Probing the details of the HIV-1 Rev–Rev-responsive element interaction: effects of modified nucleotides on protein affinity and conformational changes during complex formation

Suzanne B. RENWICK,*‡, Andrew D. CRITCHLEY,*§, Chris J. ADAMS,* Sharon M. KELLY,†, Nicholas C. PRICE† and Peter G. STOCKLEY*||

*Department of Genetics, University of Leeds, Leeds LS2 9JT, U.K., and †Department of Biological and Molecular Sciences, University of Stirling, Stirling FK9 4LA, U.K.

The solution structure of the human immunodeficiency virus type 1 (HIV-1) Rev-responsive element (RRE) has been investigated by enzymic and chemical structural probing of a 71 nt RRE transcript. The minimum sequence information required to maintain recognition by the Rev protein has previously been mapped to a 29 nt stem-loop structure, known as minSLIIB. The key recognition target is a single-stranded RNA bubble at the base of the RNA stem. The fine details of RNA recognition have been probed using chemically synthesized minSLIIBs containing variant base or sugar residues at sites within the bubble. These have been analysed by gel retardation assays and their relative affinities for Rev protein determined. Complex formation between the wild-type minSLIIB RRE and Rev protein was also

monitored using CD spectroscopy, which suggests a change in RNA conformation upon Rev binding. The spectral change is consistent with localized melting of RNA, leading to a decrease in the level of base stacking and/or a change in base tilting, during formation of the complex. Deoxynucleotide substitutions on just one side, the 5' side, of the bubble inhibit the conformational change detected by CD. The data are consistent with a dynamic interaction between Rev and its target site. The contact points between Rev and the RRE were probed directly using photo-cross-linking with either ribo-5-bromouridine- or ribo-4-thiouridine-substituted minSLIIBs. The data are consistent with protein–RNA contacts at the bottom of the bubble.

INTRODUCTION

Replication of human immunodeficiency virus type 1 (HIV-1) depends upon the expression of Rev protein [1]. The interaction between this 116 amino acid protein and its RNA target site, the Rev-responsive element (RRE), leads to the cytoplasmic expression of incompletely spliced viral mRNAs [2,3]. The precise mechanism of Rev action is as yet undetermined, although two separate mechanisms have been proposed. According to one mechanism, Rev is involved in the transport of incompletely spliced HIV-1 mRNAs from the nucleus to the cytoplasm [2]. The alternative proposal suggests that binding of Rev to RRE blocks formation of the spliceosome thereby inhibiting splicing and allowing it to be transported to the cytoplasm [4]. Rev has also been shown to be involved in the efficient translation of RRE-containing mRNAs [5].

The RRE was initially defined as a complex 234 nt structure located within the *env* gene of HIV-1 [6]. Mutational analysis identified a region within RRE, termed stem-loop II, as being sufficient to bind Rev *in vitro* [7]. Further work utilizing chemical interference studies has shown that a synthetic 29 nt RNA encompassing the base of the helix of stem IIB acts as a minimal, high-affinity Rev-binding site [8]. The minimal RRE site (minSLIIB) shows identical diethylpyrocarbonate (DEPC) interference patterns with larger RRE sites, such as the 66 nt SLII, although it apparently has a lower affinity as judged by competition experiments using gel retardation assays. *In vivo*, Rev function requires binding of multiple copies of Rev protein to the extended RRE site [9], SLII showing only partial activity [7].

Such multimerization of Rev on the RRE would now appear to be the result of sequential binding of Rev monomers to existing RNA–protein complexes [8,10], rather than the binding of multimeric Rev to the RRE as previously suggested [11,12].

Here we report the use of chemically synthesized variants of the minSLIIB, which have enabled the fine details of the Rev–RRE interaction to be studied by gel retardation assays, CD spectroscopy and photo-cross-linking.

MATERIALS AND METHODS

Transcription and radiolabelling of RNAs

Run-off transcription reactions were performed *in vitro* on plasmid pGEM-SLII linearized with *Hinc*II. All transcription reactions, 5' radiolabelling and purification of RNA transcripts were performed as described previously [13].

Radiolabelling at the 3'-end was performed as described [14], using T4 RNA ligase (BRL), 50–100 pmole RNA and 50 μ Ci of 5'-[32 P]pCp (New England Nuclear). All labelling reactions were stopped by the addition of 10 μ l of denaturing loading buffer (10 M urea, 2 mM EDTA, 0.05 % (w/v) xylene cyanol and Bromophenol Blue).

RNA sequencing

The sequences of radioactively labelled RNAs were determined by the sequence-specific enzymic digestion procedure [15].

Abbreviations used: RRE, Rev-responsive element; r5BrU, ribo-5-bromouridine; r4SU, ribo-4-thiouridine; HIV-1, human immunodeficiency virus type 1; DEPC, diethylpyrocarbonate; ENU, ethylnitrosourea; MPE, methidiumpropyl-EDTA; DTT, dithiothreitol; DMT, 2'-silyl-5'-dimethoxytrityl; GST, glutathione-S-transferase; SPR, surface plasmon resonance.

Present addresses: ‡Structural Biology Section, Institute of Cancer Research, Cotswold Road, Sutton, Surrey SM2 5NG, U.K., and §Department of Human Genetics, Molecular Genetics Unit, Ridley Building, University of Newcastle-upon-Tyne, Claremont Place, Newcastle-upon-Tyne NE1 7RU, U.K.

|| To whom correspondence should be addressed.

Chemical and enzymic structural analysis

Modification of $\approx 50\,000$ c.p.m. of 3' radiolabelled RNA with DEPC (Sigma) under native, semi-denaturing and denaturing conditions was performed as described [13].

5' or 3' [^{32}P]-radiolabelled RNAs ($\approx 50\,000$ – $100\,000$ c.p.m.) were alkylated [16] by the addition of 5 μl of a freshly saturated ethanolic solution (≈ 750 mM) of ethylnitrosourea (ENU) to 20 μl of the appropriate reaction buffer. Modification under native conditions was in a buffer containing 300 mM sodium cacodylate (pH 8.0), 20 mM MgCl_2 , 100 mM KCl and 2 mM EDTA, at 20 °C for 3 h. Modification under denaturing conditions was in 300 mM sodium cacodylate, pH 8.0, 2 mM EDTA, at 80 °C for 2 min. Following alkylation, 10 μl of tRNA and 6 μl of 2.5 M sodium acetate, pH 6.0, were added and the RNA precipitated with ethanol, twice. Phosphotriester groups were cleaved by dissolving the pellets in 10 μl of 0.1 M Tris/HCl, pH 9.0, and incubating at 50 °C for 5 min, followed by re-precipitation with ethanol. Pellets were then resuspended in denaturing loading buffer.

Methidiumpropyl-EDTA (MPE) was a gift from Professor P. Dervan (Caltech). The reagent [17] was activated by adding 10 μl of a freshly prepared 6 mM $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$ solution to 50 μl of MPE (1 mg/ml in water) and the solution diluted to a final MPE concentration of 20 μM . Radiolabelled RNA ($\approx 50\,000$ c.p.m.) was resuspended in 8 μl of buffer and the reaction initiated by the addition of 1 μl of the MPE reagent and 1 μl of dithiothreitol (DTT). Cleavage under native conditions was in 50 mM Tris-HCl (pH 8.0)/100 mM KCl/1 mM MgCl_2 , at 20 °C for 60 min and was terminated by the addition of 10 μl of denaturing loading buffer.

Aliquots of either 5'- or 3'-[^{32}P]RNA were digested with either RNase V1 (Pharmacia), which cleaves 5' to double-stranded or structured bases, RNase T2 (BRL) (3' single-strand specific) and RNase T1 (BRL) (3' single-strand G specific) at 24 °C for 30 min. Reactions were performed as described in [13]. Enzymic sequencing reactions and formamide ladders were electrophoresed in parallel with the experimental samples.

Chemical synthesis, purification and characterization of oligoribonucleotides

Chemical synthesis of oligoribonucleotides was performed by the solid-phase method using 2'-silyl-5'-dimethoxytrityl (DMT) phosphoramidites [18,19]. All syntheses were performed on a 1 μmole scale, using an Applied Biosystems 391PCR-Mate DNA synthesizer. The DMT-ribonucleotide-phosphoramidites were synthesized as described [20,21] or purchased from ChemGenes Corporation (MA, U.S.A.) or Glen Research Corporation (VA, U.S.A.).

Initial purification of the RNA was conducted by reverse-phase HPLC on a C_{18} 5 μm LiChrosphere (RP18) column (4 $\times$ 250 mm) (Merck). The RNA was eluted with a linear gradient of 2–10% (v/v) acetonitrile buffered with 100 mM ammonium acetate [22]. Peaks corresponding to full length and fully deprotected fragments were pooled and characterized by the sequence-specific RNase method [15,19] and in some cases by complete base composition analysis as described [22]. The synthetic RNAs for gel retardation assays were then radiolabelled at the 5'-end and gel purified as described above. Thermal melting profiles were recorded using a Perkin-Elmer Lambda II Spectrophotometer attached to a Peltier temperature programmer (PTP-1). The CD spectrum of each of the variant RNAs used was identical between 240 and 320 nm (see below). The presence of the thio grouping in the ribo-4-thiouridine (r4SU)

derivatives was confirmed by the characteristic absorbance band between 320 and 350 nm.

Purification of Rev

Purification of Rev, expressed in *Escherichia coli* as a glutathione-S-transferase fusion protein (GST-Rev), was performed as described [10].

Gel retardation assays

Gel retardation assays between GST-Rev and minSLIIBs were carried out based on the protocol of Malim et al. [7]. Various concentrations of GST-Rev (up to 32.4 μM) were incubated with radiolabelled RRE (≈ 7500 cpm; 2.5–8 nM) for various times on ice. The samples were then electrophoresed in 10% (w/v) [29:1 (w/w) acrylamide to bisacrylamide] non-denaturing polyacrylamide gels using 50 mM Tris-HCl/50 mM glycine (pH 8.8) as the running buffer. The gels were dried under vacuum and the Rev-RRE complexes visualized by autoradiography. The relative intensities of bands on autoradiograms, corresponding to the Rev-RRE complex, were determined by densitometry. Binding assays were carried out across a range of protein concentrations such that the amount of retarded complex became saturated (not necessarily when 100% of input RNA had shifted). It was assumed that the intensity of bands at the plateau concentrations corresponded to 100% saturation of complex formation. The percentage saturation for each of the other protein concentrations was calculated relative to this figure.

CD measurements

CD spectra were measured using a JASCO model J-600 spectropolarimeter. RNA concentrations were 50 $\mu\text{g/ml}$ ($= 5.2$ μM). Samples were prepared in gel retardation assay buffer, omitting the BSA, DTT, yeast tRNA and RNA guard. Spectra were recorded at 20 °C from 320 to 240 nm using a 0.5 cm path-length cuvette. When GST-Rev (up to 31.2 μM) was added to RNA samples, the small contribution to the CD made by the protein was corrected for by running appropriate blanks. Molar ellipticity values (Θ) were measured using a value of 9570 for the M_r of minSLIIB.

UV cross-linking studies

Samples of RNA and protein were prepared as for gel retardation assay and allowed to equilibrate on ice for 30 min before being split into two aliquots, the first of which was used in a standard gel retardation assay to confirm that complexes were present. The second aliquot was irradiated at a distance of 1 cm for 5–60 min, with short wavelength (254 nm) UV light for the ribo-5-bromouridine (r5BrU)-containing samples and with long wavelength (266 nm) UV light for r4SU-containing versions with a Model UVGL-25 Mineralight (Ultraviolet Products Ltd., Cambridge, U.K.). After irradiation, SDS was added (final concentration 4% w/v) and the samples electrophoresed on a gel identical with those used for retardation assays but containing 0.1% (w/v) SDS.

RESULTS

It has been reported that complex formation between Rev and RRE leads to conformational changes within the RRE site, allowing additional molecules of Rev to bind [23]. Tiley et al [8]

Figure 1 Results obtained for enzymic and chemical probes

(a) Enzymic structural probing of the RRE. Sequencing gel showing the results of enzymic digestions of a 5'-radiolabelled SP6 transcript of the SLII fragment. Lanes on the left are the sequence-specific enzymic sequencing ladder [15]. F, Formamide ladder; I, incubation control; V1, RNase V1 (0.005, 0.01 and 0.1 units respectively); T2, RNase T2 (0.1, 0.5 and 1.0 units respectively); and T1, RNase T1 (0.001, 0.01 and 0.1 units respectively). (b and c) Chemical structural probing of the RRE. (b) DEPC modification [24] of 3'-radiolabelled transcripts encompassing the SLII site. Lanes are: I, incubation control; D, denaturing conditions (90 °C); S, semi-denaturing conditions; N, native conditions. (c) ENU modification of 3'-radiolabelled transcripts encompassing the SLII site under native (lanes 1 and 2) or denaturing (lanes 3 and 4) conditions. Odd numbered lanes were incubation controls. RNase T1 (EG) and formamide (F) ladders are shown on the left. (d) Summary of the structure probing data. The secondary structure of the RRE SLII proposed by Malim et al. [7]. Symbols and their number represent the reactivity of the various sites towards the structural probes: ●, RNase T1; ○, RNase T2; □, RNase V1; and △, DEPC (under native conditions). Regions with reduced accessibility towards ENU under native conditions are indicated by broken lines; the solid line indicates the region of SLII not probed, except for A98, during the experiments described here.

Table 1 Affinities of the minSLIIB variants for Rev and the results of CD titrations

The affinity of each minSLIIB variant was determined as described in the text. The values quoted for 50% complex saturation are the means of three experiments, except for the (GC) variant where $n = 1$. nd, not determined; + = 30–40% decrease in Θ_{265} ; – = < 5% decrease in Θ_{265} .

Variant	GST–Rev concn (μ M) at 50% saturation	Relative affinity (minSLIIB = 1)	CD result
minSLIIB	2.0 $\pm$ 0.18	1	+
minSLIIB(GC)	2.2	0.91	nd
r5BrU45	2.3 $\pm$ 0.23	0.87	–
r4SU45	2.2 $\pm$ 0.48	0.91	nd
dG46	3.3 $\pm$ 0.66	0.61	–
dG47	0.8 $\pm$ 0.06	2.50	–
dG48	1.8 $\pm$ 0.25	1.11	–
r5BrU60	1.9 $\pm$ 0.23	0.87	+
r4SU60	2.1 $\pm$ 0.41	0.95	nd
dG70	5.4 $\pm$ 0.49	0.37	+
dG71	2.5 $\pm$ 0.50	0.80	+
dU72	2.0 $\pm$ 0.30	1.0	+
r4SU72	2.2 $\pm$ 0.26	0.91	nd
dT72	1.5 $\pm$ 0.05	1.33	+
dA73	1.5 $\pm$ 0.13	1.33	+
dI73	No complex	–	+
d7deazaA73	2.8 $\pm$ 0.03	0.71	nd

showed that, *in vitro*, there is a single, high-affinity site for Rev binding, the minSLIIB, which is also recognized by the M4 mutant of Rev (Asp-23 to Tyr), which is unable to multimerize. Previous work in this laboratory has used a series of RNA structure probes to study the solution structure of the HIV-1 TAR stem-loop [13]. We have applied these techniques to study the solution structure of RNA transcripts encompassing the entire RRE SLII region in order to examine the conformation of the initial recognition target. Examples of the results for both enzymic and chemical probes are shown in Figures 1(a), 1(b) and 1(c) and are summarized in Figure 1(c). The cleavages by both the structured/base-paired specific and the single-strand specific enzymes are consistent with the bulk of the secondary structure shown in Figure 1(c). Experiments carried out at several temperatures suggest that the structure does not alter significantly. There are detectable cleavages by RNase T1 within the bubble region (G46–G48 and G70–G71) but they are very weak. The interpretation of these cleavages is further complicated by the results of phosphate ethylation under ‘native’ or denaturing conditions. Densitometry of the gels shown in Figure 1(c) suggests that there is some tertiary structure leading to relative protection of the phosphate backbone in the regions around the bubble (Figure 1d). Such tertiary structure might well inhibit RNase T1 digestion of otherwise single-stranded guanosine residues. However, cleavage with the intercalating MPE probe [17] occurs throughout the bubble region whereas cleavage is noticeably reduced around the single-stranded loop regions. This is consistent with the bubble being highly structured.

Previously we have shown that DEPC modification can be a sensitive indicator of adenines in sites of intercalation or non-Watson–Crick base-pairs [24]. Adenines expected to be single-stranded in the secondary structure, shown in Figure 1(d), are indeed modified by DEPC under native conditions with A68 and A58 somewhat hyper-reactive compared with A88 and A85. This may well indicate that A68 intercalates into the base-paired stem, although this is probably not important for recognition by Rev

since DEPC modification of this residue is not deleterious for Rev-binding [8]. The bubble adenine at A73 is only weakly reactive under native conditions, indicative of some accessibility to the N7 position. In contrast, A75 is unreactive under identical conditions but weakly accessible under semi-denaturing conditions, consistent with the formation of the U45–A75 base-pair.

In order to probe the details of Rev binding we then produced a series of minSLIIBs substituted at defined sites with either deoxynucleotides, variant bases or both (Table 1 and Figure 2a). Gel retardation assays were used to estimate the relative affinities of the variants for GST–Rev under identical conditions. A single retarded species (c1) was the major product when complexes were allowed to form over a period of 15 min, but multiple species were formed after longer incubations and could be visualized upon longer exposures as shown in Figure 2(c). Protein–protein cross-linking experiments with the reagent dimethyl suberimide and GST–Rev at higher concentrations (1 mg/ml) suggested that the fusion protein did contain some oligomers, possibly due to dimerization via the GST domain (results not shown, see Discussion). Competition experiments were used to show that the retarded complexes formed under all conditions were sequence-specific (Figure 2c). The concentration of fusion protein required to produce 50% saturation of the c1 complex was taken as a rough estimate of the affinity (Figure 2d). Addition of an extra G–C base-pair at the end of the bubble had a negligible effect on the affinity, confirming the observation that the 29mer is the minimal RRE site. Deoxyribose substitution at specific sites was silent (dG48, dG71 and dU72), slightly deleterious (dG46, dG70) or beneficial (dG47, dA73). However, DEPC-modification of these residues led to decreased Rev binding [8], suggesting that they are all important for recognition by Rev. Substitutions of r5BrU residues at U-45 or U-60 and r4SU residues at U45, U60 and U72 were also essentially silent.

The importance of the base functional groups for recognition is illustrated by the variant dI73, which failed to form a detectable retarded species, although the dA73 variant bound more tightly than the wild type. Substitution with deoxy-7-deaza-adenosine (d7deazaA73) was also deleterious whilst incorporation of dT72 was beneficial. Incorporation of differing functional groups could lead to gross distortion of the minSLIIB conformation. The variants were therefore examined by UV thermal-melting profiles, which in all cases, including the dI73 derivative, gave single unfolding transitions with mid-point of thermal melting (T_m) values similar to the wild-type minSLIIB ($T_m = 82^\circ\text{C}$).

CD spectroscopy has previously been used to monitor complex formation between Rev and the entire RRE site [23]. These authors reported a small (approx. 10%) decrease in the intensity of the CD peak at 265 nm on addition of Rev, which was proposed to represent localized melting of the double-stranded structure. This conformational change may facilitate the binding of additional Rev molecules or cellular factors. In order to determine whether the changes in the CD spectrum were due to conformational changes taking place within the minimal high-affinity RRE site we carried out similar experiments with the minSLIIB variants. The results are shown in Figures 3 and 4. The CD spectra of wild-type minSLIIB (Figure 3a) and all the variants tested were typical of double-stranded RNA with a positive peak at 265 nm and a smaller negative peak at 295 nm [25]. On addition of GST–Rev to the wild-type minSLIIB there was a marked (30%) decrease in the CD peak at 265 nm. The change was proportionately greater than that observed by Daly et al. [23] because we were using a much smaller RNA fragment. The change in CD appeared to saturate at a molar ratio of roughly one GST–Rev to one minSLIIB (Figure 4).

Table 1 summarizes the results of similar CD titrations for the

Figure 3 CD spectra of (a) minSLIIB showing the decrease in intensity at 265 nm on the addition of GST-Rev and (b) dG46-minSLIIB which shows a much reduced effect

Molar ratios of RNA:GST-Rev for both (a) and (b) were: RNA alone (solid line); 1:1 ratio (broken line). RNA concentrations were $\approx 50 \mu\text{g/ml}$ ($= 5.2 \mu\text{M}$).

minSLIIB variants. CD spectra and thermal melting data on the variants suggested that all the samples had similar unliganded three-dimensional conformations. However, the deoxyribose variants on the 5' leg of the bubble showed, if any, only a very small (approx. 5%) decrease in the absorption band at 265 nm on addition of GST-Rev (Figures 3b and 4), although they all formed complexes as determined by the gel retardation assay. In contrast, the dI73 variant showed the larger CD effect, with an apparent molar saturation of 1:1, although it did not form a retarded complex on gels. Presumably, this difference reflects the relative stabilities of the complexes formed, complexes with short half-lives dissociating during electrophoresis into gels.

In an attempt to localize the protein-RNA interactions in the complex we used the r5BrU- and r4SU-containing minSLIIBs in photo-cross-linking experiments. No cross-linked products were observed with either of the r5BrU derivatives (results not shown). However, the r4SU derivatives were more successful, with the 4-

Figure 2 Probing the details of Rev binding

(a) Secondary structure of the minimal RRE (minSLIIB) [8], indicating the variant sequences synthesized; Δ indicates positions of deleted nucleotides relative to the wild-type sequence. (b and c) Gel retardation analysis of the minSLIIB. (b) Lane 1, RNA control; lanes 2-11, a constant amount (3 nM) of [^{32}P]-labelled RNA was titrated with decreasing concentrations of GST-Rev and complexes were allowed to form for 15 min. The lanes contained the following concentrations of protein (in μM): lane 2, 10.8; lane 3, 8.1; lane 4, 5.4; lane 5, 4.0; lane 6, 2.7; lane 7, 2.0; lane 8, 1.4; lane 9, 1.0; lane 10, 0.7; lane 11, 0.5. (c) Competition assay with minSLIIB. A constant amount of GST-Rev (2.7 μM in each sample) was incubated with unlabelled minRRE, tRNA or 5S RNA, before the addition of [^{32}P]-labelled minRRE ($\approx 3 \text{ nM}$) and the incubation continued for 30 min. Lane 1, RNA control; lanes 2-7 increasing amounts of unlabelled minRRE (0.2-2.7 μg); lanes 8 and 9, 1 μg and 1.5 μg of tRNA respectively; lanes 10 and 11, 500 ng and 1 μg of 5S RNA respectively. (d) Binding curves of the minSLIIB ($\bullet$) and the dG47-minSLIIB variant ($\blacktriangle$).

Figure 4 Complex formation between Rev and the RRE site

Relative intensity of the ellipticity at 265 nm upon addition of GST-Rev to dG71-minSLIIB (○) and dG46-minSLIIB (□).

Figure 5 UV illumination used to generate a cross-link between GST-Rev protein (5 μ M) and radiolabelled r4SU45 minSLIIB RRE ($\approx$ 5 nM)

SDS (final concentration 4% w/v) was added to all samples at the end of the incubation. Lane 1, control (no illumination); lanes 2–10, 5, 10, 15, 20, 25, 30, 40, 50 and 60 min of illumination respectively; lane 11, free RNA.

SU45 variant leading to the apparent formation of a cross-linked product after 5 min of illumination (Figure 5). No increase in the extent of cross-linking was seen after 25 min of illumination. There were no such products for the r4SU60 or r4SU72 variants, although in non-denaturing gels complexes were clearly present. The results suggest that there are strong protein–RNA contacts at the base of the bubble.

DISCUSSION

The minimal, high-affinity Rev-binding site previously identified [8] is contained within the much larger RRE site [2]. RRE has the potential to fold into extensive secondary structure elements (as shown by the CD spectrum, which is consistent with a high degree of base-pairing); these elements presumably interact to generate a defined tertiary structure. It is within this context that the minimal RRE site is recognized by Rev. The solution structure probing of transcripts encompassing just the SLII region has largely confirmed the secondary structure originally suggested by Malim et al. [2]. Unfortunately, the structure probing enzymes did not cleave readily in the bubble region. Comparison of the extent of phosphate ethylation in buffers, allowing tertiary interactions or designed to prevent such interactions, suggest that even within the SLII fragment, tertiary interactions lead to

relative shielding of the bubble region. It is not possible to carry out enzymic cleavages of the transcripts under the equivalent of 'denaturing conditions', so the weak cleavages by RNase T1 within the bubble could be due either to breathing of a structured region or to cleavage of single-stranded residues, which are only moderately accessible to the bulky enzyme. A number of authors [26,27] have proposed several non-Watson–Crick base-pair interactions in this region (G48:G71 and G47:A73). The enzymic structure-probing data described here provide little, if any, definitive information on such interactions. However, the bubble region is clearly recognized by the intercalating agent MPE, although not as well as the Watson–Crick base-paired stems on either side, consistent with a structure which maintains base-stacking interactions. Solution structure probing data for RRE or various sub-fragments have been reported previously [28]. The data presented here are largely in agreement with the published data.

Previously, Tiley et al. [8], showed that the minimal RRE site could be substituted by a chemically synthesized ribo-oligonucleotide 29 nt long, encompassing the region U45 to A75 with deletion of the A56–U62 base-pair. (Even with this minimal fragment, multiple retarded species were detected by the gel retardation assay, although it was not clear whether species above c1 were due to multimerization with GST-Rev protein alone or GST-Rev–minSLIIB complexes.) We therefore prepared a series of such minSLIIBs having modified sugar or base positions in order to probe sites of possible RNA–protein interaction. Similar experiments with the RNA bacteriophage MS2 coat protein–translational operator complex [29] have shown that putative contacts to RNA in solution identified in this way correlate very well with the major sequence-specific contacts seen in the crystal structure [30]. For the MS2 operator, deoxyribose substitution of sites can be silent [19,29]. This was not the case here, deoxyribose substitution resulting in a range of effects from deleterious (e.g. dG70) to beneficial (dG47). In the MS2 operator complex, deletion of direct hydrogen bonds between protein side-chains and base functional groups leads to decreases in affinity of at least 10-fold. The magnitude of the effects seen here are much smaller and so these effects could be due either to a loss of a direct interaction between the protein and, for instance, the 2' hydroxyl group or to an effect of sugar substitution on the conformation of the RNA backbone [31].

The most dramatic effect of substitution on Rev affinity occurred for the dI73 variant, dA substitution at this position being mildly beneficial. No complex could be detected by gel retardation of the inosine variant, although CD spectra suggested that a complex does form with an affinity similar to that for the wild-type minSLIIB. The CD experiments do not include competitor RNA, which is present in the gel retardation assay, suggesting that the CD effect could be due to non-specific binding. However, there was no effect on the CD spectrum of the MS2 translational operator [30] upon addition of GST-Rev. Furthermore, preliminary experiments using surface plasmon resonance (SPR) [32] to monitor the GST-Rev interaction with the minSLIIB variants in the presence of competitor tRNA do show specific binding to the dI73 variant. This variant also produces a significant blue-shift in the intrinsic fluorescence emission spectrum of the GST-Rev [33], similar to wild type, which the MS2 operator does not (M. Farrow and P. G. Stockley, unpublished work). These results are not consistent with the proposed G47–A73 heteropurine base-pair [26,27], since the C-6 amino group of adenine is replaced by a carboxyl group in inosine and could not hydrogen-bond to the O-6 of guanosine [31]. Recent NMR studies suggest that the G47–A73 base-pair may not be present in the unliganded RNA, consistent with the

DEPC reactivity of A73, only making a stable contact when a complex with Rev or its fragments has formed [34]. This suggests that complexes should also form with the dI73 variant but be less stable, consistent with the gel retardation result.

The conformational subtlety of the minSLIIB-Rev interaction is highlighted by the CD results of deoxyribose variants reported here. The variants on the 5' leg of the bubble do not undergo the conformational change on binding GST-Rev which is seen for wild-type minSLIIB and the variants on the 3' leg. The conformational change in the latter cases is consistent with localized melting of the RNA double-helix, leading to changes in the degrees of stacking and/or tilting of the bases [35,36]. By contrast, the variants on the 5' leg (dGs 46, 47 and 48) form complexes with a variety of affinities but only show a very small CD effect. This suggests that the sugar residues in this region of RRE play a crucial role in the conformational change associated with formation of the liganded state. The apparent molar saturation at a GST-Rev:minSLIIB stoichiometry of 1:1 must be interpreted with care since multimeric proteins and complexes are almost certainly present under these conditions. The data, however, would suggest that single RNA fragments are bound by single Rev domains, assuming that all the molecules are available for binding. It will be of interest to explore how the conformational change is related to the effects of Rev-RRE interaction *in vivo*.

No photo-cross-linked products were observed for any of the substituted single-stranded residues. However, the r4SU45 variant did produce a photo-product that was stable to denaturation with SDS. Thio-nucleotides are useful photo-probes of protein-RNA complexes since their unique absorbance bands between 320 and 350 nm are well away from the absorbance bands of the other protein and nucleic acid components. The result is that photo-cross-links can be formed in high yield without significant photo-damage to the macromolecules, enabling the precise positions of the cross-links to be identified. Experiments to do this for the r4SU45 variant are in hand.

We thank Professor Bryan Cullen for gifts of the RRE transcript vectors and the GST-Rev fusion constructs and for many helpful discussions. We thank Dr. Hemant K. Tewary and James Murray for help with the synthesis and purification of synthetic oligonucleotides and Mark Farrow for preliminary SPR binding data. This work was supported by grants from the SERC to P.G.S. and N.C.P., and also from the Wellcome Trust to P.G. S.S.B.R. and A.D.C. were supported by SERC Studentships.

REFERENCES

- Vaishnav, Y. N. and Wong-Staal, F. (1991). *Annu. Rev. Biochem.* **60**, 577-630
- Malim, M. H., Hauber, J., Le, S.-Y., Maizel, J. V. and Cullen, B. R. (1989). *Nature (London)* **338**, 254-257
- Heaphy, S., Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Karn, J., Lowe, A. D., Singh, M. and Skinner, M. A. (1990). *Cell* **60**, 685-693
- Chang, D. D. and Sharp, P. A. (1989). *Cell* **59**, 789-795
- Cochrane, A. W., Jones, K. S., Beidas, S., Dillon, P. J., Skalka, A. M. and Rosen, C. A. (1991). *J. Virol.* **B**, 5305-5313
- Rosen, C. A., Terwilliger, E., Dayton, A., Sodroski, J. G. and Haseltine, W. A. (1988). *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2071-2075
- Malim, M. H., Tiley, L. S., McCarn, D. F., Rusche, J. R., Hauber, J. and Cullen, B. R. (1990). *Cell* **60**, 675-683
- Tiley, L. S., Malim, M. H., Tewary, H. K., Stockley, P. G. and Cullen, B. R. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 758-762
- Heaphy, S., Finch, J. T., Gait, M. J., Karn, J. and Singh, M. (1991). *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7366-7370
- Malim, M. H. and Cullen, B. R. (1991). *Cell* **65**, 241-248
- Olsen, H. S., Cochrane, A. W., Dillon, P. J., Nalin, C. M. and Rosen, C. A. (1990). *Genes Dev.* **4**, 1357-1364
- Zapp, M. L., Hope, T. J., Parslow, T. G. and Green, M. R. (1991). *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7734-7738
- Critchley, A. D., Haneef, I., Cousens D. J. and Stockley P. G. (1993). *J. Mol. Graphics* **11**, 92-97
- Peattie D. A. (1979). *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1760-1764
- Donis-Keller, H. (1980). *Nucleic Acids Res.* **8**, 3133-3141
- Ehresmann, C., Baudin, F., Mougél, M., Romby, P., Ebel, J.-P. and Ehresmann, B. (1987). *Nucleic Acids Res.* **5**, 9109-9128
- Hertzberg, R. P. and Dervan, P. B. (1984). *Biochemistry* **23**, 3934-3945
- Usman, N., Ogilvie, K. K., Jiang, M.-Y. and Cedergren, R. J. (1987). *J. Am. Chem. Soc.* **109**, 7845-7854
- Talbot, S. J., Goodman, S., Bates, S. R. E., Fishwick, C. W. G. and Stockley, P. G. (1990). *Nucleic Acids Res.* **18**, 3521-3528
- Goodman, S. T. S. (1992). PhD Thesis, University of Leeds
- Adams, C. J., Murray, J. B., Arnold, J. R. P. and Stockley, P. G. (1994). *Tetrahedron Lett.* **35**, 765-768
- Murray, J. B., Collier, A. K. and Arnold, J. R. P. (1994). *Anal. Biochem.* **218**, 177-184
- Daly, T. J., Rusche, J. R., Maione, T. E. and Frankel, A. D. (1990). *Biochemistry* **29**, 9791-9795
- Talbot, S. J., Medina, G., Fishwick, C. W. G., Haneef, I. and Stockley, P. G. (1991). *FEBS Lett.* **283**, 159-164
- Tinoco I. and Cantor C. R. (1970). *Methods Biochem. Anal.* **18**, 81-203
- Bartel, D. P., Zapp, M. L., Green, M. R. and Szostak, J. W. (1991). *Cell* **67**, 529-536
- Iwai, S., Pritchard, C., Mann, D. A., Karn, J. and Gait, M. J. (1992). *Nucleic Acids Res.* **20**, 6465-6472
- Kjems, J., Brown, M., Chang, D. D. and Sharp, P. A. (1991). *Proc. Natl. Acad. Sci. U.S.A.* **88**, 683-687
- Stockley, P. G., Stonehouse, N. J., Murray, J. B., Goodman, S. T. S., Talbot, S. J., Adams, C. J., Liljas, L. and Valegård, K. (1995). *Nucleic Acids Res.*, in the press
- Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J. and Liljas, L. (1994). *Nature (London)* **371**, 623-626
- Saenger, W. (1984). in *Principles of Nucleic Acid Structure*, p. 56 Springer-Verlag, New York
- Karlsson, R., Roos, H., Fagerstam, L. and Persson, B. (1994). *Methods: A Companion to Methods in Enzymology*, vol. 6, pp. 99-110, Academic Press, New York
- Casa-Finetti, J. R., Afonina, F., Gulnik, S., Adachi, Y., Paulakis, G. N. and Erikson, J. W. (1994). *Biophys. J.* **66**, pA358
- Battiste, J. L., Tan, R., Frankel, A. D. and Williamson, J. R. (1994). *Biochemistry* **33**, 2741-2747
- Samejima, T., Hashizume, H., Imahori, K., Fujii, I. and Miura, K.-I. (1968). *J. Mol. Biol.* **34**, 39-48
- Aboul-ela, F., Varani, G., Walker, G. T. and Tinoco, I. (1988). *Nucleic Acids Res.* **16**, 3559-3572