Costabilization of Peptide and RNA Structure in an HIV Rev Peptide−RRE Complex†

Ruoying Tan and Alan D. Frankel*

Department of Biochemistry and Biophysics and Gladstone Institute of Virology and Immunology, University of California, San Francisco, P.O. Box 419100, San Francisco, California 94141

Received July 29, 1994; Revised Manuscript Received September 30, 1994®

ABSTRACT: An arginine-rich peptide corresponding to amino acids 34−50 of the human immunodeficiency virus Rev protein has been shown to bind specifically to its RNA-binding site (RRE) when the peptide is in an α-helical conformation. Mutation of any one of six amino acids (Thr34, Arg35, Arg38, Arg39, Asn40, or Arg44) was shown to strongly decrease specific RNA-binding affinity in vitro, suggesting that these residues may contact specific bases or distinct structural features of the RNA. We now show that the four arginine side chains, and not just their charge, are important for specific binding in vivo, and present evidence that three additional arginines (Arg46, Arg48, and Arg50) may make electrostatic contacts to the RRE. RNA-binding specificity of the Rev peptide is temperature-dependent in vitro, correlating with α-helix unfolding. Circular dichroism experiments indicate that the peptide helical structure is stabilized when bound specifically to the RRE and that the RNA undergoes a conformational change upon binding. Because the structures of the peptide and RNA in this model system appear to be mutually stabilized upon binding, it is suggested that the entire complex may be viewed as a single folding unit.

From a biological perspective, the study of RNA−protein recognition is interesting because many cellular functions, including transcription, RNA splicing, and translation, depend on the specific interaction of proteins and RNA. From a macromolecular perspective, the problem is interesting because RNAs can fold into a wide range of tertiary structures, although so far only a few structures have been solved. The three-dimensional structure of tRNA has been known for quite some time, and the cocrystal structures of three tRNA synthetase−tRNA complexes are now known (Rould et al., 1989, 1991; Ruff et al., 1991; Cavarelli et al., 1993; Biou et al., 1994). While the basic shapes of the tRNAs in the three protein−RNA complexes are similar, important recognition features are quite different. For example, protein contacts occur largely in the RNA minor groove in the glutaminyl-tRNA synthetase complex (Rould et al., 1989, 1991), largely in the major groove in the aspartyl complex (Ruff et al., 1991; Cavarelli et al., 1993), and largely to the backbone of a long variable tRNA stem in the seryl complex (Biou et al., 1994). The tertiary folds of the synthetase proteins are very different, further illustrating how structurally diverse RNA−protein recognition is likely to be.

Fortunately, several common amino acid motifs have recently been identified in RNA-binding proteins, allowing them to be grouped into families (Mattaj, 1993; Burd & Dreyfuss, 1994). It was anticipated that related family members would use similar protein frameworks and amino acid contacts to recognize RNA, as observed, for example, with certain members of the helix−turn−helix DNA-binding family (Pabo et al., 1990). However, the wide diversity of RNA structures recognized by closely related RNA-binding proteins suggests that the "rules" for RNA recognition may be less obvious. For example, binding studies with RNP domain-containing proteins show that the same protein architecture, or even the same protein, can be used to recognize rather different RNA structures (Mattaj, 1993). Conversely, the diversity of tRNA synthetase structures indicates that it is possible for very different proteins to recognize the same basic RNA fold, although some small domains may be common to several synthetases (Schimmel et al., 1993). To further complicate the picture, RNA-binding domains from proteins even within the same family may adopt different conformations. For example, an arginine-rich domain from the human immunodeficiency virus (HIV) Rev protein recognizes its RNA site only when in an α-helical conformation (Tan et al., 1993) whereas related domains from the HIV and bovine immunodeficiency virus (BIV) Tat proteins appear to recognize their sites as unstructured or nonhelical peptides (Calnan et al., 1991a,b; Chen & Frankel, 1994; L. Chen and A.D.F., submitted; R.T. and A.D.F., in preparation). In this study, we further examine RNA recognition by the α-helical Rev peptide.

Rev is an essential regulatory protein encoded by HIV. Although its mechanism of action is not yet fully determined, Rev is believed to activate the nucleocytoplasmic transport of unspliced and incompletely spliced HIV mRNAs either directly, by facilitating their access to a nuclear RNA transport pathway (Emerman et al., 1989; Malim et al., 1989), or indirectly, by inhibiting their interaction with cellular splicing factors (Chang & Sharp, 1989; Kjems et al., 1991b). The unspliced and incompletely spliced mRNAs encode the viral structural proteins needed to assemble infectious virus particles.

To function, Rev must interact with a cis-acting RNA element, the Rev response element (RRE), located within the *env* gene. A single hairpin within the relatively large

† Supported by NIH Grants GM39589 and AI29135. R.T. is a Scholar of the American Foundation for AIDS Research (70405-15-RF).

* Address correspondence to this author at Gladstone Institute/ Biochemistry and Biophysics, UCSF, P.O. Box 419100, San Francisco, CA 94141. Phone: 415-695-3816. Fax: 415-826-8449.

® Abstract published in *Advance ACS Abstracts,* November 1, 1994.

0006-2960/94/0433-14579$04.50/0

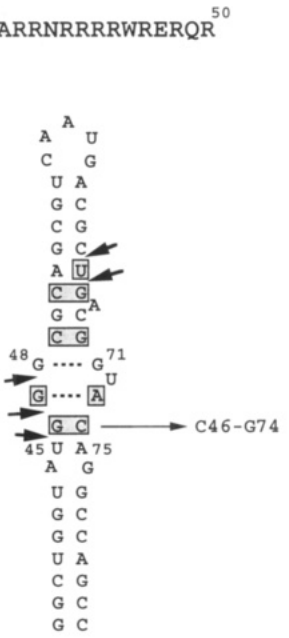

FIGURE 1: Sequence of Rev_{34-50} and RRE IIB RNA. The boxed nucleotides in IIB were shown to be important for Rev binding by mutagenesis, chemical interference, and in vitro RNA selection experiments (Bartel et al., 1991; Kjems et al., 1991a, 1992; Tiley et al., 1992). Arrows indicate positions of phosphates whose modification interferes with binding (Kjems et al., 1992). Two non-Watson−Crick base pairs (G48−G71 and G47−A73) are indicated by dashed lines. The mutant used in this study (C46−G74) was previously shown to decrease Rev binding affinity (Bartel et al., 1991; Tan et al., 1993).

RRE, known as stem-loop IIB, has been identified as the high-affinity Rev-binding site (Bartel et al., 1991; Cook et al., 1991; Heaphy et al., 1991; Kjems et al., 1992; Tiley et al., 1992). In vitro RNA selection, chemical substitution, and NMR experiments have shown that RRE IIB contains an asymmetric internal bulge with two non-Watson−Crick base pairs (Bartel et al., 1991; Iwai et al., 1992; Battiste et al., 1994; Peterson et al., 1994). The determinants of Rev binding in IIB are located primarily in the major groove in and near the asymmetric bulge and an adjacent single-nucleotide bulge (Figure 1) (Bartel et al., 1991; Heaphy et al., 1991; Kjems et al., 1992; Tiley et al., 1992). It has been shown that a 17-amino acid peptide spanning the arginine-rich region of Rev, Rev_{34-50}, binds specifically to the RRE in the same manner as the intact protein, provided that the peptide is in an α-helical conformation (Kjems et al., 1992; Tan et al., 1993). Binding experiments with mutant peptides identified six amino acids (Thr34, Arg35, Arg38, Arg39, Asn40, and Arg44) as essential for specific RNA binding in vitro (Tan et al., 1993). Here we examine the roles of individual arginine residues in vivo and present evidence that Arg35, Arg38, Arg39, and Arg44 may make specific contacts to bases or to sequence-dependent structural features of the RNA, while Arg46, Arg48, and Arg50 may make electrostatic contacts. Circular dichroism experiments demonstrate that specific binding to IIB RNA significantly stabilizes the peptide α-helical conformation and induces conformational changes in the RNA. The results are consistent with a model in which the orientation of the α-helix in the RNA major groove is determined largely by the tertiary structure of IIB RNA, with the structures of both the peptide and the RNA stabilized in the complex.

MATERIALS AND METHODS

Construction of Plasmids and CAT Assays. HIV-LTR-CAT reporters were constructed by oligonucleotide cassette mutagenesis in which the TAR stem-loop was replaced by the RRE IIB stem-loop or mutant RRE IIB(C46-G74) stem-loop [as in Tan et al. (1993)]. Tat−Rev peptide fusion proteins were constructed by cloning cassettes encoding Rev amino acids 3−70 or shorter Rev peptides into the *tat* gene of pSV2tat72 after Tat amino acid 48 [as in Tan et al. (1993)]. Mutations were confirmed by dideoxynucleotide sequencing. Reporter plasmids (50 ng) and Tat−Rev fusion protein plasmids (100 ng) were transfected into HeLa cells, and CAT activity was assayed after 48 h as described (Calnan et al., 1991a). Total plasmid concentrations were adjusted to 1 μg of total DNA with nonspecific pUC18 DNA.

RNA and Peptide Synthesis and Purification. RNAs were transcribed in vitro using T7 RNA polymerase and synthetic oligonucleotide templates (Milligan & Uhlenbeck, 1989). All RNAs contained GG at the 5′ end, which increases the efficiency of transcription, and CC at the 3′ end to base pair with the G's. To prepare randomly labeled RNAs, [α-^{32}P]-CTP (NEN, 3000 Ci/mmol) was included in the transcription reaction. RNAs were purified on 20% polyacrylamide/8 M urea gels, eluted from the gels in 0.5 M sodium acetate, pH 6.0, 1 mM EDTA, and 0.1% SDS, and ethanol precipitated. Purified RNA was resuspended in sterile deionized water. The concentrations of radiolabeled RNAs were determined from the specific activity of [^{32}P]CTP incorporated into the transcripts. Unlabeled RNAs were quantitated by spectrophotometry. RNAs were renatured by incubating in renaturation buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl) for 2 min at 90 °C followed by slow cooling to room temperature.

Peptides were synthesized on a Milligen/Biosearch model 9600 peptide synthesizer as previously described (Tan et al., 1993) or on an Applied Biosystems model 432A peptide synthesizer as described (Chen & Frankel, 1994). Peptides were purified on a C_4 reverse-phase HPLC column (Vydac) using acetonitrile gradients of 0.2%/min or 0.5%/min in 0.1% trifluoroacetic acid. Absorption spectra were used to determine peptide concentrations by tryptophan absorbance or using known peptides as standards. Purity and concentrations were confirmed by native gel electrophoresis (20% polyacrylamide in 30 mM sodium acetate, pH 4.5) in which peptides were visualized by Coomassie blue staining. Peptide molecular weights were confirmed by electrospray mass spectrometry (University of Michigan Protein and Carbohydrate Structure Facility).

RNA-Binding Gel Shift Assays. Gel shift assays were performed at 4 °C unless described otherwise. Peptide and RNA were incubated together for 15 min on ice in 10-μL binding reactions containing 10 mM HEPES-KOH, pH 7.5, 100 mM KCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 1 mM dithithreitol, 50 μg/mL *Escherichia coli* tRNA, and 10% glycerol. To determine relative binding affinities, 0.5 nM radiolabeled IIB RNA (or IIB mutant) was titrated with peptide. Peptide−RNA complexes were resolved on 10% polyacrylamide, 0.5× TBE gels that had been prerun for 1 h and allowed to cool to 4 °C. Gels were electrophoresed at 220 V for 3 h at 4 °C.

Circular Dichroism and Absorbance Melting Curves. Circular dichroism (CD) spectra were measured using an

Aviv model 62DS spectropolarimeter. Samples were prepared in 10 mM potassium phosphate buffer, pH 7.5, and 100 mM KF, and temperature was maintained at 0 °C unless described otherwise. Peptide concentrations were 20 μM, and RNA concentrations were 2−5 μM. RNA−peptide complexes were prepared at a 1:1 stoichiometry. Spectra were recorded using a 1-cm pathlength cuvette, and the signal was averaged for 2 s. Scans were repeated five times and averaged. CD melting curves were recorded with degassed, stirred samples at a heating rate of 1 °C/min, and samples were prepared in 10 mM potassium phosphate buffer, pH 7.5, and 100 mM KCl. Mean molar ellipticity for the peptide alone and peptide−RNA complexes was calculated per amino acid residue, and values for the RNA alone were normalized to the peptide−RNA complexes to allow comparison. Absorbance melting curves were recorded on an Aviv model 14DS spectrophotometer using conditions similar to the CD experiments. RNA concentrations were 15 μg/mL.

RESULTS

Specific Binding of the Rev Peptide to RRE IIB RNA in Vivo. In a previous study (Tan et al., 1993), gel shift experiments were used to measure the in vitro RNA-binding specificities of Rev_{34-50} derivatives by comparing apparent binding affinities to RRE IIB RNA and to a C46−G74 base pair mutant (Figure 1) known to decrease binding affinity of the intact Rev protein (Bartel et al., 1991). To confirm that the C46−G74 mutant also diminished peptide binding affinity in vivo, we constructed HIV long terminal repeat (LTR)-chloramphenicol acetyltransferase (CAT) reporters in which the TAR stem-loop was replaced by the RRE IIB stem-loop or by the mutant RRE IIB(C46-G74) stem-loop (Figure 2A). HIV Tat expression vectors were constructed in which Rev RNA-binding peptides, or a larger portion of Rev (amino acids 3−70), were fused to the activation domain (residues 1−48) of Tat (Figure 2B). With this system, Rev_{3-70} and Rev peptide binding to IIB or mutant IIB RNAs could be monitored, using transcriptional activation by Tat as a readout (Tan et al., 1993).

Rev_{3-70} or AAAA-Rev_{34-50}-AAAAR (the alanine termini help stabilize the helical conformation of the Rev peptide; Tan et al., 1993) fused to the C-terminus of Tat gave high levels of activation through the IIB reporter, but little activation through the mutant IIB reporter (Figure 2C). Thus, mutation of the G46−C74 base pair (to C46−G74) eliminates specific Rev peptide binding both in vivo and in vitro. A proline at position 42 eliminates peptide helix formation and specific IIB binding in vitro, and activation through IIB in vivo (Tan et al., 1993), and similarly shows no activity through the mutant IIB reporter (Figure 2C). A Rev protein truncated at amino acid 71 has been shown to bind RRE RNA with similar affinity and specificity as wild-type Rev and appears to correspond roughly to a structural (and highly helical) domain (Daly et al., 1993; Auer et al., 1994); thus Rev_{3-70} serves as a control in which the arginine-rich RNA-binding domain might be expected to be completely structured because it is placed within a larger protein context. The Tat-AAAA-Rev_{34-50}-AAAAR fusion protein has only slightly lower activity (~2-fold) than the Tat-Rev_{3-70} fusion protein, suggesting that the Rev peptide is largely structured in vivo and that other parts of the Rev protein are not necessarily required to stabilize the helical conformation.

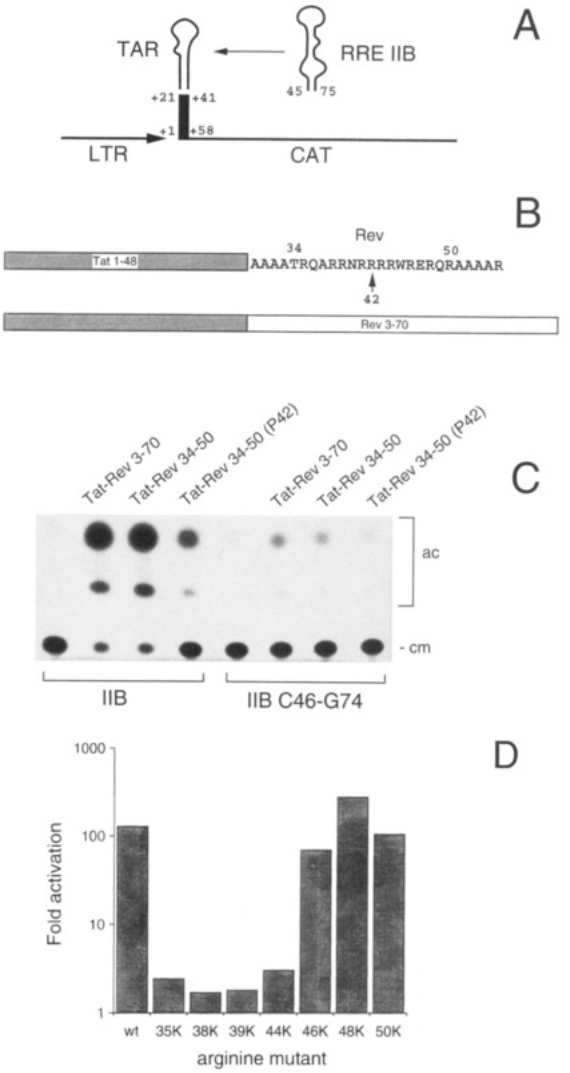

FIGURE 2: Binding of Rev peptides and arginine mutants to IIB RNA in vivo. (A) HIV LTR-CAT reporter plasmids were constructed in which the top part of the TAR stem-loop was replaced by the RRE IIB hairpin or the mutant IIB (C46−G74) hairpin. TAR nucleotide numbers are relative to the transcription start site (+1) and IIB nucleotide numbers correspond to those in Figure 1. (B) Plasmids expressing the Tat−Rev_{3-70} and Tat−Rev peptide fusions were constructed encoding the sequences shown. The Rev segments were fused to the activation domain of Tat (amino acids 1−48), which is required for transcriptional activation by Tat. (C) CAT assays of Tat−Rev_{3-70} and Tat−Rev peptide fusions. Assays were carried out using reporter plasmids containing the wild-type and mutant IIB RNA hairpins as indicated. Tat−Rev_{34-50}(P42) is a mutant in which Arg42 was changed to proline in the context of the Tat−Rev peptide fusion shown in panel B. Acetylated (ac) and unacetylated (cm) [^{14}C]chloramphenical are indicated. (D) Activation of the IIB CAT reporter by Tat−Rev peptide fusions encoding arginine mutants. Arginines at positions 35, 38, 39, 44, 46, 48, and 50 were mutated to lysines and CAT activities were determined as in panel C. Previous alanine scanning mutagenesis indicated that these are the only arginines required to maintain full RNA-binding specificity of the Rev peptide in vitro (Tan et al., 1993). Fold activation represents the increase in CAT activity relative to the reporter plasmid alone.

It also had been shown that substitution of Thr34, Arg35, Arg38, Arg39, Asn40, or Arg44 with alanine strongly decreased the RNA-binding specificity of Rev_{34-50} in vitro, that substitution of Arg46, Arg48, or Arg50 slightly decreased specificity, and that substitution of Arg41, Arg42, or Arg43 had no effect (Tan et al., 1993). All mutants had helical contents similar to the wild-type peptide. Thus it

was reasoned that Thr34, Arg35, Arg38, Arg39, Asn40, Arg44 might make specific contacts to the RNA and that Arg46, Arg48, and Arg50 might make nonspecific contacts. To more directly assess which of the seven arginines contribute specific contacts (presumably through hydrogen bonding to bases or to distinct structural features of the backbone) and which provide ionic contacts, we individually mutated each arginine in the Tat-AAAA-Rev_{34-50}-AAAAR fusion protein to lysine (instead of alanine) and measured activities of the mutants in vivo. Mutation of Arg35, Arg38, Arg39, or Arg44 strongly reduced activation through the RRE IIB site, whereas mutation of Arg46, Arg48, or Arg50 had little effect (Figure 2D). Thus, mutations that strongly decrease RNA-binding affinity in vitro correlate with a strict requirement for the arginine side chain whereas mutations that weakly decrease affinity correlate with the need for a positively charged side chain. The results are consistent with the proposal that four arginines (at positions 35, 38, 39, and 44) may interact with specific bases and/or sequence-dependent structural features of the RNA backbone, and that three additional arginines (at positions 46, 48, and 50) may provide stabilizing electrostatic interactions. The results further confirm that the peptide binding specificities measured in vitro accurately reflect RNA binding in vivo.

Temperature-Dependent Specificity. Because the stability of short α-helical peptides is often strongly dependent on temperature (Marqusee et al., 1989) and because specific binding of Rev_{34-50} to IIB RNA is thermodynamically coupled to α-helix formation (Tan et al., 1993), it was anticipated that specific RNA-binding affinity might be temperature dependent. Binding might also be temperature dependent if the RNA undergoes conformational changes or if the interaction has a significant enthalpic component. To examine possible contributions of α-helix formation, we first determined the CD melting profile of the suc-Rev_{34-50}-am peptide at 222 nm (a characteristic miminum for α-helix). This peptide has succinylated and amidated termini and a helical content of ~25% at 4 °C (Tan et al., 1993). As expected for an incompletely folded peptide, the spectra showed a significant decrease in helix content with increasing temperature and by 25 °C the helical content had decreased to ~10% (Figure 3A). A control peptide with unmodified termini (N-Rev_{34-50}-C) adopts a random structure (Tan et al., 1993) and showed no significant temperature-dependent change in structure (Figure 3A). The melting profile of IIB RNA was determined by monitoring the hyperchromic shift at 260 nm (Figure 3B). One major cooperative transition was observed with a $T_m = 73$ °C, a small pretransition was observed between 40 and 60 °C, and no apparent change was observed at lower temperatures.

RNA-binding gel shift experiments were performed with suc-Rev_{34-50}-am at 4 and 25 °C to determine binding specificities (Figure 3C). At 4 °C, discrimination between wild-type and mutant IIB RNAs was ~50-fold (additional binding titrations were performed at lower peptide concentrations than that shown in Figure 3C to determine accurate binding affinities), whereas at 25 °C, discrimination was only ~5-fold. Nonspecific RNA-binding affinities were similar at both temperatures. Thus, although the decrease in binding specificity was somewhat greater than the decrease in helix content over this temperature range (~10-fold versus ~2.5-fold), it appears that melting of the α-helix at least qualitatively correlates with loss of specific RNA binding.

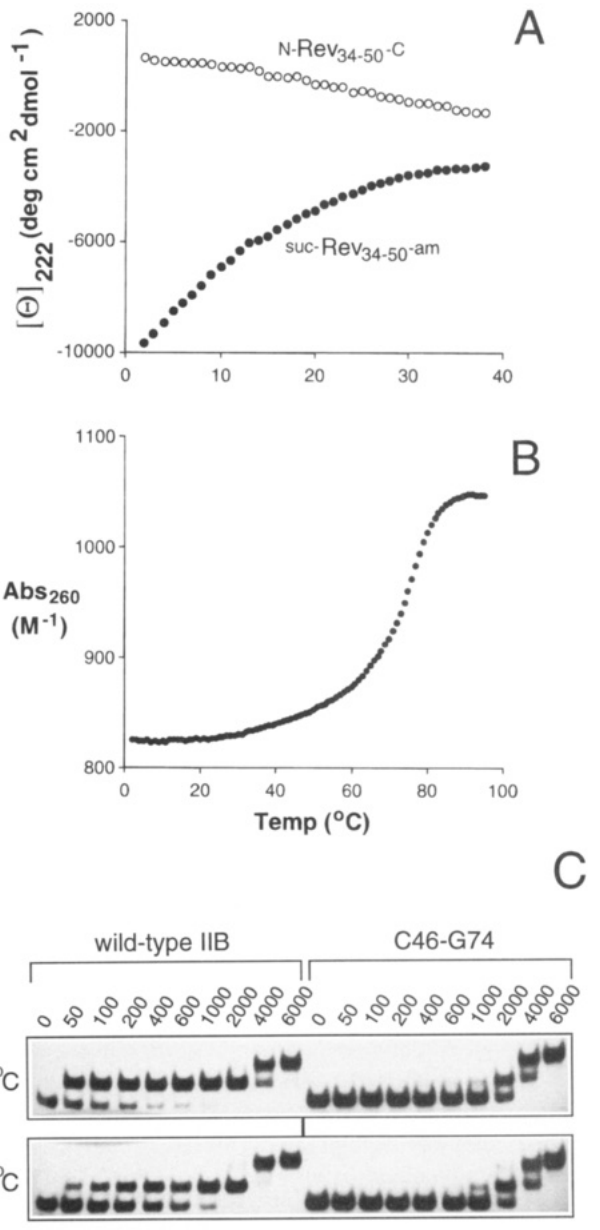

FIGURE 3: α-Helix content and specific RNA binding are temperature dependent. (A) Melting curve of the α-helix as monitored by CD ellipticity at 222 nm. Curves are shown for the mostly unstructured N-Rev_{34-50}-C peptide and for the partially helical suc-Rev_{34-50}-am peptide. The suc-Rev_{34-50}-am peptide has a helical content of ~25% at 4 °C, based on $[\theta]_{222} = -33\,500$ deg cm^2 dmol^{-1} calculated for a 100% helical 17-amino acid peptide (Chen et al., 1974). (B) Melting curve of wild-type IIB RNA as monitored by the hyperchromic shift at 260 nm. Absorbance has been calculated per molar concentration of nucleotide. (C) RNA-binding gel shift assays with the suc-Rev_{34-50}-am peptide at 4 and 25 °C. Peptide was titrated at the concentrations indicated (nM) using wild-type and mutant IIB RNAs. Gel temperatures were maintained as close as possible to 4 or 25 °C during electrophoresis.

While many explanations are possible for the lack of a quantitative correlation, it may be that the enthalpy of RNA binding is greater than the enthalpy of helix formation or that the RNA structure changes subtly with temperature (and is not detected in the RNA melting profile of Figure 3B).

Helical Conformation of the Rev Peptide Is Stabilized by Specific RNA Binding. The Rev_{34-50} peptide on its own displays partial α-helical character that can be stabilized by modification of the peptide termini to interact favorably with the helix macrodipole (Tan et al., 1993). Because specific

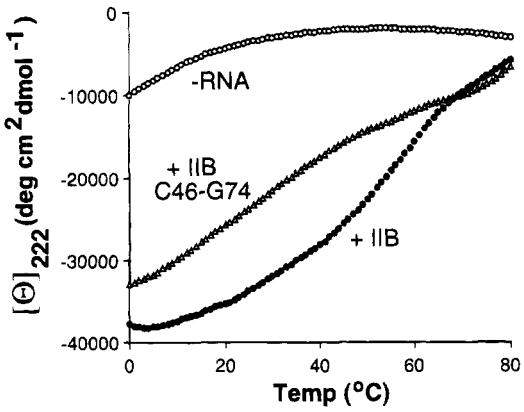

FIGURE 4: Helical conformation of the Rev peptide is stabilized by specific RNA binding. Melting curves for the suc-Rev_{34-50}-am peptide were determined by monitoring CD ellipticity at 222 nm for the peptide alone (−RNA) or for the peptide complexed at a 1:1 stoichiometry to wild-type IIB RNA (+IIB) or mutant IIB RNA (+IIB C46−G74). The CD melting curve of each RNA alone was subtracted from the melting curve of the corresponding complex. Note that the absolute value of $[\theta]_{222}$ for the wild-type complex is slightly greater than that expected for a fully helical peptide (see Figure 3 legend) and may reflect contributions from RNA conformational changes.

RNA-binding affinity correlates with α-helical content, it was expected that binding to RNA would influence the extent and/or stability of helix formation. As described above, the helix content of the suc-Rev_{34-50}-am peptide was ~25% at 4 °C in the absence of RNA and became mostly unstructured by 30 °C in the melting profile (also see Figure 4). When bound in a 1:1 complex with RRE IIB RNA, a sigmoidal melting transition was observed at 222 nm, suggestive of cooperative helix unfolding, with a T_m = 58 °C (Figure 4). It is assumed that the helical content of the peptide in the complex can be reasonably estimated from $[\theta]_{222}$ because the free RNA contributes little CD signal at this wavelength (Daly et al., 1990; data not shown); however, we cannot rule out small contributions resulting from changes in RNA structure (see the CD difference spectra below). We estimate that the α-helical content of the peptide in the complex increases to ~100% at 0 °C. The unfolding transition is relatively broad, perhaps reflecting dissociation of the complex over this temperature range. When bound in a 1:1 complex with mutant RRE IIB C46−G74, the α-helical content of the peptide increased almost as much as with wild-type IIB at 0 °C, but the T_m of the helix was substantially lower (28 °C; Figure 4). These results indicate that the helical content and stability of the Rev_{34-50} peptide are increased significantly upon specific RNA binding.

Circular Dichroism of IIB RNA and Peptide Complexes. A significant change in HIV TAR RNA structure (the binding site for Tat) has been observed by CD (Calnan et al., 1991a; Tan & Frankel, 1992) and NMR (Puglisi et al., 1992) spectroscopy upon specific binding of Tat peptides or arginine. RNase mapping, gel shift, and CD experiments suggested that the RRE also may change conformation upon Rev binding (Daly et al., 1990; Kjems et al., 1991a). To examine the effect of Rev_{34-50} peptide binding on the structure of the RRE IIB hairpin, we recorded CD spectra of wild-type and mutant IIB RNAs with the helical suc-Rev_{34-50}-am peptide and with a random coil mutant having a substitution of Arg42 to Pro, which virtually eliminates RNA-binding specificity (Tan et al., 1993). The CD difference spectrum (RNA−peptide complex minus free RNA) with the helical suc-Rev_{34-50}-am peptide and wild-type IIB showed significantly larger changes than with the C46−G74 IIB mutant (Figure 5A). The double minima observed at 208 and 222 nm are consistent with increased peptide α-helix. The minima are not entirely distinct because changes in RNA structure contribute to the spectrum below 222 nm. The CD difference observed near 280 nm in the wild-type IIB−peptide complex probably reflects differences in RNA base stacking (Tan & Frankel, 1992). A change in the environment of Trp45 may also contribute to the difference at 280 nm; however, experiments with a Trp45 to Ala mutant, which binds IIB RNA with wild-type affinity (Tan et al., 1993), gave similar CD results (data not shown). Difference spectra with the random coil suc-Rev_{34-50}-am mutant (Arg42 to Pro) showed considerably smaller changes than with the helical peptide and showed little difference between wild-type and mutant IIB RNAs (Figure 5B). Thus, it appears that conformational changes in RRE IIB occur upon specific binding of the α-helical Rev_{34-50} peptide,

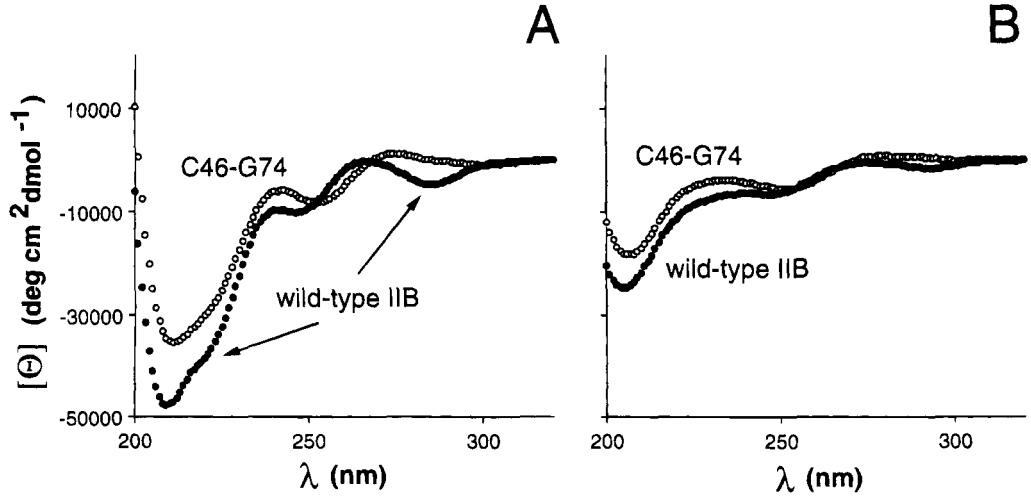

FIGURE 5: Difference CD of peptide−RNA complexes. Spectra of free RNAs and peptide−RNA complexes (1:1 stoichiometry) were recorded, and the free RNA spectra were subtracted from the spectra of the complexes. (A) Difference spectra with wild-type and mutant IIB RNAs using the helical suc-Rev_{34-50}-am peptide. (B) Difference spectra with wild-type and mutant IIB RNAs using the random coil suc-Rev_{34-50}-am (Arg42→Pro) mutant peptide.

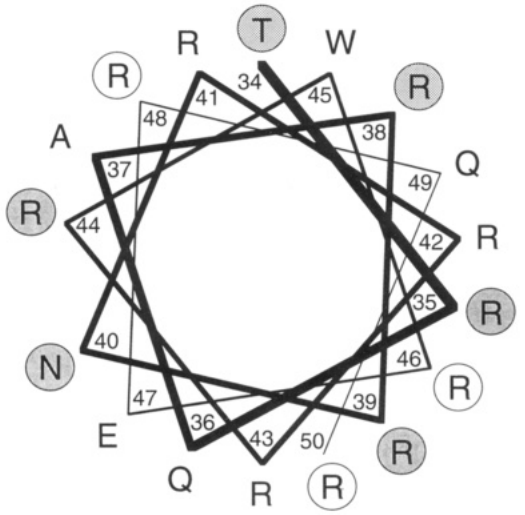

FIGURE 6: Helical projection of the Rev_{34-50} peptide. Mutation of any of six amino acids (shaded circles) to alanine strongly decreases RNA-binding specificity (Tan et al., 1993), suggesting that these residues make direct specific contacts to the RNA. The four shaded arginines cannot be replaced by lysines. Specific contacts may include interactions with bases or with sequence-dependent structural features of the backbone. Three additional arginines (open circles) can be replaced by lysines, suggesting that these residues make electrostatic contacts.

consistent with NMR studies of the free and bound RNA (Battiste et al., 1994; Peterson et al., 1994; discussed below).

DISCUSSION

It has been shown previously that the Rev_{34-50} peptide binds specifically to the major groove of RRE IIB RNA as a monomeric α-helix (Kjems et al., 1992; Tan et al., 1993). Substitution of Thr34, Arg35, Arg38, Arg39, Asn40, or Arg44 to alanine was found to strongly decrease specific RNA-binding affinity, whereas substitution of Arg46, Arg48, or Arg50 weakly decreased affinity (Tan et al., 1993). To help establish which of the seven arginines make specific contacts to the RNA (perhaps through guanidinium−base or guanidinium−backbone hydrogen bonds) and which make electrostatic contacts, we mutated each arginine to lysine and measured RNA-binding affinities in vivo. Mutation of arginines 35, 38, 39, or 44 strongly decreased IIB RNA-binding activity, suggesting that these residues make specific contacts, whereas mutation of arginines 46, 48, or 50 had no effect, suggesting that charged residues at these positions make electrostatic contacts. As discussed below, electrostatic contacts may contribute to RNA-binding specificity by helping to stabilize the bound RNA conformation. It also is possible that one or more basic residues at positions 46, 48, or 50 help neutralize the glutamic acid at position 47.

When placed in a helical projection, the nine amino acids important for binding are found distributed around the α-helix (Figure 6). At first glance this was surprising because one might have anticipated that side chains which contact the RNA might be located on a single face of the helix. However, the location of important nucleotides and phosphates in IIB (Figure 1) suggests that the RNA-binding site may be bipartite, and gel mobility (Kjems et al., 1992; Tan et al., 1993; R.T. and A.D.F., unpublished results), NMR (Battiste et al., 1994; Peterson et al., 1994), and molecular modeling (Leclerc et al., 1994) experiments suggest that the RNA may be distorted or bent in the complex, perhaps allowing the RNA to wrap around the α-helix. In this arrangement, it might be possible for several amino acids to contact the asymmetric internal bulge (comprising G46−C49 and G70−C74) and for other amino acids to contact nucleotides C51, U66, and G67, adjacent to the A68 bulge. As proposed by Weeks and Crothers (1991, 1993), bulges are likely to be important sites of protein recognition because they widen the major groove of adjacent A-form double helices, making the normally narrow and deep groove more accessible. Bulge regions also can contain non-Watson−Crick base pairs, such as the G•G and G•A pairs observed in RRE IIB (Bartel et al., 1991; Battiste et al., 1994; Peterson et al., 1994), or tertiary interactions, such as the base triple observed in TAR (Puglisi et al., 1992). It has been proposed that nonhelical geometries of the phosphate backbone near bulge regions may favor specific hydrogen bonding arrangements with the arginine guanidinium group (Calnan et al., 1991b), and it has been suggested that arginine−phosphate interactions at the bulge regions of IIB RNA might help orient the Rev α-helix in the major groove (Tan et al., 1993).

One interesting feature of the Rev peptide−IIB RNA complex is that both the peptide and RNA structures appear to be stabilized upon binding. The unbound peptide shows only partial α-helix formation at low temperature but becomes fully helical and relatively thermostable when bound to IIB RNA (Figure 4). Similar increases in α-helical content and thermostability have been observed upon binding of short basic peptides from GCN4 to specific DNA sites (Talanian et al., 1990; O'Neil et al., 1990; Weiss et al., 1990; Weiss, 1990) and upon binding of model basic peptides to nonspecific DNA (Johnson et al., 1994). Indeed, there appear to be many examples in which localized protein folding is coupled to specific DNA recognition (Spolar & Record, 1994). In the Rev peptide−RNA interaction, the conformation of IIB also changes upon peptide binding and involves stabilization of the structure of the bulge regions. NMR experiments indicate that the asymmetric internal bulge and base pairs adjacent to the A68 bulge are relatively disordered in the absence of peptide and that binding induces formation of two Watson−Crick (C49−G70 and G50−C69) and two non-Watson−Crick (G47−A73 and G48−G71) base pairs (Battiste et al., 1994; Peterson et al., 1994). A specific conformational change is supported by gel mobility shift (Kjems et al., 1992; Tan et al., 1993; R.T. and A.D.F., unpublished results), RNase footprinting (Kjems et al., 1992), and CD (this paper) experiments. Computer modeling studies also have suggested the protein−RNA interface may undergo structural adjustments upon binding (Leclerc et al., 1994). Because the conformations of the peptide and RNA are relatively unstable on their own, we suggest that the complex, rather than the individual components, may be viewed as a folding unit. It should be emphasized, however, that while the peptide−RNA interaction displays many of the sequence-specific properties of the full Rev−RRE interaction (Kjems et al., 1992; Tan et al., 1993), it is not known if the arginine-rich region forms a stable helix in the intact protein, if identical RNA conformational changes occur in the context of the larger RRE (conformational changes have been detected by RNase footprinting and CD experiments; Daly et al., 1990; Kjems et al., 1991a; Auer et al., 1994), or if binding of additional Rev monomers to the RRE influences interactions with the high-affinity IIB site. It will be interesting to determine whether the coupled folding

observed in the peptide model system extends to the intact Rev protein.

The results of peptide mutagenesis experiments suggest that a cooperative set of amino acid–RNA contacts are needed to stabilize the bound IIB RNA structure. Mutation of any one of six amino acids (Thr34, Arg35, Arg38, Arg39, Asn40, and Arg44) almost entirely eliminates RNA-binding specificity (Tan et al., 1993), and differences in gel mobility suggest that the RNA conformations formed with mutant peptides differ from that of the wild-type complex (Kjems et al., 1992; Tan et al., 1993). It is presumed that these six amino acids make critical specific contacts to bases or to sequence-dependent structural features of the RNA. Three additional arginines (Arg46, Arg48, and Arg50) may make electrostatic contacts, but even these appear to contribute to RNA-binding specificity since substitution with alanine slightly reduces binding specificity (Tan et al., 1993). We propose that electrostatic contacts in the Rev peptide–IIB RNA complex help stabilize the bound RNA conformation by neutralizing unfavorable phosphate interactions that may occur near the bulges. A similar role for electrostatic interactions has been proposed in the Tat–TAR interaction in which basic amino acids surrounding a single arginine increase the binding specificity to TAR (Tao & Frankel, 1993). The mutual stabilization of peptide and RNA structure may be a common feature of arginine-rich peptide–RNA interactions, and perhaps of other RNA–protein interactions.

ACKNOWLEDGMENT

We thank L. Chen for help with several plasmid constructions, Rene Tipton and Derek Hudson for peptide synthesis, members of the laboratory for helpful suggestions, and Susan Marqusee and James Williamson for comments on the manuscript.

REFERENCES

Auer, M., Gremlich, H. U., Seifert, J. M., Daly, T. J., Parslow, T. G., Casari, G., & Gstach, H. (1994) *Biochemistry 33*, 2988–2996.

Bartel, D. P., Zapp, M. L., Green, M. R., & Szostak, J. W. (1991) *Cell 67*, 529–536.

Battiste, J. L., Tan, R., Frankel, A. D., & Williamson, J. R. (1994) *Biochemistry 33*, 2741–2747.

Biou, V., Yaremchuk, A., Tukalo, M., & Cusack, S. (1994) *Science 263*, 1404–1410.

Burd, C. G., & Dreyfuss, G. (1994) *Science 265*, 615–621.

Calnan, B. J., Biancalana, S., Hudson, D., & Frankel, A. D. (1991a) *Genes Dev. 5*, 201–210.

Calnan, B. J., Tidor, B., Biancalana, S., Hudson, D., & Frankel, A. D. (1991b) *Science 252*, 1167–1171.

Cavarelli, J., Rees, B., Ruff, M., Thierry, J. C., & Moras, D. (1993) *Nature 362*, 181–184.

Chang, D. D., & Sharp, P. A. (1989) *Cell 59*, 789–795.

Chen, L., & Frankel, A. D. (1994) *Biochemistry 33*, 2708–2715.

Chen, Y. H., Yang, J. T., & Chau, K. H. (1974) *Biochemistry 13*, 3350–3359.

Cook, K. S., Fisk, G. J., Hauber, J., Usman, N., Daly, T. J., & Rusche, J. R. (1991) *Nucleic Acids Res. 19*, 1577–1583.

Daly, T. J., Rusche, J. R., Maione, T. E., & Frankel, A. D. (1990) *Biochemistry 29*, 9791–9795.

Daly, T. J., Rennert, P., Lynch, P., Barry, J. K., Dundas, M., Rusche, J. R., Doten, R. C., Auer, M., & Farrington, G. K. (1993) *Biochemistry 32*, 8945–8954.

Emerman, M., Vazeux, R., & Peden, K. (1989) *Cell 57*, 1155–1165.

Heaphy, S., Finch, J. T., Gait, M. J., Karn, J., & Singh, M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 7366–7370.

Iwai, S., Pritchard, C., Mann, D. A., Karn, J., & Gait, M. J. (1992) *Nucleic Acids Res. 20*, 6465–6472.

Johnson, N. P., Lindstrom, J., Baase, W. A., & von Hippel, P. H. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 4840–4844.

Kjems, J., Brown, M., Chang, D. D., & Sharp, P. A. (1991a) *Proc. Natl. Acad. Sci. U.S.A. 88*, 683–687.

Kjems, J., Frankel, A. D., & Sharp, P. A. (1991b) *Cell 67*, 169–178.

Kjems, J., Calnan, B. J., Frankel, A. D., & Sharp, P. A. (1992) *EMBO J. 11*, 1119–1129.

Leclerc, F., Cedergren, R., & Ellington, A. D. (1994) *Nature Struct. Biol. 1*, 293–297.

Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V., & Cullen, B. R. (1989) *Nature 338*, 254–257.

Marqusee, S., Robbins, V. H., & Baldwin, R. L. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5286–5290.

Mattaj, I. W. (1993) *Cell 73*, 837–840.

Milligan, J. F., & Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51–62.

O'Neil, K. T., Hoess, R. H., & DeGrado, W. F. (1990) *Science 249*, 774–778.

Pabo, C. O., Aggarwal, A. K., Jordan, S. R., Beamer, L. J., Obeysekare, U. R., & Harrison, S. C. (1990) *Science 247*, 1210–1213.

Peterson, R. D., Bartel, D. P., Szostak, J. W., Horvath, S. J., & Feigon, J. (1994) *Biochemistry 33*, 5357–5366.

Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., & Williamson, J. R. (1992) *Science 257*, 76–80.

Rould, M. A., Perona, J. J., Söll, D., & Steitz, T. A. (1989) *Science 246*, 1135–1142.

Rould, M. A., Perona, J. J., & Steitz, T. A. (1991) *Nature 352*, 213–218.

Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C., & Moras, D. (1991) *Science 252*, 1682–1689.

Schimmel, P., Giege, R., Moras, D., & Yokoyama, S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8763–8768.

Spolar, R. S., & Record, M. T., Jr. (1994) *Science 263*, 777–784.

Talanian, R. V., McKnight, C. J., & Kim, P. S. (1990) *Science 249*, 769–771.

Tan, R., & Frankel, A. D. (1992) *Biochemistry 31*, 10288–10294.

Tan, R., Chen, L., Buettner, J. A., Hudson, D., & Frankel, A. D. (1993) *Cell 73*, 1031–1040.

Tao, J., & Frankel, A. D. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 1571–1575.

Tiley, L. S., Malim, M. H., Tewary, H. K., Stockley, P. G., & Cullen, B. R. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 758–762.

Weeks, K. M., & Crothers, D. M. (1991) *Cell 66*, 577–588.

Weeks, K. M., & Crothers, D. M. (1993) *Science 261*, 1574–1577.

Weiss, M. A. (1990) *Biochemistry 29*, 8020–8024.

Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C., & Struhl, K. (1990) *Nature 347*, 575–578.