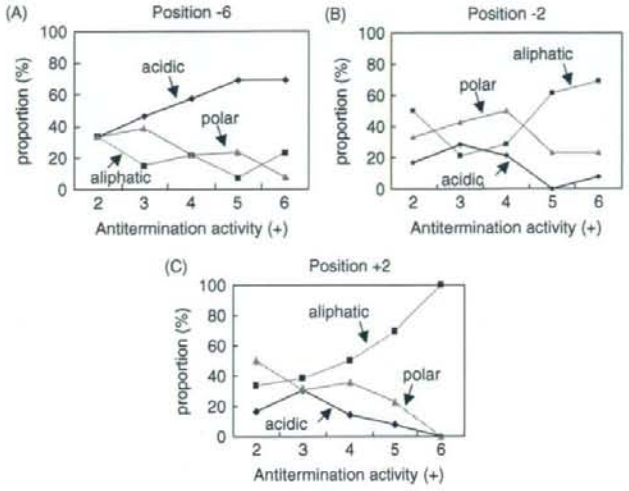

Figure 5 The relationship between the antitermination activity and the distribution of amino acids at each randomized position of the glutamine-containing peptides. (A), (B), and (C) indicate the distribution of amino acids at positions −6, −2, and +2 relative to the glutamine residue, respectively. The *y*-axis indicates the proportion of acidic amino acids (diamonds and solid lines), polar amino acids (triangles and gray lines), and aliphatic amino acids (squares and dotted lines) in percent (%), and the *x*-axis indicates antitermination activity (+).

Table 2 Antitermination activities and dissociation constants of RRE–peptide interactions

Peptide	Sequence[a]	Antitermination activity (X-gal[b])	K_d (nM)
DLA	RRRDRRLRQRARRR	6	0.5
RLA	RRRRRRLRQRARRR	3.5	30
DNA	RRRDRRNRQRARRR	4	2
DLE	RRRDRRLRQRERRR	4	1.5
R_8QR_5	RRRRRRRRQRRRRR	0	
DLA Q9N	RRRDRRLRNRARRR	0	
Rev	TRQARRNRRRRWRR	3	25
RSG-1.2	RDRRRRGSRPSGAERRRR	4	
K1	DRRQRRRDRQRRRRR	6	
BIV Tat	MSGPRPRGTRGKGRRIRR	0	

[a] Amino acid positions differing from the DLA peptide are underlined.
[b] The number of plusses in the colony color assay using RRE reporters.

binding, with the DLA peptide binding 50-fold more strongly to the RRE than the HIV Rev peptide (Table 2).

The Effect of Individual Amino Acid Substitutions at the Nonarginine Positions of the DLA Peptide on RRE-binding

In order to understand the role of the three nonarginine positions −6, −2, and +2 in more detail, and to exclude the effect of possible interactions between these positions, all 20 amino acid substitutions at each of these three nonarginine positions were prepared, and their antitermination activities were determined. The relationship between antitermination activity, which correlates with binding affinity, and (i) α-helical propensity [31], (ii) residue volume [32], and (iii) hydrophobicity of the amino acid residue was analyzed. Figure 6 shows the relationship between antitermination activity and α-helix propensity (panels A, B, and C) or residue volume (panels D, E, and F) for positions −6, −2, and +2. First, in the case of α-helical propensity, no clear correlation with antitermination was observed for position −6, while for positions −2 and +2 a fairly good correlation was observed, with amino acids with higher α-helix propensities showing high antitermination activities. The only exceptions were tryptophan in the case of position −2 and tyrosine in the case of position +2, suggesting that these amino acids may be sterically interfering with RNA-binding. Indeed, when antitermination activities were plotted against the residue volumes of the amino acid side chains, low activities were observed when the large tryptophan residue was present at positions −6 and −2, and when a tyrosine residue was present at position +2. These results show that most amino acid residues could be accommodated into all three positions, and that amino acid propensity was important at positions −2 and +2. No correlation seemed to exist between hydrophobicity and antitermination activity, and data is not shown.

Alanine- and Lysine-scanning of Arginine Residues in the DLA Peptide

In order to identify arginine residues within the DLA peptide that are important for RRE-binding, each

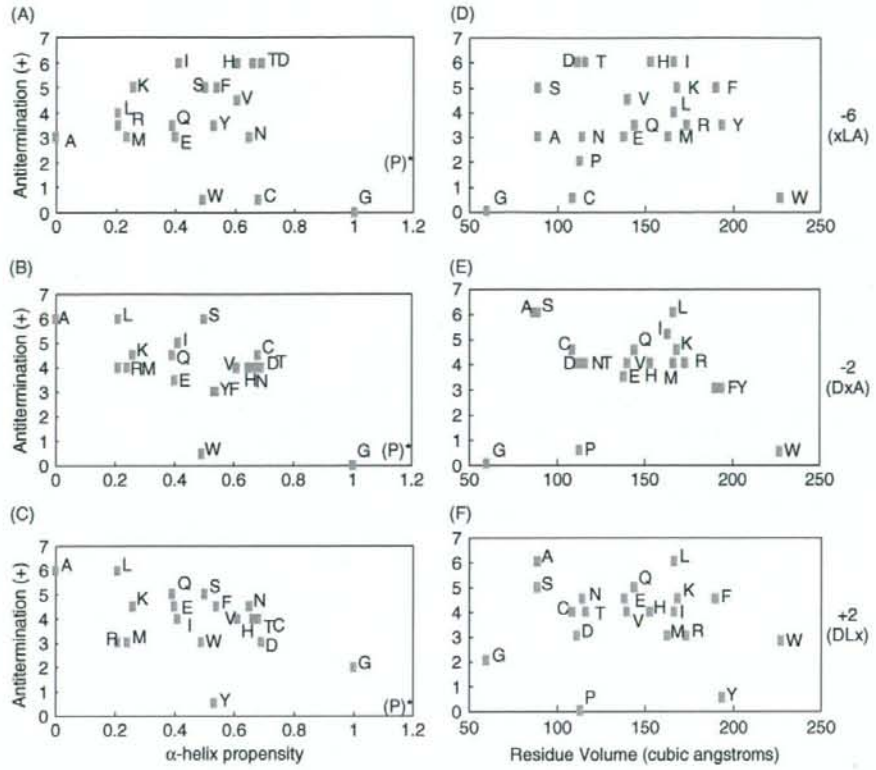

Figure 6 The relationship between antitermination activity and α-helix propensity or residue volume of amino acids at positions −6, −2, and +2 of the DLA peptide, relative to the glutamine residue. The α-helix propensity of proline has a value of 3.16 and lies outside of the region shown in the graph as indicated by the asterisk. This figure is available in colour online at www.interscience.wiley.com/journal/jpepsci.

arginine residue was substituted to either alanine or lysine residues, and antitermination activities were analyzed by the colony color assay (Table 3). Substitution of arginine residues with alanine residues resulted in a considerable decrease in antitermination activity to less than 4+. In particular, the R5A and R10A substitutions showed no activity (0+), and R6A, R8A, and R13A showed low activity (0.5+ ~ 2+). On the other hand, in the case of substitution with lysine residues, while R5K and R10K substitutions completely abolished activity (0+) as with the alanine substitutions, and the R6K, R8K, and R12K substitutions led to a moderate loss of activity (3+ to 4+), the R1K, R2K, R4K, R13K, and R14K substitutions maintained activity.

DISCUSSION

Selection of RRE-binding Peptides from Arginine-rich Peptide Libraries 1 ~ 3 (ARPL1~3)

In a previous study, selection of RRE-binding peptides from the arginine-based library ARPL1, a 15 mer polyarginine codon-mutagenized at the *N*-terminal 10 positions and flanked by alanine residues, resulted solely in the identification of peptides with a conserved glutamine at the 10th position [25]. In this study, in an attempt to identify a wider variety of sequences, possibly with a higher affinity toward the RRE, a second arginine-rich library, ARPL2, doped at all 15 positions within the same 15 mer polyarginine context as ARPL1, was constructed. In addition, since the glutamine-containing peptides were predicted to bind in an α-helical conformation, a third library, ARPL3, where the same codon-mutagenized 15 mer polyarginine as ARPL2 was flanked by glycine residues instead of alanine residues, was constructed. Since glycine residues are known to destabilize α-helix conformation, it was anticipated that peptides that bind to the RRE in a conformation other than an α-helix might be identified.

As a result, the peptides selected from the newly constructed ARPL2 and ARPL3 libraries were similar to those identified in the ARPL1 selection, and possessed a glutamine at the 9th or 10th amino acid from the

Table 3 Alanine- and lysine-scanning of the DLA peptide

	Peptide sequences[b]	Antitermination activity (X-gal[a])	
		X[b]	
		A	K
R1X	XRDRRRLRQRARRR	3	5
R2X	RXDRRRLRQRARRR	3	6
R4X	RRDXRRLRQRARRR	3	6
R5X	RRDRXRLRQRARRR	0	0
R6X	RRDRRXLRQRARRR	0.5	3
R8X	RRDRRRLXQRARRR	1	4
R10X	RRDRRRLRQXARRR	0	0
R12X	RRDRRRLRQRAXRR	4	4
R13X	RRDRRRLRQRARXR	2	6
R14X	RRDRRRLRQRARRX	4	6
DLA	RRDRRRLRQRARRR	6	
Rev	TRQARRNRRRRWRR	3	
RSG-1.2	RDRRRRGSRPSGAERRRR	4	
BIV Tat	MSGPRPRGTRGKGRRIRR	0	

[a] The number of plusses in the colony color assay using RRE reporters.
[b] Xs indicate the positions substituted for alanine or lysine residues.

N-terminus, and nonarginine residues at positions −9, −6, −2, or +2 relative to the glutamine residue (Table 1). Since these peptides resemble R6QR7, another peptide selected from a Rev-based library [24], it was assumed that the peptides bind to the RRE in an α-helical conformation, and that the glutamine residue recognizes the G–A base pair in a way analogous to the asparagine residue of the Rev peptide. When survival rates in the second round of the ARPL2 and ARPL3 selections are compared, substitution of the flanking alanines to glycines lead to considerable reduction from 2.7 to 0.9%, indicating that the proportion of RRE-binders in the ARPL3 library are reduced, supporting the assumption that the glutamine-containing peptides are binding to the RRE in an α-helical conformation. These results also show that simply substituting the alanines on the ends of the peptides to glycines alone was not sufficient to identify nonhelical peptides.

The Role of Nonarginine Residues in the DLA Peptide upon RRE-binding

In order to understand the role of the nonarginine residues in RRE-binding, the selected peptides were depicted in an axial view as an α-helical wheel (Figure 4). Interestingly, the alanine, asparatic acid, and glutamic acid were found to be located opposite of the glutamine residue, which is assumed to bind deeply into the major groove of the RRE and hydrogen bond to the critical G–A base pair. Therefore, the nonarginine residues do not seem likely to make specific contacts in the major groove, and are presumably important in the folding and stabilization of the α-helical conformation of the peptides. Indeed, alanine is known to have a high α-helical propensity [31], and the carboxylate group of asparatic and glutamic acid may be expected to form salt bridges with arginine and lysine side chains [33]. Most of selected peptides screened from the Q3L library that exhibited high activity toward the RRE had an acidic amino acid (aspartic or glutamic acid) at position −6, and aliphatic amino acids at positions −2 (primarily leucine) and +2 (primarily alanine) relative to glutamine (Figure 5). In the peptides selected from the Q3L library, an increased requirement for nonarginine residues at positions −6, −2, and +2 was observed compared to the K1 peptide. In addition, the preference for glutamic acid at position −6 in the case of the K1 peptide is diminished. These differences may be due to the slightly shorter length of the Q3L library compared to the ALPL1~3 libraries.

Next, the antitermination activity of all individual substitutions at positions −6, −2, and +2 were investigated, and a fairly clear correlation with α-helical propensity was observed at positions +2 and −2 (Figure 6), again suggesting that these residues are important in stabilization of the α-helix. On the other hand, the antitermination activities of amino acid substitutions at position −6 did not show any correlation with α-helix propensity or residue volume. Furthermore, while the aspartic acid substitution at position −6 showed high activity (6+), that of glutamic acid was considerably lower (3+) which does not correlate with the preference for glutamic acid observed in the K1 peptide context. This low activity of the clone with glutamic acid at the −6 position may rule out the involvement of salt bridge formation with basic amino acid side chains in the stabilization of the α-helix [34]. Other possibilities for the role of the aspartic acid residue at position −6 include stabilizing the macrodipole, and induction of RNA folding upon interaction with the negatively charged RNA.

The Role of the Arginine Residues in the DLA Peptide upon RRE-binding

The arginine residues can be thought to play two main roles upon high affinity and highly specific binding to the RRE. The first role involves the interaction of the positively charged guanidinium group of arginine forming an ionic bond with the negatively charged phosphates of the RNA backbone. The second role is the specific interaction of the guanidinium group with the base moiety of the RNA, a typical example being the base pair formation at the Hoogstein site of guanine from the major groove. When the activity of alanine and lysine substitutions at these positions were compared, the activity of the lysine substitution

DOI: 10.1002/psc

was substantially higher for positions R1, R2, R4, R13, and R14, suggesting that these arginine residues are involved in ionic interactions with the negatively charged phosphates of the RRE-binding. On the other hand, the activity of the lysine substitution is only moderately higher for R6 and R8, implying that the identity of the amino acid at this position is somewhat important. In contrast, both alanine and lysine substitutions at positions R5 and R10 showed no activity, suggesting that these two positions are critically important for specific recognition, most likely, of RRE bases in the major groove.

A Model for the Interaction of the Glutamine-containing α-Helical Arginine-rich Peptides with the RRE

While structural studies are needed to determine the detailed contacts between the glutamine-containing arginine-rich peptides and the RRE, and to understand the basis for high affinity binding, a general model for the interaction of the glutamine-containing arginine-rich peptide DLA with the RRE as shown in Figure 7 can be derived from the results above. First, amino acid requirements at the nonarginine residues strongly suggest that the peptides bind to the RRE in an α-helical conformation as depicted by the α-helical wheel. Next, based on the similarity of the glutamine-containing peptides obtained in this study with the R7QR8 peptide identified from a previous selection [24], the conserved glutamine is assumed to recognize to the G–A base pair of the RRE from the major groove. This is also supported by the similar nucleotide requirements for the binding of the glutamine-containing K1 peptide to the RRE with those of Rev-binding to the RRE. This places the three nonarginine residues at positions −6, −2, and +2 opposite the conserved glutamine residue in the helical conformation so that these amino acid side chains are likely to be exposed to solvent. As discussed above, the amino acids at positions +2 and −2 relative to the glutamine appear to be stabilizing the α-helical conformation of the peptide, while the role of the amino acid at position −6 relative to the glutamine is not clear.

Alanine- and lysine-scanning experiments indicated that the identity of the two arginine residues at position 5 (R5) and 10 (R10) were absolutely required for RRE binding, suggesting that these two residues were making specific contacts with the bases in the major groove. The guanidinium group of R5 and R10 may bind to the Hoogsteen site of guanine bases, as is observed in the Rev–RRE interaction, as well as those of the HIV and BIV Tat–TAR interactions and in the interaction of the RRE with the selected peptide RSG-1.2 [6,10,26,27,35]. In the case of the Rev peptide, an interaction between R35 and R39 of Rev with G67 and G70 of the RRE, respectively, was observed, while in the case of RSG-1.2, a single interaction between R14 of RSG-1.2 and G39 was observed. There are several other candidate guanine residues within the RRE that may make specific contacts with the arginine side chains identified in this study, including the G48–G71 base pair of which the Hoogstein site of G48 most likely faces the major groove. The presence of an extra interaction between the G–G base pair and an arginine residue may explain the increased affinity of the K1 and DLA peptides with the RRE compared to the Rev peptide. On the other hand, loss of activity upon substitution by alanine could be partially or completely recovered by substitution by lysine at positions 1, 2, 4, 6, 8, 13, and 14, indicating that these positions were primarily important for ionic interactions.

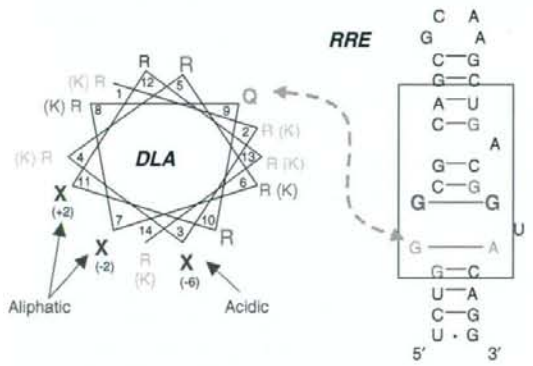

Figure 7 A model for the interaction of the DLA peptide and the RRE. The DLA peptide is depicted as an α-helical wheel with the conserved glutamine shown in light blue, the arginines important for RRE-binding shown in red, the arginines that can be replaced by lysine shown in green, and the three nonarginine residues shown in blue. The secondary structure of the RRE with nucleotides important for peptide binding boxed, the critical G–A base pair in light blue, and guanine residues that may specifically bind to arginine residues are indicated in red.

CONCLUSIONS

Several attempts to identify arginine-rich peptides that bind to the RRE have suggested that the RRE is particularly suited for accommodating α-helical peptides, with the RSG-1.2 peptide being a slight anomaly in that it binds in an extended-turn-helix conformation [26,27]. Combinatorial analysis of the amino acids important for the high affinity binding of the RRE and glutamine-containing arginine-rich peptides has led to a consensus peptide, the DLA peptide, with optimal binding affinity toward the RRE. Since the α-helical propensity of the nonarginine residues appears to directly correlate with RNA-binding affinity, it was shown that substitution of solvent exposed nonarginine residues provides a simple means

to adjust binding affinity. In another study, the substitution of the G–G base pair with an A–A base pair has been shown to lead to a considerable loss of binding to the K1 peptide, while this substitution is tolerated for the Rev and RSG-1.2 peptides. This suggests that the G–G base pair may be directly interacting with an amino side chain of the DLA peptide, and may account for at least part of the increased RRE affinity compared to the Rev peptide.

The DLA peptide ($K_d = 0.5$ nM) was found to bind to the RRE with an affinity ~50-fold stronger than the Rev peptide ($K_d = 25$ nM), as determined by a gel shift assay, and to our knowledge is the tightest RRE-binding peptide known. Since it has been shown that RRE-binding peptides can inhibit HIV Rev function in mammalian cells [23], the DLA and related peptides are attractive candidates for the development of peptide-based drug for the inhibition of HIV replication. The amino acid requirements elucidated in this study will serve as guidelines in the improvement of the binding affinity of these peptides through artificial modification, for example by stabilization of the α-helical conformation of the peptide [36]. While it is not clear whether the glutamine-containing α-helical peptides characterized in this study represent a general class of RNA-binding peptides, we believe that studies such as those conducted here in collaboration with structural studies of these complexes will lead to an increased understanding of RNA–polypeptide interactions.

MATERIALS AND METHODS

Construction of Combinatorial Libraries

Degenerate oligonucleotide libraries encoding ARPL2 and ARPL3 were synthesized using a codon-based mutagenesis procedure as previously described [23] using an Expedite System Nucleic Acids Synthesizer MODEL 8909 (PerSeptive Biosystems) with a prototype sequence encoding 15 arginines (5′-GAATCCCCATGGCCNNKNNKCGTCGGCGTAGGCGTCGG CGTAGGCGTCGGCGTAGGCGTCGGCGTGCAGCTGCTGCGAA TGCAGCAAATCC-3′ (ARPL2) and 5′-GAACCCATGGGTVNKVN KCGTCGGCGTAGGCGTCGGCGTAGGCGTCGGCGTAGGCGT CGGCGTGGTGGCGGGAATGA-3′ (ARPL3)] and a randomized sequence (5′-GAATCCCCATGGCC NNKNNK(VVK)$_{15}$GCAGCTG CTGCGAATGCAGCAAATCC-3′ (ARPL2) and 5′-GAA TCCCCAT GGGTVNKVNK(VVK)$_{15}$GGTGGCGGGAATGCA-3′ (ARPL3)], where N is an equimolar mixture of A, C, G and T, V is an equimolar mixture of A, C, and G, and K is an equimolar mixture of G and T. The library was synthesized at a 1:1 ratio of arginine: VVK codons as previously described [30]. The degenerate oligonucleotides were annealed to a primer (5′-GGATTTGCTGCATTC-3′ (ARPL2) and 5′-CATTCCCGCCACC-3′ (ARPL3)]. The synthetic library oligonucleotides were converted to double-stranded DNA in a total volume of 500 µl (2.6 µM ARPL, 1.8 µM primer, 0.3 mM each deoxynucleotide triphosphate, 4.5 mM $MgCl_2$, 1× PCR buffer) using Taq polymerase, and 120 ng was digested with *Nco* I and *Bsm* I. A second degenerate oligonucleotide library encoding Q3L, (5′-GAATCCCCATGGCCCGCCGTNNKCGTC GCCGTNNKCGTCAGCGTNNKCGTCGCCGTGCAGCTGCTGCG AATGCAGCAAATCC-3′) was annealed to a primer (5′-GGATTTGCTGCATTC-3′), and double-stranded DNA was synthesized by Taq polymerase, and product was digested with *Nco*I and *Bsm*I.

Selection of RRE-binding Peptides from the ARPL2 Library

RRE-specific peptides were identified in four steps using the antitermination system [25]. For the primary screen, the ARPL3 library insert (40 ng) was ligated into pBRN$^-$ expressor plasmid (1.2 µg) in a 1-ml volume using T4 DNA ligase (24 000 units, New England Biolabs (NEB)). Ligation mixtures were phenol extracted, and concentrated to 10 µl using MontagePCR (Millipore). Plasmids were electroporated into N567/pACK-RRE cells (17 × 80 µl) in 1-mm cuvettes at 2.0 kV using 1 µl of the above solution per electroporation. SOC medium [0.5% (w/v) yeast extract, 2% (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 20 mM MgSO4, 20 mM glucose] (5 ml) was added immediately after electroporation; cells were allowed to recover by incubating at room temperature (23 °C) for 6 h on a rotating shaker (40 rpm). Isopropyl-β-D-thiogalactopyranoside (IPTG) was added after 4 h of recovery to a final concentration of 0.2 mM to induce expression from the tac promoters. Cells were then plated onto tryptone plates containing ampicillin (50 µg/ml), chloramphenicol (20 µg/ml), IPTG (0.05 mM), and kanamycin (2.5 µg/ml) and incubated at room temperature (23 °C). A total of 1.8×10^7 transformants were obtained, representing 1.0×10^7 independent clones as estimated by plating cells before the 6-h recovery period on plates lacking kanamycin. After incubation at room temperature (23 °C) for 63 h on the kanamycin plates, surviving colonies were scraped off and plasmid DNA was isolated. Because both the pAC-RRE and pBR library plasmids were present, *Eco*RV digestion was performed to selectively digest pACK-RRE. Following phenol extraction and ethanol precipitation, the DNA was redissolved in H_2O (15 µl).

For the secondary screen, kanamycin selection was repeated by electroporating 1 µl of the above plasmid DNA solution (20 ng) into N567/pACK-RRE cells (80 µl) two times, spreading the cells on plates containing 2.5 µg/ml kanamycin, and incubating at room temperature (23 °C) for 68 h. Plasmid DNA was isolated as above and following *Eco*RV digestion, the pBR plasmid DNA was redissolved in 15 µl H_2O (50 ng/µl).

For the tertiary screen, N567/pAC-RRE cells (100 µl) were transformed using 2 µl of the pBR plasmid DNA (200 ng) from the secondary screen by heat shock, and incubated in the tryptone medium (1 ml) at 37 °C for 1 h. A portion of the culture was spread onto tryptone plates (150-mm diameter) containing 100 µg/ml ampicillin, 20 µg/ml chloramphenicol, 0.05 mM IPTG, 80 µg/ml of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), incubated at 37 °C for 19 h, and the proportion of blue colonies was scored [22].

For the quaternary screen, RNA-binding specificities of 48 of the darkest blue colonies were assessed. The pBR plasmid DNAs were isolated and activities were monitored using the LacZ colony color assay with N567/reporter cells containing pAC-RRE or pAC-BTAR plasmids.

Selection of RRE-binding Peptides from the ARPL3 Library

RRE-specific peptides were identified in four steps using the antitermination system [25]. For the primary screen, the ARPL3 library insert (120 ng) was ligated into pBRN$^-$ expressor plasmid (1.2 μg) in a 1-ml volume using T4 DNA ligase (24 000 units, NEB). Ligation mixtures were phenol extracted, ethanol precipitated using sodium acetate, and dissolved in 6 μl H_2O. Plasmids were electroporated into N567/pACK-RRE cells (6 × 80 μl) in 1-mm cuvettes at 2.0 kV using 1 μl of the above solution per electroporation. SOC medium (5 ml) was added immediately after electroporation; cells were allowed to recover by incubating at room temperature (23 °C) for 6 h on a rotating shaker (40 rpm). IPTG was added after 4 h of recovery to a final concentration of 0.2 mM to induce expression from the tac promoters. Cells were then plated onto tryptone plates containing ampicillin (50 μg/ml), chloramphenicol (20 μg/ml), IPTG (0.05 mM), and kanamycin (2.5 μg/ml) and incubated at room temperature (23 °C). A total of 1.1×10^8 transformants were obtained, representing 1.6×10^7 independent clones as estimated by plating cells before the 6-h recovery period on plates lacking kanamycin. After incubation at room temperature (23 °C) for 62.5 h on the kanamycin plates, surviving colonies were scraped off and plasmid DNA was isolated. Because both the pAC-RRE and pBR library plasmids were present, *EcoRV* digestion was performed to selectively digest pACK-RRE. Following phenol extraction, the DNA was desalted and concentrated to 15 μl using Montage PCR (Millipore).

For the secondary screen, kanamycin selection was repeated by electroporating 1 μl of the above plasmid DNA solution (50 ng) into N567/pACK-RRE cells (80 μl) two times, spreading the cells on plates containing 2.5 μg/ml kanamycin, and incubating at room temperature (23 °C) for 62 h. Plasmid DNA was isolated as above, and following *EcoRV* digestion, the pBR plasmid DNA was redissolved in 15 μl H_2O (50 ng/μl).

For the tertiary screen, N567/pAC-RRE cells (100 μl) were transformed using 0.5 μl of the pBR plasmid DNA (25 ng) from the secondary screen by heat shock, and incubated in the tryptone medium (1 ml) at 37 °C for 1 h. A portion of the culture was spread onto tryptone plates (150-mm diameter) containing 100 μg/ml ampicillin, 20 μg/ml chloramphenicol, 0.05 mM IPTG, 80 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), and incubated at 37 °C for 21 h, and the proportion of blue colonies was scored [22].

For the quaternary screen, RNA-binding specificities of 28 of the darkest blue colonies were assessed. The pBR plasmid DNAs were isolated and activities were monitored using the LacZ colony color assay (Harada and Frankel, 1998) with N567/reporter cells containing pAC-RRE or pAC-BTAR plasmids [29].

Screening of RRE-binding Peptides from the Q3L Library

For the primary screen, the Q3L oligonucleotide insert (6 ng) was ligated into 180 ng of pBR N$^-$ in a total volume of 150 μl using T4 DNA ligase (3600 units, NEB). Ligation mixtures were phenol extracted, ethanol precipitated using sodium acetate, and dissolved in 5.5 μl H_2O. Plasmids (1.5 μl) were electroporated into N567/pAC RRE (80 μl) and plated onto 50 X-gal plates, yielding ~6.6×10^5 transformants. Individual blue colonies (1248) were then grown to saturation in 96-well plates containing tryptone and antibiotics, cultures were pooled, and plasmid DNA was isolated.

In the secondary screen, pooled plasmid DNAs (1 μl) were then electroporated into fresh RRE-reporter cells (80 μl). A portion of the culture was spread onto X-gal plates incubated at 37 °C for 19 h, and the proportion of blue colonies was scored (Harada *et al.*, 1996). The blue colonies (70, 6+ ~ 2+) were picked, plasmid DNA was isolated, and the library region was sequenced.

Peptide-binding Assays

Internally labeled RNAs were transcribed *in vitro* using T7 RNA polymerase and [α-^{32}P]CTP (Perkin Elmer, 3000 Ci/mmol) as described [15]. Gel shift assays were carried out by incubating the internally labeled RNA (0.1 nM) with varying amounts of Rev, DLA, RLA, DNA, or DLE peptide at 4 °C in 10-μl binding mixtures containing 10 mM HEPES-KOH (pH 7.5), 100 mM KCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 1 mM dithiothreitol (DTT), 50 μg/ml tRNA, and 10% glycerol, and resolving RNA and RNA-peptide complexes on 10% polyacrylamide, 0.5× TBE (45 mM Tris/45 mM boric acid/1 mM EDTA, pH 8) gels at 4 °C. Apparent K_d is defined as the concentration of peptide required to shift 50% of the free RNA into the complex.

The Preparation and Analysis of Expression Plasmids for Substitutions at the Nonarginine Positions and for Alanine and Lysine Scanning at Arginine Positions

Expression plasmid coding for single amino acid substitution of the DLA peptides were prepared in the following manner. Oligonucleotides used are listed in Supplementary Table 2. First, DNA inserts with single randomized codons at one of the three nonarginine positions of the DLA peptide, XLA, DXA, and DLX were prepared. A solution containing one of the three oligodeoxynucleotides (8.8 mM), the primer 5′-CATTCGCAGCAGCTGC-3′ (8.8 mM), 0.3 mM each deoxynucleotide triphosphate, 4.5 mM $MgCl_2$, 1× PCR buffer, and Taq polymerase was heated at 90 °C for 5 min, slow cooled to room temperature, then heated at 72 °C for 10 min. The double-stranded DNA was digested with NcoI and BsmI, and introduced into the NcoI and BsmI sites of pBR N plasmid using T4 DNA ligase. The ligation mixtures were used to transform N567 cells, and cells were spread onto tryptone plates containing ampicillin and tetracycline. Several colonies were picked, pBR plasmid isolated, and the region coding the DLA peptide was sequenced (DYEnamic ET Terminater Cycle Sequencing Kit on an ABI PRIZM 377). Next, amino acid substitutions that could not be identified using the above method were obtained by introducing the DNA inserts prepared from the oligodeoxynucleotides, WLA, SLA, FLA, 2LA (F or Y at position −6), 3LA (M or K at position −6), DFA, DKA, D2A (P or H at position −2), D3A (Y, H, or D at position −2), DLK, DLP, DLW, DL3 (F or Y at position +2), DL4 (P or H at position +2) into the pBR plasmid as described above. Expression plasmids encoding the DLA peptide with either alanine or lysine substitutions of the ten arginine residues were prepared as described above using oligodeoxynucleotides containing substitutions of the arginine codon by either the alanine codon GCG or the

lysine codon AAG. Antitermination activity of each expression plasmid was carried out using the colony color assay using tryptone plates containing X-gal (80 mg/ml).

Supplementary Material

Supplementary electronic material for this paper is available in Wiley InterScience at: http://www.interscience.wiley.com/jpages/1075-2617/suppmat/

Acknowledgements

We thank Taiichi Sakamoto and Andrea Gamarnik for comments on the manuscript. This work was supported by a Grant-in-Aid for Scientific Research in Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan.

REFERENCES

1. Patel DJ. Adaptive recognition in RNA complexes with peptides and protein modules. *Curr. Opin. Struct. Biol.* 1999; **9**: 74–87.
2. Frankel AD. Fitting peptides into the RNA world. *Curr. Opin. Struct. Biol.* 2000; **10**: 332–340.
3. Cheng AC, Calabro V, Frankel AD. Design of RNA-binding proteins and ligands. *Curr. Opin. Struct. Biol.* 2001; **11**: 478–484.
4. Das C, Frankel AD. Sequence and structure space of RNA-binding peptides. *Biopolymers* 2003; **70**: 80–85.
5. Tan R, Chen L, Buettner JA, Hudson D, Frankel AD. RNA recognition by an isolated α-helix. *Cell* 1993; **73**: 1031–1040.
6. Battiste JL, Mao H, Rao NS, Tan R, Muhandiram DR, Kay LE, Frankel AD, Williamson JR. α-Helix major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* 1996; **273**: 1547–1551.
7. Legault P, Li J, Mogridge J, Kay LE, Greenblatt J. NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* 1998; **93**: 289–299.
8. Cai Z, Gorin A, Frederick R, Ye X, Hu W, Majumdar A, Kettani A, Patel DJ. Solution structure of P22 transcriptional antitermination N peptide-boxB RNA complex. *Nat. Struct. Biol.* 1998; **5**: 203–212.
9. Chen L, Frankel AD. A peptide interaction in the major groove of RNA resembles protein interactions in the minor groove of DNA. *Proc. Natl. Acad. Sci. U.S.A.* 1995; **92**: 5077–5081.
10. Puglisi JD, Chen L, Blanchard S, Frankel AD. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* 1995; **270**: 1200–1203.
11. Calnan BJ, Biancalana S, Hudson D, Frankel AD. Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. *Genes Dev.* 1991; **5**: 201–210.
12. Aboul-ela F, Karn J, Varani G. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J. Mol. Biol.* 1995; **253**: 313–332.
13. Heaphy S, Finch JT, Gait MJ, Karn J, Singh M. Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich "bubble" located within the rev-responsive region of viral mRNAs. *Proc. Natl. Acad. Sci. U.S.A.* 1991; **88**: 7366–7370.
14. Tiley LS, Malim MH, Tewary HK, Stockley PG, Cullen BR. Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type 1 Rev protein. *Proc. Natl. Acad. Sci. U.S.A.* 1992; **89**: 758–762.
15. Chen L, Frankel AD. An RNA-binding peptide from bovine immunodeficiency virus Tat protein recognizes an unusual RNA structure. *Biochemistry* 1994; **33**: 2708–2715.
16. Weeks KM, Crothers DM. Major groove accessibility of RNA. *Science* 1993; **261**: 1574–1577.
17. Frankel AD, Smith CA. Induced folding in RNA-protein recognition: more than a simple molecular handshake. *Cell* 1998; **92**: 149–151.
18. Kjems J, Brown M, Chang DD, Sharp PA. Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element. *Proc. Natl. Acad. Sci. U.S.A.* 1991; **88**: 683–687.
19. Kjems J, Calnan BJ, Frankel AD, Sharp PA. Specific binding of a basic peptide from HIV-1 Rev. *EMBO J.* 1992; **11**: 1119–1129.
20. Bartel DP, Zapp ML, Green MR, Szostak JW. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* 1991; **67**: 529–536.
21. Giver L, Bartel D, Zapp M, Pawul A, Green M, Ellington AD. Selective optimization of the Rev-binding element of HIV-1. *Nucleic Acids Res.* 1993; **21**: 5509–5516.
22. Harada K, Martin SS, Frankel AD. Selection of RNA-binding peptides in vivo. *Nature* 1996; **380**: 175–179.
23. Harada K, Martin SS, Tan R, Frankel AD. Molding a peptide into an RNA site by in vivo peptide evolution. *Proc. Natl. Acad. Sci. U.S.A.* 1997; **94**: 11887–11892.
24. Tan R, Frankel AD. A novel glutamine-RNA interaction identified by screening libraries in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 1998; **95**: 4247–4252.
25. Peled-Zehavi H, Horiya S, Das C, Harada K, Frankel AD. Selection of RRE RNA binding peptides using a kanamycin antitermination assay. *RNA* 2003; **9**: 252–261.
26. Gosser Y, Hermann T, Majumdar A, Hu W, Frederick R, Jiang F, Xu W, Patel DJ. Peptide-triggered conformational switch in HIV-1 RRE RNA complexes. *Nat. Struct. Biol.* 2001; **8**: 146–150.
27. Zhang Q, Harada K, Cho HS, Frankel AD, Wemmer DE. Structural characterization of the complex of the Rev response element RNA with a selected peptide. *Chem. Biol.* 2001; **8**: 511–520.
28. Iwazaki T, Li X, Harada K. Evolvability of the mode of peptide-binding by an RNA. *RNA* 2005; **11**: 1364–1373.
29. Sugaya M, Nishimura F, Katoh A, Harada K. Tailoring the peptide-binding specificity of an RNA by combinations of specificity-altering mutations. *Nucleosides, Nucleotides, Nucleic Acids*, 2008 (in press).
30. Harada K, Frankel AD. *In vivo* selection of specific RNA-binding polypeptides using a transcriptional antitermination reporter assay. In *RNA: Protein Interactions: A Practical Approach*, Smith CWJ (ed.). Oxford University Press: Oxford, 1998; 217–236.
31. Pace CN, Scholtz JM. A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* 1998; **75**: 422–427.
32. Zamyatin AA. Protein volume in solution. *Prog. Biophys. Mol. Biol.* 1972; **24**: 107–123.
33. Ghosh T, Garde S, Garcia AE. Role of backbone hydration and salt-bridge formation in stability of α-helix in solution. *Biophys. J.* 2003; **85**: 3187–3193.
34. Olson CA, Spek AJ, Shi Z, Vologodskii A, Kallenbach NR. Cooperative helix stabilization by complex Arg-Glu salt bridges. *Proteins* 2001; **44**: 123–132.
35. Puglisi JD, Tan R, Calnan BJ, Frankel AD, Williamson JR. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* 1992; **257**: 76–80.
36. Mills NL, Daugherty MD, Frankel AD, Guy RK. An alpha-helical peptidomimetic inhibitor of the HIV-1 Rev-RRE interaction. *J. Am. Chem. Soc.* 2006; **128**: 3496–3497.

Nucleosides, Nucleotides and Nucleic Acids, 27:534–545, 2008

ISSN: 1525-7770 print / 1532-2335 online
DOI: 10.1080/15257770801944493

Taylor & Francis
Taylor & Francis Group

TAILORING THE PEPTIDE-BINDING SPECIFICITY OF AN RNA BY COMBINATIONS OF SPECIFICITY-ALTERING MUTATIONS

Maki Sugaya,[1] Futoshi Nishimura,[2] Akira Katoh,[1] and Kazuo Harada[2]

[1]*Department of Materials and Life Science, Faculty of Science and Technology, Seikei University, Musashino, Tokyo, Japan*
[2]*Department of Life Sciences, Tokyo Gakugei University, Koganei, Tokyo, Japan*

□ *In this study, the ability to tailor the peptide-binding specificity of an RNA was investigated. First, variants of the Rev-response element (RRE) RNA with different specificities toward the natural binding partner, Rev, and two RRE-binding aptamers, the RSG-1.2 and the Kl peptides, were identified. Next, hybrid RRE mutants with combinations of two sets of specificity-altering substitutions were tested for peptide-binding specificity. It was shown that in most cases the results of the combination of individual mutations were of an additive nature, therefore providing a way to manipulate the peptide-binding specificity of an RNA in a predictable manner.*

Keywords HIV RRE RNA; arginine-rich peptide; bacterial reporter system; peptide-binding specificity; rational design

INTRODUCTION

Biochemical and structural studies on the interaction of arginine-rich peptides and their RNA sites has considerably increased our understanding of RNA-polypeptide interactions.[1,2] In particular, the interaction of the HIV Rev peptide and the RRE RNA has been extensively studied. The Rev peptide, a short peptide corresponding to the RNA-binding domain of the HIV Rev protein (Figure 1B), has been shown to bind to an internal loop region of the RRE consisting of U45 to C51 and U66 to A75 of the RRE (Figure 1A).[3–5] In vitro selection experiments suggested the presence of a non-canonical G48-G71 base pair in the internal loop that can be replaced by an isostructural A48-A71 base pair (Figure 1A).[6,7] The NMR structure

Received 25 December 2007; accepted 24 January 2008.
This work was supported by a Grant-in-Aid for Scientific Research in Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan (KH).
Address correspondence to Kazuo Harada, Department of Life Sciences, Tokyo Gakugei University, Nukuikita-machi 4-1-1, Koganei, Tokyo 184-8501, Japan. E-mail: harada@u-gakugei.ac.jp

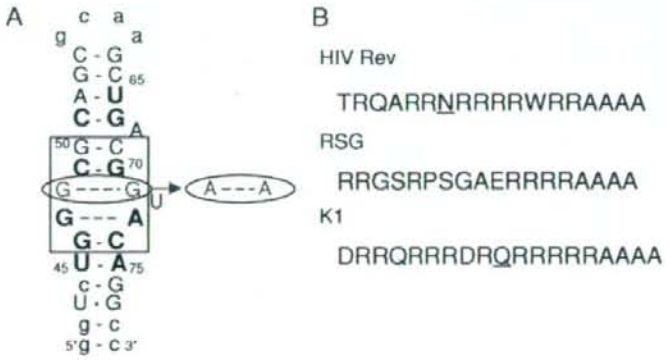

FIGURE 1 (A) The secondary structure of the RRE with nucleotides important for Rev-binding are shown in bold and those important for RSG-binding are boxed. The G-G base pair that may covary to an isostructural A-A base pair, and which is important for Kl-binding is shown in gray. (B) The amino acid sequences of the HIV Rev peptide and two selected RRE-binding peptides, RSG and Kl. The critical aspargine residue of Rev and the conserved glutamine residue of Kl are underlined.

of the Rev-RRE complex later confirmed the formation of the G48-G71 base-pair as well as a G47-A73 base pair that binds to a critical asparagine residue of the Rev peptide.[8]

Several peptides that bind to a similar region of the RRE, but with higher affinities and using a different binding strategy as the Rev peptide have been identified from random peptide libraries.[9] The RSG-1.2 peptide, which binds to the RRE several times more strongly than the Rev peptide, was selected from a relatively simple random library consisting of the three amino acids, arginine (R), serine (S), and glycine (G), followed by mutagenesis and reselection (Figure 1B).[10,11] While the helical Rev peptide binds along the major groove of the RRE in a manner similar to the binding of α-helices to the major groove of double-stranded DNA,[8] the RSG-1.2 peptide was found to bind to a similar region of the RRE in an unstructured-turn-helix conformation with the helix axis of the peptide almost perpendicular to that of the RRE.[12,13] In another selection of RRE-binding peptides from a complex library consisting of a polyarginine doped by codon-based mutagenesis using 12 amino acids, a putative α-helical high-affinity RRE-binding peptide, the Kl peptide, was identified (Figure 1B).[14] The Kl peptide contained a conserved glutamine residue that may bind to the G47-A73 base pair of the RRE in a manner analogous to the asparagine residue of the Rev peptide.

On the other hand, we have previously shown that single nucleotide substitutions or similarly small changes in the HIV Rev-response element (RRE) RNA sequence that result in well-defined secondary structural changes can dramatically alter the specificity of the RRE toward the Rev and RSG-1.2 peptides.[15] When the RRE nucleotides important for RSG-1.2-binding were determined, a considerable difference was observed compared to those

of the Rev peptide as illustrated in Figure 1A. Strikingly, several nucleotide substitutions in the upper stem region of the RRE, where the nucleotide requirement for Rev- and RSG-binding differed, were found to switch the peptide-binding specificity of the RRE from a bifunctional Rev- and RSG-1.2-binding mode to either a Rev-specific or a RSG-1.2-specific mode.

In this study, we expanded our studies on the peptide-binding specificity of RRE mutants to include the Kl peptide, and found that RRE mutants could be classified into a number of groups depending on their specificity toward the Rev, RSG-1.2, and Kl peptides. We then examined the peptide-binding specificity of hybrid RNAs containing combinations of these specificity-altering base substitutions. It was shown that in most cases the effect of the individual base substitutions were of an additive nature, therefore providing a way to manipulate the peptide-binding specificity of an RNA in a predictable manner.

RESULTS AND DISCUSSION

Determination of Nucleotide Requirements for the Binding of the RRE to the Kl Peptide

In this study, in order to expanded our studies on the peptide-binding specificity of RRE mutants to include the Kl peptide in addition to the Rev and RSG-1.2 peptides, we first determined the nucleotide requirements for the binding of RRE to the Kl peptide. This was carried out by selecting for Kl-binding RRE variants from a doped RRE library using a bacterial reporter system (Figure 2) and comparing the selected sequences shown

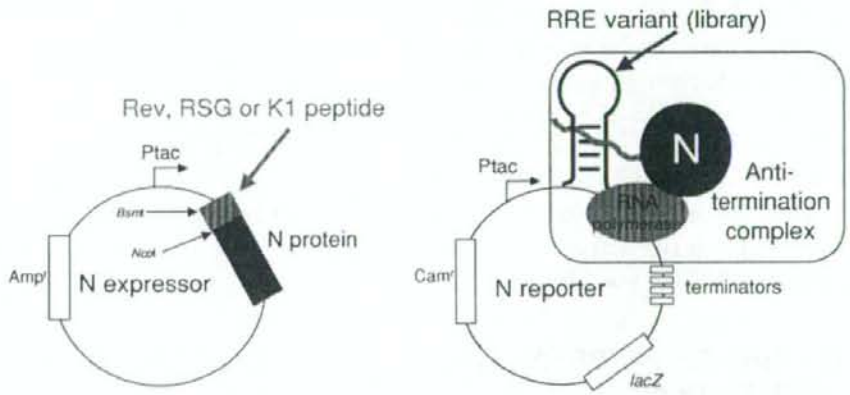

FIGURE 2 The bacterial two-plasmid reporter system for measuring transcriptional antitermination by heterologous RNA-binding polypeptides fused to the λ N protein. The RNA-binding peptide, in this case the Rev, RSG-1.2, or Kl peptide, fused to the N protein, is expressed under the control of a tac promoter from a pBR plasmid. The reporter plasmid contains the LacZ gene, also under the control of the tac promoter, so that binding of the Rev, RSG-1.2, or Kl peptide to the RNA site of interest, in this case RRE variants of a doped RRE library, results in reporter gene expression.

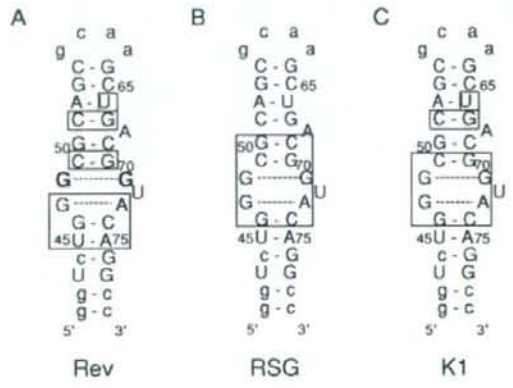

FIGURE 3 A comparison of RRE nucleotides important for Rev, RSG, and Kl peptide binding. Nucleotides within the RRE that are important for binding to the Rev (A), RSG (B), and Kl (C) peptides are boxed. The G-G base-pair highlighted in bold can be replaced by an A-A base-pair for Rev and RSG-binding.

in Table 1.[15] The differences in the nucleotide requirements of the RRE toward the Rev, RSG, and Kl peptides are summarized in Figure 3.

When the clones with Kl- and Rev-binding activities similar to those of the wildtype RRE were compared (Table 1, upper 22 clones), the nucleotide requirements for Kl-binding were found to be similar to those of the Rev peptide, except that the G-G base pair did not covary to an A-A base pair as in the case of the Rev and RSG peptides (Figure 1A). This suggested that the Kl peptide may be directly contacting the G-G base pair, unlike in the case of the Rev and RSG peptides, where the G-G to A-A base pair substitution is tolerated.[6,7,15] Indeed, clone **Gl***, a G-G to A-A variant identified in a previous study,[15] was found to have considerably lower activity toward the Kl peptide (Table 1). Next, when RRE variants with reduced activity toward the Rev peptide were compared (clones 18, 19, and 21), unlike in the case of Rev-binding, the U45-A75 base-pair was not conserved. In order to confirm the importance of this base pair for Rev-binding, and not for Kl-binding, a number of previously identified RRE-variants with substitutions at these positions were tested for Rev- and Kl-binding. As a result, a Kl-specific RRE variant, **G16***, derived from clone **G16** (Table 1) with a single nucleotide substitution (A75G) was identified.

The Specificity of RRE Variants Toward the Rev, RSG-1.2, and Kl Peptides

Various Rev- and/or RSG-specific RRE variants **Gl***, **G9**, **G30** and **R26** that were identified in a previous study, as well as the Kl-specific RRE variant **G16*** identified in this study were tested for their ability to bind to the Rev, RSG, and Kl peptides using the bacterial reporter system (Figure 2),[10]

TABLE 1 Kl-binding RRE variants and their antitermination activity against Kl, Rev, and BIV Tat peptides

Clone	doped region[b]	Colony color assay (X-gal)[a] Kl	Rev	Tat
RRE	GGUCUGGGGCGCAGCGCAAGCUGACGGUACAGGCC	6	3	0
40	GGUCUGGGGCGCA**A**CG**A**AAG**U**UGACGGUACAGGCC	6	4	0
3	GGUCUGGGGCGCAG**U**GC**U**AGCUGACGGUACAGGCC			
36	GGUCUGGGGCGCA**A**CGCAAG**U**UGACGGUACAGGCC			
49	GGUCUGGGGCGCA**C**CGCA**C**GG**U**GACGG**A**ACAGGCC	6	4	0
1	GGUCUGGGGCGCAGC**U**CAAGCUGACGG**A**ACAGGCC			
8	GGUCUGGGGCGCAGCGCA**U**GCUGACGG**A**ACAGGCC			
32	GGUCUGGGGCGCAGC**U**CA**U**GCUGACGGUACAGGCC	6	4	0
41	GGUCUGGGGCGCAGC**UG**AAGCUGACGGUACAGGCC			
43	GGUCUGGGGCGCAGCGC**UC**GCUGACGGUACAGGCC			
12	GGUCUGGGGCGCAGCGC**C**AGCUGACGGUACAGGCC	6	4	0
42	GGUCUGGGGCGCAGCGCA**U**GCUGACGGUACAGGCC			
46	GGUCUGGGGCGCAGCGC**G**AGCUGACGGUACAGGCC			
25	GGUCUGGGGCGCAGCGC**C**A**U**CUGACGGUACAGGCC	6	3	0
7	GGUCUGGGGCGCAGCGCA**G**GCUG**U**CGGUACAGGCC	6	3	0
4	GGUCUGGGGCGCAG**U**GCAA**U**CUGACGGUACAGGCC	6	4	0
9	GGUCUGGGGCGCAGCGCA**CA**CUGACGGUACAGGCC	6	4	0
2	GGUCUGGGGCGCAG**U**GCAAGCUGACGGUACAGGCC			
10	GGUCUGGGGCGCAG**AG**G**T**A**U**CUGACGGUACAGGCC	6	3	0
5	GGUCUGGGGCGCAGC**A**CAA**U**CUGACGGUACAGGCC	6	3	0
31	GGUCUGGGGCGCAGCGCAA**A**CUGACGGUACAGGCC			
20	GGUCUGGGGCGCAGCGC**C**AGCU**CG**CGGUACAGGCC	6	3	0
48	GGUCUGGGGCGCAG**U**GG**G**AAGCUG**CA**GGUACAGGCC	6	4	0
18	GGUC**G**GGGGCGCAGCGCA**C**GCUGA**G**GGUAC**C**GGCC	6	1	0
15	GGUCUGGGGCGCAG**UU**C**U**AGCUGA**U**GGUACAGGCC	5	4	0
14	GGUCUGGGGCGCAGCGC**C**AGCUG**UU**GGUACAGGCC	5	2	0
19	GGUC**G**GGGGCGCAGCG**A**AAGCUGACGGUAC**U**GGCC	6	2	0
16	GGUCUGGGGCG**U**GCGCAAGCUGACGGUACAGGCC	5	3	0
21	GGUC**G**GGGGCGCAGCGCAAGCUGACGGUAC**U**GGCC	5	0	0
17	GGUCUGGGGCGCAGCG**U**AAG**G**UGACGGUACAGGCC	5	5	0
6	GGUCUGGGGCGCAG**U**GCA**U**GCUG**U**CGG**G**ACAGGCC	4	2	0
G1*	GGUCUGG**A**CGCAGCGCAAGCUGACG**AC**ACAGGCC	1	5	-
G16	GGUCUGGGGCGCAG**U**GCA**U**GCUGACGGUAC**G**GGCC	0	6	ND

[a]X-gal colony color assays were performed with the Rev, Kl, and BIV Tat (negative control) peptides as described.[14] Numbers represent + 's used to score blue color intensity.

[b]The nucleotide sequence of the RRE used as a starting point for doped library design with the nucleotides known to be important for Rev-binding are shadowed. The nucleotide sequences of the selected RRE variants are shown with those differing from the wildtype RRE in bold and underlined.

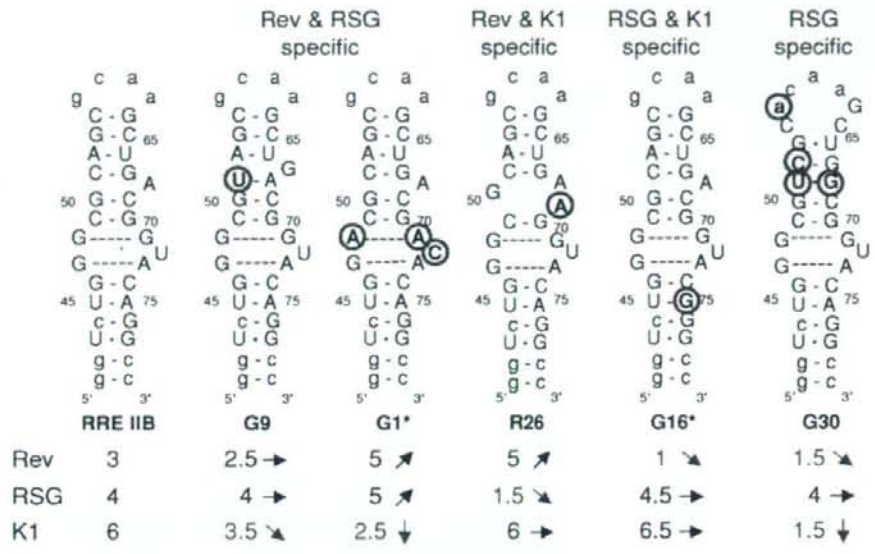

FIGURE 4 The predicted secondary structures of the RRE and RRE mutants with distinct peptide-binding specificities and their in vivo peptide-binding activities. The numbers represent the intensity of colony color scored using the LacZ reporter system. The arrows indicate the increase or decrease of in vivo activity of the mutant RRE relative to the wildtype RRE, where horizontal arrows represent no change (±0.5 units), diagonal arrows represent significant differences of up to 3 units, and vertical arrows indicate large changes of greater than 3 units. Significant decreases in β-galactosidase expression are indicated in gray, and lead to four different classes of peptide-binding specificities as indicated.

and the results are summarized in Figure 4. In the bacterial reporter system, antitermination complex formation mediated by the RNA-peptide interaction results in β-galactosidase expression. Although the orientation of the RNA-peptide complex appears to influence the stability of the antitermination complex and the resulting antitermination activity, when comparing related sets of RNA-peptide complexes, β-galactosidase expression scored by intensity of colony color on tryptone/agar plates containing 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (Xgal) has been shown to correspond to binding affinity. Those colonies with blue color similar to that of the wildtype RRE and Rev peptide (Kd = 60 nM), the RSG peptide (Kd = 12 nM), and the Kl peptide (Kd = 0.5 nM) scored as 3+, 4+, and 6+, respectively, with one plus unit generally corresponding to 3- to 10-fold differences in binding affinity.[16]

The base substitutions in clone **G9** and **Gl*** lead to loss of binding to, for the most part, only the Kl peptide, while those in **R26** and **G16*** resulted in loss of binding to the RSG and Rev peptides, respectively, and those in **G30** lead to the loss of binding to both Rev and Kl peptides. As a result, the mutants could be classified into a total of four peptide-specificity groups, the Rev- & RSG-specific **G9** and **Gl***, the Rev- & Kl-specific **R26**, the RSG-1.2- & Kl-specific **G16***, and the RSG-specific **G30.** In the case of mutant **Gl*** and

R26, an increase in the binding toward the Rev peptide, and in the case of **Gl***, an increase was also observed for the RSG-1.2 peptide.

The Peptide-Binding Specificity of Hybrid RRE Mutants with Combinations of Specificity-Altering Substitutions

To examine whether the base substitutions in clones **G9**, **Gl***, **R26**, **G16***, and **G30** could be combined to create new specificities, for example a Rev-specific clone or a Kl-specific clone, in a predictable manner, hybrid RRE mutants with all possible combinations of two sets of substitutions were tested for peptide-binding specificity using the bacterial reporter system. Of the ten combination of two sets of mutations, the **G9** and **G30** mutations overlapped and could not be combined, leaving nine combinations that could be tested.

Antitermination activities of the nine possible hybrid RREs against the Rev, RSG-1.2, and Kl peptides were determined, and the results are shown in Table 2. In the **G9/R26** and **R26/G30** hybrids, the pairs of base substitutions both resulted in distict secondary structural changes in the upper stem

TABLE 2 In vivo peptide-binding activities of hybrid RREs containing combinations of specificity-altering base substitutions[a]

		Rev / RSG	Rev / K1	RSG / K1	RSG
	Clone # (RRE) Rev 3 RSG 4 K1 6	G1* 5 (+2) 5 (+1) 2.5 (-3.5)	R26 5 (+2) 1.5 (-2.5) 6 (+)	G16* 1 (-2) 4.5 (+0.5) 6.5 (+0.5)	G30 1.5 (-1.5) 4 (-) 1.5 (-4.5)
	G9 2.5 (-0.5) 4 (-) 3.5 (-2.5)	4 (+1) 4.5 (+0.5) 0.5 (-3.5)	1.5 1 1.5	2 (-) 4 (-) 3 (-)	-
Rev / RSG	G1* 5 (+2) 5 (+1) 2.5 (-3.5)		1 1 1	4 (+1) 7 (+3) 3 (-)	1 2 1
Rev / K1	R26 5 (+2) 1.5 (-2.5) 6 (+)			2 (-) 0.5 (-3.5) 6 (-)	1.5 0.5 0
RSG / K1	G16* 1 (-2) 4.5 (+0.5) 6.5 (+0.5)				0.5 (-2.5) 4 (-) 1.5 (-1.5)

[a]Combinations that lead to loss of activity against all three peptides are shown with a white background. Those combinations that lead to hybrid RREs with new binding specificities are shown with the gray background. The arrows indicate the increase or decrease of in vivo activity of the mutant and hybrid RREs relative to the wildtype RRE, where horizontal arrows represent no change (±0.5 units), diagonal arrows represent significant differences of up to 3 units, and vertical arrows indicate large changes of greater than 3 units. The first values represent the intensity of colony color scored using the LacZ reporter system, and the values in parenthesis represent the change in intensity compared to wildtype RRE.

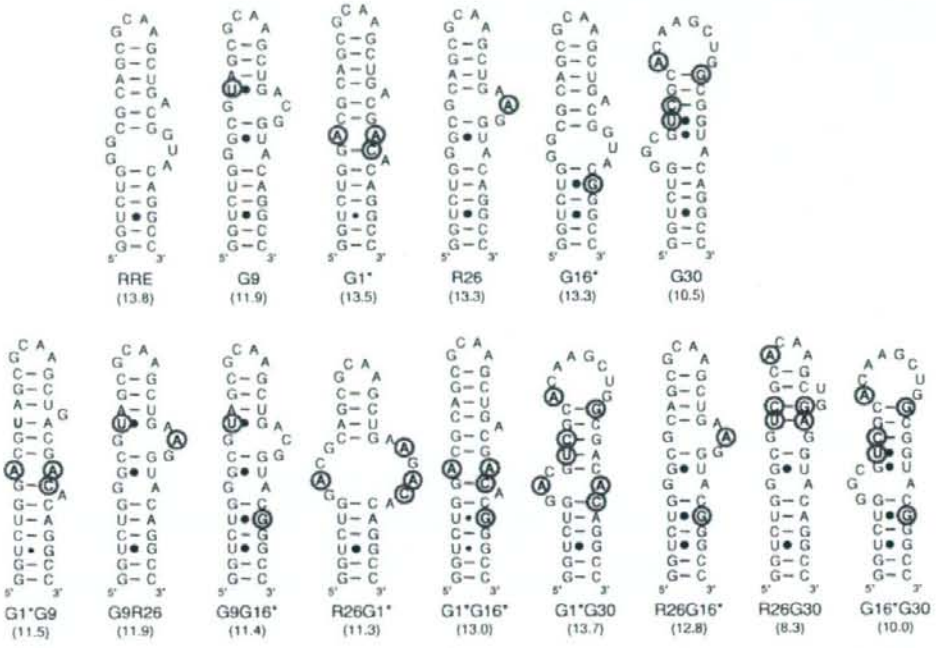

FIGURE 5 The lowest free-energy structures of mutant and hybrid RREs and their calculated free energy ($-\Delta G^{\circ}_{37}$) as determined by MulFold version 2.

regions and would be expected to alter or destabilize the conformation of the other. Therefore, not surprisingly, an almost complete loss of binding to all three peptides was observed. It should be added, however, that the **R26/G30** hybrid would be expected to have no activity against all three peptides if the peptide-binding specificities of the two sets of substitutions are assumed to be additive.

Of the remaining seven hybrid RREs, the peptide-binding specificities of the five hybrids highlighted in gray in Table 2 appear to be of an additive nature, with the sum of the specificity change of the individual sets of substitutions agreeing well with the experimental values. For example, combination of the **G9** substitution, where the change in antitermination activity against the Rev, RSG, and Kl peptides was (−0.5, 0, −2.5), and the **Gl*** substitution, which was (2, 1, −3.5), resulted in a hybrid RRE **G9/G1*** with a specificity change of (1, 0.5, −5.5), which is similar to the calculated specificity of (1.5, 1, −6). Since both the **G9** and **Gl*** substitutions were Rev & RSG-specific, combination of the two resulted in a hybrid RRE with reduced Kl-binding, and as a consequence, improved Rev- and RSG-1.2-specificity. A similar additivity was observed for the remaining four hybrids **G9/G16***, **G1*/G16***, **R26/G16***, and **G16*/G30**, and the difference between the calculated values and the experimental values were almost all within 1 colony color unit. The only exceptions were the 1.5 colony color

unit differences observed for the Rev-binding activity of **G9/G16*** and the RSG-binding activity of **G1*/G16***. In the case of **G9/G16***, this appeared to be because the negative effect of **G16*** on Rev-binding was not dominant, whereas in the case of **G1*/G16***, the positive effect that the **G1*** mutation has against the RSG-1.2 peptide appeared to be dominant over the negative effect of **G16***. It is particularly worth noting that the combination of the Rev & Kl-specific mutation of **R26** with the RSG-1.2- & Kl-specific mutation of **G16*** resulted in the creation of a novel highly Kl-specific hybrid RRE. In addition, combination of the RSG-1.2- & Kl-specific mutation of **G16*** with the RSG-1.2-specific mutation of **G30** resulted in a decrease in Rev- and Kl-binding to yield a more specific RSG-1.2-binding hybrid RRE.

The low-binding affinity of the **G1*/G30** toward all three peptides may be due to the formation of a stable alternative secondary structure as predicted by an RNA folding algorithm, MulFold version 2.0. The majority of the hybrid RREs including those that exhibited the predicted activities such as the **G1*/G9** hybrid folded into an overall secondary structure similar to the wildtype RRE, and appeared to require minimum rearrangement in the internal loop region to form the peptide-binding site. On the other hand, the **G1*/G30** hybrid folded into a considerably different secondary structure with an alternative upper stem- loop that is expected to be difficult to rearrange to form the peptide-binding site. The **G1*/R26** hybrid, which was anticipated to yield a new Rev-specific hybrid, did not show activity toward all three peptides, suggesting that the two mutations, although separated by one base pair, are altering the conformation of the other mutants.

CONCLUSIONS

In conclusion, the results show that the peptide-binding specificity of an RNA can be manipulated in a semi-rational manner by combinations of specificity-altering base substitutions, provided that the pairs of substitutions are sufficiently separated in space and do not result in the formation of stable alternative structures. A similar strategy is also likely to be applicable to aptamers so that their ligand-binding specificity may be modified to create novel binding specificities or fine-tuned by reducing nonspecific binding. In general, the effects of the individual substitutions were of an additive nature except for the positive effect of the **G1*** mutation against the Rev and RSG-1.2 peptides, which was dominant over the negative effect of the second mutation. While further structural studies will be necessary to understand the exact nature of the changes in specificity observed in this study, the results demonstrate the adaptability and modularity of RNA structure in creating new binding specificities, which is in contrast to the case of peptides, where a switch in RNA-binding specificity has been shown to require several amino acid substitutions to support.[17] This difference may be attributed to the ability that a given RNA has to adopt a multitude

of relatively stable structures predominantly through alternate Watson-Crick base pairing patterns, while retaining sufficient flexibility to adapt these structures to different binding partners.[18] This multidimensionality of conformational space that RNA has for peptide binding was likely to have been utilized in the evolution of RNA-protein interactions, and an understanding of this mechanism is expected to provide insights on such processes as the emergence of novel RNA viruses and drug resistance.

EXPERIMENTAL

In Vivo Selection of Kl Peptide-Binding RRE Variants

A DNA cassette prepared in a previous study encoding a library of RRE variants, in which a 26 nucleotide region that included U45-C54 and G64-A75 of the wildtype sequence and a GCAA tetraloop cap (Figure 1A) was doped at a frequency of 24% per residue, was used.[15] This level of doping was expected to result in an average of 5–6 substitutions per sequence, and a representation of most of the RRE variants with 2 substitutions, including the G-G to A-A substitution observed in previous in vitro selections[6,7] when 10^5 sequences were screened.

In the primary selection, this doped library DNA was introduced into the pACK plasmid carrying the NPT II reporter and transformed into Kl peptide expressor cells (N567/pBR Kl) (Figure 2).[14] Of the estimated 1.8×10^6 transformants, 4.2% survival was observed on tryptone plates containing kanamycin (5 μg/ml). In a secondary screen to eliminate reporter plasmid-related false positives, the PCR-amplified library insert was reintroduced into the pAC plasmid carrying the LacZ reporter, retransformed into N567/pBR Kl cells, and 40.2% of the colonies exhibited blue color on plates containing X-gal. Reporter gene expression was scored by intensity of colony color,[11,16] with those colonies with blue color similar to that of the wildtype RRE and Kl peptide scored as 6+, and those similar to wildtype RRE and Rev peptide scored as 3+. Fifty-nine colonies with blue colors scored as 4+ to 6+ were picked, plasmid isolated, and the library portion of the pAC plasmid was sequenced. After excluding overlapping clones, a total of 30 unique clones were obtained. Representative clones were tested for antitermination activity against the Kl and Rev peptides, as well as the BIV Tat peptide (Table 1). None of the clones showed antitermination activity in BIV Tat expressor cells, indicating that false positives had been eliminated.[11]

The Preparation of Expression Plasmids for RRE Variants and in vivo Peptide-Binding Analysis

The double-stranded DNA insert encoding the RRE variant sequence was prepared by annealing two oligonucleotides complementary at the

3′-ends, and second strand synthesis using Taq polymerase. The oligonucleotides used are listed in Supplementary Table 1. A solution containing the corresponding sense oligodeoxynucleotide (8.8 μM), and antisense oligodeoxynucleotide (8.8 μM), 0.3 mM dNTP, 4.5 mM $MgCl_2$, 1 × PCR buffer, and Taq polymerase was heated at 90°C for 5 minutes, slow cooled to room temperature, then incubated at 72°C for 10 minutes. The double stranded DNA was digested with PstI and BamHI, and introduced into the PstI and BamHI sites of pAC plasmid using T4 DNA ligase. The ligation mixtures were used to transform N567 cells, and cells were spread onto tryptone plates containing ampicillin and tetracycline. Several colonies were picked, pAC plasmid isolated, and the region coding the RRE variant was sequenced on an ABI 3130 sequencer using BigDye Terminator v3.1 Cycle sequencing Kit (ABI). Antitermination activity of each expression plasmid was carried out using the colony colour assay using tryptone plates containing X-gal (80 μg/ml).

REFERENCES

1. Patel, D.J. Adaptive recognition in RNA complexes with peptides and protein modules. *Curr. Opin. Struct. Biol.* **1999**, 9, 74–87.
2. Frankel, A.D. Fitting peptides into the RNA world. *Curr. Opin. Struct. Biol.* **2000**, 10, 332–340.
3. Kjems, J.; Brown, M.; Chang, D.D.; Sharp, P.A. Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element. *Proc. Natl. Acad. Sci. USA* **1991**, 88, 683–687.
4. Kjems, J.; Calnan, B.J.; Frankel, A.; Sharp, P.A. Specific binding of a basic peptide from HIV-1 Rev. *EMBO J.* **1992**, 11, 1119–1129.
5. Tiley, L.S.; Malim, M.H.; Tewary, H.K.; Stockley, P.G.; Cullen, B.R. Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type 1 Rev protein. *Proc. Natl. Acad. Sci. USA* **1992**, 89, 758–762.
6. Bartel, D.P.; Zapp, M.L.; Green, M.R.; Szostak, J.W. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **1991**, 67, 529–536.
7. Giver, L.; Bartel, D.; Zapp, M.; Pawul, A.; Green, M.; Ellington, A.D. Selective optimization of the Rev-binding element of HIV-1. *Nucleic Acids Res.* **1993**, 21, 5509–5516.
8. Battiste, J.L.; Mao, H.; Rao, N.S.; Tan, R.; Muhandiram, D.R.; Kay, L.E.; Frankel, A.D.; Williamson, J.R. α-Helix major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* **1996**, 273, 1547–1551.
9. Das, C.; Frankel, A.D. Sequence and structure space of RNA-binding peptides. *Biopolymers* **2003**, 70, 80–85.
10. Harada, K.; Martin, S.S.; Frankel, A.D. Selection of RNA-binding peptides in vivo. *Nature* **1996**, 380, 175–179.
11. Harada, K.; Martin, S.S.; Tan, R.; Frankel, A.D. Molding a peptide into an RNA site by in vivo peptide evolution. *Proc. Natl. Acad. Sci. USA* **1997**, 94, 11887–11892.
12. Gosser, Y.; Hermann, T.; Majumdar, A.; Hu, W.; Frederick, R.; Jiang,, R.; Xu, W.; Patel, D.J. Peptide-triggered conformational switch in HIV-1 RRE RNA complexes. *Nat. Struct. Biol.* **2001**, 8, 146–150.
13. Zhang, Q.; Harada, K.; Cho, H.S.; Frankel, A.D.; Wemmer, D.E. Structural characterization of the complex of the Rev response element RNA with a selected peptide. *Chem. Biol.* **2001**, 8, 511–520.
14. Peled-Zehavi, H.; Horiya, S.; Das, C.; Harada, K.; Frankel, A.D. Selection of RRE RNA binding peptides using a kanamycin antitermination assay. *RNA* **2003**, 9, 252–261.
15. Iwazaki, T.; Li, X.; Harada, K. Evolvability of the mode of peptide-binding by an RNA. *RNA* **2005**, 11, 1364–1373.

16. Sugaya, M.; Nishino, N.; Katoh, K.; Harada, K. Amino acid requirement for the high affinity binding of a selected arginine-rich peptide with HIV Rev-response element RNA. *J. Pept. Sci.* **2008** (in press).
17. Franklin, N.C. Morphing molecular specificities between Arm-peptide and NUT-RNA in the antitermination complexes of bacteriophages lambda and P22. *Mol. Microbiol.* **2004**, 52, 815–822.
18. Leulliot, N.; Varani, G. Current topics in RNA-protein recognition: control of specificity and biological function through induced fit and conformational capture. *Biochemistry* **2001**, 40, 7947–7956.

SUPPLEMENTAL TABLE 1 Oligonucleotides used to construct expressor plasmids encoding RRE variants

G9/G30-sense	5′-gggCTGCAGGTCGACGCTCTTAAAAATTAAGGTCTGGGCGT-3′;
G16*/R26-sense	5′-gggCTGCAGGTCGACGCTCTTAAAAATTAAGGTCTGGGCGCAGCGCAAGCTG-3′
G9-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCTGTACCGTCAGCTTGCGCTACGCCCAGACCTTAA-3′
G30-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCTGTACCGCCAGCTTGTGCGACGCCCAGACCTTAA-3′
G16*-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCCGTACCGTCAGCTTGCGCTGCGC-3′
R26-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCTGTACCTTCAGCTTGCGCTGCGC-3′
G1*-sense	5′-gggCTGCAGGTCGACGCTCTTAAAAATTAAGGTCTGGACG-3′
G1*/G9-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCTGTGTCGTCAGCTTGCGCTACGTCCAGACCTTAA-3′
G1*/G16-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCCGTGTCGTCAGCTTGCGCTGCGTCCAGACCTTAA-3′
G1*/G30-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCTGTGTCGCCAGCTTGTGCGACGTCCAGACCTTAA-3′
G1*/R26-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCTGTGTCTTCAGCTTGCGCTGCGTCCAGACCTTAA-3′
G9/G16*-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCCGTACCGTCAGCTTGCGCTACGCCCAGACCTTAA-3′
G9/R26-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCTGTACCTTCAGCTTGCGCTACGCCCAGACCTTAA-3′
G16*/G30-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCCGTACCGCCAGCTTGTGCGACGCCCAGACCTTAA-3′
G16*/R26-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCCGTACCTTCAGCTTGCGCTGCGC-3′
R26/G30-anti	5′-cggGGATCCCTGCTTTGAATGCTGGCCTGTACCTCCAGCTTGTGCGACGCCCAGACCTTAA-3′

doi:10.1111/j.1440-1746.2007.05076.x

HEPATOLOGY

Inhibition of hepatitis C virus infection and expression *in vitro* and *in vivo* by recombinant adenovirus expressing short hairpin RNA

Naoya Sakamoto,*,† Yoko Tanabe,* Takanori Yokota,‡ Kenichi Satoh,§ Yuko Sekine-Osajima,* Mina Nakagawa,*,† Yasuhiro Itsui,* Megumi Tasaka,* Yuki Sakurai,* Chen Cheng-Hsin,* Masahiko Yano,¶ Shogo Ohkoshi,¶ Yutaka Aoyagi,¶ Shinya Maekawa,†† Nobuyuki Enomoto,†† Michinori Kohara§ and Mamoru Watanabe*

Departments of *Gastroenterology and Hepatology, †Hepatitis Control, and ‡Neurology and Neurological Science, Tokyo Medical and Dental University, §Department of Microbiology and Cell Biology, The Tokyo Metropolitan Institute of Medical Science, Tokyo, ¶Gastroenterology and Hepatology Division, Graduate School of Medical and Dental Sciences, Niigata University, Niigata, and ††First Department of Medicine, Yamanashi University, Yamanashi, Japan

Key words
adenovirus vector, hepatitis C virus, RNA interference.

Accepted for publication 12 April 2007.

Correspondence
Dr Naoya Sakamoto, Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan.
Email: nsakamoto.gast@tmd.ac.jp

NS and YT have contributed equally to this paper.

Abstract

Background and Aim: We have reported previously that synthetic small interfering RNA (siRNA) and DNA-based siRNA expression vectors efficiently and specifically suppress hepatitis C virus (HCV) replication *in vitro*. In this study, we investigated the effects of the siRNA targeting HCV-RNA *in vivo*.

Methods: We constructed recombinant retrovirus and adenovirus expressing short hairpin RNA (shRNA), and transfected into replicon-expressing cells *in vitro* and transgenic mice *in vivo*.

Results: Retroviral transduction of Huh7 cells to express shRNA and subsequent transfection of an HCV replicon into the cells showed that the cells had acquired resistance to HCV replication. Infection of cells expressing the HCV replicon with an adenovirus expressing shRNA resulted in efficient vector delivery and expression of shRNA, leading to suppression of the replicon in the cells by $\sim 10^{-3}$. Intravenous delivery of the adenovirus expressing shRNA into transgenic mice that can be induced to express HCV structural proteins by the Cre/*lox*P switching system resulted in specific suppression of virus protein synthesis in the liver.

Conclusion: Taken together, our results support the feasibility of utilizing gene targeting therapy based on siRNA and/or shRNA expression to counteract HCV replication, which might prove valuable in the treatment of hepatitis C.

Introduction

Hepatitis C virus (HCV), which affects 170 million people worldwide, is one of the most important pathogens causing liver-related morbidity and mortality.[1] The difficulty in eradicating HCV is attributable to limited treatment options against the virus and their unsatisfactory efficacies. Even with the most effective regimen with pegylated interferon (IFN) and ribavirin in combination, the efficacies are limited to less than half of the patients treated.[2] Given this situation, the development of safe and effective anti-HCV therapies is one of our high-priority goals.

RNA interference (RNAi) is a process of sequence-specific, post-transcriptional gene silencing that is initiated by double-stranded RNA.[3,4] Because of its potency and specificity, RNAi rapidly has become a powerful tool for basic research to analyze gene functions and for potential therapeutic applications. Recently, successful suppression of various human pathogens by RNAi have been reported, including human immunodeficiency viruses,[5,6] poliovirus,[7] influenza virus,[8] severe acute respiratory syndrome (SARS) virus[9] and hepatitis B virus (HBV).[10–13]

We and other researchers have reported that appropriately designed small interfering RNA (siRNA) targeting HCV genomic RNA can efficiently and specifically suppress HCV replication *in vitro*.[14–19] We have tested siRNA designed to target the well-conserved 5′-untranslated region (5′-UTR) of HCV-RNA, and identified the most effective target, just upstream of the translation initiation codon. Furthermore, transfection of DNA-based vectors expressing siRNA was as effective as that of synthetic siRNA in suppressing HCV replication.[14]

In this study, we explored the further possibility that efficient delivery and expression of siRNA may be effective in suppression and elimination of HCV replication and that delivery of such