2592–2600 Nucleic Acids Research, 1994, Vol. 22, No. 13

Methylphosphonate mapping of phosphate contacts critical for RNA recognition by the human immunodeficiency virus tat and rev proteins

Clare E.Pritchard+, Jane A.Grasby, François Hamy§, Anthony M.Zacharek¶, Mohinder Singh, Jonathan Karn and Michael J.Gait*
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Received March 22, 1994; Revised and Accepted May 25, 1994

ABSTRACT

The HIV-1 regulatory proteins tat and rev are both RNA binding proteins which recognize sequences in duplex RNA which are close to structural distortions. Here we identify phosphate contacts which are critical for each binding reaction by use of a new method. Model RNA binding sites are constructed carrying substitutions of individual phosphodiesters by uncharged methylphosphonate derivatives isolated separately as *Rp* and *Sp* diastereoisomers and tested for protein binding by competition assays. In the binding of tat to the *trans*-activation response region (TAR), three phosphates, P21 and P22 which are adjacent to the U-rich bulge and P40 on the opposite strand, are essential and in each case both isomers inhibit binding. Similarly, in the interaction between the HIV-1 rev protein and the rev-responsive element (RRE) both methylphosphonate isomers at P103, P104, P124 and P125 interfere with rev binding. At P106, only the *Rp* methylphosphonate isomer is impaired in rev binding ability and it is proposed that the *Rp* oxygen is hydrogen-bonded to an uncharged amino acid or to a main chain hydrogen atom. Synthetic chemistry techniques also provide evidence for the conformations of non-Watson – Crick G_{106}:G_{129} and G_{105}:A_{131} base-pairs in the RRE 'bubble' structure upon rev binding. Almost all functional groups on the 5 bulged residues in the bubble have been ruled out as sites of contact with rev but, by contrast, the N^7-positions of each G residue in the flanking base-pairs are identified as sites of likely hydrogen-bonding to rev. The results show that both tat and rev recognize the major groove of distorted RNA helixes and that both proteins make specific contacts with phosphates which are displaced from the sites of base-pair contact.

INTRODUCTION

Gene expression in human immunodeficiency virus (HIV) requires two viral proteins: the *trans*-activator, tat, and the regulator of virion expression, rev (1, 2). Tat stimulates transcriptional elongation from the viral long terminal repeat (LTR), whereas rev is required for the efficient cytoplasmic expression of mRNAs encoding the viral enzymes and structural proteins.

Tat interacts with the *trans*-activation responsive region (TAR), a 59-residue stem-loop structure that is located at the 5′-end of all mRNAs (Figure 1) (3–5). Recombinant tat protein binds TAR RNA stoichiometrically and with high efficiency (K_d 1–3 nM) (6, 7). The binding site for tat surrounds a distortion in the TAR RNA duplex created by a trinucleotide bulge (7–12). Residues essential for tat recognition include the first residue in the bulge, U_{23}, and the two base-pairs immediately above the bulge, G_{26}:C_{39} and A_{27}:U_{38} (7, 8, 11, 12). The two base-pairs below the bulge also make a contribution to tat binding affinity (11, 12). The other residues in the bulge, U_{24} and U_{25}, appear to act predominantly as spacers and may be replaced by other nucleotides or even by non-nucleotide linkers (10).

Rev is also an RNA binding protein which binds directly to a second genetic element encoded by HIV, the rev-response element (RRE) (13–17). The complete biologically active RRE is 351 nucleotides long and forms a complicated secondary structure (18). The binding reaction is complex and involves an initial interaction with a high affinity site followed by the co-operative addition of rev monomers to lower affinity sites (14, 19–21). Initial high affinity interaction of rev with the RRE RNA occurs at a unique 'bubble' structure (Figure 1) (22–25) which is stabilized by cross-strand non-Watson–Crick G:G and G:A pairs (19, 24, 25). In preliminary work, it was shown that the N^7-position of G_{104} flanking the bubble is essential to rev binding, suggesting that, like tat, rev also recognizes the major groove of an RNA duplex at a site of a local distortion (25) [NB.

*To whom correspondence should be addressed

Present addresses: +Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK, §Ciba-Geigy AG, CH-4002 Basel, Switzerland and ¶Department of Biochemical Sciences, Harvard University, Cambridge, MA, USA

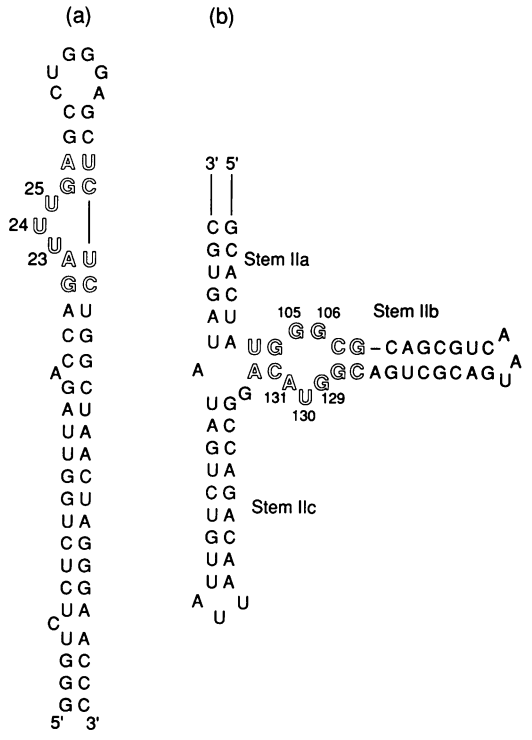

Figure 1. (**a**) Secondary structure of the *trans*-activation responsive region (TAR) RNA. Highlighted residues define the region involved with tat recognition. (**b**) Partial secondary structure for the stem-loop II region of the rev-response element (RRE) RNA. Highlighted residues define the region involved with initial high affinity rev recognition.

Numbers for nucleotide residues are +70 compared to our previous publication (25)]. At higher rev concentrations, additional rev molecules become associated with the RRE, packing around selected regions of primarily double stranded RNA. The packing of rev around the RRE correlates well with rev activity *in vivo* (18).

In addition to base recognition, specific contacts through phosphate residues are often essential elements in protein–nucleic acid interactions. For example in the high resolution structural model of the interaction of the Zif268 complex with DNA, interactions are seen between phosphate and serine residues as well as between phosphate and arginine or histidine residues (26). In the interaction of helix 3 of Hin recombinase with the major groove of a DNA duplex, there are simultaneous contacts of phosphate 8 with a tyrosine and an arginine and also of phosphate 9 with a threonine and the main chain amide of a glycine (27). Whereas contacts to uncharged amino acids are hydrogen bonds, contacts to the side-chains of basic residues may have both hydrogen-bonding as well as some electrostatic character, due to the presence of the opposing charges.

The most commonly used method for mapping phosphate contacts in nucleic acid–protein interactions is ethylation interference, and this technique has been used to define positions which are important for both the tat–TAR and the rev–RRE interactions. Ethylation of either of two adjacent phosphates in TAR, between G_{21} and A_{22} (P21) or between A_{22} and U_{23} (P22), interfered with the binding of a short basic peptide (28) as well as of the amino acid arginine (29). Ethylation interference studies of TAR using the complete tat protein revealed additional sites of interference at five phosphates on the opposite strand (P36–P40) (12). Ethylation interference analysis using rev and a rev-derived peptide has identified three phosphates on one side of the bubble and three phosphates on the other strand distant from the bubble as important to the interaction (22).

Although ethylation interference analysis has proved valuable for suggesting possible phosphate contacts in both tat–TAR and rev–RRE interactions, the method has some limitations. First, the introduction of bulky ethyl groups by treatment of RNA with ethylnitrosourea (ENU) is likely to cause substantial steric hindrance to interactions with a protein. Furthermore, ethylation results in the non-stereo-specific introduction of an *O*-ethyl group to give a mixture of both *R*p and *S*p ethyl phosphotriesters. Because both diastereoisomers are present, the ENU interference data does not allow distinctions to be made between electrostatic contacts, hydrogen-bonding contacts, or combinations of these two types of contact. Thus, ENU footprints should not be regarded as definitive evidence for sites of phosphate–protein contact.

Recently we have begun to explore an alternative approach to the study of phosphate residues in RNA important for protein binding. Synthetic chemistry is used to replace individual phosphodiesters in the RNA by methylphosphonates (30). In these derivatives, a methyl group replaces one of the phosphate oxygens to give an uncharged isosteric analogue of phosphate. The small methyl group is much less likely to interfere with protein binding than a bulky *O*-ethyl group. Here we describe the separation of oligonucleotides carrying the individual *R*p and *S*p methylphosphonate isomers and assembly into RNA models representing TAR and RRE. The results show the effects of substitution of individual oxygen atoms of phosphate residues and give important insights into the location and nature of phosphate contacts in both the rev–RRE and the tat–TAR interactions.

MATERIALS AND METHODS

Chemical synthesis of oligoribonucleotides

Oligoribonucleotides were synthesized by the phosphoramidite method on a 1 mmol scale using an ABI 380B DNA/RNA Synthesizer (25, 30, 31). Ribonucleoside phosphoramidites, 2′-deoxyncleoside methylphosphonoamidites (dT, dC, dG and dA), spacer phosphoramidite C3, N^7-deaza-dA and N^7-deaza-dG phosphoramidites were obtained from Glen Research (via Cambio). A sample of dU methylphosphonoamidite was a gift from Dr Nanda Sinha (Millipore Corporation). N^6-Benzoyl-2′-deoxyadenosine and N^4-benzoyl-2′-deoxycytidine phosphoramidites were obtained from Cruachem (UK) and N^2-phenoxyacetyl-2′-deoxyguanosine was obtained from Pharmacia. Other reagents were from Applied Biosystems Ltd or Cruachem. O^6-Methylguanosine phosphoramidite was obtained as previously described (31). 2′-Deoxy-2′-fluoroadenosine phosphoramidite was a gift from F.Eckstein (32). Oligonucleotides were cleaved from the support and base-deprotected with saturated methanolic ammonia at room temperature for 24 h (or 48 h in the case of benzoyl removal from N^7-deazaA). O^6-MethylG-containing oligonucleotides were deprotected with 10% DBU in methanol as previously described (31). 2′-*O*-t-Butyldimethylsilyl groups were removed by treatment with 1 M tetrabutylammonium fluoride (Aldrich) in tetrahydrofuran (1 ml) at room temperature for 24–48 h. After desalting, oligonucleotides were purified by strong anion exchange HPLC on a 1×25 cm Partisil-10 SAX column (Whatman) as previously described (25, 30, 31) and

checked for purity by reversed-phase HPLC on a μ-Bondapak C18 column (0.4×25 cm) (Millipore). Typical yields for 15-mers after purification were 10 A_{260} units (*ca* 5% overall yield). For oligonucleotides containing single 2′-deoxy-3′-methylphosphonate substitutions, individual diastereoisomers were resolved using a shallow gradient of acetonitrile (generally 7.5−17.5% in 25 min, Figure 3). The product in each peak was collected and desalted.

Analysis of oligonucleotides

Nucleoside analysis was carried out by digestion of the oligonucleotides with snake venom phosphodiesterase (Boehringer) and alkaline phosphatase (Boehringer) followed by HPLC analysis on a μ-Bondapak C18 reversed-phase column (0.4×25 cm) as previously described (25). Retention times (0−15% acetonitrile gradient in 25 min) for unmodified nucleosides were: C, 6.2; U, 8.4; G, 16.0; A, 19.4; dG, 18.2; dA, 22.5 min. Retention times for nucleosides containing modified nucleosides were: O^6-methylG, 22.9; 2′-fluoro-2′-dA, 24.8; N^7-deaza-2′-dA, 22.7; N^7-deaza-2′-dG, 19.5 min. Oligonucleotides containing single 2′-deoxy-3′-methylphosphonates were digested under the same conditions and the resultant methylphosphonate dimers separated by reversed-phase HPLC using a 0−20% acetonitrile gradient. For example, digestion of the fast eluting isomer (25.1 min) of CGUGUGGdGmprCGCAGC (methylphosphonate at P_{106}, i.e. between G_{106} and C_{107}, Figure 2) gave methylphosphonate dimer (dGmprC) isomer 1 eluting at 22.8 minutes, whereas the slow eluting isomer (25.4 minutes) of the same oligonucleotide gave methylphosphonate dimer isomer 2 eluting at 23.8 min. All other methylphosphonate-containing oligonucleotides were checked in this way. In each case the slow eluting oligomer always produced the slow eluting dimer after nuclease digestion.

The positions of 2′-deoxynucleoside substitutions were checked by ^{32}P end-labelling of the oligonucleotides, by treatment with [γ-^{32}P]ATP and T4 polynucleotide kinase, followed by partial alkaline hydrolysis and polyacrylamide gel electrophoresis as previously described (30). The absence of a radioactive band signified the presence of a 2′-deoxynucleoside at that position, since such residues are resistant to alkaline hydrolysis. By contrast, 2′-deoxy-3′-methylphosphonate-containing oligonucleotides were cleaved at the methylphosphonate linkage under partial alkaline hydrolysis (data not shown) and produced a radioactive band which migrated slower than the equivalent band obtained for hydrolysis of the corresponding inter-ribonucleotide linkage. This band was therefore diagnostic for the presence of the methylphosphonate. No migration differences were observed between individual *Rp* and *Sp* isomers.

Tat binding assays

Tat protein was prepared and binding assays were carried out as previously described (12). ^{32}P-Labelled unmodified duplexes were formed by annealing a 1.5-fold molar excess of 17-mer strand with 14-mer strand labelled with [γ-^{32}P]ATP and T4 polynucleotide kinase. Competition binding assays (25 μl) included 0.1% Triton X100, 50 mM Tris−HCl (pH 7.4), 20 mM KCl, 0.1 M DTT, 7.5 nM ^{32}P-labelled unmodified duplex RNA (20 000 c.p.m.), 20−30 nM tat protein and various concentrations of unlabelled competitor duplex RNA. Binding reactions were incubated on ice for 30 min and non-denaturing polyacrylamide gel electrophoresis carried out at 4°C (8% gel, 20×20 cm) in 0.5× TB (TB = 44.5 mM Tris base, 44.5 mM boric acid), 0.1% Triton for 35 min at 12 W. Gels were dried and complexes detected by autoradiography. Relative binding constants were calculated after scanning of the resultant autoradiographs by densitometry.

Rev binding assays

Rev protein was prepared as described (19, 34). Competition binding reactions (20 μl) shown in Figure 5 contained 0.5% glycerol, 0.05% Triton X100, 40 mM Tris−HCl (pH 7.9), 20 mM KCl, 3.5 nM ^{32}P-labelled unmodified duplex RNA, 28 nM rev protein and various concentrations of unlabelled competitor duplex RNA. Other competition binding reactions (20 μl) (Table 3) were carried out similarly except that 25 nM ^{32}P-labelled duplex RNA and 70 nM rev protein concentrations were used. Non-denaturing gel electrophoresis was carried out with 4% (Figure 5, Table 2) or 7.5% (Table 3) polyacrylamide gels and running buffer 0.5× TBE (TBE = 44.5 mM Tris base, 44.5 mM boric acid, 1 mM EDTA), 0.1% Triton, 1% glycerol.

RESULTS

Synthetic RNA models for TAR RNA and RRE RNA containing methylphosphonate diastereomer substitutions

Both the binding site for tat on TAR and the high affinity site for rev on the RRE can be mimicked by short synthetic RNA duplexes (Figure 2) (10, 25, 30). Models of this type lend themselves to studies where modified bases are incorporated into specific sites. This approach has been used to identify the N^7-positions of both G_{26} and A_{27} and the N^3-H and/or O^4-positions of U_{23} as crucial to tat binding (10, 30). Similarly, functional group substitution experiments at the high affinity rev binding site have suggested specific base-pairing schemes for the non-Watson−Crick base-pairs in the 'bubble' and indicated that rev interacts with the N^7 of G_{104} (25).

Single 2′-deoxy-3′-methylphosphonate linkages were incorporated into these model RNA duplexes by chemical synthesis. Because methylphosphonate linkages 3′ to ribose units are highly unstable, a ribonucleoside phosphoramidite was replaced by the corresponding 2′-deoxy-3′-methylphosphonoamidite at the appropriate coupling step. Oligonucleotides having the same sequence but with the corresponding 2′-deoxynucleotide incorporated at the equivalent site were synthesized as controls for the possible effects of the sugar substitution on protein binding.

Oligonucleotides containing single methylphosphonate isomers were resolved by careful high performance liquid chromatography (HPLC) using a reversed-phase column and shallow gradients of acetonitrile eluents. A typical example of diastereoisomer separation by HPLC is shown in Figure 3.

It has been shown recently that for all 16 pairs of di-2′-deoxynucleoside 3′→5′-methylphosphonates isomers the *Rp* isomer elutes faster and the *Sp* isomer slower on reversed-phase HPLC (35). Therefore in order to distinguish the stereochemistry at the methylphosphonate for the synthesized oligonucleotides, we digested each pair of the separated oligonucleotides with snake venom phosphodiesterase followed by alkaline phosphatase. This yields the four nucleosides and a dinucleoside 3′→5′-methylphosphonate. These products were analysed by reversed-phase HPLC (data not shown). As expected, in each case the dinucleoside 3′→5′-methylphosphonate released from the fast eluting oligonucleotide product was eluted faster than the corresponding dinucleoside 3′→5′-methylphosphonate obtained from the slow

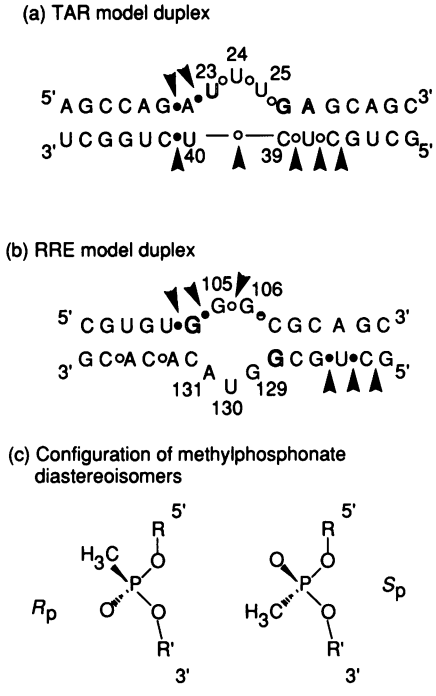

Figure 2. (**a**) Structure of a TAR RNA model prepared by the annealing of two synthetic oligoribonucleotides. Highlighted residues indicate where specific functional groups on bases have been implicated as sites of hydrogen-bonding contact with tat (10, 29). Arrow heads indicate phosphate residues mapped by ENU interference studies as possible sites of interaction with tat protein (12, 29). Filled circles: both methylphosphonate isomers interfere with tat binding, open circles: neither methylphosphonate isomer (this work) nor mixed isomers (30) interfere with tat binding. (**b**) Structure of a RRE RNA model prepared by annealing two synthetic oligoribonucleotides. Highlighted residues indicate the two purines where N^7-positions have been implicated as sites of hydrogen-bonding contact with rev (25 and this work). Arrow heads indicate phosphate residues mapped by ENU interference studies (22). Filled circles: both methylphosphonate isomers interfere with rev binding, half-filled circles: one isomer interferes, open circles: neither isomer interferes with rev binding. (**c**) Structures of the *Rp* and *Sp* configurations of methylphosphonate diastereoisomers.

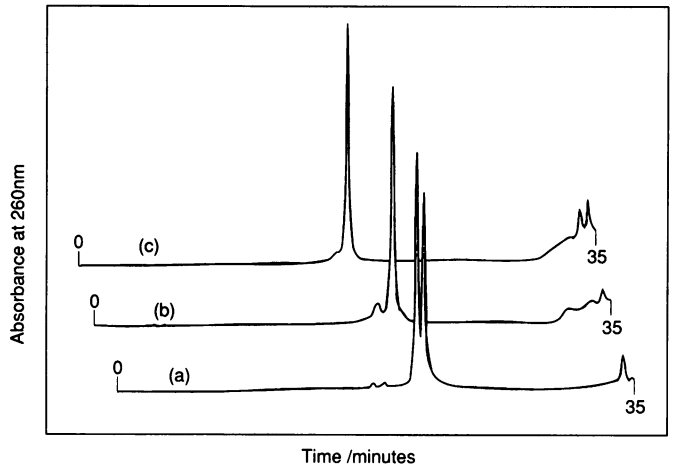

Figure 3. Reversed-phase HPLC chromatograms showing the purification of the two isomers of the 14-mer rCGUGdTmprGGGCGCAGC. (**a**) Oligonucleotide product peak material after ion exchange purification, (**b**) purified slow isomer (*Sp*), (**c**) purified fast isomer (*Rp*). The elution times in (a) are slightly different to those in (b) and (c) since the separations were carried out on different days with different batches of elution buffers. Slight variation in compositions of elution buffers have a noticeable effect on elution times due to the use of very shallow acetonitrile gradients.

eluting oligonucleotide. Thus we believe that for each pair of oligonucleotides containing a single methylphosphonate linkage, the fast eluting oligonucleotide product corresponds to the *Rp* isomer and the slow eluting product to the *Sp* isomer.

Identification of phosphates in TAR RNA required for tat binding

In the previous studies using unfractionated oligonucleotides carrying methylphosphonates, we observed that substitution at position P22 (i.e. 3′ to A_{22}) strongly inhibited tat binding, but methylphosphonates were well-tolerated at P23, P24 and P25 (30). We have now prepared stereo-specific methylphosphonate substitutions at positions P21 and P22. We have also prepared substitutions at P37, P38, P39 and P40 on the other strand, since these sites have all been implicated by ethylation interference analysis as possible sites of phosphate contact with tat (12, 29).

The ability of TAR analogues carrying methylphosphonates to bind tat was measured by competition gel mobility shift assays (Figure 4 and Table 1). Competition assays are more sensitive than direct binding assays to the relatively small differences in binding constants (10–50-fold) observed between specific and non-specific RNA–protein complex formation (12, 30). Briefly, tat was bound to unmodified TAR RNA which had been 5′-^{32}P-labelled in the presence or absence of modified competitor TAR RNA. Derivatives with high affinities for tat are strong competitors for binding in this assay, whereas derivatives which fail to bind tat with high affinity fail to compete. The data can be expressed in terms of $D_{1/2}$, the concentration of competitor required to reduce by 50% tat binding to the labelled wild-type duplex. However, for ease of comparison between experiments, we have expressed the binding data in Table 1 and subsequent tables as K_{rel}, the ratio of the $D_{1/2}$ values obtained from competitions carried out with modified duplexes and with the control wild-type duplex.

The control duplexes containing single deoxynucleoside substitutions at the equivalent positions (dG_{21}, dA_{22} and dU_{40}) were not impaired in their ability to compete for tat binding (Table 1). TAR RNA duplexes containing methylphosphonates at any of the three positions P21, P22 or P40 showed dramatically reduced affinities compared to the unmodified duplex. At each position, both the *Rp* and *Sp* isomers behaved similarly. By contrast, the methylphosphonate isomers at positions P37 and P39 showed a modest 2-fold reduction in their ability to compete for tat binding compared to the wild type duplex (4-fold compared to the corresponding 2′-deoxynucleoside control duplexes). Again, both the *Rp* and *Sp* isomers behaved similarly. P38 was difficult to score unambiguously because the control (dU_{38}) was 4-fold impaired in tat binding ability. However, binding of both *Rp* and *Sp* methylphosphonate isomers appeared to be further reduced.

Identification of phosphates in the RRE high affinity site required for rev binding

Methylphosphonate substitutions were also introduced into the model of the RRE high affinity binding site. Ethylation interference analysis has previously suggested that P103, P104, P105, P124 and P125 contact rev (22). Methylphosphonates were introduced into these positions as well as three sites which were not previously implicated, P106, P133 and P135 (Figure 2b).

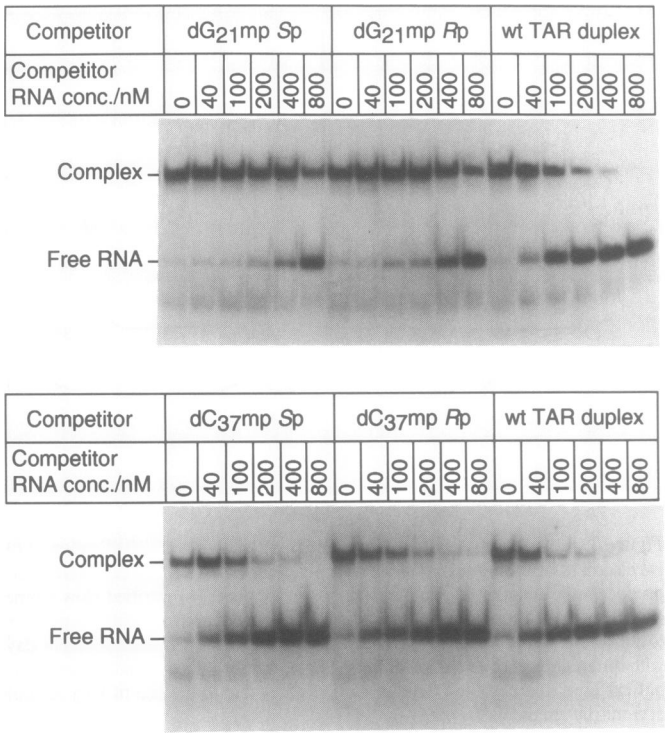

Figure 4. Competition band-shift assay for tat binding. Each lane contains 7.5 nM ^{32}P-labelled wild-type TAR duplex RNA and 20–30 nM tat protein. Concentrations of unlabelled competitor RNA are as shown: top panel, dG_{21}mp *Sp*, dG_{21}mp *Rp* and wild-type RNA duplex; bottom panel, dC_{37}mp *Sp*, dC_{37}mp *Rp* and wild-type RNA duplex.

In each case the *Rp* and *Sp* diastereoisomers were readily resolved by use of reversed-phase HPLC.

The effect of the methylphosphonates on rev binding was measured in competition gel-mobility shift assays similar to those carried out to study the tat–TAR interaction (Figure 5 and Table 2). At three of the positions mapped by ethylation interference, P103 and P104 on the top strand and P125 on the bottom strand, both methylphosphonate isomers strongly inhibited rev binding more than 8-fold, whilst the corresponding 2′-deoxynucleoside-containing control oligonucleotides were hardly affected. At positions P103 and P125 the methylphosphonate substitutions were carried out using commercially available dT–methylphosphonoamidite derivative. However, in both cases the control dT-containing duplexes were unaffected in rev binding ability. By contrast, at position P124 there was a significant effect of 2′-deoxy substitution (dC_{124}) and, although the methylphosphonate appeared to additionally reduce binding, it is difficult to score this position unambiguously.

At P105, which had previously been shown to inhibit rev binding in ethylation interference experiments, neither methylphosphonate isomer nor the control dG_{105} was impaired in their ability to compete for rev binding. Methylphosphonate substitution at either position P133 or P135, which are each located well away from the bubble, also had no effect on competitive rev binding ability. The corresponding 2′-deoxynucleotide-containing controls (dA_{133} and dA_{135}) were able to compete normally for rev binding.

An example of stereo-specific inhibition by methylphosphonates is provided by P106, which is located within the 'bubble'. The *Rp* methylphosphonate isomer competed poorly for rev binding whereas the *Sp* isomer competed as well as the control RNA duplex and the dG_{106} control.

Table 1. Competition tat binding for methylphosphonate-modified TAR RNA duplexes

Wildtype		Methylphosphonate	
Duplex RNA	**K_{rel}[a]**	**Duplex RNA**	**K_{rel}**
unmodified	1.0		
dG_{21}	1.5	dG_{21}mp (*Rp*)	5.6
		dG_{21}mp (*Sp*)	6.4
dA_{22}	1.3	dA_{22}mp (*Rp*)	>8
		dA_{22}mp (*Sp*)	>8
dC_{37}	0.4	dC_{37}mp (*Rp*)	1.6
		dC_{37}mp (*Sp*)	1.8
dU_{38}	4.3	dU_{38}mp (*Rp*)	>8
		dU_{38}mp (*Sp*)	>8
dC_{39}	0.6	dC_{39}mp (*Rp*)	2.1
		dC_{39}mp (*Sp*)	2.5
dU_{40}	2.2	dU_{40}mp (*Rp*)	>8
		dU_{40}mp (*Sp*)	>8

[a]K_{rel} is the concentration (nM) of mutant competitor duplex required to reduce by 50% tat binding to ^{32}P-labelled wild-type duplex divided by the concentration (nM) of wild-type unmodified duplex required to reduce by 50% tat binding to ^{32}P-labelled wild-type duplex.

U_{130} acts as a spacer

We have previously reported that U_{130} in the bubble may be replaced by dU, 5-bromodU or by N^3-methylU (25). We therefore suggested that this nucleoside is not likely to be involved in cross-strand base-pairing. However, use of these analogues did not address the issue of potential recognition of U_{130} by rev (e.g. at O^2 or at O^4). We therefore replaced U_{130} by a C3-spacer and observed the effect on rev binding. The C3-spacer is a propyl linker that maintains the correct number and type of atoms between two neighbouring phosphate residues in one oligonucleotide strand, yet the sugar and base is entirely absent. C-3 spacers have previously been used in assessing the role of U residues in TAR (11). The results showed that there was no effect of the C-3 spacer on rev binding (Table 3). Thus U_{130} appears to play no part either in cross-strand hydrogen-bonding or in rev recognition.

Configuration of non-Watson–Crick base-pairs in the RRE bubble

Previous studies have suggested that the bubble in the RRE is stabilized by non-Watson–Crick base-pairs between G_{106} and G_{129} and between G_{105} and A_{131} (19, 24, 25). We have now completed chemical substitution experiments (Table 3) which, taken together with our previous studies (25), provide strong evidence for the conformations of the proposed base-pairs shown in Figure 6. The data also suggest that the non-Watson–Crick base-pairs are unlikely to form specific contacts to rev, but instead they are essential to form a unique structural element for rev binding.

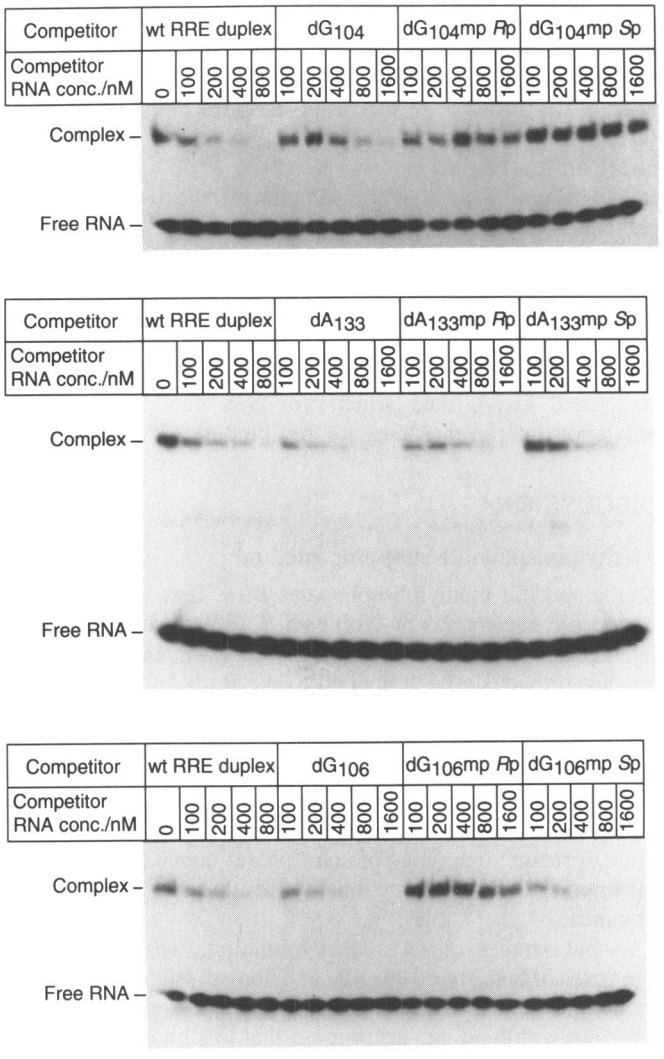

Figure 5. Competition band-shift assay for rev binding. Each lane contains 25 nM ^{32}P-labelled wild-type RRE duplex RNA and 70 nM rev protein. Concentrations of unlabelled competitor RNA are as shown: top panel, wild-type RRE duplex RNA, dG_{128} and N^7-deaza-dG_{128}; middle panel, wild-type RRE duplex RNA, dA_{133} and N^7-deaza-dA_{133}; bottom panel, wild-type RRE duplex RNA, dG_{129} and N^7-deaza-dG_{129}.

Rev binding was tolerant of N^7-deaza-dG substitution at position G_{106} and also tolerant of simultaneous loss of the exocyclic amino groups at G_{106} and G_{129} (25). These results are consistent only with a symmetrical O^6 to N^1-H configuration of this pair as proposed by Bartel *et al.* (24). The results also rule out the possibility that either the N^7-position or the 2′-OH of G_{129} interacts with rev. By contrast, substitution of G_{129} by a C-3 spacer caused a dramatic loss of rev binding suggesting that the nucleoside moiety of G_{129} is required for an important structural role or possibly to make a contact to rev via a less obvious position (i.e. $O^{4'}$ or N^3 positions). Surprisingly, O^6-methyl-G substitution of G_{106} produced only a small loss of rev binding (Table 3), suggesting that there may be some flexibility in the G_{106}:G_{129} base-pair. Recent *in vitro* selection experiments have shown that this pair may be replaced by A_{106}:A_{129} or by C_{106}:A_{129}, which can each form isosteric hydrogen bonds (36).

Table 2. Competition rev binding data for methylphosphonate-modified RRE RNA duplexes

Wildtype		Methylphosphonate	
Duplex RNA	**K_{rel}**	**Duplex RNA**	**K_{rel}**
unmodified	1.0		
dT_{103}	2.1	dT_{103}mp (*Rp*)	>8
		dT_{103}mp (*Sp*)	>8
dG_{104}	1.9	dG_{104}mp (*Rp*)	>18
		dG_{104}mp (*Sp*)	>18
dG_{105}	0.8	dG_{105}mp (*Rp*)	0.3
		dG_{105}mp (*Sp*)	0.8
dG_{106}	1.0	dG_{106}mp (*Sp*)	1.0
		dG_{106}mp (*Rp*)	>18
dC_{124}	6.0	dC_{124}mp (*Rp*)	>9
		dC_{124}mp (*Sp*)	>9
dT_{125}	1.7	dT_{125}mp (*Rp*)	10.6
		dG_{125}mp (*Sp*)	>16
dA_{133}	0.6	dA_{133}mp (*Rp*)	0.8
		dA_{133}mp (*Sp*)	1.3
dA_{135}	1.0	dA_{135}mp (*Rp*)	1.9
		dA_{135}mp (*Sp*)	0.6

Table 3. Competition rev binding data for base-modified RRE RNA duplexes.

Unmodified		Base-modified	
Duplex RNA	**K_{rel}**	**Duplex RNA**	**K_{rel}**
unmodified	1.0		
		O^6-methyl-G_{105}	5.4
		O^6-methyl-G_{106}	1.4
dG_{108}	0.3	N^7-deaza-dG_{108}	0.4
		N^7-deaza-dA_{110}	1.0
dG_{128}	1.6	N^7-deaza-dG_{128}	10.3
dG_{129}	9.3	C-3 spacer129	10.1
		N^7-deaza-dG_{129}	0.8
		C-3 spacer130	0.9
dA_{131}	17.8	N^7-deaza-dA_{131}	2.4
		2′-deoxy-2′-fluoroA_{131}	1.6
		G_{131}	14.5
dA_{133}	0.6	N^7-deaza-dA_{133}	0.3

Our previous studies suggested two possible configurations for the G_{105}:A_{131} base-pair (25). Each would contain hydrogen bonds between G_{105}-O^6 and the exocyclic amino group of A_{131}. However in one case the G_{105} N^1-H would interact with A_{131}-N^1 and in the other with A_{131}-N^7. To distinguish these possibilities, we substituted A_{131} by N^7-deaza-dA and found no effect on rev binding (Table 3). This result points to the G:A pair containing a A_{131}-N^1 hydrogen bond to G_{105}-O^6 being the more likely and that neither the N^7 nor the 2′-OH of A_{131} is involved in rev recognition. Further evidence for the presence of the G:A pair

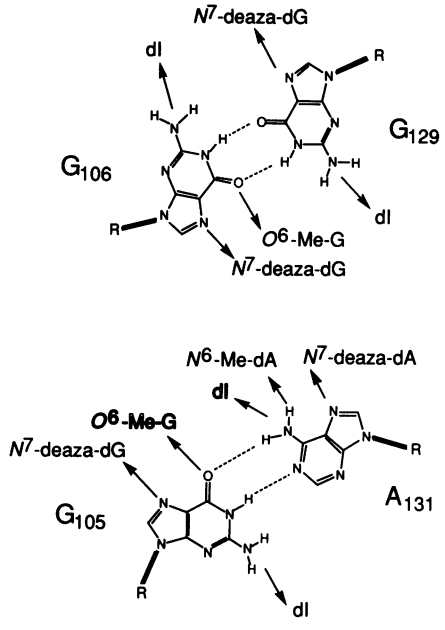

Figure 6. Proposed non-Watson–Crick base-pairing schemes for G_{106}:G_{129} and G_{105}:A_{131} showing functional group substitutions that are either unaffected in rev binding (bold) or significantly reduced in rev binding (outlined).

comes from the finding that substitution of G_{105} by O^6-methyl-G resulted in a moderate loss of rev binding ability (Table 3). Severe loss of binding ability resulted when A_{131} was replaced by G (Table 3) or by dI (25). The same structure for the G:A pair was also suggested from recent chemical probing studies on an *in vitro* selected RRE model RNA (37).

It is interesting to note that substitution by the corresponding 2′-deoxynucleoside at either G_{129} and A_{131} inhibited rev binding, whereas substitution by the corresponding N^7-deaza-2′-deoxynucleoside had no effect on rev binding. One interpretation of these results is that rev binding is greatly influenced by the conformations of the sugars at these two positions and that these conformations are affected in turn by the identities of the corresponding bases. Thus perhaps the 2′-deoxynucleoside sugars, which prefer a $C^{2'}$-*endo* configuration, force alternative base-pairing schemes less favourable for rev binding, whereas for the corresponding N^7-deaza-2′-deoxynucleosides, alternative base-pairing is disallowed and the sugars are forced into the $C^{3'}$-*endo* configuration. To address this question, we substituted A_{131} by 2′-deoxy-2′-fluoroA and found no loss of competition for rev binding (Table 3). Since it has been shown that a 2′-deoxy-2′-fluoronucleoside adopts predominantly the $C^{3'}$-*endo* configuration commonly found for ribonucleosides (32), this result suggests that a $C^{3'}$-*endo* configuration at A_{131} is essential for the maintenance of the correct bubble structure to ensure good rev binding.

Rev recognition in the major groove of the RRE

In previous studies it had been found that N^7-deaza-dG-substitution of G_{104}, located in a base-pair flanking the bubble in the model duplex, resulted in a severe loss of rev binding (25) suggesting that rev may interact in the major groove of the RNA duplex. 2′-Deoxy-N^7-deaza-substitutions were examined at G_{108}, A_{110}, G_{128} and A_{133} (Figure 2b). Competition band shift analysis (Table 3) showed that 2′-deoxy-N^7-deaza substitution at G_{128} immediately flanking the bubble dramatically reduced rev binding whereas the 2′-dG substitution alone had no effect. By contrast, none of the 2′-deoxy-N^7-deaza-substitutions in the three purines further away from the bubble, G_{108}, A_{110} and A_{133}, had any effect on rev binding.

The results show that the function of the bubble structure principally seems to be to create a local distortion in the RNA duplex allowing accessibility of the major groove at each of the flanking base-pairs. Major groove recognition of the RNA duplex occurs in both purines flanking the bubble but this does not extend to more distant purines. Instead, rev seems to interact with the N^7-positions of each of the guanines in the flanking base-pairs (G_{104} and G_{128}), data which are consistent with carboxyethylation interference data for rev binding (22, 37).

DISCUSSION

Methylphosphonate mapping method

Stereo-specific methylphosphonates were first introduced into oligodeoxynucleotides by Noble *et al.* (38) for the study of the interaction of a DNA duplex with *lac* repressor, but the technique has not previously been applied to the study of RNA–protein interactions. The results of our methylphosphonate mapping on the TAR binding site and the RRE high affinity site suggest that the method has several advantages over ethylation interference mapping. In particular, because there is less risk of steric clash and because stereochemically pure substitution can be obtained, more precise mapping of phosphate contacts as well as information on the directionality of the interaction can be obtained.

In our studies of tat and rev binding, we have obtained examples of both stereo-specific and non-specific interference by methylphosphonates. Where both methylphosphonate isomers interfere with binding, we propose that this phosphate interacts with a charged partner (e.g. Arg or Lys), since in addition to the interference with hydrogen-bonding, the removal of the phosphate charge would be expected to be disruptive of the electrostatic component of the interaction independent of the stereochemistry of the modified phosphate. An alternative but perhaps less likely explanation for non-stereo-specific interference by methylphosphonates is that each oxygen atom is hydrogen-bonded to a different partner. When only one methylphosphonate isomer interferes, as in the case of P106 in RRE, we suggest that there is a single hydrogen bond to an uncharged amino acid side chain or to the main chain. The hydrogen bond must therefore be to the oxygen atom affected by methylphosphonate substitution (i.e. pro-*R* for *R*p substitution). We have also seen examples of sites which were identified as potential sites of phosphate contacts to protein by ethylation interference mapping but where methylphosphonate substitutions are well tolerated. The most likely explanation for these divergent results is that the ethylation interfered with binding through a steric clash, rather than through removal of a hydrogen-bonding contact.

The tat–TAR complex

Recent NMR data has suggested that free TAR RNA contains a continuous A-form helix which extends throughout the bulge region (39, 40). However, the molecule is highly flexible and can be stabilized in an 'open' configuration with a widened major groove after binding to basic peptides, or even to the amino acid

derivative argininamide. On the basis of these observations, Puglisi *et al.* (39, 40) suggested that a base triple is formed between U_{23} and A_{27}:U_{38}, which creates a binding pocket and allows arginine to insert into the major groove by forming two hydrogen bonds with G_{26} as well as contacts to P_{21} and P_{22}. Although this work has demonstrated the feasibility of NMR studies of the tat–TAR interaction, there is still insufficient NMR data to unambiguously position the argininamide in the complex, to define the width of the major groove, or to directly demonstrate the existence of the base triple.

Preliminary NMR work in this laboratory (G.Varani, personal communication) has focussed on studies of complexes between a model peptide carrying both the 'core' and basic regions of tat and TAR RNA. This 37-residue peptide (ADP-1) more closely resembles tat in its binding specificity (12) than do the short basic peptides and argininamide studied by Puglisi *et al.* (39, 40). In agreement with Puglisi *et al.*, we have found that TAR is rearranged after peptide binding and that there are NOEs linking one of the arginines in the peptide to the base moieties of U_{23} and the adjacent A_{22}. However, we have also found that ADP-1 makes a number of additional contacts to TAR RNA below the TAR bulge.

Methylphosphonate substitution mapping of the TAR RNA structure is of particular value since current NMR techniques give only limited and indirect information as to phosphate–protein contacts. The model of Puglisi *et al.* (39, 40) includes direct hydrogen bonds from argininamide to both P21 and P22 inferred from ethylation interference studies (28, 29). Here we report that both the *Rp* and *Sp* isomers of methylphosphonate substituted at either P21 or P22 interfere with tat binding. The absence of stereo-specificity for these sites is consistent with interaction of each of these phosphate residues with a charged residue, such as arginine. However, the data does not distinguish between a single contact, as postulated by Puglisi *et al.* for the argininamide complex (39, 40), or a model for the tat–TAR complex in which there are separate contacts to the two phosphates made by two different charged amino acid side chains.

In addition to P21 and P22, phosphates P36 to P40 from the opposite strand of TAR RNA have been implicated by ENU interference as sites of tat interaction (12). Peptides which bind TAR RNA specifically (such as ADP-1, residues 37–72) show a strong ENU footprint involving phosphates P36–P40, whereas peptides which bind with lower specificity (such as ADP-3, residues 48–72) do not interact strongly with these phosphates (12). In agreement with these observations, the insertion of either *Rp* or *Sp* methylphosphonate isomers at P40 strongly interfered with tat binding to TAR RNA and it seems likely that P40 is therefore a site of interaction with a charged amino acid side chain of tat. By contrast, the methylphosphonate isomers at P37, P38 and P39 have a weaker effect on tat binding (2- to 4-fold), suggesting that none of these phosphates forms strong hydrogen-bonding contacts with tat. It is probable that tat protein closely approaches these residues on this side of the RNA duplex, with the result that ethylation of P36–P40 phosphates produces substantial steric hindrance. Electrostatic contacts have also been suggested by Tao and Frankel to be of importance for the P23, P24, and P25 phosphates located in the TAR RNA bulge (41), although it should be noted that methylphosphonate substitutions are well-tolerated at these positions (30).

The picture of the tat–TAR complex which is emerging from both chemical and NMR studies is that there is a conformational change in TAR RNA following binding by tat. However, the precise nature of the contacts made between tat and TAR RNA, and the structure of TAR RNA in the bound configuration, remain to be unequivocally determined.

The rev–RRE high affinity complex

The high affinity binding site for rev, like the binding site for tat, is also characterized by a local distortion of an RNA duplex. In the TAR structure, the distortion is produced by the bulged uridines, whereas in the RRE structure, the distortion is due to the presence of a single bulged uridine (U_{130}) and to two non-Watson–Crick base-pairs (G_{106}:G_{129} and G_{105}:A_{131}). NMR studies suggest that the rev binding site, like the tat binding site on TAR, is flexible and can be stabilized by binding to small ligands such as peptides (42). The precise structure of the binding site has not yet been determined, but model building studies based on the structures of a series of variant rev binding sites (aptamers) have suggested that the major groove is substantially widened by the presence of the non-Watson–Crick base-pairs (43).

Our functional group substitution experiments have allowed us to define the likely configuration of the G_{106}:G_{129} and G_{105}:A_{131} base-pairs. The G_{106}:G_{129} pair appears to be in a symmetrical O^6 to N^1-H configuration, whereas in the G_{105}:A_{131} base-pair, G_{105}-O^6 hydrogen bonds with the exocyclic amino group of A_{131}, and G_{105} N^1-H interacts with A_{131}-N^1. These pairing schemes are consistent with NMR data obtained from studies of RRE–peptide complexes (42). Assuming bond angles for the sugars and phosphates normally observed in duplex RNA, G_{105}, A_{131} and G_{106} would be expected to be in *anti* configurations and G_{129} in a *syn* configuration (25, 43). However recent NMR studies predict an unusual *anti* conformation for G_{129} rather than the expected *syn* configuration for this base (42). Although our chemical substitution experiments do not directly address whether the G_{129} base is in a *syn* or *anti* conformation, our finding that O^6-methylG substitution at G_{106} was not harmful to rev binding suggests that the G_{106}:G_{129} pair may be rather flexible, even in the rev-bound state. Therefore it is possible that G_{129} may be able to adopt either *syn* or *anti* conformations depending on the ligand bound to the RRE RNA.

The methylphosphonate substitution experiments reported here have established that rev makes important contacts to the high affinity RRE binding site at P103 and P104 on one side of the bubble and at P124 and P125 on the other strand, somewhat distant to the bubble. Evidence that P103, P104, P124 and P125 play a critical role in rev recognition was also obtained in ethylation interference studies of a RRE high affinity site generated by *in vitro* selection (37). At each of the four sites of methylphosphonate interference both isomers are impaired in rev binding ability. These results are consistent with these phosphates being involved in hydrogen-bonding interactions with charged amino acids, or with two different ligands. Tan *et al.* have proposed that the basic region of rev takes up an α-helical structure and that 6 amino acids (Thr-34, Arg-35, Arg-38, Arg-39, Asn-40 and Arg-44) make specific contacts with the RNA (43). Our results are consistent with this proposal since an α-helix could easily bridge the distance between P103/P104 and P124/P125 and some of these identified arginines could indeed be responsible for such contacts.

The discovery that a single methylphosphonate isomer (*Rp*) at position P106 was impaired in rev binding ability is particularly informative. We interpret this result to indicate that the pro-*R* oxygen atom of this phosphate makes a hydrogen bond with rev

either to an uncharged amino acid side chain (e.g. Thr-34 or Asn-40) or alternatively to a main chain amide functionality. This phosphate was not previously predicted by ethylation interference mapping as a site of rev interaction.

Surprisingly, we have not obtained any evidence for direct contacts between rev and the five residues in the 'bubble'. Each of the functional groups tested on these bases is either non-essential or appears to be involved in base-pairing. By contrast, major groove recognition of the RNA duplex occurs at the N^7-positions of both purines flanking the bubble (G_{104} and G_{128}), but this does not extend to more distant purines. Thus, many of the specific contacts made by rev to the high affinity binding site appear to be with phosphates rather than to the bases.

CONCLUSION

Studies of the interactions of tat and rev with RNA have provided new insights into the chemistry of nucleic acid recognition. There are several examples of RNA binding proteins that recognize bases displayed in single-stranded loop and bulge structures (45). By contrast, both tat and rev associate primarily with functional groups on Watson–Crick base-pairs that are exposed within the major groove of a distorted double-stranded RNA (9, 12, 20). In addition to base-specific contacts, our methylphosphonate mapping shows that both proteins make specific phosphate contacts on both strands of the flanking RNA duplexes. The data on both tat binding to TAR and rev binding to the RRE suggest that there are numerous contacts made between the amino acid side chains of tat and rev and functional groups displayed in the RNA binding sites. In the absence of high-resolution structural data, chemical modification studies, such as those described here are providing our most precise images of the structures of these two important RNA–protein complexes.

ACKNOWLEDGEMENTS

We thank Terry Smith, Jan Fogg and Richard Grenfell for help with synthesis of oligoribonucleotides, Christina Lamont and Sheila Green for help in preparation of tat protein, Anne Kelley for provision of rev protein and Gabriele Varani and Jo Butler for helpful discussions. We are very grateful to Professor Fritz Eckstein (Max-Planck-Institüt für Experimentelle Medizin, Göttingen) for provision of 2′-deoxy-2′-fluoroadenosine and its corresponding phosphoramidite and to Dr Nanda Sinha (Millipore Corporation, Bedford, Massachusetts) for 2′-deoxuridine methylphosphonoamidite. This work was supported by a grant from the MRC AIDS Directed Programme.

REFERENCES

1. Karn, J. (1991) *Current Opinion in Immunol.*, **3**, 526–536.
2. Gait, M.J. and Karn, J. (1993) *TIBS*, **18**, 255–259.
3. Cullen, B.R. (1986) *Cell*, **46**, 973–982.
4. Rosen, C.A., Sodroski, J.G. and Haseltine, W.A. (1985) *Cell*, **41**, 813–823.
5. Muesing, M.A., Smith, D.H. and Capon, D.J. (1987) *Cell*, **48**, 691–701.
6. Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Heaphy, S., Karn, J., Lowe, A.D., Singh, M., Skinner, M.A. and Valerio, R. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6925–6929.
7. Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Heaphy, S., Karn, J., Lowe, A.D., Singh, M. and Skinner, M.A. (1990) *EMBO J.*, **9**, 4145–4153.
8. Roy, S., Delling, U., Chen, C.-H., Rosen, C.A. and Sonenberg, N. (1990) *Genes & Devel.*, **4**, 1365–1373.
9. Weeks, K.M. and Crothers, D.M. (1991) *Cell*, **66**, 577–588.
10. Sumner-Smith, M., Roy, S., Barnett, R., Reid, L.S., Kuperman, R., Delling, U. and Sonenberg, N. (1991) *J. Virol.*, **65**, 5196–5202.
11. Delling, U., Reid, L.S., Barnett, R.W., Ma, M.Y.-X., Climie, S., Sumner-Smith, M. and Sonenberg, N. (1992) *J. Virol.*, **66**, 3018–3025.
12. Churcher, M., Lamont, C., Dingwall, C., Green, S.M., Lowe, A.D., Butler, P.J.G., Gait, M.J. and Karn, J. (1993) *J. Mol. Biol.*, **230**, 90–110.
13. Daly, T.J., Cook, K.S., Gary, G.S., Maione, T.E. and Rusche, J.R. (1989) *Nature*, **342**, 816–819.
14. 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.
15. Malim, M.H., Hauber, J., Le, S.-Y., Maizel, J.V. and Cullen, B.R. (1989) *Nature*, **338**, 254–257.
16. Malim, M.H., Tiley, L.S., McCarn, D.F., Rusche, J.R., Hauber, J. and Cullen, B.R. (1990) *Cell*, **60**, 675–683.
17. Zapp, M.L. and Green, M.R. (1989) *Nature*, **342**, 714–716.
18. Mann, D.A., Mikaélian, I., Zemmel, R., Green, S.M., Lowe, A.D., Kimura, T., Singh, M., Butler, P.J.G., Gait, M.J. and Karn, J. (1994) *J. Mol. Biol.*, in press.
19. Heaphy, S., Finch, J.T., Gait, M.J., Karn, J., and Singh, M. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 7366–7370.
20. Kjems, J., Brown, M., Chang, D.D. and Sharp, P.A. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 683–687.
21. Malim, M.H. and Cullen, B.R. (1991) *Cell*, **65**, 241–248.
22. Kjems, J., Calnan, B.J., Frankel, A.D. and Sharp, P.A. (1992) *EMBO J.*, **11**, 1119–1129.
23. Tiley, L.S., Malim, M.H., Tewary, H.K., Stockley, P.G. and Cullen B.R. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 758–762.
24. Bartel, D.P., Zapp, M.L., Green, M.R. and Szostak, J.W. (1991) *Cell*, **67**, 529–536.
25. Iwai, S., Pritchard, C., Mann, D.A., Karn, J. and Gait, M.J. (1992) *Nucl. Acids Res.*, **20**, 6465–6472.
26. Pavlevitch, N.P. and Pabo, C.O. (1991) *Science*, **252**, 809–817.
27. Feng, J.-A., Johnson, R.C. and Dickerson, R.E. (1994) *Science*, **263**, 348–355.
28. Calnan, B.J., Tidor, B., Biancalana, S., Hudson, D. and Frankel, A.D. (1991) *Science*, **252**, 1167–1171.
29. Tao, J. and Frankel, A.D. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 2723–2726.
30. Hamy, F., Asseline, U., Grasby, J., Iwai, S., Pritchard, C., Slim, G., Butler, P.J.G. Karn, J. and Gait, M.J. (1993) *J. Mol. Biol.*, **230**, 111–123.
31. Grasby, J.A., Butler, P.J.G. and Gait, M.J. (1993) *Nucl. Acids Res.*, **21**, 4444–4450.
32. Pieken, W.A., Olsen, D.B., Benseler, F., Aurup, H. and Eckstein, F. (1991) *Science*, **253**, 314–317.
33. Grasby, J.A., Pritchard, C.E., Mersmann, K., Seela, F. and Gait, M.J. (1993) *Coll. Czech. Chem. Commun.*, **58**, 154–157.
34. Karn, J., Churcher, M.J., Rittner, K., Kelley, A.C., Butler, P.J.G. , Mann, D.A. and Gait, M.J. (1994) In HIV: A Practical Approach, J. Karn (ed.), Oxford University Press, Oxford, in press.
35. Lebedev, A.V., Frauendorf, A., Vyazovkina, E.V. and Engels, J.W. (1993) *Tetrahedron*, **49**, 1043–1052.
36. Giver, L., Bartel, D., Zapp, M., Pawul, A., Green, M. and Ellington, A.D. (1993) *Nucl. Acids Res.* **21**, 5509–5516.
37. Jensen, K.B., Green, L., MacDougal-Waugh, S. and Tuerk, C. (1994) *J. Mol. Biol.*, **234**, 235–247.
38. Noble, S.A., Fisher, E.F. and Caruthers, M.H. (1984) *Nucl. Acids Res.* **12**, 3387–3404.
39. Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. and Williamson, J.R. (1992) *Science*, **257**, 76–80.
40. Puglisi, J.D., Chen, L., Frankel, A.D. and Williamson, J.R. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 3680–3684.
41. Tao, J. and Frankel, A.D. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 1571–1575.
42. Battiste, J.L., Tan, R., Frankel, A.D. and Williamson, J.R. (1994) *Biochemistry*, **33**, 2741–2747.
43. Leclerc, F., Cedergren, R. and Ellington, A.D. (1994) *Nature Structural Biol.*, **1**, 293–300.
44. Tan, R., Chen, L., Buettner, J.A., Hudson, D. and Frankel, A.D. (1993) *Cell*, **73**, 1031–1040.
45. Nagai, K. (1992) *Curr. Opinion Struct. Biol.*, **2**, 131–137.