Arginine-rich motifs present multiple interfaces for specific binding by RNA

TRAVIS S. BAYER,[1] LAUREN N. BOOTH,[2] SCOTT M. KNUDSEN,[2] and ANDREW D. ELLINGTON[2]

[1]Division of Biology, California Institute of Technology, Pasadena, California 91125, USA
[2]Department of Chemistry and Biochemistry, Institute for Cell and Molecular Biology, University of Texas at Austin, Austin, Texas 78712, USA

ABSTRACT

A number of proteins containing arginine-rich motifs (ARMs) are known to bind RNA and are involved in regulating RNA processing in viruses and cells. Using automated selection methods we have generated a number of aptamers against ARM peptides from various natural proteins. Aptamers bind tightly to their cognate ARMs, with K_d values in the nanomolar range, and frequently show no propensity to bind to other ARMs or even to single amino acid variants of the cognate ARM. However, at least some anti-ARM aptamers can cross-recognize a limited set of other ARMs, just as natural RNA-binding sites have been shown to exhibit so-called "chameleonism." We expand upon the number of examples of cross-recognition and, using mutational and circular dichroism (CD) analyses, demonstrate that there are multiple mechanisms by which RNA ligands can cross-recognize ARMs. These studies support a model in which individual arginine residues govern binding to an RNA ligand, and the inherent flexibility of the peptide backbone may make it possible for "semi-specific" recognition of a discrete set of RNAs by a discrete set of ARM peptides and proteins.

Keywords: aptamer; arginine-rich motif; ARM; in vitro selection; peptide; specificity

INTRODUCTION

RNA-binding proteins with high affinity and specificity for their targets underlie many cellular and viral regulatory processes (Weiss and Narayana 1998). One common RNA binding domain, the arginine-rich motif (ARM), is characterized by a preponderance of arginine residues but little other sequence similarity. ARMs appear to have arisen independently throughout phylogeny. For example, an arginine-rich domain is implicated in RNA binding by human ribosomal protein L7 (Hemmerich et al. 1997). Similarly, the ARM of the lambdoid bacteriophage N protein binds hairpin loops in nascent RNAs to regulate antitermination (Su et al. 1997), while the ARM of the HIV-1 Rev protein binds to a 5′ region of the HIV genome, the Rev-responsive element (RRE), and regulates mRNA transport and splicing (Frankel and Young 1998).

In many instances, ARMs seem to function as independent recognition domains, separate from the protein in which they are found, and have proven to be excellent models for studying protein:RNA interactions (Smith et al. 2000b). For example, chemically synthesized peptides corresponding to the Rev ARM (residues 34–50 of HIV-1 Rev) can adopt an α-helical conformation and bind to the Rev-binding element (RBE), a short stem-internal loop-stem structure within the RRE (Battiste et al. 1996). Other synthetic ARM peptides, such as residues 65–81 from BIV Tat, can also specifically interact with their cognate RNA binding elements (Puglisi et al. 1995).

It would not be surprising if protein domains with numerous arginine residues bound RNA nonspecifically. However, in those instances where interactions between ARMs and RNA molecules have been examined, they have been shown to be specific. For example, the ARM peptide from phage λ binds its own boxB RNA 16-fold better than the related phage P22 boxB RNA (Tan and Frankel 1995), and the RBE and selected anti-Rev aptamers can discriminate by 10-fold or more between Rev ARM variants that have single amino acid substitutions (Tan et al. 1993; Xu and Ellington 1996).

However, other experiments have shown that a given ARM can exhibit "chameleon-like" behavior, binding to multiple, different RNA targets. The Jembrana disease virus (JDV) Tat protein can bind both the BIV and HIV-1 TAR RNAs (as well a hybrid TAR containing elements of both the HIV-1 TAR and BIV TAR), despite the fact that

the BIV and HIV Tat ARMs assume a β-hairpin structure and an extended conformation, respectively. Indeed, the JDV Tat ARM was shown to bind its own TAR and BIV TAR via a β-hairpin structure, while it bound HIV TAR in an extended conformation (Smith et al. 2000a). Structural characterization of complexes with natural or selected ARM-binding RNAs and ARM peptides has also revealed the flexibility of ARM structures. Solution structures are known for the HTLV-1 Rex ARM complexed with an anti-Rex aptamer (Jiang et al. 1999) and for the Rev ARM complexed with either the RBE (Battiste et al. 1996) or anti-Rev aptamers (Ye et al. 1996, 1999). The Rev ARM binds the RBE as an α-helix that fits into the widened groove of a purine-rich internal loop region. However, the same ARM peptide binds other anti-Rev aptamers in an extended conformation.

Taken together, these structural studies seem to indicate that RNA architectures can dictate the folds of ARM peptides. In attempting to describe the general features of ARM:RNA interactions, binding and structural results at first glance seem to contradict one another. On the one hand, ARMs can bind specifically to their natural RNA targets; on the other, they can bind to multiple RNA targets. The obvious resolution of this conundrum is that ARMs can bind several targets specifically, but do not in general bind with high affinity to most RNA targets. The question thus becomes whether this is a common feature of ARM:RNA interactions: Do arginine contacts with RNA foster high-specificity interactions, or do these contacts intrinsically allow enough flexibility so that there is cross-recognition between targets?

To begin to answer these questions, in vivo and in vitro peptide selection methods have previously been used to identify what peptides can bind a given RNA. For example, Roberts and coworkers used mRNA displays to generate a diverse set of peptides that can bind lambda N boxB RNA (Barrick et al. 2001). Here, we selected aptamers against a variety of natural ARMs, characterized the cross-recognition of these aptamers, examined the residues used in

mediating specific interactions, and characterized the structural changes that occur upon binding.

RESULTS AND DISCUSSION

Specificities of natural ARM:RNA interactions

We examined the affinity of a number of known, natural ARM:RNA interactions (sequences shown in Table 1) in order to provide a baseline for comparing the specificity and flexibility of ARM:RNA interactions generated through in vitro selection. RNA gel shifts were carried out as a function of ARM peptide concentration (initially from 4 to 1024 nM) in the presence of limiting RNA. Once the approximate affinity of an ARM peptide for its RNA ligand was determined, a more limited set of peptide concentrations was used to more accurately estimate the dissociation constant of the interaction (Table 2).

For the most part, the affinities revealed by the RNA gel-shift assays were consonant with results obtained in other studies. For example, Frankel and coworkers also used gel-shift assays and determined that the P22 N ARM:P22 boxB interaction had a K_d of 160 nM (Tan and Frankel 1995); here we estimate a K_d of 230 nM. Similarly, a K_d range of 200–800 nM was reported for the HIV Tat:TAR interaction by gel-shift assay (Tao and Frankel 1993; Smith et al. 1998); we estimate a K_d of 600 nM. Finally, dissociation constants of 1.3 and 8 nM were previously found for the BIV Tat:TAR interaction using mobility shift assays (Smith et al. 1998, 2000a), while we estimate a K_d of 60 nM. This apparent difference in affinity may be due to the differing Mg^{2+} concentrations used in the binding buffers in these different studies (5 mM here vs. 1 mM in the Smith et al. 1998 and Tao and Frankel 1993 studies). The higher magnesium concentration would more effectively compete with the basic peptide for electrostatic interactions with the RNA. Although our results are generally consistent with other gel-shift data, different methods and techniques have been known to yield different affinities. For example, when P22

TABLE 1. Sequences of arginine rich motif peptides and selected natural RNA ligands

Protein	ARM sequence	RNA sequence
HIV-1 Rev	TRQARRNRRRRWRERQR	GGUCUGGGCGCAGCGCAAGCUGACGGUACAGACC
Rev N7D	TRQARRDRRRRWRERQR	
Rev R11Q	TRQARRNRRRQWRERQR	
Rev R10Y	TRQARRNRRYRWRERQR	
P22 N	NAKTRRHERRRKLAIER	GGUGCGCUGACAAAGCGCGCC
λ N	MDAQTRRRERRAEKQAQWKAAN	GGGCCUGAAGAAGGGCCC
BMV Gag	KMTRAQRRAAARRNRWTAR	
CCMV Gag	KLTRAQRRAAARKNKRNTR	
BIV Tat	SGPRPRGTRGKGRRIRR	GGCUCGUGUAGCUCAUUAGCUCCGAGCC
HIV-1 Tat	SYGRKKRRQRRRPPQ	CCAGAUCUGAGCCUGGGAGCUCUCUGG
HTLV-1 Rex	MPKTRRRPRRSQRKRP	GCUCAGGUCGAGGUACGCAAGUACCUCCCUUGGAGC

TABLE 2. Specificity of natural ARM ligands

	Rev	RevN7D	RevR11Q	RevR10Y	P22N	λ N	BIV Tat	HIV-1 Tat	Rex
HIV-1 RBE	**160**	*240*	*250*	*>512*					*340*
P22 boxB					**230**				
λ boxB					*550*	**90**			
BIV TAR							**60**		
HIV-1 TAR								**600**	
HTLV XBE									**270**

Dissociation constants for ARM peptides with their natural RNA ligands. Wild-type interactions are shown in bold, and noncognate ligand binding is shown in italics. Blank fields indicate that no binding was observed at 1024 nM.

boxB RNA was fluorescently labeled, and the kinetics of interactions with the P22 N ARM complex were examined as a function of salt concentration, the K_d was found to be 200 pM (Austin et al. 2002). Although this dissociation constant is 800–1200-fold lower than those reported by gel-shift assays (above), it should be noted that the studies by Austin et al. (2002) were performed in the absence of magnesium, and the observed dissociation constants appeared to be exquisitely sensitive to salt concentration, with roughly a twofold difference in monovalent salt concentration yielding a 40-fold difference in apparent K_d. In this regard, our studies are especially useful as they are the first large comparison of ARM:RNA interactions carried out under a uniform set of binding conditions.

A survey of the data in Table 2 shows that many ARM:RNA interactions are specific, with a given RNA generally binding tightly only to its cognate ARM. However, some related ARMs show modest cross-recognition, again consistent with what was previously observed. For example, the P22N peptide bound the P22 boxB element RNA with a dissociation constant of 230 nM, and the λ boxB RNA with a dissociation constant of 550 nM. The P22N ARM was previously reported to bind both the P22 and λ boxB RNAs, with K_d values of 5 pM and 850 pM, respectively (Austin et al. 2002). In addition, the observed cross-recognition between HIV-1 Rev and HTLV Rex ARM peptides and binding elements is similar to what was reported previously (Felber et al. 1989; Solomin et al. 1990; Unge et al. 1991). In previous studies, it was shown that Rex could functionally substitute for Rev, but not vice versa. Likewise, we see here that the Rex peptide will bind HIV-1 RRE RNA (K_d of 340 nM), albeit with lower affinity than Rev (K_d of 160 nM) does, but that Rev peptide does not appear to bind Rex-binding element (XBE) RNA at the concentrations we surveyed.

In vitro selection of anti-ARM aptamers

In order to determine whether the ARM:RNA specificities that have been observed with natural RNA ligands were unique to natural biopolymers, or whether there were many possible ways in which ARMs and RNA mole-cules could interact specifically, we carried out in vitro selection experiments against some of the same ARM peptide targets that were known to bind natural ligands. The target set also included several ARMs with single amino acid substitutions, in order to examine the extent to which ARM sequence would influence RNA-binding specificity.

The use of in vitro selection experiments to examine the specificity of molecular recognition for an entire class of binding domains is novel, and may prove to be more useful than any number of individual binding experiments with natural ligands. It is generally difficult to assign a quantitative meaning to the term "specificity," because the very notion of specificity implies an exhaustive knowledge of the relative binding affinities of all molecular species for a given target. In a selection experiment, all nucleic acid species of a given length compete for binding to a target, and the most successful competitors are recovered. By comparing the sequences of binding species selected against similar targets (as in this study) it should be possible to observe the extent to which known target:ligand interactions overlap within the space of all possible target:ligand interactions.

Selection experiments were facilitated by the development of protocols for automated in vitro selection (Cox et al. 1998, 2002b; Cox and Ellington 2001). Peptide targets were synthesized with a C-terminal biotin and immobilized on streptavidin beads. From 12 to 18 cycles of automated selection and amplification were carried out on a robotic workstation starting from an RNA pool that contained 30 random sequence positions (N30). Any resultant RNA aptamers should therefore be roughly the size of most of the natural ligands that bind to ARMs. The RNA and DNA pools amplified well in each round and yielded products of the proper length. The affinities of the selected pools for ARM peptides typically increased 5–10-fold over the initial (round zero) pool by the conclusion of the selection. Although this does not seem like a large increase in affinity, this is in part because the ARMs have a high nonspecific affinity for the initial RNA pool. We previously observed that small aggregate increases in affinity typically yield individual aptamers with greatly improved affinities and specificities for their targets (Giver et al. 1993); this was again observed here.

Sequence analyses reveal diverse RNA ligands to similar targets

At the conclusion of the selection, aptamers from each of the pools were cloned and sequenced. In general, each selection yielded only a few major families of related aptamer sequences, as is typically the case in successful selection

experiments. Clones from major families typically represented from one-fourth to one-half of sequenced pools; the primary families for each ARM target are shown in Table 3. These families bear no resemblance to one another or to wild-type RNA ligands (Table 1). Again, this is expected, as in vitro selections frequently yield RNA molecules that differ from natural ligands. The predicted secondary structures of aptamers from the major families were also dissimilar to one another, to natural RNA ligands, and to previously selected aptamers (Baskerville et al. 1995; Ye et al. 1996). Despite the fact that aptamers frequently do not resemble natural ligands, it was nonetheless interesting to observe that even aptamers to extremely closely related peptide targets (the Rev ARM peptide variants, which differed by a single amino acid from the wild-type Rev ARM) showed no greater similarity to one another than did aptamers to completely dissimilar peptide targets. Overall, the aptamers represent an independent set of RNA ligands that have been optimized for a particular binding buffer, and that can therefore be uniquely used to directly compare the specificities of ARM:RNA interactions.

Affinities and specificities of selected ARM:aptamer interactions

As described above, gel-shift experiments have previously been used to quantitatively characterize interactions between ARMs and their target RNAs (Tao and Frankel 1993; Smith et al. 1998, 2000a). Given that we were able to obtain affinity and specificity results with natural ligands that were similar to those previously reported, we used the same techniques for the characterization of selected ARM:RNA interactions. Data for a typical set of gel-shift experiments are shown in Figure 1, and the affinities of the selected interactions are shown in Table 4.

The measured affinities of the selected ARM:RNA interactions often paralleled the values determined for natural interactions. For example, the Rev ARM bound anti-Rev aptamer 5 with a K_d of 190 nM, while the same peptide bound the RBE with a K_d of 160 nM. The affinity of the Rex peptide:XBE interaction was virtually equivalent to that of the Rex peptide:anti-Rex aptamer interaction, being measured at 270 nM and 260 nM, respectively. The dissociation constants for the P22 N:P22 boxB and P22 N:anti-P22 N aptamer interactions were also similar, being 230 nM and 300 nM, respectively. Even when the affinities differed, they did not differ greatly: BIV TAR bound BIV Tat peptide with a dissociation constant of 60 nM, while the anti-BIV Tat aptamer bound with a dissociation constant of 280 nM. In general, with the buffer conditions that we used ARMs appear to bind to their cognate nucleic acids with dissociation constants in the hundreds of nanomolar. It is possible to bind even more tightly, but it is likely that such reports are again highly dependent on the nature of the binding buffer and reaction conditions.

Furthermore, we wished to determine whether selections that only targeted affinity also yielded aptamers with a high specificity, or whether, as has been observed for a subset of the natural ARM:RNA interactions, cross-recognition would be observed. In order to quickly screen for cross-recognition between ARMs and aptamers, high concentrations of noncognate ARMs (1024 nM) were mixed with a representative member of each of the major aptamer families. In most instances, no shift was observed. In those instances where a shift was observed, the concentration dependence of the shift was examined, and an approximate K_d value for the complex was determined, as shown in Table 4.

In all cases, selected RNAs bound the ARM peptide they were selected against with higher affinity than any noncognate ARM peptide (Table 4). Very little cross-recognition was observed. This could be considered surprising, given the fact that the aptamers were selected only for their ability to bind to one target, not for the ability to discriminate between multiple different targets.

Dissecting the specificity of Rev ARM:aptamer interactions

Some of the most compelling examples of specificity occurred among natural or selected RNAs that bound the Rev

---

TABLE 3. Sequences of major families for anti-ARM selection

Target	Sequence
	GGGAAUGGAUCCACAUCUACGAAUUC-**N30**-UUCACUGCAGACUUGACGAAGCUU
HIV-1 Rev (5)	**GUUGGAAGAGAGGAAAAUUGACAGCGCGA**
HIV-1 Rev (11)	**GGAAGAUCAUCAUCAUGGGCAAAUCGGCGAC**
Rev N7D	**GACAAGUUGGUCCGCACAGUUGCGAGGUGU**
Rev R11Q	**CGCUUAUGGUCAUUGAGUAUCUCCUGCCGA**
P22 N	**GGAAAAAAGCAACGUCACACUUCGAGCAAGC**
BMV Gag	**GGCGGUGGGUUUGGAAAACGGUAACGGGCA**
CCMV Gag	**GUAAGCGUCAAUCAGUCACCAUCGGUUAGC**
BIV Tat	**GAGGCGGUGAACUUGGAAUCCCACAAGGGCG**
HTLV-1 Rex	**CGGAUCGUGCACGACGGUUGGUUUGCUUGG**

Sequences for the predominant aptamers selected against various peptides are shown. Pool-derived static regions are shown in regular font and sequences selected from the random region are indicated for each target in bold.
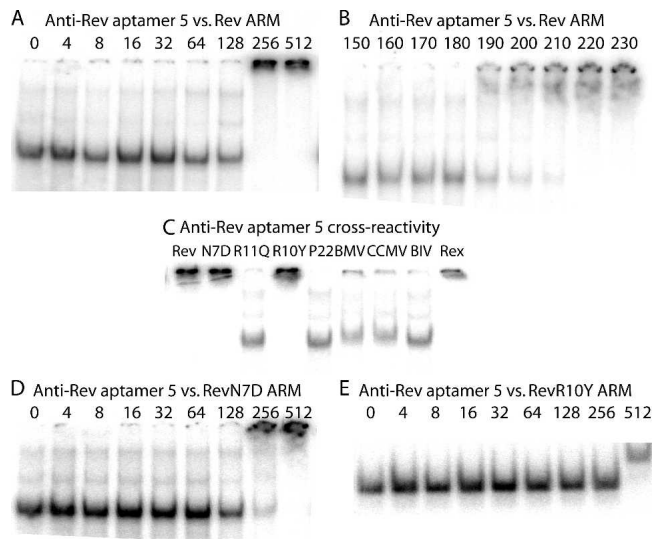
FIGURE 1. Nondenaturing gel shifts for anti-Rev aptamer 5. Limiting aptamer RNA was incubated with various concentrations (shown in nanomolar) of (*A,B*) Rev peptide (broad and narrow concentration ranges, respectively), (*D*) RevN7D, and (*E*) RevR10Y peptides. Cross-recognitions (*C*) were assessed with 1024 nM of each peptide.

ARM. The RBE, a previously selected anti-Rev aptamer, and a ribozyme activated by the Rev ARM show at least a 10-fold loss in binding or catalytic activity in the presence of peptides containing single amino acid substitutions. In the current work, we selected aptamers against both the Rev ARM and two single amino acid variants of the Rev ARM (RevN7D and RevR11Q). The anti-RevR11Q aptamer did not bind to any other ARMs, and other aptamers could discriminate against at least one noncognate peptide that differed by only a single amino acid.

In order to further determine to what extent the binding pockets formed by the selected Rev-binding RNAs were similar to or differed from one another, we carried out gel-shift analyses with a series of Rev ARM peptide variants that contained alanine residues at each of the positions in the peptide. In brief, each RNA ligand that was shown to

bind the Rev peptide (both aptamers and wild-type ligands) was assayed against the set of 17 "alanine-scanning" peptides. Two assays were performed per RNA ligand, one with a peptide concentration that was twice that of the original interaction K_d, and one with a peptide concentration that was half that of the original K_d. The "high"-concentration experiments revealed which alanine mutations abolished binding, while the "low"-concentration experiments served as controls for these improved signals (or sometimes revealed alanine mutations that actually increased binding affinities).

The results of alanine-scanning experiments are shown in Table 5. Aptamers appear to recognize their peptide ligands primarily via arginine residues, but different aptamers will recognize the same ARMs via different arginines. For example, the natural Rev-binding element recognizes arginine residues 2, 5, 6, and 11 of the Rev ARM, while the selected anti-Rev aptamer 5 recognizes arginine residues 2, 8, and 9. Another anti-Rev aptamer from the same selection, anti-Rev 11, recognizes arginine residues 8, 9, 13, and 17 and also appears to recognize glutamine 16. The majority aptamer selected against the variant RevN7D ARM recognizes yet another set of arginine residues (2, 9, and 10). It should be noted that residues other than the arginines are also important for binding, if only to present the arginines in an appropriate structural context. For example, the BIV Tat ARM contains arginine residues spaced at positions 4, 11, and 12, similar to the 2, 9, and 10 arginine spacing recognized by the anti-RevN7D aptamer, yet the anti-RevN7D aptamer was not found to bind the Tat peptide.

A number of alanine substitutions also appeared to increase the affinity of ARMs for their cognate RNAs. It is significant that the alanine substitutions that increased affinity tended to flank the arginine residues that were shown by substitution experiments to be critical for affinity. For example, in anti-Rev aptamer 5 the replacement of an asparagine residue on the N-terminal side of two critical arginine residues (residues 8 and 9) increased affinity compared to the wild-type peptide, as did changing the threonine on the N-

TABLE 4. Specificity of selected aptamers

	Rev	RevN7D	RevR11Q	RevR10Y	P22N	BMV Gag	CCMV Gag	BIV Tat	Rex
anti-Rev 5	**190**	*220*		*310*					*420*
anti-RevN7D	*250*	**120**							
anti-RevR11Q			**100**						
anti-P22N					**300**				
anti-BMV Gag						**20**	*260*		
anti-CCMV Gag						*480*	**280**		
anti-BIV Tat								**280**	
anti-Rex	*380*	*420*	*450*						**260**

Dissociation constants for ARM peptides binding to RNA aptamers. The selected interaction is in bold, and noncognate ligand binding is in italics. Blank fields indicate that no binding was observed at 1024 nM.

TABLE 5. Results from alanine scanning

RBE	T	**R**	Q	A	**R**	**R**	N	R	R	R	**R**	W	R	E	R	Q	R
anti-Rev 5	*T*	**R**	Q	A	R	R	*N*	**R**	**R**	R	R	W	R	E	R	Q	R
anti-Rev 11	*T*	R	Q	A	R	R	*N*	**R**	**R**	R	R	*W*	**R**	E	R	**Q**	**R**
anti-N7D	*T*	**R**	*Q*	A	R	R	*N*	*R*	**R**	**R**	*R*	*W*	R	E	R	Q	R
anti-Rex	*T*	R	Q	A	R	**R**	*N*	*R*	**R**	R	*R*	*W*	R	E	R	Q	R

Alanine scanning experiments with the Rev peptide. Residues that, upon substitution, decrease affinity are shown in bold; residues whose substitution increases affinity are underlined and italicized.

terminal side of arginine 2. Results from alanine scans from anti-RevN7D aptamer even show that while arginines 9 and 10 appear to be critical for binding, replacing adjacent arginines (8 or 11) with alanine improves binding. These results can potentially be explained by assuming that the alanine substitutions relieve steric hindrance brought about by the close approach of the critical arginine residues to their binding sites on the selected RNAs.

Cross-recognition between ARMs and aptamers

Overall, in vitro selection for affinity can also generate extremely specific nucleic acid binding species for ARM peptides. This was previously observed for anti-Rev ARM peptide aptamers, which did not bind HIV-1 Tat or HTLV Rex ARM peptides (Xu and Ellington 1996). The aptamers we selected also displayed the ability to differentiate between closely related ARMs. However, a number of notable examples of cross-recognition were observed with closely related peptides. Most of the cross-recognition interactions displayed in Table 4 occurred among the peptide variants of the HIV-1 Rev ARM, which differed from the wild-type by single amino acid substitution and from one another by two amino acid residues. These results are consistent with what is known about contacts between the ARMs and the aptamers. For example, N7A and R10A substitutions in the Rev ARM did not disrupt binding of the anti-Rev ARM aptamer, and this aptamer could also bind to N7D and R10Y Rev ARM variants.

Aside from the variant Rev ARMs, there were two major cases of cross-recognition displayed between selected aptamers and ARM peptides. First, the anti-BMV Gag and anti-CCMV Gag ARM aptamers each cross-recognize one another's peptide target. This is not surprising given the extensive sequence similarity between the two peptides, which have 15 of 19 amino acids in common. Second, an anti-Rev aptamer will bind to the Rex ARM, and the corresponding anti-Rex aptamer will bind to the Rev ARM and variants. Curiously, these results partially mirror observations with the natural binding elements, in that the Rev-responsive element has been shown to bind to the Rex protein (Unge et al. 1991). However, the anti-Rex ARM aptamers must form a binding pocket that is substantially different than the natural Rex-binding element (XBE), as they can bind to the Rev ARM while the XBE is not known to bind Rev.

Structural analysis of cross-recognized pairs

We wanted to determine whether the specificity or cross-recognition of aptamers with their ARM ligands stemmed from the sequences of the ARMs, or whether there was a structural component, as well. Circular dichroism (CD) is a technique which has previously been used to elucidate structural details about peptide complexes with RNA. The magnitude and location of changes in CD spectra can be used to identify structural changes in both the RNA and the peptide, with changes below 250 nm generally corresponding to changes in peptide structure (Berova et al. 2000), and changes above 250 nm corresponding to changes in RNA structure. For example, negative ellipticities at 208 and 222 nm correspond to α-helical structures, while positive ellipticities at ~265 nm are directly proportional to the degree of base stacking (Gilligan and Schwarz 1976).

The CD spectra of the peptides alone (not in complex with RNA) reveal none of the characteristic shapes of protein secondary structure, suggesting that the peptides are either unstructured or adopt a noncanonical structure. This result has been observed before for ARM peptides, most of which only assume a stable structure in complexes with their RNA ligands. The exception is the Rev ARM peptide, which can assume an α-helical structure in solution as long as it contains C- and N- terminal modifications (Tan et al. 1993); these modifications were not included on the Rev ARM used in the present study.

The CD spectra for the peptide:RNA complexes were dominated by the RNA signal. This is likely because of the large size difference between the two; the aptamers (~27 kD) were about 10 times as massive as the peptides (2–3 kD). This is a larger difference than has previously been the case for CD studies of ARM:ligand interactions. For example, when Frankel and his coworkers analyzed the interaction of the RBE with the Rev ARM, the relative weights were 15.4 and 2.4 kD, respectively (Tan and Frankel 1994). Nonetheless, to the extent that contributions from the peptides could be seen, in general there are no signals that appear to be due to the formation of particular peptide secondary structures. Thus, it is likely that the selected aptamers recognize their ARMs in an extended conformation (as has been observed for ARM interactions with an anti-Rev aptamer and with HIV-1 TAR) (Puglisi et al. 1992; Ye et al. 1999) and that the change in signal upon complex formation is due largely to the RNA

component. To analyze changes that occurred during complex formation, we have presented CD data as difference spectra—that is, the spectra corresponding to the RNAs alone have been subtracted from the spectra of the complexes.

Although it is difficult to determine fine structure using CD, we expect that similar changes in CD spectra are diagnostic of similar structural changes. A striking example of such similarities is seen for the anti-Rev aptamer 5 and the anti-Rex aptamer with their respective ARM peptides. The anti-Rex aptamer showed cross-recognition with both the Rex and Rev ARMs, and these complexes appear to adopt similar structures, as the difference spectra are almost superimpos-able (Fig. 2A). Likewise, the different spectra for the anti-Rev aptamer 5 bound by both the Rev and Rex ARMs have the same shapes (Fig. 2B). Interestingly, the difference spectra for both the anti-Rev and anti-Rex aptamers bound to the RevN7D peptide are unlike the spectra with the Rev and Rex ARMs. One possible explanation is that these RNAs must undergo additional structural rearrangements in order to accommodate the negative charge on the aspartate substitution. This is anecdotally supported by the observation that the replacement of N7 with an alanine improves the binding of the Rev ARM to both anti-Rev and anti-Rex aptamers, indicating that there may already be a slightly unfavorable interaction at this position that would be exacerbated by an aspartate substitution.

The cross-recognition between the BMV and CCMV ARM peptides and their respective aptamers was also investigated by CD. The difference spectra for the anti-CCMV aptamer with both ARMs have similar shapes, again indicating that recognition may be achieved via similar structures (Fig. 2D). However, the difference spectra for the anti-BMV aptamer and these ARMs were distinctly different in shape (Fig. 2E). The large increase at 265 nm seen with the anti-BMV aptamer bound to the BMV ARM is not seen when the same RNA binds the CCMV ARM. These results are consistent with binding assays that show K_ds with the BMV and CCMV ARMs of 20 and 260 nM, respectively. The tighter-binding, cognate ARM gives the larger spectral change. It may be that the BMV ARM makes multiple contacts that are not available to the CCMV ARM, and that making these contacts causes a more significant structural rearrangement in the RNA and complex.

Implications for ARMs as RNA-binding motifs

Frankel and coworkers previously discussed "chameleonism", the notion that the same peptide can bind to two different RNAs by assuming different conformations. In particular, the JDV Tat ARM could assume different conformations upon binding to different RNAs (Smith et al. 2000a). Likewise, structural studies have shown that the Rev ARM adopts an α-helical
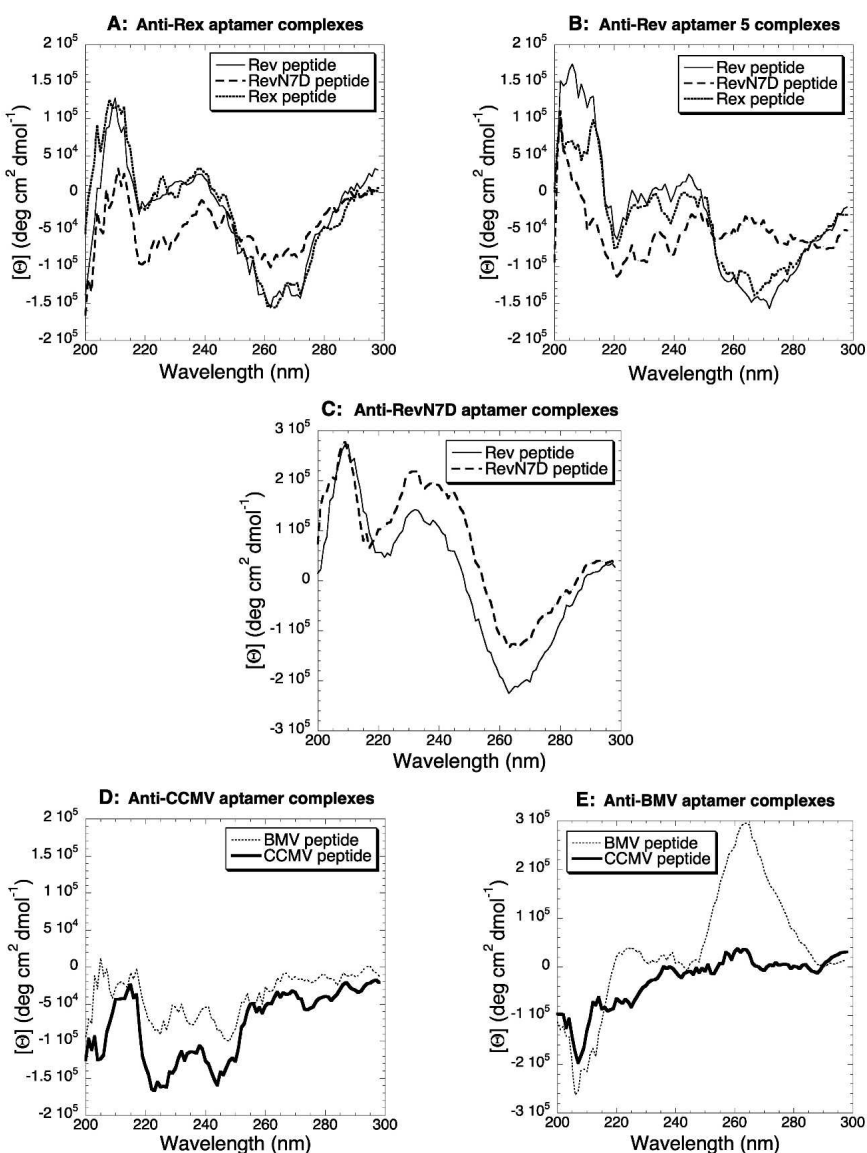
FIGURE 2. Circular dichroism spectra. Spectra are grouped according to peptide binding to the (*A*) anti-Rex aptamer, (*B*) anti-Rev aptamer 5, (*C*) anti-RevN7D aptamer, (*D*) anti-CCMV Gag aptamer, and (*E*) anti-BMV Gag aptamer. Spectra are shown as the difference between the complex and the RNA alone.

conformation when binding the RBE and an extended conformation when binding to an anti-Rev aptamer (Battiste et al. 1996; Ye et al. 1999).

In the present work we observed additional examples of cross-recognition between RNA molecules and ARM peptides. In characterizing the structural mechanisms underlying cross-recognition, all of our structural interpretations will be couched in terms of the complexes formed, as it should be recalled that the CD difference spectra cannot on their own be used to determine whether the RNA or peptide component of a complex has undergone conformational changes upon binding. As with chameleonism, the structural changes can differ between different peptides or RNA molecules. The anti-BMV aptamer complexes appear to undergo quite different structural transitions depending on whether the BMV ARM or CCMV ARM is bound (Fig. 2E). The structural transitions of complexes between the RevN7D ARM and anti-Rev and anti-Rex aptamers are quite different than the transitions that occur with the Rev or Rex ARMs (Fig. 2A,B), suggesting that this peptide may drive these RNAs into unique conformations.

In contrast, there also seem to be many instances when similar structural changes occur upon binding. When the anti-Rev aptamer selected here binds to either the Rev ARM or the Rex ARM, the complex appears to undergo similar conformational changes (Fig. 2B), as do complexes between the anti-Rex aptamer and these two peptides (Fig. 2A). Similarly, anti-CCMV aptamer complexes undergo similar spectral shifts and presumably similar conformational changes irrespective of whether the CCMV or BMV ARM peptide is being bound (Fig. 2D). Finally, anti-RevN7D aptamer complexes appear to undergo similar conformational changes irrespective of whether the RevN7D ARM or the wild-type Rev ARM is bound (Fig. 2C).

At root, cross-recognition is a result of the propensity for structural adaptation by the peptide or RNA partner or both, irrespective of whether different or similar conformational changes occur. It can be hypothesized that although in general there is specificity of ARMs for their cognate RNAs, some cross-recognition is observed in part because large numbers of interactions are possible between the multiple arginine residues of the peptide and the poly-anionic RNA backbone or hydrogen-bond donors and acceptors. In support of this hypothesis, our alanine-scanning experiments indicated that particular arginine residues in an ARM were required for binding to a given RNA. Moreover, since the entire complement of arginines did not necessarily participate in binding, it can be imagined that unique subsets of arginines are important for forming different structural complexes. For example, the Rev ARM peptide binds to both anti-Rev and anti-Rex aptamers via different sets of arginine residues. It is possible that any ARM that can present multiple arginine residues in a productive orientation may be bound by a given RNA partner. In this regard, it is interesting to note that selection experiments for peptides that bind to RNAs such as the

Rev-binding element or the λ boxB RNA yield multiple arginine-rich peptides, most of which do not resemble the wild-type ARM (Harada et al. 1996; Barrick et al. 2001).

ARMs and their RNA partners may be privileged in their abilities to form limited numbers of cross-recognized complexes. The wealth of arginine residues allows multiple interactions to potentially be formed, and peptide and RNA structural flexibility allows a subset of productive interfaces to be realized. Although structural work has hinted at this before (for reviews, see Frankel 2000; Das and Frankel 2003), we now show systematically that ARMs are versatile binding domains, able to bind diverse RNA molecules via different mechanisms. Overall, the ARM appears to be a much more adaptable and flexible RNA binding element than other RNA binding domains. For example, in vitro selections against the RRM-containing U1A protein yield conserved aptamer motifs throughout a given selection and between independent selections (Tsai et al. 1991; Cox et al. 2002a). That is, selection yields a single RNA "solution" to a recognition problem. The simplicity and adaptability of ARMs may be one reason that they are consistently chosen during evolution as recognition elements in viral proteins (Ellington et al. 1995, 1996).

One reason that ARM:RNA cross-recognitions may not be even more widespread is that the fits between RNAs and peptides are influenced not only by positive contacts but also by negative steric or electrostatic interactions. For example, the anti-RevN7D aptamer bound the Rev ARM peptide twofold worse than it bound the RevN7D peptide (Table 4). However, changing the asparagine to an alanine (N7A) restored binding (Table 5). Similarly, the anti-RevN7D aptamer did not bind RevR11Q. However, when arginine 11 was changed to alanine (R11A) the RNA bound better than it did to the wild-type Rev ARM peptide. In additional examples, anti-Rev aptamer 5 showed a greater affinity for the T1A and N7A variants of the Rev ARM peptide than for the wild-type ligand it was selected against; likewise anti-Rev aptamer 11 preferred the T1A, N7A, and W12A variants. For cross-recognition, the anti-Rex and anti-RevN7D aptamers bound more tightly to five and six different alanine replacement variants, respectively, of the Rev ARM peptide.

Overall, it is not unreasonable to think of ARM peptides as platforms for arginine recognition where all other residues play a limited (and potentially negative) role. From this perspective, it is noteworthy that none of the alanine substitutions was able to elicit tighter binding from the RBE. The primary difference between this naturally evolved system and the synthetically evolved system described herein is that in the former the peptide and RNA co-evolve their recognition interface, while in our system the peptide sequence is static. It seems likely that the selected RNA aptamers adapted as best they could to the interface, including altering the peptide conformation in order to do so. However, steric constraints on binding may remain that can only be resolved by further evolution of the peptide. In this regard, it is interesting to note that "perfect" interfaces are

the product of the conformational and steric contributions of each partner, and that neither partner may be able to completely optimize an interface on its own.

MATERIALS AND METHODS

Peptides and RNA preparation

Peptides were synthesized and purified (Biosynthesis) with a C-terminal pendant biotin. A pool containing 30 random residues (N30) flanked by static primer regions was used in all selections (Lato et al. 1995). RNAs were transcribed from synthetic DNA oligonucleotides that contained a T7 RNA polymerase promoter, using Ampliscribe T7 Transcription kits (Epicentre) according to the manufacturer's instructions, except that transcription was carried out at 42°C rather than 37°C. RNAs were purified on 8%–20% acrylamide (depending on RNA size), 7 M urea gels, eluted, and ethanol precipitated. RNAs for gel-shift experiments were dephosphorylated with calf intestinal phosphatase, phenol:chloroform and chloroform extracted, ethanol precipitated, and labeled at their 5′ ends using polynucleotide kinase and γ-^{32}P ATP. Kinase reactions were desalted using Centri-Spin 20 size exclusion columns (Princeton Separations).

Automated in vitro selection

Selection was performed on a Beckman Biomek 2000 workstation (Beckman-Coulter), using an automated selection protocol that has been described in detail (Cox et al. 1998, 2002b; Cox and Ellington 2001). Biotinylated peptides were loaded on streptavidin beads (DynaBeads M280; Dynal) at a ratio of 1:10 peptide per streptavidin site (according to the manufacturer's instructions). Free streptavidin sites were then blocked with biotin. Each round utilized 400 mg of peptide-derivatized beads, and the selection was initiated with 5 μg (~182 pmoles) of RNA. The selection binding buffer was 20 mM HEPES (pH 7.5), 100 mM KCl, 1 mM DTT, and 5 mM MgCl$_2$. Wash buffers were the same as binding buffer with the exception that the selection stringency was progressively increased by increasing the KCl concentration: 100 mM KCl for rounds 1 and 2, 200 mM for rounds 3 and 4, 400 mM for rounds 5 and 6, 600 mM for rounds 7 and 8, 800 mM for rounds 9 and 10, and 1 M for rounds 11–18. Selections were carried out against Rev, RevN7D, RevR11Q, P22 N, BIV Tat, CCMV Gag, and BMV Gag peptides. This scheme did not yield anti-Rex aptamers, and a different wash regime was then successfully employed: 100 mM KCl for rounds 1–6, 250 mM for rounds 7–9, 500 mM for rounds 10–12, and 1 M for rounds 13–18.

Aptamer cloning, sequencing, and conformational analysis

Aliquots (1 μL) of RT-PCR reactions from a given round of selection were further amplified and then ligated into a thymidine-overhang vector (TA Cloning Kit; Invitrogen). Templates for sequencing reactions were generated from individual colonies via PCR with M13 forward (−40) and M13 reverse primers. Cycle sequencing reactions were carried out with a CEQ DTCS Quick Start Kit (Beckman-Coulter) and the vendor's modified M13

sequencing primer (primer −47 seq) and were then developed on a CEQ 2000XL eight-channel capillary DNA sequencer (Beckman-Coulter). Aptamer secondary structures were predicted using RNA structure 3.6 (Mathews et al. 1999).

Binding assays

RNAs and peptides were incubated together for 30 min on ice in 10-μL reactions containing 1x selection buffer, 100 nM *E. coli* tRNA, and 10% glycerol. A final concentration of 1 nM radiolabeled RNA was mixed with increasing amounts of peptide (2–1024 nM). RNA:peptide complexes were separated at 200 V on 10% (100:1 acrylamide:bisacrylamide) polyacrylamide gels in 0.5x TBE buffer at 4°C for 1 h. Orange G was run in empty lanes to monitor the progress of the gel. Gels were then dried and bands were visualized using a PhosphorImager. Dissociation constants were estimated from PhosphorImager data.

Gel-shift experiments were initially carried out at a series of standard peptide concentrations (2–1024 nM) to determine approximately where the shift occurred. Once the approximate K_d value was determined, an additional series of experiments was performed with a more limited range of peptide concentrations (within 10 nM) to more precisely estimate the critical concentration required for the shift. An example of this analysis can be seen in Figure 1A,B.

Circular dichroism (CD) measurements

CD spectra were measured using an OLIS DSM 16 UV/VIS spectrophotometer (On Line Instrument Systems). Samples were prepared in 10 mM K$_2$HPO$_4$ (pH 7.55; H$_3$PO$_4$), 100 mM KCl, and 5 mM MgCl$_2$. Spectra were obtained at 5 μM RNA or peptide (for individual spectra) or an equimolar (5 μM) mix for the complexes in a 1-mM path-length cuvette. RNA samples were heat-denatured and allowed to refold prior to adding buffer and/or peptide. Spectra were recorded from 300 to 180 nm at 1 nm intervals with an integration time of 5 sec. CD scans were repeated five times, then averaged; data were smoothed over 5 points and plotted using Kaleidagraph.

REFERENCES

Austin, R.J., Xia, T., Ren, J., Takahashi, T.T., and Roberts, R.W. 2002. Designed arginine-rich RNA-binding peptides with picomolar affinity. *J. Am. Chem. Soc.* **124:** 10966–10967.

Barrick, J.E., Takahashi, T.T., Ren, J., Xia, T., and Roberts, R.W. 2001. Large libraries reveal diverse solutions to an RNA recognition problem. *Proc. Natl. Acad. Sci.* **98:** 12374–12378.

Baskerville, S., Zapp, M., and Ellington, A.D. 1995. High-resolution mapping of the human T-cell leukemia virus type 1 Rex-binding element by in vitro selection. *J. Virol.* **69:** 7559–7569.

Battiste, J.L., Mao, H., Rao, N.S., Tan, R., Muhandiram, D.R., Kay, L.E., Frankel, A.D., and Williamson, J.R. 1996. Alpha helix-RNA

major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **273:** 1547–1551.

Berova, N., Nakanishi, K., and Woody, R. 2000. *Circular dichroism: Principles and applications.* Wiley, New York.

Cox, J.C. and Ellington, A.D. 2001. Automated selection of anti-protein aptamers. *Bioorg. Med. Chem.* **9:** 2525–2531.

Cox, J.C., Rudolph, P., and Ellington, A.D. 1998. Automated RNA selection. *Biotechnol. Prog.* **14:** 845–850.

Cox, J.C., Hayhurst, A., Hesselberth, J., Bayer, T.S., Georgiou, G., and Ellington, A.D. 2002a. Automated selection of aptamers against protein targets translated in vitro: From gene to aptamer. *Nucleic Acids Res.* **30:** e108.

Cox, J.C., Rajendran, M., Riedel, T., Davidson, E.A., Sooter, L.J., Bayer, T.S., Schmitz-Brown, M., and Ellington, A.D. 2002b. Automated acquisition of aptamer sequences. *Comb. Chem. High Throughput Screen.* **5:** 289–299.

Das, C. and Frankel, A.D. 2003. Sequence and structure space of RNA-binding peptides. *Biopolymers* **70:** 80–85.

Ellington, A.D., Symensma, T.L., Giver, L., and Baskerville, S. 1995. A simple code for protein: RNA interactions. *Nucleic Acids Symp. Ser.* **33:** 156–159.

Ellington, A.D., Leclerc, F., and Cedergren, R. 1996. An RNA groove. *Nat. Struct. Biol.* **3:** 981–984.

Felber, B.K., Derse, D., Athanassopoulos, A., Campbell, M., and Pavlakis, G.N. 1989. Cross-activation of the Rex proteins of HTLV-I and BLV and of the Rev. protein of HIV-1 and nonreciprocal interactions with their RNA responsive elements. *New Biol.* **1:** 318–328.

Frankel, A.D. 2000. Fitting peptides into the RNA world. *Curr. Opin. Struct. Biol.* **10:** 332–340.

Frankel, A.D. and Young, J.A. 1998. HIV-1: Fifteen proteins and an RNA. *Annu. Rev. BioChem.* **67:** 1–25.

Gilligan 3rd, T.J. and Schwarz, G. 1976. The self-association of adenosine-5′-triphosphate studied by circular dichroism at low ionic strengths. *Biophys. Chem.* **4:** 55–63.

Giver, L., Bartel, D., Zapp, M., Pawul, A., Green, M., and Ellington, A.D. 1993. Selective optimization of the Rev-binding element of HIV-1. *Nucleic Acids Res.* **21:** 5509–5516.

Harada, K., Martin, S.S., and Frankel, A.D. 1996. Selection of RNA-binding peptides in vivo. *Nature* **380:** 175–179.

Hemmerich, P., Bosbach, S., von Mikecz, A., and Krawinkel, U. 1997. Human ribosomal protein L7 binds RNA with an α-helical arginine-rich and lysine-rich domain. *Eur. J. BioChem.* **245:** 549–556.

Jiang, F., Gorin, A., Hu, W., Majumdar, A., Baskerville, S., Xu, W., Ellington, A., and Patel, D.J. 1999. Anchoring an extended HTLV-1 Rex peptide within an RNA major groove containing junctional base triples. *Structure Fold. Des.* **7:** 1461–1472.

Lato, S.M., Boles, A.R., and Ellington, A.D. 1995. In vitro selection of RNA lectins: Using combinatorial chemistry to interpret ribozyme evolution. *Chem. Biol.* **2:** 291–303.

Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288:** 911–940.

Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D., and Williamson, J.R. 1992. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* **257:** 76–80.

Puglisi, J.D., Chen, L., Blanchard, S., and Frankel, A.D. 1995. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* **270:** 1200–1203.

Smith, C.A., Crotty, S., Harada, Y., and Frankel, A.D. 1998. Altering the context of an RNA bulge switches the binding specificities of two viral Tat proteins. *Biochemistry* **37:** 10808–10814.

Smith, C.A., Calabro, V., and Frankel, A.D. 2000a. An RNA-binding chameleon. *Mol. Cell* **6:** 1067–1076.

Smith, C.A., Chen, L., and Frankel, A.D. 2000b. Using peptides as models of RNA-protein interactions. *Methods Enzymol.* **318:** 423–438.

Solomin, L., Felber, B.K., and Pavlakis, G.N. 1990. Different sites of interaction for Rev, Tev, and Rex proteins within the Rev-responsive element of human immunodeficiency virus type 1. *J. Virol.* **64:** 6010–6017.

Su, L., Radek, J.T., Hallenga, K., Hermanto, P., Chan, G., Labeots, L.A., and Weiss, M.A. 1997. RNA recognition by a bent α-helix regulates transcriptional antitermination in phage λ. *Biochemistry* **36** 12722–12732.

Tan, R. and Frankel, A.D. 1994. Costabilization of peptide and RNA structure in an HIV Rev. peptide-RRE complex. *Biochemistry* **33:** 14579–14585.

Tan, R. and Frankel, A.D. 1995. Structural variety of arginine-rich RNA-binding peptides. *Proc. Natl. Acad. Sci.* **92:** 5282–5286.

Tan, R., Chen, L., Buettner, J.A., Hudson, D., and Frankel, A.D. 1993. RNA recognition by an isolated α helix. *Cell* **73:** 1031–1040.

Tao, J. and Frankel, A.D. 1993. Electrostatic interactions modulate the RNA-binding and transactivation specificities of the human immunodeficiency virus and simian immunodeficiency virus Tat proteins. *Proc. Natl. Acad. Sci.* **90:** 1571–1575.

Tsai, D.E., Harper, D.S., and Keene, J.D. 1991. U1-snRNP-A protein selects a ten nucleotide consensus sequence from a degenerate RNA pool presented in various structural contexts. *Nucleic Acids Res.* **19:** 4931–4936.

Unge, T., Solomin, L., Mellini, M., Derse, D., Felber, B.K., and Pavlakis, G.N. 1991. The Rex regulatory protein of human T-cell lymphotropic virus type I binds specifically to its target site within the viral RNA. *Proc. Natl. Acad. Sci.* **88:** 7145–7149.

Weiss, M.A. and Narayana, N. 1998. RNA recognition by arginine-rich peptide motifs. *Biopolymers* **48:** 167–180.

Xu, W. and Ellington, A.D. 1996. Anti-peptide aptamers recognize amino acid sequence and bind a protein epitope. *Proc. Natl. Acad. Sci.* **93:** 7475–7480.

Ye, X., Gorin, A., Ellington, A.D., and Patel, D.J. 1996. Deep penetration of an α-helix into a widened RNA major groove in the HIV-1 rev peptide-RNA aptamer complex. *Nat. Struct. Biol.* **3:** 1026–1033.

Ye, X., Gorin, A., Frederick, R., Hu, W., Majumdar, A., Xu, W., McLendon, G., Ellington, A., and Patel, D.J. 1999. RNA architecture dictates the conformations of a bound peptide. *Chem. Biol.* **6:** 657–669.