. After reaction at 37°C for 30 min, the display efficiency of the wild-type monobody on the mRNA was analyzed by urea SDS polyacrylamide gel electrophoresis (SDS–PAGE), followed by gel analysis using a fluorescence gel imager. The display efficiencies of the monobody were similar when we used PuL with 2–5 SPC18 units (10%–11%; Fig. 2b, lanes 6, 8, 10, and 12), whereas the PuL with one SPC18 unit yielded a slightly lower efficiency (8.6%; Fig. 2b, lane 4). The length of the linker between the 5-methyl group on the modified thymine base and the amino group of the puromycin

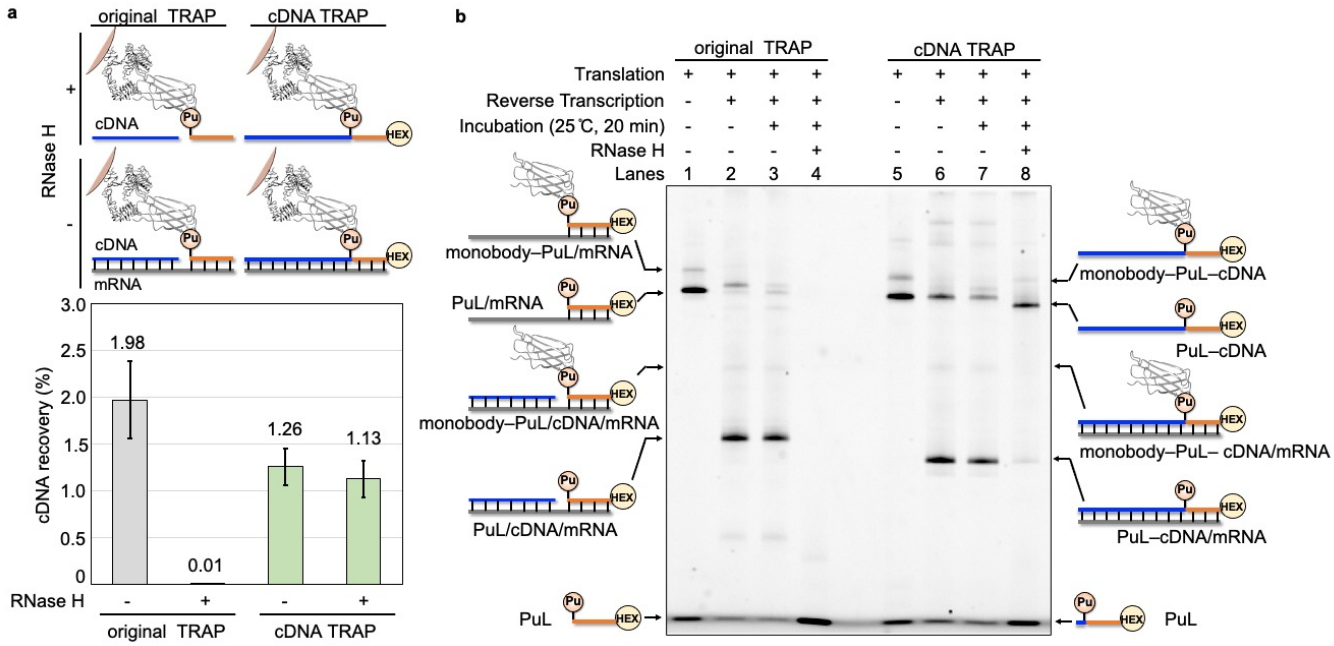


Fig. 3 Model selection of anti-EGFR1 monobody against its target using the original TRAP display or the cDNA TRAP display. (a) Recovery of cDNA in the pull-down experiments. After the preparation of the monobody–PuL/cDNA/mRNA or monobody–PuL–cDNA/mRNA complexes, RNase H was added to the solution mixture and the mixture was incubated at 25°C for 20 min. The complex was pulled down from the solution using EGFR1-immobilized magnetic beads, and the recovered cDNA was quantified by real-time PCR. The recovery of cDNA was calculated by dividing the amount of recovered cDNA by the theoretical amount of mRNA/PuL (1 μM). (b) Urea SDS–PAGE analysis of the monobody–PuL/cDNA/mRNA and monobody–PuL–cDNA/mRNA complexes in each solution mixture. PuL was visualized by fluorescence from the HEX attached at the 5′ end. Error bars, standard deviation of each experiment (in triplicate). Abbreviation: RNase H, ribonuclease H

was ~8.5 nm for the PuL with one SPC18, and was incremented by 2.4 nm per SPC18 unit. Conversely, the distance between the modified thymine base in an3H18M and the 3′ end of the peptidyl-tRNA at the P-site was estimated at about 10 nm by overexposing it to the crystal structure (PDB:4V6F)[29]; thus, one SPC18 unit would not be sufficient to localize the puromycin at the 3′ end of the peptidyl-tRNA in the model structure. Experimentally, two SPC18 units were required to reach the maximal display efficiency; therefore, a PuL with two SPC18 units was used in subsequent experiments.

The display efficiency of the monobody on the PuL with two SPC18 units was slightly lower than that of the monobody on the PuL linker that was used in the original TRAP display (Fig. 2b, lane 2). This might be caused by the lower affinity of an3H18M for the PuL compared with that of an21M for the mRNA. To increase the mRNA/PuL duplex stability, next we used L-αTNA, which reportedly has a high affinity for DNA and RNA (Fig. S5, ESI)[30, 31]. The PuL with an3H18T ("T" represents L-αTNA) was annealed with the monobody mRNA, and the resulting complex was added to the cell-free translation system. As expected, the portion corresponding to the PuL/mRNA complex was increased from 79.6% for the PuL with an3H18M to 85.3% when we used the PuL with an3H18T (Fig. S5, ESI). However, the display efficiency was slightly decreased (by 8.7%; Fig. S5, lane 4, ESI); thus, we used a PuL with an3H18M in subsequent experiments.

Using the optimized PuL [an3H18M-$(SPE18)_2$-CC(puromycin)], we finally demonstrated a model selection using a cDNA TRAP display. An anti-EGFR1 monobody (clone MC; K_D = 1.27 nM)[18] was used as a model ALP. The template DNA was added to the TRAP cell-free translation system containing either the original PuL or the new PuL. After the translation and reverse-transcription reactions, buffer exchange was performed via gel filtration, and the resulting solution was incubated with/without RNase H. A pull-down experiment was performed using EGFR1-immobilized magnetic beads, and the recovered cDNA was quantified by real-time PCR. The results showed that the amount of recovered cDNA in the presence of RNase H decreased from 1.98% (absence of RNase H) to 0.01% in the original TRAP display (Fig. 3a). We also analyzed the products of each step using urea SDS–PAGE. The monobody–PuL/cDNA/mRNA complex that was produced via reverse transcription remained after the incubation without RNase H, whereas it was decomposed after incubation with RNase H (Fig. 3b, lanes 2, 3, and 4). The degradation of mRNA led to the disruption of the monobody–PuL/cDNA/mRNA complex, which resulted in the very low recovery of cDNA in the EFGR1 pull-down experiment.

The quantity of cDNA recovered in the absence of RNase H (1.26%; Fig. 3a) was slightly lower than the original TRAP display, which is possibly due to the reverse transcription efficiency influence (69% for the original TRAP display and 59% for the cDNA TRAP display; Fig. 3b, lanes 2 vs. 6). On the other hand, the amount of recovered cDNA in the presence of RNase H (1.13%; Fig. 3a) was similar to that recorded in the absence of RNase H in the cDNA TRAP display. RNase H degraded the mRNA in the monobody–PuL–cDNA/mRNA complex, whereas the monobody–PuL–cDNA complex remained intact in the presence of RNase H because of the covalent linkage between PuL and cDNA (Fig. 3b, lanes 6 vs. 8). These results demonstrate the applicability of the cDNA TRAP display to an RNase-containing sample.

For monobody identification, biotin-Phe was incorporated at the protein's *N*-terminal by reprogramming the genetic code using biotin-Phe-tRNA^{ini} prepared with the flexizyme system[32], and the products were detected by Western blot analysis. Biotin was detected on the PuL/mRNA (Fig. S6, lanes 1 and 7, I, ESI), PuL–cDNA/mRNA (Fig. S6, lanes 2 and 8, II, ESI), and PuL–cDNA (Fig. S6, lanes 3 and 9, III, ESI) monobody complexes, and at the bottom of the gel (probably as drop-off products). Moreover, DNase I treatment resulted in the disappearance of band III (Fig. S6, lanes 9 vs. 10, ESI), providing further evidence that band III was the monobody–PuL–cDNA complex.

In this study, we developed a cDNA TRAP display for the rapid and stable selection of ALPs. We first optimized the 3' end of the an21 in the PuL for the efficient production of a PuL–cDNA molecule. We next optimized the length of the SPC spacer and the backbone structure of the an21 of the PuL. Moreover, using the optimized PuL, we established a model cDNA TRAP display selection method. The cDNA of an anti-EGFR1 monobody was successfully recovered even in the presence of RNase H. Because the cDNA TRAP display can prepare a stable ALPs–cDNA complex more rapidly than mRNA display[13, 14] and cDNA display[21], it would be useful for the rapid selection of ALPs under a variety of conditions, thus facilitating the future production of new ALPs against various targets.

This study was financially supported by AMED (grant numbers 20he0622010h0001 and 20fk0108293s0101 to H.M.), a Grant-in-Aid for Scientific Research on Innovative Areas (grant number 20H04704 to G.H.), and a Grant-in-Aid for Early-Career Scientists (grant number 18K14332 to T.F.) from the Japan Society for the Promotion of Science, and a donation from H. Murakami.

Conflicts of interest

There are no conflicts to declare.

1. Y. C. Huang, M. M. Wiedmann and H. Suga, *Chem. Rev.*, 2019, **119**, 10360-10391.
2. T. Fujino and H. Murakami, *Chem. Rec.*, 2016, **16**, 365-377.
3. S. Horiya, J. K. Bailey and I. J. Krauss, *Meth. Enzymol.*, 2017, **597**, 83-141.
4. M. S. Newton, Y. Cabezas-Perusse, C. L. Tong and B. Seelig, *ACS Synth. Biol.*, 2020, **9**, 181-190.
5. C. Hamers-Casterman, T. Atarhouch, S. Muyldermans, G. Robinson, C. Hammers, E. B. Songa, N. Bendahman and R. Hammers, *Nature*, 1993, **363**, 446.
6. K. Nord, E. Gunneriusson, J. Ringdahl, S. Stahl, M. Uhlen and P. A. Nygren, *Nat. Biotechnol.*, 1997, **15**, 772-777.
7. A. Koide, C. W. Bailey, X. Huang and S. Koide, *J. Mol. Biol.*, 1998, **284**, 1141-1151.
8. H. K. Binz, M. T. Stumpp, P. Forrer, P. Amstutz and A. Plückthun, *J. Mol. Biol.*, 2003, **332**, 489-503.
9. D. Grabulovski, M. Kaspar and D. Neri, *J. Biol. Chem.*, 2007, **282**, 3196-3204.
10. Y. Z. Wang, Z. Fan, L. Shao, X. W. Kong, X. J. Hou, D. R. Tian, Y. Sun, Y. Z. Xiao and L. Yu, *Int. J. Nanomedicine*, 2016, **11**, 3287-3302.
11. M. Gebauer and A. Skerra, *Annu. Rev. Pharmacol. Toxicol.*, 2020, **60**, 391-415.
12. Y. Nagumo, K. Fujiwara, K. Horisawa, H. Yanagawa and N. Doi, *J. Biochem.*, 2016, **159**, 519-526.
13. R. W. Roberts and J. W. Szostak, *Proc. Natl. Acad. Sci.*, 1997, **94**, 12297-12302.
14. N. Nemoto, E. Miyamoto-Sato, Y. Husimi and H. Yanagawa, *FEBS Lett.*, 1997, **414**, 405-408.
15. C. A. Olson, H. I. Liao, R. Sun and R. W. Roberts, *ACS Chem. Biol.*, 2008, **3**, 480-485.
16. T. Ishizawa, T. Kawakami, P. C. Reid and H. Murakami, *J. Am. Chem. Soc.*, 2013, **135**, 5433-5440.
17. T. Kawakami, T. Ishizawa, T. Fujino, P. C. Reid, H. Suga and H. Murakami, *ACS Chem. Biol.*, 2013, **8**, 1205-1214.
18. T. Kondo, Y. Iwatani, K. Matsuoka, T. Fujino, S. Umemoto, Y. Yokomaku, K. Ishizaki, S. Kito, T. Sezaki, G. Hayashi and H. Murakami, *Sci. Adv.*, 2020, **6**, eabd3916.
19. M. Kurz, K. Gu, A. Al-Gawari and P. A. Lohse, *Chembiochem*, 2001, **2**, 666-672.
20. I. Tabuchi, S. Soramoto, N. Nemoto and Y. Husimi, *FEBS Lett.*, 2001, **508**, 309-312.
21. J. Yamaguchi, M. Naimuddin, M. Biyani, T. Sasaki, M. Machida, T. Kubo, T. Funatsu, Y. Husimi and N. Nemoto, *Nucleic Acids Res.*, 2009, **37**, 13.
22. T. Suzuki, Y. Mochizuki, S. Kimura, Y. Akazawa-Ogawa, Y. Hagihara and N. Nemoto, *Biochem. Biophys. Res. Commun.*, 2018, **503**, 2054-2060.
23. T. Terai, T. Koike and N. Nemoto, *Molecules*, 2020, **25**, 12.
24. G. Krupp, *Nucleic Acids Res.*, 1989, **17**, 3023-3036.
25. R. Padilla and R. Sousa, *Nucleic Acids Res.*, 2002, **30**, e138.
26. Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa and T. Ueda, *Nat. Biotechnol.*, 2001, **19**, 751-755.
27. H. Ohashi, Y. Shimizu, B. W. Ying and T. Ueda, *Biochem. Biophys. Res. Commun.*, 2007, **352**, 270-276.
28. P. C. Reid, Y. Goto, T. Katoh and H. Suga, *Methods Mol. Biol.*, 2012, **805**, 335-348.
29. L. B. Jenner, N. Demeshkina, G. Yusupova and M. Yusupov, *Nat. Struct. Mol. Biol.*, 2010, **17**, 555-U548.
30. K. Murayama, H. Kashida and H. Asanuma, *Chem. Commun.*, 2015, **51**, 6500-6503.
31. H. Kashida, K. Murayama and H. Asanuma, *Polym. J.*, 2016, **48**, 781-786.
32. H. Murakami, A. Ohta, H. Ashigai and H. Suga, *Nat. Methods*, 2006, **3**, 357-359.