ChemComm

Design of a nucleobase-conjugated peptide that recognizes HIV-1 RRE IIB RNA with high affinity and specificity

Tsuyoshi Takahashi,[a] Keita Hamasaki,[a] Ichiro Kumagai,[a] Akihiko Ueno[a] and Hisakazu Mihara*[ab]

[a] *Department of Bioengineering, Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Nagatsuta, Yokohama 226-8501, Japan. E-mail: hmihara@bio.titech.ac.jp*

[b] *Form and Function, PRESTO, Japan Science and Technology Corporation, Tokyo Institute of Technology, Nagatsuta, Yokohama 226-8501, Japan*

Received (in Cambridge) 6th December 1999, Accepted 26th January 2000

A nucleobase-conjugated peptide, derived from HIV-1 Rev, was designed and synthesized, and the peptide which contained the cytosine moiety at the L-α-amino acid side chain bound RRE IIB RNA with high affinity and high specificity.

RNA–protein interactions play important roles in nature. Many cellular functions, including transcription, RNA splicing and translation, depend on the specific interaction of proteins and RNA. RNA-binding proteins seldom target fully double-stranded tracts for recognition, but interact with secondary structural domains such as hairpin loops, internal loops, and bulges of RNA.[1] In most cases, an RNA-binding domain of a protein forms a suitable conformation such as an α-helix or a β-strand to recognize a structured RNA, and amino acids orientated exactly in the protein structure make specific contacts to the RNA backbone and bases, resulting in high affinity and high specificity in the RNA binding. The study of the design and synthesis of novel molecules that recognize the structure of RNA might lead to drugs targeting RNA.

Peptide nucleic acid (PNA) is a DNA mimic with the nucleobases on a pseudopeptide backbone composed of *N*-(2-aminoethyl)glycine units.[2] A PNA molecule has the ability to efficiently and sequence-specifically bind both single-stranded DNA and RNA as well as double-stranded DNA. However, a simple PNA molecule itself may not recognize highly structured RNA with high specificity, because PNA alone would not have the ability to form the various conformations that proteins or peptides do. We have attempted to design peptides containing nucleobases that combine the advantages of peptides (formation of rigid structure such as α-helices) and nucleobases (specific recognition of a base in RNA). For this purpose, we have utilized an artificial L-α-amino acid, with a nucleobase at the side chain, in peptides for keeping the peptide conformation.

In order to demonstrate the applicability of a nucleobase amino acid (NBA) in peptides, we have chosen the regulatory protein of virion expression (Rev) of human immunodeficiency virus type-1 (HIV-1)[3–6] from among the enormous range of examples of proteins which bind RNA specifically. The Rev protein binds the corresponding response region of HIV-1 mRNA (RRE), and this protein–RNA interaction plays a key role in HIV-1 virus replication. The arginine-rich domain (34–50) of the Rev protein binds specifically the stem-loop IIB region of RRE RNA (Fig. 1) by forming an α-helix conformation.[7] The α-helix potential of the Rev_{34-50} peptide affects the binding affinity and specificity of the peptide to RRE IIB RNA.[7,8] Moreover, it has been proposed that the Gln36 residue in the Rev_{34-50} peptide is placed close to the guanine-48 base in the internal loop region of the RRE IIB RNA by NMR structural analyses.[9]

On the basis of these informations, we designed and synthesized the nucleobase-conjugated peptide, $Q36C_{NBA}$, in which L-α-amino γ-cytosine butanoic acid (cytosine nucleobase amino acid: C_{NBA}) was introduced instead of Gln36 in Rev_{34-50} (Fig. 1), and examined the effect of the introduction of the C_{NBA} unit on the peptide conformation and the RNA-binding affinity. It was expected that the cytosine base in $Q36C_{NBA}$ could interact with the guanine-48 base in RRE IIB RNA and that the interaction may lead to an increase in affinity and specificity of the peptide for the RNA. Furthermore, the chiral C_{NBA} unit possibly gives a rigid conformation in the peptide. For the elucidation of the structural importance of $Q36C_{NBA}$, $Q36C_{PNA}$, in which Gln36 and Ala37 were replaced by Nielsen type cytosine PNA (C_{PNA}),[2] was also designed (Fig. 1). The two amino acids were replaced by C_{PNA}, because a PNA monomer has the length equivalent of two amino acids in the main chain. Q36A and Q36GG were used to evaluate the effect of introduction of the cytosine moiety in $Q36C_{NBA}$ and $Q36C_{PNA}$, respectively (Fig. 1). To increase the stability of the α-helix structure, the N-terminal amino and C-terminal carboxy groups of the peptides were succinylated and amidated, respectively.[7,8] To introduce the C_{NBA} unit in the peptide, (2*S*)-4-(N^4-benzyloxycarbonylcytosin-1-yl)-2-(fluoren-9-ylmethoxycarbonyl)aminobutanoic acid [Fmoc-C_{NBA}(Z)-OH] was synthesized according to the method in ref. 10 with some modifications. Peptides were synthesized by the solid phase method using an Fmoc-strategy[11] and purified by HPLC with high purity ($>98\%$ on analytical HPLC). The peptides $Q36C_{NBA}$ and $Q36C_{PNA}$ gave a molecular ion peak at *m/z* 2602.3 [$(M + H)^+$] (calc. = 2603.9) and 2588.6 [$(M + H)^+$] (calc. = 2589.9), respectively, by matrix assisted laser desorption ionization time-of-flight mass spectrometry. RRE IIB RNA was prepared by *in vitro* transcription from a synthetic DNA template using T7 RNA polymerase and purified by polyacrylamide gel electrophoresis.[12]

Circular dichroism studies revealed that none of the peptides including Rev_{34-50} had a rigid structure in aqueous buffer (pH

(*a*)

	36
Rev_{34-50}	suc-TR**Q**ARRNRRRRWRERQR-NH_2
$Q36C_{NBA}$	suc-TR**C_{NBA}**ARRNRRRRWRERQR-NH_2
$Q36C_{PNA}$	suc-TR**C_{PNA}**RRNRRRRWRERQR-NH_2
Q36A	suc-TR**A**ARRNRRRRWRERQR-NH_2
Q36GG	suc-TR**GG**RRNRRRRWRERQR-NH_2

(*b*) RRE IIB

5'-GGCUGGUCUGG**G**CGCAGCGUCA A
3'-CCGACCGGACAGGCGUCGCAGU
U A (48)

(*c*) C_{NBA} C_{PNA}

Fig. 1 (*a*) Amino acid sequences of Rev_{34-50} and designed peptides; (*b*) the secondary structure of RRE IIB RNA; (*c*) structures of C_{NBA} and C_{PNA}.

This journal is © The Royal Society of Chemistry 2000

Table 1 α-Helix contents of the peptides in TFE solution, and dissociation constants (K_d) of the peptides with RRE IIB RNA

Peptide	α-Helicity (%)[a]	K_d/10^{-9} mol dm^{-3}
Rev$_{34-50}$	59	3.4 ± 0.3
Q36C$_{NBA}$	60	1.7 ± 0.1
Q36C$_{PNA}$	41	10.9 ± 0.7
Q36A	76	3.4 ± 0.1
Q36GG	52	9.9 ± 1.1

[a] α-Helix contents were estimated from $[\theta]_{222}$ according to the method of ref 14.

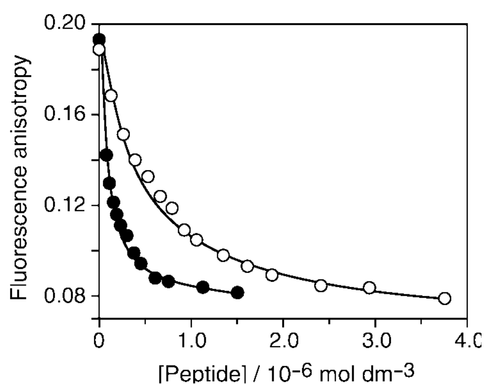

Fig. 2 Fluorescence anisotropy of Rhod-Rev containing RRE IIB RNA as a function of Q36C$_{NBA}$ (●) and Q36C$_{PNA}$ (○) concentration in 1.0 × 10^{-2} mol dm^{-3} Tris·HCl buffer (pH 7.5) containing 1.0 × 10^{-1} mol dm^{-3} KCl, 1.0 × 10^{-3} mol dm^{-3} $MgCl_2$, and 5.0 × 10^{-4} mol dm^{-3} EDTA at 25 °C. [Rhod-Rev] = 1.0 × 10^{-8} mol dm^{-3} and [RRE IIB] = 2.5 × 10^{-8} mol dm^{-3}.

7.5). On the contrary, in TFE, which is known to be an α-helix forming solvent,[13] the peptides showed an α-helix CD pattern. The α-helicity[14] (Table 1) of Q36C$_{NBA}$ was 60%, similar to that of Rev$_{34-50}$ (59%), suggesting that the C$_{NBA}$ unit had a similar potential to the Gln36 residue in the Rev peptide for forming an α-helix structure. However, the α-helicity of Q36C$_{PNA}$ was decreased to 41% despite it being in TFE solution. These results indicate that the α-helix structure is disturbed by the flexible and achiral PNA unit, but not by the rigid and chiral C$_{NBA}$.

The binding properties of the peptides with RRE IIB RNA were evaluated by competition assay[15] using the Rev peptide modified with 5-carboxytetramethylrhodamine at the N-terminal (Rhod-Rev) as a fluorescence tracer. The dissociation constant (2.1 × 10^{-9} mol dm^{-3}) of Rhod-Rev with RRE IIB RNA was calculated from the anisotropy increase of the tracer upon the addition of the RNA using an equation with 1:1 stoichiometry.[15] In the mixture of Rhod-Rev (1.0 × 10^{-8} mol dm^{-3}) and RRE IIB RNA (2.5 × 10^{-8} mol dm^{-3}), fluorescence anisotropy values were decreased by the addition of the designed peptides as a competitor, affording the free Rhod-Rev (Fig. 2). This competition assay revealed that Rev$_{34-50}$ bound RRE IIB RNA strongly with a dissociation constant (K_d) of 3.4 x 10^{-9} mol dm^{-3} (Table 1). Interestingly, Q36C$_{NBA}$ showed a K_d value of 1.7 × 10^{-9} mol dm^{-3}. Even though Q36C$_{NBA}$ has almost the same α-helix potential as Rev$_{34-50}$, it could bind RRE IIB RNA 2.0-fold stronger than Rev$_{34-50}$. These results suggest that the cytosine moiety of Q36C$_{NBA}$ contributes to the interaction of the peptide with the RNA by making a contact such as a hydrogen bond to a nucleobase of the RNA. Q36A showed the same affinity for the RNA (K_d = 3.4 × 10^{-9} mol dm^{-3}) as Rev$_{34-50}$. This result indicates that the higher α-helix content (76% in TFE) of Q36A enables it to maintain the same binding affinity as the Rev$_{34-50}$. On the other hand, the binding affinity of Q36C$_{PNA}$ for the RNA was significantly decreased (K_d = 10.9 × 10^{-9} mol dm^{-3}) to a level similar to Q36GG (K_d = 9.9 × 10^{-9} mol dm^{-3}), even if it has the same base in its sequence. These findings suggest that the introduction of the C$_{PNA}$ residue at the Gln36 position does not improve the binding affinity with RNA. Furthermore, the decrease of the binding affinity is also attributed to the lower α-helix potential of Q36C$_{PNA}$ (41% in TFE).

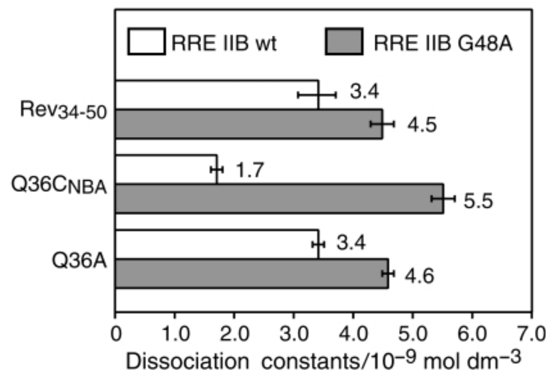

Fig. 3 Comparison of dissociation constants of the peptides with the wild-type and the mutant RNA.

In order to elucidate the selectivity of the designed peptide Q36C$_{NBA}$ for RRE IIB RNA, a mutant RNA (G48A), in which the guanine-48 base of RRE IIB RNA was replaced by adenine, was also prepared. Rev$_{34-50}$ bound G48A RNA with a dissociation constant of 4.5 × 10^{-9} mol dm^{-3}, comparable to that of the wild-type RRE IIB RNA (1.3-fold decrease) (Fig. 3). Q36A bound the mutant RNA with an affinity similar to Rev$_{34-50}$ (K_d = 4.6 × 10^{-9} mol dm^{-3}). Q36C$_{NBA}$ bound the mutant RNA with K_d = 5.5 × 10^{-9} mol dm^{-3}, 3.2-fold weaker than that of the wild-type RNA, that is, the affinity of Q36C$_{NBA}$ to the mutant RNA was lower than that of the Rev$_{34-50}$ and Q36A peptides. It seems that the cytosine base in Q36C$_{NBA}$ may interact with the guanine-48 moiety in RRE IIB RNA, and the interaction enhances the binding specificity to RRE IIB RNA more than Gln36 in Rev$_{34-50}$ does.

In conclusion, the nucleobase-conjugated peptide, Q36C$_{NBA}$, derived from HIV-1 Rev was successfully designed and synthesized. Introduction of the cytosine moiety as an artificial L-α-amino acid to the Rev$_{34-50}$ peptide increased the binding affinity and specificity to RRE IIB RNA, without changing the conformational properties of the peptide. This study should lead to a new strategy applicable to the construction of molecules that recognize a specific structure of an RNA molecule using various nucleobase amino acids.

Notes and references

1 I. W. Mattaj, *Cell*, 1993, **73**, 837.
2 P. E. Nielsen and G. Haaima, *Chem. Soc. Rev.*, 1997, **73**; H. Knudsen and P. E. Nielsen, *Nucleic Acids Res.*, 1996, **24**, 494.
3 C. Jain and J. G. Belasco, *Cell*, 1996, **87**, 115.
4 M. H. Malim, J. Hauber, S.-Y. Le, J. V. Maizel and B. R. Cullen, *Nature*, 1989, **338**, 254.
5 B. K. Felber, M. Hadzopoulou-Cladaras, C. Cladaras, T. Copeland and G. N. Pavlakis, *Proc. Natl. Acad. Sci. U.S.A.*, 1989, **86**, 1495.
6 M. Emerman, R. Vazeux and K. Peden, *Cell*, 1989, **57**, 1155.
7 R. Tan, L. Chen, J. A. Buettner, D. Hudson and A. D. Frankel, *Cell*, 1993, **73**, 1031.
8 R. Tan and A. D. Frankel, *Biochemistry*, 1994, **33**, 14579.
9 J. L. Battiste, H. Mao, N. S. Rao, R. Tan, D. R. Muhandiram, L. E. Kay, A. D. Frankel and J. R. Williamson, *Science*, 1996, **273**, 1547.
10 A. Lenzi, G. Reginato and M. Taddei, *Tetrahedron Lett.*, 1995, **10**, 1713.
11 E. Atherton and R. C. Sheppard, *Solid Phase Peptide Synthesis: A Practical Approach*,IRL Press, Oxford, 1989.
12 J. F. Milligan and O. C. Uhlenbeck, *Methods Enzymol.*, 1989, **180**, 51.
13 N. E. Zhou, B.-Y. Zhu, C. M. Kay and R. S. Hodges, *Biopolymers*, 1992, **32**, 419.
14 J. M. Scholtz, H. Qian, E. J. York, J. M. Stewart and R. L. Baldwin, *Biopolymers*, 1991, **31**, 1463.
15 K. Hamasaki, J. Killian, J. Cho and R. R. Rando, *Biochemistry*, 1998, **37**, 656.

Communication a909579i