Journal of Peptide Science
J. Pept. Sci. 2008; **14**: 924–935
Published online 20 March 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psc.1027

JOURNAL OF PEPTIDE SCIENCE

Amino acid requirement for the high affinity binding of a selected arginine-rich peptide with the HIV Rev-response element RNA

MAKI SUGAYA,[a] NORIKAZU NISHINO,[b] AKIRA KATOH[a] and KAZUO HARADA[c]*

[a] Department of Materials and Life Science, Faculty of Science and Technology, Seikei University, Musashino, Tokyo 180-8633, Japan
[b] Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Wakamatsu, Kitakyushu 808-0196, Japan
[c] Department of Life Sciences, Tokyo Gakugei University, Koganei, Tokyo 184-8501, Japan

Received 17 December 2007; Revised 26 January 2008; Accepted 1 February 2008

Abstract: The arginine-rich motif is a class of short arginine-rich peptides that bind to specific RNA structures that has been found to be a versatile framework for the design and selection of RNA-binding peptides. We previously identified novel peptides that bind to the Rev-response element (RRE) RNA of the HIV from an arginine-rich polypeptide library (ARPL) consisting of a polyarginine (15 mer) randomized at the *N*-terminal 10 positions. The selected peptides bound more strongly to the RRE than the natural binding partner, Rev, and contained glutamine residues that were assumed to be important for recognition of the G–A base pair. In addition, the peptides were predicted to bind to the RRE in an α-helical conformation. In this study, in order to understand the mechanism of the interaction between the RRE and the putative α-helical glutamine-containing peptides, the amino acid requirements for high affinity binding were analyzed by a combinatorial approach using a bacterial system for detecting RNA–peptide interactions. A consensus peptide, the DLA peptide, was elucidated, which consists of a single glutamine residue within a polyarginine context with the glutamine residue flanked at specific positions by three nonarginine residues, two of which appear to be important for α-helix stabilization. In addition, the DLA peptide was found to bind extremely tightly to the RRE with an affinity 50-fold higher than that of the Rev peptide as determined by a gel shift assay. A working model for the interaction of the DLA peptide to the RRE is proposed, which should aid in the development of peptide-based drugs that inhibit HIV replication, as well as in our understanding of polypeptide–RNA interactions. Copyright © 2008 European Peptide Society and John Wiley & Sons, Ltd.

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

Keywords: HIV RRE RNA; arginine-rich peptide; combinatorial analysis; antitermination system; anti-HIV drugs

INTRODUCTION

RNA–protein interactions play important roles in gene regulation and in the assembly of functional RNA–protein complexes such as the ribosome. Complexes between arginine-rich peptides and their RNA sites have been useful model systems in increasing our understanding of RNA–protein interactions [1–4]. Arginine-rich peptides have been shown to bind to their respective RNAs in a variety of conformations such as an α-helix in the case of HIV Rev and phage N peptides [5–8], a β-hairpin structure in the case of the bovine immunodeficiency virus (BIV) Tat peptide [9,10], and an extended conformation in the case of the HIV Tat peptide [11,12]. The RNA sites to which these peptides bind have been shown to be located at bulged and looped regions within relatively structured RNAs [7,8,13–15], where the narrow major groove of the RNA double helix is widened [16]. Adaptive binding involving unstructured to structured transitions of peptide and/or RNA upon complex formation has also been found to be a general feature [1,17].

The Rev peptide, a short peptide corresponding to the RNA-binding domain of the HIV Rev protein, has been shown to bind to an RNA structure within the HIV genome referred to as the Rev-response element (RRE) (Figure 1). Mutagenesis and chemical modification experiment showed that the Rev peptide binds to an internal loop region of the RRE consisting of G46 to C51 and U66 to C745 of the RRE (Figure 1(A)) [14,18,19]. *In vitro* selection experiments suggested the presence of a noncanonical G48–G71 base pair in the internal loop that can be replaced by an isostructural A48–A71 base pair (Figure 1(A)) [20,21]. The NMR structure of the Rev–RRE complex later confirmed the formation of the G48–G71 base pair as well as a G47–A73 base pair that results in the widening of an otherwise narrow major groove, allowing the binding of the α-helical Rev peptide deep into the RNA major groove [6]. Four specific amino acid side chain–nucleotide base interactions, an asparagine amide group (N40) binding to the G47–A73, the arginine guanidinium groups of R35, R39, and R44 binding

* 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

to G67, G70, and U45, respectively, along with several arginine guanidinium–phosphate interactions, and a number of hydrophobic interactions were observed [6].

A number of novel RRE-binding peptides have been identified from random peptide libraries, showing that multiple strategies exist for the recognition of the same RNA site [22–25]. In one such selection experiment using a bacterial reporter assay for detecting RNA–polypeptide interactions, RRE-binding peptides were selected from a randomized library consisting of three amino acids (R, S, and G), then further evolved by mutagenesis and reselection, resulting in the identification RSG-1.2 peptide, which bound to the RRE several times more strongly and specifically than the Rev peptide (Figure 1(B)) [22,23]. Strikingly, while the helical Rev peptide bound 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 [6], 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 [26,27]. The details of the two interaction have been found to be considerably different, as determined by selection of RRE-variants that bind to RSG-1.2, and comparison of the resulting sequences [28].

In another selection of RRE-binders from a Rev-based library, where the six amino acids important for RRE-binding (Figure 1(B), amino acids in bold) were completely randomized, peptides with a conserved glutamine residue at a position corresponding to the critical asparagine of Rev were identified using a mammalian screening system [24]. The RRE-binding affinity of the glutamine-containing peptides correlated with the α-helical propensity of the peptides, strongly suggesting that this class of peptides binds to the RRE in a manner similar to the Rev peptide. It was therefore concluded that the critical glutamine binds to the G47–A73 base pair of RRE in a manner analogous to the asparagine residue of the Rev peptide [24]. Strikingly, substitution of the glutamine residue to an asparagine led to loss of binding, as did substitution of the asparagine residue to a glutamine residue in the Rev context.

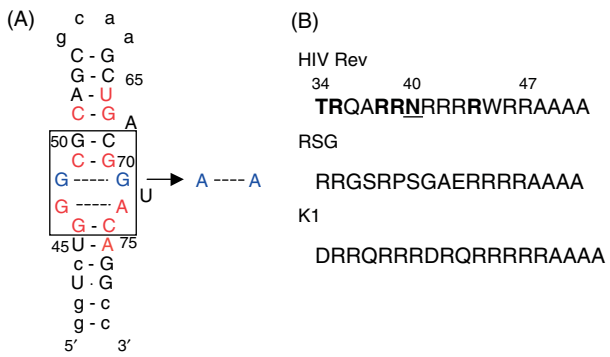

Figure 1 HIV RRE and selected RRE-binding peptides. (A) The secondary structure of the RRE with nucleotides important for Rev-binding shown in red, and those important for RSG-binding boxed. The G–G base pair that may covary to an isostructural A–A base pair, and which is important for K1 binding is shown in blue. (B) Amino acid sequences of Rev peptide and selected peptides, RSG-1.2 and K1. Amino acids within the Rev peptide important for RRE binding are indicated in bold, and the critical aspargine residue that contacts the G–A base pair is underlined.

In an attempt to further identify novel RRE-binders, selection from a complex arginine-rich peptide library, arginine-rich peptide library 1 (ARPL1), which consisted of a 15 mer polyarginine mutagenized at the *N*-terminal 10 positions using 12 nonhydrophobic amino acids at a rate of 50% was carried out using the bacterial screening system mentioned above [25]. As a result, peptides with activities considerably higher than the natural binding partner Rev, and which contained a conserved glutamine residue at a position similar to the glutamine-containing peptides in the above-mentioned selection were obtained [24].

While these results suggested that the glutamine-containing peptides were binding to the RRE in a manner similar to the Rev peptide, the nucleotide base requirements for RRE binding to the Rev peptide and a glutamine-containing peptide, K1, were found to be somewhat different [29]. In particular, the replacement of the characteristic G–G base pair to the isostructural A–A base pair, which is tolerated in the case of the Rev and RSG-1.2 peptide, lead to a loss of binding in the case of the K1 peptide, suggesting that the G–G base pair was being recognized by the K1 peptide.

In this study, in order to understand the mechanism for the high affinity binding of the glutamine-containing arginine-rich peptides with the RRE, combinatorial analysis of amino acids important for binding was carried out. First, in order to understand the sequence variability in all 15 residues of the peptide and possibly identify even stronger binders, two new ARPLs ARPL2 and ARPL3, where all 15 residues were randomized were constructed, and RRE-binding peptides were identified using the bacterial screening system (Figure 2). Comparison of RRE-binding peptides from the two libraries ARPL2 and ARPL3 resulted in the determination of a consensus RRE-binding sequence where the critical glutamine is flanked at a number of specific positions by nonarginine residues within a polyarginine framework. In order to determine the role of the nonarginine positions, a new library, Q3L, was designed to exhaustively screen for optimal amino acids in these positions, resulting in the identification of a consensus RRE-binding peptide, the DLA peptide. The importance of the remaining arginine residues was also determined by alanine- and lysine-scanning experiments. On the basis of these results, a model for the binding of the DLA peptide to the RRE is proposed, and the basis for the high affinity RRE-binding by the DLA peptide is discussed.

DOI: 10.1002/psc

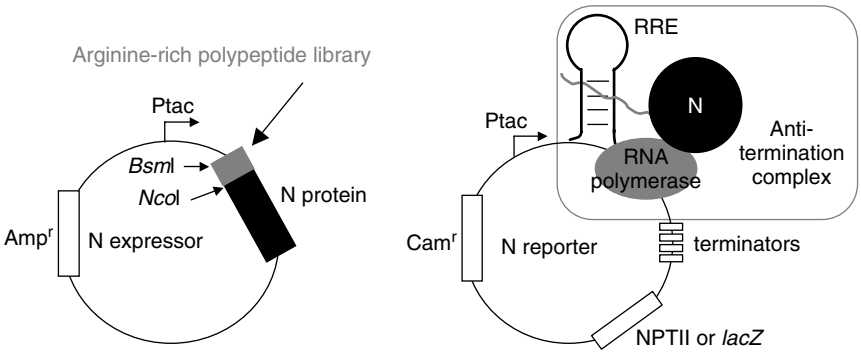

Figure 2 The two-plasmid system for detecting RNA–polypeptide interactions by λ N-mediated transcriptional antitermination. The arginine-rich polypeptide libraries ARPL2 and ARPL3 were fused to the N protein and expressed under the control of a tac promoter from a pBR322-based plasmid (N expressor). The pACYC-based reporter plasmid (N reporter) contains the *LacZ* gene or *NPT II* gene, also under the control of the tac promoter, so that binding of library peptides to the RRE results in reporter gene expression.

RESULTS

Selection of RRE-binding Peptides from the ARPL2 and ARPL3 Libraries

The glutamine-containing peptides identified in the previous study were identified from the ARPL1 library, which was randomized at the first 10 positions within a 15 mer polyarginine framework [25]. In order to determine those positions within the 15 residue peptide that are important for RRE-binding, two new libraries, ARPL2 and ARPL3, that were mutagenized at all 15 arginine residues at the codon level using 12 nonhydrophobic amino acids were constructed [30] and RRE-binding peptides were identified using a bacterial two-plasmid system (Figure 2) [22,25]. The doped polyarginine library was flanked at the *N*-terminus by two randomized amino acids: at the *C*-terminus by alanine residues, which stabilize α-helices in the case of ARPL2, and glycine residues, which destabilize α-helices in the case of ARPL3 (Figure 3).

Both selections were carried out using a four-step procedure to enrich RRE-binding sequences while eliminating reporter-related false positives and nonspecific positives [25]. In the first round, library plasmids encoding the ARPL2 and ARPL3 libraries were transformed into RRE-reporter cells carrying the kanamycin resistance reporter to yield 3.7×10^7 and 1.1×10^8 total transformants on plates containing 5 μg/ml of kanamycin, of which a total of 1.1×10^5 (0.29%) and 3.1×10^5 colonies (0.3%) survived, respectively. In the second round, pooled library plasmid DNA from each of the two first round survivors was then transformed into RRE-reporter cells, and the cells were spread on plates containing 10 μg/ml kanamycin to further enrich for tight binders and to reduce the background from reporter-related

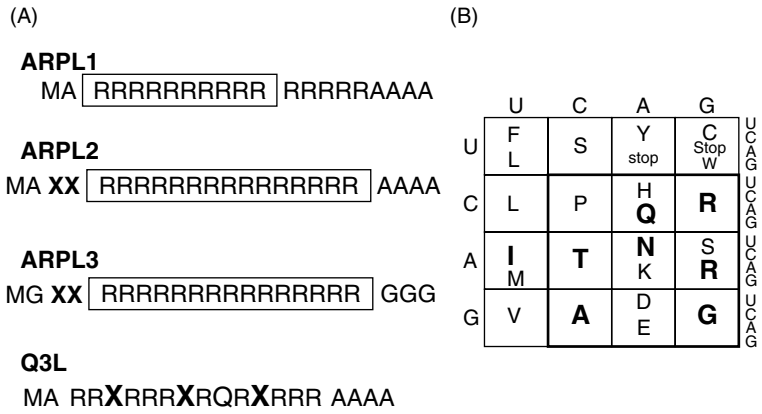

Figure 3 (A) Arginine-rich peptide libraries (ARPLs) used in this study. The doped regions of ARPLs are boxed. ARPL1 consisted of a polyarginine (15 mer) codon-mutagenized at 10 positions using 12 nonhydrophobic amino acids at a rate of 50% [25,29]. ARPL2 consisted of a 15 mer polyarginine codon-mutagenized at all 15 positions using 12 nonhydrophobic amino acids at a rate of 50%, and an additional 2 randomized amino acids at the *N*-terminal, indicated by X. ARPL3 consisted of the same doped region as ARPL2, but was flanked by glycines instead of alanines. Q3L consisted of a polyarginine (14 mer) mutagenized at 3 positions using 20 amino acids and a glutamine at position 9. (B) The genetic code viewed from the perspective of arginine rich peptides. The 12 amino acids used in the libraries are encoded by 18 codons, as indicated by the boldfaced box in the codon chart.

DOI: 10.1002/psc

false positives. Of the $\sim 3.6 \times 10^7$ total transformants for the ARPL2 selection and the $\sim 5.6 \times 10^7$ total transformants for the ARPL3 selection, 1.1×10^6 (3.1%) and 5.1×10^5 (0.9%) colonies survived, respectively. In the third round, pooled plasmid DNA from each of the second round survivors was transformed into pAC-RRE reporter cells carrying the *lacZ* reporter and the cells were plated onto X-gal plates. In the case of the ARPL2 selection, 99% of colonies that survived two rounds of kanamycin selection were blue, while in the case of the ARPL3 selection only 51% of colonies that survived two rounds of kanamycin selection were blue. In the fourth round, plasmid DNA from 28 of the darkest blue colonies in each selection were individually tested against the RRE reporter and the BIV TAR reporter. All clones showed β-galactosidase expression in with the RRE reporter and not with the BIV TAR reporter. In the ARPL2 selection, sequencing of the 28 clones resulted in the identification of 11 unique clones encoding peptides that contained a glutamine residue at either the 9th or 10th position of the polyarginine stretch, similar to RRE-binders from the ARPL1 [25], were identified (Table 1). In the case of the ARPL3 selection, six unique clones encoding peptides that contained a glutamine residue at either the 9th or 10th position of the polyarginine stretch, similar to RRE-binders from the ARPL1 and ARPL2 libraries [25], were identified (Table 1).

Analysis of RRE-binding Peptides Identified from ARPL1~3 Libraries, and Reselection from the Q3L Library

Upon alignment of RRE-binding peptides obtained from ARPL2 and ARPL3 selections at the glutamine residue, while the majority of positions appeared to require arginine residues for binding, three positions −6, −2, and +2 relative to glutamine contained aspartic acid, glutamic acid, and alanine residues at high frequencies (Table 1). Since the selected peptides were predicted to bind to the RRE in an α-helical conformation based on the similarity to the glutamine-containing peptides obtained in a previous selection [24], axial views of the putative α-helical peptides are shown in Figure 4. As

Table 1 RRE-binding peptides selected from ARPL2 and ARPL3

	Clone No.			Peptide sequences[a]		Antitermination activity (X-gal[b])	
						HIV RRE	(BIV TAR)
ARPL2							
Class I[c]	2–17	MA	SA	RRRRRRRRRQRARRR	AAAA	7.5	0
	2–28	MA	SD	PRRQRRRRRQRARRR	AAAA	6	0
Class II[c]	2–07	MA	NP	ARRARRRCRQRERRR	AAAA	7	0
	2–11	MA	SL	RRNRRRSRQRARRRE	AAAA	6	0
	2–06	MA	IL	ARRQRRRGRQRERRR	AAAA	6	0
Class III[c]	2–42	MA	ND	RRRRRRTRQRDRRRN	AAAA	6.5	0
	2–04	MA	QR	ERRERRRERQRNRRK	AAAA	6.5	0
	2–09	MA	YA	DRRKRRRERQRRRRA	AAAA	5	0
	2–26	MA	LL	PRRERRRRRQRERRR	AAAA	6	0
	2–39	MA	SR	DRTRRRRTRQRSRRR	AAAA	6.5	0
	2–08	MA	VA	TRRERRRRRQRTRRR	AAAA	5	0
ARPL3							
Class I[c]	3–07	MG	AN	ARRRRRRARQRARRR	GGG	6	0
	3–13	MG	DT	RRRRRHARQRERRRR	GGG	5	0
Class II[c]	3–14	MG	LS	PRRDKRRRRQRARRR	GGG	5	0
Class III[c]	3–01	MG	PH	RRQRRRDRQRRRRRA	AAA	6	0
	3–04	MG	LE	RRRRRRRDRQRKRRS	GGG	5	0
	3–12	MG	LL	RRDRRRQRQRQRRRR	GGG	6	0
Rev				TRQARRNRRRRWRR		3	0
RSG-1.2				RDRRRRGSRPSGAERRRR		4	0
K1		MA		DRRQRRRDRQRRRRR	AAAA	6	0
BIV Tat				MSGPRPRGTRGKGRRIRR		0	3

[a] The conserved glutamine residue is shown in bold, and the nonarginine residues are underlined.
[b] The number of plusses indicates blue color scored by the colony color (X-gal) assays that were performed as described [25].
[c] Selected peptides were classified into three groups based on the types of amino acids observed in the nonarginine positions. Class I peptides contained mostly hydrophobic amino acids. Class II peptides contained both hydrophobic amino acids and acidic amino acids. Class III peptides contained mostly acidic amino acids.

DOI: 10.1002/psc

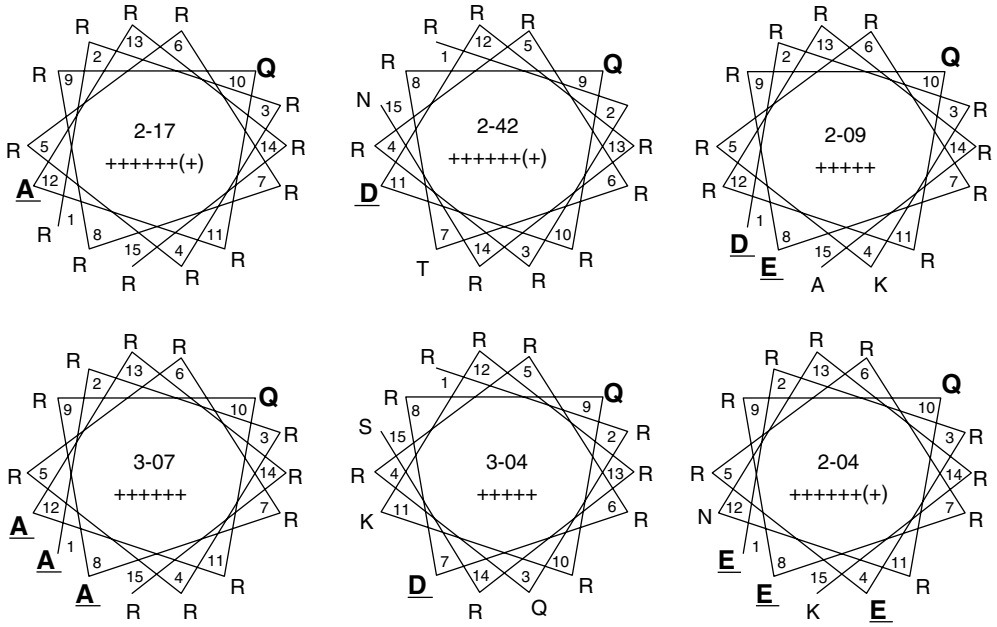

Figure 4 Axial view of the putative α-helical conformations of representative RRE-binding peptides. The nonarginine residues (alanines and acidic amino acids) are underlined, and the critical glutamine (Q) residue is shown in bold. The number of plusses indicates blue color scored by the colony color (X-gal) assay.

a result, the three nonarginine residues are positioned opposite to the conserved glutamine residue. In order to determine the role of amino acids at the three nonarginine positions and to identify the peptides that bind more strongly to the RRE, a new library, Q3L, was designed. The Q3L library consisted of 14 mer polyarginines with the glutamine residue important for RRE-binding at the 9th position, and positions −6, −2, and +2 relative to the glutamine completely randomized with 32 codons encoding the 20 amino acids, and resulting in a total of $32^3 = 32\,768$ codon sequences (Figure 3(A)).

In the case of the Q3L library, RRE-binding peptides were screened using RRE-LacZ reporter cells. In the first round, $\sim 6.6 \times 10^5$ colonies were screened, and 20% of the colonies were blue. Of these blue colonies, 1248 colonies with varying blue intensity were pooled, and pBR library plasmid DNA was isolated. In the second round, this library plasmid was retransformed into RRE-reporter cells, and 72% of colonies were blue. Plasmids were isolated from 70 clones of varying blue intensity, and scored as $2+ \sim 6+$, relative to the Rev peptide (3+) and RSG-1.2 peptide (4+). Of these 70 clones, 59 were found to be unique sequences. The sequence of the clones exhibiting blue intensities of 6+, 5+, 4+, 3+, and 2+ are shown in Supplementary Table 1, and the relationship between antitermination activity and amino acid identity at the three positions is shown in Figure 5. At position −6 relative to glutamine, clones with high antitermination activities had a preference for acidic amino acids, in particular, aspartic acid (D) (Figure 5(A), Supplementary Table 1). At position −2 and +2, clones with high antitermination activities both had a preference for aliphatic amino acids, in particular, leucine (L) and alanine (A), respectively (Figure 5(B) and (C), Supplementary Table 1). This tendency diminished at each position as the intensity of blue decreased, and in the case of 3+ and 2+ clones, no clear preference could be observed (Figure 5(A–C)). On the basis of these results, the DLA peptide, which contains an asparagine (D) at position −6, a leucine (L) at position −2, and an alanine (A) at position +2, was chosen as the optimal sequence for RRE binding.

In order to confirm the importance of these positions for RRE-binding, three mutant DLA peptides that were expected to show decreased activity, with a D to R substitution at position −6 (RLA peptide), an L to N substitution at position −2 (DNA peptide), and an A to E substitution at position +2 (DLE peptide), as well as a peptide where all three positions were substituted by arginines (R8QR5) were tested *in vivo* using the antitermination assay. As a result, all three peptides with single amino acid substitutions showed blue color intensities of $3.5+ \sim 4+$, and substitution of all three positions led to a complete loss of activity (Table 2). On the other hand, a peptide with a substitution of the conserved glutamine to an asparagine residue (Table 2, DLA Q9N) did not show activity, as in the case of a previously reported RRE-binding peptide containing glutamine [24]. Gel shift analysis of the affinity of RRE-binding by the DLA peptide and the three mutant peptides RLA, DNA, and DLE showed that the activities reported *in vivo* reflect the strength of RNA-peptide

DOI: 10.1002/psc

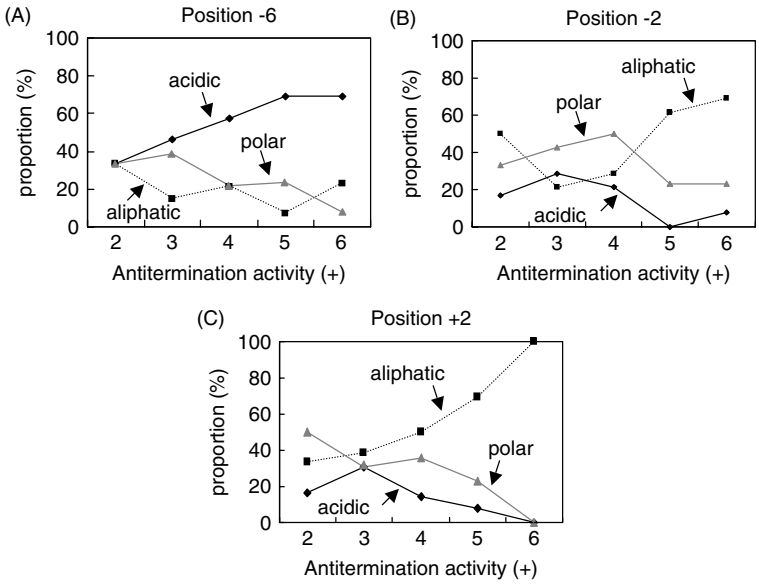

Figure 5 The relationship between the antitermination activity and the distribution of amino acids at each randomized position of the glutamine-containing peptides. (A), (B), and (C) indicate the distribution of amino acids at positions −6, −2, and +2 relative to the glutamine residue, respectively. 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

DOI: 10.1002/psc

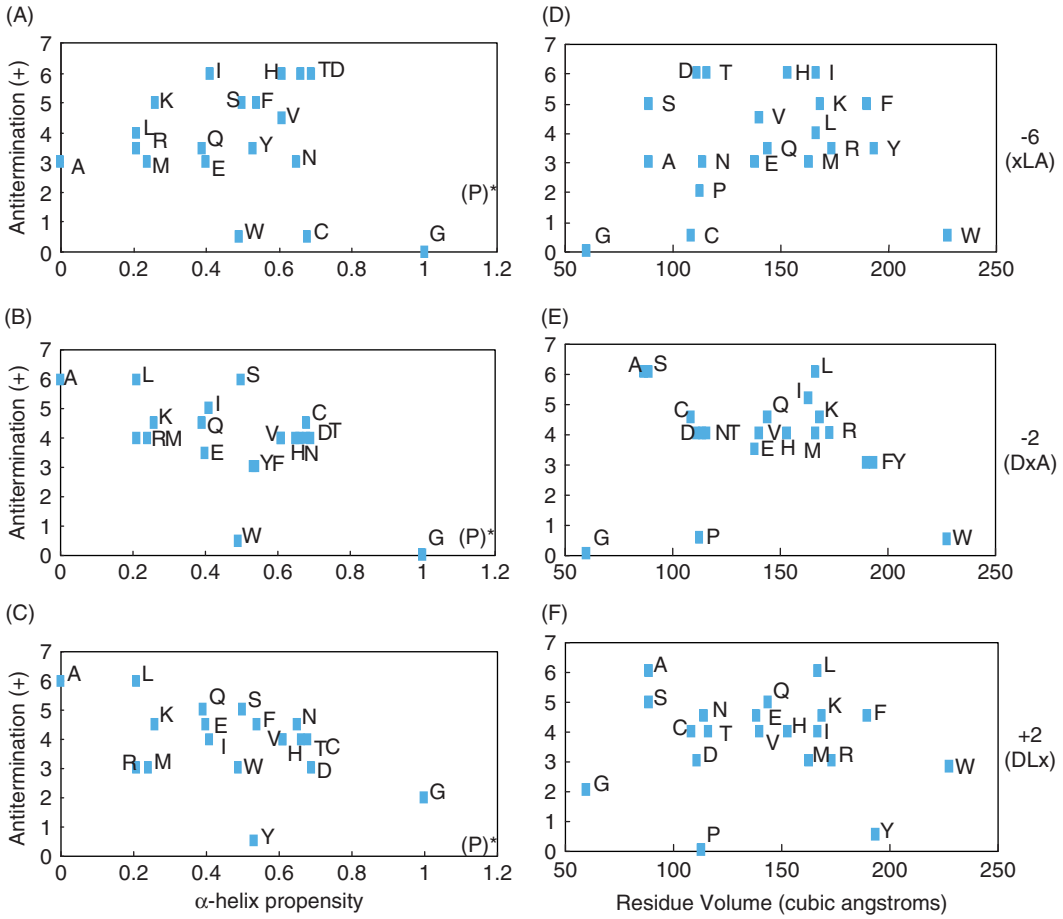

Figure 6 The relationship between antitermination activity and α-helix propensity or residue volume of amino acids at positions -6, -2, and $+2$ of the DLA peptide, relative to the glutamine residue. The α-helix propensity of proline has a value of 3.16 and lies outside of the region shown in the graph 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+ $\sim$ 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 $\sim$ 3 (ARPL1$\sim$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

DOI: 10.1002/psc

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.

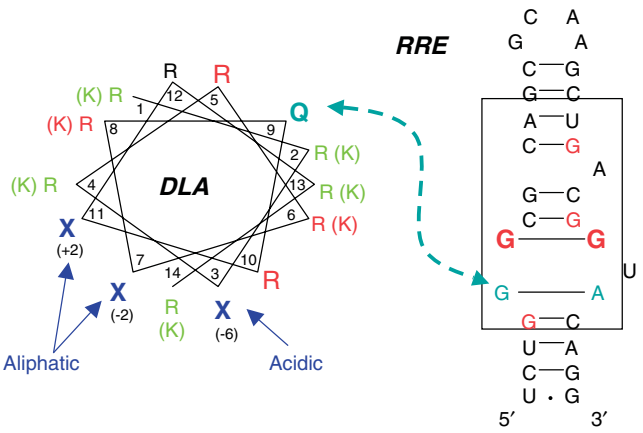

Figure 7 A model for the interaction of the DLA peptide and the RRE. The DLA peptide is depicted as an α-helical wheel with the conserved glutamine shown in light blue, the arginines important for RRE-binding shown in red, the arginines that can be replaced by lysine shown in green, and the three nonarginine 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.

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.

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

DOI: 10.1002/psc

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)}, wh ere 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.

DOI: 10.1002/psc

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, *Eco*RV 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 *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 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 $\sim 6.6 \times 10^5$ 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

DOI: 10.1002/psc

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.