JOURNAL OF VIROLOGY, Feb. 1992, p. 1139–1151
0022-538X/92/021139-13$02.00/0

Vol. 66, No. 2

Extensive Sequence-Specific Information throughout the CAR/RRE, the Target Sequence of the Human Immunodeficiency Virus Type 1 Rev Protein

ELAHE T. DAYTON,[1] DANIELLE A. M. KONINGS,[2]† DOUGLAS M. POWELL,[1,3] BRUCE A. SHAPIRO,[2] LUCA BUTINI,[1]‡ JACOB V. MAIZEL,[2] AND ANDREW I. DAYTON[1]*

Laboratory of Immunoregulation, Building 10, Room 11C112, National Institute of Allergy and Infectious Diseases, 9000 Rockville Pike, Bethesda, Maryland 20892[1]; Laboratory of Mathematical Biology, Division of Cancer Biology & Diagnosis, National Cancer Institute, Frederick, Maryland 21701[2]; and Genetics Department, George Washington University, Washington, D.C. 20052[3]

Received 11 June 1991/Accepted 4 November 1991

The significance and location of sequence-specific information in the CAR/RRE, the target sequence for the Rev protein of the human immunodeficiency virus type 1 (HIV-1), have been controversial. We present here a comprehensive experimental and computational approach combining mutational analysis, phylogenetic comparison, and thermodynamic structure calculations with a systematic strategy for distinguishing sequence-specific information from secondary structural information. A target sequence analog was designed to have a secondary structure identical to that of the wild type but a sequence that differs from that of the wild type at every position. This analog was inactive. By exchanging fragments between the wild-type sequence and the inactive analog, we were able to detect an unexpectedly extensive distribution of sequence specificity throughout the CAR/RRE. The analysis enabled us to identify a critically important sequence-specific region, region IIb in the Rev-binding domain, strongly supports a proposed base-pairing interaction in this location, and places forceful constraints on mechanisms of Rev action. The generalized approach presented can be applied to other systems.

Human immunodeficiency virus type 1 (HIV-1), the etiologic agent of AIDS, is a complex retrovirus whose life cycle is characterized by a period of regulated latency that underlies the clinical course of the disease and its insidious spread throughout the population (for reviews, see references 18 and 72). In tissue culture, the key viral mediator of the emergence of HIV from latency and the transition from early to late viral synthesis following acute infection is the *rev* axis of HIV autoregulation (for a review, see reference 8). As this apparatus regulates the cytoplasmic accumulation of virion genomic and structural mRNAs (2, 17, 20, 35, 38, 44, 52) and is required for completion of the virus life cycle (62, 65), it represents both a window onto the workings of critical cellular processes and a legitimate therapeutic target.

Central to the *rev* axis of viral regulation is the RNA target sequence for the viral Rev protein called CAR (*cis* antirepressor [12, 14, 55, 62]) or RRE (Rev-responsive element [20, 23, 44]) located in an approximately 220-bp region of *env*, just 3′ to the amino terminus of the gp41 coding region (2, 12, 23, 24, 29, 44, 55). Computational and mutational analysis has determined it to consist of a long central stem, I/I′ (Fig. 1A), from which branches a series of stem-loops, II through VI (14, 40, 44, 45). Interaction of the CAR/RRE with the Rev protein in the nucleus (9, 44, 51, 68) overcomes the constitutive repression of viral genomic and structural RNA expression mediated by the viral *cis* repressor sequences (55) to directly stabilize and promote export to the cytoplasm of viral genomic and structural mRNAs or to directly inhibit splicing of viral messages (5, 17, 19, 20, 24, 38, 44).

The extent and significance of sequence information in the CAR/RRE have been controversial. In theory, the information content of the CAR/RRE should be partitioned between sequence and higher-order structure. Suggestions that the CAR/RRE is sequence-specific have frequently been based on findings that antisense CAR/RRE RNA neither binds Rev protein in vitro nor functions in vivo (6, 10, 26, 49, 73). Antisense RNA may adopt a structure different from that of sense RNA. RNA folding utilizes non-Watson-Crick base pairing that is not conserved in the conversion from sense to antisense. Even when the antisense molecules adopt the same set of secondary structures as the sense, the secondary structures will differ in their disposition with respect to the 5′-3′ orientation of the molecule, except for the degenerate case in which they lie symmetrically distributed about a dyad. In fact, antisense RNA structures often retain much of the sense sequence-specific information content in the helical stem regions but present it in a location different from that of the sense molecule. On the other hand, the considerable tolerance the CAR/RRE shows to sequence alterations has suggested a lack of sequence specificity in the recognition events (6, 49). Whether this tolerance derives from a lack of sequence recognition or from the presence of sequence redundancy has not previously been addressed. The ability of Rev to oligomerize with itself and with CAR/RRE RNA in vitro (6, 10, 28, 45, 47–49) suggests the possibility of multiple Rev-binding sites and underscores the possibility of redundancy. To date, only one structure-conserving sequence-altering mutation has been reported to strongly inhibit CAR/RRE function. This is a mutation of the entire stem IIa (28), which also inhibits Rev binding. A systematic examination of the rest of the molecule has not yet been reported.

The structure requirements for CAR/RRE function are only partly understood. Disruption of any of the stem

* Corresponding author.
† Present address: Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309.
‡ Present address: Servizio di Immunologia Clinica, Istituto di Clinica Medica, Ancona, Italy.

structures of the CAR/RRE severely impairs function, although individual removal of many of them is often without effect (3, 6, 12, 14, 28, 45, 49). The element retains considerable in vivo function in the absence of the stem-loop IV through VI regions. Under most circumstances, it does not retain function in the absence of either stem-loop region II or III (3, 6, 12–14, 28, 45, 49). These latter stem-loop regions are the predominant, if not exclusive, locus of Rev binding in band-shift and protection assays (3, 6, 10, 26, 28, 45, 49, 69, 73).

To assemble a coherent picture of the interactions among the CAR/RRE and the proteins it binds, Rev and (presumably) cellular factors, it is necessary to analyze the distribution of sequence-specific information as well as the location of structure-specific information in the molecule. Towards this end, we present here a comprehensive experimental and computational approach combining mutational analysis, phylogenetic comparison, and thermodynamic structure calculations with a systematic strategy for distinguishing sequence-specific information from secondary structural information. This generalized approach can be applied to other systems.

(D.M.P. contributed to this work in partial fulfillment of the Ph.D. requirements of the Graduate Genetics Program at the George Washington University.)

MATERIALS AND METHODS

Plasmid design and preparation. The Rev-dependent indicator plasmid, III-*gag*-CAR, described previously (14) consists of a 367-bp CAR/RRE fragment of HXB-2 placed in a "gutted" HXB-2 proviral construct in PTZ19R (U.S. Biochemicals) to create a plasmid with the general structure long terminal repeat (LTR)-*gag-prot*-CAR-LTR. A simian virus 40 replication origin elsewhere on the plasmid drives amplification in COS cells. As a source of *rev*, another simian virus 40 *ori*-containing plasmid, pSV-CMV-*rev*-C, was used. This vector is similar to pSV-CMV-*rev* previously described except for the insertion of a 21-base sequence from upstream of the *tat* initiation codon, GTCGACAGAGGAGAGCAAGAA, to boost translation efficiency (58). A polymerase error during this cloning introduced a single extra C-terminal Cys on the Rev protein. All of the plasmids used in this work were isolated on Qiagen columns, by using the kits provided by the manufacturer. Their concentrations were determined spectrophotometrically and confirmed by agarose gel electrophoresis of restriction digests of carefully measured aliquots.

Construction of mutants. Mutants were constructed by either the polymerase chain reaction assembly and ligation of short oligodeoxynucleotides or site-directed mutagenesis in the context of the entire expression vector, III-*gag*-CAR, as described previously (14). To avoid the potential artifact of unanticipated second-site mutations outside of the CAR/RRE, all mutant CAR/RRE fragments made by direct mutagenesis of III-*gag*-CAR were eventually transferred by subcloning into a virgin III-*gag*-CAR vector which had never been through in vitro mutagenesis. All of these transferred mutations were tested to confirm the results of the initial transfections of untransferred mutations. The proviral mutants were made by first mutagenizing a short *Bgl*II-*Hin*dIII proviral fragment containing the CAR/RRE. This short fragment was then transferred into the provirus by a series of subcloning steps. The sequence of the entire mutagenized fragment was confirmed after its transfer into the provirus, as were the sequences across each restriction site used in the transfers.

Sequence confirmation of mutations. The sequence of the entire minimal CAR/RRE region, as determined previously (14), of each mutant was confirmed by dideoxy sequencing. The identity of each plasmid preparation was also checked by sequencing to avoid inadvertent switches.

Cells, transfections, and assays. COS-1 cells in six-well plates were transfected as described previously (14) with 0.5 μg of indicator plasmid with or without either 0.25 or 0.5 μg of SV-CMV-*rev*-C as a source of Rev. For this series of experiments, 5×10^5 cells were plated per well. Cell monolayers were harvested and tested for $p24^{gag}$ by enzyme-linked immunosorbent assay as described previously (14). The results presented have not been normalized to either cotransfected CAT expression plasmids or recovered protein. Attempts to do so consistently resulted in greater variability of the results of replicate cultures (data not shown). In our hands, variability was most effectively reduced by careful technique and multiple experiments.

Statistical analysis of Rev response ratios. Wild-type values determined on each of 13 different days were centered at 100%, and corresponding shifts were made in the experimental data. Control values were then pooled, and normality was confirmed by the Shapiro-Wilk test (60). Data collected for each mutant were pooled, and the mean value was compared with the control (wild type) mean by a one-tailed *t* test. Comparisons between pairs of mutants also utilized the *t* test.

Design of mutants. The HIV-1 wild-type sequence and all the mutants presented here, both base-switched and point mutants (though not the substitution mutants denoted by SUB), were studied in detail for features present in their secondary structures. Optimal and suboptimal foldings of the sequences were calculated through minimization of free energy by using the algorithm described by Zuker (74). For each sequence, large sets of suboptimal foldings were generated with different window sizes (i.e., distance of structural difference between subsequently generated suboptimal foldings) ranging from 0 to 10, including all structures having a folding energy within 10% of the minimum. The consideration of suboptimal foldings is essential in terms of checking for possible alternative configurations, from both a methodological and a biological point of view. This analysis was performed by using the improved set of free-energy values given by Jaeger et al. (33). The design of the base-switched CAR/RREs followed in principle the forward complement of the wild type. However, this way of designing mutants needs additional alterations for the occurrence of new, undesirable base-pairing schemes possible in the forward complement sequences. New base-pairing interactions can be due to G-U pairings in the wild type which would be changed to the nonpairing C-A or to new potential G-U pairings from C's and A's in the wild type. The G-U base pairings present in the CAR/RRE were switched to either a C-G, U-A, or U-G pair, depending on the required stability and potential interference elsewhere at each location, to prevent undesired interactions. Base switching only limited regions of the molecule can similarly present extensive new pairing interactions. All such potential alternatives detected were able to be eliminated by changing essential parts of single-stranded regions to nucleotides other than their Watson-Crick pairing partners, the final purpose of these mutants being to preserve structure but change base order. It should be noted that by studying optimal and suboptimal folding of designed mutants by free-energy minimization, one at least can eval-

uate the most likely base-pairing schemes with reference to those of the wild-type sequence. All potential base-pairing interactions of a particular mutant that are certainly not desired can subsequently be avoided by modifying the sequence within the limits of the planned mutational test.

Phylogenetic analysis of structure. For the phylogenetic comparison of the multi-stem-loop regions II and III, CAR/RRE sequences of primate immunodeficiency viruses present in the data base were compared by the multiple-sequence alignment method TRIALS (27) and by individual minimal-energy folding (as described above). Similarity patterns obtained with the two independent analyses, the one of primary structure and the other of secondary structure, were then compared. The comparison included 25 HIV-1 (40), 6 HIV-2, and 13 simian immunodeficiency virus (SIV) CAR/RRE sequences.

RESULTS

Sequence-specific information is widely distributed throughout the CAR/RRE. Considering reports that attributed only a minor functional role to sequence specificity in the CAR/RRE (6, 49), we first asked if there were any sequence-specific information at all by analyzing a target sequence analog, base-switched CAR/RRE. "Base switching" is a term we propose to represent the following procedure. First, every base pair is switched and every unpaired base is changed to its Watson-Crick pairing partner according to the secondary structure predicted by previous analyses (14, 40, 45) (Fig. 1B). This results in a sequence which differs from the forward complement of the wild type only at the positions of atypical base pairs. Second, extensive computational analysis (see Materials and Methods) is performed to search for new alternative conformations which might result from new sequence patterns. Additional base substitutions are then introduced to eliminate alternative structures which might compete with the wild-type secondary structures. Thus, base switching retains the same secondary structure as that of the wild type yet presents a new base at every position. Alterations in function caused by base switching thus can result only from alterations in information other than secondary structure. Accordingly, we operationally define sequence-specific information as any kind of non-secondary structure information.

The base-switched CAR/RRE has a predicted secondary structure identical to that of the wild-type CAR/RRE but a sequence that differs from that of the wild type at every position. This base-switched CAR/RRE, BS(CAR/RRE), was inserted into our *gag* expression plasmid and tested for *rev* responsiveness in COS cells. The levels of *gag* expression supported in the absence and presence of Rev are presented in Table 1, and the ratios of expression (as a percentage of the wild-type ratio) are summarized in Fig. 1 and Table 1. The base-switched CAR/RRE was almost totally inactive (Fig. 1B and Table 1). This is consistent with the presence of sequence specificity in the CAR/RRE.

To localize the sequence specificity, we then instituted a systematic exchange of fragments between the wild-type CAR/RRE and the inactive base-switched CAR/RRE. Each of the exchange mutants presented was also computationally analyzed and altered, if necessary, to eliminate potential alternative structures which might compete with the wild-type structures. The logical first step, a reciprocal exchange of the Rev-binding domains (II and III) produced two analogs, each with impaired function. Base switching everywhere but domains II and III, BS(I/I', IV, V, & VI) (Fig. 1C), reduced function to about 31% of that of the wild type. Base switching just domains II and III, BS(II+III) (Fig. 1D), completely eliminated function. This is consistent with sequence-specific information lying both inside and outside the Rev-binding domain.

TABLE 1. $p24^{gag}$ production by base-switched and substitution mutants of the CAR/RRE in the absence and presence of Rev

Mutation	$p24^{gag}$ production[a]		Rev response ± SE[b]	P[c]
	−Rev	+Rev		
Wild type	3	100	100 ± 2.3	
BS(CAR/RRE)	2	2	1 ± 0.4	<0.001
BS(I/I', IV, V, & VI)	3	37	31 ± 2.1	<0.001
BS(II+III)	3	4	5 ± 2.2	<0.001
BS(I/I')	2	137	157 ± 3.2	NS[d]
BS(V)	4	111	80 ± 5.0	<0.001
BS(VI)	5	126	58 ± 3.6	<0.001
BS(IIa)	3	62	48 ± 2.0	<0.001
BS(IIc+III)	3	20	16 ± 2.5	<0.001
BS(IIc)	4	89	63 ± 6.5	<0.001
BS(III)	4	61	50 ± 4.6	<0.001
BS(IIc&III-tips)	7	234	136 ± 18	NS
SUB7391–7393	2	2	2[e]	<0.001
SUB7394–7395	3	3	3 ± 0.2	<0.001
SUB7413–7421	4	2	1 ± 0.1	<0.001

[a] Production is expressed as a percentage of $p24^{gag}$ synthesized by the wild type stimulted by Rev and represents the averages of multiple experiments. This is appropriate for assessing whether baseline values (i.e., −Rev) are changed by the mutations.

[b] For the Rev response, the ratios of +Rev and −Rev for each mutant as a percentage of the wild-type ratio were determined for each experiment. The data presented are the averages of these ratios. This method of calculation is more appropriate for analyzing the Rev response and differs slightly from determining the ratio of the first two columns.

[c] *P*, probability that the mean is not less than that of the wild type.

[d] NS, not significant.

[e] Result of one experiment.

Analysis of subregions outside the Rev-binding domain has not yet localized the apparent sequence specificity. Base switching domain I/I', BS(I/I') (Fig. 1E), or V, BS(V) (Fig. 1F), did not impair function. Base switching domain VI, BS(VI) (Fig. 1G), partially impaired function.

A more detailed analysis of the Rev-binding domain suggested it too contains considerable sequence-specific information, much of which is redundant. Base switching either all of domain IIa (Fig. 1H) or all of III (Fig. 1J) reduced function to about 50% of normal. Base switching domain IIc reduced function somewhat less, to 63%. However, simultaneous base switching of domains IIc and III, BS(IIc+III) (Fig. 1I), reduced function to approximately 16% of that of the wild type. Simultaneously base switching the tips of the stems, BS(IIc&III-tips) (Fig. 1L), caused no impairment of function, confirming a lack of sequence specificity in the loop regions. In contrast to the results of some previous investigators (6, 45, 49), all multiple base substitutions (SUB) of either the 5' side of domain IIb, SUB7391–7393 and SUB7394–7395, or the 3' side, SUB(7413–7421) (Fig. 1M), eliminated function.

Phylogenetic conservation of base pairing in the IIb region. To further analyze the apparent sequence specificity in the *rev*-binding domain, it was necessary to understand as much as possible about its in vivo secondary structure. Imperfections in the helix we originally suggested for the IIb region have promoted a widespread view of the IIb region as being mostly single stranded, not base paired (42, 44, 49). To

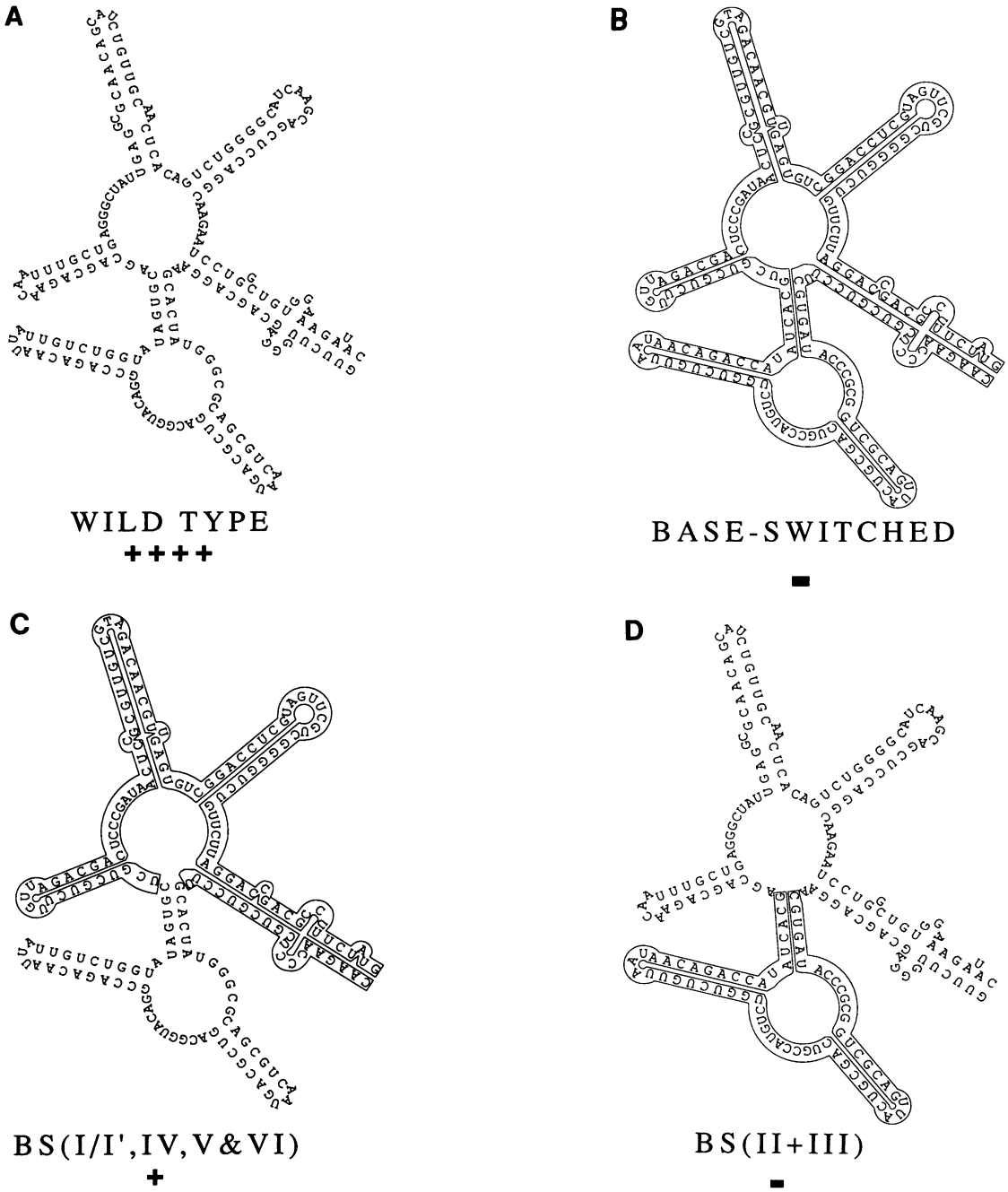


FIG. 1. Structure and Rev response of CAR/RRE mutants. (A through L) Individual mutants; (M) three separate mutants as indicated. Rev response data were obtained from Tables 1 and 2. All mutants are diagrammed only from HIV-1 position 7363 to 7559 (53). The CAR/RRE sequence used in all of the constructs started at position 7357. Wild-type 5′ ends thus are preceded by UCCUUGG. Base-switched 5′ ends are preceded by AGGAACC. BS(CAR/RRE) and BS(I/I′) terminated at position 7567. Thus, they are followed by GGUUUCCT. All wild-type 3′ ends were at position 7723. The data presented are the results of multiple transfection experiments. ++++, ≥80% function; +++, 55 to 80%; ++, 35 to 55%; +, 15 to 35%; +/−, 5 to 15%; −, ≤5%. BS, base switched; SUB, substitution.

resolve this issue, we performed a phylogenetic analysis of the conservation of base pairing in the IIb region. Twenty-five HIV-1 (40), 6 HIV-2, and 13 SIV CAR/RRE sequences were compared by multiple sequence alignment (27) of primary structure and by individual minimal-energy folding by using the suboptimal RNA folding program of Zuker (74) to consider a large set of alternative possible structures (39). The summary of the comparison of these two independent analyses for the IIb region is presented in Fig. 2.

The overlap of the CAR/RREs with a highly conserved domain of gp41 allowed for an unambiguous primary structure alignment. Two strikingly conserved sequence elements, AUGGG and CA, are shown in boldface in Fig. 2 to serve as reference points in the phylogenetic comparison.

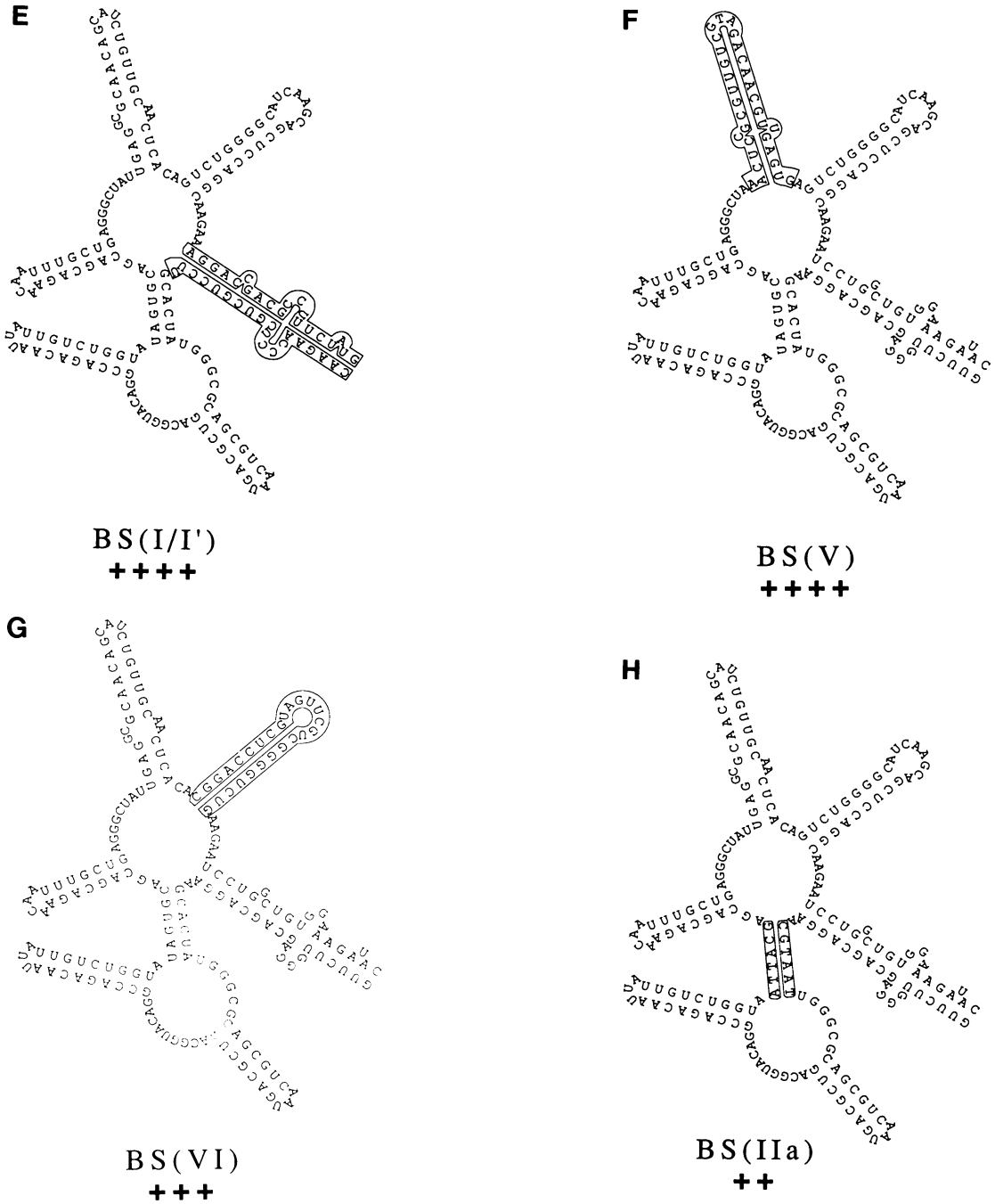


FIG. 1—*Continued.*

The individual structures presented all represent the optimal (minimal-energy) foldings of the depicted part of the CAR/RRE except for the proposed G-A base pairings. These odd base pairs are not treated as hydrogen bonded in the presently used free-energy rules. Taking into consideration the other possible base pairings in the multi-stem-loop of HIV-1, the base-pairing scheme proposed here for region IIb clearly fits both the phylogenetic data and the minimal-energy folding of HIV-1 the best (Fig. 2).

Analysis of point mutations in region IIb supports the proposed base-pairing scheme and demonstrates that sequence-specific information in region IIb is particularly critical. Consistent with the proposed region IIb base-pairing scheme, none of the mutants that retain full activity interrupt the proposed region IIb base pairing (Fig. 3A and Table 2). This includes mutations at positions 7413(?), 7414, 7417, 7421, 7442, and 7443. The information at these positions is, therefore, not base specific. Mutations 7392G-U, 7418A-U, 7419C-U, and 7420A-G conserve the proposed base pairing yet impair function. This strongly suggests that the information at these positions is base specific.

Also consistent with the proposed base pairing is the finding that all mutants which cause major disruptions in the pairing scheme severely impair function (Fig. 3B and Table

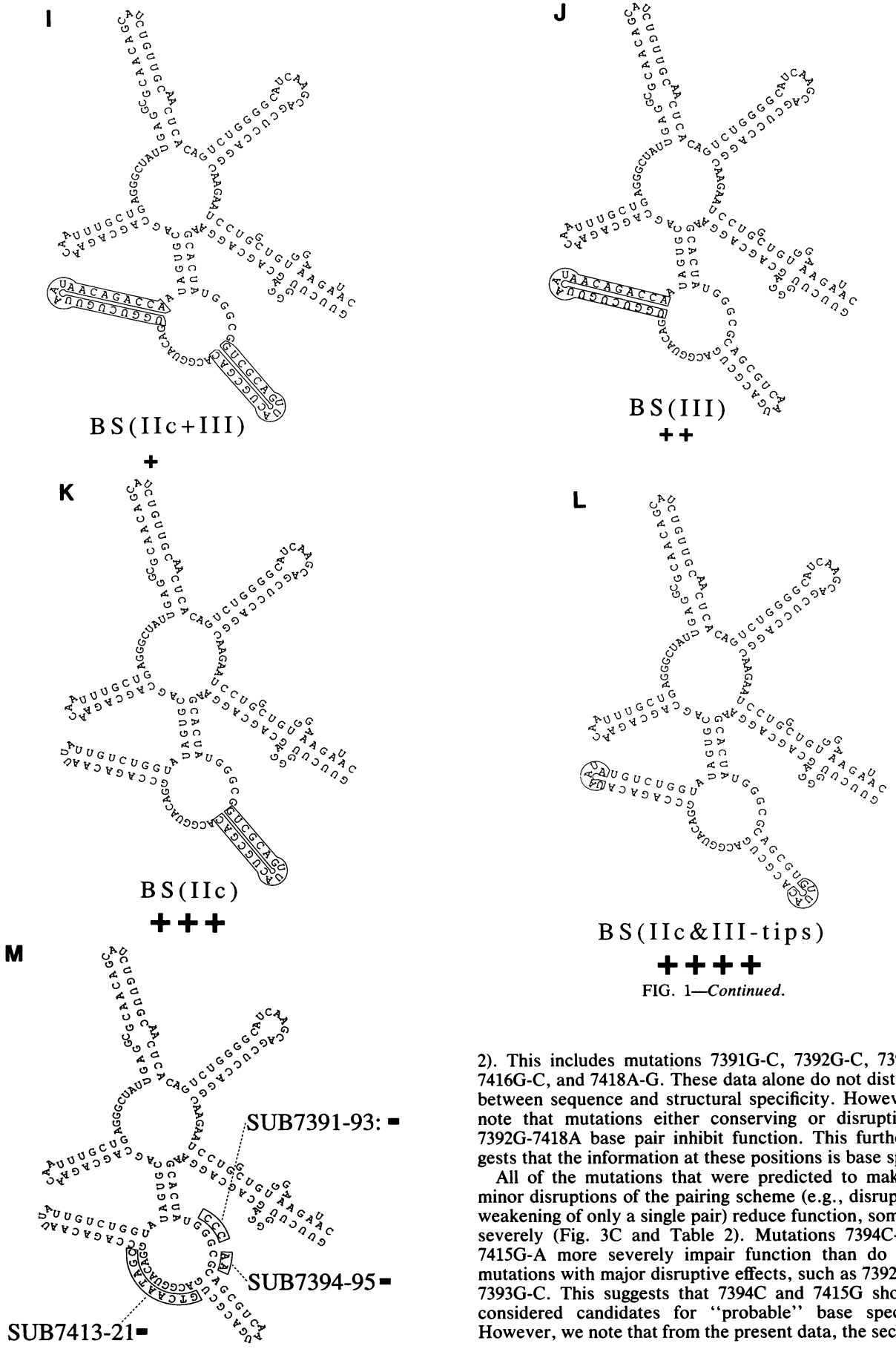


FIG. 1—*Continued.*

2). This includes mutations 7391G-C, 7392G-C, 7393G-C, 7416G-C, and 7418A-G. These data alone do not distinguish between sequence and structural specificity. However, we note that mutations either conserving or disrupting the 7392G-7418A base pair inhibit function. This further suggests that the information at these positions is base specific.

All of the mutations that were predicted to make only minor disruptions of the pairing scheme (e.g., disruption or weakening of only a single pair) reduce function, sometimes severely (Fig. 3C and Table 2). Mutations 7394C-A and 7415G-A more severely impair function than do nearby mutations with major disruptive effects, such as 7392G-C or 7393G-C. This suggests that 7394C and 7415G should be considered candidates for "probable" base specificity. However, we note that from the present data, the secondary

TABLE 2. $p24^{gag}$ production by point mutants of the CAR/RRE

Mutation	$p24^{gag}$ production[a]		Rev response ± SE[b]	P[c]
	−Rev	+Rev		
Wild type	3	100	100 ± 2.3	
7391G-C	3	5	4 ± 0.8	<0.001
7392G-U	3	45	51 ± 5.6	<0.001
7392G-C	2	27	39 ± 5.6	<0.001
7393G-C	3	17	16 ± 2.2	<0.001
7394C-G	1	1	6 ± 1.5	<0.001
7394C-A	3	7	6 ± 1.2	<0.001
7395G-A	3	54	47 ± 4.7	<0.001
7413A-G	3	86	80 ± 7.7	<0.001
7414C-T	2	92	105 ± 14	NS[d]
7415G-A	3	5	5 ± 0.5	<0.001
7415G-C	2	7	6 ± 1	<0.001
7416G-A	3	37	29 ± 5.6	<0.001
7416G-C	3	8	7 ± 0.8	<0.001
7417U-C	3	105	89 ± 6.4	<0.05
7418A-G	3	10	8 ± 2.4	<0.001
7418A-T	3	13	9 ± 1.8	<0.001
7419C-T	3	8	6 ± 0.5	<0.001
7420A-G	3	64	64 ± 7.2	<0.001
7421G-T	3	110	88 ± 20	NS
7442U-G	3	116	70 ± 5.1	<0.001
7443A-G	3	123	101 ± 12	NS

[a] See Table 1, footnote *a*.
[b] See Table 1, footnote *b*.
[c] *P*, probability that the mean is not less than that of the wild type.
[c] NS, not significant.

fine structure at these sites is not clear in terms of alternative base pairings.

Impairment of viral replication by a CAR/RRE structure-conserving point mutation in region IIb. To assess the validity of the HIV *gag* expression system used in our assays, we constructed proviral derivatives of HXB-2 with single (7418A-T), structure-conserving, or double nonreciprocal (7418A-T + 7415G-T) point mutations in the CAR/RRE. These point mutations were made in the third position of degenerate codons so as not to alter any envelope amino acids and were chosen to create new codons that are frequently used by HIV. Mutant and wild-type proviral clones were transfected into A3.01, a human $CD4^+$ T-cell line. Both proviral mutants exhibited major, though not absolute, impairment of growth in this system (Fig. 4). Both mutants also supported reduced levels of virion production after transfection into COS cells (data not shown). These data clearly validate the *gag* expression assay used in these studies as being relevant to HIV biology.

DISCUSSION

Extensive distribution of sequence specificity. The data presented here demonstrate that sequence specificity is surprisingly widespread throughout the CAR/RRE. What is the significance of this specificity? The base-switching approach used rigorously distinguishes between secondary structural and nonsecondary structural information. Formally, the latter includes base-specific contact points for both inter- and intramolecular interactions (e.g., regulatory factor binding and tertiary structure formation, respectively). Much of the sequence-specific information identified here is quantitatively and possibly mechanistically redundant and is located in base-paired helices. This information is, therefore, relatively unlikely to make critical contributions to tertiary structure. More likely, the sequence specificity derives from sequence recognition by proteins. This sequence recognition by proteins may involve direct contact with specific bases or recognition of specific features of the phosphate backbone which are determined by stacking interactions.

The distribution of sequence-specific information in the CAR/RRE Rev-binding domain, loosely considered regions II and III (26, 28, 45, 49), was unexpected. The data presented here suggest the existence of redundant information roughly evenly distributed among stems IIa, IIc, and III. Although these stems are in the Rev-binding domain, some of the sequence-specific information here may be required for the binding of critical cellular factors prior to the occupancy of the region by Rev. However, it is more likely that the sequence specificity of this region derives from sequence-specific recognition by Rev. The work of Holland et al. (28) showed that stem IIa is a critical binding site for Rev. Their suggestion that the recognition site extends also into stem IIb was, unfortunately, based on mutations that did not distinguish between sequence alterations and secondary structural alterations capable of disturbing neighboring binding sites. Furthermore, their failure to detect a requirement for stems IIc and III may derive from the existence of a 1-bp mismatch (bulge) found at the base of stem III in the pNL43 CAR/RRE used in their studies. The existence of this mismatch may render the pNL43 CAR/RRE partially defective to begin with. The protection data of Kjems et al. (37) suggested three binding sites in this region, each binding site consisting of one stem with additional material extending into part of stem IIb. These data did not determine, however, the relative importance of the binding sites. Nor did they distinguish between bases that were protected and bases that were required for binding and presumably function. The data of Cook et al. (7) suggested the minimal Rev-binding element consists of stems IIc and IIb and the 5′ strand of stem IIa. The data presented here demonstrate in vivo that stems IIa and IIc are roughly equivalent in significance and suggest that stem III may play a relatively secondary role in CAR/RRE function. Previous data have demonstrated that deletion of either stem IIc or III abolishes function (13, 14). Our data also support previous work suggesting that Rev needs to multimerize on the CAR/RRE to have function (43, 48). Entirely unexpected was the finding of sequence-specific information within bases in stem IIb which by protection analysis are not bound by Rev. We discuss this more fully immediately below.

Sequence specificity in the IIb region is unusually critical. The sensitivity of CAR/RRE function to structure-conserving mutations in the IIb region clearly demonstrates that this region is unique. Nowhere else in the molecule is sequence information so crucial. At least four bases in this region, 7392G, 7418A, 7419C, and 7420A, contain critical base-specific information (summarized in Fig. 5). Mutating any of these bases to bases which either disrupt or conserve the proposed base pairing reduces function. Interestingly, two of these bases, 7392G and 7418A, pair with each other. The information at positions 7418 and 7419 is particularly critical. Consistent with the exquisite sequence specificity of 7418A and 7419C, we note that with *gag* expression vectors, such as the one used here, all previously reported deletions of these bases in the CAR/RRE severely impair the Rev response (14, 28) or Rev binding (7). Two bases in this region, 7394C and 7415G (which may pair with one another), are listed as probably sequence specific (Fig. 5) because mutations can be made at these positions which, while only mildly

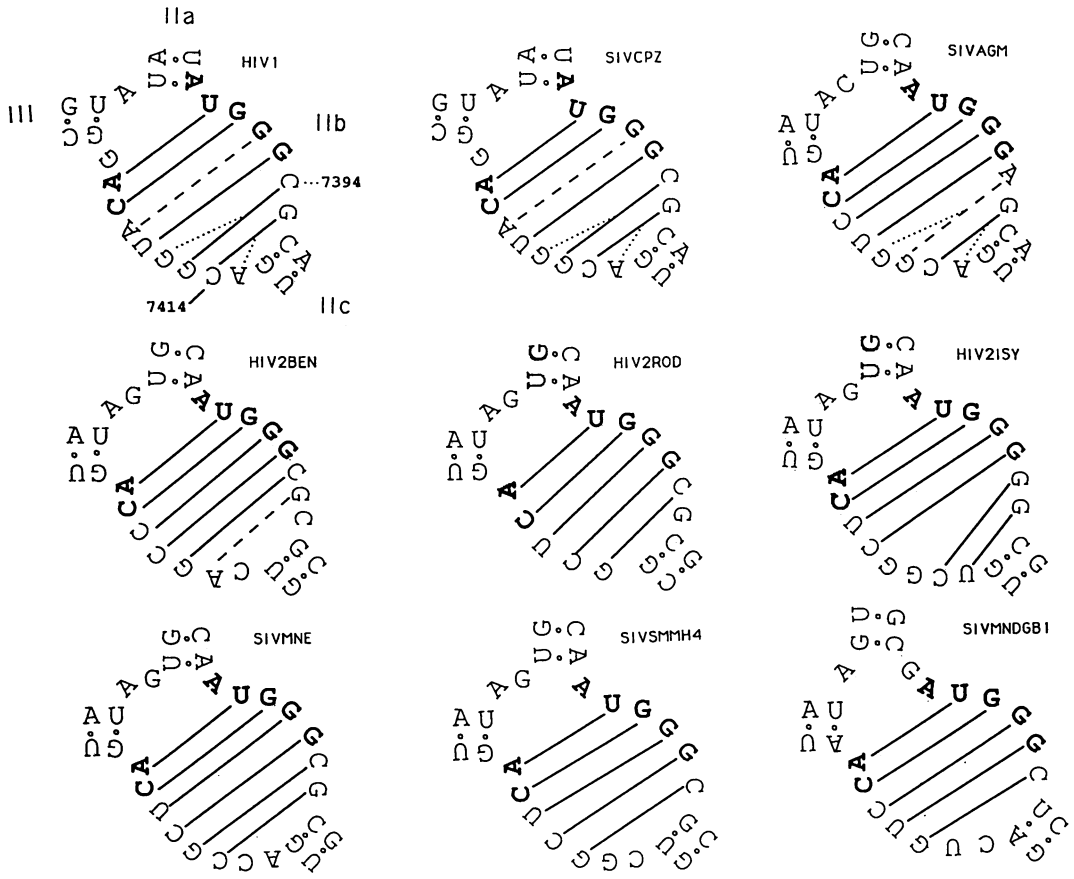
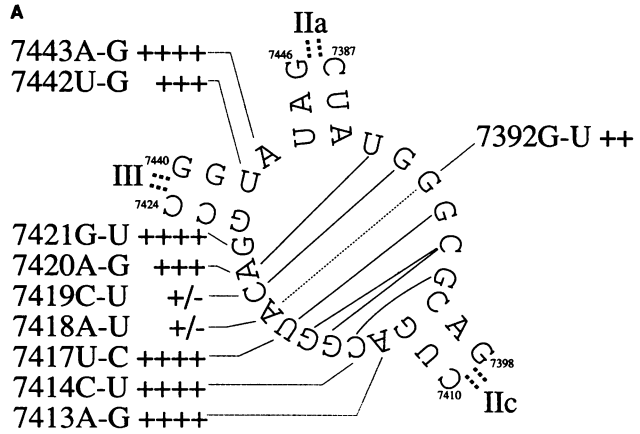
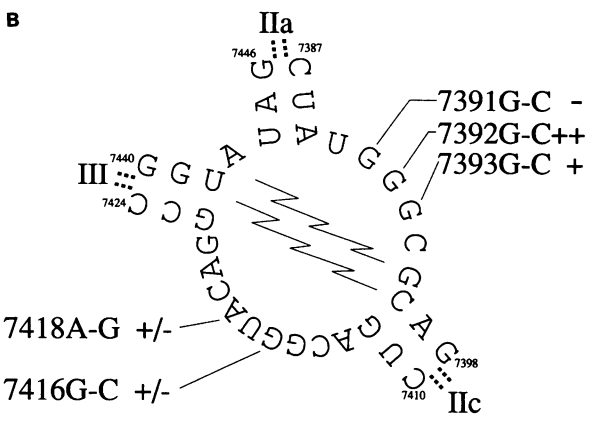


FIG. 2. Phylogenetic conservation of base pairing in the IIb region of the CAR/RRE of primate immunodeficiency viruses. All of the individual base-pairing schemes shown represent optimal free-energy foldings, except for the proposed G-A base pairings (see Materials and Methods and Discussion). The different pairing schemes of the HIV-2 and SIV strains shown represent the major variations of the theme of sequence and base-pairing patterns within these strains. The conserved sequence elements AUGGG and CA are shown as reference points. Dotted lines indicate alternative base pairings. On the basis of both phylogenetic data and known thermodynamic contributions, there is no strong reason to favor one of the two alternative C-G pairings (positions 7394 to 7415 or 7394 to 7416 in HIV-1). Furthermore, the possibility that in conjunction with a two-nucleotide bulge (positions 7414 and 7415 in HIV-1), the IIc stem is extended by an additional G-A base pair must be considered as another alternative conformation. This latter conformation is also reflected in the base-pairing schemes of the simian immunodeficiency viruses.

disruptive of the proposed structure, seriously impair the function. However, it is difficult to factor out sequence specificity from structural specificity with the mutations analyzed at these positions. Furthermore, the proposed fine structure at these points is unclear with respect to alternative base pairings. At least five positions, 7414, 7417, 7421, 7442, and 7443, in and around stem IIb have no or minimal base-specific information. Position 7413 may also lack significant base specificity. Structure-conserving substitutions can be made at these positions with no or minimal effect on function. The presence or absence of base specificity at positions 7390, 7391, 7393, 7395, 7413, and 7416 remains to be determined.

The exquisite sequence specificity of the IIb region forcefully constrains mechanistic models of Rev action. The IIb region is centrally located with respect to the stem regions protected by Rev protein bound to the CAR/RRE in vitro (37). IIb sequences close to stems IIa, IIc, and III are protected by Rev binding, but the functionally critical bases at positions 7418 and 7419 actually become hypersensitive to RNase T_2 upon Rev binding (37). This hypersensitivity has an unknown physical basis, but it should preclude the possibility that Rev remains bound to these bases directly. This suggests two testable alternative possibilities for the role of the central IIb region. One possibility is that it is a site for sequence-specific recognition by critical cellular factors either before or after Rev binding. In the latter case, Rev may bind to regions IIa, IIc, and III, melt out the IIb region, and present it for cellular factor binding. In support of this possibility, we note the following two observations: (i) the IIb region is contained within a fragment of the CAR/RRE reported to specifically recognize a nuclear cellular factor (67) and (ii) upon binding Rev, CAR/RRE RNA undergoes a conformational change consistent with a partial unfolding (11). Another possibility is that the sequence-specific information within region IIb may serve a critical role in Rev binding. For 7418 and 7419 to be critical Rev contact points, however, *rev* would have to bind and then migrate away from positions 7418 and 7419. The requirement that Rev multimerize on the CAR/RRE in order to function properly (43, 48) suggests the attractive possibility that IIb might nucleate Rev binding. A small number of sequence-specific-





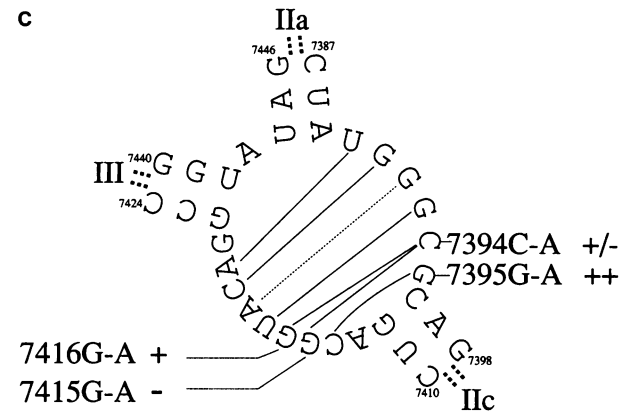


FIG. 3. Analysis of point mutations in and around the IIb region. Dotted lines indicate G-A base pairings. All other proposed base pairings are indicated by solid lines. (A) Mutations that conserve the proposed base pairing in region IIb; (B) mutations that cause major disruptions in the pairing scheme; (C) mutations that cause only minor disruptions without grossly altering the proposed base-pairing scheme. Stability calculations were done as for the base-switched mutants (see Materials and Methods). Quantitation is the same as for Fig. 1.

binding events in the contiguous regions could extend into a nonspecific polymerization of Rev along an entire RNA molecule, protecting it from the nuclear splicing machinery (cf. references 34 and 70).

Conservation of sequence in region IIb. Because the CAR/RRE is in a conserved and functionally important region of gp41, it is often difficult to attribute sequence conservation in this area to either protein-coding constraints or CAR/RRE requirements. Conservation of base 7418A, however, is unambiguously due to CAR/RRE requirements. This position is the third base in a valine codon (Fig. 5). Regardless of the significance of the proposed base pairing, valine could readily be encoded by G (structure disruptive) or U (structure conserving) at this position. These mutations would still use codons frequently used by HIV-1.

Conservation of secondary structure in region IIb. The strongest evidence supporting the base-pairing scheme proposed here for region IIb is that the base pairing is highly conserved among all primate immunodeficiency viruses, even in the presence of envelope coding sequence heterogeneity. The data presented support the proposed base pairing. All mutations which preserve wild-type activity preserve the region IIb base-pairing scheme. All mutations which disrupt this base pairing impair function. The exquisite sequence specificity in the IIb region has unfortunately prevented us from obtaining functional double reciprocal mutants which could prove the proposed structure (13). Although the conservation is proposed to include G-A base pairs, they are not unusual. After the Watson-Crick and G-U base pairings occurring in folded RNA structures, the G-A is considered to be the most regular among the non-Watson-Crick base pairs. G-A pairs have strong stacking properties and can show hydrogen bonding (57, 64). They are frequently found both flanking helical ends and closing internal loops (22, 66, 71). The proposed helicity of this region is consistent with the protection data reported by Kjems et al. (37), which show a lack of reactivity of these nucleotides to single-strand-specific reagents. The seemingly inconsistent poor reactivity of this region with double-strand-specific reagents, such as RNase CV, may derive from tertiary shielding by the surrounding stems or by the bulges and imperfect base pairings present in region IIb. RNase CV is known to fail to recognize documented RNA helices (50).

Sequence specificity outside of the *rev*-binding domain. At present, the presence of sequence specificity outside of the *rev*-binding domain is indicated primarily by the low level of function remaining after base switching together all of domains I/I′, IV, V, and VI. Admittedly, however, computationally analyzing such a large and complex region is more difficult than analyzing the effects of base switching a single stem. Nevertheless, we note that base switching this large region reduces function by about the same amount as deleting all of domain IV-VI (14). None of the information uniquely present in the I/I′, IV, V, or VI domain seems sufficient to account for all of the information present in the region taken as a whole. Base switching domain I/I′ may actually boost Rev-responsiveness. However, the small boost observed in this construct may derive from the lack of sequences 3′ to the minimal CAR/RRE, which sequences have been deleted during the construction of this truncated CAR/RRE analog (14) (see Materials and Methods). Vectors similar to ours have been shown elsewhere to tolerate deletion of domain IV with no loss of Rev responsiveness (28). Base switching domain V has a minimal effect on function. Base switching domain VI moderately impairs function but not enough to account for the impairment in BS(I/I′, IV, V, & VI), $P < 0.01$. Furthermore, deleting all of domain VI has no discernible effect on function (14). Although more extensive mutagenic analysis of this region needs to be done and is under way, the sequence information

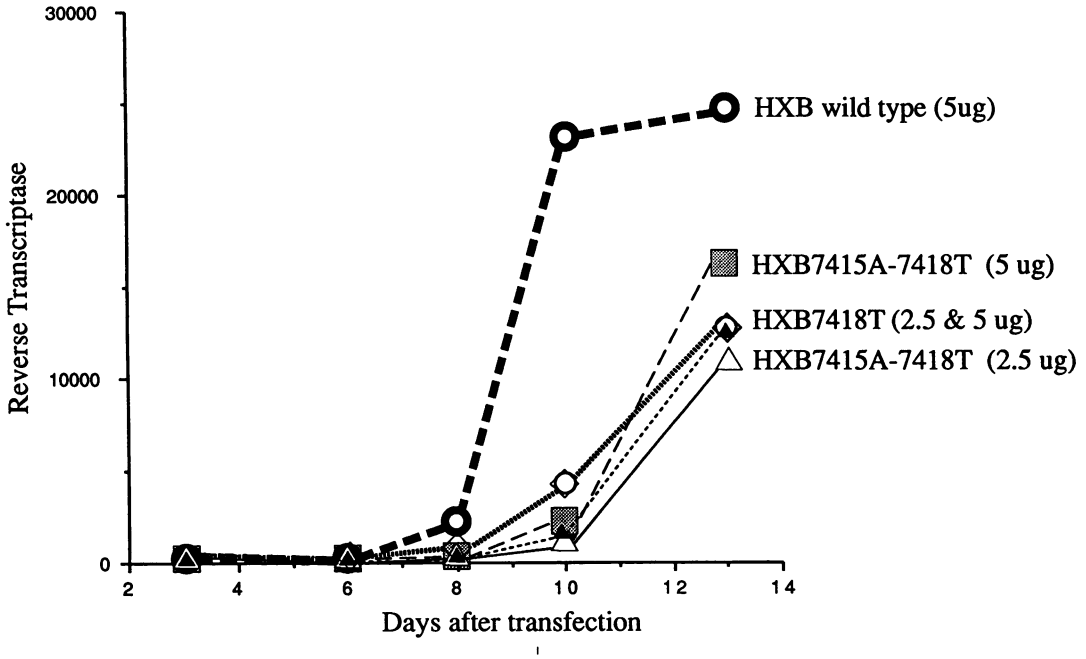


FIG. 4. Replication in T4$^+$ cells of HXB-2 provirus clones with point mutations in the CAR/RRE. Two different concentrations of each mutant (2.5 or 5 μg of DNA per 10 × 10^6 cells) were transfected into A3.01 cells at day 0. Aliquots were removed for reverse transcriptase at the indicated times.

outside of the Rev-binding domain seems to be quantitatively redundant. This information could determine minor binding sites for Rev as suggested by the protection assays of Kjems et al. (37) or could determine binding by critical cellular factors.

The interaction of the Rex protein of human T-cell leukemia virus type I with the CAR/RRE supports the notion that critical cellular factors recognize regions outside the Rev-binding domain. Rev and Rex share virtually no amino acid homologies yet they are teleologically and mechanistically almost identical (4, 15, 16, 21, 25, 30–32, 36, 59, 61). Although both functionally interact with the CAR/RRE (21, 41, 54, 56, 63), they interact with different regions: Rev interacts primarily with regions II and III and Rex interacts primarily with regions IV(?), V, and VI, both seeming to require various amounts of region I/I′ for full activity (1, 3, 29, 63). It is likely that HIV has pirated via the CAR/RRE a complex cellular system involved in RNA splicing and/or transport. In this cellular system, the RNA target would interact with multiple cellular factors, some of which are presumably regulated to control the fate of the RNA to which they are attached. We speculate (Fig. 6) that Rev and Rex have each mimicked one set of cellular factors, supplanting their regulation but operating in conjunction with at least some of the other factors. This suggests the existence of at least two sets of cellular factors, of which one set is Rev-like and binds in conjunction with Rex while the other is Rex-like and binds in conjunction with Rev. According to the model suggested in Fig. 6, the lack of synergy between Rev and Rex (1) could be explained by an intermediate set of proteins which bind to regions of the CAR/RRE falling within both the Rev domain and the Rex domain if Rev and Rex each mimic the intermediate proteins as well.

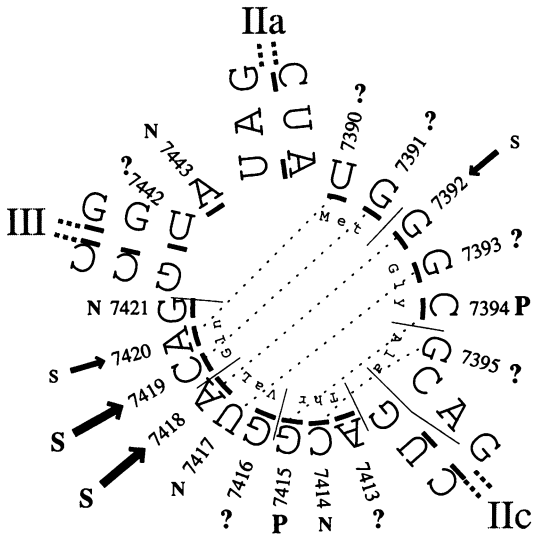


FIG. 5. Sequence specificity and conservation in and around the IIb region of the CAR/RRE. Underlined bases are conserved among all the HIV-1 strains analyzed. Selected envelope amino acids are listed near their codons. S, strongly sequence specific; s, mildly sequence specific; P, probably sequence specific; N, not sequence specific; ?, presently undetermined.

The impaired growth of proviral clones with point mutations in the CAR/RRE validates the Rev-dependent *gag* expression assay. Discrepancies in the literature concerning the biological effects of mutating various regions of CAR/RRE exist. In general, Rev-dependent expression systems that use heterologous genetic material in the regulated transcriptional unit, such as heterologous splice sites (44) or bacterial indicator genes (55), are more tolerant of mutations than are HIV-1 *gag* expression systems similar to the one

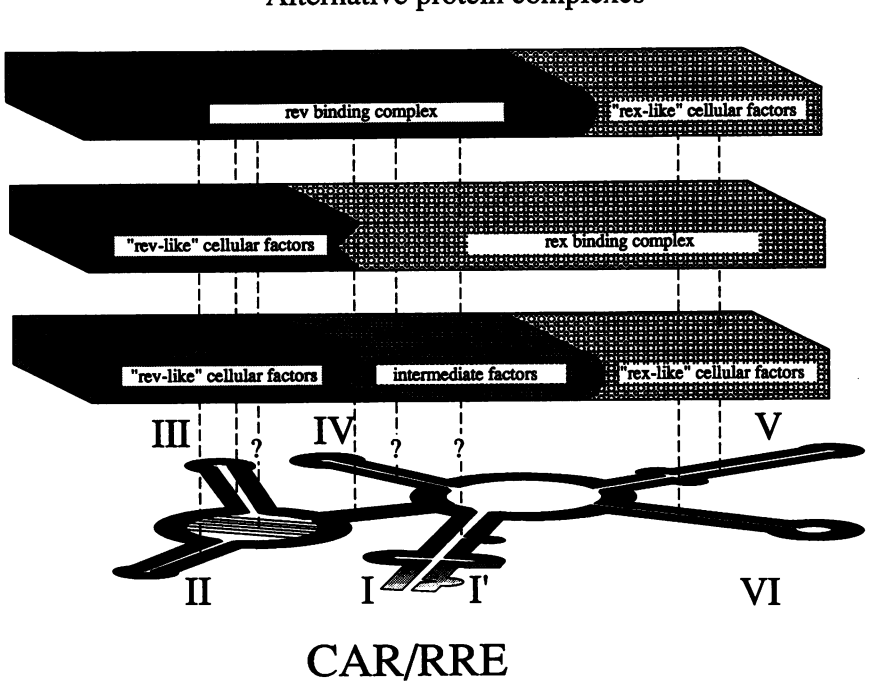


FIG. 6. Protein-RNA interactions proposed for maximal Rev response. Only factors required for maximal response are diagrammed. Presumably other cellular factors bind to undetermined locations, possibly region IIb. Suboptimal responses can presumably occur in the absence of some of the diagrammed factors.

used here (12, 14, 28), which offer the advantage of containing only HIV genetic material in the regulated transcriptional unit. Point mutations which impair the function of the CAR/RRE in our assay system also impair viral growth in tissue culture. This convincingly validates the use of the *gag* expression assay as a measure of CAR/RRE function. The various alternative assays are by no means invalid, however. Rev is called upon to function at multiple points during the processing of RNA in the nucleus. To rescue full-length genomic RNA and *gag-pol* mRNA, Rev must intervene before the first splicing event. To rescue intermediately spliced RNAs, Rev must intervene at various points beyond the first splicing event. The kinetic requirements of these diverse events should be expected to differ from one another. It is likely that the diversity of assay sensitivities mirrors the diversity of roles for the Rev axis of autoregulation.

ACKNOWLEDGMENTS

We thank Alicia Buckler-White for providing oligodeoxynucleotides and help with sequencing, David Alling for statistical analysis, and particularly Anthony Fauci for support, advice, and a reading of the manuscript.

REFERENCES

1. **Ahmed, Y. F., S. M. Hanly, M. H. Malim, B. C. Cullen, and W. C. Green.** 1990. Structure-function analysis of the HTLV-I Rex and Rev RNA response elements; insights into the mechanism of Rex and Rev action. Genes Dev. **4**:1014–1022.
2. **Arrigo, S. J., S. Weitsman, J. D. Rosenblatt, and I. S. Y. 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. **Benko, D. M., R. Robinson, L. Solomin, M. Mellini, B. K. Felber, and G. N. Pavlakis.** 1990. Binding of trans-dominant rev protein of human immunodeficiency virus type 1 to cis-acting rev-responsive element does not affect the fate of viral mRNA. New Biol. **2**:1111–1122.
4. **Black, A. C., I. S. Y. Chen, S. Arrigo, C. T. Ruland, T. Allogiamento, E. Chin, and J. D. Rosenblatt.** 1991. Regulation of HTLV-II gene expression by rex involves positive and negative cis-acting elements in the 5′ long terminal repeat. Virology **181**:433–444.
5. **Chang, D. D., and P. A. Sharp.** 1989. Regulation by HIV rev depends upon recognition of splice sites. Cell **59**:789–795.
6. **Cochrane, A. W., C. H. Chen, and C. A. 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.
7. **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.
8. **Cullen, B. R., and W. C. Green.** 1989. Regulatory pathways governing HIV-1 replication. Cell **58**:423–426.
9. **Cullen, B. R., J. Hauber, K. Campbell, J. G. Sodroski, W. A. Haseltine, and C. A. Rosen.** 1988. Subcellular localization of the human immunodeficiency virus *trans*-acting *art* gene product. J. Virol. **62**:2498–2501.
10. **Daly, T., K. Cook, G. Gray, T. Maione, and J. Rusche.** 1989. Specific binding of HIV-1 recombinant rev protein to the rev-responsive element in vitro. Nature (London) **342**:816–819.
11. **Daly, T. J., J. R. Rusche, T. E. Maione, and A. D. Frankel.** 1990. Circular dichroism studies of the HIV-1 rev protein and its specific RNA binding site. Biochemistry **29**:9791–9795.
12. **Dayton, A. I., E. F. Terwilliger, J. Potz, M. Kowalski, J. G. Sodroski, and W. A. Haseltine.** 1988. Cis-acting sequences responsive to the rev gene product of the human immunodeficiency virus. J. Acquired Immune Defic. Syndr. **1**:441–452.
13. **Dayton, E. T., and A. I. Dayton.** Unpublished observations.
14. **Dayton, E. T., D. M. Powell, and A. I. Dayton.** 1989. Functional analysis of CAR, the target sequence for the rev protein of HIV-1. Science **246**:1625–1629.

15. **Derse, D.** 1988. *Trans*-acting regulation of bovine leukemia virus mRNA processing. J. Virol. **62**:1115–1119.
16. **Dokhelar, M. C., H. Pickford, J. Sodroski, and W. A. Haseltine.** 1989. HTLV-1 p27rex regulates gag and env protein expression. J. Acquired Immune Defic. Syndr. **2**:431–440.
17. **Emerman, M., R. Vazeux, and K. Peden.** 1989. The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. Cell **57**:1155–1165.
18. **Fauci, A. S.** 1988. The human immunodeficiency virus; infectivity and mechanisms of pathogenesis. Science **239**:617–622.
19. **Feinberg, M. B., R. F. Jarrett, A. Aldovini, R. C. Gallo, and F. Wong-Staal.** 1986. HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. Cell **46**:807–817.
20. **Felber, B. K., D. Derse, A. Athanassopoulos, M. Cambell, and G. N. Pavlakis.** 1989. Cross-activation of the Rex proteins of HTLV-1 and BLV and of the Rev protein of HIV-1 and non-reciprocal interactions with their RNA responsive elements. New Biol. **1**:318–328.
21. **Felber, B. K., M. Hadzopoulou-Cladaras, C. Cladaras, T. Copeland, and G. N. Pavlakis.** 1989. The 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.
22. **Gutell, R. R., and G. E. Fox.** 1988. Compilation of large subunit RNA sequences presented in a structural format. Nucleic Acids Res. **16**(Suppl.):r175–r269.
23. **Hadzopoulou-Cladaras, M., B. K. Felber, C. Cladaras, A. Athanassopoulos, A. Tse, and G. N. Pavlakis.** 1989. The *rev* (*trs/art*) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a *cis*-acting sequence in the *env* region. J. Virol. **63**:1265–1274.
24. **Hammarskjold, M. L., J. Heimer, B. Hammarskjold, I. Sangwan, L. Albert, and D. Rekosh.** 1989. Regulation of human immunodeficiency virus *env* expression by the *rev* gene product. J. Virol. **63**:1959–1966.
25. **Hanly, S. M., L. T. Rimsky, M. H. Malim, J. H. Kim, J. Hauber, M. D. Dodon, S.-Y. Le, J. V. Maizel, B. R. Cullen, and W. C. Greene.** 1989. Comparative analysis of the HTLV-1 Rex and HIV-1 Rev trans-regulatory proteins and their RNA response elements. Genes Dev. **3**:1534–1544.
26. **Heaphy, S., C. Dingwall, I. Ernberg, M. J. Gait, S. M. Green, J. Karn, A. D. Lowe, M. Singh, and M. A. Skinner.** 1990. HIV-1 regulator of virion expression (Rev) protein binds to an RNA stem-loop structure located within the Rev response element region. Cell **60**:685–693.
27. **Hogeweg, P., and B. Hesper.** 1984. The alignment of sets of sequences and construction of phyletic trees: an integrated method. J. Mol. Evol. **21**:323–333.
28. **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.
29. **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.
30. **Inoue, J., M. Seiki, and M. Yoshida.** 1986. The second pX product p27x-III of HTLV-I is required for gag gene expression. FEBS Lett. **209**:187–190.
31. **Inoue, J., M. Yoshida, and M. Seiki.** 1987. Transcriptional (p40x) and post-transcriptional (p27x-III) regulators are required for the expression and replication of human T-cell leukemia virus type I genes. Proc. Natl. Acad. Sci. USA **84**:3653–3657.
32. **Itoh, M., J. Inoue, H. Toyoshima, T. Akizawa, M. Higashi, and M. Yoshida.** 1989. HTLV-1 rex and HIV-1 rev act through similar mechanisms to relieve suppression of unspliced RNA expression. Oncogene **4**:1275–1279.
33. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1989. Improved predictions of secondary structures for RNA. Proc. Natl. Acad. Sci. USA **86**:7706–7710.
34. **Karn, J., C. Dingwall, J. T. Finch, S. Heaphy, and M. J. Gait.** 1991. RNA binding by the tat and rev proteins of HIV-1. Biochimie **73**:9–16.
35. **Kim, S. Y., R. Byrn, J. Groopman, and D. Baltimore.** 1989. Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. J. Virol. **63**:3708–3713.
36. **Kiyokawa, T., M. Seiki, S. Iwashita, K. Imagawa, F. Shimizu, and M. Yoshida.** 1985. p27XIII and p21XIII proteins encoded by the pX sequence of human T-cell leukemia virus type I. Proc. Natl. Acad. Sci. USA **82**:8359–8363.
37. **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.
38. **Knight, D. M., F. A. Flomerfelt, and J. Ghrayeb.** 1987. Expression of the art/trs protein of HIV and study of its role in viral envelope synthesis. Science **236**:837–840.
39. **Konings, D. A. M., and A. I. Dayton.** Unpublished data.
40. **Le, S.-Y., M. H. Malim, B. R. Cullen, and J. V. Maizel.** 1990. A highly conserved RNA folding region coincident with the Rev response element of primate immunodeficiency viruses. Nucleic Acids Res. **18**:1613–1623.
41. **Lewis, N., J. Williams, D. Rekosh, and M.-L. Hammarskjöld.** 1990. Identification of a *cis*-acting element in human immunodeficiency virus type 2 (HIV-2) that is responsive to the HIV-1 *rev* and human T-cell leukemia virus types I and II *rex* proteins. J. Virol. **64**:1690–1697.
42. **Malim, M. H., S. Bohnlein, J. Hauber, and B. R. Cullen.** 1989. Functional dissection of the HIV-1 rev trans-activator-derivation of a trans-dominant repressor of rev function. Cell **58**:205–214.
43. **Malim, M. H., and B. R. Cullen.** 1991. HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: implications for HIV-1 latency. Cell **65**:241–248.
44. **Malim, M. H., J. Hauber, S.-Y. Le, J. V. Maizel, and B. R. 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.
45. **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.
46. **Muesing, M. A., D. H. Smith, C. D. Cabradillo, D. V. Benton, L. A. Laskey, and D. J. Capon.** 1985. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. Nature (London) **313**:450–458.
47. **Nalin, C. M., R. D. Purcell, D. Antelman, D. Mueller, L. Tomchak, B. Wegrzynbski, E. McCareny, V. Toome, R. Kramer, and M.-C. Hsu.** 1990. Purification and characterization of recombinant Rev protein of human immunodeficiency virus type 1. Proc. Natl. Acad. Sci. USA **87**:7593–7597.
48. **Olsen, H. S., A. W. Cochrane, P. J. Dillon, C. M. Nalin, and C. A. Rosen.** 1990. Interaction of the human immunodeficiency virus type 1 rev protein with a structured region in env mRNA is dependent on multimer formation mediated through a basic stretch of amino acids. Genes Dev. **4**:1357–1364.
49. **Olsen, H. S., P. Nelbock, A. W. Cochrane, and C. A. Rosen.** 1990. Secondary structure is the major determinant for interaction of HIV rev protein with RNA. Science **247**:845–848.
50. **Parker, K. A., and J. A. Steitz.** 1987. Structural analyses of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base pairing with pre-rRNA. Mol. Cell. Biol. **7**:2899–2913.
51. **Perkins, A., A. W. Cochrane, S. M. Ruben, and C. A. Rosen.** 1989. Structural and functional characterization of the human immunodeficiency virus rev protein. J. Acquired Immune Defic. Syndr. **2**:256–263.
52. **Pomerantz, R. J., D. Trono, M. B. Feinnberg, and D. Baltimore.** 1990. Cells nonproductively infected with HIV-1 exhibit an aberrant pattern of viral RNA expression; a molecular model for latency. Cell **61**:1271–1276.
53. **Ratner, L., W. Haseltine, R. Patarca, K. L. Livak, B. Starcich, S. F. Josephs, E. R. Doran, J. A. Rafalski, E. A. Whitehorn, K. Baumeister, L. Ivanoff, S. R. Petteway, M. L. Pearson, J. A.**

Lautenberger, T. S. Papas, J. Ghrayeb, N. T. Chang, R. C. Gallo, and F. Wong-Staal. 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature (London) **313**:277–283.

54. **Rimsky, L., J. Hauber, M. Dukovich, M. H. Malim, A. Langlois, B. R. Cullen, and W. C. Greene.** 1988. Functional replacement of the HIV-1 rev protein by the HTLV-1 rex protein. Nature (London) **335**:738–740.
55. **Rosen, C. A., E. Terwilliger, A. I. Dayton, J. G. Sodroski, and W. A. Haseltine.** 1988. Intragenic cis-acting art gene-responsive sequence of the human immunodeficiency virus. Proc. Natl. Acad. Sci. USA **85**:2071–2075.
56. **Sakai, H., R. Shibata, T. Miura, M. Hayami, K. Ogawa, T. Kiyomasu, A. Ishimoto, and A. Adachi.** 1990. Complementation of the *rev* gene mutation among human and simian lentiviruses. J. Virol. **64**:2202–2207.
57. **SantaLucia, J., R. Kierzek, and D. H. Turner.** 1990. Effects of GA mismatches on the structure and thermodynamics of RNA internal loops. Biochemistry **29**:8813–8819.
58. **Schwartz, S., B. K. Felber, D. M. Benko, E.-M. Fenyö, and G. N. Pavlakis.** 1990. Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. J. Virol. **64**:2519–2529.
59. **Seiki, M., J. Inoue, M. Hidaka, and M. Yoshida.** 1988. Two cis-acting elements responsible for posttranscriptional trans-regulation of gene expression of human T-cell leukemia virus type 1. Proc. Natl. Acad. Sci. USA **85**:124–128.
60. **Shapiro, S. S., and M. B. Wilk.** 1965. An analysis of variance test for normality (complete samples). Biometrika **52**:591–611.
61. **Siomi, H., H. Shida, S. H. Nam, T. Nosaka, M. Maki, and M. Hatanaka.** 1988. Sequence requirements for nucleolar localization of human T cell leukemia virus type I pX protein, which regulates viral RNA processing. Cell **55**:197–209.
62. **Sodroski, J., W. C. Goh, C. A. Rosen, A. I. Dayton, E. Terwilliger, and W. A. Haseltine.** 1986. A second post-transcriptional transactivator gene required for the HTLV-III replication. Nature (London) **321**:412–417.
63. **Solomin, L., B. K. Felber, and G. N. Pavlakis.** 1990. Different sites of interaction for Rev, Tev, and Rex proteins within the Rev-responsive element of human immunodeficiency virus type 1. J. Virol. **64**:6010–6017.
64. **Sugimoto, N., R. Kierzek, and D. H. Turner.** 1987. Sequence dependence for the energetics of terminal mismatches in ribooligonucleotides. Biochemistry **26**:4559–4562.
65. **Terwilliger, E., R. Burghoff, R. Sia, J. Sodroski, W. A. Haseltine, and C. Rosen.** 1988. The *art* gene product of the human immunodeficiency virus is required for replication. J. Virol. **62**:655–658.
66. **Traub, W., and J. L. Sussman.** 1982. Adenine-guanine base pairing in ribosomal RNA. Nucleic Acids Res. **10**:2701–2708.
67. **Vaishnav, Y. N., M. Vaishnav, and F. Wong-Staal.** 1991. Identification and characterization of a nuclear factor that specifically binds to the Rev response element (RRE) of human immunodeficiency virus type 1 (HIV-1). New Biol. **3**:142–150.
68. **Venkatesh, L. K., S. Mohammed, and G. Chinnadurai.** 1990. Functional domains of the HIV-1 rev gene required for trans-regulation and subcellular localization. Virology **176**:39–47.
69. **Walley, A. J., S. Heaphy, and I. M. Jones.** 1989. Direct RNA binding by the rev product of HIV-1. AIDS **3**:859–861.
70. **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.
71. **Woese, C. R., R. Gutell, R. Gupta, and H. F. Noller.** 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol. Rev. **47**:621–669.
72. **Wong-Staal, F., and R. C. Gallo.** 1985. Human T-lymphotropic retrovirus. Nature (London) **317**:395–403.
73. **Zapp, M. L., and M. R. Green.** 1989. Sequence-specific RNA binding by the HIV-1 Rev protein. Nature (London) **342**:714–716.
74. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244**:48–52.