UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Molecular analysis of HIV-1 Rev

Permalink
https://escholarship.org/uc/item/9h09p0xf

Author
McDonald, David,

Publication Date
1993

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

MOLECULAR ANALYSIS OF HIV-1 REV
BY
DAVID MCDONALD

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

IMMUNOLOGY

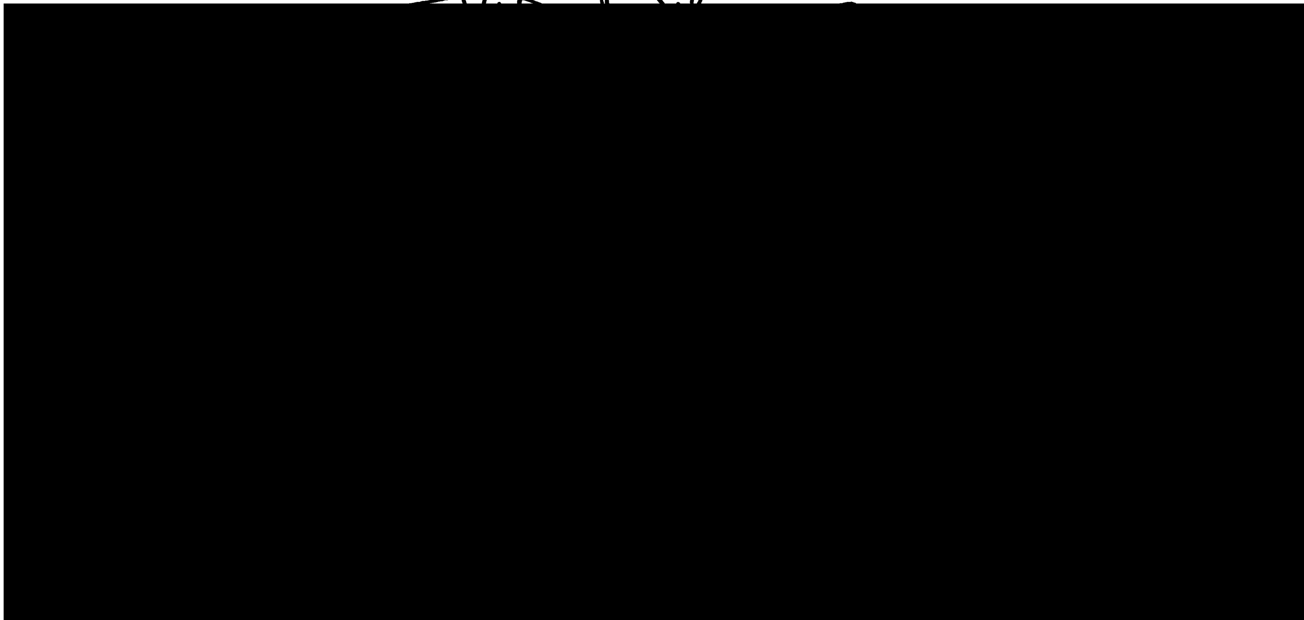
in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco



to Mom and Dad
for their unconditional love and understanding
always

ACKNOWLEDGMENTS

I would like to express my thanks to Tris Parslow, for providing an especially friendly work environment, for lending untold hours of help and encouragement, and for never losing the faith. I also thank Tom Hope, an outstanding colleague and a rare friend.

My collaborators for these studies were T. G. Parslow and T. J. Hope (Chapters II-VI), X. Huang (Chapters II-III), B. L. Bond (Chapters IV-V), J. Low (Chapter III), K. Grahl (Chapter IV) and N. P. Klein (Chapter V). I received financial support from NIH Training grant T32-AI07334 and NIH Research grant R01-AI29313.

UCSF LIBRARY

MOLECULAR ANALYSIS OF HIV-1 REV

by

David McDonald

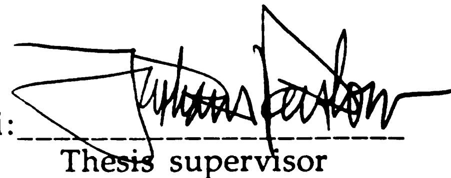
The *rev* gene of the Human Immunodeficiency Virus Type-1 (HIV-1) encodes an essential regulatory molecule belonging to a class of nuclear RNA-binding proteins found in complex retroviruses, including Human T-cell Leukemia Virus (HTLV) Types-I and -II, Simian immunodeficiency viruses (SIV) and Visna viruses. These proteins act post-transcriptionally by binding their target RNAs in the nucleus of the cell and allowing the cytoplasmic expression of the unspliced and incompletely spliced viral mRNAs encoding the structural proteins of the virion.

We have performed mutagenic analyses of HIV-1 Rev and its RNA target sequence, the Rev-response element (RRE). Mutagenesis of Rev suggested that it is organized in a two-domain structure: an N-terminal, arginine-rich RNA-binding domain, and a C-terminal effector domain which has been postulated to interact with an unidentified cellular factor required for transactivation. We identified a functionally homologous effector domain in the HTLV-I Rex protein by fusing Rex to a Rev mutant which lacks an effector domain. Successive truncations of the Rex moiety revealed that a 52-amino acid internal segment from Rex could functionally replace the Rev effector domain, and that both Rev and Rex contain an essential tetrapeptide motif which can also be found in other viral Rev-like molecules.

The target of Rev-binding, the RRE, was originally described as a 240-base RNA sequence within the HIV-1 *env* coding region which has the potential to form substantial secondary structure. We found that much of the RRE

could be deleted without eliminating function, and that a subregion of the RRE containing a single high-affinity Rev-binding site was sufficient to mediate transactivation, suggesting that the sole function of the RRE is to bind Rev. Two tandem copies of this sequence conferred wild-type RRE function, consistent with other reports suggesting that multiple Rev molecules need to be bound to a single RNA for efficient transactivation.

These findings suggested a mechanism of Rev action whereby multiple effector domains are tethered to the regulated RNA, and the remaining portions of the protein, along with the viral RNA response element, serve only to facilitate such linkage and are potentially dispensable. We tested this hypothesis by fusing Rev to the RNA-binding coat protein of bacteriophage MS2. We found that the resulting chimera, called Rev/MS2, could transactivate efficiently in the absence of any RRE sequences when two MS2-binding sites were present on the reporter RNA. Transactivation was dependent on an intact effector domain, but mutations in Rev which abolish RRE-specific binding did not diminish activity through the MS2 operator. We extended this analysis to include a similar Rex/MS2 fusion. This chimera also functioned efficiently through the MS2-operators, and mutation of the Rex effector domain eliminated function. Taken together, these results support the hypothesis that Rev and Rex function by tethering effector-domains to the regulated RNA, and that these proteins act through the same pathway, possibly by interacting with the same hypothetical cellular effector protein.



Abstract approved:____________________
Thesis supervisor

UCSF LIBRARY

TABLE OF CONTENTS

page

LIST OF TABLES..........ix

LIST OF FIGURES..........x

CHAPTER

I: The Molecular Biology of the Human Immunodeficiency Virus Type-1 Rev Protein..........1

II: A Novel Transfection Assay for Rev Function..........12

Materials and Methods..........15

Results..........16

Discussion..........18

III: Mutational Analysis of HIV-1 Rev..........24

Materials and Methods..........24

Results..........26

Discussion..........39

IV: A Minimal Rev-Response Element for HIV-1..........45

Materials and Methods..........46

Results and Discussion..........46

UCSF LIBRARY

CHAPTER page

V: Effector Domains of HIV-1 Rev and HTLV-I Rex are Functionally Interchangeable and Share an Essential Peptide Motif.....58

Materials and Methods.....58

Results.....59

Discussion.....75

VI: Posttranscriptional Regulation by HIV-1 Rev and HTLV-I Rex Through a Heterologous RNA-Binding Site.....76

Materials and Methods.....77

Results.....77

Discussion.....100

VII: Conclusions and Discussion.....98

REFERENCES.....107

APPENDIX A: Nucleotide Sequence of pDM128.....119

UCSF LIBRARY

LIST OF TABLES

Table page

1. Functional characterization of RRE mutants in the cotransfection assay.....49
2. Transactivation by RevΔ/Rex chimeras through the RRE or the XRE.....68
3. Effects of mutations in the conserved tetrapeptide motif on the ability of Rex and of Rev/Rex chimeras to transactivate through the XRE.......73
4. Function of Rex/MS2 through the bul-MS2A response element.............86

UCSF LIBRARY

LIST OF FIGURES

Figure page

1. Genomic organization and mRNA production by HIV-1......................2
2. Sequence and predicted secondary structure of the HIV-1 RRE............6
3. A transient transfection assay for Rev activity......................................13
4. Quantitative analysis of Rev transactivation in CV1 cells...................19
5. Functional analysis of mutant HIV-1 Rev proteins...............................27
6. Mutations near the N-terminus can functionally inactivate Rev.......29
7. Defective subcellular localization of Rev proteins that contain N-terminal mutations...32
8. Function and subcellular localization of mutant Rev/GR fusion proteins in the presence of steroid hormone.......................................35
9. Lack of *trans*-dominant inhibitory activity in N-terminal mutants of the Rev protein...37
10. Conserved structure of an essential N-terminal motif in Rev proteins of immunodeficiency viruses HIV-1, HIV-2, and SIV........42
11. Rev reporter constructs and the RRE...47
12. Alternative models for the secondary structure of RRE SLIIA.............52
13. An 88-base truncated RRE is sufficient for Rev transactivation...........55
14. A Rev/Rex fusion protein transactivates through either the RRE or the XRE..61
15. Fusion with N-terminal Rev sequences confers RRE specificity onto Rex..64

UCSF LIBRARY

Figure page

16. Mutational analysis of the Rex moiety in RevΔ/Rex............................66

17. A conserved and essential effector motif in posttranscriptional transactivator proteins...71

18. Reporters and RRE derivatives for analysis of the Rev response......80

19. Function of the Rev/MS2 fusion protein through a modified response element...82

20. A functional Rev derivative with defective nucleolar localization..85

21. Function of Rev/MS2-MB3 and Rex/MS2 in the absence of the viral response elements...90

22. Northern blot analysis of reporter transcripts in cytoplasmic RNA..92

23. Mutagenesis of the arginine-rich domain in Rev/MS2........................94

UCSF LIBRARY

CHAPTER I:
THE MOLECULAR BIOLOGY OF THE HUMAN IMMUNODEFICIENCY VIRUS TYPE-1 REV PROTEIN

Human Immunodeficiency Virus type-1 (HIV-1), the causative agent of acquired immune deficiency syndrome (AIDS) (1,2), is a retrovirus with an 8.5-kilobase genome that comprises a complex array of genetic elements. In addition to the structural proteins Gag/Pol and Env, HIV-1 encodes at least six other known protein products (3-5), including two important *trans*-acting regulators of gene expression called Tat and Rev (Fig. 1). The mRNAs corresponding to these various gene products are all derived from a single primary transcript which spans the entire HIV-1 genome, and which can be differentially spliced to produce three classes of mRNAs. Unspliced message encodes the Gag and Gag/Pol polypeptides (3). A single splicing event, initiating 5' to the Gag translational start and removing nearly the entire Gag/Pol coding region, gives rise to Env message, or, alternatively, to mRNAs encoding Vif, Vpu or Vpr, depending on the 3' splice acceptor site utilized (3,4). A second splice, eliminating the *env* coding sequence, creates the multiply spliced messages encoding Tat, Rev or Nef proteins (5).

All of the unspliced and singly spliced mRNAs contain signals for further splicing reactions, and can therefore be considered incompletely spliced messages. This is an unusual circumstance for alternatively spliced mRNAs. Most cellular genes that undergo alternative splicing do so by exon skipping or by alternative splice-site selection, resulting in deletion of one or both

UCSF LIBRARY

Figure 1. Genomic organization and mRNA production by HIV-1. Schematic representation of the 9 kilobase proviral genome. Coding sequences for the viral proteins are depicted by rectangles. Tat and Rev are each specified by two coding exons separated by the same intron. Sequences encoding the four known splice donor (D1-D4) and six known splice acceptor (S1-S6) sites are indicated. Representative unspliced, singly spliced, and multiply spliced mRNAs are depicted below, with the proteins they encode listed at right. Dotted lines indicate the regions excised by splicing to produce a given mRNA. LTR = Long terminal repeat; R = Vpr; U = Vpu.

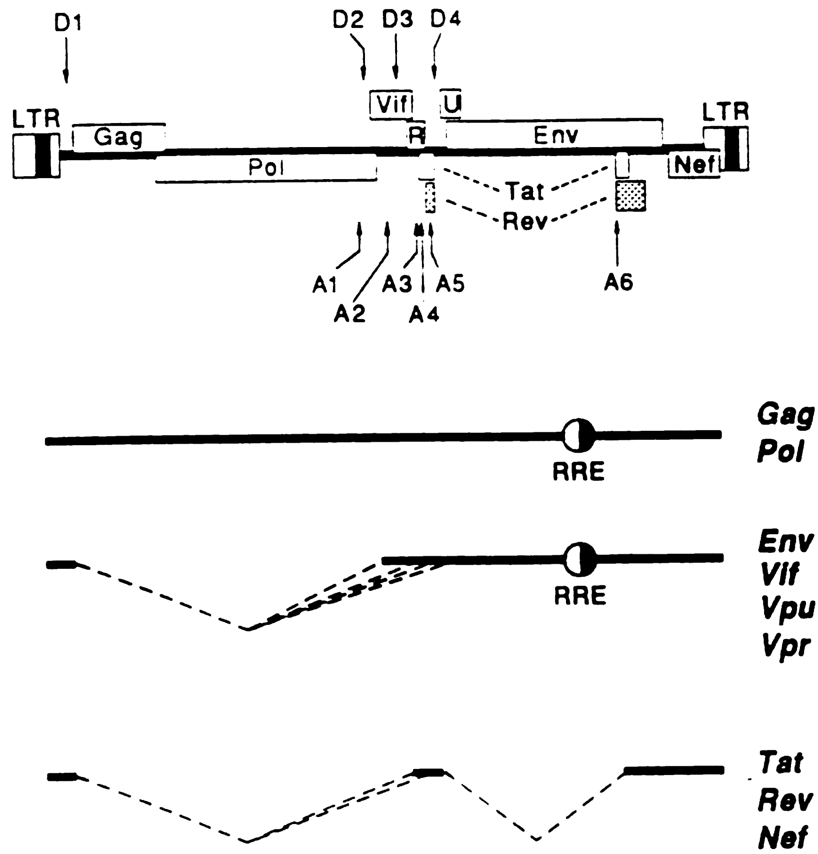


Figure 1

UCSF LIBRARY

splice sites (6,7). Cellular messages that contain intact splicing signals are retained in the nucleus, and so are not allowed access to the cytoplasmic translation system until they are fully spliced. Because of the overlapping organization of genes in HIV-1 and many other retroviral genomes, however, the distinction between introns and exons is blurred: for example, the *env* exon can be treated as an intronic sequence that must be removed from the doubly-spliced messages, and, likewise, the *gag/pol* exon is an intron removed from both singly- and doubly-spliced mRNAs. According to the usual rules of cellular gene expression, then, the incompletely spliced HIV-1 RNAs should normally be retained in the nucleus until splicing is complete. To resolve this problem of cytoplasmic access, HIV-1 encodes the posttranscriptional regulatory protein Rev--a small, nuclear protein encoded by multiply-spliced HIV-1 mRNAs (8).

Proviruses that harbor mutations in the *rev* gene are incapable of producing the structural proteins of the virion, and so are defective for virus replication (8-10). Transcription of *rev*-defective provirus proceeds normally: following activation by cellular transcription factors, Tat protein is produced, and it acts on the LTR to enhance transcription to very high levels. These transcripts are processed in the nucleus, and the three classes of RNAs accumulate there, but only the fully spliced mRNAs appear in the cytoplasm (11-13). The incompletely spliced mRNAs encoding the virion structural proteins are trapped in the nucleus, as would be expected of other incompletely spliced mRNAs. Addition of Rev protein in *trans*, however, relieves this nuclear retention, and results in cytoplasmic expression and translation of the structural messages. Rev, therefore, is an essential, posttranscriptional regulator of viral gene expression that allows the

UCSF LIBRARY

cytoplasmic expression of incompletely spliced viral mRNAs.

Rev acts through a *cis*-acting RNA sequence within the *env* coding region known as the Rev-response element (RRE). The RRE was initially described as a 240-base sequence which has the potential to form a highly complex secondary structure, comprising a major double-stranded stem (Stem 1) crowned by a single-stranded loop that includes four smaller stem-loops designated SL2-SL5 (Fig. 2). Deletion of the entire RRE renders viral messages unresponsive to Rev, and regulation is restored only when the RRE is reinserted in the correct orientation. These observations led Malim *et al.* to postulate that Rev binds specifically to the RRE, and that binding of Rev in the nucleus of the cell allows cytoplasmic expression of the incompletely spliced structural mRNAs (13). This hypothesis was supported by subsequent studies demonstrating RRE-specific RNA binding by Rev in vitro (14-18). Rev was shown to bind with high affinity to one specific stem-loop structure within the RRE, called SL2, as well as to lower-affinity sites located throughout the RRE. Although much of the RRE, including the low affinity sites, can be deleted, SL2 is critically required for RRE function (16, 17).

Interestingly, the RRE is not the sole requirement for Rev regulation of a given mRNA. When the RRE was inserted into a beta-globin intron, the message was processed normally, and no unspliced mRNA was seen in the cytoplasm in response to Rev expression (19). However, replacement of either of the globin splice sites with the corresponding HIV-1 signals conferred Rev regulation. Moreover, mutations of the globin splice sites which were known to decrease the processing efficiency of the message also conferred Rev responsiveness. Globin RNAs which were rendered Rev-responsive were seen to accumulate in the nucleus as unspliced precursor in

UCSF LIBRARY

Figure 2. Sequence and predicted secondary structure of the HIV-1 RRE. The sequence depicted corresponds to residues 7770-8011 of the viral genome, and lies within the Env coding region. The broken circle indicates the primary binding site for Rev. SLII-SLV = Stem loops I to V. Redrawn from ref. 13 with modifications from refs. 91 and 92.

UCSF LIBRARY

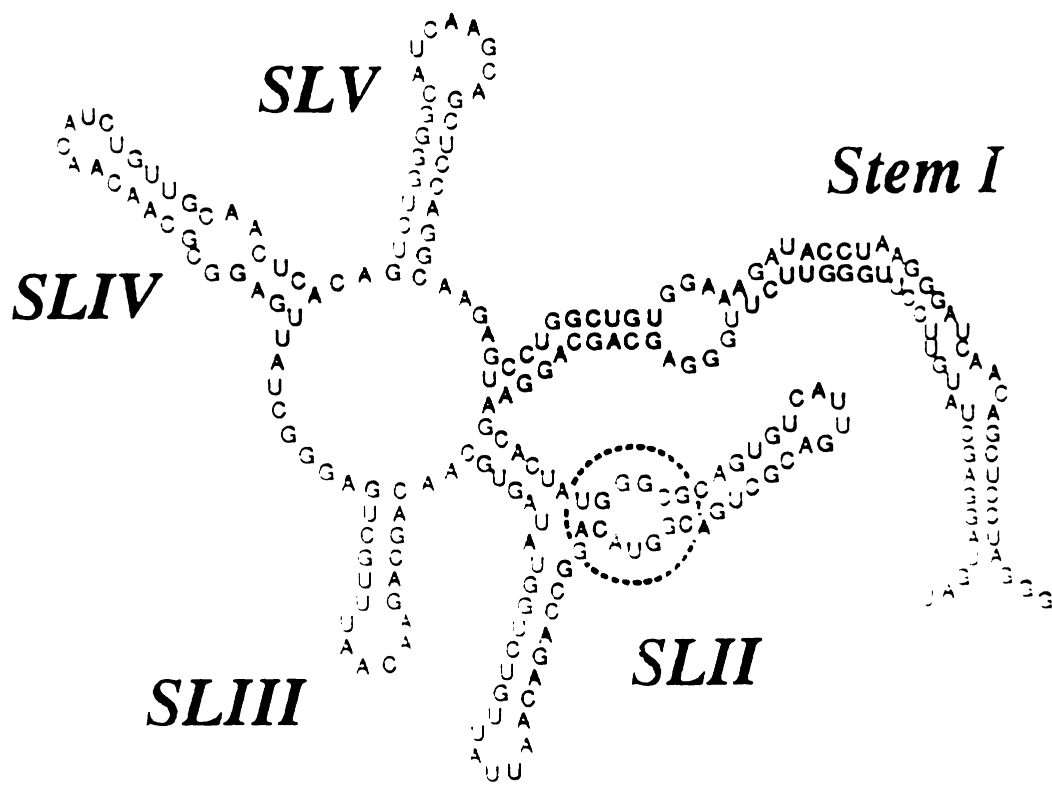


Figure 2

UCSF LIBRARY

the absence of Rev, whereas the RRE-containing message with wild-type globin signals did not accumulate. The authors interpreted these data as evidence that Rev regulation requires engagement by the cellular splicing machinery, and that Rev might function, in part, by inhibition of splicing (19). An alternative interpretation of these data is that there is a kinetic requirement for Rev function: unspliced, RRE-containing precursor must persist within the nucleus in order for Rev to act on it. When, as in the case of wild-type splice sites, the splicing reaction is efficient, RRE-containing substrate is efficiently processed, and there is little unspliced product for Rev to act upon.

Although there are no known cellular counterparts of Rev, proteins with Rev-like function are encoded by a number of evolutionarily distinct retroviruses, including the simian (20) and feline (21) immunodeficiency viruses, sheep Visna virus (22), and the Human T-cell Leukemia Virus (HTLV) types I and II (23, 24). Although these proteins have widely divergent sequences, each appears to function in a manner analogous to that of HIV-1 Rev. For example, the HTLV-I protein, Rex, acts through its operator, the Rex response element (XRE), to induce the cytoplasmic expression of HTLV-I structural mRNAs (23). Indeed, Rimsky *et al.* made the remarkable observation that, although Rev and Rex share virtually no amino acid sequence similarity, Rex can functionally replace Rev to rescue the production of virions from a *rev*-defective HIV-1 provirus (24). It was subsequently discovered that Rex regulates HIV-1 expression by binding to the RRE at a site distinct from the Rev high-affinity site (25).

Further evidence of a common pathway for Rev and Rex transactivation comes from mutational analysis of the two proteins. Rev is divided into two

UCSF LIBRARY

distinct functional domains. The larger, N-terminal domain includes an arginine-rich tract (residues 34-50) that is critical for function. Mutations in this region affect the nuclear transport, nucleolar localization, and RRE-specific binding activity of the protein (26,27). Additionally, residues at positions 14-20 are also required for Rev function, and appear to contribute to RNA binding and multimerization (28, 29). The second, more carboxy-terminal domain of Rev contains essential residues that do not contribute to RNA binding or subcellular localization. Instead, mutations in this region, which is called the effector domain, are thought to interfere with critical functional contacts between Rev and an unidentified (and still hypothetical) cellular cofactor--perhaps a component of the splicing or transport apparatus. Some mutations in the effector domain can give rise to a dominant-negative Rev mutant--a molecule which is not only inactive, but can block the function of wild-type Rev when expressed in excess in the same cell (26).

Although the domain structure of Rex is not as simple as that of Rev, residues important for nuclear and nucleolar localization, and for XRE-specific binding have been identified by mutagenic analysis of the protein (30-33). Dominant-negative Rex mutants were also derived in these studies, and these molecules were found to inhibit Rev transactivation through the RRE as well (31,32). Further analysis revealed that dominant-negative Rex mutants can inhibit Rev activity even when the RRE is modified to abolish Rex-specific binding, and that dominant-negative Rev mutants can also inhibit Rex function without binding the response-element (25). This reciprocal inhibition in the absence of response-element binding by the dominant-negative proteins implies that the inhibition does not occur by competition for the regulated RNAs, and supports the hypothesis that the

UCSF LIBRARY

inhibition can occur by competition for a common cellular cofactor that is involved in transactivation by both Rev and Rex.

In the following pages, I describe a series of studies designed to characterize the functional domains of Rev and Rex, and to investigate the contribution of these domains, and of their RNA response elements, to this novel form of gene regulation. Chapter two describes the development of a simplified assay for Rev and Rex function based on the forced regulation of the bacterial chloramphenicol acetyl-transferase (CAT) gene in cell culture transfection experiments. Each of the following four chapters correspond to individual publications resulting from these studies which can be found in references 28, 34, 35 and 36. In particular, chapter three describes initial mutagenesis of the Rev protein, and the identification of critical amino-acid residues N-terminal to the arginine-rich tract. Chapter four deals with the mutagenesis of the Rev-response element, resulting in the derivation of a minimal-length RRE that is active in vivo. The fifth chapter describes the use of a Rev/Rex chimeric fusion protein to further map functional domains of the two proteins, demonstrating that Rex contains an effector domain that is functionally equivalent to that of Rev, and that they are likely to share structural similarity. In chapter six, fusion of Rev and Rex to the RNA-binding bacteriophage MS2 coat protein is shown to redirect the target specificity of the two proteins. The chimeras transactivate efficiently through the MS2-binding site, and do not require any RRE or XRE sequences to function. Mutation of the Rev/MS2 chimera reveals that residues outside the effector domain are dispensable for function, supporting a model for Rev transactivation in which the essential contribution of the RNA-binding domain and of the RRE is to tether effector domains onto the regulated RNA.

UCSF LIBRARY

The final chapter contains a summary and discussion of the results presented. My collaborators for each of these studies are listed in the acknowledgements on page iv.

UCSF LIBRARY

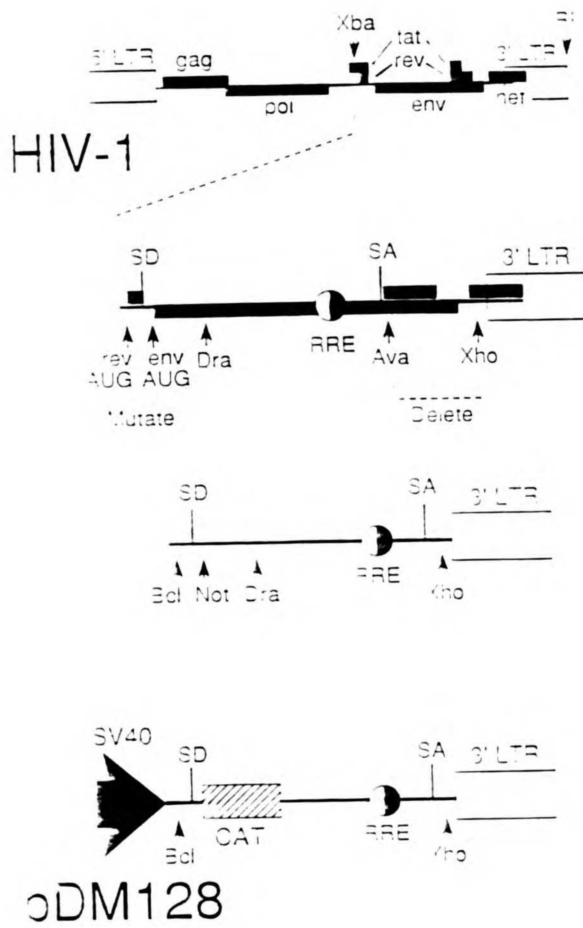
CHAPTER II:
A NOVEL TRANSFECTION ASSAY FOR REV FUNCTION

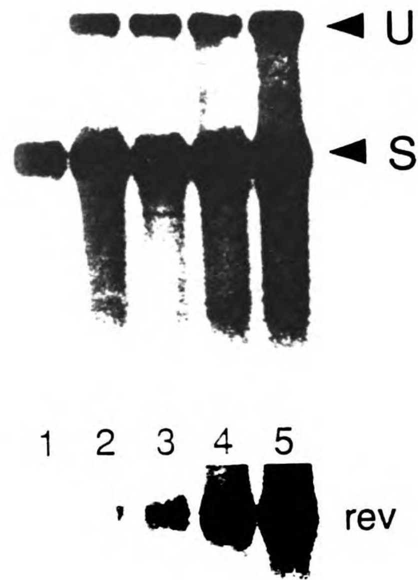
Rev was first identified functionally by its ability to rescue virion production from a *rev*-defective provirus when infected cells were cotransfected with a Rev expression construct (8). While this proviral assay can provide a sensitive measure of Rev activity, it involves the production of a potentially lethal pathogen, and also can be influenced by the expression of other proteins encoded in the HIV-1 genome. We wished to simplify the assay of Rev function, and to eliminate potential complications that might result from the expression of other viral gene products. To this end, we developed a reporter plasmid, called pDM128, derived from the 3′ half of the HIV-1 genome (Fig. 3A). This construct contains unique HIV-1 donor and acceptor splice sites flanking the *env* exon, and is driven by the simian virus 40 (SV40) promoter. In addition, a portion of the *env* exon is replaced by the bacterial chloramphenicol acetyl-transferase (CAT) coding sequence to create a CAT-encoding unspliced message which mimics Rev-regulated Env expression. In the absence of Rev, the unspliced message encoding CAT is retained in the nucleus of the cell, and no CAT protein is produced. Coexpression of Rev allows the cytoplasmic expression of this unspliced message, and so enables CAT to be translated. CAT activity is very low in cells transfected with pDM128 alone, and increases up to 100-fold in response to functional Rev expression.

UCSF LIBRARY

Figure 3. A transient transfection assay for Rev activity. (A) Derivation of the Rev-dependent reporter plasmid pDM128 from the HIV-1 genome. The RRE and the splice donor (SD) and splice acceptor (SA) sites are indicated. (B) Cytoplasmic RNA expression in COS7 cells transfected with 0.5 μg of pDM128 either alone (lane 1) or with 2, 5, 10, or 20 μg of pRSV-Rev (lanes 2-5). Duplicate 5-μg aliquots of RNA were probed for spliced (S) and unspliced (U) pDM128 transcripts by using a 3′ LTR probe (Upper) or for *rev* coding sequences (Lower). The 3′ LTR probe was HIV-1 nt 8915-9575; the Rev probe was a full-length cDNA. (C) CAT enzyme assay of CV1 cells cotransfected with 1 μg of p128 and the indicated amounts of wild-type (lanes 7-11) or mutant forms of pRSV-Rev (lanes 12 and 13). FS = frameshift at Rev codon 58 as described (13); Δ2 = in-frame deletion of codons 38-50.

UCSF LIBRARY





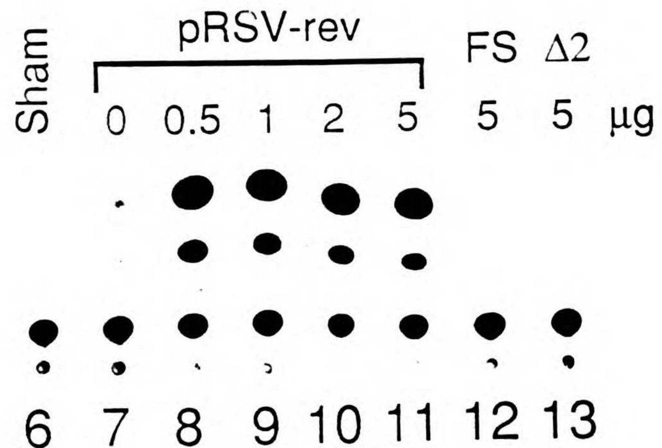


Figure 3

UCSF LIBRARY

MATERIALS AND METHODS

Plasmid constructions. The reporter plasmid pDM128 was derived from pSV9B, which transcribes the genome of HIV-1 strain SF2 under control of the simian virus-40 (SV40) promoter and enhancer (37). The 5′ half of the HIV-1 genome, encoding the *gag* and *pol* genes, was deleted to nucleotide (nt) 5863, and replaced with an *Xba* I linker. Oligonucleotide-directed mutagenesis (38) was used to introduce an *Eco* RI site at the end of the 3′ LTR (nt 9722), and to replace the *rev* and *env* start codons with unique *Bcl* I and *Not* I sites at nt 5979 and 6233, respectively. The second exon of *rev* was deleted by introduction of an *Ava* I site nt 8915 and subsequent deletion to the *Ava* I site at nt 8400. The CAT gene was placed at the *env* translational start site by subcloning the 780-base pair CAT coding sequence into the *Not* I to*Dra* III sites. The entire transcription unit (*Sal* I/*Eco* RI) was then subcloned into pUC9 to produce pDM128 (see Fig. 3). The pDM138 derivative was constructed by replacing the 1.2-kb *Stu* I-*Bsm* I fragment of the *env* intron (nt 6852-8066), which contains the RRE, with a unique *Cla* I site. pRSV-Rev contains the Rev coding sequence from HIV-1 strain BRU (39), the Rous sarcoma virus (RSV) promoter, and HIV-1 polyadenylation signals (nt 9575-9694) in pUC118. The sequences of the pDM128 and pRSV-Rev transcription units can be found in appendix A.

Transfection and Expression Assays. Transfections were performed by the calcium phosphate coprecipitation method using doubly CsCl-purified plasmids. Each 10-cm plate of cells received the indicated amounts of Rev plasmids along with 1 μg of pDM128, and sufficient pUC118 to maintain total

UCSF LIBRARY

DNA input at 10 μg. Cells were washed 16 h later and given fresh medium, and then harvested 30-36 h after transfection. For isolation of cytoplasmic RNA, cells from each transfected plate were washed once in Tris-buffered saline and then resuspended in 200μl of ice-cold 10 mM Tris (pH 8.0)-10 mM NaCl-1.5mM $MgCl_2$-.5% (wt/vol) 3-[(3-cholamidopropyl)-dimethyl-amonio]-1-propanesulfonate (CHAPS)-20% (wt/vol) sucrose. Nuclei were removed by microcentrifugation (1 min. and then 10 min.) at 4^0C. Supernatants were combined with 200 μl of Tris-buffered saline-0.5% CHAPS-0.5% sodium dodecyl sulfate and repeatedly extracted with phenol-chloroform. RNA was then precipitated with ethanol, analyzed by formaldehyde-agarose gel electrophoresis, transferred to nitrocellulose, and probed for pDM128- or Rev-specific sequences.

For CAT assays, 0.25 μg of the β-galactosidase expression vector pCH110 (40) was included in the transfected DNA to provide an internal control for transfection efficiency and for nonspecific effects on the SV40 promoter. After harvesting the transfected cells, lysates were prepared by freeze thawing three times in 0.25 M Tris (pH 7.5) followed by centrifugation. Aliquots of lysate were assayed colorimetrically for β-galactosidase activity, and volumes that contained equal β-galactosidase activity were then assayed for CAT enzyme activity, which was quantitated by thin-layer chromatography and scintillation counting.

RESULTS

We first asked whether the pDM128 reporter accurately mimicked a Rev-regulated transcript at the level of cytoplasmic mRNA expression. COS7 cells

UCSF LIBRARY

were transfected with pDM128 alone or with pRSV-Rev, and cytoplasmic RNA was isolated and analyzed by Northern blot hybridization (Fig 3B). In the absence of cotransfected Rev, only a single band, corresponding to the spliced form of the reporter, was detected in the cytoplasm. Addition of a small amount of pRSV-Rev plasmid resulted in the appearance of the unspliced transcript, containing the CAT coding sequence, in the cytoplasm. Increasing amounts of pRSV-Rev did not induce significantly more unspliced message, suggesting that under these assay conditions, maximal Rev-response is achieved by relatively low levels of protein expression.

The ability of Rev to control CAT protein expression was next tested. In this case, CV1 cells, which lack the SV40 T-antigen, were chosen for transfection. Because COS7 cells express T-antigen, SV40-based vectors, such as pDM128, are expressed at very high levels. While this is convenient for the relatively insensitive Northern blot RNA assay, we have found that transfection of pDM128 into COS7 cells results in significant expression of CAT enzyme activity in the absence of Rev, presumably due to leakage of the unspliced message from the nucleus. Transfection of pDM128 alone into CV1 cells, on the other hand, resulted in a very low level of CAT activity (Fig. 3C). When pRSV-Rev was included in the transfection, CAT activity was dramatically increased by more than 75-fold. Addition of increasing amounts of pRSV-Rev plasmid did not induce more CAT activity, again suggesting that this assay system is sensitive to low levels of Rev expression. That the induction of CAT activity is dependent on expression of a functional Rev protein is demonstrated in lanes 12 and 13, in which addition of two mutant forms of Rev, previously shown to be nonfunctional (26), did not increase CAT activity above background levels. As a control for transfection efficiency

UCSF LIBRARY

and for nonspecific effects on the SV40 promoter, the plasmid pCH110 (40), which directs the synthesis of the bacterial β-galactosidase gene under control of the SV40 promoter, is included in all transfections. After harvesting the transfected cells, volumes of lysate containing equal amounts of β-galactosidase activity are assayed for CAT activity. This internal transfection control allows accurate quantitative comparison of all samples in the transfection.

We next tested the ability of the reporter to detect limiting amounts of Rev protein. The RSV promoter in pRSV-Rev was replaced by the Herpes Simplex Virus type 1 (HSV-1) minimal thymidine kinase promoter to make ptk-Rev. This vector expresses at least 25 fold less Rev protein than pRSV-Rev, as judged by Western immunoblotting (data not shown). When ptk-Rev was tested for its ability to transactivate in this assay, we found that CAT expression was induced in a linear, dose-dependent manner. Under standard conditions, as little as 1μg of pRSV-Rev induced maximal CAT expression. One μg of ptk-Rev, on the other hand, yielded only partial transactivation, and the CAT response increased in direct proportion to the amount of ptk-Rev plasmid used, up to 5 μg (Fig. 4).

DISCUSSION

The pDM128 assay has a number of advantages over other Rev detection methods that have been reported. Because there is no HIV virion production involved, the need for biological containment is minimal. Other assays based on the regulation of *env* (12) or of a truncated form of *tat* (13) by Rev have been described, but these assays require high-level expression of the reporter

UCSF LIBRARY

Figure 4. Quantitative analysis of Rev transactivation in CV1 cells. Induction of CAT enzyme activity from reporter plasmid pDM128 by cotransfection of Rev expression vectors. Each plate of cells received 1 μg of pDM128. Western blot titrations indicate that pRSV-Rev yields at least 25-fold higher Rev protein expression than does an equivalent amount of ptk-Rev (data not shown).

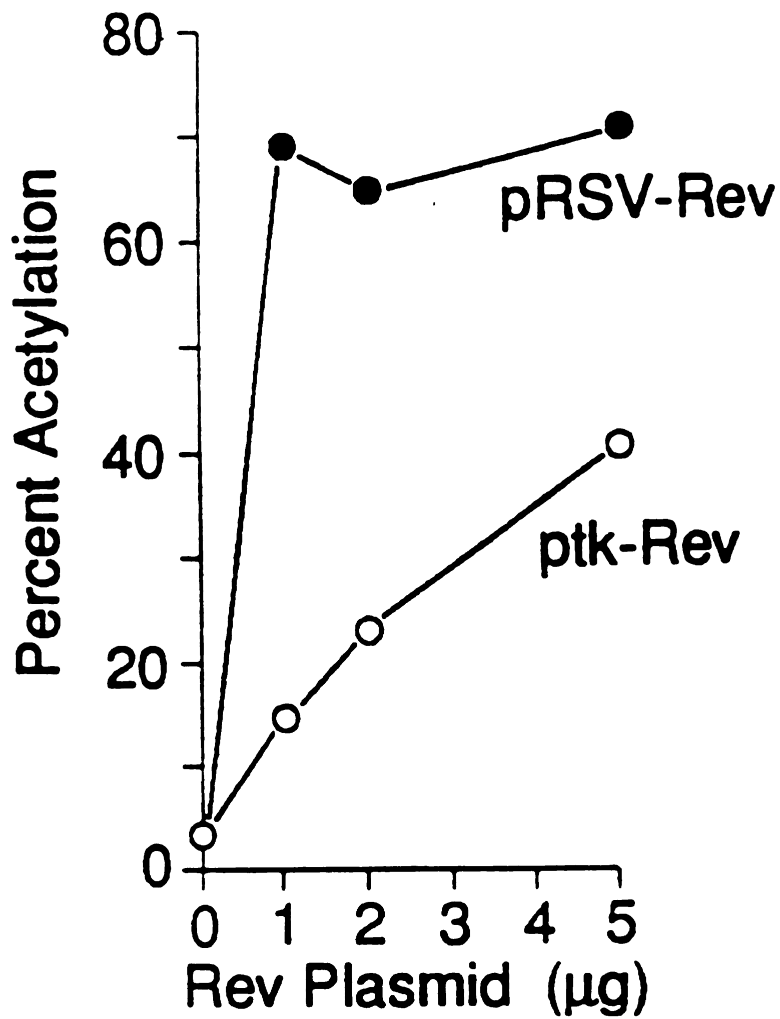


Figure 4

UCSF LIBRARY

using replicating vectors, and depend upon immunoprecipitation of biosynthetically radiolabelled products by specific antisera to detect the regulated polypeptides. Detection of the CAT protein, by comparison, is extremely sensitive, easy to perform and relatively inexpensive. Because pDM128 does not require plasmid replication, the assay can be applied to a wide variety of cell types. We and others have successfully detected pDM128 transactivation by Rev in human fibroblast and T-cell lines, mouse L-cells, and rat, chicken and quail cell lines.

We have also derived a related reporter, called pDM138, in which the Rev-response element of pDM128 is replaced by a unique *Cla* I restriction site. As expected, this reporter does not respond to Rev transactivation, but reintroduction of the RRE in the correct orientation restores the response. This plasmid has been useful for mapping important sequences in the RRE, as will be discussed in Chapter IV, and for testing response elements of other Rev-like proteins (Chapter V), as well as heterologous RNA targets (Chapter VI).

The intermediate responses seen with limiting amounts of Rev in the pDM128 assay contradict a report by Pomerantz *et al.*, who observed a non-linear, threshold effect of Rev expression required for efficient replication of a *rev*-defective provirus. The authors hypothesized that this effect might indicate an important role for Rev in the maintenance of proviral latency (41). Although the reason for these differing observations is not known, it is probably due to differences in the assays. Transactivation of pDM128 is dependent only on Rev expression, whereas the proviral rescue assay relies on the expression of HIV-1 structural proteins and assembly of virion particles. Therefore, the observed threshold effect could be at some other

UCSF LIBRARY

CHAPTER III:

MUTATIONAL ANALYSIS OF HIV-1 REV

Rev is encoded by two viral exons specifying amino acids 1 to 26 and 27 to 116, respectively. Previous studies (26) have distinguished two regions of the protein that are essential for its function *in vivo*. The first of these has been mapped by mutagenesis to amino acids 25 to 50 (26, 27) and so encompasses a prominent arginine-rich tract at residues 35 to 50; this region contains the signals that target Rev to nuclei and nucleoli (26, 27, 42) and may also participate in binding of the RRE (48). A second region near amino acid 80 is also required for transactivation but appears to have no role in localization of the protein (26). Point mutations outside these two regions have little or no effect on Rev activity. Some evidence suggests, however, that additional sequences near the N-terminus might also contribute to function. Malim *et al.* (26) have observed, for example, that deletion of amino acids 5 to 17 abolishes transactivation, and exon-shuffling experiments suggest that residues within the first 26 amino acids of Rev may be a major determinant of its specificity for the RRE (20). We therefore undertook a further mutational analysis of Rev to identify features at the N-terminus that are essential for biological activity.

MATERIALS AND METHODS

Fluorescence immunocytochemistry. COS7 cells growing on glass cover slips were transfected with derivatives of pRSV-Rev or pRSV-Rev/GR (10 μg

UCSF LIBRARY

per 3-cm plate). At 40 h after transfection, the cells were briefly washed twice in phosphate-buffered saline (PBS) and once in ice-cold methanol, fixed in methanol at -20^{0}C for 20 min, and air-dried for 1 h. After rehydration for 15 min, the specimens were incubated for 2 h in the primary antiserum, washed for 15 min, incubated for 1 h in a 1:50 dilution of rhodamine-conjugated goat anti-rabbit immunoglobulin G (Boehringer Mannheim), and then washed again for 15 min, all in PBS at room temperature. Use of the primary GR antiserum has been described (42). The primary Rev antiserum was a mixture of two rabbit antipeptide sera (each at 1:800 dilution) directed against residues 1 to 20 and 27 to 50, respectively (9), along with a third rabbit serum (a gift from M.-L. Hammarskjold and D. Rekosh; used at 1:1500 dilution) raised against an *E. coli* fusion protein (43) that contained the C-terminal half of Rev.

Western immunoblot analysis. COS7 cells transfected 40 h previously with the indicated plasmids (20 μg per 10-cm plate) were washed and then suspended in 75 μl of 100 mM Tris (pH 7.5)-50 mM NaCl-10% glycerol-1 mM dithiothreitol--1 mM phenylmethysulfonyl fluoride-1 mg each of leupeptin, aprotinin, and pepstatin per ml. Cells were lysed by three cycles of freeze-thawing, and samples containing equivalent amounts of total protein were analyzed by sodium dodecyl sulfate-polyacrilamide gel electrophoresis and then transferred to nitrocellulose by electroblotting. The membranes were then preincubated for 15 min, incubated for 1 h with the mixed primary Rev antisera described above, briefly washed three times, incubated for 1 h with 5 μCi of ^{125}I-labeled staphylococcal protein A (Amersham), and again washed three times, all at room temperature in PBS containing 2.5% (wt/vol) nonfat dry milk and 0.1% Tween-20.

UCSF LIBRARY

RESULTS

With pRSV-Rev as a template for oligonucleotide-directed mutagenesis, we prepared a series of plasmids that encode mutant forms of Rev. The mutations included some deletions and frameshifts as well as individual and clustered missense mutations at selected sites throughout the protein. Each of these mutants was then tested for its ability to induce CAT expression from pDM128 in the cotransfection assay. As seen in Fig. 5, we found that several mutations involving amino acids 27 to 50, as well as three different mutations involving the C-terminal half of the protein, completely abolished transactivation, whereas mutations in the intervening region (residues 53 to 66) had relatively little effect. The properties of these mutants was consistent with previous studies (26). We also observed that deletion of N-terminal residues 4 to 26 strongly inhibited Rev function, suggesting that critical residues might lie in this region.

The Rev protein sequence encoded by pRSV-Rev is that of HIV-1 strain BRU (39). In this sequence, positions 7 to 11 are occupied by a cluster of acidic amino acids that includes one of the two phosphoserines in Rev (26), and positively charged residues are found at positions 4, 14, 17 and 20 (Fig. 6A). A computerized secondary-structure prediction, done by the method of Finer-Moore and Stroud (44) indicated that the N-terminal region of Rev (amino acids 8 to 18) has a very strong propensity to form an α-helix.

With these features in mind, we constructed and tested six additional derivatives of pRSV-Rev (designated MA1 through MA6) that contained missense mutations in the N-terminal region. As illustrated in Fig. 6B, most of these mutations had relatively minor effects on Rev function. Mutations

UCSF LIBRARY

Figure 5. Functional analysis of mutant HIV-1 Rev proteins. Mutations were introduced into the coding region of pRSV-Rev by oligonucleotide-directed mutagenesis (38). Hatching denotes the location of the arginine-rich tract in the wild-type protein; vertical bars indicate sites of missense mutations. In-frame deletions of indicated codons are designated with Δ; fs58 is a frameshift mutation at codon 58 created by cleaving pRSV-Rev with *Bam* HI, filling in with the Klenow fragment of DNA polymerase I and religation. A short C-terminal missense coding sequence is added as a consequence of the frame-shift (filled rectangle). Naming of missense mutants indicates the native and mutant residues (in single-letter code) at the affected positions; for instance, Y 23 S refers to a substitution of serine for tyrosine at position 23. LE 78, 79 is identical to the mutant designated M10 in ref. 13 and Fig. 9. The autoradiogram depicts results of CAT enzyme assays of extracts from CV1 cells that had been cotransfected with pDM128 and wild-type or mutant pRSV-Rev; these data are from a single representative experiment. Sham, pUC118 and pCH110 (β-galactosidase control plasmid) only; None, no pRSV-Rev. Summarized at right are the relative activities of mutants compared with wild-type Rev, based on two or more tests of each construct: +++ = 50 to 100%; ++ = 25 to 50%; +, 5 to 25%; -, 0 to 5%.

UCSF LIBRARY

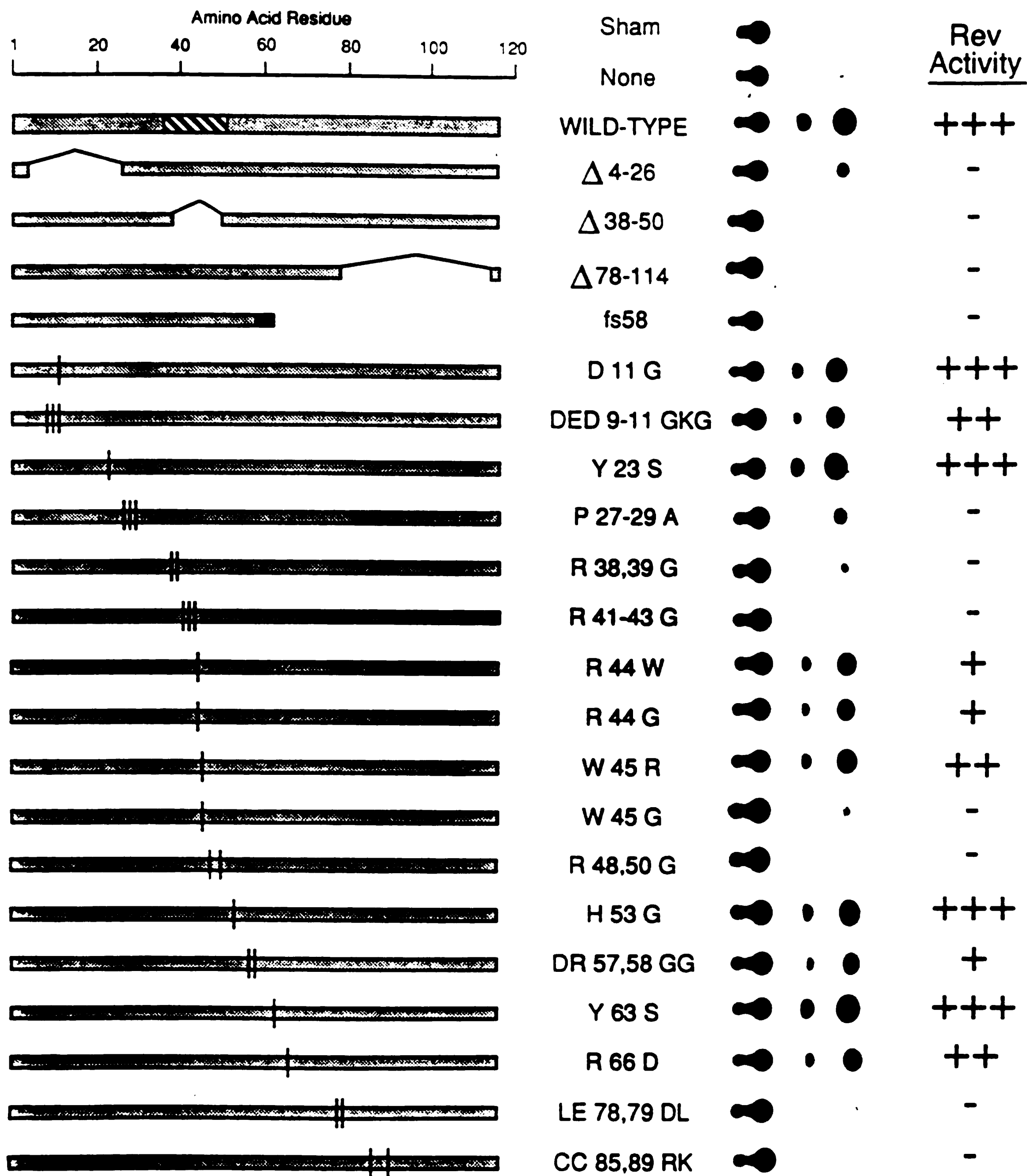
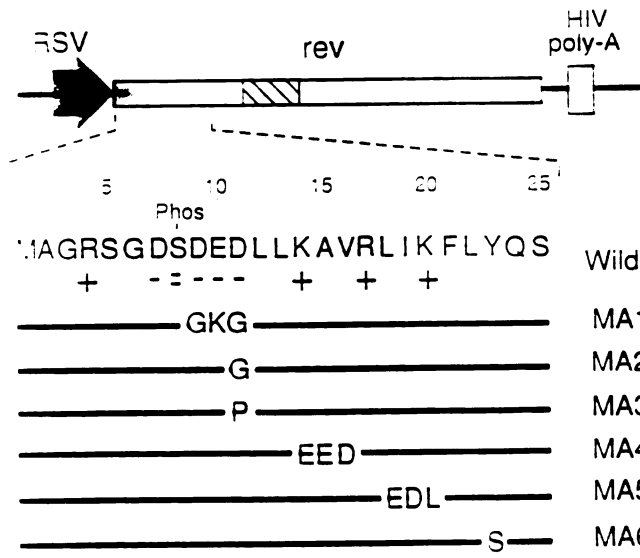


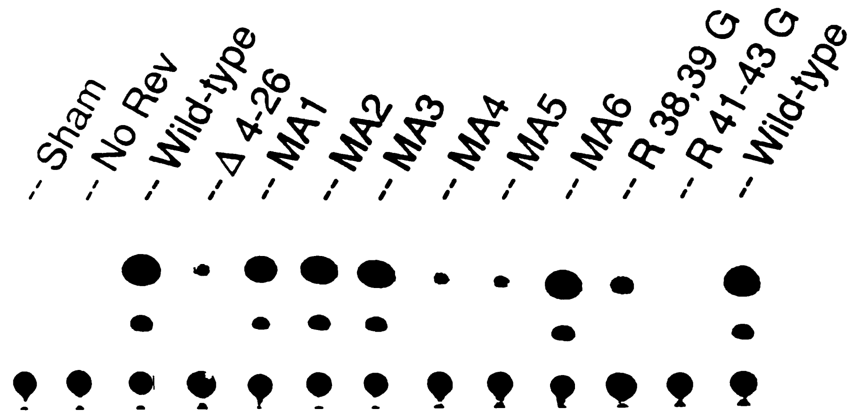
Figure 5

UCSF LIBRARY

Figure 6. Mutations near the N-terminus can functionally inactivate Rev. (A) Structures of wild-type and N-terminal mutant forms of pRSV-Rev. Hatching denotes the arginine-tract. N-terminal amino-acid sequences (in single letter code) are depicted for the wild-type and six mutants. Charged residues in the wild-type sequence are also indicated. HIV-1 poly-A, HIV-1 polyadenylation site. (B) CAT enzyme assay of CV1 cells after cotransfection of pDM128 with wild-type or mutant forms of pRSV-Rev. Sham, pUC118 and pCH110 only; No Rev, pDM128, pCH110 and pUC118. (C) Western blot detection of variant Rev proteins in transfected COS7 cells with Rev-specific antisera and radioiodinated protein A after fractionation in a 14? polyacrylamide gel. Positions of ^{14}C-labeled molecular mass standards (in kilodaltons) are indicated at right.

UCSF LIBRARY





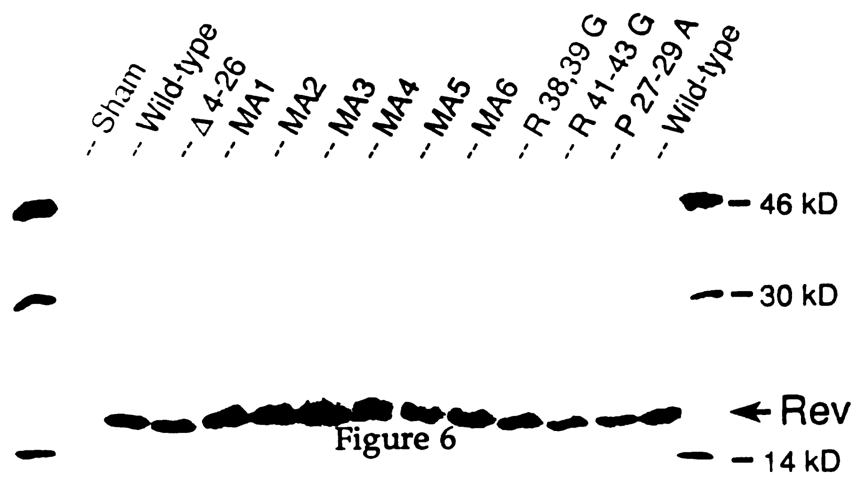


Figure 6

UCSF LIBRARY

that altered the charge of the acidic region (MA1) or introduced helix-breaking amino acids at position 11 (MA2 and MA3) produced only partial loss of activity, and replacement of tyrosine by serine at position 23 (MA6) had little effect. In striking contrast, however, two mutations (MA4 and MA5) that introduced acidic residues at positions 14 to 16 or 18 to 20, respectively, completely abolished function.

Western blot analysis was performed to see if the effect of the N-terminal mutations resulted from decreased protein stability. When lysates of transfected COS7 cells were analyzed with Rev-specific antisera, each plasmid was found to yield a single predominant immunoreactive protein (Fig 6C). The mutant proteins were expressed at levels comparable to that of the wild type and were of the expected molecular weight.

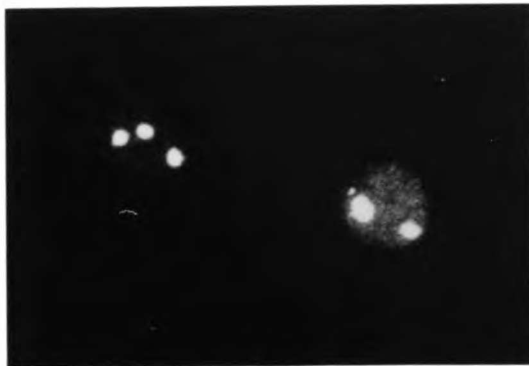
The subcellular distribution of these mutant proteins was analyzed by fluorescence immunocytochemistry in situ. When COS7 cells that had been transfected with wild-type pRSV-Rev were stained with Rev-specific antisera (Fig. 7), Rev was found to be associated primarily with nucleoli, as reported previously (9). This pattern of localization was also observed for the functional mutants MA1, MA2, MA3, and MA6 (data not shown). In contrast, the inactive mutants MA4 and MA5 were each expressed diffusely throughout both nucleus and cytoplasm (Fig. 7). Thus, the inability of these mutant proteins to transactivate is associated with improper subcellular localization in transfected cells.

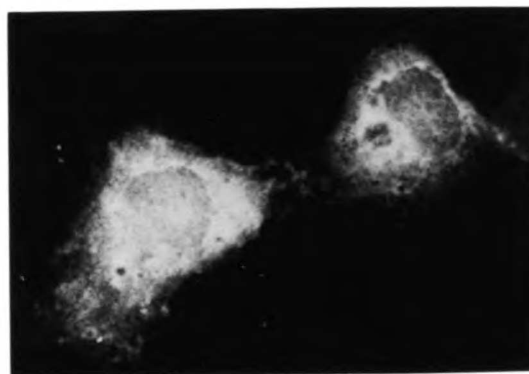
We then asked whether the function of these mutants could be restored by correcting the defect in nuclear translocation. Our lab has reported that translocation defective Rev mutants can be transported efficiently into the nucleus by fusion with a portion of the rat glucocorticoid receptor (GR) (42).

UCSF LIBRARY

Figure 7. Defective subcellular localization of Rev proteins that contain N-terminal mutations. COS7 cells that had been transfected with wild-type or mutant pRSV-Rev were fixed *in situ* and then stained with a mixture of three Rev-specific primary antisera, followed by a fluorescent second antibody. Approximately 5% of cells in each population showed detectable expression of the transfected gene by this assay. The results shown are typical of two or more independent experiments with each construct.

UCSF LIBRARY





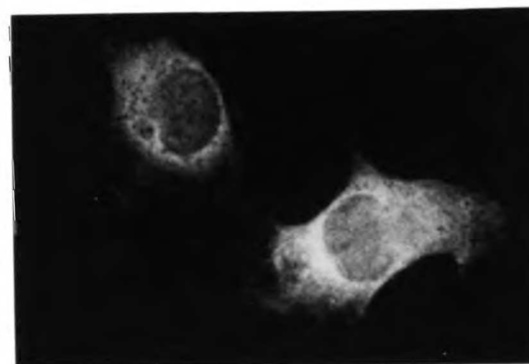


Figure 7

UCSF LIBRARY

Although the GR moiety ensures nuclear translocation of such fusion proteins when steroid is present, signals in the Rev moiety remain essential for nucleolar localization and for transactivation. To test for such signals in the N-terminus, we introduced selected mutations into the expression vector pRSV-Rev/GR, which codes for the Rev/GR fusion protein (Fig 8A). The resulting plasmids were transfected into CV1 cells growing in the presence of a steroid agonist. By Western blotting, each of the mutants was found to be strongly expressed in the transfected cells (Fig. 8D). Every mutant tested(Δ4-26, MA1, MA4 and MA5) also exhibited a wild-type pattern of localization in hormone treated cells (e.g. Fig. 8B), implying that these mutations did not inactivate the nucleolar targeting signals in Rev. Unlike the wild-type and MA1 fusion proteins, however, proteins that contained the MA4 and MA5 mutations failed to transactivate CAT expression from pDM128 (Fig. 8C). The latter mutations can therefore abolish transactivation even in the absence effect on protein localization.

Malim, *et al.* (26) have demonstrated that certain inactive mutant rev proteins can block the activity of the wild-type protein in *trans*. This dominant-negative phenotype has been observed only for mutations in the C-terminal half of Rev; mutations in the arginine rich tract have yielded only recessive-negative proteins. To determine whether MA4 or MA5 had *trans*-inhibitory activity, we tested the ability of the mutant plasmids to block the function of a wild-type Rev vector in our transfection assay. For comparison, two plasmids (Δ78-114 and M10) encoding C-terminal mutants were also tested. As illustrated in Fig. 9, cotransfection of increasing amounts of these C-terminal variants progressively inhibited the induction of CAT by

UCSF LIBRARY

Figure 8. Function and subcellular localization of mutant Rev/GR fusion proteins in the presence of steroid hormone. (A) Structure of pRSV-Rev/GR. This vector was derived by inserting cDNA sequences encoding the steroid-binding domain of the rat GR into codon 116 of pRSV-Rev in frame (42). Hatching indicates locations of the arginine-rich tract and of the MA4 and MA5 mutations; the black rectangle marks residues 75 to 90. (B) Nucleolar localization of wild-type Rev/GR and representative mutants in the presence of dexamethasone. Proteins were detected by using primary antisera specific for the Rev moiety. (C) Transactivation of CAT expression by mutant and wild-type Rev/GR. CV1 cells were cotransfected with pDM128, pCH110, and a pRSV-Rev derivative and were then grown in the continuous presence of the steroid agonist dexamethasone (10^{-5} M) for 36 h prior to harvest. As expected (42), none of the Rev/GR derivatives showed detectable activity in the absence of steroid (data not shown). (D) Western blot hybridization of Rev and Rev/GR expression in transfected COS7 cells. Proteins were detected by using Rev-specific antisera after fractionation on a denaturing 12% polyacrylamide gel. kD, Kilodaltons.

UCSF LIBRARY

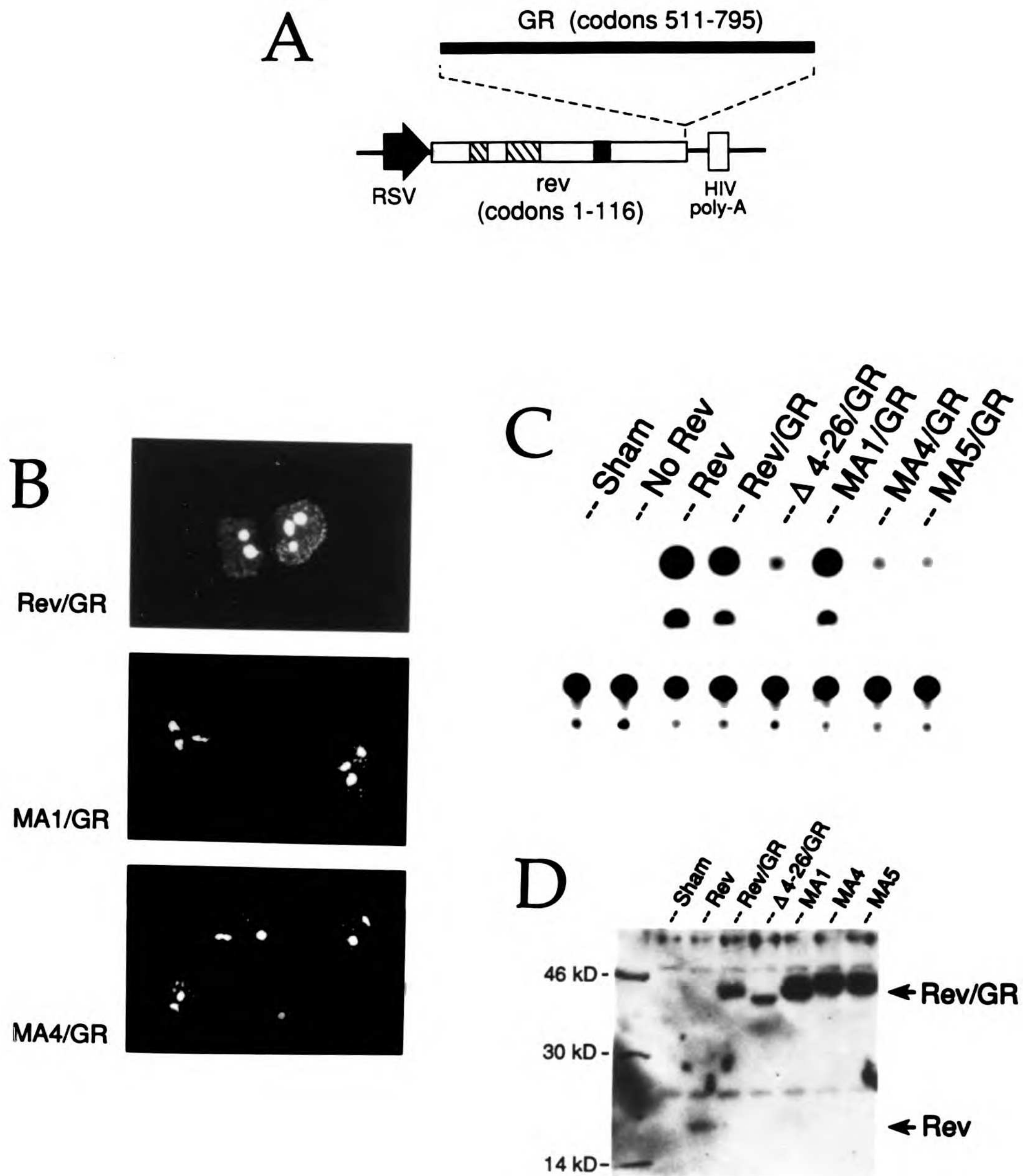


Figure 8

UCSF LIBRARY

Figure 9. Lack of *trans*-dominant inhibitory activity in N-terminal mutants of the Rev protein. CV1 cells were cotransfected with pDM128, pCH110 and 5 μg of ptk-Rev along with the indicated amounts of wild-type ptk-Rev or pRSV-Rev (competitor) plasmid. CAT expression was assayed 36 h later and is expressed relative to the value observed in the absence of any competitor plasmid. Steady-state expression of Rev protein from ptk-Rev was substantially lower than that produced by an equimolar amount of pRSV-Rev (data not shown). M10 is the pRSV-Rev equivalent of a mutant derived by Malim *et al.* (26) and is identified as LE 78, 79 DL in Fig. 5B.

UCSF LIBRARY

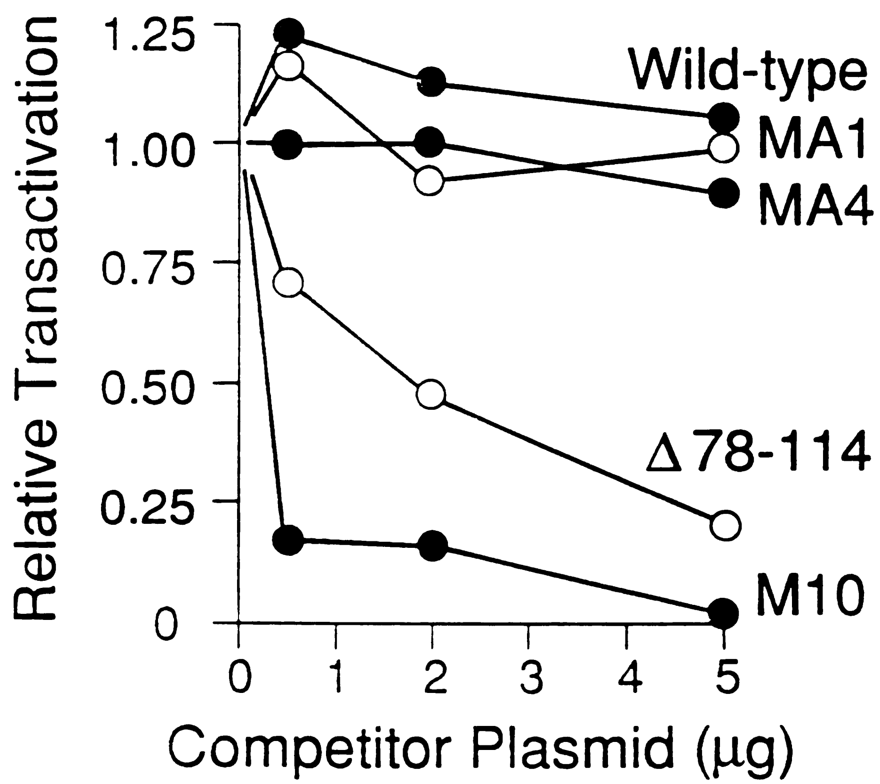


Figure 9

UCSF LIBRARY

wild-type Rev. As little as 5 μg of M10, a mutant previously shown to be strongly dominant-negative (26), was sufficient to achieve 98% inhibition of transactivation in this assay. In contrast, neither MA4 nor MA5 produced significant inhibition under any of the conditions tested. We therefore conclude that both MA4 and MA5 are recessive-negative mutants.

DISCUSSION

The Rev transactivator of HIV-1 is a small RNA-binding protein that is thought to interact with certain incompletely spliced HIV-1 transcripts in the nucleus (14,15,45). As a result of this interaction, the transcripts are selectively exported to the cytoplasm, where they direct the synthesis of essential virion proteins (11,13,46). The precise mechanism of Rev action remains unknown, but its importance in the viral life cycle is clear: HIV-1 proviruses that cannot express functional Rev protein also fail to generate infectious viral particles (8,10,47). It appears likely that this form of posttranscriptional control is critical for other primate retroviruses as well. Thus, similar Rev proteins are expressed by HIV-2 and by simian immunodeficiency virus (SIV) (20), and a structurally distinct protein (termed Rex) produced by human T-cell leukemia virus type I (HTLV-I) appears to serve a virtually identical function (25). The effect of each of these proteins is mediated by a *cis*-acting RNA target element found in transcripts from the corresponding virus. Although their sequences are unique to each virus, some of these proteins and response elements can substitute for one another under certain conditions (20, 24, 31), implying that they may act through a common pathway.

This study is an extension of earlier work (26) which delineated two

UCSF LIBRARY

essential regions in the Rev protein of HIV-1. One of these spans the arginine-rich tract at amino acids 35 to 50 and has also been shown to extend to near residue 25; the other lies in the C-terminal half of the protein. Although mutations in either region can inactivate Rev, the resulting mutant proteins have markedly dissimilar properties. Thus, mutations in the arginine-rich region tend to be recessive negative and to disrupt localization of the protein, whereas mutations near amino acid 80 impair function without affecting localization and can exhibit a dominant-negative phenotype.

Here we have presented the results of a further mutational analysis of Rev. Using a reporter plasmid to detect Rev transactivation in vivo, a series of missense mutations was tested that together involved over 40% of the residues in the N-terminal half of the protein. While generally confirming the results of earlier studies, these data also reveal that point mutations at positions 14 to 16 or 18 to 20 can completely abolish Rev function. The N-terminal domain defined by these mutations (MA4 and MA5) coincides with a relatively basic segment of the protein in which three of seven consecutive amino acids carry positive charges. Indeed, the density of basic residues in this short segment of the proteins that of the nearby arginine tract: the mutant proteins do not accumulate normally within nuclei and fail to act as *trans*-dominant inhibitors of wild-type Rev. In light of these similarities, it is reasonable to view the N-terminal locus together with the arginine-rich tract as a single contiguous domain, spanning residues 14 to 50, that is essential for the localization and function of Rev.

When transported passively into the nucleus by fusion with the GR, Rev proteins containing certain mutations in the arginine-rich region are

UCSF LIBRARY

selectively excluded from nucleoli (42), suggesting that the affected residues may target Rev to that organelle. Other mutations in this region, however, inactivate the Rev/GR fusion protein without affecting its localization (42). The latter finding provides support for the view that sequences in the arginine-rich tract form part of the RRE-binding domain of Rev, just as similar arginine-rich motifs are thought to function in a variety of eucaryotic and procaryotic RNA-binding proteins (48). In this study, we found no evidence that sequences encoded by the first exon are required for nucleolar localization of Rev/GR. The finding that mutations MA4 and MA5 inactivate this fusion protein (Fig. 8) therefore implies that these mutations interfere directly with transactivation.

A comparison of predicted Rev protein sequences from all available HIV-1 isolates (49) reveals that, although individual residues in the vicinity of positions 14 to 20 may vary, the overall structure is conserved (Fig. 10). three positively charged residues, separated from one another by pairs of nonpolar amino acids, are predicted at corresponding positions in every known HIV-1 sequence. Interestingly, Rev sequences from HIV-2 and closely related strains of SIV also predict a very similar organization in this region but are distinguished by the presence of additional charged residues in several positions. By contrast, there is no recognizable counterpart of this N-terminal motif in the HTLV-I Rex protein. One obvious possibility is that, when present, the basic side chains in this region could contribute to RNA binding through electrostatic interactions. This might account for the observation (20) that a fusion protein formed by linking the first exon of HIV-1 *rev* with the second exon of HIV-2 *rev* can transactivate through the HIV-1 RRE, whereas the HIV-2 Rev protein itself cannot. It is also consistent with reports that

UCSF LIBRARY

Figure 10. Conserved structure of an essential N-terminal motif in Rev proteins of primate immunodeficiency viruses HIV-1, HIV-2, and SIV. Predicted sequences are shown for amino-acids 12 to 23 of HIV-1 Rev and for corresponding residues in HIV-2 and SIV Rev proteins. For each viral species, a sequence from one representative strain is depicted at the top, and any alternative amino-acids found in other strains are listed below (in arbitrary order) for each position. Positively charged residues are shown by open letters, and the proposed conservation of these residues are shown by dashed-line boxes. Sequences and alignments are based on data compiled in ref. 48 and comprise 13 North American and Zairean isolates of HIV-1, seven isolates of HIV-2, and three isolates of macaque or sooty mangabey SIV.

UCSF LIBRARY

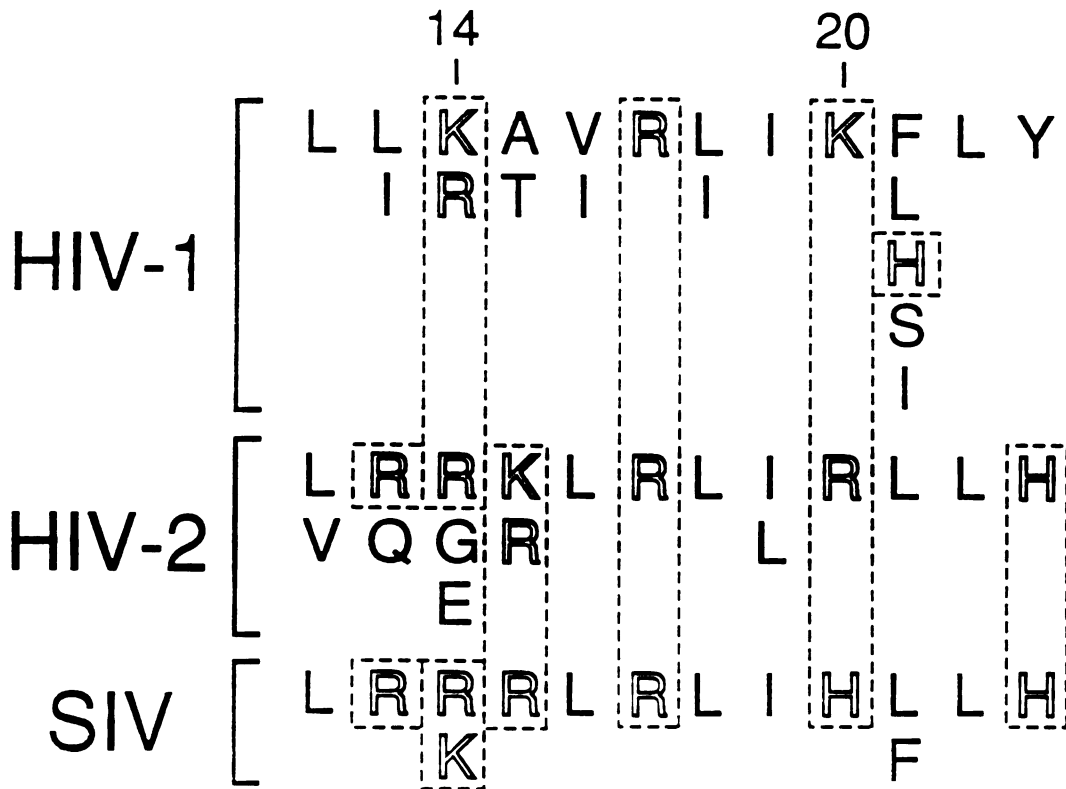


Figure 10

UCSF LIBRARY

hybrid HIV-1 proteins containing only the second rev exon are partially (50) or completely (50, 51) defective in transactivation. It is possible, then that this novel N-terminal motif may be a principal determinant of RRE-binding specificity in Rev proteins from the primate immunodeficiency viruses.

The development of in vitro RRE-binding assays (14-17, 45) will permit a direct test of this hypothesis. A more speculative possibility is that this conserved motif mediates specific protein-protein interactions that may be essential for transactivation in vivo. The existence of such interactions, however, has not yet been demonstrated.

UCSF LIBRARY

CHAPTER IV:
A MINIMAL REV-RESPONSE ELEMENT FOR HIV-1

Rev transactivation requires direct binding of Rev protein to the Rev-response element (RRE) present in the regulated mRNA. On the basis of functional mapping, Malim et al. (13) postulated that the RRE coincides with a 240-base region of RNA secondary structure in the *env* coding region; subsequent studies have confirmed both the existence of such a structure in vivo (52,53) and its importance for transactivation (15-17, 44, 54). Rev protein binds specifically to a discrete 30- to 71-base region within the RRE (16, 17), and mutations that disrupt this binding region abolish transactivation (17, 54, 55), The importance of sequences elsewhere in the RRE, however, remains uncertain: mutations at a variety of sites in the 240-base element can severely inhibit Rev binding and transactivation (14-17, 44, 54, 55), but it is not known whether these sites have independent functions or merely affect folding or accessibility of the Rev binding locus. To address this issue, we attempted to define the minimal sequence required for RRE activity. We found that an 88-nucleotide truncated RRE encompassing the Rev binding site is sufficient to mediate transactivation in vivo and that two tandem copies of this sequence function as efficiently as the intact RRE. These findings imply that most of the RRE has no unique role in transactivation and suggest that Rev binding might be the only function of the RRE.

UCSF LIBRARY

MATERIALS AND METHODS

Plasmid constructions. pDM138 was prepared from pDM128 by replacing a 1.2-kb *Stu* I-*Bsm* I (HIV nt 6852 to 8066) fragment with a *Cla* I linker. The 318-bp DNA fragment encompassing the RRE was prepared from pDM128 by the polymerase chain reaction with a pair of oligonucleotide primers with the sequences 5′-GGG-GGATCCATCGATAGGCAAAGAGAAGAG-3′ and 5′-GGGGGATCCATCGATGC-AAATGAGTTTTT-3′; the amplified fragment corresponded to HIV-1 nucleotides 7725 to 8043, flanked on either side by a 15-bp linker containing *Cla* I and *BamH* I sites (underlined). This fragment was trimmed with *BamH* I and cloned into pUC118 for oligonucleotide-directed mutagenesis (38); wild-type or mutant RRE derivatives were then excised with *Cla* I and subcloned into the unique *Cla* I site of pDM138. DNA encoding the bul1 mutant and its derivatives was synthesized with two partially complementary oligonucleotides that together encompassed the mutant along with flanking *Cla* I sites; these oligonucleotides were annealed and extended with Klenow polymerase, trimmed with *Cla* I, cloned in pUC118 and subsequently subcloned into the pDM138 *Cla* I site.

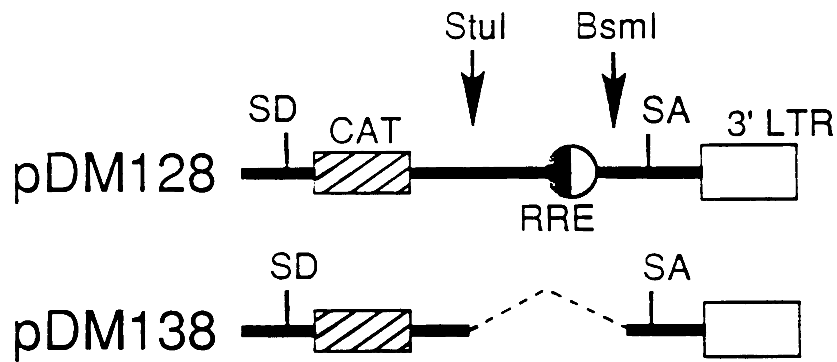
RESULTS AND DISCUSSION

We derived a defective reporter (pDM138) by deleting 1.2-kb of intron sequence from pDM128 which contains the RRE (Fig. 11A). Various RRE derivatives were then compared for their ability to restore transactivation when inserted at the site of the deletion. Insertion of a 318-bp DNA fragment encompassing the entire 240-base RRE conferred a 23-fold CAT induction in response to Rev (Table 1). The failure of this insertion to restore the full

UCSF LIBRARY

Figure 11. Rev reporter constructs and the RRE. (A) A 1.2-kb *Stu* I-*Bsm* I fragment (HIV-1 nt 6852 to 8066) was excised from pDM128 and replaced by a *Cla* I linker to produce pDM138. The unique splice donor (SD) and acceptor (SA) sites and the RRE are indicated. LTR, Long terminal repeat.
(B) Predicted secondary structure of the RRE from HIV-1 strain SF2, redrawn from ref. 13. Note that the modifications of the structure referred to in Fig. 2 are not included here.

UCSF LIBRARY



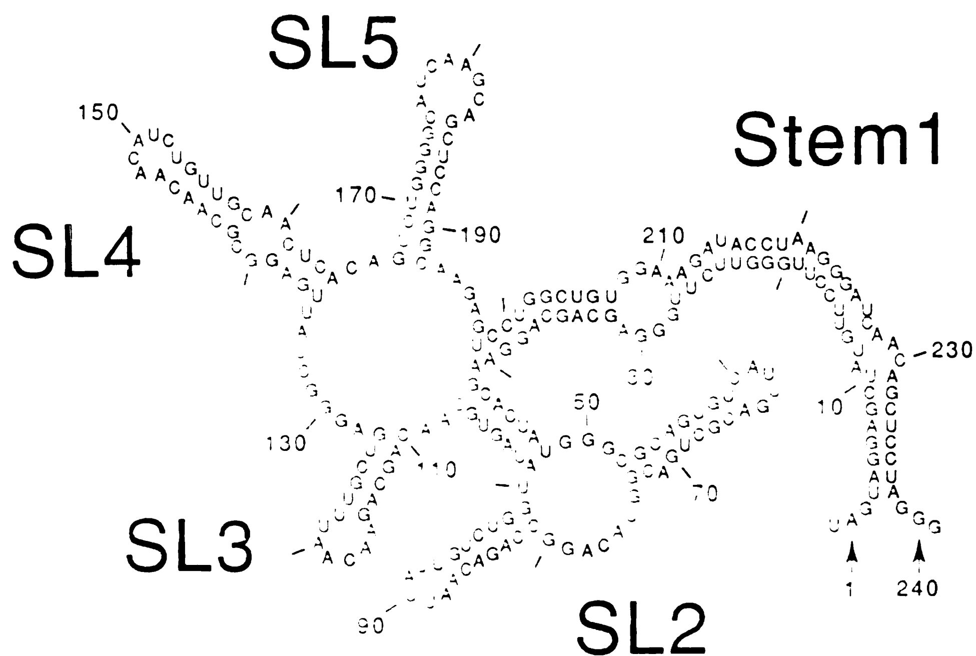


Figure 11

UCSF LIBRARY

Table 1. Functional characterization of RRE mutants in the cotransfection assay. (a) Wild-type or mutant RREs were inserted into the pDM138 *Cla* I site. Numbers in parenthes are the nucleotide positions deleted. (b) Relative CAT activity is expressid as % acetylation normalized to the β-glactosidase internal control. (c) Values exceeded linear range of the assay and were not used in calculating mean response.

UCSF LIBRARY

Table 1

Construct[a]	Presence of Rev	Relative CAT activity[b]				Mean Rev response (fold)
		Expt 1	Expt 2	Expt 3	Expt 4	
pDM128	+			43.5, 53.0		110
	-			0.5, 0.4		
pDM138	+			6.2, 3.0		1
	-			4.4, 3.0		
240-base RRE	+	87.7[c]	94.6[c]	69.0, 47.6	72.1	23
	-	3.5	6.0	3.1, 1.9	3.2	
ΔSL2 (43-106)	+	22.6	69.0			2
	-	11.5	37.0			
ΔSL2A (82-98)	+	4.4	12.8			2
	-	2.1	7.3			
ΔSL2B (53-72)	+	28.4	94.7			8
	-	3.5	11.2			
ΔSL3 (111-126)	+	56.4	77.3			186
	-	0.5	0.3			
ΔSL4 (137-163)	+	6.2	4.0			9
	-	0.4	1.0			
ΔSL5 (168-191)	+	13.6	54.5			22
	-	0.6	2.5			
bul1RRE						
1 copy	+			16.5, 10.2	29.8	6
	-			2.9, 2.2	4.5	
2 copies	+			47.7, 26.0	25.2	22
	-			1.4, 1.4	1.8	
4 copies	+			23.4, 17.7	21.7	7
	-			2.4, 2.8	5.9	
ΔSL2A (82-98)	+			5.2	1.9	<1
	-			5.5	5.1	
ΔSL2B (53-72)	+			5.1	2.6	<1
	-			4.2	5.6	

UCSF LIBRARY

100-fold response seen with pDM128 may suggest the presence of additional regulatory sequences elsewhere in the deleted region. For this study, however, a 23-fold CAT induction was considered the wild-type response attributable to the RRE.

The 240-base RRE has been proposed (13) to form a double stranded stem (Stem 1) crowned by four stem-loop structures (SL2 to SL5), one of which gives rise to two additional stem-loops, called SL2A and SL2B (Fig. 11B). Consistent with previous reports (52, 54), we found that mutants of the 240-base RRE lacking any of four stem-loops (SL2B, SL3, SL4 or SL5) could support transactivation, but that precise deletion of SL2 or SL2B alone reduced transactivation to very low levels (Table 1). These data confirm that SL2B is the region most critical for RRE function. Like other investigators (52), we observed that some RRE mutations also affected basal CAT activity in the absence of Rev, for reasons that are unknown.

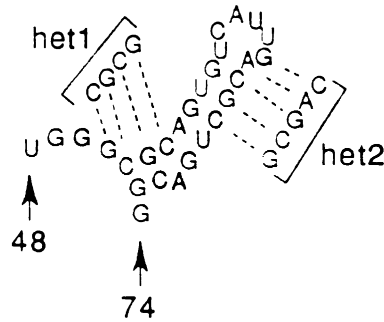
Although point mutations in and around SL2B can inhibit RRE function, compensatory mutations designed to preserve the folding pattern (Fig. 12, structure A) predicted by Malim et al. (13) have been found to restore both Rev binding and transactivation (17, 55). This suggests that the conformation of this region is a major determinant of Rev binding, although primary sequence is also known to play a role (54). A comparison of predicted RRE conformations among different strains, however, led Heaphy et al. (16) to propose an alternative structure for this region (Fig 12, structure B). To test the validity of this alternative model, we constructed two additional mutants of the 240-base RRE. The first (called het1) contained four consecutive point mutations that would completely disrupt structure B but would only partially destabilize structure A. This mutant was found to retain substantial RRE

UCSF LIBRARY

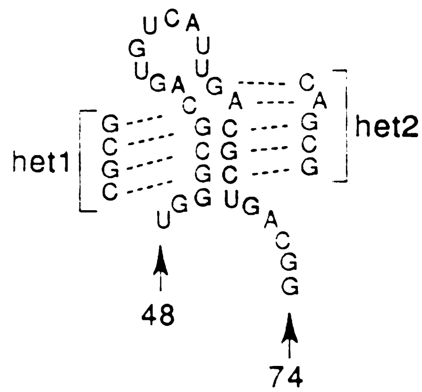
Figure 12. Alternative models for the secondary structure of RRE SLIIB. The structures predicted by Malim *et al.* (13) and Heaphy *et al.* ()for residues 48 to 74 of the RRE are shown (residues are numbered as described in the legend for Fig. 11A) Also shown are two sets of clustered point mutations (het1 and het2) introduced into this region of the 240-base RRE. Under model B, the het2 mutations would be expected to complement het1, restoring all putative base pairing while preserving an unpaired adenosine. Results of a representative CAT assays of reporters containing the wild-type (wt) and mutant RREs are depicted below the models.

LIBRARY

A



B



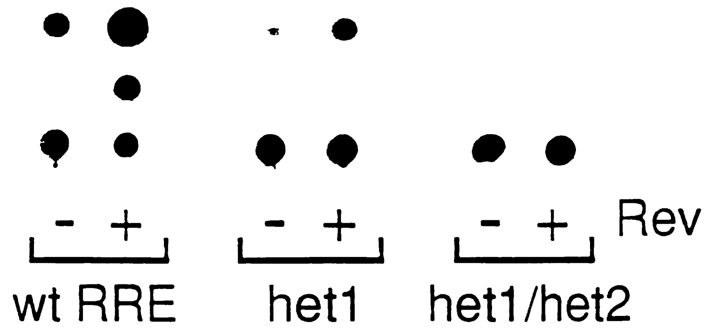


Figure 12

activity (Fig. 12). The second mutant (het1/het2) contained the het1 mutations along with four additional mutations that were predicted either to restore the base pairing of structure B or to further disrupt structure A. As depicted in Fig. 12, the het1/het2 mutant was completely unresponsive to Rev. This limited analysis thus favored structure A for the critical SL2B subdomain.

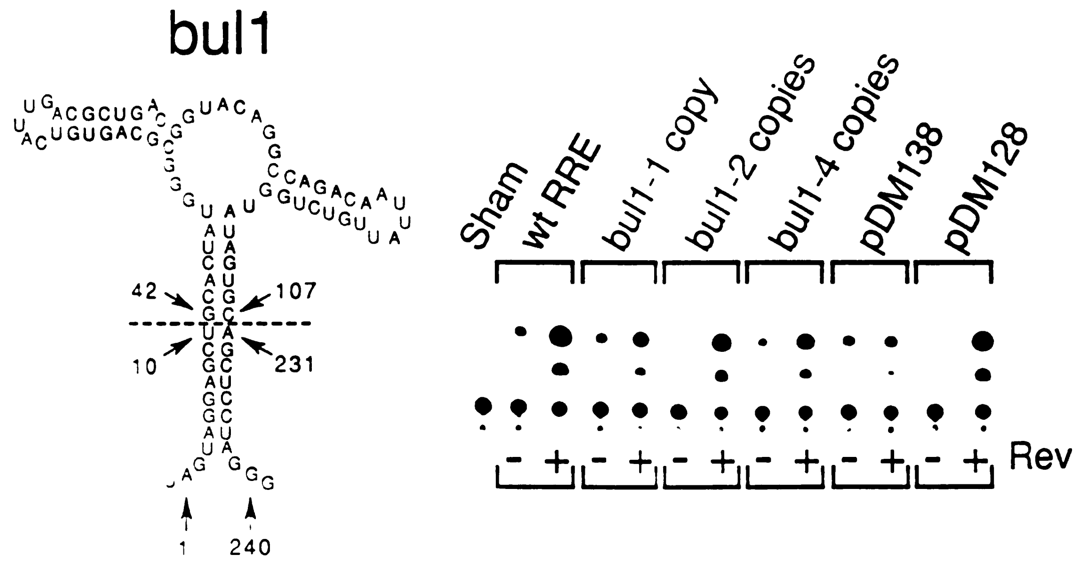
Having confirmed that SL2 is indispensable in the context of the 240-base RRE, we examined whether this element might be sufficient for RRE function. An 88-base truncated RRE (designated bul1) was constructed in which the wild-type SL2 sequence was fused to an 8-base sequence from the base of stem 1, with all intervening and flanking residues deleted. The likely secondary structure of bul1 is depicted in Fig. 13A. When inserted into pDM138 and tested in the cotransfection assay, a single copy of bul1 yielded sixfold transactivation of CAT in response to Rev (Fig. 13A and Table 1). Insertion of two tandem copies of bul1 yielded 22-fold transactivation, a level of responsiveness comparable to that of the 240-base RRE. Curiously, four tandem copies of bul1 showed a degree of responsiveness similar to that observed with a single copy; this might reflect the increased likelihood that adjacent copies will anneal with one another to form aberrant, nonfunctional secondary structures. Two truncated forms of bul1, designated bul1-ΔA and bul1-ΔB, that lacked SL2A and SL2B, respectively (Fig. 13B), failed to respond to Rev when present in a single copy (Table 1).

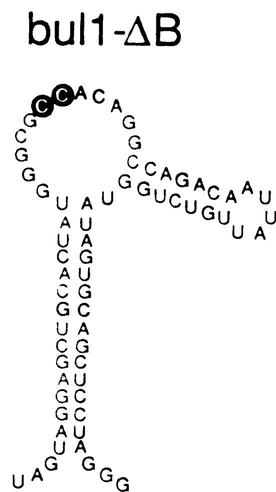
This study extends previous mapping of the essential features of the RRE. Although earlier reports have shown that SL2 is indispensable for function, mutations in other regions of the RRE have also been reported to inhibit Rev

UCSF LIBRARY

Figure 13. An 88-base truncated RRE is sufficient for Rev transactivation. (A) Sequence and predicted structure of the bul1 mutant RRE (numbered as described in the legend for Fig. 11B), with representative CAT assay results for pDM128, pDM138 and pDM138 derivatives containing either the 240-base RRE or the indicated number of tandem copies of bul1. Lane Sham, Cells transfected with no reporter plasmid. Residues from stem 1 were included to promote proper folding. (B) Two nonfunctional derivatives of the bul1 RRE. Each mutant was constructed from synthetic oligonucleotides as described. Point mutations (open lettering) introduce a unique *Stu* I or *Nar* I site at the site of deletions in bul-1ΔA and bul-1ΔB, respectively.

UCSF LIBRARY





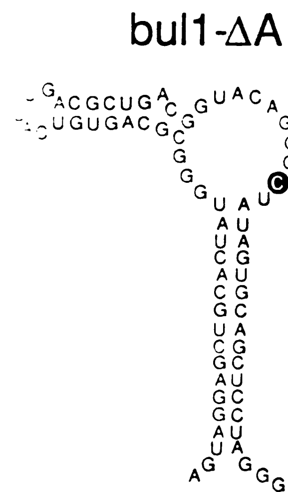


Figure 13

binding (14, 44, 55) and transactivation (17, 52, 55). This has prompted speculation that the latter regions might serve as binding sites for unidentified cellular factors that are need for transactivation or that Rev binding might require two or more noncontiguous regions of the RRE (14, 16, 55). This study has shown that an 88-base truncated RRE (bul1) derived almost entirely from SL2 is sufficient to mediate Rev transactivation in vivo and that tandem copies of bul1 can mediate a degree of responsiveness comparable to that of the full-length RRE. Sequences from the RRE that are not present in bul1 thus have no unique role in transactivation and make no more than a quantitative contribution to RRE function.

The SL2 sequence contained in bul1 is the minimal region known to bind Rev protein in vitro (16, 17). While this suggests that the ability to bind Rev may be the only requirement for RRE function, we cannot exclude the possibility that SL2 might also contain binding sites for unknown accessory factors or that Rev might act by displacing such factors from this region of the RRE. The finding that two tandem copies of bul1 are needed to fully replace the wild-type response element suggests that some regions in the RRE may be functionally redundant, perhaps serving as secondary sites for cooperative binding of Rev or other factors. For example, deletion of the SL2A arm completely inactivates bul1 but only partially inactivates the 240-base RRE (Table 1), suggesting that other sequences can compensate for the loss of SL2A in its normal context. The comparatively small size of bul1 may make it a useful model for further investigation of these issues as well as for structural analysis of a biologically active RRE.

UCSF LIBRARY

CHAPTER V:

EFFECTOR DOMAINS OF HIV-1 REV AND HTLV-I REX ARE FUNCTIONALLY INTERCHANGEABLE AND SHARE AN ESSENTIAL PEPTIDE MOTIF

HIV-1 Rev and HTLV-I Rex are comparable in many respects: both are nuclear phosphoproteins that associate preferentially with nucleoli, and each acts by binding directly to its target transcripts at its response element sequence, the RRE or the XRE. Indeed, Rex can functionally replace Rev under certain conditions (24, 32), suggesting that both proteins may act through a common pathway. Yet there is little apparent sequence similarity between these two proteins that could account for their similar effects. To search for peptide domains within Rev and Rex that serve related functions, we have constructed and characterized a series of fusion proteins containing sequences from both transactivators. Here we demonstrate that short effector domains which mediate the regulatory activities of Rev and Rex are functionally interchangeable and we identify a tetrapeptide present in both domains that may be a core effector motif for transactivators of this type.

MATERIALS AND METHODS

Plasmid constructions. All transactivator expression vectors (Fig. 14A) were derivatives of pRSV-Rev. To construct Rex and Rev/Rex vectors, *rev* codon 116 in pRSV-Rev was mutated to a *Bgl* II site (Asp replaces Glu), and a Rex cDNA with appropriate flanking restriction sites was either inserted in

place of the *rev* sequence (*Sac* I-*Bgl* II) or fused to the 3′ end of *rev* in frame (*Bgl* II-*Bgl* II). In all fusions, Leu replaces Met at *rex* codon 1. The XRE fragment was prepared from cloned HTLV-I by polymerase chain reaction with primers 5′-CGCGGATCCATCGATCGATATAAACTAGCAGGAGTC-TAT-3′ and 5′-CGCGGATCCATCGATCGATCTCGAGAGTTGAGCAAGCA-GGGTC-3′; the amplified fragment was trimmed with *Bam* HI and cloned in pUC 118, and then subcloned into the Cla I site of pDM138.

Western immunoblot analysis. Immunoblots were performed as described in Chapter III. Rex proteins were detected with an antipeptide serum (provided by W.C. Greene, Duke University) against the Rex C-terminus; because this Rex antiserum reacted with a background of cellular proteins in the 30- to 46-kDa range, the anti-Rev serum was used to detect fusion proteins.

RESULTS

Previous studies have delineated at least two essential functional domains in the 116-amino acid HIV-1 Rev protein. Sequences in the N-terminal half of Rev provide signals for nuclear and nucleolar localization and also mediate high-affinity binding of the RRE, whereas the C-terminal region known as the Rev effector domain is believed to interact with unidentified cellular factors required for transactivation (26-29, 42). The functional organization of the 189-amino acid Rex protein of HTLV-I is less well characterized, but protein localization signals have been identified at the N- and C-termini (30, 31), and mutations at several other sites within the protein have been found to abolish function (31-33). To extend the mutational analysis of Rev and Rex, and to search for peptide sequences within the two

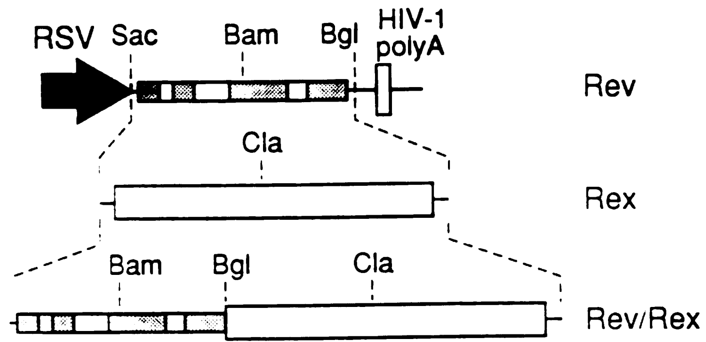
proteins that might serve related functions, a construct encoding a chimeric Rev/Rex fusion protein was prepared and tested for its ability to transactivate through the Rev- or Rex-responsive elements (Fig. 14A).

Under the conditions of this transfection assay, each wild-type transactivator functioned efficiently only through its cognate response element (Fig. 14C, construct 1 and 4); in particular, Rex yielded little or no CAT induction through the RRE. This relatively strict specificity of Rex contrasts with findings in some other transfection systems (24, 56) but is in accord with the observations of at least one other group (57). Although the technical features of an assay that maintain such specificity have not been identified, we speculate that our use of nonreplicating expression vectors may be contributory. As expected, we found that deletion of residues 89 to 116 from Rev had no functional effect (Fig. 14C, construct 2), but that a slightly larger deletion (residues 78 to 114) which removed the effector domain abolished activity completely (construct 3). By fusing the full-length Rex sequence to the C-terminus of Rev, a chimeric protein (construct 5) that could transactivate efficiently through either the RRE or the XRE was produced. Remarkably, the dual activity of this fusion protein was not affected by deletion of the Rev effector domain (construct 6, designated RevΔ/Rex), nor by point mutations that are known (26, 28) to inactivate this domain (constructs 7 and 8). These findings indicated that fusion with the N-terminal portion of Rev could confer RRE specificity onto Rex in this assay and also implied that fusion with Rex could complement loss of the effector domain from Rev.

To map the sequences required to confer RRE specificity, selected missense

UCSF LIBRARY

Figure 14. A Rev/Rex fusion protein transactivates through either the RRE or the XRE. (A) Transactivator plasmids. The expression vector pRSV-Rex was constructed by replacing the Rev coding sequence of pRSV-Rev with that of Rex. pRSV-Rev/Rex was constructed by in-frame fusion of the Rex coding sequence onto the C-terminus of Rev. Open rectangles in the *rev* sequence denote the essential N-terminal region, arginine-rich tract, and C-terminal effector domain. RSV, Rous Sarcoma virus 5′ LTR; Sac, *Sac* I; Bam, *Bam* HI; Cla, *Cla* I. (B) Functional characterization of wild-type, mutant and chimeric transactivators in the cotransfection assay. CV1 cells were transfected with either pDM128 (RRE) or a pDM138 derivative containing the XRE (XRE) along with the indicated transactivators and assayed for CAT enzyme activity 36 h later. Δ indicates in-frame deletion of rev codons; Δ89-116 contained a stop codon at residue 89; LE 78,79 DL is the mutation designated M10 by Malim *et al*. (26). Sham, pUC118 and pCH110 only; None, pUC118, reporter and pCH110 only.



RRE

XRE

Rev

Rex

Sham

None

1 Rev

2 Δ89-116

3 Δ78-114

4 Rex

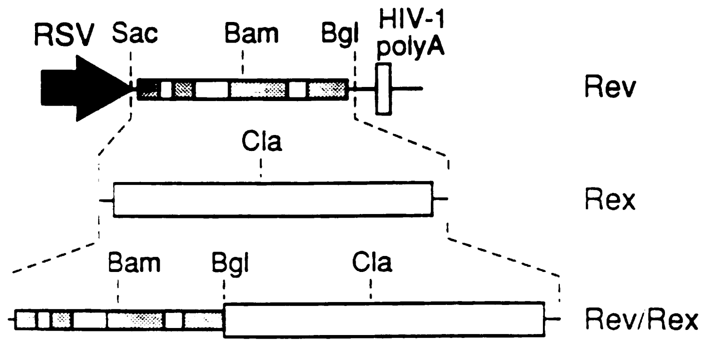
5 Rev / Rex

6 RevΔ/ Rex

7 LE 78,79 DL

8 CC 85,89 RK

Figure 14



RRE | XRE

Rev | Rex

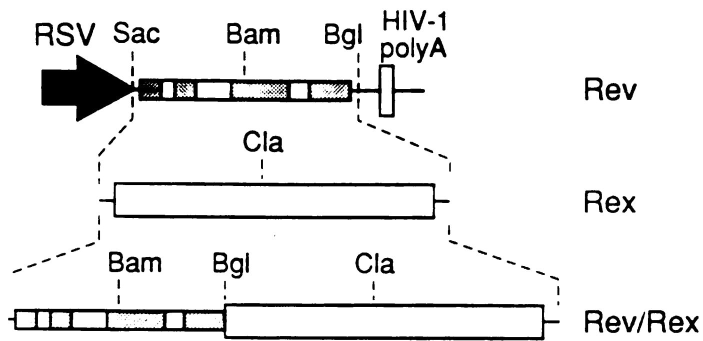
Sham
None
1 Rev
2 Δ89-116
3 Δ78-114
4 Rex
5 Rev / Rex
6 RevΔ / Rex
7 LE 78,79 DL
8 CC 85,89 RK

Figure 14

mutations were introduced into the Rev moiety of RevΔ/Rex (Fig. 15, constructs 9 to 15). Each mutation had previously been found to inhibit binding of Rev to the RRE in vitro (29, 58, 59) and to inactivate Rev in vivo (26, 28). In the context of the fusion protein, five of these mutations selectively eliminated transactivation through the RRE; the latter mutations each involved sequences within the arginine-rich tract in Rev (amino acids 35 to 50) that is known to be critical for RRE binding (29, 58, 59). By contrast, two other mutations (involving Rev amino acids 14 to 20) had no detectable effect, suggesting that the function of these sequences was complemented by Rex. Indeed, we found that large simultaneous deletions could be introduced at both Rev termini (amino acids 4 to 26 and 78 to 114) without diminishing RRE-specific transactivation (construct 16). Parallel testing confirmed that all seven mutants could transactivate through the XRE, implying that the proteins were stably expressed and that signals in the intact Rex moiety could ensure appropriate localization within the nucleus. Fusion with a 51-amino acid region that includes the arginine-rich tract of Rev is therefore necessary and sufficient to confer RRE specificity onto Rex in this assay.

Mutational analysis was next performed on the Rex moiety in RevΔ/Rex to map the sequences that could complement deletion of the Rev effector domain. Representative functional data are shown in Fig. 16A and Table 2; the size and stability of selected mutant proteins were confirmed by immunoblot analysis of the transfected cells (Fig. 16B). Missense mutations involving known protein localization signals (30, 31) at the N- and C-termini of Rex (constructs 17 and 22) did not inactivate RevΔ/Rex, presumably because equivalent signals are present in the Rev moiety (26, 42). Three other

Figure 15. Fusion with N-terminal Rev sequences confers RRE specificity onto Rex. Reporter plasmids containing the RRE or the XRE were cotransfected with vectors encoding Rev, Rex, Rev/Rex, or various derivatives of RevΔ/Rex. Naming of missense mutations in the Rev moeity of RevΔ/Rex indicates native and mutant amino acids (in single-letter code) at the affected positions. Δ4-26 denotes in-frame deletion of *rev* codons 4 to 26. Mutant vectors were prepared by replacing the 5′ *Sac* I-*Bam* HI fragment of construct 5 (Fig. 14B) with corresponding fragments from pRSV-Rev mutants described previously (Chapter 3). None, reporter with no transactivator.



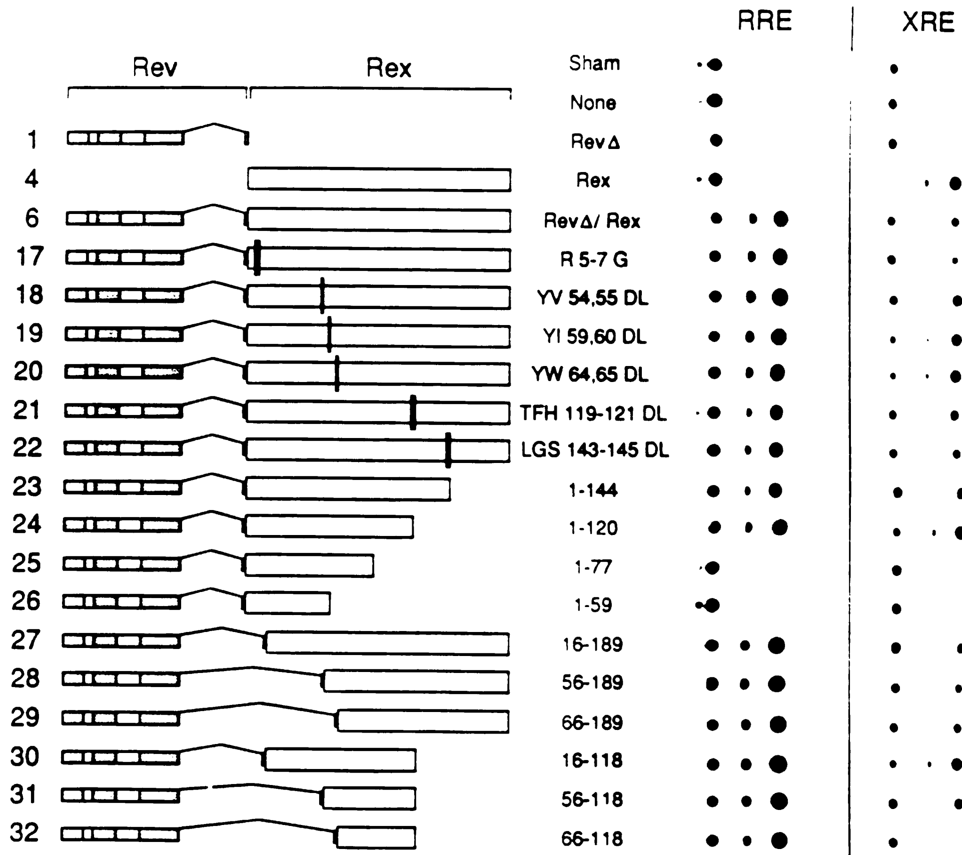
RRE XRE

Rev Rex

Sham
None
1 Rev
2 Δ89-116
3 Δ78-114
4 Rex
5 Rev / Rex
6 RevΔ/ Rex
7 LE 78,79 DL
8 CC 85,89 RK

Figure 15

Figure 16. Mutational analysis of the Rex moiety in RevΔ/Rex. (A) Functional analysis. CAT enzyme activity was assayed in CV1 cells that had been transfected with indicated transactivator constructs together with reporter plasmids containing either the RRE or the XRE. Mutations were prepared by oligonucleotide-directed mutagenesis of a cloned Rex cDNA and were then subcloned into expression vectors. Missense mutations in the Rex moiety are named as described in the legend to Fig. 15; each introduces a unique *Bgl* II site. The mutations in constructs 17 and 19 to 22 are similar or identical to those designated M1, M6, M7, M13, and M15, respectively, in a previous study (31). Truncated mutants (constructs 23 to 32) were prepared by cleaving *rex* at the *Bgl* II site introduced by a missense mutation or at an endogenous *Cla* I site (codons 78 and 79) and are named according to the *rex* amino-acid residues present. (B) Immunoblot analysis of mutant proteins. Extracts (60 µg of total protein per lane) of transfected COS7 cells were fractionated on a denaturing sodium dodecyl sulfate-14% polyacrylamide gel and probed with antipeptide antisera specific for the N-terminal half of Rev. Sham, pUC118 and pCH110 only; None, pUC118, pCH110 and reporter only.



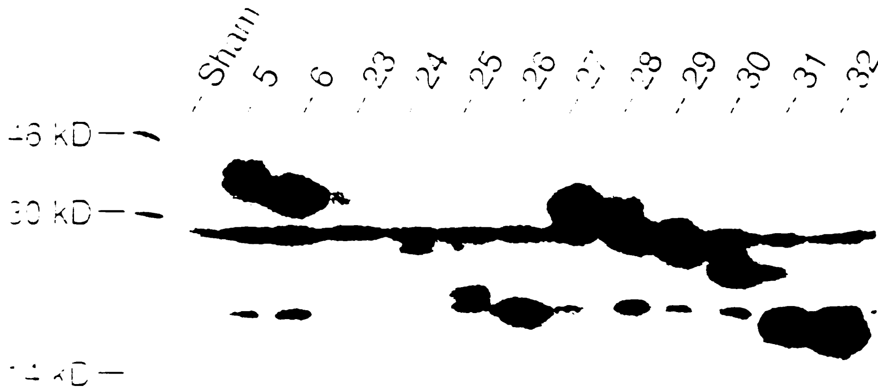


Figure 16

Table 2. Transactivation by RevΔ/Rex chimeras through the RRE or the XRE. (a) Constructs are numbered as in Fig. 17. (b) Ratio of CAT expression in CV1 cells transfected with and without each indicated construct. Data are from two independent experiments. Acetylation of chloramphenicol was assayed by thin-layer chromatography and quantified by scintillation counting. All data were corrected for the trace acetylation produced by extracts from untransfected cells and were normalized within each experiment relative to the β-galactosidase internal control.

Table 2

Construct[a]	Induction (fold) of CAT activity[b]			
	RRE		XRE	
	Expt 1	Expt 2	Expt 1	Expt 2
3		2.8		1.8
4	2.1	4.1	38.0	40.4
6	44.2	89.6	10.9	4.7
17	52.9	102	5.1	2.6
18	52.3	131	9.5	7.7
19	57.7	137	17.6	8.9
20	59.1	97.3	17.5	10.4
21	36.9	53.5	8.1	6.0
22	34.7	58.4	5.8	4.8
23	29.4	58.3	8.8	2.7
24	43.7	83.3	12.4	7.3
25	1.4	1.9	0.8	0.6
26	1.7	2.2	1.3	1.0
27	58.7	97.2	6.3	2.9
28	54.1	110	7.0	4.9
29	59.0	116	9.6	3.7
30	54.0	120	11.2	9.8
31	53.3	118	5.4	4.1
32	50.8	84.0	1.3	1.7

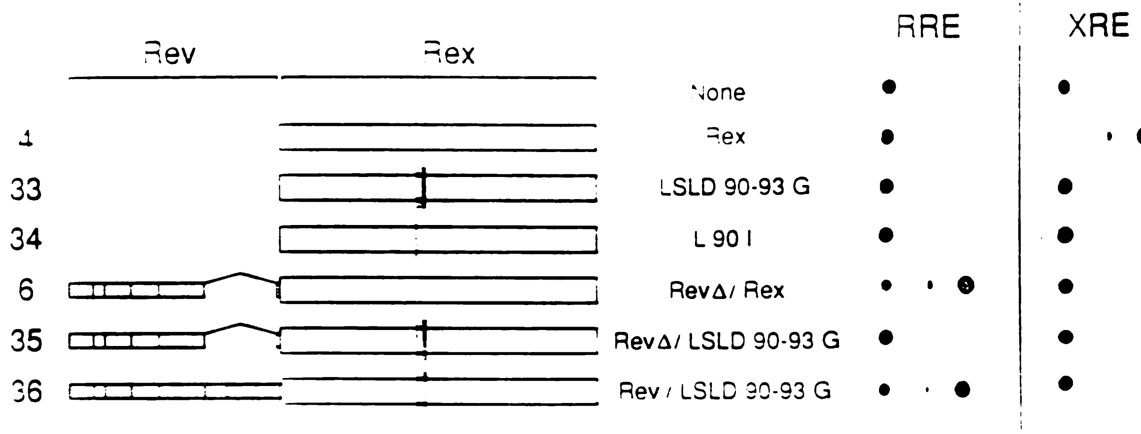
missense mutations that have been reported (31, 32) to abolish effector function in Rex failed to inactivate the fusion protein (constructs 19 to 21). Progressive truncations of the Rex moiety (constructs 23 to 29) revealed that up to 79 C-terminal or 65 N-terminal residues could be deleted without eliminating function, although simultaneous deletion of both of these regions profoundly inhibited XRE-specific transactivation (construct 32). Most importantly, two chimeras that lacked the central portion of Rex failed to transactivate through either response element (construct 25 and 26), whereas a fusion protein comprising only amino acids 66 to 118 of Rex (construct 32) transactivated through the RRE as efficiently as wild-type Rev.

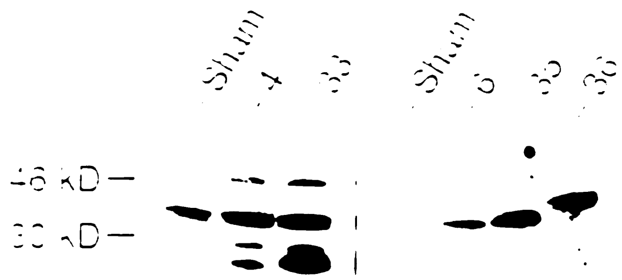
These findings indicated that a 52 amino-acid internal segment from Rex could functionally replace the effector domain of Rev. Comparison of the of these two regions (Fig. 17A) revealed the shared tetrapeptide sequence Leu, Thr or Ser, Leu, Asp. Moreover, related (though not identical) sequences are present at comparable locations in a number of other known retroviral transactivators of this type, including HTLV-II Rex, Rev proteins from two related families of primate immunodeficiency lentiviruses (simian immunodeficiency virus type sm SIV_{sm}/HIV-2 and SIV_{agm}), and the functionally equivalent OrfL/Rev-V protein of Visna, an ungulate lentivirus (22, 60). In the HTLV-I and Visna proteins, this motif encompasses the site of mutations that have been reported (31, 61) to inhibit transactivation.

We therefore asked whether this tetrapeptide was required for Rex effector activity. As shown in Fig. 17B and Table 3, Rex was completely inactivated when all four residues in the tetrapeptide were replaced by glycines (construct 33) and that even a highly conservative mutation (isoleucine replacing leucine) at the first position of the tetrapeptide strongly inhibited function

Figure 17. A conserved and essential effector motif in posttranscriptional transactivator proteins. (A) Conserved tetrapeptide in HIV-1 Rev, HTLV-I Rex and equivalent proteins from other complex retroviruses. Sequences of effector domains from the HIV-1 (BRU strain) and HTLV-I proteins are aligned with respect to the proposed conservation (solid rectangle). Partial sequences of HTLV-II Rex, Visna OrfL/Rev-V (22, 60) protein, and Rev proteins from representative strains of HIV-2 (strain ROD), SIV_{sm} (strain 251), and SIV_{agm} (strain TYO) are also shown. In proteins in which the tetrapeptide is most divergent from the consensus sequence, a second copy may be present (dashed rectangle). Mutations at the nine underlined sites in HTLV-I Rex have no functional effect (31, 32). (B) Mutation of the tetrapeptide in Rex eliminates function and is complemented by the effector domain of Rev. CAT assays with CV1 cells transfected with the indicated transactivator and reporter plasmids were performed. (C) Immunoblot analysis of transfected COS7 cells expressing the indicated constructs. Extracts (100 µg of total protein per lane) were fractionated on denaturing sodium dodecyl sulfate-14% polyacrylamide gels and blotted with the indicated antisera. Sham, pUC118 only; None, reporter only.

HIV-1	69- EPVPLQLPPLERLTLD CNEDCG -90
HTLV-I	66- PPVQS I RSPGTPSMDALSAQLYSS LSLD SPPSPPREPLRPSRSLPRQSL I QPP -118
HTLV-II	67- PPVQSTNSPGTPSMDALSALLSNT LSLA SPPSPPREPQGPSRSLPLPPLLSPP -119
Visna	100- PNMESNMVGMEN LTLE TQLEDNALYN -125
HIV-2	67- SPLDQTIQHLQG LTIQ ELPDPPTHLP -92
SIV_{sm}	67- TPLDLAIQQLQN LAIE SIPDPPTNTP -92
SIV_{agm}	60- YAVDRLADEAQH LAIQ QLPDPPHSA -64





Anti-Rex Anti-Rev

Figure 17

Table 3. Effects of mutations in the conserved tetrapeptide motif on the ability of Rex and of Rev/Rex chimeras to transactivated through the XRE. (a) Constructs are numbered as in Fig. 17. (b) Ratio of CAT expression in CV1 cells transfected with and without each indicated construct. Data derived and expressed as described in Table 2, footnote (b).

Table 3

Construct[a]	Induction (fold) of CAT activity[b]				
	Expt 1	Expt 2	Expt 3	Expt 4	Mean +/- SEM
4	50.6	30.6	29.5		36.9 +/- 9.7
33	1.1	1.6	1.1	0.6	1.1 +/- 0.4
34	1.2	1.7	2.0	1.4	1.6 +/- 0.3
6	9.7	9.4	10.1	6.8	9.0 +/- 1.3
35			0.4	0.4	0.4
36			3.9	6.2	5.1 +/- 1.2

(construct 34). In contrast, data from previous studies (31, 32) indicate that mutations at nine nearby sites (Fig. 17A, underlined) have no effect on Rex activity, suggesting that the tetrapeptide is the only essential sequence in the Rex effector domain. We also found that fusion with full-length Rev partially restored the ability of a Rex tetrapeptide mutant to transactivate through the XRE (construct 36) but that deletion of the Rev effector domain prevented this complementation (construct 35). Immunoblot analysis confirmed that these effects were not due to differences in protein stability (Fig. 17C). Thus, the tetrapeptide in Rex is essential for transactivation and is interchangeable with a region from Rev that contains this same peptide motif.

DISCUSSION

These studies extend earlier mutational analyses and provide new insights into the functional architecture of Rev and Rex. Under conditions of the in vivo assay, in which each of these proteins can transactivate only through its cognate response element, it was shown that sequences in and around the arginine-rich tract of Rev are sufficient to confer RRE specificity onto Rex. This observation is compatible with earlier evidence (29, 58, 59) that mutations throughout this region in Rev inhibit RRE binding in vitro and that similar arginine-rich tracts together with their flanking sequences dictate the target specificities of several other RNA-binding proteins (48, 62). Our data suggest, however, that the structural determinants of XRE specificity may be very different: although an arginine-rich tract is present at amino acids 1 to 15 in Rex, several RevΔ/Rex mutants which lacked this region (constructs

27 to 31) remained fully competent for transactivation through the XRE. A similar interpretation has been suggested by Hofer et al. (33), who found that the arginine-rich sequence in Rex could be replaced by that of Rev without affecting recognition of the XRE. Sequences critical for target discrimination must therefore be located in the remainder of Rex, which includes relatively few arginines or other basic residues but can nevertheless confer XRE specificity onto Rev in vivo (e.g., construct 27).

This study also reveals that amino acids 66 to 118 of Rex comprise an autonomous effector domain that is functionally interchangeable with the effector domain of Rev. The shared function of these two domains appears to depend upon a previously unrecognized tetrapeptide motif found in both Rev and Rex, as well as in other known retroviral transactivators of this type (Fig. 17A). Identification of this conserved effector motif provides a structural basis for the view (63, 64) that, while the posttranscriptional transactivators of various complex retroviruses have markedly different sequences and target specificities, they may nevertheless act through a common effector pathway. Although the details of this pathway are presently unknown, transactivation is believed to result from specific interactions of the viral protein not only with the target mRNA but also with one or more cellular components of the RNA splicing or transport apparatus (26, 31, 63, 65). The existence of a conserved and essential effector motif suggests that posttranscriptional transactivators from diverse retroviruses may exert their effects through contact with a common cellular factor.

CHAPTER VI:

POSTTRANSCRIPTIONAL REGULATION BY HIV-1 REV AND HTLV-I REX THROUGH A HETEROLOGOUS RNA BINDING SITE

Binding of Rev or Rex to their viral response element is critical for their biological effects. While much of the RRE can be deleted without eliminating either Rev protein binding in vitro or Rev responsiveness in vivo, we have found that a discrete 77-base region known as stem-loop 2, which serves as the primary Rev binding site in the RRE, is both necessary and sufficient for biological activity (35). Rev binding appears to be accompanied by conformational changes in both the protein and the RRE (66, 67), but it is not known whether these changes are important for triggering the response. It also is not known whether cellular factors must interact with the RRE or XRE, or whether these elements make other unrecognized contributions to the response. The fact that both Rev and Rex tend to localize in nucleoli suggests that this organelle might have an important role in the response, but it has been difficult to evaluate the significance of this localization conclusively because mutations that inactivate the nucleolar localization signals in Rev or Rex (13, 30, 31, 42, 68) also abolish RNA binding (29, 58, 69, 70).

To dissect the roles of these retroviral proteins and their RNA response elements, we have created a series of chimeric proteins in which Rev and Rex are fused to the coat protein of bacteriophage MS2. The MS2 coat protein has no intrinsic regulatory activity in eucaryotic cells but binds as a dimer to a

specific 21-base RNA operator (71-73). Here we demonstrate that Rev and Rex can exert their effects when tethered by this phage protein to a target RNA containing the MS2 operator. Under these conditions, nucleolar localization, the viral RNA response elements, and the specific RNA binding sequences in Rev can all be shown to be unnecessary for inducing release of unspliced RNAs to the cytoplasm.

MATERIALS AND METHODS

Plasmid construction. Response elements were synthesized as overlapping oligonucleotides with flanking *Cla* I sites and were ligated into the *Cla* I site of pDM138. The reporter used for Northern RNA blot experiments was derived from pDM138 by inserting MS2.4 and replacing the SV40 promoter with the cytomegalovirus immediate-early gene promoter (Mokarsky); this variant yielded higher levels of mRNA expression than does its SV40 counterpart but gave proportionally similar CAT responses (data not shown). The pRSV-Rev and pRSV-Rex expression plasmids were modified to create the fusion proteins by converting the last codon of *rev* or *rex* to a *Bgl* II site (encoding Asp-Leu) by oligonucleotide-directed mutagenesis (38), and a polymerase chain reaction fragment comprising codons 1 to 130 of MS2 was then inserted in frame.

RESULTS

As described in chapter four, we found that much of the 240-base RRE is not required for Rev-responsiveness, and that the 77-base region known as stem-loop 2 is sufficient for RRE activity. This finding suggests that Rev binding might be the sole requirement for RRE function. To test this

hypothesis, the Rev expression vector, pRSV-Rev, was modified to encode a chimeric protein, called Rev/MS2, in which the MS2 phage coat protein is fused to the C-terminus of Rev (Fig. 18A). We reasoned that fusion with the coat protein would enable Rev to bind indirectly to target RNAs containing the MS2 operator, and that such indirect binding might support function.

To test the chimera, a target response element was first designed based on the minimal RRE bul1. Sequences that are essential for Rev binding were replaced with the MS2 operator, producing a modified response element called bul-MS2A (Fig. 18D). Because earlier studies had shown that duplicate copies of bul1 are needed for maximal (23-fold) CAT induction by Rev, we initially inserted two tandem copies of bul-MS2A into the pDM138 reporter.

When tested in the cotransfection assay (Fig. 19B), Rev/MS2 induced CAT expression through the full-length RRE as efficiently as did Rev. Unlike Rev, however, Rev/MS2 also functioned through bul-MS2A, implying that fusion with the MS2 coat protein could extend the target range of Rev. To confirm that Rev protein sequences were necessary for this response, a series of missense mutations within the two functional domains of Rev were constructed. Each of the mutations tested had previously been shown to abolish Rev function through the RRE (28, 42). Four separate mutations involving the N-terminal domain (MA4, MA5, MB3, and MB8), each of which eliminates the RRE binding and oligomerization activities of Rev in vitro (29), eliminated Rev/MS2 function through the RRE but not through bul-MS2A. In contrast, two different mutations (M10 and DN2) that have been shown to inactivate the Rev effector domain (26, 28) each abolished function through both response elements. Immunoblots confirmed the sizes and stabilities of all six mutant proteins (Fig. 19C). Thus, transactivation

Figure 18. Reporters and RRE derivatives for analysis of the Rev response. (A) Structure of pRSV-Rev and pRSV-Rev/MS2. The Rev/MS2 expression construct was generated by in-frame insertion of the MS2-coat protein coding sequence at the 3′ end of *rev* (codon 116 of *rev* was converted to a *Bgl* II site encoding Asp-Leu). Open rectangles indicate the N-terminal RNA-binding and C-terminal effector domains of *rev*. RSV, Rous sarcoma virus 5′ LTR; HIV-1 polyA, HIV-1 polyadenylation site; Bam, *Bam* HI; Bgl, *Bgl* II; Sac, *Sac* I. (B) Structure of reporter pDM 138. SV40, Simian virus 40 early promoter; SD and SA, HIV-1 splice donor and acceptor sites, respectively; Cla, *Cla* I. (C) Sequence and possible structure of bul1, the minimal response element (described in Chapter IV) encompassing stem-loop 2 of the RRE. (D) Sequence and possible structure of the bul-MS2A variant of bul1, with the MS2 operator (open lettering) indicated. Response elements were ligated into the *Cla* I site of pDM138.

UCSF LIBRARY

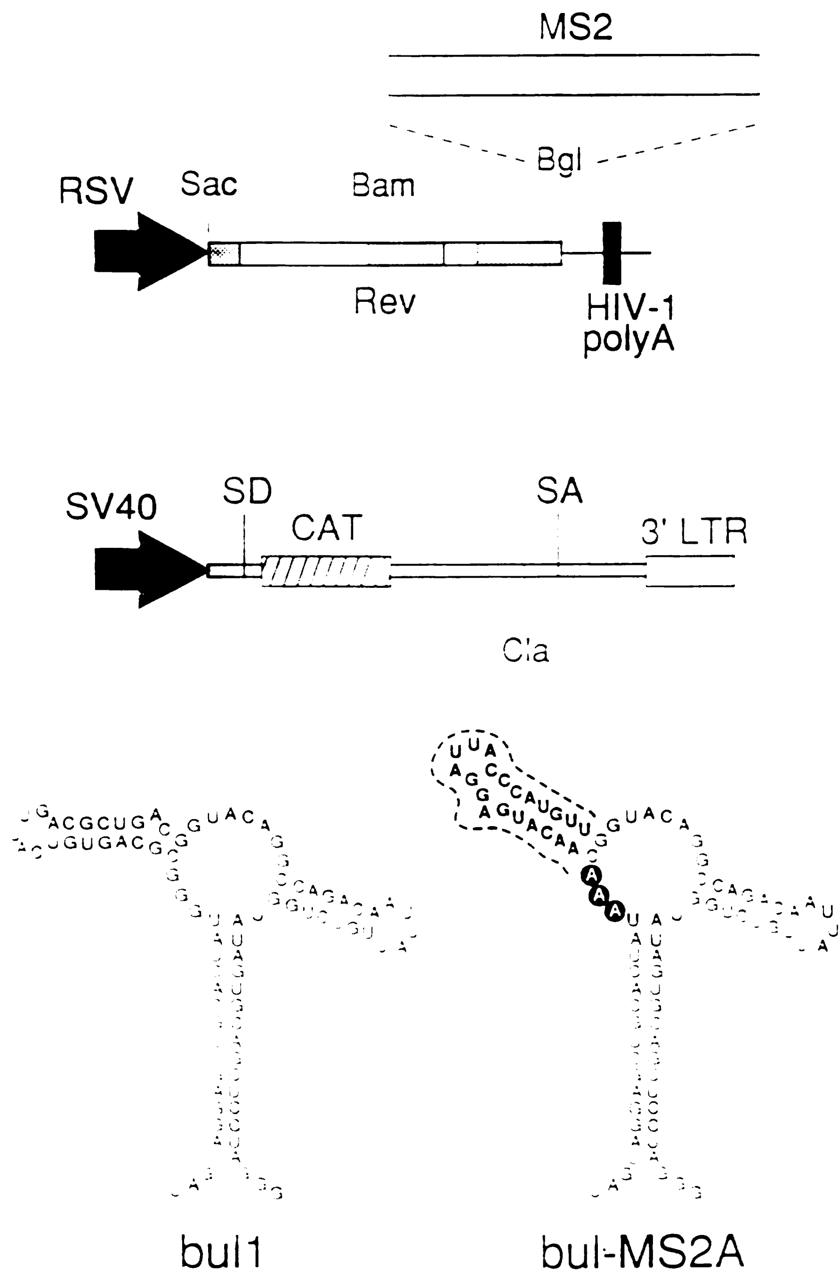
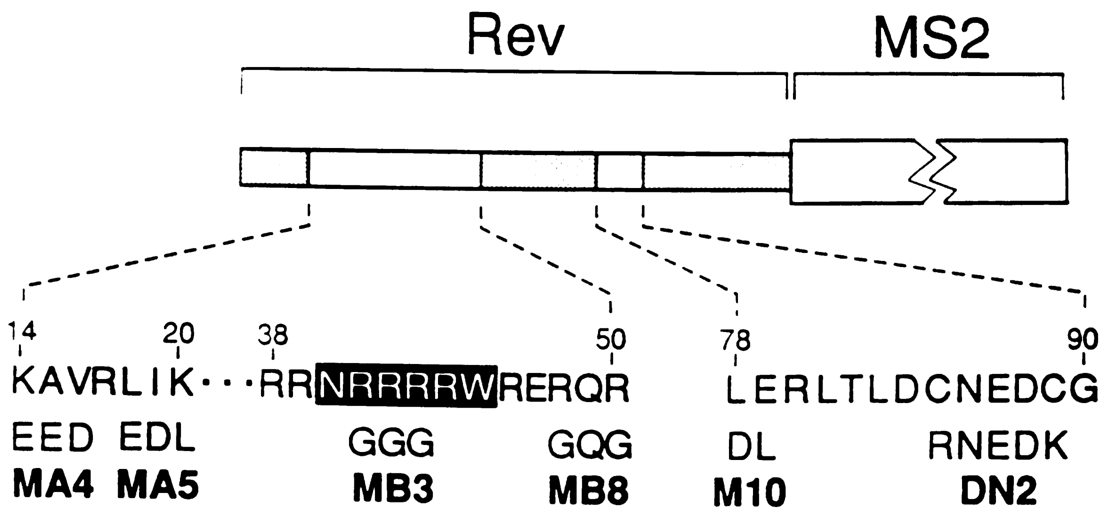


Figure 18

Figure 19. Function of the Rev/MS2 fusion protein through a modified response element. (A and B) Structures and activities of wild-type and mutant forms of Rev/MS2. (A) Schematic view of Rev/MS2. Open rectangles represent the N-terminal and effector domains of *rev*; partial sequences of each are shown, along with the six mutations tested below. Open lettering indicates a hexapeptide that is required for nuclear and nucleolar localization (42, 68). (B) CAT expression in CV1 cells after transfection with a pRSV-Rev derivative (4μg) encoding the indicated proteins, a reporter containing either the full-length RRE or two tandem copies of bul-MS2A, and pCH110. The RRE reporter was pDM128. None, reporter with no transactivator. (C) Immunoblot analysis of mutant proteins. Positions of molecular weight standards are shown at left; unfused MS2-coat protein is 13.7 kDa.

UCSF LIBRARY



RRE

bul-MS2A (2 copies)

None

Rev

Rev/MS2

Rev/MS2 - MA4

Rev/MS2 - MA5

Rev/MS2 - MB3

Rev/MS2 - MB8

Rev/MS2 - M10

Rev/MS2 - DN2

Figure 19

through bul-MS2A required the coat protein in conjunction with an intact Rev effector domain but was not inhibited by mutations that eliminate RRE binding.

Mutation MB3 was of particular interest, as it is known also to prevent nucleolar localization of Rev (42). Using in situ immunoflourescence, we found that wild-type Rev/MS2 was expressed throughout the nucleoplasm of transfected cells and was especially abundant in the nucleoli (Fig. 20A); significant amounts of the protein were also detected in the cytoplasm, suggesting that Rev/MS2 may be translocated into the nucleus somewhat less efficiently than Rev. In contrast, the mutant Rev/MS2-MB3 also was found throughout the nucleoplasm and cytoplasm but was selectively excluded from nucleoli (Fig 20B). Nevertheless, this mutant functioned at least as efficiently as Rev/MS2 through bul-MS2A (Fig. 19B). Preferential nucleolar localization is therefore not essential for Rev/MS2 activity.

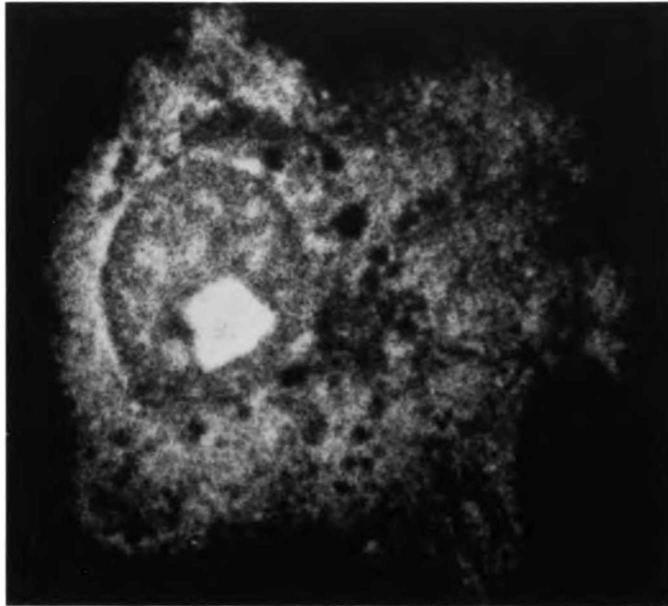
These observations were extended in tests of a similar fusion of the MS2 protein with HTLV-I Rex (Table 4). Rex contains an effector domain that is functionally interchangeable with that of Rev (34, 74), but the two proteins have significantly different RNA binding specificities; in particular, Rex cannot interact functionally with the portion of the RRE contained in bul1 (25, 69). As shown in Table 4, however, Rex/MS2 functioned almost as efficiently through two copies of bul-MS2A as it did through the XRE and that a mutation (M510) known to inactivate the Rex effector domain (34) eliminated this response. Thus, fusion with the MS2 coat protein extended the target ranges of both Rev and Rex, but only if their effector domains were intact.

To determine whether the MS2 operator alone could mediate the response

UCSF LIBRARY

Figure 20. A functional Rev derivative with defective nucleolar localization. In situ immunolocalization of Rev/MS2 (A) and Rev/MS2-MB3 (B) was performed in transfected COS7 cells, using an antiserum specific for the C-terminal half of Rev. The patterns shown typified essentially all immunoreactive cells in each population.

A



B

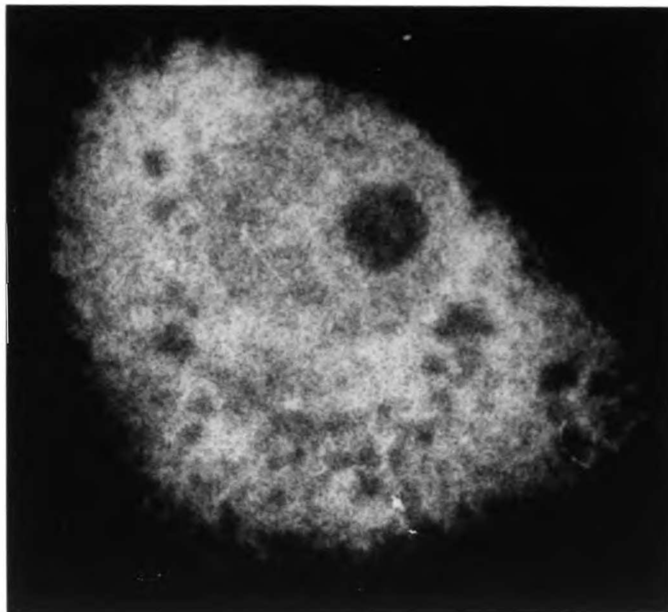


Figure 20

Table 4. Function of Rex/MS2 through the bul-MS2A response element. (a) In CV1 cells transfected with a pDM138 derivative containing either a single XRE or two copies of bul-MS2A, along with plasmids encoding the indicated transactivators. Data are means +/- standard errors of the means from triplicate transfections. Mutation M510 replaces Rex residues 90 to 93 with glycines (34). The stability of Rex/MS2-M510 was confirmed by immunoblot (data not shown).

Table 4

Response element	Transactivator	CAT activity (% acetylation)[a]	Fold induction
XRE	None (reporter alone)	0.9 +/- 0.1	
	Rex	54.3 +/- 0.4	60
	Rex/MS2	25.2 +/- 2.5	28
	Rex/MS2-M510	1.1 +/- 0.1	1
Bul-MS2A (2 copies)	None	1.3 +/- 0.1	
	Rex	1.0 +/- 0.1	1
	Rex/MS2	31.1 +/- 0.6	24
	Rex/MS2-M510	1.2 +/- 0.1	1

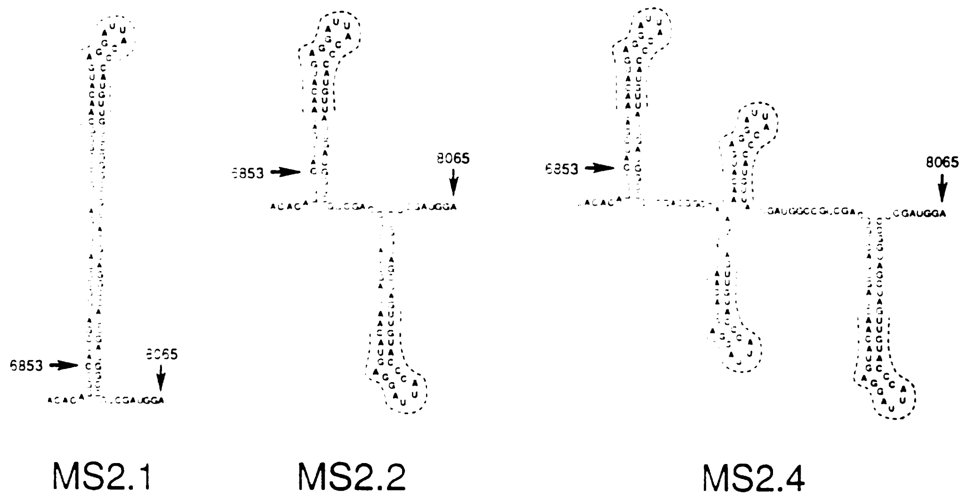
to these fusion proteins, we then designed three completely heterologous elements (Fig. 21A) that included no XRE or RRE sequences but contained one, two or four operators, respectively. Each was inserted into pDM138 and tested for responsiveness to Rex/MS2 and the Rev/MS2-MB3 mutant (Fig. 21B). Reporters containing only one operator (MS2.1 or a single copy of bul-MS2A) gave little or no response, but those containing two or more operators (MS2.2 and MS2.4) responded strongly to both fusion proteins. Indeed, the CAT responses achieved by Rex/MS2 through MS2.4 were equal in magnitude to those it produced through the XRE (Table 4). The response thus required no XRE or RRE sequences, provided that two or more MS2 operators were present.

Northern blots of cytoplasmic RNA from the transfected cells confirmed that the CAT inductions observed were due to changes in mRNA expression (Fig. 22). CV1 cells transfected with an MS2.4-containing derivative of pDM138 expressed spliced reporter transcripts constitutively in the cytoplasm. Cotransfection with a Rex/MS2 expression vector reproducibly led to the accumulation of unspliced reporter transcripts in the cytoplasm, and the M510 effector domain mutation completely eliminated this response.

An arginine-rich peptide from the Rev N-terminal domain has been reported to inhibit splicing of RRE-containing RNAs in vitro (75). We found that while much of the N-terminal domain could be deleted from Rev/MS2 without inhibiting function through MS2.4, selective deletion of the arginine-rich sequence completely inactivated the protein (Fig. 23). Activity was fully restored, however, by substitution of a different arginine-rich sequence (Scram) that lacks the ability to inhibit splicing in the in vitro assay (75). This finding implies that while basic amino acids may be required at this

Figure 21. Function of Rev/MS2-MB3 and Rex/MS2 in the absence of the viral response elements. (A) Sequences and possible structures of heterologous elements containing the MS2 operator (outlined). Each is shown as an insertion between HIV-1 nucleotides 6853 and 8065 (arrows) in the pDM138 transcript. MS2.4 is two tandem copies of MS2.2; lower energy conformations are possible. (B) Functional analysis of heterologous elements. CV1 cells were transfected with pDM138 containing the indicated response elements, either alone (None) or with a plasmid encoding Rev/MS2-MB3 or Rex/MS2. CAT expression (relative to that of the β-galactosidase control minus the activity of sham-transfected cell) in triplicate transfections was assayed at 36 h and quantified by scintillation counting. Error bars indicate standard errors of the means (SEM) exceeding +/- 0.4%; numbers above bars indicate responses expressed as fold CAT induction.

Figure 21



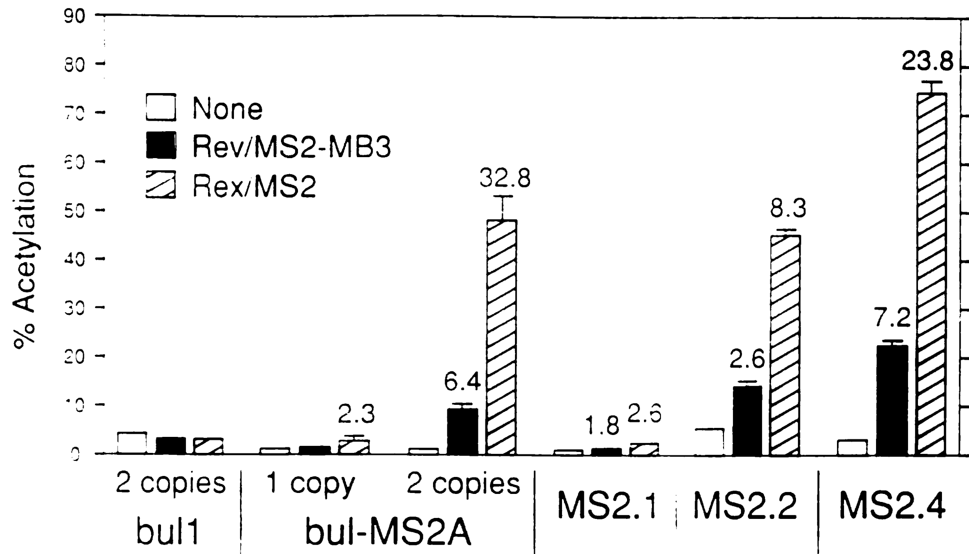


Figure 22. Northern blot analysis of reporter transcripts in cytoplasmic RNA. CV1 cells were transfected with the MS2.4-containing reporter (10 μg) either alone (None) or with a vector encoding Rex/MS2 or Rex/MS2-M510 (5 μg). Total cytoplasmic RNA was isolated 48 h after transfection, and 5-μg samples were analyzed on a 1% agarose-formaldehyde gel and probed for HIV-1 long terminal repeat sequences not present in the transactivator transcripts. Positions of the 18S and 28S rRNAs are indicated. Sham, 10 μg of pUC118 alone.

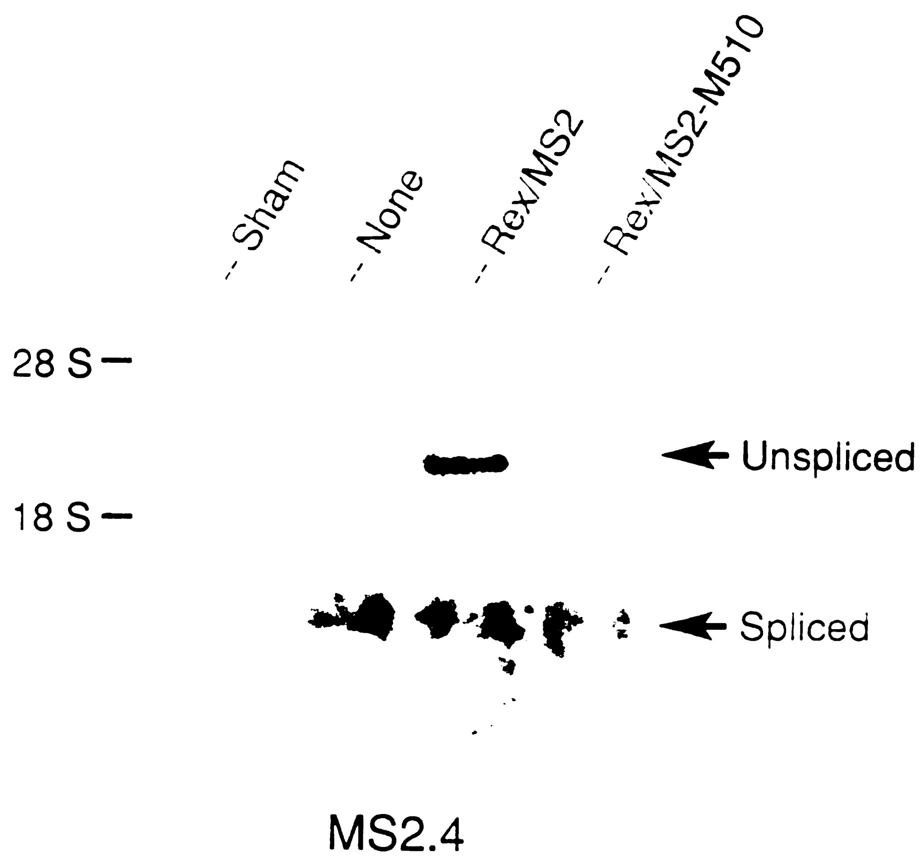


Figure 22

Figure 23. Mutagenesis of the arginine-rich domain in Rev/MS2. CAT enzyme expression in CV1 cells transfected with plasmids encoding the indicated Rev/MS2 variant together with the pDM138 derivative containing MS2.4. Δ, deletion of residues from Rev; Scram, replacement of residues 34 to 50 with the sequence shown. Properties of the Scram peptide have been described elsewhere (75). None, reporter alone. Rev/MS2-Scram does not function through the RRE (data not shown).

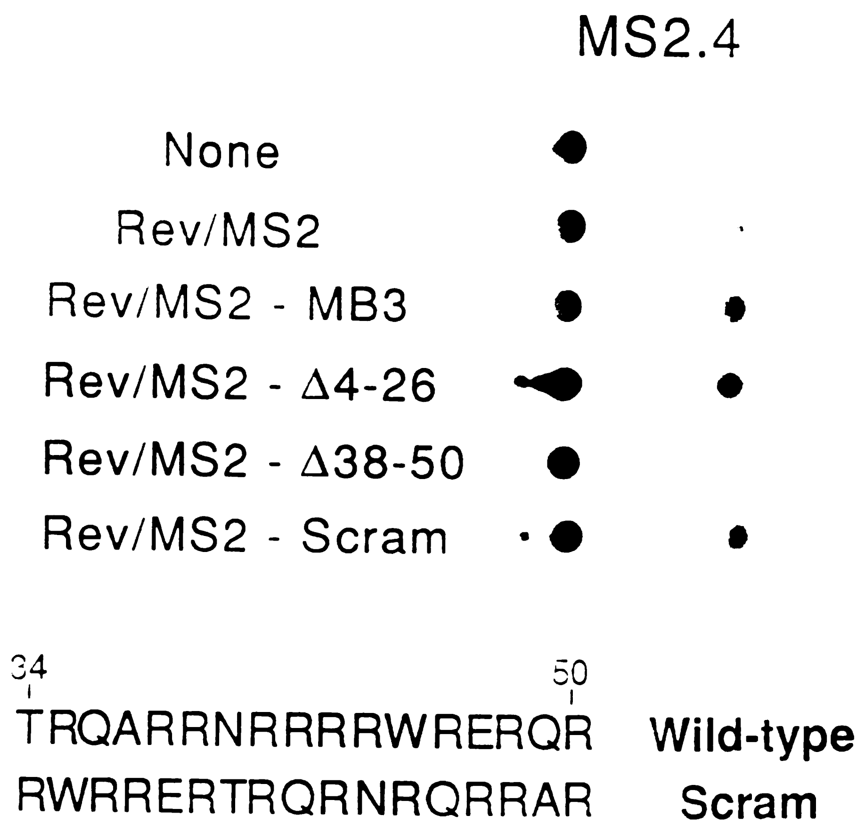


Figure 23

location in the fusion protein, the arginine-rich domain of Rev makes no sequence-specific contribution to the responses observed.

DISCUSSION

The results of this study reveal that fusion with a bacteriophage RNA-binding protein can redirect the target specificities of Rev and Rex in vivo. This finding confirms the strong inference from earlier data that RNA binding is critical for the response to these proteins (14-17, 45, 69, 76, 77). This data also sheds new light on the minimal requirements for this response. Most notably, we find that when tethered to RNA by a bacteriophage protein, Rev and Rex can each induce cytoplasmic expression of intron-containing RNAs which contain no RRE or XRE sequences. The responses achieved under these conditions (by using the heterologous target element MS2.4) are 35 to 50% as strong as those produced by unfused Rev or Rex through their native viral response elements. Thus, at least a substantial proportion of the response does not depend upon any unique structural features of the RRE or XRE or on recognition of these elements by cellular proteins, except insofar as these elements might be required for binding a particular transactivator. Although other specific properties of the viral elements (75) might be needed for a maximal quantitative response, the ability to bind Rev or Rex is the only essential requirement for RRE or XRE function in vivo.

Rev and Rex each contain sequences that function as nucleolar localization signals (30, 31, 42, 68). Early reports showed that mutations in these sequences profoundly inhibit transactivation, but the subsequent finding that these mutations also prevent RNA binding has left the significance of this localization unresolved (29, 58, 69, 70, 76, 77). This study

now demonstrates that when RNA binding is maintained by the phage protein, nucleolar localization is not required for Rev function. This finding suggests that the localization of Rev and Rex simply reflects their affinity for nucleolar constituents that are not required for the response, a view supported by recent in vitro evidence of Rev's interactions with nucleolar proteins (78).

The effector domains from Rev and Rex proved essential and interchangeable for function of the fusion constructs, as they are in the native proteins. By contrast, mutations throughout the N-terminal Rev domain did not diminish function through the phage operator, implying that sequence-specific contacts between this domain and the target RNA are not needed for the responses observed. Moreover, as the N-terminal mutations tested also inactivate the oligomerization signals in Rev, oligomerization per se also does not appear essential. This finding weighs against the possibility that oligomerization of Rev through its N terminus might be required for allosteric activation of the effector domain.

These findings strongly suggest, moreover, that multiple effector domains must be linked to each target RNA in order to trigger a response: assuming that all operator sites are accessible and competent for binding and that the fusion proteins bind as MS2 dimers (72), the data in Fig. 3B suggest a threshold requirement for three or four effector domains per transcript. This possibility is consistent with reports (15, 18, 67) that a single RRE can bind up to eight copies of Rev and that a single bound Rev is not sufficient for function (58). The data suggest the further conjecture that HTLV-I transactivation may require binding of multiple copies of Rex to the XRE.

Taken together, these studies suggest a model in which the effect of Rev or

Rex on RNA localization depends upon the linkage of a critical number of effector domains to a target RNA. For this aspect of transactivation, the remaining portions of each protein, and the viral RNA response elements themselves, serve only to facilitate such linkage and are potentially dispensable. The nucleolar localization of Rev and Rex can be viewed as reflecting their affinity for nucleolar constituents that have no obligatory role in releasing unspliced viral mRNAs from the nucleus.

CHAPTER VII:
CONCLUSIONS AND DISCUSSION

The initial mutagenesis of HIV-1 Rev, conducted in other labs, revealed two distinct regions of the protein that are critical for its function: the N-terminal RNA binding domain, and the C-terminal effector domain. The most prominent feature of the N-terminal domain was a 17 amino-acid arginine-rich region (residues 34 to 50) which was shown to be critical for RRE-specific binding (29, 70). Similar arginine-rich motifs have also been implicated in specific RNA recognition by HIV-1 Tat (79), and by the bacteriophage antiterminator N-proteins (48). All of these proteins recognize short, stem-loop RNA structures (RRE-SLIIB for Rev, TAR for Tat, and *nut* for N-proteins), with specificity for both structural determinants and for sequences in unpaired regions of the RNA (reviewed in (80)). In the case of Rev and Tat, peptides encompassing the arginine motif have been shown to bind specifically to their RNA targets *in vitro*. Tat binding to TAR RNA appears to depend little on the amino-acid sequence or intrinsic secondary structure of the binding peptide, since only a single arginine residue need be present in the peptide to retain full binding specificity (81), and since peptides which bind to TAR RNA appear to be unstructured in solution, as determined by circular dichroism spectroscopy (82). Indeed, the Rev arginine peptide binds well to TAR RNA *in vitro*, and, in fact, the Rev sequence can be substituted for the arginine motif in Tat without diminishing Tat activity *in vivo* (83). RRE binding, on the other hand, appears to depend more

stringently upon the sequence and three-dimensional structure of the arginine-rich region. Several amino-acid residues within the Rev arginine-rich motif have been shown to be critical for specific binding, both in the intact protein (29), and in isolated peptides (84, 85). Circular dichroism studies have revealed that the Rev peptide takes on a partial α-helical conformation in solution even though it is highly charged, and modifications of the N- and C- termini which increase the α-helical content also increase the binding affinity *in vitro* and *in vivo* . These findings suggest that specific recognition of the RRE is mediated by critical residues within the arginine-rich region in the context of an α-helical structure (85).

Our mutational analysis extended the N-terminal domain of Rev beyond the arginine-rich tract to include essential residues at positions 14 to 20. *In vitro* binding studies suggested that these residues are also important for RRE-specific binding (29). Consistent with the arginine peptide binding data, however, these residues are dispensable for RRE recognition in some circumstances. For instance, when Rev was fused to Rex, mutations in the Rev arginine-tract abolished RRE-specific transactivation as expected, but mutation or deletion of residues 14 to 20 had no effect on RRE binding (Chapter V, Fig. 15). This suggested that these residues might serve some function other than specific RNA contact. We have proposed that the region could be involved in protein-protein interactions. These might include homologous Rev: Rev interactions either on or off the RRE; in support of this, mutations in this region block *in vitro* multimerization of bacterially expressed Rev protein. Up to eight Rev monomers bind the full-length RRE *in vitro* under saturating conditions (18), and it has been proposed that

multiple Rev molecules bound to a single RNA is a requirement for efficient transactivation (58). Alternatively, we can't rule out involvement of the region in interactions with cellular proteins which might be involved in Rev function. Interestingly, we have observed that the amino-acid sequence in the region has a strong propensity to form an α-helical structure when analyzed in a protein secondary-structure prediction program (Chapter III). This suggests multimerization might involve interactions between hydrophobic faces of identical helices in Rev monomers, as occurs in the helix-loop-helix family of DNA binding proteins. Rev does not, however, contain classical helix-loop-helix or leucine zipper motifs.

The second, more C-terminal domain of Rev has been termed the effector domain because it is not involved in RNA binding, but is nevertheless essential for function. The finding that some mutations in this region give rise to a dominant-negative phenotype suggested that the domain might interact with an as-yet-unidentified cellular factor, presumably in the nucleus of the cell, and that functional interaction is required for Rev transactivation. Dominant-negative Rev mutants are presumed to be unable to functionally interact with the cellular factor, and could, in principle, interfere with wild-type Rev function either by competing for RRE binding or by forming inactive protein complexes. Our lab has investigated the function of such dominant-negative Rev molecules using the Rev/Glucocorticoid Receptor (Rev/GR) fusion protein. Fusion to the GR permits the control of cellular localization: in the absence of the hormone dexamethasone, Rev/GR is excluded from the nucleus, and unable to transactivate target RNAs; addition of the hormone allows rapid nuclear accumulation and subsequent transactivation (42). We found that dominant-negative mutants retained

their inhibitory activity when fused to the GR (DN Rev/GR), and the inhibition was independent of cellular localization. This suggested that inhibition was not a result of competition for RRE substrates in the nucleus, but rather might be a result of heteromultimer formation. Consistent with this, introduction of a mutation in the proposed multimerization domain of the DN Rev/GR (at amino acids 14 to 16) eliminated dominant-negative activity. The interaction between Rev and DN Rev/GR was confirmed using an immunoflourescence colocalization assay: overexpression of the DN Rev/GR caused unfused Rev to be retained in the cytoplasm in the absence of hormone and translocated to the nucleus when hormone was added. As expected, introduction of the mutation at residues 14 to 16 into the DN Rev/GR abolished the colocalization of unfused Rev. We concluded from this evidence that *trans*-dominant inhibition of Rev activity does not require competition for nuclear substrates, but may instead reflect the ability of a mutant to form inactive complexes with the wild-type Rev protein (86).

The finding that Rex can take the place of Rev in the regulation of HIV-1 gene expression (24), and that dominant-negative Rex mutants can inhibit Rev activity (29) suggested that the two proteins act through the same pathway. This hypothesis is supported by our description of a polypeptide sequence within the Rex protein that can functionally replace the Rev effector domain (Chapter V). Although the Rev and Rex effector domains share little sequence similarity, they each contain a core tetrapeptide motif that is essential for function. Using the same approach as described in Chapter V, our lab has mapped the effector domains of a number of other retroviral Rev-like proteins, and we have found that all contain a similar core motif. This is strong evidence that all of these proteins act through the same pathway, and

suggests that each forms a similar effector domain structure which could interact with the same hypothetical cellular effector-binding factor.

Interestingly, some mutations of Rex which give rise to dominant-negative molecules lie outside of the Rex effector domain, which is located between residues 80 and 100 (31, 42). A number of dominant-negative Rex proteins have been derived by mutation or deletion of amino acids 57 to 67. The mutant proteins retain their ability to bind the Rex-response element *in vitro*, but are apparently defective in their ability to oligomerize, as wild-type Rex activity can be restored by replacing the region with the putative Rev multimerization domain (87). This is in contrast to the recessive-negative phenotypes of Rev multimerization mutations. Although most multimerization-defective Rev mutants do not bind the RRE, one mutation has been described which blocks multimer formation but not RRE binding. This mutant protein was also found to have a recessive-negative phenotype (58). The ability of Rex multimerization-defective mutants to block wild-type Rex activity may reflect a greater ability of Rex monomers to bind their target RNAs *in vivo* than that of Rev monomers, and further implies the importance of oligomeric complex formation on the XRE.

The two-domain structure of Rev, and the existence of a family of discrete, modular effector domains in all known Rev-like proteins suggests a mechanism of action whereby the viral effector domains are tethered to target RNAs and presented to the cellular machinery responsible for mRNA export from the nucleus. Further, the apparent requirement for multiple Rev or Rex proteins bound to a single RNA suggests a need for several effector domains arrayed in close proximity for efficient transactivation. This model implies that both the RNA-binding domain of Rev and the RRE sequence are

dispensable for function, and that they could be replaced by some other RNA-binding protein and its target RNA. Our results, presented in Chapter VI, support this assumption. Fusion of Rev to the bacteriophage MS2-coat protein redirected the specificity of Rev to the MS2 RNA operator, and relieved the need for specific RNA-binding by the Rev moiety. The Rev/MS2 fusion protein mediated function through the MS2 operator in the absence of any RRE sequences in the target RNA, provided that two or more MS2-binding sites were present. As predicted, transactivation was dependent on an intact Rev effector domain, but missense mutations in Rev known to affect multimerization and RRE-specific binding did not diminish MS2-mediated transactivation, and, in fact, in all cases significantly increased the response. We speculate that the multimerization domain of Rev may interfere with the proper dimerization of the MS2-coat protein moiety, and that mutation of the Rev domain eliminates this interference.

The significant complexity of the RRE, despite the relatively minor sequence requirement for Rev-specific binding, has led others to speculate that the RRE might serve some other function in addition to Rev-binding (88). Although extensive mutagenesis of the RRE has revealed that only the high-affininty Rev-binding site, SLIIB, is essential for function, other regions of the RRE may contribute to the magnitude of the Rev-response, possibly by binding cellular factors (88). To this end, a nuclear factor which binds specifically to the RRE has been described (89). While our data cannot rule out the possible quantitative contribution of such a factor to the Rev-response, we have demonstrated that the RRE is not essential for transactivation when an alternative RNA-binding domain is provided.

The finding that a single MS2-operator did not confer efficient

transactivation by Rev/MS2 or Rex/MS2 supported the hypothesis that multiple effector domains must be tethered to the regulated RNA. Since the MS2-coat protein binds its operator as a dimer, we concluded that a minimum of three or four effector domains are needed. This is consistent with the finding that the minimal functional Rev-response element (described in Chapter IV) can bind up to three Rev monomers (18). We found that this minimal response-element could mediate transactivation, but that two copies were needed to confer maximal response. Interestingly, when two MS2-binding sites were separated by over 200 nucleotides on the RNA, the response to Rev/MS2 and Rex/MS2 was virtually identical to the reporter with tandemly arrayed operators, even when the Rev multimerization domain was disrupted (data not shown). This suggests that the response does not require the effector domains to be in immediate proximity to one another, which would be expected if multiple juxtaposed domains are needed to create a single factor binding site or to increase the affinity of interaction with the cellular effector-binding protein. Rather, it might indicate that multiple effector-binding proteins must be targeted to the RNA, but need not be directly adjacent to one another. One possibility is that the viral effector domains act to nucleate binding of the target RNA by a cellular low-affinity RNA binding protein, which could then multimerize on the RNA. Possible candidates for such a cellular factor are the heteronuclear RNA-binding proteins (hnrnps), which are known to bind nuclear RNAs nonspecifically and coat all mRNAs prior to export from the nucleus (90).

Although our data largely supported the tethering model for Rev activity, we found that we could not entirely eliminate the arginine-rich domain of Rev/MS2. Kjems *et al.* have reported that peptides encompassing the

arginine-domain specifically inhibit the in *vitro* splicing of RRE-containing transcripts, and have suggested that this inhibition is essential for Rev function (75). To address this issue, we substituted the Rev arginine-domain with a scrambled version of the polypeptide sequence which did not inhibit splicing in the *in vitro* assay (Chapter VI, Fig. 23). This Rev/MS2 chimera did not transactivate through the RRE, but functioned efficiently through the MS2-operator. We concluded from this evidence that the inhibition of splicing seen in the *in vitro* assay is not an essential feature of the Rev-response. However, we could not rule out some other essential contribution of the arginine residues in that region which is independent of peptide sequence. One possibility is that multiple arginines increase the binding affinity non-specifically through electrostatic interactions. Alternatively, they could mediate interactions with other nuclear proteins involved in the response.

REFERENCES

1) Barre-Sinoussi, F., Chermann, J.C., Rey, F.,, Nugeyre, M.T., Chamaret, S., Gruest, T., Dauguet, C., Axler-Blin, C., vexin-Brun, F., Rouzioux, C., Rosenbaum, W., and Montagniew, L. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk of acquired immune deficiency syndrome (AIDS). Science **220**: 868-873.

2) Popovic, M., Sarngadharan, M.G., Read, E., and Gallo, R.C. (1983) Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. Science **224**: 497-501.

3) Muesing, M.A., Smith, D.H., Cabradilla, C.D., Benton, C.V., Lasky, L.A., and Capon, D.J. (1985) Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. Nature **313**: 450-458.

4) Guatelli, J.C., Gingeras, T.R., and Richman, D.D. (1990). Alternative splice acceptor utilization during human immunodeficiency virus type 1 infection of cultured cells. J. Virol. **64**: 4093-4098.

5) Robert-Guroff, M., Popovic, M., Gartner, S., Markham, P., Gallo, R.C., and Reitz, M.S. (1990). Structure and expression of tat-, rev-, and nef-specific transcripts of human immunodeficiency virus type 1 in infected lymphocytes and macrophages. J. Virol. **64**: 3391-3398.

6) Green, M.R. (1991). Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell Biol. **7**: 559-599.

7) Maniatis, T. (1991). Mechanisms of alternative pre-mRNA splicing. Science **251**: 33-34.

8) Feinberg, M.B., Jarrett, R.F., Aldovini, A., Gallo, R.C., and Wong-Staal, F. (1986). HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. Cell **46**: 807-817.

9) Cullen, B.R., Hauber, J., Campbell, K., Sodroski, J.G., Haseltine, W.A., and Rosen, C.A. (1988). Subcellular localization of the human immunodeficiency virus trans-acting art gene product. J. Virol. **62**: 2498-2501.

10) Sodroski, J., Goh, W.C., Rosen, C., Dayton, A., Terwilliger, E., and Haseltine, W.A. (1986). A second post-transcriptional trans-activator gene required for HTLV-III replication. Nature **321**: 412-417.

11) Rosen, C. A., Terwilliger, E., Dayton, A., Sodroski, J.G., and Haseltine, W.A. (1988). Intragenic cis-acting art gene-responsive sequences of the human immunodeficiency virus. Proc. Natl. Acad. Sci. U.S.A. **85**: 2071-2075.

12) Hadzopoulou-Cladaras, M., Felber, B.K., Cladaras, C., Athanassopoulos, A., Tse, A., and Pavlakis, G.N. (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.

13) Malim, M.H., Hauber, J., Le, S.-Y., Maizel, J.V., and Cullen, B.R. (1989). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature **338**: 254-257.

14) Zapp, M.L., and Green, M.R. (1989). Sequence-specific RNA binding by the HIV-1 Rev protein. Nature **342**: 714-716.

15) Daly, T. J., Cook, K.S., Gray, G. S., Maione, T.E., and Rusche, J.R. (1989). Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element in vitro. Nature **342**: 816-819.

16) Heaphy, S., Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Karn, J., Lowe, A.D., Singh, M., and Skinner, M.A. (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.

17) Malim, M.H., Tiley, L.S., McCarn, D.F., Rusche, J.R., Hauber, J., and Cullen, B.R. (1990). HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. Cell **60**: 675-683.

18) Cook, K.S., Fisk, G.J., Hauber, J., Usman, N., Daly, T.J., and Rusche, J.R. (1991). Characterization of the HIV-1 Rev protein: binding stoichiometry and minimal RNA substrate. Nucleic Acids Res. **19**: 1577-1583.

19) Chang, D.D., and Sharp, P.A. (1989). Regulation by HIV Rev depends upon recognition of splice sites. Cell **59**: 789-795.

20) Malim, M.H., Böhnlein, S., Fenrick, R., Le, S.-Y., Maizel, J.V., and Cullen, B.R. (1989). Functional comparison of the Rev trans-activators encoded by different primate immunodeficiency virus species. Proc. Natl. Acad. Sci. USA **86**: 8222-8226.

21) Kiyomasu, T., Miyazawa, T., Furuya, T., Shibata, R., Sakai, H., Sakuragi, J.-I., Fukasawa, M., Maki, N., Hasegawa, A., Mikami, T., and Adachi, A. (1991). Identification of feline immunodeficiency virus rev gene activity. J. Virol. **65**: 4539-4542.

22) Tiley, L.S., Brown, P.H., Le, S.-Y., Maizel, J.V., Clements, J.E., and Cullen, B.R. (1990). Visna virus encodes a post-transcriptional regulator of viral structural gene expression. Proc. Natl. Acad. Sci. U.S.A. **87**: 7497-7501.

23) Hidaka, M., Inoue, J., Yoshida, M., and Seiki, M. (1988). Posttranscriptional regulator (rex) of HTLV-I initiates expression of viral structural proteins but suppresses expression of regulatory proteins. EMBO J. **7**: 519-523.

24) Rimsky, L., Hauber, J., Dukovich, M., Malim, M.H., Langlois, A.,

Cullen, B.R., and Greene, W.C. (1988). Functional replacement of the HIV-1 Rev protein by the HTLV-1 Rex protein. Nature **335**: 738-740.

25) Ahmed, Y.F., Hanly, S.M., Malim, M.H., Cullen, B.R., and Greene, W.C. (1990). Structure-function analyses of the HTLV-I Rex and HIV-1 Rev RNA response elements: insights into the mechanism of Rex and Rev action. Genes Devel. **4**: 1014-1022.

26) Malim, M.H., Böhnlein, S., Hauber, J., and Cullen, B.R. (1989). Functional dissection of the HIV-1 Rev trans-activator--derivation of a trans-dominant repressor of Rev function. Cell **58**: 205-214.

27) Perkins, A., Cochrane, A., Ruben, S., and Rosen, C. (1989). Structural and functional characterization of the human immunodeficiency virus Rev protein. J. AIDS **2**: 256-263.

28) Hope, T. J., McDonald, D., Huang, X., Low, J., and Parslow, T.G. (1990). Mutational analysis of the human immunodeficiency virus type 1 Rev transactivator: Essential residues near the amino terminus. J. Virol. **64**: 5360-5366.

29) Zapp, M. L., Hope, T.J., Parslow, T.G., and Green, M.R. (1991). Oligomerization and RNA binding domains of the HIV-1 Rev protein: A dual function for an arginine-rich binding motif. Proc. Natl. Acad. Sci. USA **88**: 7734-7738.

30) Siomi, H., Shida, H., Nam, S. H., Nosaka, T., Maki, M., and Hatanaka, M. (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.

31) Rimsky, L., Duc Dudon, M., Dixon, E.P., and Greene, W.C. (1989). Trans-Dominant inactivation of HTLV-I and HIV-1 gene expression by

mutation of the HTLV-I Rex transactivator. Nature **341**: 453-456.

32) Böhnlein, S., Pirker, F.P., Hofer, L., Zimmermann, K., Bachmayer, H., Böhnlein, E., and Hauber, J. (1991). Transdominant repressors for human T-cell leukemia virus type I Rex and human immunodeficiency virus type 1 Rev function. J. Virol. **65**: 81-88.

33) Hofer, L., Weichselbraun, I., Quick, S., Farrington, G.K., Böhnlein, E., and Hauber, J. (1991). Mutational analysis of the human T-cell leukemia virus type I trans-acting rex gene product. J. Virol. **65**: 3379-3383.

34) Hope, T.J., Bond, B.L., McDonald, D., Klein, N.P., and Parslow, T.G. (1991). Effector domains of human immunodeficiency virus type 1 Rev and human T-cell leukemia virus type I Rex are functionally interchangeable and share an essential peptide motif. J. Virol. **65**: 6001-6007.

35) Huang, X., Hope, T.J., Bond, B.L., McDonald, D., Grahl, K., and Parslow, T.G. (1991). Minimal Rev-response element for type 1 human immunodeficiency virus. J. Virol. **65**: 2131-2134.

36) McDonald, D., Hope, T.J., and Parslow, T.G. (1992). Posttranscriptional regulation by the Human Immunodeficiency Virus Type 1 Rev and Human T-Cell Leukemia Virus Type I Rex proteins through a heterologous RNA binding Site. J. Virol. **66**: 7232-7238.

37) Peterlin, B.M., Luciw, P.A., Barr, P.J. and Walker, M.D. (1986). Elevated levels of mRNA can account for the trans-activation of Human Immunodeficiency Virus. Proc. Natl. Acad. Sci. USA **83**: 9734-9738.

38) Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**: 488-492.

39) Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R., Shaw G., Wong-Staal F. and Reddy, E.P. (1985). HTLV-III env gene products

synthesized in E. coli are recognized by antibodies present in the sera of AIDS patients. Cell **41**: 979-986.

40) Meyer M.E., Gronemeyer H., Turcotte B., Bocquel M.T., Tasset, D. and Chambon P. (1989). Steroid hormone receptors compete for factors that mediate their enhancer function. Cell **57**: 433-442.

41) Pomerantz, R.J., Trono, D., Feinberg, M.B., and Baltimore, D. (1990). Cells nonproductively infected with HIV-1 exhibit an aberrant pattern of viral RNA expression: A molecular model for latency. Cell **61**: 1271-1276.

42) Hope, T.J., Huang, X.,. McDonald, D., and Parslow, T.G. (1990). Steroid-receptor fusion of the HIV-1 Rev transactivator: Mapping cryptic functions of the arginine-rich motif. Proc. Natl. Acad. Sci. U.S.A. **87**: 7787-7791.

43) Smith, D.B. and Johnson, K.S. (1989). Single-step purification of polypeptides expressie in *E. Coli* as fusions with glutathione S-transferase. Gene **67**: 31-40.

44) Finer-Moore , J. and Stroud, R.M. (1984). Amphipathic analysis and possible formation of the ion channel in acetylcholine receptor. Proc. Natl. Acad. Sci. USA **81**: 155-159.

45) Cochrane, A.W., Chen, C.-H., and Rosen, C.A. (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.

46) Emerman, R., Vazeux, R., and Peden, K. (1989). The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. Cell 57: 1155-1165.

47) Terwilliger, E., Burghoff, R., Sia, R., Sodroski, J., Haseltine, W., and Rosen, C. (1988). The art gene product of human immunodeficiency virus is required for replication. J. Virol. **62**: 655-658.

48) Lazinski, D., Grzadzielska, E., and Das, A. (1989). Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. Cell **59**: 207-218.

49) Myers, G., Berzofsky, J.A., Rabson, A.B., Smith, T.F. and Wong-Staal, F. (ed.) (1990). Human retroviruses and AIDS. Los Alamos Natl. Lab., Los Alamos, NM.

50) Benko, D.M., Schwartz, S., Pavlakis, G.N. and Felber, B.K. (1990). A novel HIV-1 protein, *tev*, sharaes sequences with *tat*, *env* and *rev* proteins. J. Virol. **64**: 2505-2518.

51) Salfeld, J., Gottlinger, H.G., Sia, R.A., Park, R.E., Sodroski, J.G. and Haseltine, W.A. (1990). A tripartite HIV-1 *tat-env-rev* fusion protein. EMBO J. **9**: 965-970.

52) Dayton, E.T., Powell, D.M., and Dayton, A.I. (1989). Functional analysis of CAR, the target sequence for the Rev protein of HIV-1. Science **246**: 1625-1628.

53) Le, S.-Y., Malim, M.H., Cullen, B.R. and Maizel, J.V. (1990). A highly conserved RNA folding region coincident with the Rev response element of primate immunodeficiency vireses. Nucleic Acids Res. **18**: 1613-1623.

54) Holland, S.M., Ahmad, N., Maitra, R.K., Wingfield, P. and Venkatesan, S. (1990). HIV-1 Rev protein recognizes a target sequence in Rev-responsive element RNA within the context of RNA secondary structure. J Virol. **64**: 5966-5975.

55) Olsen, H.S., Nelbock, P., Cochrane, A.W., and Rosen, C.A. (1990). Secondary structure is the major determinant for interaction of HIV rev protein with RNA. Science **247**: 845-848.

56) Solomin, L., Felber, B. K., and Pavlakis, G.N. (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.

57) Sakai, H., Siomi, H., Shida, H., Shibata, R., Kiyomasu, T., and Adachi, A. (1990). Functional comparison of transactivation by human retrovirus rev and rex genes. J. Virol. **64**: 5833-5839.

58) Malim, M.H., and Cullen, B.R. (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.

59) Sakai, H., Siomi, H., Shida, H., Shibata, R., Kiyomasu, T., and Adachi, A. (1990). Functional comparison of transactivation by human retrovirus rev and rex genes. J. Virol. **64**: 5833-5839.

60) Davis, J.L., and Clements, J.E. (1989). Characteriztion of a cDNA clone encoding the visna virus transactivating protein. Proc. Nat. Acad. Sci. USA **86**: 414-418.

61) Tiley, L.S., Malim, M.H., and Cullen, B.R. (1991). Conserved functional organization of the human immunodeficiency virus type 1 and visna virus Rev proteins. J. Virol. **65**: 3877-3881.

62) Weeks, K.M., Ampe, C., Shultz, C.S., Steitz, T., and Crothers, D. (1990). Fragments of HIV-1 Tat protein specifically bind TAR RNA. Science **249**: 1281-1284.

63) Cullen, B. R. (1991). Human immunodeficiency virus as a prototypic complex retrovirus. J. Virol. **65**: 1053-1056.

64) Hanly, S. M., Rimsky, L. T., Malim, M.H., Kim, J. H., Hauber, J., Duc Dodon, M., Le, S.-Y., Maizel, J.V., Cullen, B.R., and Greene, W.C. (1989). Comparative analysis of the HTLV-I Rex and HIV-1 Rev trans-regulatory proteins and their RNA response elements. Genes Dev. **3**: 1534-1544.

65) Cullen, B.R., and Greene, W.C. (1989). Regulatory pathways governing HIV-1 replication. Cell **58**: 423-426.

66) Daly, T.J., Rusche, J.R., Maione, T.E., and Frankel, A.D. (1990) Circular dichroism studies of the HIV-1 Rev protein and its specific RNA binding site. Biochemistry **29**: 9791-9795.

67) Kjems, J., Brown, M., Chang, D.D., and Sharp, P.A. (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.

68) Cochrane, A.W., Perkins, A., and Rosen, C.A. (1990). Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: Relevance of nucleolar localization to function. J. Virol. **64**: 881-885.

69) Bogerd, H.P., Huckaby, G.L., Ahmed, Y., Hanly, S.M., and Greene, W.C. (1991). The type I human T-cell leukemia virus (HTLV-I) Rex trans-activator binds directly to the HTLV-I Rex and the type 1 human immunodeficiency virus Rev RNA response elements. Proc. Natl. Acad. Sci. USA **88**: 5704-5708.

70) Olsen, H. S., Cochrane, A.W., Dillon, P. J., Nalin, C.M., and Rosen, C.A. (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.

71) Romaniuk, P.J., Lowary, P., Wu, H.-N.,Stormo, G., and Uhlenbeck, O.C. (1987). RNA binding site of R17 coat protein. Biochemistry **26**: 1563-1568.

72) Valgard, K., Liljas, L., Fridborg, K., and Unge, T. (1990). The three-dimensional structure of the bacterial virus MS2. Nature **345**: 36-41.

73) Selby, M.j., and Peterlin, B.M. (1990). *Trans*-activation by HIV-1 Tat via a heterologous RNA binding protein. Cell **62**: 769-776.

74) Weichselbraun, I., Farrington, G.K., Rusche, J.R., Böhnlein, E., and Hauber, J. (1991) Definition of the human immunodeficiency virus type 1 Rev and human T-cell leukemia virus type I Rex protein activation domains by functional exchange. J. Virol. **66**: 2583-2587.

75) Kjems, J., Frankel, A.D., and Sharp, P.A. (1991). Specific regulation of mRNA splicing in vitro by a peptide from HIV-1 Rev. Cell **67**: 169-178.

76) Ballaun, C., Farrington, G.K., Dobrovnik, M., Rusche, J., Hauber, J., and Böhnlein, E. (1991). Functional analysis of human T-cell leukemia virus type I rex-response element: Direct RNA binding of Rex protein correlates with in vivo activity. J. Virol. **65**: 4408-4413.

77) Grassmann, R., Berchtold, S., Aepinus, C., Ballaun, C., Böehnlein, E., and Fleckenstein, B. (1991). In vitro binding of human T-cell leukemia virus rex proteins to the rex-response element of viral transcripts. J. Virol. **65**: 3721-3727.

78) Fankhauser, C., Izaurralde, E., Adachi, Y., Wingfield, P., and Laemmli, U.K. (1991). Specific complex of human immunodeficiency virus type 1 Rev and nucleolar B23 proteins: Dissociation by the Rev response element. Mol. Cell. Biol. **11**: 2567-2575.

79) Subramanian T., Govindarajan R., and Chinnadurai G. (1991). Heterologous basic domain substitutions in the HIV-1 Tat protein reveal an arginine-rich motif required for transactivation. Embo J., 8: 2311-2318.

80) Mattaj, I.W. (1993). RNA recognition: a family matter? Cell **73**: 837-840.

81) Calnan B.J., Tidor B., Biancalana S., Hudson D., and Frankel A.D. (1992). Arginine-mediated RNA recognition: the arginine fork. Science **252**: 1167-1171.

82) Calnan B.J., Biancalana S., Hudson D., and Frankel A.D. (1991).

Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. Genes and Develop. **5**: 201-210.

83) Subramanian T., Kuppuswamy M., Venkatesh L., Srinivasan A., and Chinnadurai, G. (1990) Functional substitution of the basic domain of the HIV-1 trans-activator, Tat, with the basic domain of the functionally heterologous Rev. Virology **176**: 178-183.

84) Kjems J., Calnan B.J., Frankel A.D., and Sharp P.A. (1992). Specific binding of a basic peptide from HIV-1 Rev. Embo J. **11**: 1119-1129.

85) Tan R., Chen L., Buettner J.A., Hudson D., and Frankel A.D. (1993). RNA recognition by an isolated alpha helix. Cell **73**: 1031-1040.

86) Hope T.J., Klein N.P., Elder M.E., and Parslow T.G. *Trans*-dominant inhibition of human immunodeficiency virus type 1 Rev occurs through formation of inactive protein complexes. J. Virol. **66**: 1849-1855.

87) Weichselbraun I., Berger J., Dobrovnik M., Bogerd H., Grassmann R., Greene W.C., Hauber J., Bohnlein E. (1992). Dominant-negative mutants are clustered in a domain of the human T-cell leukemia virus type I Rex protein: implications for trans dominance. J. Virol. **66**: 4540-4545.

88) Dayton E.T., Konings D.A., Powell D.M., Shapiro B.A., Butini L., Maizel J.V., Dayton A.I. (1992). Extensive sequence-specific information throughout the CAR/RRE, the target sequence of the human immunodeficiency virus type 1 Rev protein. J. Virol. **66**: 1139-1151.

89) Vaishnav, Y.N., Vaishnav, M., and Wong-Staal, F. (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.

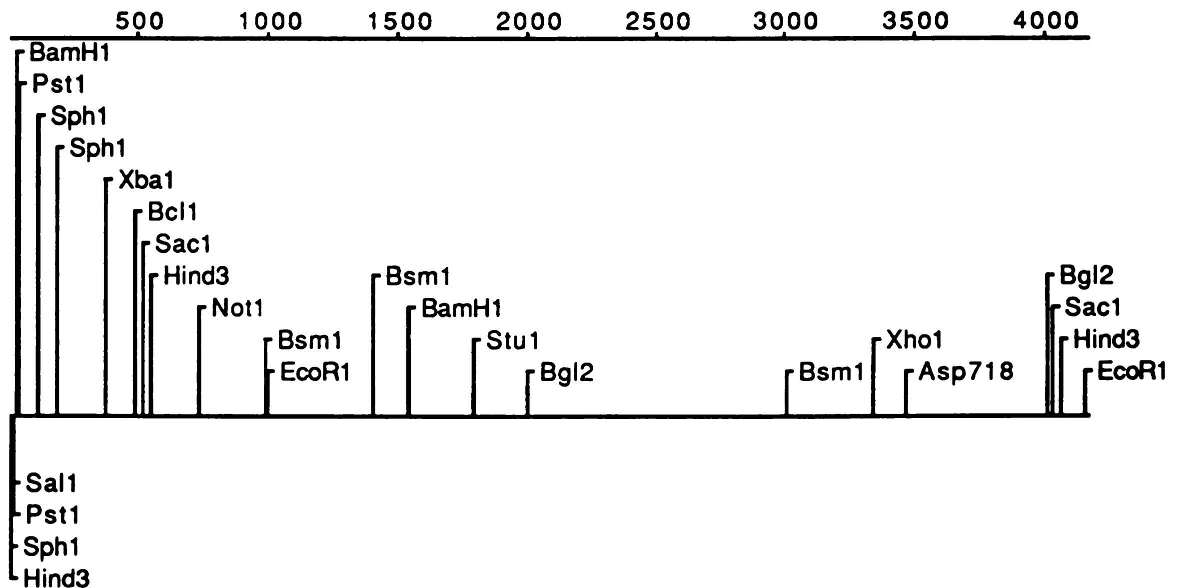
90) Pinol-Roma S., and Dreyfuss G. (1992). Shuttling of pre-mRNA

binding proteins between nucleus and cytoplasm. Nature **355**: 730-732.

91) Tiley, L.S., Malim, M.H., Tewary, H.K., Stockley, P.G., and Cullen, B.R. (1992) Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type 1 Rev protein. Proc. Natl. Acad. Sci. USA 89, 758-762.

92) Bartel, D.P., Zapp, M.L., Green, M.R., and Szostak, J.W. (1991) HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. Cell **67**: 529-536.

pDM 128
Insert in pUC9



1. SV-40 immediate early promoter, Sal1-Xba1
2. HIV-1 (ARV-2) nucleotides 5851 (Xba1)to 9730 (EcoR1)
3. rev ATG mutated to Bcl1
4. Splice donor at 561
5. env ATG mutated to Not1
6. CAT sequence into Not1-(Dra3 (blunted))
7. Splice Acceptor at 3331
8. rev deleted Ava1 (8399) to Xho1(8915)
9. EcoR1 introduced (9730) after polyA site in LTR

pDM128 Restriction Enzyme Sites

(5 cut sites or less)

Enzyme	#Cuts	Positions					NONCUTT
Aat1	1	2067					Aat2
Acc1	1	20					Aos1
Acc2	3	626	720	4040			ApaL1
Acc3	3	623	1269	4111			Asp700
Acy1	4	123	237	258	4210		AspA1
Afl2	1	4319					Asu2
Afl3	2	718	2713				Ban3
Aha1	3	138	625	2115			BssH2
Aha2	4	123	237	258	4210		BstB1
Aha3	4	1193	1532	1831	3786		BstE2
AlwN1	3	305	630	4223			Cla1
Aoc1	4	790	2546	3204	3729		Dra3
Apa1	1	630					Esp1
Aqu1	2	3615	4098				Fdi2
Ase1	1	968					Fsp1
Asp718	1	3733					Hpa1
Ava1	2	3615	4098				Mst1
Ava3	2	379	451				Nde1
Avr2	2	666	3221				Nhe1
Bal1	1	1540					Nru1
BamH1	2	295	1808				Pvu1
Ban1	5	122	236	257	1739	3733	Rsr2
Bbe1	3	126	240	261			Sac2
Bcl1	1	754					Sma1
Bgl1	1	582					SnaB1
Bgl2	2	2274	4275				Spe1
Bsm1	3	1266	1673	3281			Spl1
BspH1	1	181					Sst2
BspM1	1	6					Tth111I

pDM128 Restriction Enzyme Sites

(5 cut sites or less)

BspM2	3	623	1269	4111				Xma1	
BstU1	3	626	720	4040				Xmn1	
BstX1	2	307	624						
Bsu36I	4	790	2546	3204	3729				
Ccr1	1	3615							
Cfr1	4	139	271	1006	1538				
Cfr101	2	260	269						
Cvn1	4	790	2546	3204	3729				
Dra1	4	1193	1532	1831	3786				
Dra2	3	147	627	2555					
Eae1	4	139	271	1006	1538				
Eag1	1	1006							
Eco47III	1	178							
Eco52I	1	1006							
Eco81I	4	790	2546	3204	3729				
EcoN1	3	47	755	2854					
EcoO109I	3	147	627	2555					
EcoR1	2	1273	4423						
EcoR5	1	3917							
EcoT22I	2	379	451						
FnuD2	3	626	720	4040					
Hae2	4	126	180	240	261				
Hga1	5	31	270	3059	4046	4199			
HgiA1	4	87	788	2462	4293				
Hinc2	2	21	2224						
Hind2	2	21	2224						
Hind3	3	1	811	4333					
Kpn1	1	3737							
Mae2	3	1360	1535	2151					
Mlu1	1	718							
Mro1	3	623	1269	4111					
Mst2	4	790	2546	3204	3729				
Nae1	1	271							
Nar1	3	123	237	258					
Nci1	3	138	625	2115					

pDM128 Restriction Enzyme Sites

(5 cut sites or less)

Nco1	2	536	1574			
Not1	1	1006				
Nsi1	2	379	451			
Nsp(7524)1	5	11	112	377	449	2717
NspH1	5	11	112	377	449	2717
PaeR7I	1	3615				
Pfim1	1	1505				
PflM1	1	1505				
Ple1	4	33	797	3189	4373	
PpuM1	1	2555				
Pss1	3	150	630	2558		
Pst1	2	17	304			
Pvu2	4	305	1173	2318	4237	
Sac1	2	788	4293			
Sal1	1	19				
Sau1	4	790	2546	3204	3729	
Sca1	4	922	1690	2006	4118	
Sfi1	1	582				
Sph1	4	11	112	377	449	
Ssp1	4	862	943	1585	2783	
Sst1	2	788	4293			
Stu1	1	2067				
Taq1	4	20	617	3616	4138	
Tha1	3	626	720	4040		
Xba1	1	635				
Xho1	1	3615				
Xma3	1	1006				

pDM128 Sequence

```
                    >Sal1
                      |
Hind3                 |         >BamH1     >Pst1
|                     |            |          |
|                     |       25   |          |             50
|                     |        *   |          |              *
AAGCT TGCAT GCCTG CAGGT CGACN NGGAT CCNNC AGCTG TGGAA TGTGT

                              75                           100
                               *                             *
GTCAG TTAGG GTGTG GAAAG TCCCC AGGCT CCCCA GCAGG CAGAA GTATG

                             125                           150
                               *                             *
CAAAG CATGC ATCTC AATTA GTCAG CAACC AGGTG TGGAA AGTCC CCAGG

                             175                           200
                               *                             *
CTCCC CAGCA GGCAG AAGTA TGCAA AGCAT GCATC TCAAT TAGTC AGCAA

                             225                           250
                               *                             *
CCATA GTCCC GCCCC TAACT CCGCC CATCC CGCCC CTAAC TCCGC CCAGT

                             275                           300
                               *                             *
TCCGC CCATT CTCCG CCCCA TGGCT GACTA ATTTT TTTTA TTTAT GCAGA

                             325                           350
                               *                             *
GGCCG AGGCC GCCTC GGCCT CTGAG CTATT CCAGA AGTAG TGAGG AGGNN

                             375                           400
                               *                             *
RAGAT CCNNN CCCTG CTCTA GAGCC CTGGA AGCAT CCAGG AAGTC AGCCT

                             425                           450
                               *                             *
AGGAC TGCTT GTAAC AATTG CTATT GTAAA AAGTG TTGCT TTCAT TGCTA

                                          >Bcl1
                                            |
                             475            |              500
                               *            |                *
CGCGT GTTTC ACAAG AAAAG GCTTA GGCAT CTCCT TGATC AGGAA GAAGC
```

pDM128 Sequence

```
                   >Sac1                        >Hind3
                     |                            |
                     |    525                     |    550
                     |      *                     |      *
GGAGA CAGCG ACGAA GAGCT CCTCA GGACA GTCAG ACTCA TCAAG CTTCT

                          575                          600
                            *                            *
CTATC AAAGC AGTAA GTAGT AAATG TAATG CAATC TTTAC AAATA TTAGC

                          625                          650
                            *                            *
AATAG TATCA TTAGT AGTAG TAGCA ATAAT AGCAA TAGTT GTGTG GACCA

                          675                          700
                            *                            *
TAGTA CTCAT AGAAT ATAGG AAAAT ATTAA GACAA AGAAA ATAGA CAGAT

                                          >Not1
                                            |
                          725               |          750
                            *               |            *
TAATT GATAG AATAA GAGAA AAAGC AGAAG ACAGT GGCGG CCGCA AGGGG

                          775                          800
                            *                            *
ATCTG AGCTT GGCGA GATTT TCAGG AGCTA AGGAA GCTAA AATGG AGAAA

                          825                          850
                            *                            *
AAAAT CACTG GATAT ACCAC CGTTG ATATA TCCCA ATGGC ATCGT AAAGA

                          875                          900
                            *                            *
ACATT TTGAG GCATT TCAGT CAGTT GCTCA ATGTA CCTAT AACCA GACCG

                          925                          950
                            *                            *
TTCAG CTGGA TATTA CGGCC TTTTT AAAGA CCGTA AAGAA AAATA AGCAC

                          975                         1000
                            *                            *
AAGTT TTATC CGGCC TTTAT TCACA TTCTT GCCCG CCTGA TGAAT GCTCA

  >EcoR1
     |
     |                   1025                         1050
     |                      *                            *
TCCGG AATTC CGTAT GGCAA TGAAA GACGG TGAGC TGGTG ATATG GGATA
```

pDM128 Sequence

```
                              1075                    1100
                                 *                       *
GTGTT CACCC TTGTT ACACC GTTTT CCATG AGCAA ACTGA AACGT TTTCA

                              1125                    1150
                                 *                       *
TCGCT CTGGA GTGAA TACCA CGACG ATTTC CGGCA GTTTC TACAC ATATA

                              1175                    1200
                                 *                       *
TTCGC AAGAT GTGGC GTGTT ACGGT GAAAA CCTGG CCTAT TTCCC TAAAG

                              1225                    1250
                                 *                       *
GGTTT ATTGA GAATA TGTTT TTCGT CTCAG CCAAT CCCTG GGTGA GTTTC

                              1275                    1300
                                 *                       *
ACCAG TTTTG ATTTA AACGT GGCCA ATATG GACAA CTTCT TCGCC CCCGT

                              1325                    1350
                                 *                       *
TTTCA CCATG GGCAA ATATT ATACG CAAGG CGACA AGGTG CTGAT GCCGC

                              1375                    1400
                                 *                       *
TGGCG ATTCA GGTTC ATCAT GCCGT CTGTG ATGGC TTCCA TGTCG GCAGA

                              1425                    1450
                                 *                       *
ATGCT TAATG AATTA CAACA GTACT GCGAT GAGTG GCAGG GCGGG GCGTA

                              1475                    1500
                                 *                       *
ATTTT TTTAA GGCAG TTATT GGTGC CCTTA AACGC CTGGT GCTAC GCCTG

                                                >BamH1
                                                 |
                              1525               |    1550
                                 *               |       *
AATAA GTGAT AATAA GCGGA TGAAT GGCAG AAATT CGCCG GATCC TCTAG

                              1575                    1600
                                 *                       *
TCTGT GTTAC TTTAA ATTGC ACTGA TTTGG GGAAG GCTAC TAATA CCAAT

                              1625                    1650
                                 *                       *
AGTAG TAATT GGAAA GAAGA AATAA AAGGA GAAAT AAAAA ACTGC TCTTT
```

pDM128 Sequence

```
                                1675                           1700
                                   *                              *
CAATA TCACC ACAAG CATAA GAGAT AAGAT TCAGA AAGAA AATGC ACTTT

                                1725                           1750
                                   *                              *
TTCGT AACCT TGATG TAGTA CCAAT AGATA ATGCT AGTAC TACTA CCAAC

                                                            >Stu1
                                                                |
                                1775                           1800
                                   *                             |*
TATAC CAACT ATAGG TTGAT ACATT GTAAC AGATC AGTCA TTACA CAGGC

                                1825                           1850
                                   *                              *
CTGTC CAAAG GTATC ATTTG AGCCA ATTCC CATAC ATTAT TGTAC CCCGG

                                1875                           1900
                                   *                              *
CTGGT TTTGC GATTC TAAAG TGTAA TAATA AAACG TTCAA TGGAA AAGGA

                                1925                           1950
                                   *                              *
CCATG TACAA ATGTC AGCAC AGTAC AATGT ACACA TGGAA TTAGG CCAAT

                                1975                           2000
                                   *                              *
AGTGT CAACT CAACT GCTGT TAAAT GGCAG TCTAG CAGAA GAAGA GGTAG

   >Bgl2
      |
      |                         2025                           2050
      |                            *                              *
TAATT AGATC TGACA ATTTC ACGAA CAATG CTAAA ACCAT AATAG TACAG

                                2075                           2100
                                   *                              *
CTGAA TGAAT CTGTA GCAAT TAACT GTACA AGACC CAACA ACAAT ACAAG

                                2125                           2150
                                   *                              *
AAAAA GTATC TATAT AGGAC CAGGG AGAGC ATTTC ATACA ACAGG AAGAA

                                2175                           2200
                                   *                              *
TAATA GGAGA TATAA GAAAA GCACA TTGTA ACATT AGTAG AGCAC AATGG
```

pDM128 Sequence

```
                                2225                        2250
                                   *                           *
AATAA CACTT TAGAA CAGAT AGTTA AAAAA TTAAG AGAAC AGTTT GGGAA

                                2275                        2300
                                   *                           *
TAATA AAACA ATAGT CTTTA ATCAA TCCTC AGGAG GGGAC CCAGA AATTG

                                2325                        2350
                                   *                           *
TAATG CACAG TTTTA ATTGT AGAGG GGAAT TTTTC TACTG TAATA CAACA

                                2375                        2400
                                   *                           *
CAACT GTTTA ATAAT ACATG GAGGT TAAAT CACAC TGAAG GAACT AAAGG

                                2425                        2450
                                   *                           *
AAATG ACACA ATCAT ACTCC CATGT AGAAT AAAAC AAATT ATAAA CATGT

                                2475                        2500
                                   *                           *
GGCAG GAAGT AGGAA AAGCA ATGTA TGCCC CTCCC ATTGG AGGAC AAATT

                                2525                        2550
                                   *                           *
AGTTG TTCAT CAAAT ATTAC AGGGC TGCTA TTAAC AAGAG ATGGT GGTAC

                                2575                        2600
                                   *                           *
AAATG TAACT AATGA CACCG AGGTC TTCAG ACCTG GAGGA GGAGA TATGA

                                2625                        2650
                                   *                           *
GGGAC AATTG GAGAA GTGAA TTATA TAAAT ATAAA GTAAT AAAAA TTGAA

                                2675                        2700
                                   *                           *
CCATT AGGAA TAGCA CCCAC CAAGG CAAAG AGAAG AGTGG TGCAG AGAGA

                                2725                        2750
                                   *                           *
AAAAA GAGCA GTGGG AATAG TAGGA GCTAT GTTCC TTGGG TTCTT GGGAG

                                2775                        2800
                                   *                           *
CAGCA GGAAG CACTA TGGGC GCAGT GTCAT TGACG CTGAC GGTAC AGGCC
```

pDM128 Sequence

```
                              2825                           2850
                                 *                              *
AGACA ATTAT TGTCT GGTAT AGTGC AACAG CAGAA CAATT TGCTG AGGGC

                              2875                           2900
                                 *                              *
TATTG AGGCG CAACA ACATC TGTTG CAACT CACAG TCTGG GGCAT CAAGC

                              2925                           2950
                                 *                              *
AGCTC CAGGC AAGAG TCCTG GCTGT GGAAA GATAC CTAAG GGATC AACAG

                              2975                           3000
                                 *                              *
CTCCT AGGGA TTTGG GGTTG CTCTG GAAAA CTCAT TTGCA CCACT GCTGT

                              3025                           3050
                                 *                              *
GCCTT GGAAT GCTAG TTGGA GTAAT AAATC TCTGG AAGAC ATTTG GGATA

                              3075                           3100
                                 *                              *
ACATG ACCTG GATGC AGTGG GAAAG AGAAA TTGAC AATTA CACAA ACACA

                              3125                           3150
                                 *                              *
ATATA CACCT TACTT GAAGA ATCGC AGAAC CAACA AGAAA AGAAT GAACA

                              3175                           3200
                                 *                              *
AGAAT TATTA GAATT GGATA AGTGG GCAAG TTTGT GGAAT TGGTT TAGCA

                              3225                           3250
                                 *                              *
TAACA AACTG GCTGT GGTAT ATAAA GATAT TCATA ATGAT AGTAG GAGGC

                              3275                           3300
                                 *                              *
TTGGT AGGTT TAAGA ATAGT TTTTG CTGTG CTTTC TATAG TGAAT AGAGT

                                                         >Xho1
                                                           |
                              3325                         3350
                                 *                         |  *
TAGGC AGGGA TACTC ACCAT TGTCA TTTCA GACCC GCCTC CCAGT CCTCG

                              3375                           3400
                                 *                              *
AGACC TGGAA AAACA TGGAG CAATC ACAAG TAGCA ATACA GCAGC TACTA
```

pDM128 Sequence

```
                              3425                              3450
                                 *                                 *
ATGCT GATTG TGCCT GGCTA GAAGC ACAAG AGGAG GAAGA GGTGG GTTTT

                              3475                              3500
                                 *                                 *
CCAGT CAGAC CTCAG GTACC TTTAA GACCA ATGAC TTACA AGGCA GCTTT
                              3525                              3550
                                 *                                 *
AGATA TTAGC CACTT TTTAA AAGAA AAGGG GGGAC TGGAA GGGCT AATTT

                              3575                              3600
                                 *                                 *
GGTCC CAAAG AAGAC AAGAG ATCCT TGATC TGTGG ATCTA CCACA CACAA

                              3625                              3650
                                 *                                 *
GGCTA CTTCC CTGAT TGGCA GAATT ACACA CCAGG GCCAG GGATC AGATA

                              3675                              3700
                                 *                                 *
TCCAC TGACC TTTGG ATGGT GCTTC AAGCT AGTAC CAGTT GAGCC AGAGA

                              3725                              3750
                                 *                                 *
AGGTA GAAGA GGCCA ATGAA GGAGA GAACA ACAGC TTGTT ACACC CTATG

                              3775                              3800
                                 *                                 *
AGCCT GCATG GGATG GAGGA CGCGG AGAAA GAAGT GTTAG TGTGG AGGTT

                              3825                              3850
                                 *                                 *
TGACA GCAAA CTAGC ATTTC ATCAC ATGGC CCGAG AGCTG CATCC GGAGT

                              3875                              3900
                                 *                                 *
ACTAC AAAGA CTGCT GACAT CGAGC TTTCT ACAAG GGACT TTCCG CTGGG

                              3925                              3950
                                 *                                 *
GACTT TCCAG GGAGG CGTGG CCTGG GCGGG ACTGG GGAGT GGCGT CCCTC

                              3975                              4000
                                 *                                 *
AGATG CTGCA TATAA GCAGC TGCTT TTTGC CTGTA CTGGG TCTCT CTGGT
```

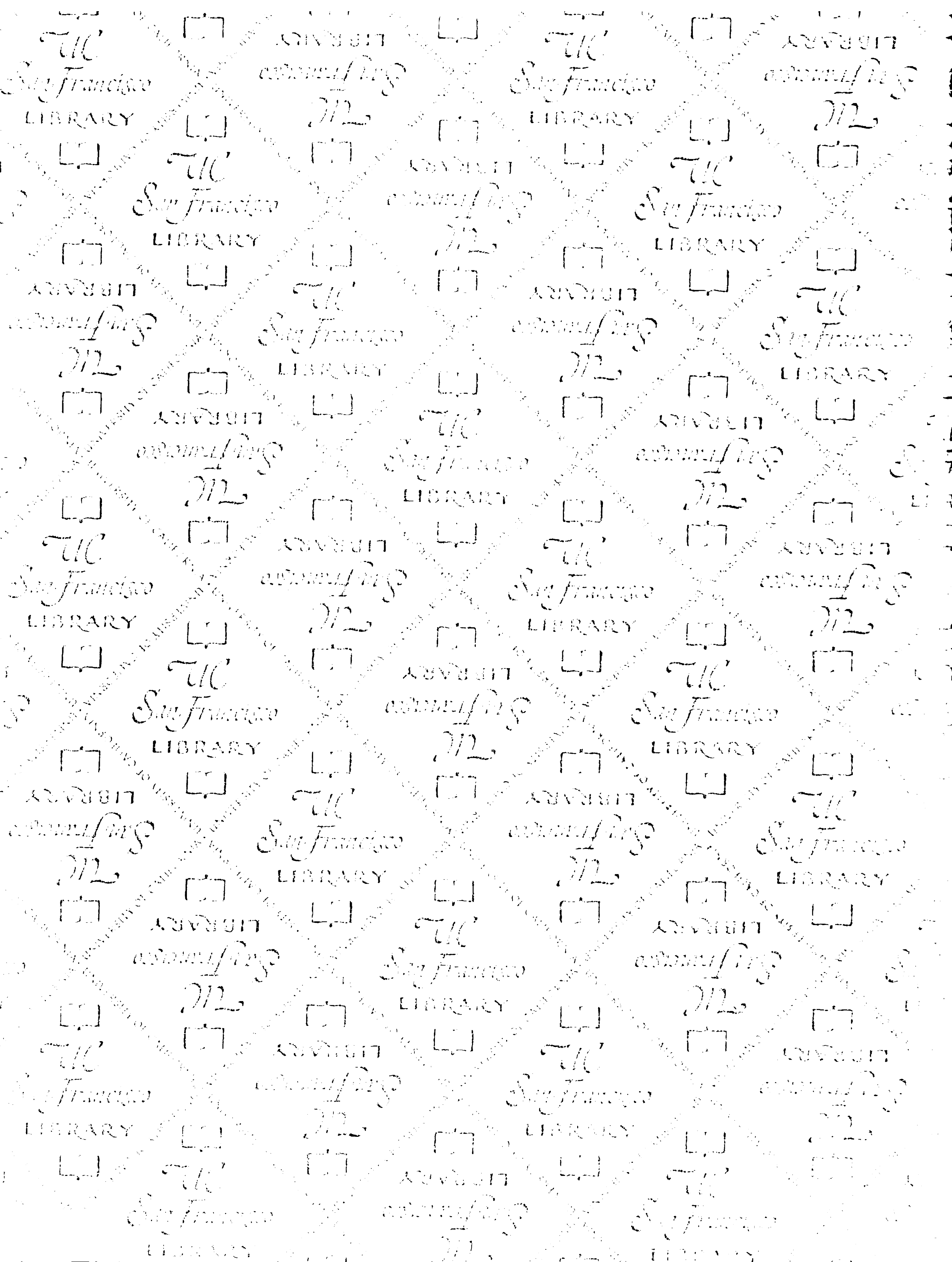
pDM128 Sequence

```
     >Bgl2                >Sac1
       |                    |
       |                4025                          4050
       |                    *                            *
TAGAC CAGAT CTGAG CCTGG GAGCT CTCTG GCTAA CTAGG GAACC CACTG

              >Hind3
                |
                |       4075                          4100
                |           *                            *
CTTAA GCCTC AATAA AGCTT GCCTT GAGTG CTTCA AGTAG TGTGT GCCCG

                        4125                          4150
                            *                            *
TCTGT TGTGT GACTC TGGTA ACTAG AGATC CCTCA GACCC TTTTA GTCAG

  >EcoR1
    |
TGTGG AATTC
```



For reference

Not to be taken from the room.

621220

3 1378 00621 2206

UC San Francisco LIBRARY