JOURNAL OF VIROLOGY, June 1992, p. 3699–3706
0022-538X/92/063699-08$02.00/0

Vol. 66, No. 6

A Specific Sequence with a Bulged Guanosine Residue(s) in a Stem-Bulge-Stem Structure of Rev-Responsive Element RNA Is Required for *trans* Activation by Human Immunodeficiency Virus Type 1 Rev

STEVEN M. HOLLAND,† MARGARITA CHAVEZ, SUSAN GERSTBERGER, AND SUNDARARAJAN VENKATESAN*

Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892

Received 14 January 1992/Accepted 23 March 1992

We demonstrate that both the in vitro RNA binding and in vivo *trans* activation functions of human immunodeficiency virus type 1 Rev regulatory protein Rev require the presence of a 9-nucleotide 5′-CACUAUGGG-3′ RNA motif on its cognate target, the Rev-responsive element RNA. For optimal Rev recognition, this sequence must be presented as a stem-bulge-stem structure and must contain at least two G's, one of which must be unpaired, and include some or all of the CACUAU sequence upstream of the three G's. Distal mutations which result in the base pairing of the G's eliminate the Rev response. The first G is crucial, but changes at the other G's are tolerated if at least one G is unpaired. The secondary structure or the three-dimensional orientation of the B1 and B2 stem-loops of the Rev-responsive element are not relevant as long as the 5′-CACUAUGGG-3′ sequence is preserved, with at least one bulged G residue.

Regulatory proteins interact with target DNA in a sequence-specific manner over one or both strands of double-stranded DNA (23, 30). In contrast, the targeting of RNA binding proteins appears less well defined, because single-stranded RNA self-anneals into multiple conformations that may or may not expose the critical sequence for protein binding (43). Although several sequence-specific RNA-protein interactions have been identified, the sequence elements are presumed to be recognized in the context of an RNA secondary structure (7, 16, 26, 34).

Human immunodeficiency virus (HIV) has evolved complex genetic regulatory schemes that are excellent paradigms for mechanisms involving RNA-protein interactions (11, 36). HIV replication is dependent upon the function of two small RNA binding proteins, Tat and Rev. Tat-mediated transcriptional enhancement from the HIV long terminal repeat (LTR) involves Tat recognition of a critical 3-nucleotide bulge sequence embedded in the stem-loop structure of a 22-nucleotide *trans*-acting-responsive region element at the 5′ end of nascent HIV transcript (5, 6, 19, 35, 39). Rev modulates splicing (8), extranuclear transport (17, 18, 27), and/or translational utilization (1, 2) of unspliced or partially spliced viral RNAs containing a highly structured 244-nucleotide sequence referred to as the Rev-responsive element (RRE) in the *env* open reading frame. These actions of Rev are contingent upon specific binding to the RRE (9, 12, 13, 15, 21, 41).

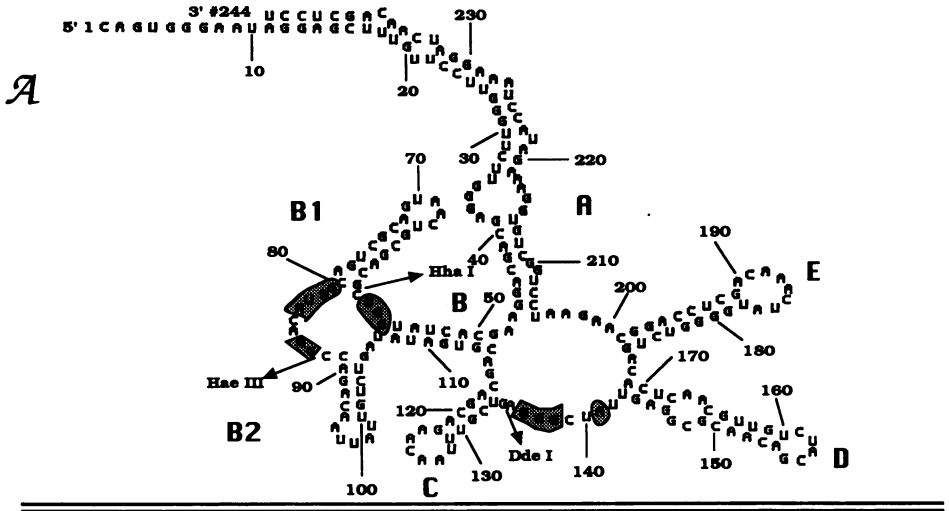
The computer-predicted secondary structure of the RRE RNA, composed of four stem-loops (A, C, D, and E) and a branched stem-loop (B/B1/B2) surrounding a central loop sequence (Fig. 1), has been largely confirmed experimentally (24, 37). Large portions of the RRE were demonstrated to be dispensable for a Rev response, and a minimal structure composed of the branched stem-loop B/B1/B2 and part of the incoming stem A retained a Rev response both in vitro and in vivo (15, 21, 28, 32). B/B1/B2 alone was capable of in vitro Rev binding (10, 12, 20, 21, 28), and mutations in this region which obliterate Rev binding of the complete RRE have the same phenotype in the context of the B/B1/B2 subdomain (21).

Although there is consensus about the identity of the branched stem-loop B/B1/B2 as the critical subdomain for the Rev response, disagreements exist as to whether Rev recognizes a unique sequence within this region in certain conformational contexts or recognizes a preferred RNA secondary structure irrespective of the sequence content. We have previously identified a 12-nucleotide 5′(50)CAC UAUGGGCGC(61)3′ RRE subsequence as the binding domain for Rev and further suggested that the two bulged guanylate residues may represent a critical contact point for Rev (21). This 12-nucleotide sequence is conserved in all HIV type 1 (HIV-1) isolates except for the substitution of G for U at position 55, and the distal 8 nucleotides of this sequence are also conserved in many of the HIV type 2 and simian immunodeficiency virus isolates. Recent in vitro studies (3) have suggested that Rev recognizes unique distortions of the sugar phosphate backbone of the RNA imposed by noncanonical base-pairing schemes rather than a specific sequence, but our studies propose a 5′-CACUA UGGG-3′ sequence containing at least 1 bulged G nucleotide as the minimal Rev recognition sequence both in vitro and in vivo.

MATERIALS AND METHODS

Construction of HIV-1 *gag* expression plasmids containing RRE derivatives. The construction of a Rev-responsive, HIV-1 LTR-linked wild-type (wt) RRE containing Gag reporter plasmid pRMK4-RRE has been described previously (21). The RRE fragment was cloned at the *Asp*I site of M13mp18 replicative form. M13-RRE phage DNA was used

* Corresponding author.
† Present address: Laboratory of Host Defenses, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892.



B

Sequence Change	In vitro REV Binding	p24 GAG ELISA % of wt RRE (SD)	
		+ REV	− REV
RRE 1-244	++++	100	2.2 (1.4)
ERR (Reverse RRE) 244'-1'	−	2.9 (1.8)	1.6 (0.7)
B/B1/B2 49-115	+++	49 (7.8)	12.0 (2.8)
B/B1/B2 + A 31-115/204-221	ND	44 (7.0)	8.1 (4.4)
B/B1/B2 + a 39-115/204-213	+++	84 (3.8)	9.9 (5.9)
B/B1/B2 + a 3'-5' 204-213/49-115/39-49	+++	44 (6.8)	5.7(4.8)
MS2 -B/B1/B2 3As/MS2/3As /49-115	ND	137 (31.8)	23.7 (9.5)
B/3G/MS2/B2 49-58/MS2/83-115	ND	24 (6.5)	7.7(2.5)
B/3G/MS2/B2 + a 39-58/MS2/83-115/204-213	+++	53 (16.8)	10.2(6.8)

FIG. 1. (A) RRE RNA secondary structure predicted by the Zuker algorithm (42, 43) and verified by nuclease digestions (16). The shaded areas denote the nucleotides shielded from nucleases by Rev binding (38). (B) In vitro and in vivo behavior of RRE subdomains. Rev-responsive *gag* expression values were from eight independent experiments with two plasmid preparations for all but the last three plasmids. The last three plasmids were transfected five times. B/B1/B2 refers to the *gag* expression plasmid containing the B/B1/B2 subdomain (nucleotides 49 to 115). MS2-B/B1/B2 denotes a construct which placed the MS2 operator stem-loop sequence upstream of the B/B1/B2 subdomain of RRE. In B/3G/MS2/B2, the B1 stem-loop was replaced by the MS2 sequence, but the three G's were left intact. B/3G/MS2/B2+a denotes the addition of the RRE stem A sequence to B/3G/MS2/B2. Different portions of the stem A sequence were added to the various mutants, in the order indicated by the nucleotide numbers under each mutant. Levels of Rev binding are as described in Materials and Methods. SD, standard deviation.

to engineer mutations by using a commercial mutagenesis protocol (Muta-Gene; Bio-Rad Laboratories). The mutated RREs were recovered from the M13 replicative form DNA by *Asp*I digestion and exchanged for the wt RRE in the pRMK4-RRE vector. To construct (MS2)+B/B1/B2, the B/B1/B2 subdomain (positions 49 to 115) of the RRE was initially cloned into M13mp18. The MS2 sequence flanked by three A's was then inserted upstream of the B/B1/B2 motif. M13 containing B/B1/B2 was also mutagenized to replace the B1 stem-loop with the MS2 sequence (5'-GGUUGAG-GAUUACCCAACC-3'). To generate B/3G/MS2/B2+a, RRE sequences between positions 31 and 49 and 204 and 221 were inserted upstream and downstream, respectively, of the B/3G/MS2/B2 recombinant.

RNA synthesis and Rev binding. T7-promoter-tagged RRE templates were generated by polymerase chain reaction with primer pairs corresponding to the ends of the RRE, except that the 5′ primer was tagged with the bacteriophage T7 RNA polymerase promoter (21). T7-promoter-tagged, polymerase chain reaction-amplified RRE DNA fragments (0.5 pmol) were used as templates to generate unlabeled or uniformly labeled ([α-^{32}P]UMP) RRE transcripts by using a commercial kit from Promega Biotec (Milwaukee, Wis.). All RNAs were purified by gel filtration on Sephadex G-100 columns followed by urea–10% acrylamide gel electrophoresis, if necessary. The Rev-binding affinities of the different RNAs were quantitated from saturation binding curves by nitrocellulose filter binding assay (21). wt RRE binding to Rev was expressed as ++++, 100% of wt binding; +++, 50 to 60% of wt binding; ++, 20% of wt binding; +, 10% of wt binding; and −, levels of binding 2 orders of magnitude less than that of wt RRE. Rev used in these experiments was expressed in *Escherichia coli* and purified to near homogeneity (40).

Transient expression assay. Rev-responsive *gag* expression was measured under transient conditions. Duplicate sets of individual *gag*-RRE plasmids (4 μg) with or without HIV-1 LTR-linked Rev plasmid (pHIV-REV, 2 μg) were cotransfected with pHIVLTR-TAT (1.5 μg) and pHIVLTR-CAT (2 μg) into HeLa cells by using the calcium phosphate method. The cells were harvested between 48 and 72 h posttransfection and disrupted by three freeze-thaw cycles, and the extracts were clarified by low-speed centrifugation. *gag* expression was quantitated by p24 enzyme-linked immunosorbent assay (ELISA) of cell extracts and clarified culture fluid (Coulter Diagnostics). p24^{gag} expression levels were normalized to the constant values of chloramphenicol acetyltransferase (CAT) activity expressed from HIV LTR-linked CAT (21). We determined the absolute values of *gag* production both in the presence and absence of Rev. Instead of expressing the mutant values in terms of relative fold activation by Rev, we have tabulated the absolute values obtained for each mutant in the presence of Rev in relation to the values obtained with the wt RRE plasmid. Our method enabled us to develop a consistent basal line, allowing the detection of marginal Rev responders.

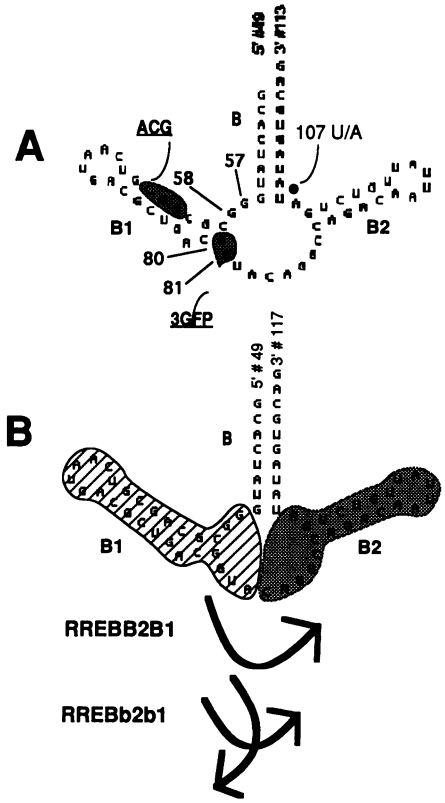
RESULTS

Rev recognizes a modified subdomain of the RRE both in vitro and in vivo. To examine the Rev response in vivo, *gag* expression from HIV-1 LTR-linked *gag* plasmids containing different RRE derivatives, positioned downstream of the *pol* open reading frame, was measured in the presence or absence of Rev coexpression. Rev-responsive *gag* expression was correlated with the Rev-binding potential of the respective RRE derivatives by in vitro filter binding assays. A *gag* plasmid containing just the B/B1/B2 subdomain (bases 49 to 113) of the RRE was about half as Rev responsive as the 244-nucleotide wt RRE (Fig. 1). The Rev response of a B/B1/B2 dimer was not notably superior to that of the monomer, in contrast to the results of a recent report (22). When stem B was extended by adding varying lengths of RRE stem A either in the sense configuration (B/B1/B2+A and B/B1/B2+a) or in an inverse complementary configuration (B/B1/B2+a 3′-5′), we observed slight improvements in the Rev response which sometimes approached the levels obtained with the full RRE (Fig. 1). The primary structure of the added components was not crucial as long as they formed a stable stem. In support of this, a recombinant which extended the B stem of B/B1/B2 by the addition of the MS2 phage translational operator stem sequence (MS2+B/B1/B2) had an even better Rev response than the RRE itself (Fig. 1). Curiously, both the B/B1/B2 and MS2+B/B1/B2 constructs had high basal levels of *gag* expression which were further enhanced by Rev.

Structure-compensated mutations in the B1 stem-loop preserved the Rev response (15, 21, 28, 32), and portions of the B1 stem-loop may be excised without the loss of Rev recognition (15, 21). Further, the entire B1 stem-loop in the B/B1/B2 may be replaced by the MS2 operator sequence, and the Rev response of the resulting chimera, B/3G/MS2/B2, which retained the three G's at nucleotides 56 to 58 was about 50% of that of B/B1/B2. Extending the B stem of this chimera by adding the RRE stem A sequence created B/3G/MS2/B2+a, with a moderately improved Rev response (Fig. 1). The MS2 operator sequence which replaced the RRE sequence between nucleotides 58 and 83 formed a Watson-Crick base-paired helix. These results and the earlier mutational analyses of the B2 helix (15, 21, 32) suggested that the sequence between nucleotides 58 and 83 and 90 and 114 may be largely dispensable for Rev recognition.

A subsequence within the B/B1/B2 subdomain containing bulged guanylates is recognized by Rev in different potential conformations when at least one G is unpaired. We have shown before that a 12-nucleotide 5′(50)CACUAUGG GCGC(61)3′ sequence within the B/B1/B2 subdomain may contain a binding domain for Rev and have suggested that two of the three G residues in this sequence in a bulged structure may represent a critical contact point for Rev (21). Kjems et al. (24) have demonstrated that Rev specifically shielded discrete regions of the RRE containing bulged G residues flanked by a stem sequence from chemical or RNase digestion. By using partial RNase digestions of RRE RNA, we too have demonstrated that Rev shielded the three G's at nucleotides 56 to 58, two G's at 80 to 81, three G's at 136 to 138, and the A's at 83 and 141 (Fig. 1). Rev did not shield mutant RRE RNAs unable to bind or to be *trans* activated by Rev (38). These data argued for the requirement of the bulged G structures of the RRE in Rev recognition.

Mutation of the sequence 62-AGC-64 to ACG in the B1 stem eliminated both in vitro Rev binding and the in vivo Rev response (Fig. 2). Computer-predicted RNA folding of the ACG mutant showing the G at position 56 (56 G) paired with 107 U, 57 G paired with 89 C, and 58 G paired with 88 C was verified by partial RNase digestions (not shown). This suggested that base pairing of all three G's may be incompatible with Rev recognition. To address this issue more directly, we constructed a mutant (3GFP) in which the two G residues at nucleotides 80 and 81 were changed to -CC- to facilitate the potential base pairing of 57 G with 81 C and 58 G with 80 C. This construct was also devoid of Rev binding and an in vivo Rev response. In contrast, a U-to-A mutant at 107 (107 U/A) which predicted that the three G's at nucleotides 56 to 58 would remain unpaired, had a wt Rev response (Fig. 2). Since in the wt RRE, G at 56 wobbles with U at 107, only one or two of these G's may have to be unpaired for Rev recognition. We then examined the conformational constraints on the interaction of the three-G motif with Rev. A mutant (RREBB2B1) which flipped the orientation of the B1 (nucleotides 57 to 83) and B2 (nucleotides 84 to 106) stem-loops (nucleotides 56 to 103) with respect to the B stem was nearly half as responsive as Rev. In RREBB2B1, the secondary structures of the B1 and B2 stem-loops were preserved and the two G's at nucleotides 57 to 58 remain unpaired, albeit at a different reference point with respect to



Sequence Change	In vitro REV Binding	p24 GAG ELISA % of wt RRE (SD)	
		REV +	REV −
A			
ACG AGC to ACG (62-64)	-	3.0 (1.5)	12.9 (4.1)
3GFP CGG to CCC (79-81)	-	3.0 (2)	8.6 (3.8)
107 U/A AUA to AAA (103-105)	++++	103 (12)	5.6 (2.7)
35 UCC GGGAG to UCCAG (35-39)	++++	91 (9.9)	9.3 (3.5)
35 UCCΔA GGGAG to UCCG (35-39)	++++	94.5(12.6)	7.8 (3.1)
67 ΔCAA GUCAAUGA to GUUGA (67-72)	++++	79 (8.9)	7.7 (2.2)
B			
RREBB2B1 1-56/106-84/83-57/107-244	ND	44 (3.3)	11.1 (4.2)
RREBb2b1 1-56/84-106/57-83/107-244	ND	11	6.2 (2.7)

FIG. 2. (A) Schematic diagram of distal mutations which force the three G's in the B1 stem to base pair (ACG and 3GFP) or remain unpaired (104 U/A). The mutations were engineered in the full RRE (27), although the figure illustrates only the B/B1/B2 subdomain. Mutations that change the three G's in the A stem to -UCC- (35 UCC) and delete the A at 38 (35 UCCΔA) are not illustrated. (B) Mutants which rotate the B1 and B2 stem-loops around the B stem. All mutants were constructed in the context of the full RRE. The numbers under each mutant refer to the respective coordinates within the RRE sequence. The in vitro Rev binding and the in vivo Rev response of the mutants are tabulated on the right. Results are from eight independent experiments for all mutants except RREBb2b1, which gave poor yields and was genetically unstable; therefore, only three transfections were undertaken. The experimental conditions were as described in Materials and Methods. HeLa cells were cotransfected with the *gag*-RRE, HIV LTR-CAT, and HIV Tat plasmids with or without the HIV LTR-linked Rev expression plasmid by using the calcium phosphate procedure. Nonresponders to Rev are in boldface type.

stem B. We also constructed a flip-flop mutant, RREBb2b1, which rotated and inverted the B1 and B2 stem-loops. The flip-flop mutant had a negligible Rev response (Fig. 2). While the transposed G's at nucleotides 57 to 58 were maintained in a bulge sequence in RREBB2B1, near the 3′ strand of stem B, RREBb2b1 was predicted to fold into a novel secondary structure with no bulged G residues.

Functional analysis of base substitutions at the three guanylates in the B/B1/B2 subdomain of RRE. Bulged nucleotides have been reported to play a critical role in protein binding to RNA. To examine the contribution of each of the three G's in the Rev recognition sequence, we designed a comprehensive set of mutants in this region (Table 1). The insertion of an additional G (GGGG) was of no consequence, but an AGGG mutant had a reduced Rev response. Deleting one of the G's (GG) resulted in a structure with a single G in a bulge sequence, which exhibited a Rev response significantly reduced (44%) compared with wt levels. The deletion of two or all three G's resulted in a complete loss of binding and *trans* activatability (Table 1). Homopolymeric substitutions of the three G's (AAA, CCC, and UUU) preserved the RRE secondary structure but were nonresponsive to Rev (21). Individual positions in the three-G bulge were substituted with the other nucleotides either singly or in combinations. The replacement of 56 G (AGG, CGG, and UGG) led to modest improvement in basal expression but a drastic reduction of Rev response (Table 1). Mutation of 57 G to A or C (GAG and GCG) gave 67 and 90% of wt activity, respectively, but 57 G-to-U (GUG) transversion led to losses of binding and *trans*-activation. The substitution of 58 G with A or U (GGA and GGU) resulted in a slight reduction of the wt phenotype; however, GGC was markedly debilitated for both binding and the in vivo Rev response. The substitution of both G's at nucleotides 56 to 57 (AAG, CCG and UUG) was incompatible with the Rev response. The replacement of 57-GG-58 with AA or UU (GAA and GUU) resulted in a modest diminution of wt activity. Substitutions at both the proximal and distal G's led to a modest (UGU, 60%) or a significant (AGA, 34%; CGC, 21%) decrease in the Rev response.

Although homopolymeric substitutions and deletions of the three G's resulted in a predictable loss of Rev binding, the behavior of the mutants with changes at one or two of the three G's was not as straightforward. Whereas substitutions at the third G (GGA and GGU) were tolerated (with the exception of GGC), changes at the first and double mutations at the first and second G's were not as forgiving. Solitary changes at the second G yielded mixed results. While GAG and GCG retained wt Rev responses, GUG was consistently negative. The nonfunctional GUG RNA was predicted to

TABLE 1. In vitro and in vivo behavior of RRE mutants at the three G's[a]

Sequence change[b]	In vitro Rev binding[c]	$p24^{gag}$ ELISA (% of wt RRE [SD])[d]	
		Rev+	Rev−
GGGG	++++	86 (6.8)	8.5 (3.1)
AGGG	++	38 (12)	7.8 (4.6)
GG	ND	44 (14)	10.1 (3.7)
1G	−	8.0 (5)	4.0 (1.5)
0G	−	7.0 (2.8)	2.2 (0.9)
AAA	−	8 (3.0)	4.3 (1.1)
CCC	−	11 (3.7)	5.1 (2.8)
UUU	−	3.6 (0.9)	1.8 (1.4)
AGG	+	20 (8.8)	11 (5.8)
CGG	+	12 (5.8)	9.8 (3.3)
UGG	+	10 (4.8)	10.4 (2.8)
GAG	++	67 (9.8)	11.4 (6.6)
GCG	++++	90 (10)	13.6 (5.8)
GUG	−	14 (6.2)	12.3 (4.1)
GGA	++++	76 (9.7)	6.8 (4.4)
GGC	++	15 (6.6)	7.7 (2.8)
GGU	++++	62 (11)	14.3 (6.8)
AAG	+	18 (9.8)	10 (3.2)
CCG	+	13 (4.6)	11 (2.1)
UUG	+	8 (4.8)	8.7 (3.8)
GAA	+++	36 (8.7)	9.8 (5.3)
GCC	ND	ND	ND
GUU	+++	33 (6.1)	8.4 (4.2)
AGA	+	34 (7.8)	10.1 (3.2)
CGC	++	21 (5.3)	8.8 (1.2)
UGU	ND	60 (7.5)	17 (11.2)

[a] All mutations were engineered into the full RRE.
[b] Nonresponders and marginal responders to Rev are in boldface type and underlined, respectively.
[c] In vitro Rev binding levels as explained in Materials and Methods were quantitated by filter binding (21). ND, not determined.
[d] In vivo Rev responses were determined by transfections with calcium phosphate. Values are from 12 independent experiments with two or three plasmid preparations. SD, standard deviation; Rev+, with Rev plasmid; Rev−, without Rev plasmid; ND, not determined.

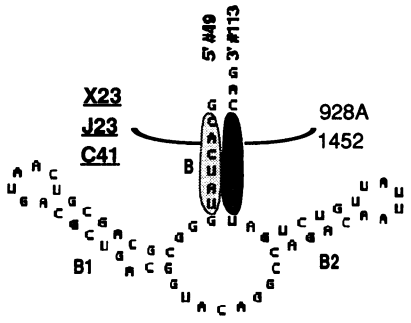
fold into a secondary structure, wherein the G's at nucleotides 56 and 58 remained base paired. Similarly, the differential responses of the other G substitutions probably relate to the secondary structure context of the remaining G's in the respective mutations.

Contribution of the stem B sequence upstream of the three G's for Rev recognition. Mutational analysis of the B stem (49-GCACUAU-55 paired with 114-CGUGAUA-108) has detected a potential site for Rev recognition. Deletion of the 5′ strand of the B stem abolished Rev recognition, while removal of the 3′ strand reduced but did not eliminate the Rev response (21). Since deletions may distort the secondary structure and inadvertently mask an authentic binding site, base substitutions of the B stem were examined. Mutations that exchanged the 50-CACUAU-55 sequence for its complementary (X23) or inverse complementary (J23) sequence or an unrelated sequence (C41), which each disrupted the B stem, were nonresponsive to Rev (Fig. 3) (21). To examine whether changes in the 3′ strand result in a similar phenotype, we introduced a complementary (928A) or an inverse complementary (1452) mutation in the 3′ strand. Both of these mutants, with a predicted disruption of the B stem but an intact 5′ sequence (50-CACUAU-55), maintained Rev responsiveness at levels of 79% (928A) and 100% (1452) of that of wt RRE (Fig. 3). Changes in the 5′ strand (X23) and the 3′ strand (928A and 1452) of stem B resulted in high levels of basal expression. Rev expression reduced the basal levels of X23 but enhanced expression from 928A and 1452. These observations, together with our earlier report showing that mutants with structure-compensated strand inversions of the 6 bp from the B stem were also unresponsive to Rev (21), suggested that a Rev-sensitive site was present in the 5′ strand of the B stem. It is possible that the Rev response requires only a few bases upstream of the three G's rather than the entire CACAUAU sequence.

DISCUSSION

Our direct approach of comparing the in vitro Rev binding and the in vivo Rev response of several discrete RRE mutants demonstrates that RRE RNA secondary structure alone is insufficient for Rev recognition; instead, a unique 50-CACUAUGGG-58 sequence is the functional determinant. Since we have not verified the predicted RNA secondary structures for all of our mutations, it is possible that some or all of our mutations may have disrupted the structure in unpredicted ways. Within these limitations, some conclusions may be drawn from our study. The 50-CACUA UGGG-58 sequence is recognized by Rev even when packaged as a chimeric RNA with the MS2 sequence replacing the B1 stem-loop (B/3G/MS2/B2) or as a conformational mutant with B1 and B2 stem-loops rotated by 180°C (RREBB2B1). The importance of the three G's was corroborated by the loss of Rev binding and the in vivo Rev response associated with the homopolymeric substitutions at the three G's of the B/3G/MS2/B2 construct. The requirement of unpaired G's at nucleotides 56 to 58 for Rev recognition was enunciated by the phenotype of the ACG mutation at nucleotides 62 to 64. This mutation, which altered the RNA sequence at a site distal to the three G's at nucleotides 56 to 58, folded into a structure with base-paired G's and was nonresponsive to Rev. Extensive mutations of the B1 stem-loop have suggested that this region may have no sequence-specific information (15, 21, 28, 32) and that portions of this stem-loop may be excised (as in 67 ΔCAA, Fig. 2) without seriously impairing the Rev response (15, 21). Therefore, the loss of Rev binding and function by the ACG mutation was unlikely to be due to the alteration of the sequence at nucleotides 62 to 64 but may have resulted from the forced pairing of the three G's at nucleotides 56 to 58. The Rev-responsive phenotype of the conformational mutant, RREBB2B1, which reoriented the B1 and B2 stem-loops but presumably left two of the three G's unpaired, further supported the importance of the bulged G's for the Rev response. The behavior of the RREBB2B1 mutant was reminiscent of that of the GG mutant which emphasized the requirement of at least two G's, one of which should be unpaired.

Kjems et al. (24) have suggested that Rev binding of RRE may be mediated by multiple sites on the RNA. However, the loss of the Rev responses of mutants in the CACU AUGGG sequence implies that this one Rev binding site is essential for RRE *trans* activation. The other sites identified by chemical or ribonuclease protection may not represent primary functional sites. Our 3GFP mutation was at one of these redundant sites (-CGG- at nucleotides 79 to 81) and



Sequence Change	In vitro REV Binding	p24 GAG ELISA % of wt RRE (SD)	
		REV +	REV −
X23 CACUAU/GUGAUA (50-55)	−	11 (3.7)	26.8 (7.7)
J23 CACUAU/AUAGUG (50-55)	−	3.6 (0.9)	12.9 (3.9)
C41 CACUAU/CCCGGG (50-55)	−	3 (0.7)	4.9 (2.3)
928A AUAGUG/UAUCAC (108-113)	ND	79 (10.2)	22.3 (6.6)
1452 AUAGUG/CACUAU (108-113)	ND	100 (9.1)	28.4 (9.8)

FIG. 3. RRE mutations in the 5′ and 3′ strands of stem B that disrupt base pairing are illustrated in the context of the B/B1/B2 structure. In vitro Rev binding and the in vivo Rev response were determined as described in Materials and Methods. HeLa cells were cotransfected with the indicated *gag*-RRE plasmids, HIV LTR-CAT, and HIV Tat with or without the Rev plasmid. The normalization of data and statistical analysis are as described previously (21). SD, standard deviation. Nonresponders to Rev are in boldface type.

was devoid of the Rev response. However, this was probably due to the forced pairing of the G's at nucleotides 56 to 58 in the 3GFP mutant. The binding sites outside the B/B1/B2 domain are unable to mediate the Rev effect. Mutating one of these in stem A (35 UCC, 35UCCΔA, Fig. 2) did not diminish the Rev response. In contrast, an intact nucleotides 35 to 37 three-G sequence was not able to support Rev function in the context of mutations at the G's at nucleotides 56 to 58. This may be due to the low affinity of these ancillary sites or to an a priori requirement for binding of Rev to the sequence at nucleotides 50 to 58 prior to binding at the other sites.

The sequence-specific recognition of RRE RNA by Rev resembles that by other RNA binding proteins such as HIV Tat–(4–6, 19, 35), λ phage N protein (26), RNA phage coat protein (34), RNA virus capsid proteins (14, 29), ribosomal proteins (16), small nuclear ribonucleoproteins (31, 33), and iron-responsive element-binding proteins (7), all of which recognize discrete cognate sequences with a bulged nucleotide(s) in the context of an RNA stem or stem-loop structure. Our analysis has underscored the importance of bulged G's, but the paucity of mutants precludes a clear statement as to whether Rev has a stringent requirement for a conserved primary structure in the B stem. Changes in the 3′ strand of stem B (928A and 1452) are allowed for the Rev response at the expense of stem disruption, but structure-compensating mutations were not permitted in the 5′ strand (21). Our observations, which emphasize the requirement of the 5′CACUAU3′ sequence of stem B, are somewhat at odds with those of the in vitro studies showing that a generic guanylate-rich bubble flanked by a nonspecific stem sequence constitutes the Rev recognition motif (20). Rev binding to generic guanylate bubbles under saturating conditions probably reflects a core reaction which may be further strengthened by the inclusion of stem B components, and it is doubtful whether the simple guanylate bubbles can support Rev function in vivo. Therefore, we propose that the CACUAUGGG sequence presented as a stem-bulge-stem structure is the functional determinant for Rev. Our proposals are compatible with the recent finding of a 38-nucleotide RRE fragment between positions 50 and 87 as the irreducible RRE motif capable of sustaining Rev binding (10) and the Rev response in an in vitro splice inhibition assay (25). 5′ truncation of the 38-nucleotide sequence deletes the 50-CACU-53 sequence, and 3′ truncation may induce the G's at nucleotides 56 to 58 to become base paired. Although the RRE subdomain between positions 50 and 87 appears to function well in the in vitro splice inhibition assay (25), we have been unable to demonstrate an in vivo Rev response for a GAG expression vector containing this subsequence (21; unpublished data). It is possible that in the context of the large *gag-pol* mRNA in vivo, the subsequence above may not fold appropriately for Rev binding to the critical primary structure.

Our postulates have to be reconciled with a recent claim that the Rev-response element is a stem-bulge-stem structure containing many conserved bases of the B1 stem-loop and stabilized by a noncanonical G-G or a homopurine base pair(s) and that Rev recognizes distortions of the sugar phosphate backbone imposed by these base pairs (3). Among our three G mutants, substitutions at the second (GAG and GCG), the third (GGA and GGU), or the second and third positions (GAA and GUU) were remarkably Rev responsive, although they obliterated the putative noncanonical pairing schemes (3). A mutant which replaced the B1 stem-loop sequence (between nucleotides 58 and 83) with a Watson-Crick base-paired MS2 operator sequence (B/3G/MS2/B2+a) was Rev responsive in vitro and in vivo. Deletion of most of the B1 stem-loop (positions 65 to 82) reduced but did not eliminate the Rev response (21). These observations are inconsistent with the recent proposal (3), and from our data we predict that single-base changes in the 3′ strand

of the B1 bulge will lose the Rev response if they lead to base pairing of the three G's, as in the 3GFP mutant. The noncanonical base-pair model was based on an elegant random mutagenesis scheme (3), but numerical arguments show that the authors could not have tested a majority of the two base changes in RRE and may have missed many of the subtle changes that we tested which preserved the Rev response. Further, Bartel et al. (3) examined the Rev binding of RRE RNAs randomly mutated at multiple sites and tended to ignore the effects of multiple mutations on the RNA secondary structure. The model also predicts a novel Watson-Crick base-paired stem II D structure between nucleotides upstream of the three G's with the bases across the B1/B2 bulge. On the basis of the covariant analysis, the authors have proposed that for Rev binding, changes in the CACUAU motif must be accompanied by compensatory changes across the bulge to facilitate canonical base pairing of stem II D (3). Whereas this may be a valid explanation for some of our null mutants in the B stem, other mutants (X23 and 23X) which preserved alternative Watson-Crick pairing schemes for generating the putative stem II D were still devoid of a Rev response both in vitro and in vivo.

Finally, it is as yet unclear whether any host factors are required for in vivo Rev function in the context of HIV RNA. These factors may be RRE-binding cellular proteins or Rev-binding factors or ancillary helper factors. Candidate proteins for some of these activities have been tentatively identified (36). Whereas the role(s) of these factors in Rev function remains to be determined, it is possible that unidentified cellular factors may substitute for Rev or function as helper factors. Among our library of *gag*-RRE mutants, some were constitutively expressed in the absence of Rev. Coexpression of Rev further enhanced *gag* expression from some of the constitutive mutants, but other constitutive mutants were repressed by Rev. For instance, X23, which was nonresponsive to Rev both in vitro and in vivo, had a high level of basal expression which was repressed by Rev in vivo. X23 was probably *trans* activated in vivo by a cellular Rev surrogate, and Rev coexpression may have squelched limiting amounts of a common nuclear helper factor required for the resolution of RRE-Rev or RRE-nuclear factor complexes. Further studies with the constitutive mutants may enlighten us as to some of the complex interactions between cellular Rev surrogates and helper factors.

ACKNOWLEDGMENTS

We thank Alicia Buckler-White for oligonucleotide synthesis. We are grateful to Paul Wingfield of the Protein Expression Laboratory, Office of the Director, NIH, for the purified Rev protein expressed in *E. coli*. We thank former research fellows Nafees Ahmad and Ratan K. Maitra for preliminary experiments. We thank Damien Purcell and Frank Malderilli of LMM, NIAID, for critically reviewing the manuscript.

This work was supported in part by the Intramural AIDS targeted Anti-Viral program administered by the Office of the Director, National Institutes of Health.

REFERENCES

1. **Arrigo, S. J., and I. S. Chen.** 1991. Rev is necessary for translation but not cytoplasmic accumulation of HIV-1 vif, vpr, and env/vpu 2 RNAs. Genes Dev. **5**:808–819.
2. **Arrigo, S. J., S. Weitsman, J. D. Rosenblatt, and I. S. Chen.** 1989. Analysis of *rev* gene function on human immunodeficiency virus type 1 replication in lymphoid cells by using a quantitative polymerase chain reaction method. J. Virol. **63**:4875–4881.
3. **Bartel, D., M. Zapp, M. Green, and J. W. Szostak.** 1991. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. Cell **67**:529–536.
4. **Berkhout, B., A. Gatignol, J. Silver, and K. Jeang.** 1990. Efficient trans-activation by the HIV-2 Tat protein requires a duplicated TAR RNA structure. Nucleic Acids Res. **18**:1839–1846.
5. **Berkhout, B., R. H. Silverman, and K. T. Jeang.** 1989. Tat trans-activates the human immunodeficiency virus through a nascent RNA target. Cell **59**:273–282.
6. **Calnan, B. J., S. Biancalana, D. Hudson, and A. D. Frankel.** 1991. Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. Genes Dev. **5**:201–210.
7. **Casey, J. L., D. M. Koeller, V. C. Ramin, R. D. Klausner, and J. B. Harford.** 1989. Iron regulation of transferrin receptor mRNA levels requires iron-responsive elements and a rapid turnover determinant in the 3′ untranslated region of the mRNA. EMBO J. **8**:3693–3699.
8. **Chang, D. D., and P. A. Sharp.** 1989. Regulation by HIV Rev depends upon recognition of splice sites. Cell **59**:789–795.
9. **Cochrane, A., C. Chen, and C. Rosen.** 1990. Specific interaction of the human immunodeficiency virus Rev protein with a structured region in the env mRNA. Proc. Natl. Acad. Sci. USA **87**:1198–1202.
10. **Cook, K. S., G. J. Fisk, J. Hauber, N. Usman, T. J. Daly, and J. R. Rusche.** 1991. Characterization of HIV-1 REV protein: binding stoichiometry and minimal RNA substrate. Nucleic Acids Res. **19**:1577–1583.
11. **Cullen, B. R.** 1991. Regulation of HIV-1 gene expression. FASEB J. **5**:2361–2368.
12. **Daefler, S., M. E. Klotman, and S. F. Wong.** 1990. Trans-activating rev protein of the human immunodeficiency virus 1 interacts directly and specifically with its target RNA. Proc. Natl. Acad. Sci. USA **87**:4571–4575.
13. **Daly, T., K. Cook, G. Gray, T. Maione, and J. R. Rusche.** 1989. Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element in vitro. Nature (London) **342**:816–819.
14. **Darlix, J., C. Gabus, M. Nugeyre, F. Clavel, and F. Barre-Sinoussi.** 1990. *cis* elements and *trans*-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. J. Mol. Biol. **216**:689–699.
15. **Dayton, E., D. Powell, and A. Dayton.** 1989. Functional analysis of CAR, the target sequence for the Rev protein of HIV-1. Science **246**:1625–1629.
16. **Douthwaite, S., T. Powers, J. Y. Lee, and H. F. Noller.** 1989. Defining the structural requirements for a helix in 23 S ribosomal RNA that confers erythromycin resistance. J. Mol. Biol. **209**:655–665.
17. **Emerman, M., R. Vazeur, and K. Peden.** 1989. The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. Cell **57**:1155–1165.
18. **Felber, B., M. Hadzopoulou-Cladaras, C. Cladaras, T. A. Athanasopoulos, and G. Pavlakis.** 1989. Rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. Proc. Natl. Acad. Sci. USA **86**:1495–1499.
19. **Feng, S., and E. Holland.** 1988. HIV-1 tat trans-activation requires the loop sequence within the TAR. Nature (London) **334**:165–167.
20. **Heaphy, S., J. T. Finch, M. J. Gait, J. Karn, and M. Singh.** 1991. Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich "bubble" located within the rev-responsive region of viral mRNAs. Proc. Natl. Acad. Sci. USA **88**:7366–7370.
21. **Holland, S. M., N. Ahmad, R. K. Maitra, P. Wingfield, and S. Venkatesan.** 1990. Human immunodeficiency virus rev protein recognizes a target sequence in rev-responsive element RNA within the context of RNA secondary structure. J. Virol. **64**:5966–5975.
22. **Huang, X., T. J. Hope, B. L. Bond, D. McDonald, K. Grahl, and T. G. Parslow.** 1991. Minimal Rev-response element for type 1 human immunodeficiency virus. J. Virol. **65**:2131–2134.
23. **Johnson, P., and S. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58**:799–839.
24. **Kjems, J., M. Brown, D. D. Chang, and P. A. Sharp.** 1991.

Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element. Proc. Natl. Acad. Sci. USA **88**:683–687.

25. **Kjems, J., A. D. Frankel, and P. A. Sharp.** 1991. Specific regulation of mRNA splicing in vitro by a peptide from HIV-1 Rev. Cell **67**:169–178.
26. **Lazinski, D., E. Grzadzielska, and A. Das.** 1989. Sequence-specific recognition of RNA hairpins by bacteriophage antitermination requires a conserved arginine-rich motif. Cell **59**:207–218.
27. **Malim, M., J. Hauber, S. Le, J. Maizel, and B. Cullen.** 1989. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature (London) **338**:254–257.
28. **Malim, M. H., L. S. Tiley, D. F. McCarn, J. R. Rusche, J. Hauber, and B. R. Cullen.** 1990. HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. Cell **60**:675–683.
29. **Marquet, R., F. Baudin, C. Gabus, J. Darlix, M. Mougel, C. Ehresmann, and B. Ehresmann.** 1991. Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanisms. Nucleic Acids Res. **19**:2349–2357.
30. **Mitchell, P., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245**:371–378.
31. **Nelson, K., and M. Green.** 1989. Mammalian U2 snRNP has a sequence-specific RNA-binding activity. Genes Dev. **3**:1562–1571.
32. **Olsen, H., P. Nelbock, A. Cochrane, and C. Rosen.** 1990. Secondary structure is the major determinant for interaction of HIV rev protein with RNA. Science **247**:845–849.
33. **Query, C., R. Bentley, and J. Keene.** 1989. A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. Cell **57**:89–101.
34. **Romaniuk, P., P. Lowary, H. Wu, G. Stormo, and O. Uhlenbeck.** 1987. RNA binding site of R17 coat protein. Biochemistry **26**:1563–1568.
35. **Roy, S., U. Delling, C. H. Chen, C. A. Rosen, and N. Sonenberg.** 1990. A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. Genes Dev. **4**:1365–1373.
36. **Vaishnav, Y. N., and S. F. Wong.** 1991. The biochemistry of AIDS. Annu. Rev. Biochem. **60**:577–630.
37. **Venkatesan, S., S. M. Holland, N. Ahmad, P. Wingfield, R. K. Maitra, and H. C. Lane.** 1991. Molecular functional studies of HIV-1 REV and NEF proteins, p. 189–201. *In* A. Kumar (ed.), Advances in molecular biology and targeted treatment of AIDS. Plenum Press, New York.
38. **Venkatesan, S., S. M. Holland, A. R. Nebreda, R. K. Maitra, N. Ahmad, H. Park, P. Wingfield, and E. Santos.** 1991. Regulation of the HIV genome: functional studies of REV, NEF and *cis* regulatory elements, p. 323–344. *In* W. A. Haseltine and F. Wong-Staal (ed.), Genetic structure and regulation of HIV. Raven Press Ltd., New York.
39. **Weeks, K. M., and D. M. Crothers.** 1991. RNA recognition by Tat-derived peptides: interaction in the major groove? Cell **66**:577–588.
40. **Wingfield, P. T., S. J. Stahl, M. A. Payton, S. Venkatesan, M. Misra, and A. C. Steven.** 1991. HIV-1 Rev expressed in recombinant *Escherichia coli*: purification, polymerization, and conformational properties. Biochemistry **30**:7527–7534.
41. **Zapp, M., and M. Green.** 1989. Sequence-specific RNA binding by the HIV-1 Rev protein. Nature (London) **342**:714–716.
42. **Zuker, M.** 1989. Computer prediction of RNA structure. Methods Enzymol. **180**:262–288.
43. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244**:48–52.