Helix–Loop–Helix Motif in HIV-1 Rev

Manfred Auer,* Hans-Ulrich Gremlich,‡ Jan-Marcus Seifert, Thomas J. Daly,§ Tristram G. Parslow,‖ Georg Casari, and Hubert Gstach

Sandoz Research Institute, Brunnerstrasse 59, 1235 Vienna, Austria, Sandoz Pharma AG, Preclinical Research, Basel, Switzerland, Repligen Corporation, One Kendall Square, Building 700, Cambridge, Massachusetts 02139, and Department of Pathology, University of California, San Francisco, California 94143

Received August 20, 1993; Revised Manuscript Received December 9, 1993®

ABSTRACT: Circular dichroism (CD) spectra of C-terminal deletion mutants of the HIV-1 Rev protein, Rev M9Δ14 (missing aa 68–112) and Rev M11Δ14 (lacking aa 92–112), indicated that Rev contains 46–49 residues in α-helical conformation within the N-terminal 71 or 95 amino acids of the 116 residue protein. Complexation with a 40-nucleotide fragment of the Rev responsive element, RRE, (G^{39} to C^{78}), containing the minimal element for Rev binding, induced an A to B form structural transition in the RRE fragment, whereas the percentage of α-helical conformation in the protein stays constant on substrate binding. When complexed to the RNA, neither mutant protein showed structural changes upon raising the temperature to 40 °C, as determined by the lack of decrease of the signal intensity at 222 nm, indicative for α-helical conformation. In contrast, Rev M9Δ14, which is shorter than Rev M11Δ14 by 24 amino acids, in the absence of RNA, lost about 60% of the spectral minima at 222 nm at the same temperature. The Rev M11Δ14 mutant, in the absence of RNA, showed a decrease of 20% in spectral intensity upon heating to 40 °C. Free and RNA-bound mutant proteins showed reversible transitions upon heating to 80 °C and subsequent cooling down to 10 °C overnight. The Rev peptide Cys 75-93, spanning the Rev transactivation domain, showed secondary structure in 40% and 60% hexafluoropropanol (HFP) solutions. The CD spectrum of the peptide in the presence of HFP showed little change upon heating to 60 °C. The CD spectrum of Rev 8-26 at a concentration of 1 mM in aqueous solution contained 28% α-helix, whereas no α-helix could be detected in a peptide solution of 0.2 mM. However, at this concentration helical conformation was induced as a function of HFP concentration with a maximum negative band intensity at 222 nm at 60% HFP. These data together with an analysis of Rev primary sequence based on general structural arguments and secondary structural calculations led to the identification of a potential $helix_1$–loop–$helix_2$ motif within the Rev N-terminus. This model proposes that between 70% and 100% of all residues in the RNA-binding and nucleolar localization domain of Rev are included in helical conformation. Three distinct hydrophobic motifs within the two helices suggest the formation of a hydrophobic $helix_1$–$helix_2$ contact between an "isoleucine motif" of $helix_2$ (I 52, I 55, and I 59) and one of the two hydrophobic motifs between residues 12 and 22 in $helix_1$.

The two regulatory gene products Tat and Rev, encoded by the human immunodeficiency virus (HIV), are essential for viral replication (Sodroski et al., 1986; Arya et al., 1985; Dayton et al., 1989; Terwilliger et al., 1988). Whereas Tat is a trans acting transcriptional activator, Rev post-transcriptionally induces cytoplasmatic accumulation of incompletely spliced viral RNA (Feinberg et al., 1986; Sodroski et al., 1986) encoding for viral structural proteins. For (full) Rev activity, two functional domains were found to be essential. The arginine-rich domain spanning the region between aa[1] 34 and 51 mediates direct and sequence-specific binding to the Rev responsive element (RRE)[2] RNA target sequence (Daly et al., 1989; Zapp and Green, 1989; Cochrane et al., 1990; Heaphy et al., 1991; Malim et al., 1990; Olson et al., 1990) and nuclear/nucleolar localization of the protein. Studies *in vivo* have shown that deletion of all or part of the RRE results in a loss of the ability of Rev to stimulate expression of viral structural proteins. The protein activation domain necessary to mediate Rev effector functions *in vivo* is located between aa 78 and 93 (Malim et al., 1989; Mermer et al., 1990; Venkatesh and Chinnadurai, 1990; Hope et al., 1990a; Malim et al., 1991; Weichselbraun et al., 1992). Both domains therefore represent attractive targets for chemotherapy (Bevec et al., 1992; Malim et al., 1992), provided that enough structural and mechanistic information is available for the application of drug design methods or the further development of lead compounds found through screening efforts.

The elucidation of the three-dimensional structure of the Rev wild-type protein using NMR and X-ray crystallography has been severely impeded by the strong tendency of the protein to aggregate (Heaphy et al., 1991; Karn et al., 1991). Information about secondary structure of the two regulatory proteins Tat and Rev has been obtained by circular dichroism (CD) spectroscopy. Whereas in early studies on recombinant Tat protein no secondary structure could be identified (Frankel et al., 1988), in a later study, helical conformation could be induced when CD measurements were performed in trifluoroethanol (Slice et al., 1992). Two groups investigated Tat peptides spanning either the activation domain or the arginine-rich regions which mediate RNA recognition by Tat. CD spectra of Tat aa 47–58 (RNA binding domain) suggest

* To whom correspondence should be addressed at Sandoz Research Institute. Telephone: 0043/86634/257. Fax: 0043/222/86634/727.
‡ Sandoz Pharma AG.
§ Repligen Corporation.
‖ University of California, San Francisco.
® Abstract published in *Advance ACS Abstracts,* February 15, 1994.
[1] Abbreviations: CD, circular dichroism; HFP, hexafluoropropanol; TFE, trifluoroethanol; RRE, Rev responsive element; aa, amino acid; MES, morpholinoethane sulfonic acid; DTT, dithiotreithol; Bz, benzoyl; Ibu, isobutyryl; Pac, phenoxyacetyl; $MgCl_2$, magnesium chloride.
[2] Designation of the RRE sequence based on Malim et al., 1990.

0006-2960/94/0433-2988$04.50/0

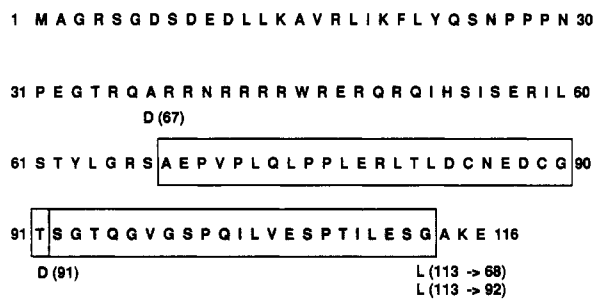

FIGURE 1: Amino acid sequence of HIV-1 Rev. Deletions of amino acids 68–112 for Rev M9Δ14 and 92–112 for Rev M11Δ14 (boxes) are indicated. The positions 67, 91, and 113 are replaced by D, D, and L, respectively.

a random coil conformation in the absence of nucleic acid (Calnan et al., 1991; Loret et al., 1991; Loret et al., 1992). Using a C-terminally extended Tat peptide, Tat (47–72), α-helical structure could be induced in 80% trifluoroethanol (Loret et al., 1992). Upon RNA binding, the peptides became partially or fully structured and induced a conformational change in the nucleic acid, which has been interpreted as a transformation from a typical A-form RNA toward B-form polynucleotide structure (Loret et al., 1992).

For HIV-1 Rev, Nalin and co-workers measured a CD spectrum characteristic for α-helical conformation and estimated between 39% and 47% α-helical content (Nalin et al., 1990). Similar results were obtained by Wingfield et al., 1991. Mixed with full-length RRE RNA, the spectrum was undistinguishable from that of the protein alone. Daly et al., 1990, showed that HIV-1 Rev at 5 °C in potassium phosphate buffer, pH 3, and 300 mM KF contains about 50% α-helix and 25% β-sheet. To localize regions with defined secondary structure within Rev more exactly, we used N- and C-terminal truncation and deletion mutants and short peptides spanning the regions which affect biological activity. A 40-nucleotide fragment of RRE (F8) which had been identified as the minimal element for Rev binding (Cook et al., 1991) was used for complexation with RNA. The correlation between functional domains and secondary structure would provide information essential to describe the interactions of Rev in more molecular detail.

Rev proteins with mutations in the N-terminus, including Rev M4 (Malim et al., 1989; Zapp et al., 1991; Malim & Cullen, 1991) and endoproteinase Lys-C cleaved wild-type Rev (Lys-C Rev), lacking the amino terminal 20 amino acids (Daly et al., 1993, personal communication) did show a 10- or 50-fold decrease in their ability to discriminate between sequence-specific RNA (RRE) and nonspecific RNA binding. Lys-C Rev displayed decreased α-helical conformation relative to wild-type Rev, as indicated from the reduction of the 222-nm signal intensity in the CD spectrum (Daly et al., 1993, personal communication, data not shown here). Therefore, for further CD studies, C-terminal deleted fragments of Rev were used (Daly et al., 1993c).

The C-terminal Rev deletion mutants Rev M9Δ14 and Rev M11Δ14 lack aa 68–112 (including the transactivation region) and aa 92–112, respectively (Figure 1). They bind to the Rev responsive element (RRE) with affinities between 2×10^{-10} and 8×10^{-10} M (Daly et al., 1993a). The dissociation constants (K_d) for binding of Rev M9Δ14 and Rev M11Δ14 to the shorter F8-RRE fragment are between 1 and 10×10^{-9} M, depending on whether filter binding or fluorescence titration methods were used for measurement. These K_d values guarantee that, at a minimal concentration of 0.025 mM of the two binding partners used in our experiments (and a 1:1 complex, Cook et al., 1991), at least 98% of Rev M9Δ14- or Rev M11Δ14-F8 complex is present in solution. In contrast to the case of mutations in the N-terminus of Rev, these mutants possess all of the determinants necessary to discriminate the RRE from nonspecific RNA. Rev M11Δ14, but not Rev M9Δ14, possesses biological activity equivalent to wild-type Rev in a cellular assay of Rev function (Daly et al., 1993c). The advantage the mutant proteins offer for biophysical studies is the reduced ability for RRE-dependent multimerization as shown by gel retardation assays. They are soluble in water and very dilute buffers up to a concentration of at least 10 mg/mL. Therefore, we focused on the N-terminal part of Rev, +/–RRE, as it contains all information for sequence-specific recognition.

MATERIALS AND METHODS

Expression and Purification of Rev M11Δ14 and Rev M9Δ14. Expression and purification of the C-terminal deletion mutant proteins has been described elsewhere in detail (Daly et al., 1993c). In short Rev M11Δ14 and Rev M9Δ14 were cloned for *E. coli* expression by translational couple to the 5′-end of the *E. coli* β-glucuronidase gene carried by expression plasmid pREV2.1 and expressed in *E. coli* strain RGN714 (Daly et al., 1993a). Purification of Rev M9Δ14 comprised lysis by Dynomil treatment, a Q-Sepharose ion-exchange step, precipitation of the protein-containing fractions with NaCl, and solubilization in 50 mM MES, pH 6.5, 8 M urea. Finally aliquots needed for experiments were refolded by dialysis into 20 mM MES, pH 6.5, or H_2O in 1:2 or 1:4 dilution steps. Purity of the refolded protein (>98%) was determined by SDS PAGE including Coomassie blue and silver staining, HPLC analysis on a RPC4 column (Vydac) in acetonitrile/H_2O/0.1% trifluoroacetic acid buffer systems, and N-terminal amino acid sequencing. Protein was quantitated by amino acid analysis and correlation with the area under the peak of the HPLC chromatogram. The protein samples were stored at –20 °C for months without change.

Purification of Rev M11Δ14 included the following procedures: lysis by Dynomil treatment, incubation with DNase I (Sigma), two cycles of resuspension and centrifugation, and precipitation with ethanol at 4 °C. After resuspension in 50 mM MES, pH 6.5, 8 M urea, 5 mM DTT column chromatographic purification started with a CM-Sepharose ion-exchange step followed by semipreparative HPLC on C4 reversed-phase material (Vydac). After lyophilization, the protein was refolded, characterized, and quantitated according to the procedures described for Rev M9Δ14.

Synthesis of Peptides. Rev peptides were synthesized using Fmoc methodology on an ABI synthesizer and purified to >95% purity by HPLC. The peptide sequence and purity was confirmed by amino acid analysis.

Synthesis of F8-RNA. The F8 RNA oligonucleotide was synthesized in either a 0.25 μmol or 1 μmol scale using an Applied Biosystems 394 A synthesizer and 2′-*O*-*tert*-butyldimethylsilyl-protected β-cyanoethyl *N*,*N*-diisopropylphosphoramidite monomers. The synthetic protocol and subsequent purification were carried out generally according to established procedures (Scaringe et al., 1990; Wu et al., 1988; Ogilvie et al., 1988; Usman et al., 1987). Sets of either A^{Pac}, G^{Pac}, C^{Bz}, and U phosphoramidite monomers from Glen Research (Sterling, VA) with phenoxyacetyl protection for A and G and benzoyl protection for C or A^{Bz}, G^{Ibu}, C^{Bz} and U phoshoramidite monomers from Milligen (Eschborn, FRG)

with isobutyryl protection for G were used. With 0.1 M amidite and 0.5 M tetrazole solution, the reagent consumption for each step at the 1 μmol scale was 15 μmol of amidite and 150 μmol of tetrazole; at the 0.25 μmol scale, consumption was ~7 μmol of amidite and 70 μmol of tetrazole. Coupling time was 12 min per cycle, and stepwise yields measured by trityl cation absorption at 498 nm were between 97% and 98%. Cleavage from the CPG support and base and phosphate protection were done according to the described method (Usman et al., 1987). Reaction times with saturated ethanolic ammonia at 55 °C were 6 h for A^{Pac}, G^{Pac}, C^{Bz}, and U or 14 h for A^{Bz}, G^{Ibu}, C^{Bz}, and U. Desilylation was done for 15 h with tetrabutylammonium fluoride/tetrahydrofuran. After being desalted by gel filtration, the oligonucleotide was purified by preparative gel electrophoresis, electroelution, and ethanol precipitation. Typical yields for the synthesis were about 150 μg of RNA for a 0.25 μmol scale about 500 μg for a 1 μmol scale. The RNA was renatured before use by heating to 70 °C for 2 min and then allowing to cool slowly for 1 h to room temperature. Concentration was determined using the absorbance at 260 nm and determination of the hypochromicity of the renatured sample before and after RNase A and T_1 digestion (25% hyperchromicity by digestion of 0.2 A_{260} units of RNA in 400 μL of 50 mM Tris, pH 7.5, 10 mM $MgCl_2$). Purity of the RNA was checked by gel chromatography and capillary gel electrophoresis. Direct saturation binding of F8-RNA to Rev (wild-type), Rev M9Δ14, and Rev M11Δ14 using nitrocellulose filter binding assays was used to determine affinities. Dissociation constants were in the range 0.5–1 nM (Daly et al., 1989; Daly et al., 1993c).

Circular Dichroism. Circular dichroism (CD) spectra were measured in 0.01 cm path length cells from 190 to 300 nm with a Jobin Yvon CD6 spectrometer. Spectra were recorded with an averaging time of 5 s and with a data-point increment of 0.5 nm. Between one and five scans were averaged to improve the signal to noise ratio. Samples were prepared in 20 mM MES, pH 6.5, at a concentration of 0.025 mM, 0.2 mM, or 1 mM. Spectra were recorded at 10 °C. For temperature studies the samples were heated to 80 °C in steps of 10 degrees using a circulating water bath. Further heating of the cuvettes to higher temperatures was impossible due to technical reasons. Equilibration time was 30 min per step. After the sample was cooled down to room temperature overnight, spectra recording was repeated to check reproducibility. The mean residue ellipticities of the protein samples were calculated per amide bond (71 for Rev M9Δ14 and 95 for Rev M11Δ14). RNA ellipticity was calculated per nucleotide (40 for F8-RNA). CD spectra of protein/RNA complexes were either expressed as per base or per amide bond in order to approximate difference spectra (Loret et al., 1992). Hexafluoropropanol was used to investigate the propensity of Rev peptides to adopt a helical conformation. From an extended series of studies on protein folding it has become clear during the last years that TFE or HFP can only induce helix in a linear peptide sequence in regions for which there is a significant preference for helical conformation (Dyson et al., 1992; Soennichsen etal., 1992; Shin et al., 1993; Waltho et al., 1993). In the absence of such a preference, no evidence for secondary structure formation, helix or otherwise, has been found.

Molecular Models. Secondary structure prediction was performed using the GCG programs package (Genetic Computer Group (1991), Version 7, 575 Science Drive, Madison, WI 53711) following the Chou and Fasman algorithm (Chou & Fasman, 1974). Model building was performed on a Silicon graphics workstation using the programs Sybyl and Insight. (Only those general structural arguments were used which were supported by experimental results from CD.)

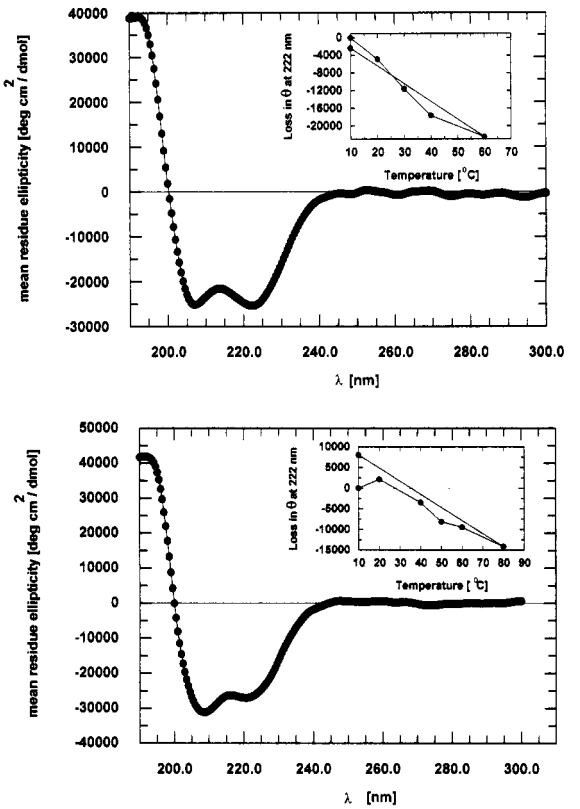

FIGURE 2: CD spectra of C-terminal Rev deletions mutant proteins M9Δ14 (A, top) and M11Δ14 (B, bottom) in 20 mM MES, pH 6.5, at a concentration of 0.025 mM, 10 °C (●). Inset: The loss of relative band intensity at 222 nm, indicative for α helical conformation, on heating in steps of 10 degrees, and regain of intensity on subsequent cooling to the starting temperature.

RESULTS

CD Spectrum of Rev M9Δ14. Biochemical experiments have shown that the first 68 amino acids of Rev plus four terminal residues (113–116) contain the sequence information required to confer affinity and specificity in RNA binding upon the Rev transactivator (Daly et al., 1993c). Peptides spanning the genetically defined RNA binding domain (residues 17–58) or a mutant protein missing the first N-terminal 20 amino acids almost completely lost their ability to discriminate sequence-specific from nonsequence-specific RNA (Daly et al., 1993, unpublished results). The high solubility of the Rev M9Δ14 protein allowed the measurement of a CD spectrum at low salt concentration (20 mM MES buffer, pH 6.5, at a protein concentration of 0.025 mM, 10 °C). These studies were aimed to obtain a better insight into the distribution of secondary structural elements essential for RNA-binding activity (specificity and activity) within the Rev protein. The spectrum was dominated by two minima at 207 and 222 nm characteristic for the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ electronic transitions seen for α-helical conformation (Figure 2A). Analysis of CD measurements using the CONTIN algorithm described by Provencher and Gloeckner (1981) yielded 65 ± 3% α-helix and 34 ± 9% random coil, which suggests that about 46 of all 71 residues adopt an α-helical conformation. For determination and comparison of thermal behavior, the protein was heated to 60 °C in steps

of 10 degrees and cooled down again to the starting temperature overnight. The peak intensity at 222 nm, in deg cm^2 dmol^{-1} per amide bond, [θ_{222}], characteristic for α-helical structure, was used as a measure of stability of the protein. Heating of the Rev M9Δ14 sample was accompanied by a gradual loss of negative band intensity (22 470 [θ_{222}] at 60 °C from a starting ellipticity of −28 590 deg cm^2 dmol^{-1}). After being cooled to 10 °C overnight, the protein had almost completely refolded (2365 deg cm^2 dmol^{-1} less than start intensity) (Figure 2A).

CD Spectrum of Rev M11Δ14. In addition to the sequence information contained in Rev M9Δ14, Rev M11Δ14 comprises the Rev transactivation domain genetically mapped between aa 78 and 93 (Figure 1). In a cellular transfection assay, this protein showed wild-type activity (Daly et al., 1993c). The CD spectrum of Rev M11Δ14, determined under the same conditions as described for Rev M9Δ14, contained approximately 40% α-helix, 21% β-sheet, and 39% random or turn potential (Figure 2B). In contrast to Rev M9Δ14, which showed a decrease of the negative band intensity at 222 nm (−28 590 deg cm^2 dmol^{-1}) by 22 472 [θ_{222}] on heating to 60 °C, M11Δ14 (mainly M9Δ14 extended by the activation domain) showed only a decrease from an ellipticity of −26 795 deg cm^2 dmol^{-1} to −17 317 deg cm^2 dmol^{-1}, corresponding to a net loss of 9478 [θ_{222}] units under the same conditions. Taking into consideration that at 60 °C no complete denaturation of the proteins had occurred and only a rough comparison of thermodynamic stability is possible, it can be concluded that helical regions in Rev M11Δ14 are approximately twofold more stable to thermal denaturation than helical structure in Rev M9Δ14. Furthermore, the spectrum of the refolded protein at 10 °C had 8040 deg cm^2 dmol^{-1} more negative intensity at 222 nm than the spectrum of the protein that had not been heated. This increase in α-helical signal is also reflected in the CONTIN fit of the CD curve with approximately 52% α-helix, 16% β-sheet, and 31% β-turn and random coil structure. An α-helix content of 52% corresponds to 49 residues in the helical form, whereas 40% would only allow 38 residues to be put into α-helical elements. As the mutant proteins were purified under denaturing conditions (reversed-phase chromatography) and refolded through dialysis, we assume that Rev M11Δ14 has several possible folding pathways, and a heating/cooling cycle is necessary to adopt a thermodynamically most stable conformation. In addition, the region from aa 68 to 92 must be responsible for a considerable enhancement in thermodynamic stability of the whole protein. In order to investigate whether this is reflected in the Rev activation domain, the peptide Rev-Cys75-93 was investigated by CD.

CD Studies of Rev Cys75-93. Rev Cys75-93 was unstructured in 20 mM MES, pH 6.5 (Figure 3A, circles). Hexafluoropropanol (HFP) preferentially induces α-helical structure in peptides which exhibit the tendency to populate backbone conformations in the α-helical region of (Φ,Ψ) space (Dyson et al., 1992; Zhou et al., 1992). In the presence of 20% HFP the Rev Cys75-93 peptide showed partial folding. It was completely folded, as observed by the unchanged spectra, at 40% and 60% HFP (Figure 3A for spectrum with 40% HFP, filled circles). Chou–Fasman secondary structure calculations proposed approximately 30% β-sheet and less than 20% helix. A temperature study performed in 40% HFP revealed unusually high thermal stability for a peptide of that length, with only minor losses of signal intensity at 222 nm up to 80 °C: −592 deg cm^2 dmol^{-1} at 20 °C, −1054 at 40 °C, −827 at

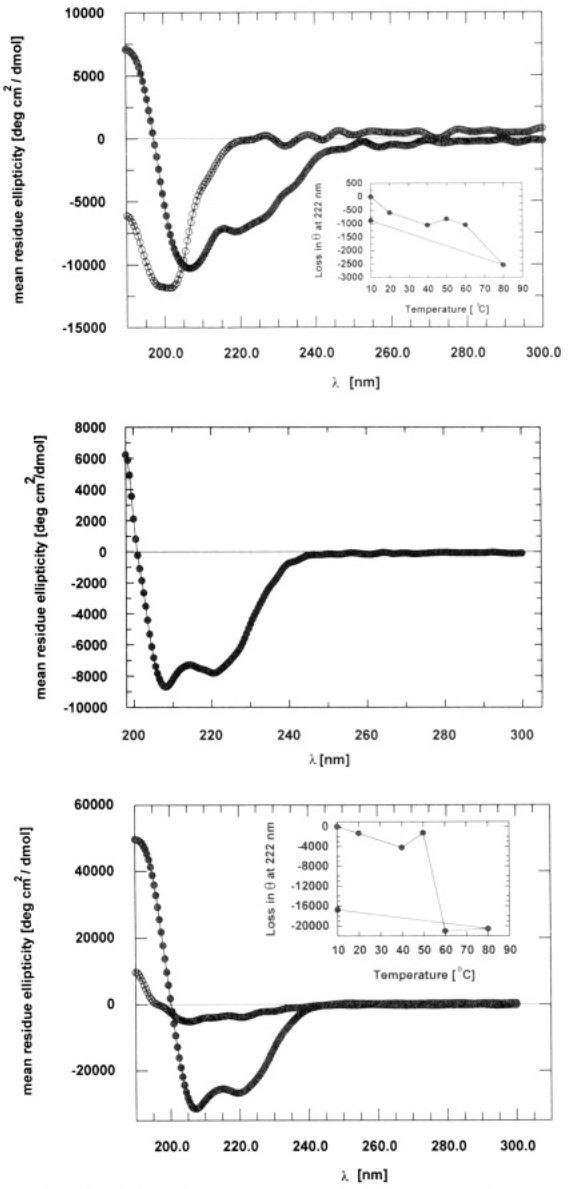

FIGURE 3: (A, top) (○) Random coil CD spectrum of the peptide Rev Cys 75–93 in 20 mM MES, pH 6.5. (●) In 40% HFP the peptide is partially structured. Inset: heating of the peptide in 40% HFP results in only 15% loss of the signal intensity at 222 nm (●). (B, middle) CD spectrum of the N-terminal peptide Rev 8–26 at a concentration of 1 mM in 20 mM MES, pH 6.5. (C, bottom) (○) CD spectrum of the N-terminal peptide Rev 8–26 at 0.2 mM without HFP added. (●) In the presence of 60% HFP, peaks at 207 and 222 nm indicate α helical structure. Inset: In 20% HFP, the structure is stable up to 50 °C; no refolding had occurred after cooling (●).

50 °C, −1043 at 60 °C, and −2538 deg cm^2 dmol^{-1} at 80 °C. After cooling, all but 881 deg cm^2 dmol^{-1} of the original band intensity [θ_{222}] (−7011 deg cm^2 dmol^{-1}) was regained. The inherent thermal stability of the Rev activation domain could therefore be responsible for the difference in stability between Rev M9Δ14 and Rev M11Δ14.

CD Studies of Rev 8-26. Among the short peptides investigated by CD, Rev 8-26 was predicted to possess the greatest amount of α-helical structure. This N-terminal fragment of Rev at low concentration (0.2 mM), and low temperature, showed little evidence for helix formation (Figure 3C, circles). However, at a concentration of 1 mM of peptide in aqueous solution, the two characteristic bands at 207 and 222 nm indicated helical structure (Figure 3B). Fitting of the data revealed a helical content of 28%. The diluted peptide showed HFP concentration-dependent induction of confor-

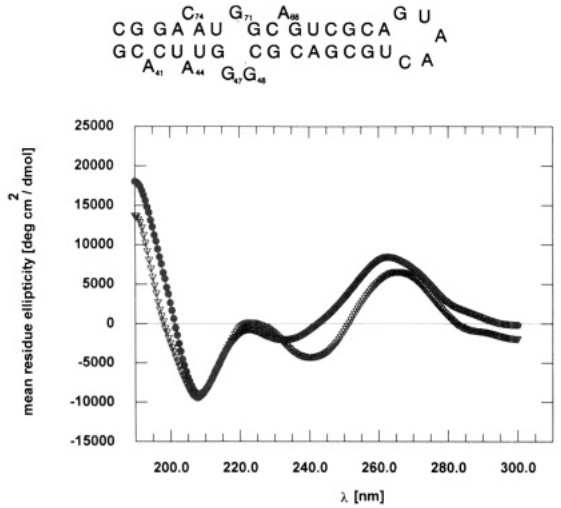

FIGURE 4: (A, top) Secondary structure of F8-RNA as calculated with the Zuker algorithm. In contrast to the case of the F8-RNA 40-mer included in the full-length RRE, the 5′ end of the sequence is shifted in the 3′-direction by two nucleotides, resulting in a pure G bubble (G^{71}, G^{47}, G^{48}). The important non-Watson–Crick base pair between G^{48} and G^{71} is conserved. U^{70} and A^{69}, which are predicted to be located in the bubble in the full RRE, become base-paired to U^{45} and G^{46}, shifting C^{74} (otherwise paired to G^{46}, and being important for Rev binding) into a bulge position. (B, bottom) (●) CD spectrum of the 40 nucleotide long F8-RNA at 10 °C and 0.025 mM. (▽) Difference spectrum of the Rev M11Δ14/F8-RNA complex minus the Rev M11Δ14 spectrum expressed as deg cm^2 dmol^{-1} per amide bond.

mation with approximately 54% helical structure in 20% HFP and 90% helical structure in 60% HFP (Figure 3C, filled circles). The thermal stability (in 20% HFP), however, was considerably lower than that of Rev Cys75-93. Although stable up to almost 50 °C, 21 017 deg cm^2 dmol^{-1} of the negative signal at 222 nm (–25 663 deg cm^2 dmol^{-1}) was lost on raising the temperature to 60 °C. Following heating, only –8932 deg cm^2 dmol^{-1} of the signal could be detected at 222 nm after cooling to 10 °C, reflecting a low refolding potential.

CD Studies of F8-RNA. Two different theoretical stem–loop structures for RRE RNA have been proposed in the literature. Early reports used the model predicted by the RNA-folding programs of Maizel (Le et al., 1990; Malim et al., 1989). However, the RNA-folding programs of Zuker predict a more stable structure (Zuker, 1989), with U^{45} and G^{46} paired to A^{73} and U^{72}, respectively. This prediction results in a bubble structure with G^{47}, G^{48}, G^{71}, U^{72}, and A^{73} either unpaired or non-Watson–Crick base paired. When the Zuker algorithm was used to calculate the secondary structure of the minimal binding element from nucleotide 39 to 78 (F8-RNA, 40-mer), defined by Cook et al., 1991, a change in the predicted base pairing was observed. G^{67}, C^{69}, G^{70}, U^{72}, A^{73}, A^{75}, G^{76}, G^{77}, and C^{78} were paired to C^{51}, G^{50}, C^{49}, G^{46}, U^{45}, U^{43}, C^{42}, C^{40}, and G^{39}. A^{41} and A^{44} opposite to C^{74} and A^{68} are unpaired, and only G^{71}, G^{47}, and G^{48} remain in the bubble structure (Figure 4A). Rev was described to be essentially a poly G binding protein (Daly et al., 1993a). At 25 °C, $\Delta G°$ for this predicted conformation is –15.2 kcal/mol. At higher temperature the bubble structure becomes more extended with various suboptimal folding predictions at 37 °C with $\Delta G°$ around –10 kcal/mol. As there is no proven three-dimensional structure either of RRE or RRE subfragment available at the moment, relative affinities for the Rev wild-type protein were used as criteria for substrate qualities of F8-RNA, which proved to be as potent in competing with ^{32}P-labelled RRE for Rev binding as stem–loop-II RNA (Cook et al., 1991).

The CD spectrum for F8-RNA alone is typical for RNA in the A-conformation of the nucleotide helix and is very similar to the CD spectrum of TAR RNA (Loret et al., 1992), as shown in Figure 4B (filled circles). The only detectable difference to the published TAR spectrum is a very low signal intensity at 222 nm. Thermal stability was studied by monitoring at wavelengths 240 and 265 nm, which are characteristic for RNA. Only a minor contribution of the peptide spectrum is anticipated at these wavelengths. The F8-RNA was stable up to 50 °C with a reduction of the (265 nm) intensity at 60 °C from 19 533 deg cm^2 dmol^{-1} to 16 571 deg cm^2 dmol^{-1} and to 11 334 deg cm^2 dmol^{-1} at 80 °C. This cooperative decrease of signal reflects asymmetric base–base interactions due to RNA secondary and probably tertiary structure (Johnson Jr., 1985). Renaturation was quantitative with an increase from –19 533 deg cm^2 dmol^{-1} to –20 768 deg cm^2 dmol^{-1} in intensity.

Complex between M9Δ14 and F8-RNA. In order to identify changes in the protein and/or nucleic acid structure due to complex formation, 1:1 mixtures of F8 and Rev M9Δ14 were formed (>98% complexed ligands) and the CD measurements analyzed by calculation of difference spectra to estimate changes in the nucleic acid and protein structure. In addition changes in thermal stability of the complex were investigated.

The CD spectrum of F8/M9Δ14 (Figure 5A) expressed as deg cm^2 dmol^{-1} per base shows a weak decrease of the band at 263 nm and an increase in the 240-nm signal relative to the case of the F8-RNA spectrum. As there are little or no influences of the protein in that region of the spectrum, these results taken together with the red shift of the crossover can be interpreted as an indication for a transition to a more B-form-like structure of the RNA. There is no strong involvement of base stacking in complex formation. In contrast to the case of M9Δ14 alone (at 40 °C, 17 644 units lost of the –28 590 deg cm^2 dmol^{-1} signal at 222 nm, where the RNA itself shows no band), the protein in the presence of the RNA is fully stable up to 40 °C. A gradual decrease in intensity to a maximum of 12 545 deg cm^2 dmol^{-1} net loss at 222 nm is observed upon heating to 60 °C (Figure 5B). Cooling the sample to 10 °C refolds the protein completely, as indicated by an increase in negative signal intensity at 222 nm relative to the start spectrum by 3525 deg cm^2 dmol^{-1} (Figure 5B). A comparison of the spectrum of F8-RNA alone with the M9Δ14 difference spectrum (M9Δ14/F8 – M9Δ14) reveals that this gain of about 10% in intensity is due to the A to B form transition of the nucleic acid (data not shown). Thus, in total, the helical content in the protein, as represented by the signal at 222 nm, stays constant on RNA binding. In the protein difference spectrum (M9Δ14/F8 – F8), however, the negative band at lower wavelength is moved from 207 to 211 nm. Together with a 5-nm red shift of the crossover, these observations are consistent with a small induction of β-sheet structure (Figure 6A).

Complex between M11Δ14 and F8. In the CD spectrum of a 1:1 complex of M11Δ14 and F8-RNA (Figure 5C), the decrease of the RNA signal at 265 nm and increase at 240 nm is slightly higher than those with the M9Δ14:F8 complex. As M11Δ14 alone is already more stable thermodynamically than M9Δ14, on a mean residue ellipticity level, differences between the complexed and the uncomplexed protein are not very marked. However, relative to the case of destabilization of free M11Δ14, the RNA-bound protein shows about 50% higher signal intensity at 60 °C (5110 deg cm^2 dmol^{-1} in comparison to 9478 deg cm^2 dmol^{-1} loss in signal at 222 nm).

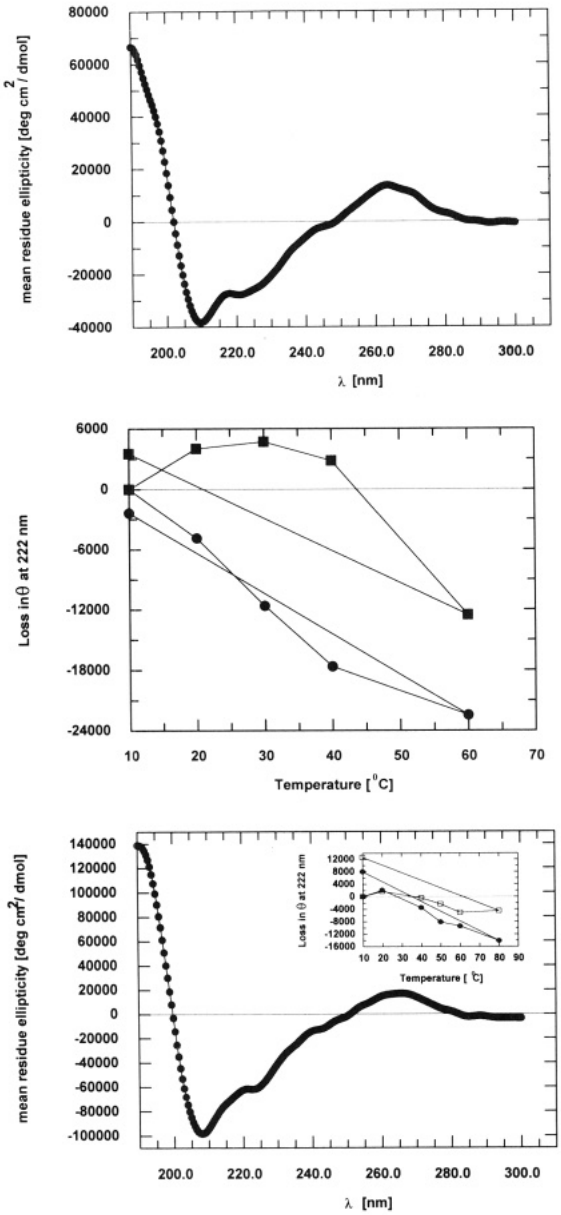

FIGURE 5: (A, top) CD spectrum of the Rev M9Δ14/F8-RNA complex at 1:1 stoichiometry. (B, middle) Comparison of the temperature dependence of the signal intensity of Rev M9Δ14 (●) and the 1:1 complex with F8-RNA (■). (C, bottom) CD spectrum of the Rev M11Δ14/F8-RNA complex at 1:1 stoichiometry. Inset: Temperature dependence of the 222-nm signal intensity of Rev M11Δ14 in the uncomplexed (●) and complexed form (□).

The helical structure is fully stable up to 40 °C. Similar to observations with the unbound protein, refolding of the RNA-bound M11Δ14 through heating and cooling results in a protein structure with enhancement in helical content. Of the 12 624 deg cm^2 dmol^{-1} increase in negative spectral intensity, which is even higher than that for the unbound protein, 30% is due to the free protein and 10% results from RNA structural changes on binding. The remaining 8% must be the result of additional protein stabilization on refolding in the presence of RNA. The M11Δ14 protein difference spectrum (Figure 6B) shows only very minor changes compared to the free protein spectrum, so that the number of residues in secondary structural elements is fairly constant. A comparison between the CD spectrum of F8-RNA (filled circles) and the approximated difference spectrum (M11Δ14/F8 – M11Δ14) for the RNA structure (triangles) is shown in Figure 4B

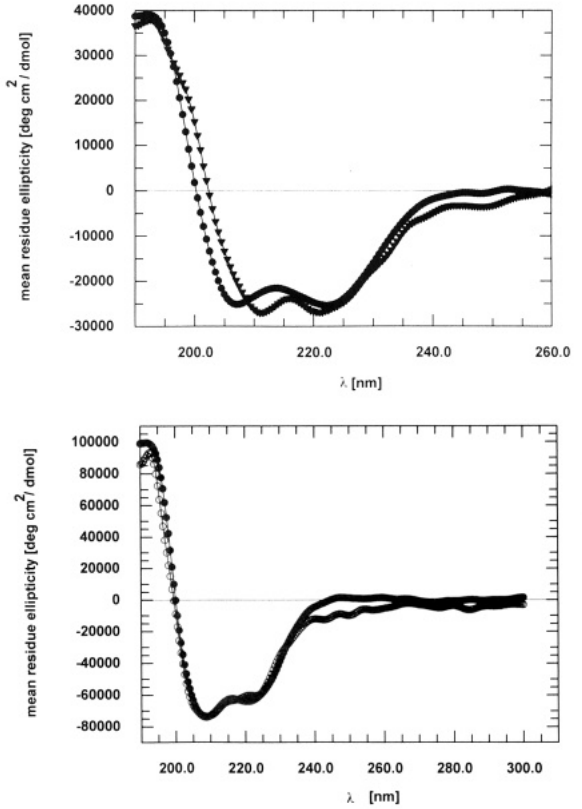

FIGURE 6: (A, top) Comparison between the CD spectrum of Rev M9Δ14 (●) and the difference spectrum between the complex Rev M9Δ14/F8-RNA and the F8-RNA spectrum expressed as mean residue ellipticity per amide bond (▼). (B, bottom) difference spectrum between the complex Rev M11Δ14/F8-RNA and the F8-RNA spectrum (○) represented along with the CD spectrum of Rev M11Δ14 as the mean residue ellipticity per amide bond (●).

demonstrating the reduction of the band intensity at 263 nm and the increase of the negative band at 240 nm.

DISCUSSION

The results presented here provide evidence that in Rev/RRE binding the conformation of RNA plays an important role in recognition by Rev and that a conformational transition in RRE takes place on formation of the complex. Similar findings were reported for Tat/TAR (Tan et al., 1992). A 40 nucleotide long sequence of stem–loop-II (F8-RNA) shows a reduction in the signal intensity at 265 nm and of the negative peak at 240 nm to a slightly larger extent than was detected for the shorter oligoribonucleotides of the TAR fragment, indicating a more pronounced structural transition in RNA.

The CD spectra of free, and F8-RNA-complexed, Rev M9Δ14 and Rev M11Δ14 indicated that between 46 and 49 amino acids are present in α-helical conformation. Thus, the structure of the N-terminal 68 aa of Rev is dominated by helical motifs, and the amount of this secondary structure is kept constant on RNA binding. A change in tertiary structure cannot, however, be excluded.

To identify the location of the helical regions in the N-terminus of Rev, a detailed analysis of the primary sequence was performed. The stretch from aa 1 to 11 is characterized by a cascade of polar residues and is followed by a periodical repeat consisting of eight hydrophobic aa: #(L(12)) #(L(13)) - #(A(15)) #(V(16)) - #(L(18)) #(I(19)) - #(F(21)) #(L(22)). This periodicity can be interpreted as overlapping hydrophobic heptads, for example #(L(12)) #(L(13))--#(V(16))--, #(L(12))--#(A(15)) #(V(16)) --, #(V(16))--#(I(19))---, and #(L(18))--#(F(21))

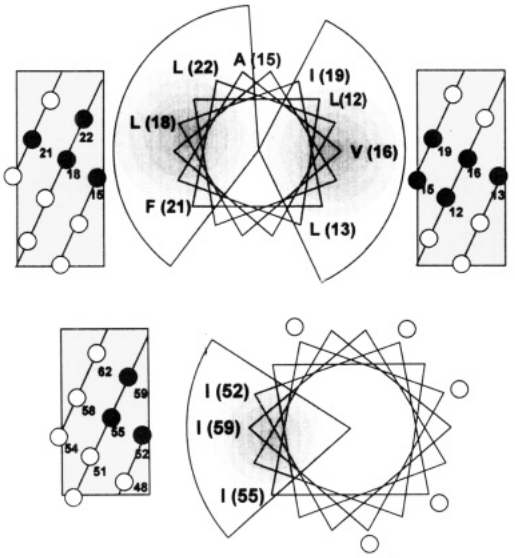

FIGURE 7: (A, top) Helical wheel representation of the periodical repeat of eight hydrophobic amino acids in the N-terminus of Rev. The two clusters (A(15), L(18), F(21), L(22)) and (L(12), L(13), V(16), I(19)) are shaded. A side view of the two helices shows the four hydrophobic residues represented as black dots at each side of the helix. (B, bottom) Representation of the "isoleucine motif" C-terminal of the RNA-binding and nucleolar localization domain as a helical wheel and as a side view. Filled circles are isoleucines; open circles are polar amino acids demonstrating the potential to form an amphiphilic helix.

#(L(22)) - -, indicating for this part of Rev a high helix-forming potential (Lesk, 1991; Gibson et al., 1993). A helical wheel presentation of residues 12–22 generates a cluster of polar aa's (K(14), R(17), K(20)) and two extended hydrophobic regions (L(12), L(13), V(16), I(19) and A(15), L(18), F(21), L(22)) (Figure 7A). V(16), L(18), I(19), and L(22) are fully conserved within all known sequences of HIV-1, HIV-2, and Visna-Rev, whereas aa's L(12), L(13), and F(21) show only slight variability. The two hydrophobic areas point in different directions, suggesting two distinct surfaces for hydrophobic contacts.

The precise N- and C-terminal end of a potential α-helical region in the amino terminus of HIV-1 Rev remains unclear. However, the first six to seven residues are most probably excluded from the helix due to two glycines at positions 3 and 6. The following sequence, S(8)-D(9)-E(10)-D(11), can be included, for example, as starting circle for the proposed highly hydrophobic helical part 12–22. Serine (S(8)) and aspartate (D(9), D(11)) occur with high frequency in the Ncap position of helices (Presta & Rose, 1988; Richardson & Richardson, 1988; Lyu et al., 1992). The same argument can be used for the residues 23–26 (Y Q S N) at the Ccap position. These residues probably form a polar C-terminal end of the helix. In model peptide systems asparagine (N) showed a strong helical-promoting effect at the Ccap position (Zhou et al., 1993). The N-terminal sequence spanning aa 8–26 of the Rev wild-type protein (Rev M11Δ14, Rev M9Δ14, and Rev 8–26) does not contain residues unlikely to be found in helices, so called helix breakers. A proposed helix from aa 8 to 26 is also in good agreement with the CD data on the linear peptide Rev 8–26, which showed about 30% α-helical conformation at 1 mM concentration in aqueous solution. More than 90% α-helix could be induced by small amounts of HFP in diluted samples. These findings are in good agreement with studies on protein folding, where secondary structures, especially helices and reverse turns, have been observed in short linear peptide fragments in aqueous solution. Mostly these peptides populate ordered helical conformations in rapid equilibrium with unfolded states throughout the whole sequence (see for a review Dyson & Wright, 1991). Also, Chou–Fasman (CF) secondary structure predictions suggest a strong helical propensity for the aa stretch 10–22. Furthermore, the existence of a helix at this location in Rev was previously suggested on genetic grounds (Hope et al., 1990b).

A clear "helix-breaking signal" starts at position 27. The sequence P-P-P-N-P-E-G is virtually impossible to be accommodated in an α-helix. It may form a loop which separates the hydrophobic signals in the helical N-terminus from the RRE binding domain which includes residues 35–50. Additional hydrophobic amino acids are clustered C-terminally to the RNA-binding region. Helical wheel representation of residues 52–59 reveals an amphiphilic helix with aa I(52), I(55), and I(59) located to one side (Figure 7B). This "isoleucine motif" represents a hydrophobic heptad with periodicity # - - # - - -. Chou–Fasman calculations predict a weak helical potential for aa 52–55, no prediction for aa 56 and 57, and a strong β-sheet potential for aa 58–64. Several of the known Rev sequences (HIV-1, HIV-2, Visna-Rev) contain glycine at either position 56 or 57. In addition, a Rev protein with aspartate-arginine (DR) at positions 57 and 58, respectively, exchanged by glycine-glycine (GG) is functional *in vivo* (Hope et al., 1990). Therefore, a stop of helix conformation at position 55 has to be taken into consideration. An arrangement of isoleucines as shown in Figure 7B generates a hydrophobic surface which could be indicative of involvement in a helix–helix contact with the usual periodicity # - - # - - - of hydrophobic residues. A potential partner for this hydrophobic signal at the end of the RNA-binding domain is one of the two hydrophobic parts on the proposed N-terminal helix 8–26.

In sum, the residues proposed to be in α-helices (8–26 and 52–55 or 52–59) can only cover 50% or 59% of the helical content in the Rev M9Δ14 mutant protein (at least 46 residues based on the CD measurements). If C-terminally of I(55) or I(59) no further contribution to the helical content of the CD spectrum of Rev M9Δ14 arises, the missing α-helical element must fully be located in the RNA-binding domain. In both cases it must include all 18 residues from aa 34 to 51.

$$[\text{N-term.}]_{\text{random}} - [8\text{–}26]_{\text{helix}} - [\text{PPPNPEG}]_{\text{loop}} - [34\text{–}55(59)]_{\text{helix}} - [\text{C-term.}]_{\text{random}}$$

However, Rev M9Δ14 is a deletion mutant bearing the original Rev(wt) sequence AKE at the C-terminus (see also Figure 1). Chou–Fasman prediction suggests that aa's 66–71 (RDLAKE) are capable of forming a helical structure, whereas the natural sequence 66–70 (SAEPV) shows a high unspecific turn potential. Assuming that the "isoleucine motif" includes I(59) in helical conformation, an additional 13 residues are needed to account for the helical content observed experimentally. The secondary structure of Rev M9Δ14 could look like

$$[\text{N-term.}]_{\text{random}} - [8\text{–}26]_{\text{helix}} - [\text{PPPNPEG}]_{\text{loop}} - [21\text{ aa between 34 and } 55(59)]_{\text{helix}} - [60\text{–}65]_{\text{random}} - [\text{C-term.}]_{\text{helix}}$$

This result is supported by Chou–Fasman calculations, suggesting a weak helical signal from aa 43–51 in Rev. Within all known Rev sequences no helix-breaking signal is present in the RNA-binding domain. Furthermore, in the highly charged part of this sequence, several helix-stabilizing salt- and H-bridges can be found.

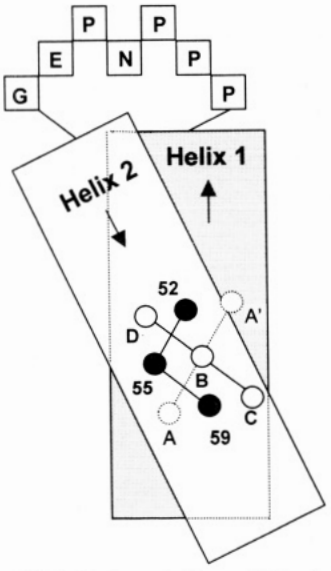

FIGURE 8: Potential helix–loop–helix motif in the N-terminus of Rev. The antiparallel helices 1 and 2 (presented as rectangles) are connected by the proline-rich loop PPPNPEG. A, B, C, D or A′, B, C, D are aa 12, 16, 13, 19 or aa 22, 18, 15, 21, respectively, forming a classical helix–helix contact where the ridges made by residues separated in sequence by three (i + 3) fit into the grooves of residues separated by four (i + 4) and vice versa.

In summary, between 70% and 100% of all residues in the central RNA-binding (aa 34–51) region can potentially adopt a helical conformation (in $helix_2$). Therefore Rev M9Δ14 contains a helix–loop–helix motif. The helix-forming potential of the arginine-rich region of the Rev protein has recently been shown on isolated, N- and C-terminally modified peptides (Tan et al., 1993).

This helix–loop–helix motif in Rev, suggested by our CD results, may be analogous to helix–loop–helix domains of DNA-binding proteins (Murre et al., 1989; Cai & Davis, 1990; Yoshida et al., 1989; Cabrera & Alonso, 1991; Cole, 1986; Evan et al., 1992). Models of the tertiary structure of these regions (Davis et al., 1990; Gibson et al., 1991; Anthony-Cahill et al., 1992; Halazonetis & Kandil, 1992; Vinson & Garcia, 1992; Gibson et al., 1993) suggest the presence of intercalating hydrophobic side chains for stabilization in the helix–helix interaction area (Figure 8). Packing antiparallel α-helices can be accomplished by fitting the ridges of side chains from one helix into the grooves between side chains of the other helix and vice versa (Lesk, 1991; Branden & Tooze, 1991). In a type 1 structure, the ridges formed by one row of residues separated in sequence by four amino acids (i + 4) fit into the same type of grooves in the other helix (i + 4), resulting in an angle of about 50° between the two helices. In our proposed model one of the two hydrophobic motifs in $helix_1$, (either aa 12, 13, 16, and 19, or 15, 18, 21, and 22) (Figure 7A) forms a "ridge and hole" structure with the isoleucine motif (I(52), (55), (59)) in $helix_2$ (Figure 7B). However, a serious steric hindrance arising due to the short loop PPPNPEG between the two helices suggests that an angle of 50° may not be accommodated.

The second packing possibility comprises the amino acids on an i + 3 ridge of one helix fitting between the holes of an i + 4 ridge of the second helix. The resulting angle of 20° can be accommodated by the "proline loop" of Rev and has also been seen in HLH structures from DNA-binding proteins. Not only is this packing motif favorable in the model due to predicted steric considerations but there is also support from mutational data as well. Replacement of isoleucine 55 by serine results in an active mutant protein (Hope et al., 1991). It can therefore be assumed that I(59) is necessarily positioned within $helix_2$ to allow the formation of a knob and hole pattern. If that is the case, an i + 3/i + 4 packing mode between the ridges 13 - 16 - 19, or 15, 18, 21, and 55 - 59 is highly favorable (Figure 8).

In transcription factors, the helix–loop–helix motif is necessary for dimerization, and this dimerization is necessary for DNA binding. Rev has been shown to bind to RRE as a monomer (Daly et al., 1993a; Cook et al., 1991), but a stoichiometry of two Rev molecules per RRE confers kinetic stability to the Rev–RRE binding event (Daly et al., 1993a,c). Besides sequence specificity, destabilization of competing structures, and steric accessibility, this kinetic stability is one of the specificity mitigating features in protein nucleic acid interactions (Weeks & Crothers, 1992). Since the N-terminal 20 amino acids of Rev are necessary to keep the off-rates of Rev from RRE equivalent to those of the wild-type protein ($T_{1/2}$ = 20–25 min) (Daly et al., 1993a,c), it is attractive to speculate that the Rev monomer, together with a second molecule, could form a four-helix bundle using the second hydrophobic contact in $helix_1$. One hydrophobic cluster would be bound intramolecularly in Rev, while the second could be available for intermolecular protein–protein contacts. This interaction may be necessary for higher order Rev–Rev interactions. Clearly, a more detailed modelling study will be necessary to propose the structure of helix–helix contacts and interfaces more precisely. With the ability to produce large quantities of homogeneous RNA, Rev mutants with reduced aggregational behavior, and effector proteins with the potential to stabilize Rev–RRE complexes, there is reason to believe that X-ray analysis or NMR studies can yield high-resolution three-dimensional information.

ACKNOWLEDGMENT

We thank Dr. Roelf Datema for critical reading of the manuscript and continuous support of this work and Othmar Hurt, Andreas Schilk, Axel Leitner, Christine Fischer, and Franz Amberger for excellent technical assistance.

REFERENCES

Anthony-Cahill, S. J., Benfield, P. A., Fairman, R., Waserman, Z. A., Brenner, S. L., Stafford, W. F., III, Altenbach, C., Hubbell, W. L., & Degrado, W. F. (1992) *Science 255*, 979–982.

Arya, S. K., Guo, C., Josephs, S. F., & Wong-Staal, F. (1985) *Science 229*, 69–73.

Bevec, D., Dobrovnik, M., Hauber, J., & Boehnlein, E. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9870–9874.

Branden, C., & Tooze, J. (1991) in *Introduction to Protein Structure*, Garland Publishing, Inc., New York and London.

Cabrera, C. V., & Alonso, M. C. (1991) *EMBO J. 10*, 2965–2973.

Cai, M., & Davis, R. W. (1990) *Cell 61*, 437–446.

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

Chou, P. Y., & Fasman, G. D. (1974) *Biochemistry 13*, 222–244.

Cochrane, A. W., Chen, C.-H., & Rosen, C. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1198–1202.

Cole, M. D. (1986) *Annu. Rev. Genet. 20*, 361–384.

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., Cook, K. S., Gray, G. S., Maione, T. E., & Rusche, J. R. (1989) *Nature 342*, 816–819.

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

Daly, T. J., Doten, R. D., Rennert, P., Auer, M., Jaksche, H., Donner, A., Fisk, G., & Rusche, J. R. (1993a) *Biochemistry 32*, 10497–10505.

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

Davis, R. L., Cheng, P.-F., Lassar, A. B., & Weintraub, H. (1990) *Cell 60*, 733–746.

Dayton, E. T., Powell, D. M., & Dayton, A. I. (1989) *Science 246*, 1625–1629.

Dyson, H. J., & Wright, P. E. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 519–538.

Dyson, H. J., Sayre, J. R., Merutka, G., Stein, H. C., Lerner, R. A., & Wright, P. E. (1992) *J. Mol. Biol. 226*, 819–835.

Evan, G. I., Wyllie, A. H., Gilbert, C. S., Littlewood, T. D., Land, H., Brooks, M., Waters, C. M., Penn, L. Z., & Hancock, D. C. (1992) *Cell 69*, 119–128.

Feinberg, M. B., Jarrett, J. F., Aldovini, A., Gallo, R. C., & Wong-Staal, F. (1986) *Cell 46*, 807–817.

Frankel, A. D., Bredt, D. S., & Pabo, C. O. (1988) *Science 240*, 70–73.

Gibson, T. J., Sibbald, P. R., & Rice, P. (1991) *J. DNA Seq. Map. 1*, 213–215.

Gibson, T. J., Thompson, J. D., & Abagayan, R. A. (1993) *Protein Eng. 6*, 41–50.

Halazonetis, T. D., & Kandil, A. N. (1992) *Science 255*, 464–466.

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

Hope, T. J., Huang, X., McDonald, D., & Parslow, T. G. (1990a) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7787–7791.

Hope, T. J., McDonald, D., Huang, X., Low, J., & Parslow, T. G. (1990b) *J. Virol. 64*, 5360–5366.

Hope, T. J., Bond, B. L., McDonald, D., Klein, N. P., & Parslow, T. G. (1991) *J. Virol. 65*, 6001–6007.

Johnson, W. C., Jr. (1985) in *Methods of Biochemical Analysis 31* (Glick, D., Ed.) pp 61–163, Wiley, New York.

Karn, J., Dingwall, C., Gait, M. J., Heaphy, S., & Skinner, M. A. (1991) in *Nucleic Acids and Molecular Biology 5* (Eckstein, F., & Lilley, D. M. J., Ed.) pp 194–218, Springer Verlag, Berlin, Heidelberg.

Lazinski, D., Gradzielska, E., & Das, A. (1989) *Cell 59*, 207–218.

Le, S.-Y., Malim, M. H., Cullen, B. R., & Maizel, J. V. (1990) *Nucleic Acids Res. 18*, 1613–1623.

Lesk, A. M. (1991) in *Protein Architecture, a Practical Approach*, IRL Press, Oxford, U.K.

Loret, E. P., Vioes, E., Ho, P. S., Rochat, H., Van Rietschoten, J., & Johnson, W. C., Jr. (1991) *Biochemistry 30*, 6013–6023.

Loret, E. P., Georgel, P., Johnson, W. C., Jr., & Ho, P. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9734–9738.

Lyu, P. C., Zhou, H. X., Jelveh, N., Wemmer, D. E., & Kallenbach, N. R. (1992) *J. Am. Chem. Soc. 114*, 6560–6562.

Malim, M. H., Böhnlein, S., Hauber, J., & Cullen, B. R. (1989) *Cell 58*, 205–214.

Malim, M. H., Tiley, L. S., McCarn, D. F., Rusche, J. R., Hauber, J., & Cullen, B. R. (1990) *Cell 60*, 675–683.

Malim, M. H., McCarn, D. F., Tiley, L. S., & Cullen, B. R. (1991a) *J. Virol. 65*, 4248–4254.

Malim, M. H., & Cullen, B. R. (1991b) *Cell 65*, 241–248.

Malim, M. H., Freimuth, W. W., Liu, J., Boyle, T. J., Lyerly, K., Cullen, B. R., & Nabel, G. J. (1992) *J. Exp. Med. 176*, 1197–1201.

Mermer, B., Felber, B. K., Campbell, M., & Pavlakis, G. N. (1990) *Nucleic Acids Res. 18*, 2037–2044.

Murre, C., Shouleber-McCaw, P., & Baltimore, D. (1989) *Cell 56*, 777–783.

Nalin, C. M., Purcell, R. D., Antelman, D., Mueller, D., Tomchak, L., Wegrzynski, B., McCorney, E., Toome, V., Kramer, R., & Hsu, M.-C. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7593–7597.

Ogilvie, K. K., Usman, N., Nicoghosian, K., & Cedergren, R. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5764–5768.

Olsen, H. S., Nelböck, P., Cochrane, A. W., & Rosen, C. A. (1990) *Science 247*, 845–848.

Presta, L. G., & Rose, G. D. (1988) *Science 240*, 1632–1640.

Provencher, S. W., & Glöckner, J. (1981) *Biochemistry 20*, 33–37.

Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652.

Scaringe, S. A., Francklyn, C., & Usman, N. (1990) *Nucleic Acids Res. 18*, 5433–5441.

Shin, H. C., Merutka, G., Waltho, J. P., Wright, P. E., & Dyson, H. J. (1993) *Biochemistry 32*, 6348–6355.

Slice, L. W., Codner, E., Antelman, D., Holly, M., Wegrzynski, B., Wang, J., Toome, V., Hsu, M.-C., & Nalin, C. M. (1992) *Biochemistry 31*, 12062–12068.

Sodroski, J., Goh, W. C., Rosen, C., Dayton, A., Terwilliger, E., & Haseltine, W. (1986) *Nature 321*, 412–417.

Soennichsen, F. D., Van Eyk, J. E., Hodges, R. S., & Sykes, B. D. (1992) *Biochemistry 31*, 8790–8798.

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.

Terwilliger, E., Burghoff, R., Sia, R., Sodroski, J., Haseltine, W., & Rosen, C. (1988) *J. Virol. 62*, 655–658.

Usman, N., Ogilvie, K. K., Jiang, M. Y., & Cedergren, R. J. (1987) *J. Am. Chem. Soc. 109*, 7845–7854.

Venkatesh, L. K., & Chinnadurai, G. (1990) *Virology 178*, 327–330.

Vinson, C. R., & Garcia, K. C. (1992) *New Biol. 4*, 396–403.

Waltho, J. P., Feher, V. A., Merutka, G., Dyson, H. J., & Wright, P. E. (1993) *Biochemistry 32*, 6337–6347.

Weeks, K. M., & Crothers, D. M. (1992) *Biochemistry 31*, 10281–10287.

Weichselbraun, I., Farrington, G. K., Rusche, J. R., Böhnlein, E., & Hauber, J. (1992) *J. Virol. 66*, 2583–2587.

Wingfield, T. P., Stahl, S. J., Payton, M. A., Venkatesan, S., Misra, M., & Steven, A. A. C. (1991) *Biochemistry 30*, 7527–7534.

Wu, T., Ogilvie, K. K., & Pon, R. T. (1988) *Nucleic Acids Res. 17*, 3501–3517.

Yoshida, K., Kuromitsu, Z., Ogawa, N., & Oshima, Y. (1989) *MGG, Mol. Gen. Genet. 217*, 31–39.

Zapp, M. L., & Green, M. R. (1989) *Nature 342*, 714–716.

Zapp, M. L., Hope, T. J., Parslow, T. G., & Green, M. R. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 7734–7738.

Zhou, N. E., Zhu, B.-Y., Sykes, B. D., & Hodges, R. S. (1992) *J. Am. Chem. Soc. 114*, 4320–4326.

Zhou, H. X., Lyu, P., Wemmer, D. E., & Kallenbach, E. R. (1993) *J. Cell. Biol. Suppl. 17C*, 222.

Zuker, M. (1989) *Science 244*, 48–52.