A 24 mer peptide derived from RecA protein can discriminate a single-stranded DNA from a double-stranded DNA

Naoki Sugimoto* and Shu-ichi Nakano

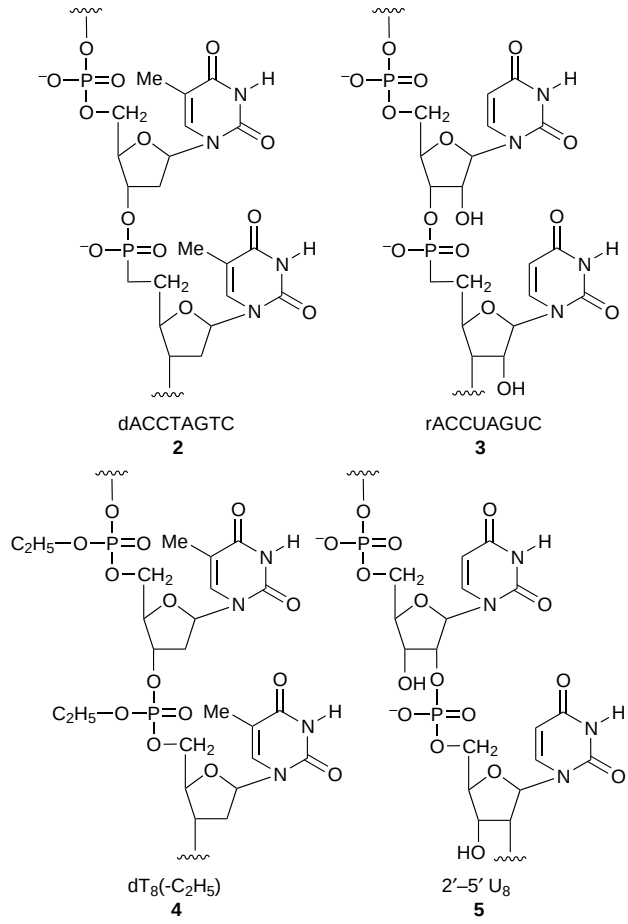
Department of Chemistry, Faculty of Science, Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe 658, Japan

A novel 24 residue peptide, Ile-Arg-Met-Lys-Ile-Gly-Val-Met-Phe-Gly-Asn-Pro-Glu-Thr-Thr-Thr-Gly-Gly-Asn-Ala-Leu-Lys-Phe-Tyr, derived from RecA protein can discriminate a single-stranded DNA (ssDNA) from a double-stranded DNA (dsDNA) and a new developed support with this peptide recognizes ssDNA but not dsDNA.

RecA protein plays an essential role for homologous DNA recombination and DNA repair in prokaryote.[1–3] The X-ray structure of 352 amino acids of *E. coli* RecA protein in the absence of DNA showed that two disordered loop domains, L1 (157–164) and L2 (195–209), composed of highly conserved amino acids project towards the polymer axis near DNA strands.[4] Based on several investigations of mutated RecA proteins, it was proposed that L1 and L2 are dsDNA and ssDNA binding domains, respectively. Helix G (210-218) lying next to loop L2 amino acids is a conserved domain including two Gly residues (211 and 212) which are absolutely invariant among bacterial RecA proteins, T4 UvsX, yeast Dmc 1 and mouse Rad 51 proteins.[5] Recently, it was reported that a 20 amino acid peptide (named FECO peptide) derived from *E. coli* RecA protein, which contains L2 and the conserved amino acids at both L2 termini including the invariant Gly residues (193–212), showed an enzymatic activity for the homologous recombination of DNA strands without any cofactors.[6] However, the FECO peptide which lacks the full helix G amino acids except both Gly residues, bound both ssDNA and dsDNA with similar affinity. Here, we have synthesized a 24 amino acid peptide including L2 and all helix G residues (195–218), Ile-Arg-Met-Lys-Ile-Gly-Val-Met-Phe-Gly-Asn-Pro-Glu-Thr-Thr-Thr-Gly-Gly-Asn-Ala-Leu-Lys-Phe-Tyr **1**, and investigated its recognition ability specific to ssDNA.[7]

Binding behaviors of the synthesized L2-G peptide **1** with four single-stranded octanucleotides, dACCTAGTC **2**, rACCUAGUC **3**, $dT_8(\text{-}C_2H_5)$ **4**, and 2′–5′ U_8 **5**, a double-stranded DNA (dsDNA) d(GTCAGGAATCTG)/d(CAGATTCCTGAC) **6**, and a hairpin-looped DNA $dCGTCA_8GACG$ **7** were investigated by surface plasmon resonance (SPR) analysis. Octamers **2** and **3** are unmodified ssDNA and ssRNA, respectively. Octamers **4** and **5** are unnatural strands; **4** has neutral charges on its backbone and **5** is linked by a phosphodiester bond between 2′ and 5′ positions. The DNA sequences of **6** and **7** were confirmed to form double-stranded and hairpin-looped structures, respectively, with an alternative secondary structure of two molecules of **7** forming an internal-looped one also being possible.

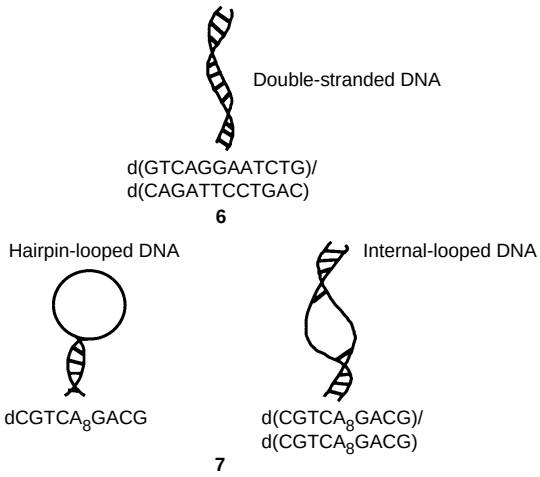
The L2-G peptide was immobilized on the SPR sensorchip,[8] and the octanucleotides were flowed across the sensorchip.† After the injection of 320 μM octanucleotides **2**–**5** to the flowcell, the response of the refractive index increased as the sample was flowed. Final intensities were approximately 1200, 400, 50 and 200 RU (resonance unit) for **2**, **3**, **4** and **5**, respectively, at 300 seconds after the injection. Although the refractive index for the ssDNA seems only three-fold greater than that for the ssRNA, the signal for **2** continued increasing even 300 seconds after the injection. Thus, affinity of **2** for the L2-G peptide is much greater than **3**. These results indicate **3** and **5** have much lower binding abilities to the peptide on the sensorchip than **2**, although these three nucleotides possess the same negative charges. Apparently, the neutral nucleotide of **4** could not bind to the L2-G peptide, suggesting the L2-G peptide recognizes nucleotides not only by their electronegativity but also by their backbone structure. A lower dissociation process for **2** was observed when the running buffer was flowed at 300 seconds after the injection, suggesting the interactions for the ssDNA with the L2-G peptide are not simple.



Ac-Ile-Arg-Met-Lys-Ile-Gly-Val-Met-Phe-Gly-Asn-Pro-Glu-Thr-Thr-Thr-Gly-Gly-Asn-Ala-Leu-Lys-Phe-Tyr-NH_2

L2 site — Helix G

1



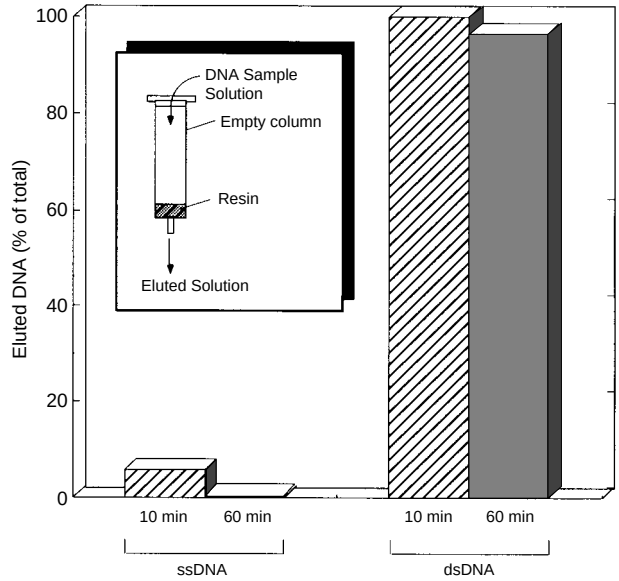


Fig. 1 Binding properties of the control and L2-G columns to ssDNA of d(TAATACGACTCACTATAGGG) and dsDNA of d(TAATACGACTCACTATAGGG)/d(CCCTATAGTGAGTCGTATTA) at room temperature. Percentages of the eluted DNAs through these columns against the total DNA amount were calculated with absorbance differences at 260 nm of the eluted solutions through the control and the L2-G columns.

When **6** was flowed over the sensorchip with the L2-G peptide, only an increment of 100 RU was observed. Even if a high concentration of this dsDNA (up to 6 mM) was examined, the response was at most 200 RU. Thus, the L2-G peptide bound to ssDNA, but not to ssRNA, unnatural nucleotides or structured DNAs. This binding affinity was also supported by circular dichroism (CD) spectra (data not shown). On the other hand, as we expected, **7** had a smaller signal (450 RU) than **2**, because this DNA strand was designed to form a hairpin-looped structure easily. However, there was a relatively larger response of **7** than that of the ssRNA **3**. This result might be due to its ability to form not only a hairpin-looped structure but also an internal-looped structure of the self-complementary duplex which would be bound by the L2-G peptide.

Furthermore, we prepared an ssDNA affinity column by using the ssDNA recognizing peptide (L2-G peptide). The peptide was attached to a polystyrene supported PEG-PS resin (PerSeptive Biosystems). This resin was put in an empty column (L2-G column), then ssDNA, d(TAATACGACTCACTATAGGG), or dsDNA, d(TAATACGACTCACTATAGGG)/d(CCCTATAGTGAGTCGTATTA), was poured into the column.‡ In addition, a control resin without attached amino acid residues on the PEG-PS resin was prepared as a control column.

Absorbance at 260 nm of the elution buffers was compared. In the case of the ssDNA, no decrement of absorbance was observed for elution through the control column. However, through the L2-G column, much less absorbance (Abs_{260} = 0.04) compared to that before pouring (Abs_{260} = 0.73) was found. The percentages of the eluted DNA strand flowing through these resins against the total DNA amount are shown in Fig. 1. This indicates that most ssDNAs bound to the L2-G resin even after mixing for 10 min. When the dsDNA was poured into these columns, no significant difference of the measured absorbance was observed between the control column and the L2-G column. Thus, there is no specific interaction of the L2-G resin with the dsDNA. These results indicate that the L2-G resin is very useful as an ssDNA binding support.

In summary, the 24 amino acid peptide (L2 and helix G domains of RecA protein) discriminated ssDNA from ssRNA, unnatural nucleotides, and dsDNA. The helix G amino acids were required for the discrimination property of the L2-G peptide. A useful support recognizing DNA secondary structures was able to be developed with the L2-G peptide derived from the DNA binding protein.

We would like to thank Drs H. Mihara (Tokyo Institute of Technology), H. Aoyagi and T. Niidome (Nagasaki Univ.) for advice on the synthesis of the peptide, Dr H. Sawai (Gunma Univ.) for providing 2′–5′ linked U_8, and Dr Y. Hayakawa (Nagoya Univ.) for providing ethoxy-dT_8. We also acknowledge Dr P. C. Bevilacqua at Penn. State University for critical reading of the manuscript and helpful comments. This work was supported in part by Grants-in-Aid from the Ministry of Education, Science, Sports and Culture, Japan, and from the 'Research for the Future' Program of the Japan Society for the Promotion of Science.

Footnotes and References

* E-mail : sugimoto@konan-u.ac.jp

† Nucleotide concentrations are expressed as base concentrations. All experiments were conducted in a buffer including 100 mM NaCl, 10 mM Na_2HPO_4 and 1 mM Na_2EDTA (pH 7.0).

‡ 100 mg of the L2-G resin and the PEG-PS resin without attached amino acid residues was prepared. The quantity of the peptide bound to the resin was calculated to be 1.7 mmol if the synthesis was completed with 100% yield. The L2-G resin or the PEG-PS resin was put into a empty column (named L2-G column or control column, respectively). The single-stranded DNA of d(TAATACGACTCACTATAGGG) or double-stranded DNA of d(TAATACGACTCACTATAGGG)/d(CCCTATAGTGAGTCGTATTA) was dissolved in a buffer including 100 mM NaCl, 10 mM Na_2HPO_4 and 1 mM Na_2EDTA (pH 7.0) to a concentration of 1.3 nmol. These DNA solutions were poured into the two columns and the percentages of the bound DNA on the resin were estimated from the absorbance at 260 nm of the eluted solution.

1 R. R. Meyer and P. Laine, *Microbiol. Rev.*, 1990, **54**, 342.
2 C. M. Radding, *J. Biol. Chem.*, 1991, **266**, 5355.
3 S. C. Kowalczykowski, *Annu. Rev. Biophys. Biophys. Chem.*, 1991, **20**, 539.
4 R. M. Story, I. T. Weber and T. A. Steitz, *Nature*, 1992, **355**, 318.
5 R. M. Story, D. K. Bishop, N. Kleckner and T. A. Steitz, *Science* 1993, **259**, 1892.
6 O. N. Voloshin, L. Wang and R. D. Camerini-Otero, *Science*, 1996, **272**, 868.
7 S. Nakano, M. Sasaki and N. Sugimoto, *Nucleic Acids Symp. Ser.*, 1995, **34**. 61.
8 N. Sugimoto and S. Nakano, *Nucleic Acids Symp. Ser.*, 1996, **35**, 303.

Received in Cambridge, UK, 1st August 1997; 7/05575G

Typeset and printed by Black Bear Press Limited, Cambridge, England