

Nucleotide Sequence of SRV-1, a Type D Simian Acquired Immune Deficiency Syndrome Retrovirus

MICHAEL D. POWER, PRESTON A. MARX, MARTIN L. BRYANT, MURRAY B. GARDNER, PHILIP J. BARR, PAUL A. LUCIW

Simian acquired immune deficiency syndrome (SAIDS) in the macaque genus of monkeys at the California Primate Research Center is apparently caused by infection by a type D retrovirus. The complete nucleotide sequence (8173 base pairs) of a molecular clone of the prototype SAIDS virus isolate, SRV-1, reveals a typical retrovirus structure with long terminal repeats (346 base pairs) and open reading frames for the *gag* (663 codons), *pol* (867 codons), and *env* (605 codons) genes. SRV-1 also has a separate open reading frame of 314 codons between the *gag* and *pol* genes that defines the viral protease gene (*prt*) and a short open reading frame of unknown significance downstream from the *env* gene. The SRV-1 protease region shows a high degree of homology to its counterpart in the hamster intracisternal A-type particle genome; both these protease genes are about twice as long as the analogous region of other retroviruses. SRV-1 has no notable similarity in either genetic organization or sequence to the human AIDS retroviruses.

SIMIAN ACQUIRED IMMUNE DEFICIENCY syndrome (SAIDS) in rhesus macaques at the California, Oregon, and Washington Primate Research Centers has been etiologically linked to a type D retrovirus related to but distinct from the Mason-Pfizer monkey virus (MPMV) (1-3). The prototype SAIDS retrovirus was called SRV-1, and strain variations at the different research centers have been defined on the basis of serology and restriction enzyme patterns (1-4). Although genetically unrelated to the retroviruses that cause the acquired immune deficiency syndrome (AIDS) in humans (5) and less restricted in cell tropism (3), SRV-1 infection leads to a depletion of both B- and T-lymphoid cells and a pathology closely resembling that of terminal AIDS (2, 6). SRV-1 is also unrelated to simian T-cell lymphotropic retrovirus type III (STLV-III), which may be causally associated with lymphoma and immunosuppression in macaques at the New England Primate Center (7). Serological evidence has excluded STLV-III as the etiological agent of SAIDS at the California Primate Research Center and has confirmed the strong association with SRV-1 (8). Furthermore, SRV-1 has now been molecularly cloned, and the cloned virus has induced SAIDS (5). To understand better the pathogenic role of these viruses and to provide information useful for development of therapeutic and preventative measures for SAIDS, we determined the complete nucleotide sequence of the SRV-1 provirus and compared its genetic organization with that of other retroviruses.

SRV-1 was molecularly cloned in bacteriophage λ (L47.1) from the unintegrated circular proviral DNA containing two long terminal repeats (LTR's) (5). Figure 1 shows the complete nucleotide sequence.

The SRV-1 provirus is 8173 base pairs (bp) long. Its overall structure resembles that of other retroviruses with sequences encoding viral proteins flanked by LTR sequences. Four long open reading frames and one short open reading frame are identified in the viral DNA (Fig. 2).

Retroviral LTR's contain sequences that function in synthesis of viral DNA, integration, and viral gene expression (9, 10). In some retroviruses, an open reading frame may extend into the rightward proviral LTR (11-17). The SRV-1 LTR is 346 bp in size. An inverted repeat of 5 bp (TGTCC) is located at each end (Fig. 1, underlined region); the dinucleotide TG in the inverted repeat is also found at the boundaries of all other retroviral LTR's (9, 18). Most other retroviruses have the dinucleotide AA upstream of the inverted repeat, but SRV-1 has CC (10, 19) (Fig. 1). Upstream of the rightward LTR, at position 7814 to 7823, is the polypurine tract (AAAAGGGTGA) that is involved in initiation of plus-strand viral DNA synthesis (9). Downstream of the leftward LTR is a 14-bp nucleotide sequence (position 351 to 364) that is complementary to the 3' stem region of tRNA_{1^{ys}} (tRNA transfer RNA) (Fig. 1, overlined region); viral minus-strand DNA synthesis is primed by a tRNA molecule (9). A tRNA_{1^{ys}} species is the primer for the human AIDS retroviruses (14-17), visna virus (19), and the mouse mammary tumor virus (MMTV) (11). A putative Goldberg-Hogness box (TATA) for promoting transcription (20, 21) is found in the sequence TATATAA (position 210) underlined in Fig. 1. Many eukaryotic promoters have a CAAT or GAAT box upstream of the TATA box (18); in the SRV-1 LTR the sequence GAAT is found at position 119 to 122. The sequence AATAAA for addition of

poly(A)⁺ tails, which occurs most commonly in eukaryotic genes and viruses (22), is not found in the SRV-1 LTR. However, a putative poly(A)⁺ signal sequence, AT-TAAA, is found 18 bp downstream of the TATA box (Fig. 1). Sequencing of viral RNA will be required to verify this and to identify the cap site. Sequencing of viral RNA will also permit delineation of the LTR into the domains designated U3, R, and U5 (9). There are no large (>10 bp) direct repeats in the SRV-1 LTR. Similarly, the enhancer core sequence [TGLLLG, where L is either A or T (23)] was not observed. Functional assessments will be necessary to delineate enhancer activity in the LTR and to examine alternative elements that may play a role in the regulation of viral transcription [such as SP-1 binding sites (24)]. Little homology was apparent when the SRV-1 LTR was compared with other retroviruses and transposons.

The first large open reading frame initiating with an ATG codon at position 503 encodes a *gag* precursor polypeptide containing 663 amino acids (Fig. 1). This ATG codon is favorable for initiation because purines are found two bases upstream and three bases downstream from the A of the start codon (25). The second amino acid, glycine, is also the second amino acid in the *gag* precursors of many other retroviruses. During viral morphogenesis, the *gag* precursor polypeptide is cleaved into several smaller polypeptides that are found in the core of the mature virus particle (26). MPMV, a closely related type D retrovirus sharing significant homology with SRV-1 (2, 3), has been shown to encode a 78-kilodalton *gag* precursor (Pr78^{gag}) that is processed into six polypeptides identified as p10, pp18 (phosphoprotein), p12, p27, p14, and p4 (27, 28). Amino acid sequencing analyses of the amino termini of five MPMV *gag* polypeptides (27) permitted localization of apparently similar proteolytic processing sites in SRV-1 (Fig. 1). The greatest degree of homology of SRV-1 with other retroviruses in the *gag* gene is in p14, the nucleic acid-binding protein (29).

The second large open reading frame has the capacity to encode a predicted polypeptide of 314 amino acids (Fig. 1). As judged from its position in the genome and its homology with other retroviruses (30, 31) and cellular proteases (31), this open reading frame encodes the viral protease gene (Fig. 3). SRV-1 is organized in this area of

M. D. Power, P. J. Barr, P. A. Luciw, Chiron Corporation, Emeryville, CA 94608.
P. A. Marx, California Primate Research Center, University of California, Davis, CA 95616.
M. L. Bryant and M. B. Gardner, Department of Medical Pathology, University of California, Davis, CA 95616.

U3->
1 CCTGTCGGAGCCGTGCTGCCCGGATGATGCTTGGCCCTGTTGCTCTAGCTCTACGCTTAAGATTCAGATGGCGAAGCTTCCCTGGTCTTCTCTGTGTTGCTTTCCCGCCGGCGCA
121 ATGTTTCCCGCTCTTAGGCTTACGTGGCTTCCAGGTTGTCAGTTGAGCATGGCCAGTACTTCCCTCCCTACTGCTGTGATATAAGACACGCATTGCCACCATTAAC
241 GAGACTGTACAGAACATGCTTGTCTCCATTTCTTGTGCTCTGTGCCCTCCAAATCCCACTCCCTCCCAAGTTCCTACTGTTGGTCCCGGGACGGGACATTGGCGCCAA
<-U5
361 CGTGGCTTGGATGACGAGGGAATTCGTGAGGAAAGACGACGGTGTGCCGCCGGATTAAAAAGAAACGAAAGTAACTTCTTCCGCCGCCGGGAGCCTGCCGCTAGGACCTG
p10->
481 ValValArgSerAspMetGlyGlnGluLeuSerGlnHisGluArgTyrValGluGlnLeuLysGlnA1aLeuLysThrArgGlyValLysValLysTyrAlaAspLeuLeuLys 38
AAAGTAAGTGGTGGCTCGGATATGGGCGAGGAAATAAGCCAGCAGCAAGCTTATGTGGAACAATAAAACAGGCTTAAAGACACGGGAGTAAAGTTAAATGCTGATCTCTTAA
601 PhePheAspPheValLysAspThrCysProTrpPheProGlnGluGlyThrIleAspIleLysArgTrpArgArgValGlyAspCysPheGlnAspTyrTyrAsnThrPheGlyProGlu 78
GTTTTTGTATTTAAAGGATACCTGCTTCCAGTTCCTCCGCAAGAGGGAACCATAGATATCAAAAGGTGGCGTAGAGTAGGCGATTGTTTCCAAGATTATTAATACTTTTGGCCCTGA
721 LysValProValThrAlaPheSerTyrTrpAsnLeuIleLysGluLeuIleAspLysLysGluValAsnProGlnValMetAlaValAlaGlnThrGluGluIleLeuLysThrSer 118
GAAAGTCCCGACTGCTTCTCACTAGGAAATTAATAAGAAATGATAGATAAGAAAGAAAGTAAACCCAGAAATGCTGCTGGCTCAAACTGAAGAAATCTTAAAACTAG
158 SerHisThrGluThrThrLysProSerGlnAsnProAspLeuAspLeuIleSerLeuAspSerAspGluGlyAlaLysGlySerSerLeuLysAspLysAsnLeuSerCysThr 158
TTCTCATACAGAGCTTACACAAAGGCTCCCAAAATCCAGACTTGGACCTTATTTCTCTGTAGTAGGACGATGAAGAGGCTAAAGGTTCCCTCCAAAGATAAAATTTATCATGTAC
pp18->
196 LysLysProLysArgPheProValLeuLeuThrAlaGlnThrSerAlaAspProGluAspProAsnProSerGlnUValAspTrpAspGlyLeuGluAspGluAlaAlaLysTyrHisAsn 198
TAAAGGCAAAAGATTCAGCTTCTTAACAGCACAACAGTAGTGGGACCTGAGGACCCCAACCCCTCAGAGGTAGACTGGGACGGATTAGAGGATGAGGACGCAAAATATCATAA
p12->
1081 ProAspTrpProPheLeuThrArgProProTyrAsnLysAlaThrProSerAlaProThrValMetAlaValValAsnProLysGluGluLeuLysGluLysIleAlaGlnLeu 238
TCCCGATTGGCTCCCTTCCAAACCGTCCACTCTCTAATAAAGCCACTCTCCGCACCCACTGTAATGGCGGTTGTTAACTCAAAAGAGGAAATTAAGAGAGAGATTCCTAAT
1201 GluGluGlnIleLysLeuGluGluLeuHisGlnAlaLeuIleSerLysLeuGlnLysLeuLysThrGlyAsnGluThrValThrSerProGluThrAlaGlyGlyPheSerArgThrPro 278
AGAGGACAGATTAATAGAGAGGTTACATCAAGACTCCTTCCAAAGTACAAAACAAAACAAAACAGGAAATGAACTAGTCCAGAACTGCAGAGGGCTTTTCTGCACACC
p27->
1321 HisTrpProGlyGlnHisIleProLysGlyLysCysCysAlaSerArgGluLysGluGlnThrProLysAspIlePheProValThrGluThrValAspGlyGlnGlyGlnAlaTrp 318
TCACTGGCCGGGCAACATATCCCTAAAGGAAATGCTGGCCAGTGGAGAAAAGGAAACAAACCCCAAAAGATATTTCCCAAGTAACTGAACTGTCGATGGCAGGGTCAAGCCTG
1441 ArgHisAsnGlyPheAspPheThrValIleLysGluLeuLysThrAlaAlaSerGlnTyrGlyAlaThrAlaProTyrThrLeuAlaIleValGluSerValAlaAspAsnTrpLeu 358
GAGGACCAATAATGTTGATTTTACCCTCAATAAAAGAAATTAAGAAAGGCTGCTCTCAATATGGGCTACTGCCCATACACATTAGCCATAGTAGAATCTGTAGCGGCAATGGCT
1561 ThrProThrAspPheAspMetLeuValArgAlaValIleLeuSerGlyGlyAspHisLeuLeuTrpLysSerGluPhePheGluAsnCysArgGluThrAlaLysArgAsnGlnAlaGly 398
TACCCTACAGATTGGAAATACACTTGTAGGGCAGTCTCTCAGGAGGAGATCATTTACTATGGAAATCTGATTTTTGAAATTTGTAGAGAACCGCTAAAGAAATCAACAGCCGG
1681 AsnGlyTrpAspPheAspMetLeuThrGlySerGlyAsnTyrSerSerThrAspAlaGlnMetGlnTyrAspProGlyLeuPheAlaGlnIleGlnAlaAlaThrAlaLysTrpArg 438
TAATGGTGGGATTTGTATGTTTACAGGCTCAGGTAACATTTCTAGCACATGATGCCACAAATGCAATATGATCCGGGATTTGTTGCTCAAAATCAAGCGGCTGCTCAAAAGCCTGGAG
1801 LysLeuProValLysGlyAspProGlyAlaSerLeuThrGlyValLysGlnGlyProAspGluProPheAlaAspPheValHisArgLeuIleThrThrAlaGlyArgIlePheGlyArg 478
AAAGTTCCTCCGTTAAGGAGGACCAAGGAGCTTCCCTTACAGGAGTCAACAGGACCCGATGAGCCATTTGACAGATTTTGTACACAGACTTATAACCACTGCTGGGAGGAAATTTTGGAA
1921 AlaGluAlaGlyValAspTyrValLysGlnLeuAlaTyrGluAsnAlaAsnProAlaCysGlnAlaAlaIleArgProTyrArgLysLysThrAspLeuThrGlyTyrIleArgLeuCys 518
TGCTGAGGCTGGTGTAGACTATGTAACCAACTAGCATATGAAACAGCTAATCCAGCTGTCAGGACAGCAATCCGCCCTATAGAAAGAAACAGATTTAAGTGGTACATCCGCTTTG
p14->
2041 SerAspIleGlyProSerTyrGlnGlnGlyLeuAlaMetAlaAlaPheSerGlyGlnThrValLysAspPheLeuAsnAsnLysAsnLysGluLysGlyGlyCysCysPheLysCys 558
TTCGGACATTGGGCTTCTTATCAACAGGGCTTAGCCATGGCCGCTTTCAGCGGACCAACTGCAAGATTTCTTAAACACAAAATAAGAAAAAGGAGGGTGTGTTTAAATG
2161 GlyArgLysGlyHisPheAlaLysAsnCysHisGluHisIleHisAsnAsnSerGluThrLysAlaProGlyLeuCysProArgCysLysArgGlyLysHisTrpAlaAsnGluCysLys 598
CGGAGGAAAGGAGCATTTTCAAAAATTTGTCATGAACATATACATAACAATTTGAAACAAAGGCTTCCGGACTCTGCTCCAGGTTAAAGAGGAGGAAACATTTGGCCCAATGAATGCA
p4->
2281 SerLysThrAspSerGlnGlyAsnProLeuProProHisGlnGlyAsnGlyLeuArgGlyGlnProGlnAlaProLysGlnAlaTyrGlyAlaValSerPheValProAlaAsnLysAsn 638
ATCCAAAATGATAGTCAAGGAAACCCACTACCACCCATCAGGAAACGGACTGAGGGGCCAGCCCAAGCCCAAGCAACAGCTTATGGGGCGTTCAGCTTTGTTCCAGCCCAACAAA 35
2401 AsnProPheGlnSerLeuProGluProProGlnGluValGlnAspTrpThrSerValProProProThrGlnTyrOC 75
CAACCCATTTCAAGCTTACCAGAGCCACCCAGGAGTGCAGGATTTGGACCTCAGTTCCACTCCACACAGATTTAACACCCGAAATGGGCCCCAAGCGTTAAGCCTGGAAATAT
2521 GlyProLeuProAsnThrPheGlyLeuIleLeuGlyArgSerSerIleThrIleLysGluLeuGlnValTyrProGlyValIleAspAsnAspTyrThrGlyGluIleLysIleMet 115
GGGCCCCTACCCTCCCAACTTTTGGATTAATCTTAGGCAAGATGAGCATTACTATAAAAGGCTCAAGATTTTCCAGGAGTAAATGATAATGACTATACTGGGAAATTAAGAAATG
2641 AlaLysAlaValAsnAsnIleValThrValProGlnGlyAsnArgIleAlaGlnLeuIleLeuLeuProLeuIleGluThrAspAsnLysValGlnGlnProTyrArgGlyGlnGlySer 155
GCAAGGCTGTCAACAAATTTGTTACTGTTCTCAAGGCAACAGGATAGCTCAATTAATCTCCCTACTCTTAATTTAGACAGACAATAAGACTACACACCCCTATAGAGGCAAGGAAAT
2761 PheGlySerSerAspIleTyrTrpValGlnProIleThrCysGlnLysProSerLeuThrLeuTrpLeuAspAspLysMetPheThrGlyLeuIleAspThrGlyAlaAspValThrIle 195
TTTGGATCTCAGACATATATTTGGTCAACCTTACTGTCAAGAGGCCCTTCAACTATGTTGATGATGATAAATGTTACAGAGTAAATCAGAGGATTAATCAGTACGGGAGCTGATGTCATC
2881 IleLysLeuGluAspTrpProProAsnTrpProIleThrAspThrLeuThrAsnLeuArgGlyIleGlyGlnSerAsnAsnProLysGlnSerSerLysTyrLeuThrTrpArgAspLys 235
ATCAAGCTAGAGGACTGGCTCCATAATGGCCATACAGATACCTTAAAGAGTATAGGCAAGGCAACACCCCTAAGCAAGTCTAATAATCTTACTTGGAGAGATARA
3001 GluAsnAsnSerGlyLeuIleLysProPheValIleProAsnLeuProValAsnLeuTrpGlyArgAspLeuLeuSerGlnMetLysIleMetMetCysSerProSerAspIleValThr 275
GAAAATAATCTGCTCAITAAACCTTTGTTATCTCAATTTACCTGTCAACCTTTGGGCGAGAGATCTCTTCTCAAAATGAAAATATGATGTAGTCTTAGTACATAGTCACT
3121 AlaGlnMetLeuAlaGlnGlyTyrSerProGlyLysGlyLeuGlyLysAsnGluAsnGlyIleLeuHisProIleProAsnGlnGlyGlnPheAspLysLysGlyPheGlyAsnPheOC 315
GCCCAATGTAGCCCAAGGCTACAGCCCGGAAAAGGATAGGAAAACGAAATGGCATCTACATCTTATCCCAATCAAGGACAAATGGACAAAAGGGATTTGGAAATTTTAA 9
3241 AlaAlaIleAspMetLeuAlaProGlnGlnCysAlaGluProIleThrTrpLysSerAspGluProValTrpValAspGlnTrpProLeuThrSerGluLysLeuAlaAlaAlaGlnGln 49
CTCGGGCATTGACATGCTTGCACCCCAAGTGTGCTGAACCCATCAGTGGAAATCAGACAGCACTGCTGGTGTGATCAGTGCCCAATTAACAGTGAAGAACTTGCCTGCTCCCAAC
3361 LeuValGlnGlnGluLeuGluAlaGlyHisIleThrGluSerAsnSerProTrpAsnThrProIlePheValIleLysLysLysSerGlyLysTrpArgLeuLeuGlnAspLeuArgAla 89
AGTTAGTGAAGACAGTATAGGAGGAGGACATATTAAGAAATTTCCCTTGGAACTCCCATATTTGTTATAAAAAGAAATCTGGTAAATGGAGGCTCTTACAAATTTAGCAG
3481 ValAsnAlaThrMetValLeuMetGlyAlaLeuGlnProGlyLeuProSerProValAlaIleProGlnGlyTyrLeuLysIleIleIleAspLeuLysAspCysPhePheSerIlePro 129
CCGTAATGCCAGTATGATTAATGGAGCTTACAACTGGATTCCTCCGCTGGCTATCCCAAGGATGATCTTAAATAATTTATGATCTCAAGATTGTTCTTTCTATCT
3601 LeuHisProSerAspGlnLysArgPheAlaPheSerLeuProSerThrAsnPheLysGluProMetGlnArgPheGlnTrpLysValLeuProGlnArgMetAlaAsnSerProThrLeu 169
CCCTTCACTTAGTCAAAAAGGTTTGCATTCAGCTTACCCTCCCAAAATTTAAGAAACCTATGCACAGCTTTTCAGTGGAAAGTTTACCAGCAAGTATGGCCACAGGCTTACCT
3721 CysGlnLysTyrValAlaThrAlaIleHisLysValArgHisAlaTrpLysGlnMetTyrIleIleHisTyrMetAspIleLeuIleAlaGlyLysAspGlyGlnGlnValLeuGln 209
TATGCCAAAATATGTTGGCCACAGCCTACATAAAGTTAGACATGCTGAAACAAATGATATATACATACATGGATGATATCCCTACAGTGGTAAAGATGGACAAACAAATTTAC
3841 CysPheAspGlnLeuLysGlnGluLeuThrIleAlaGlyLeuHisIleAlaProGluLysIleGlnLeuGlnAspProTyrThrTrpLeuGlyPheGluLeuAsnGlyProLysIleThr 249
AATGCTTTGATCAGCTCAACAAAGATTTGACTATAGCCGGTTACATATAGCCCAAGAAAATTCACATCAAGACCCCTACAGTATTTAGGATTTGAACCTAATGGTCCAAAATCA
3961 AsnGlnLysAlaValIleArgLysAspLysLeuGlnThrLeuAsnAspPheGlnLysLeuLeuGlyAspIleAsnTrpLeuArgProTyrLysLeuThrThrAlaAspLeuLysPro 289
CTAATCAAAAGCGATTATGTAAGATAAGTTGCAACCTTAAATGACTTCAAAAGCTTTAAGAGACATCAATGGCTCCGACCATCTGAACTCCTACTGCAATTTAAAC

L
T
R

g
a
g

p
r
t

p
o
l

Fig. 1. Nucleotide sequence of SRV-1 DNA. Unintegrated DNA of SRV-1 was molecularly cloned in a bacteriophage λ vector (5). Both strands of the virus were sequenced with overlapping m13 clones. DNA from viral supercoils was cloned into bacteriophage at the Bam HI site in the center of the *prt* region (5). Comparisons of the sequences of the highly homologous

SRV-2 genome in this region indicated that no viral sequences were lost at the Bam HI site during molecular cloning (40). The figure shows the 8173 nucleotides composing the SRV-1 proviral DNA. Nucleotides are numbered at the beginning of each line, and amino acids are numbered at the end of each line. The predicted amino acid sequences for the *gag*, *prt*, *pol*, and *env*

LeuPheAspThrLeuLysGlyAspSerAsnProAsnSerHisArgSerLeuSerLysGlyAlaLeuAlaLeuLeuAspLysValGluThrAlaIleAlaGluGlnPheValThrHisIle 329
 4081 CTTATTCGACACCCCTAAAGGAGGACTCTAATCCCAATTCATAGATCTTTTCAAAAGGAAGCTCTTCCCTACTTGAAGAAGTAAACAGCCATTGCAGAACAAATTTGTTACTCACA

AsnTyrSerLeuMetPheLeuIlePheAsnThrAlaLeuThrProThrGlnAsnAsnProIleMetTrpValHisLeuProAlaSerProAlaValLeu 369
 4201 TAAATATTTCATACCAATTAATGTTCTCATATTAACACAGCCCTGACGCCCATGGTATTATTTGGCAGAAATAATCCATTATGTGGTCCACCTGCCTGCATCCCCAAAAGGAT

LeuProTyrTyrAspAlaIleAlaAspLeuIleIleLeuGlyArgAspHisSerLysLysTyrPheGlyIleGluProSerValIleIleGlnProTyrSerLysSerGlnIleAspTrp 409
 4321 TACTCCCCATATAGCAGCCTATAGCAGATTTAATCATCTACAGGAGGACCATAGTAAAAAATACTTTGGAAATGAACCCCTCCGTAATCATACGCCATCTCTAAGCTCCTCAATGATT

LeuMetGlnAsnThrGluMetTrpProIleAlaCysAlaSerTyrValGlyIleLeuAspAsnHisTyrProProAsnLysLeuIleGlnPheCysLysLeuHisAlaPheIlePhePro 449
 4441 GGCTGATGCAAAACACTGAAATGTGCCAATGCTGTGCCCTTATGTTGGCATCTAGATAACCAATTAACCCACTAACCAAGCTTATCAAAATCTGCAAAATTCACATGCCCTTATTTTCC

GlnIleIleSerLysThrProLeuAsnAsnAlaLeuLeuValPheThrAspGlySerSerThrGlyMetAlaAlaTyrThrLeuAlaAspThrThrIleLysPheGlnThrAsnLeuAsn 489
 4561 CTCAAATCATTAGTAAACGCCCTTAAACAATGCTTATTAGTTTTTACTGATGGCTCTCCACTGGAATGGCCGATATACCTTGTCTGATACCTACCAATTTCAAACATAATCTTA

SerAlaGlnLeuValGluLeuGlnAlaLeuIleAlaValLeuSerAlaPheProAsnGlnProLeuAsnIleTyrThrAspSerAlaTyrLeuAlaHisSerIleProLeuLeuGluValThr 529
 4681 ATTCGGCTCAACTAGTAGAATACAAAGCTTAAATGCAAGCTTACAGCTTCCCAACCAACCTTAAACATTTACTGACAGTGTTCCTAGCCCACTCAATGCCCTGTAATGCTT

ValAlaGlnIleLysHisIleSerGluThrAlaLysLeuPheLeuGlnCysGlnGlnIleIleTyrAsnArgSerIleProPheTyrIleGlyHisValArgAlaHisSerGlyLeuPro 569
 4801 CTGTAGCACAATAAACAACATACAGAAACAGCAAGCTTCCACAGTGCACACGCTTATAACAATAGATCCATACCTTTTACATCGGACATGTCAGGGCCCAATCTGCGCTAC

GlyProIleAlaHisGlyAsnGlnLysAlaAspLeuAlaThrLysThrValAlaSerAsnIleAsnThrAsnLeuGluSerAlaGlnAsnAlaHisThrLeuHisHisLeuAsnAlaGln 609
 4921 CTGGACCTATAGCCACGBCACCAAAAGGCTGACTTGGCACTAAACCGTGGCTAGCAACATAAACCAAAACCTCGAAATCGGCTCAAAATGCTCATACCTTCAATCCCTGATGCT

ThrLeuLysLeuMetPheAsnIleProArgGluGlnAlaArgGlnIleValArgGlnCysProIleCysAlaThrTyrLeuProValProHisLeuGlyValAsnProArgGlyLeuLeu 649
 5041 AAACCTTAAACATAATGTTAACAATCCAGAGAACAGCTAGACAAATTTGTCAGACATGCCCAATATGTCACACTTACAGTCCCTCATTTAGGAGTAACTAGAGGATGTT

ProAsnMetIleTrpGlnMetAspValThrHisTyrSerGluPheGlyAsnLeuLysTyrIleHisValSerIleAspThrPheSerGlyPheLeuLeuAlaThrLeuGlnThrGlyGlu 689
 5161 TGCCGACTAATGATGAAATGGACCTTACACATTTCCGAATTTGGTAAATTAATAATATATATACATGCTTCTATAGATACCTTCAAGTGGATTCCTATAGCCACTCTAGCAGGAGAA

ThrThrLysHisValIleThrHisLeuLeuHisCysPheSerIleIleGlyLeuProLysGlnIleLysThrAspAsnGlyProGlyTyrThrSerLysAsnPheGlnGluPheCysSer 729
 5281 AAACAACAACAAATGTCATAACCAATTTACTTCAATGCTTCTTATTTGGACTCCCTAAACAATAAACAACAGATAACGGTCTGGATACCTCCAAAAATTTCAAGAAATCTGCT

ThrLeuGlnIleLysHisValThrGlyIleProTyrAsnProGlnGlyGlnGlyIleValGluArgAlaHisLeuSerLeuLysThrIleGluLysIleLysLysGlyGluTrpTyr 769
 5401 CCACCTCAAAATAAACAATGTTACTGGAATCCCTATAATCCCAAGGCCAAGGAAATGTTGAAAGAGCCCACTTATCTCTTAAACCCCACTTGAATAAATAAAGGGGGAATGTT

ProThrLysGlnIleProArgAsnIleLeuAsnHisAlaLeuPheIleLeuAsnPheLeuAsnLeuAspGlnAsnHisSerAlaAlaAspArgPheTrpHisSerAsnProArgLys 809
 5521 ACCCTACGAGGGTACCCCAAGCAATTCACATCATGCACCTTTATTCATAATTTTTAAATTTGGATGATCAAAACCCTCAGCAGCTGATCGTTTTGGCATAGCACCCCAAGAA

GlnPheAlaMetValLysTrpLysAspProLeuAspAsnThrTrpProTrpProAspValIleIleTrpGlyArgGlySerValCysValTyrSerGlnThrHisAspAlaAlaArg 849
 5641 AACAAATTTGCATGGTAAATGGAAAGATCCACTAGACAACTAGTGGCCATGGCTGATCCAGTGATATTTGGGGCAGAGGTCAGCTGTGTCTTACTCTCAAACCAATGATGCCCTA

TrpLeuProGluArgLeuValLysGlnIleProAsnAsnGlnSerArgGluOP (gp70->)
 5761 GATGGCTACCAGAACGACTAGTAAACAATAACCAATTAACCAATCCAGGGAGTGATCTCTCCCTGAGATGGCTTTCCCTTGTCTACAGAGATGCACTTCAATCATCTTCTAC

TrpSerLeuValIleIleSerGlnIlePheGlnValGlnAlaGlyPheGlyAspProGlnGluAlaLeuLeuGluIleGlnGlnLysHisGlyLysProCysAspCysAlaGlyGlyTyr 66
 5881 CTGGCAACTAGTGGTAAATCTCAAAATTTCCAAATTTCCAAAGTTCAGCCGGTTTTGGAGATCCGCGGAGGCCCTCCTAGAGATACAACAAAAACATGGTAAGCCTTGTGACTGTGAGGAGTA

ValSerSerProProThrAsnSerLeuThrThrValSerCysSerThrTyrThrAlaTyrSerValThrAsnSerLeuLysTrpGlnCysValSerThrProThrAlaSerProThr 106
 6001 TGTTCACAGCCCACTCAATTTCCCTTACACCTGCTCATGCTCTACTTATACTGCTTATTCAGTAACCAACCTCCCTAAAGTGGCAGTGTGTCTACTCCCACTACAGCCAGCCCACT

HisIleGlySerCysProSerGlnCysAsnSerGlnSerTyrAspSerValHisAlaThrCysTyrAsnHisTyrGlnGlnCysThrIleGlyAsnLysThrTyrLeuAlaThrMet 146
 6121 ACATATAGGATCTGTGCTCCAGTCAATGTAACCTCAACAATCATATGACTCTGTACATCCAGCTGCTATACCACTCACTCAACAATGTAATTTGGTAATAAGACATATCTCAGTCACT

IleArgAspLysSerProSerSerGlyAspGlyAsnValProThrIleLeuGlyAsnAsnGlnAsnLeuIleIleAlaGlyCysProGluAsnLysLysGlyGlnValValCysTrpAsn 196
 6241 GATTAGAGACAAATCCCTCCAGTGGTACGGGAACCTCCCAACAATTAGGGAAATACTCAAAACCTCATTAAGCAGGCTGCTCCGAAAATAAAGGGGCCAAGTGGTTGCTGGAA

SerGlnProSerValHisMetSerAspGlyGlyGlyProGlnAspLysValArgGluIleIleValAsnLysLysPheGluGluLeuHisLysSerLeuPheProLeuLeuSerTyrHis 226
 6361 TAGCCAACTCTGTGCTCAGTGTCTGATGGAGGAGGGCCCTCAAGATAAGGTCCTGAGGATATAGTAAATAAAGGTTTGAAGAAATGCAATAATCGCTTCCCAAGCCAAATAGCCGAAGATG

ProLeuAlaLeuProGluAlaArgGlyLysGluLysIleAspAlaHisThrPheAspLeuLeuAlaThrValHisSerLeuLeuAsnValSerSerGlnArgGlnLeuAlaGluAspCys 266
 6481 CCCCTCGGCTTTGCCCGAAGCCCGTGGTAAAGAAAAAATGATGCAACACTTTTGTATCTCCTGTCACCTGTCGATAGTCTTACTCAATGTTTCCCAAGCCAAATAGCCGAAGATG

TrpLeuCysLeuArgSerGlyAspProValProLeuAlaLeuProTyrAspAsnThrSerCysSerAsnSerThrPhePheAsnCysSerAsnCysSerCysLeuIleThrProPro 306
 6601 CTGGCTGTGCTTGGCCAGGTGATCCGCTTCTCTCGCCCTGCTTATGATAACACTGCTGCTCAACTCAACCTTTTCTTAAATGCTCAATTTGCTTGGCTTATCACCCTCC

LeuLeuValGlnProPheAsnPheThrHisSerValCysLeuTyrAlaAspTyrGlnAsnAsnSerPheAspIleAspValGlyLeuAlaGlyPheThrAsnCysSerSerTyrIleAsn 346
 6721 TTTCTTAGTACAGCCCTTAACTTCACTCATCTGTTGCTTTACGCTGATATCAAAACACTCATTTGACATAGATGTAGGCTAGCTGGCTTCACTAATGCTAGTATATATAA

IleSerLysProSerSerProLeuCysAlaProAsnSerSerValPheValCysGlyAsnAsnLysAlaTyrThrTyrLeuProThrAsnTrpThrGlySerCysValLeuAlaThrLeu 386
 6841 TATTTCAAACCTCCAGTCCCTTATGCCCCAAATAGCTAGCTTTTGTATGGTAAATAACAAGGCACTACATTACTACCCAAATGGACGGGAAGCTGTGTACTTGCCTACTCT

LeuProAspIleAspIleIleProGlySerGluProValProIleProAlaIleAspHisPheLeuGlyArgProLysArgAlaIleGlnPheIleProLeuValIleGlyLeuGlyIle 426
 6961 TTTACCCGATATAGACATTTCCAGGTAGTGAACCTGCTCCATTCAGCTATAGATCATTTTGGTGGTACCCAAAGGCAATCCAGTATTTCCCTAGTCAATAGGATAGGAT

ThrThrAlaValSerThrGlyThrAlaGlyLeuGlyValSerLeuThrGlnTyrThrLysLeuSerHisGlnLeuIleSerAspValGlnAlaIleSerSerThrIleGlnAspLeuGln 466
 7081 AACACTGCGATCTACCGGGACTGCTGGCTGGGGGTTCCCTCACTCAATACAAAAATGCTCACCACATAATCAAGATGACAAAGCTTATCTAGTATATACAGATCACTCA

AspGlnValAspSerLeuAlaGluValValLeuGlnAsnArgArgGlyLeuAspLeuThrAlaGluGlnGlyGlyIleCysLeuAlaLeuGlnGluLysCysPheTyrAlaAsn 506
 7201 AGATCAAGTAGACTCTCAGCAGAAAGTAGTACTACAAAAAGAAAGAGGATTAAGTCTGCTGACAGCAGAGCAGGAGGCACTGCTTAGCTTACAGGAAAAAATGCTGTTTACGCCAA

LysSerGlyIleValArgAspLysIleLysAsnLeuGlnAspAspLeuGluLysArgArgLysGlnLeuIleAspAsnProPheTrpThrGlyPheHisGlyLeuLeuProTyrValMet 546
 7321 CAAATCTGGAACTGCTAGAGACAAAGATTAATAAACCACAAAGATGACTTAGAAAAACGCCGAAACAACTGATCGACAAACCCCTTTGGACTGGCTTTCATGGACTCCCTCTATGTTAT

ProLeuLeuGlyProLeuLeuCysLeuLeuValLeuSerPheGlyProIleIlePheAsnLysLeuMetThrPheIleLysHisGlnIleGluSerIleGlnAlaLysProIleGln 586
 7441 GCTCTATTAAGCCCTTACTTTGCTTACTGCTTGTGTTACTTTCCGACCAATTACTTCAATAAGCTTATGACTTTTATTAACAACAAATCGAGAGCAATCAAGCCAAACCTATACA

ValHisTyrHisArgLeuGluGlnGluAspHisGlyGlySerTyrLeuAsnLeuThrAM
 7561 GGTCATATATCATGCTGCTGAACAAGAGACCATGGTGGCTCATATTTAAACTTAAACAAGACACCTCCCTCGGAGCTAAGCTGGACAGCCAAATGACGGGTAAAGAGGTGACATTTT

CACTAACCTAAGACAGGAGGGCCGTCAGAGCTACTGCCAAATCCAAAGACGGGTAAAGTGAATAAAATGATACCTCAACCTAAGACAGGCGCAGCTTCCGAGGGATTTGCTGCTGT
 7681

TTTATATATTTAAAGGGTGAACCTGCTCCGGAGCCGTGCTGCCGGATGATGCTTGGCCCTGTTTGGCTAGCTCTACGCTTAAAGATCAAGATGCGCAACTCTCGTGTCTCTCT
 7801

GTGTTGCTTTCCCGCCGCGCCGAATGTTTCCCGCTCTTAGGCTTACGTGGCTTCCCGAGTCTGCAAGTGAAGTCAAGTACCTTCCCTCCCACTTACTGCTGTGTATATAAG
 7921

ACACGCTATTGCAACATTAACGAGACTGATCAGAACACTGCTTGTCTCCTTCTGCTGCTTCTGCTCCATTCACATTCACCTCCCTCCCTCCAGGTTTCCATGCTGTGGTCCCGC
 8041

<-U5
 8161 GGGACGGGACATT

e
n
v

L
T
R

genes are presented. A putative TATA box at position 210 and a putative polyadenylation signal at position 234 are underlined. A sequence complementary to the 3' end of tRNA^{Lys} is overlined at position 351. The inverted repeats (5 bp) are also underlined at the ends of each LTR. The amino termini of several mature gag proteins, determined from amino acid sequencing data (27), and the putative start for the env glycoprotein are labeled; a

proposed processing site in env is also shown. We used M13 phage vectors (41) to clone DNA fragments produced by restriction enzyme treatment of a recombinant bacteriophage containing SRV-1 DNA (5). These M13 clones were templates for DNA sequencing by the dideoxy chain termination method with the use of oligonucleotide primers that were chemically synthesized on an Applied Biosystems 380A machine (17).

the genome in a fashion similar to human T-cell lymphotropic virus type II (HTLV-II) and bovine leukemia virus (12, 32, 33) (Fig. 4). Each of these retroviruses has a protease translation frame separate from the coding frames for *gag* and *pol*, and in each of these viruses the protease open reading frame is minus 1 bp with respect to the *gag* open

reading frame. We propose that the protease gene be designated *prt* in those retroviruses that have a separate open reading frame encoding this enzyme.

It is interesting that the recently described hamster intracisternal A-type particle (IAP-H18) sequence (34) displays more homology to the SRV-1 *prt* gene (both in amino

acid and nucleotide sequences) than to any other known retrovirus or eukaryotic transposon. Both the SRV-1 and IAP-H18 *prt* genes are about twice as long as the protease regions of all other retroviruses including HTLV-II and bovine leukemia virus. The SRV-1 and IAP-H18 *prt* genes can be divided into three regions, as judged from com-

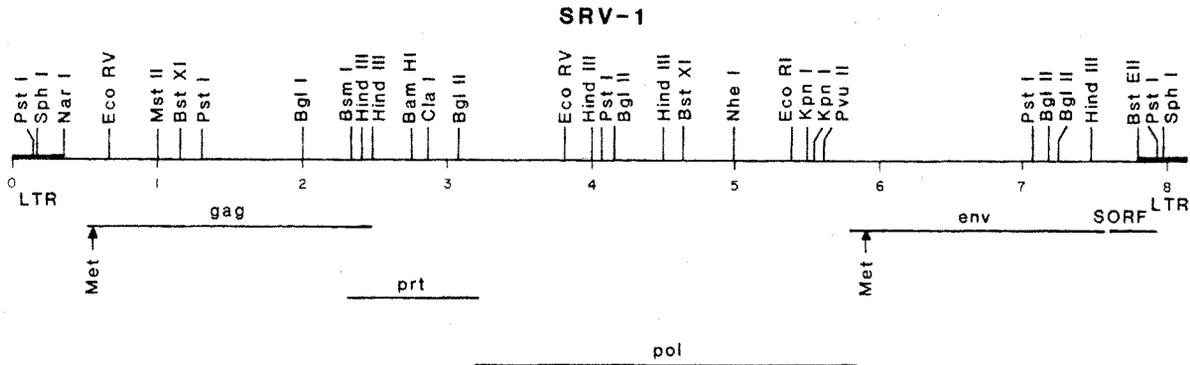


Fig. 2. Restriction endonuclease map of SRV-1. The locations of the major genetic regions of the SRV-1 genome—LTR's, *gag*, *prt*, *pol*, *env*, and the short open reading frame—are indicated. Viral genes in the same reading frame are aligned. Putative methionine initiation codons are labeled.

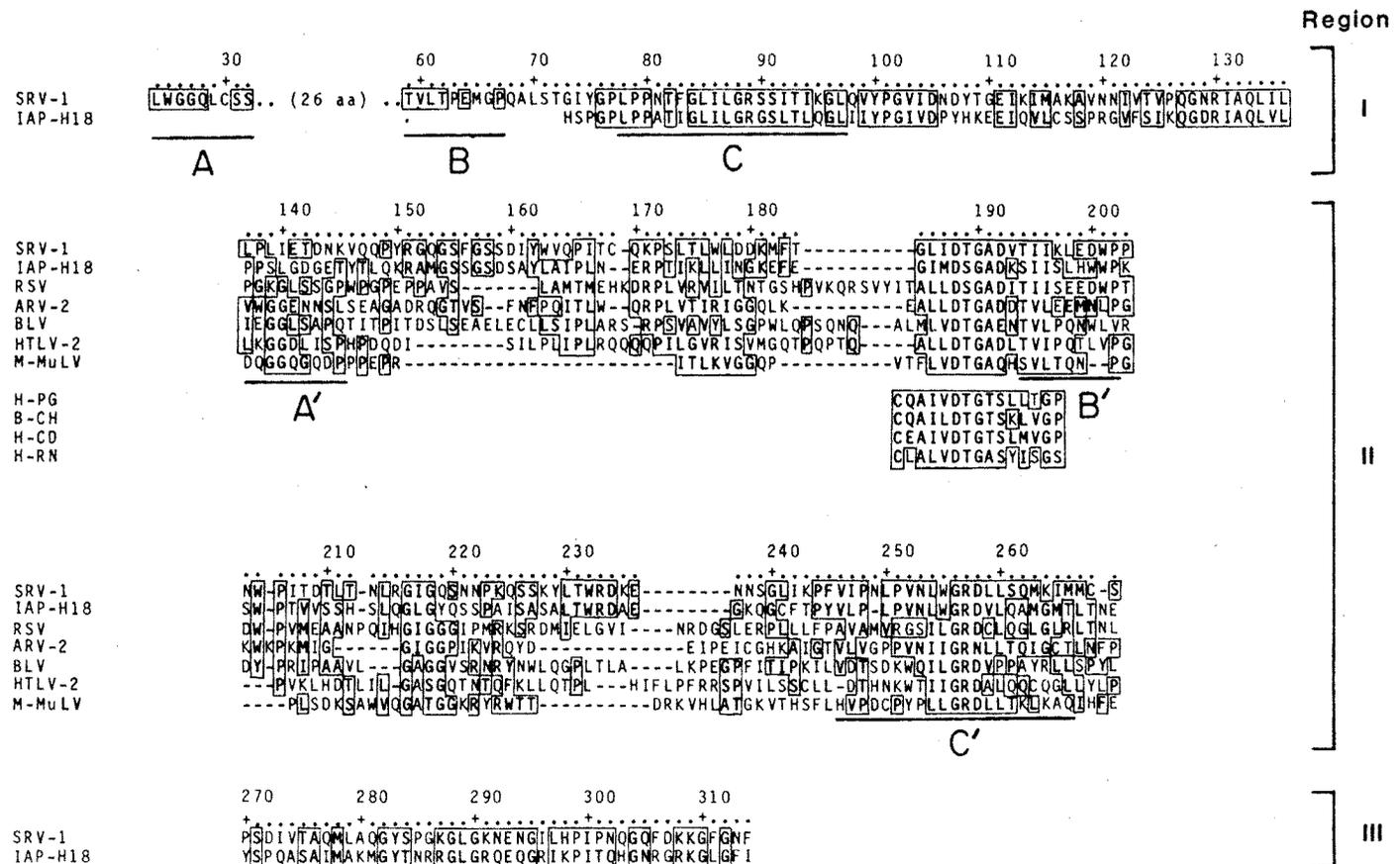


Fig. 3. Genetic homologies of several retroviral proteases and cellular aspartyl proteases. Comparisons are shown of the predicted amino acid sequences in the protease regions of SRV-1, hamster IAP-H18, RSV, ARV-2, BLV, HTLV-II, and M-MuLV. Spot homologies are shown for sequences in the carboxyl terminal portion of several cellular aspartyl proteases. Regions I and III are additional sequences found only in SRV-1 and IAP-H18. A, B, and C are short stretches that have homology to A', B', and C'. Numbering is with respect to the SRV-1 amino acid numbering scheme in

Fig. 1. Chemically similar amino acids found by inspection and with the MALIGN computer program are boxed. Gaps (-) were inserted to increase sequence similarities. Chemically similar amino acids are abbreviated and defined as follows (34): A, S, T, P, G; N, D, E, Q; H, K, R; M, L, I, V; and F, W, Y. The cellular proteases are human pepsinogen (H-PG), bovine chymosin (B-CH), human cathepsin D (H-CD), and human renin (H-RN) (31).

parisons to each other and to cellular proteases. Figure 3 shows that region I of the SRV-1 *prt* gene has three short stretches (A, B, C) that show homology to stretches (A', B', C') in region II. Comparisons with cellular proteases (Fig. 3) suggest that retroviral proteases are related to aspartyl proteases (31). These cellular proteases have two widely separated homologous stretches that contain aspartyl residues. The SRV-1 *prt* gene appears to be similarly organized because it has two widely separated homology stretches that contain aspartyl residues. Region I of SRV-1 protease is less homologous to the cellular proteases than region II (Fig. 3). Region III showed no homology to other proteases.

The *pol* gene of SRV-1 is encoded by the third large open reading frame (Fig. 1). Comparison with other retroviruses, including hepatitis B virus (35) and eukaryotic transposons [the yeast Ty element (31) and the *Drosophila* 17.6 element (31)], revealed two domains (reverse transcriptase and endonuclease) in the SRV-1 *pol* gene. Also, the hamster IAP-H18 genome shows some homology with the *pol* regions of SRV-1 and squirrel monkey virus, another type D virus (34, 36). A putative processing site for delineating the carboxyl terminus of the reverse transcriptase domain and the amino terminus of the endonuclease domain is proposed near *pol* amino acid 593 on the basis of its homology with the known processing site for Rous sarcoma virus (37).

A large open reading frame with a coding capacity of 605 amino acids is proposed to represent the SRV-1 *env* gene (Fig. 1). The ATG codon at position 5858 is a likely initiator for the *env* polypeptide precursor. This methionine codon is favorable for initiation because purines are found at minus 3 and plus 4 base pairs relative to the A of the ATG (25). The predicted amino acid sequence immediately following this codon has a hydrophobic stretch of ten amino acids; this feature is characteristic of signal peptides at the amino termini of many viral and cellular membrane proteins. An alternative possibility is that messenger RNA (mRNA) splicing may provide an ATG codon, perhaps from the *gag* gene; in Rous sarcoma virus, splicing produces an mRNA that positions the first six *gag* gene codons next to the body of the *env* gene (38). Direct protein sequence determination of the SRV-1 viral *env* precursor as well as analysis of mRNA splicing patterns will be required to resolve these issues. Homology comparisons of the predicted amino acid sequence of the SRV-1 envelope with several other retroviruses suggests the presence of two domains. Proteolytic processing may occur at the dibasic amino acid residues in the *env* se-

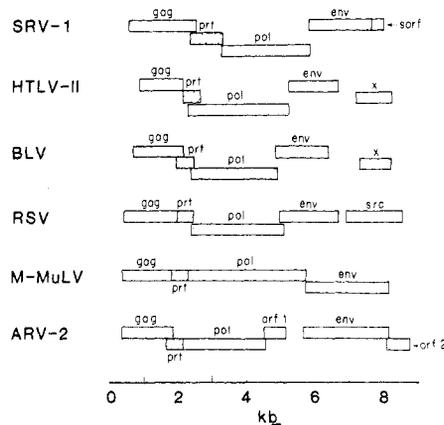


Fig. 4. Comparison of the genomic organization of SRV-1 with HTLV-II (12), bovine leukemia virus (BLV) (32, 33), Rous sarcoma virus (RSV) (38), Moloney murine leukemia virus (M-MuLV) (42), and AIDS-associated retrovirus type 2 (ARV-2) (14-17). The protease region (*prt*), shown by cross-hatching, lies in the carboxyl terminus of *gag* in RSV, in the amino terminus of *pol* in ARV-2, in the same reading frame as both *gag* and *pol* in M-MuLV, and as a separate reading frame in SRV-1, HTLV-II, and BLV.

quence Arg-Pro-Lys-Arg (position 410) to produce the relatively hydrophilic amino terminal domain. This domain has 12 potential glycosylation sites. The carboxyl terminal domain has two stretches of hydrophobic amino acids; one or both of these stretches traverse the lipid bilayer of membranes in virus particles and cells. As for other retroviruses, the carboxyl terminal *env* polypeptide serves as an anchor for the externally located amino terminal *env* domain (26). Homology assessments reveal that the carboxyl terminal domain of SRV-1 *env* is more conserved with the counterpart region of other retroviruses than is the amino terminal domain.

SRV-1 encodes a short open reading frame downstream from the *env* gene. The short open reading frame contains 104 codons and extends into the rightward LTR. Several other retroviruses (MMTV, HTLV-I, HTLV-II, and the AIDS retroviruses) encode long open reading frames beyond *env* that extend into the LTR. The SRV-1 short open reading frame is about half the size of the long open reading frame gene of human AIDS retroviruses. The counterpart region in the closely related SRV-2 (3) genome shows extensive DNA homology; however, the same translation frame in this part of SRV-2 is interrupted with five stop codons (40). Thus, the significance of the SRV-1 short open reading frame is not established.

The prototype member of the group of retroviruses to which SRV-1 belongs is MPMV. SRV-1 and MPMV are closely

related, according to nucleic acid hybridization and serological analyses (2, 3). Differences lie primarily in the viral envelope. Sequence analysis of the amino-termini of several MPMV *gag* proteins (27) permitted identification of processing sites in the SRV-1 *gag* polypeptide precursor (Fig. 1).

The open reading frame encoding the SRV-1 protease gene, designated here as *prt*, has a translation frame different from the *gag* or *pol* genes (Fig. 1). The *prt* genes of bovine leukemia virus and HTLV-II are also organized as independent open reading frames with respect to their *gag* and *pol* genes (Fig. 4). In descriptions of the HTLV-II and bovine leukemia virus genomes, the corresponding protease gene was referred to as ORF-HL and ORF-BL, respectively (30). The *prt* open reading frame in SRV-1 is much longer than the corresponding ORF-HL and ORF-BL. The SRV-1 *prt* gene shows considerable homology with the counterpart gene in the hamster IAP genome; both regions are about twice as long as the protease domains in all other retroviruses and eukaryotic transposons.

We divided the *prt* open reading frame into three regions. The second region has significant homology with other retroviral proteases and cellular proteases of the aspartyl class. The first region appears to have some homology to the second region, suggesting that the SRV-1 (and hamster IAP) *prt* region may have arisen by gene duplication. It is not known whether the unusual protease of SRV-1 plays a role in viral pathogenesis. New antiviral strategies could be directed at inhibiting this unusually large protease enzyme predicted by the genome sequence of SRV-1. The exact biogenesis of *prt* and *pol* polypeptides of these simian type D retroviruses remains to be elucidated, but mechanisms such as splicing at the RNA level or frameshifting at the translation level could be used to produce precursor polypeptides. In cells infected with MPMV, three large proteins [with molecular weights of 78,000 (78K), 95K, and 180K] containing *gag* sequences are observed (28). The 78K species is the precursor for the *gag* proteins. The genetic organization of SRV-1 revealed here supports the suggestion that the 95K and 180K polypeptides are precursors for *gag-prt* and *gag-prt-pol*, respectively.

A second strain of type D retrovirus (SRV-2) has been isolated from macaque monkeys at the Oregon Primate Center, which develop both immunodeficiency (SAIDS) and retroperitoneal fibromatosis (3). SRV-2-specific proviral sequences are present in lymphoid and retroperitoneal fibromatosis tissues but not in muscle tissue of diseased macaques (39). This virus is

closely related to SRV-1 by nucleic acid hybridization studies and serological analyses (3). The DNA sequence of molecularly cloned SRV-2 reveals a high degree of homology to SRV-1; the greatest sequence variation is in the amino terminal domain of the *env* genes (40). Genetically engineered recombinant viruses, made by exchanging portions of these related viruses, will be useful in determining which sequence of the genome of SRV-2 is associated with retroperitoneal fibromatosis. Novel vaccine strategies that use, for example, subunit envelope proteins of SRV-1 or SRV-2 expressed in recombinant microorganisms may help control SAIDS in infected primates.

Note added in proof: The location of the *gag* phosphoprotein of SRV-1 in Fig. 1 is based on the published amino terminal sequence of the MPMV phosphoprotein (pp18) (27). The amino terminal sequence of the phosphoprotein of the related SAIDS retrovirus D/W isolate, from the Washington Regional Primate Research Center (27), corresponds to *gag* position 107 in Fig. 1. Thus, SRV-1 contains DNA sequences encoding the published amino termini of both MPMV and SAIDS retrovirus D/W.

REFERENCES AND NOTES

1. D. Fine and G. Schochetman, *Cancer Res.* **38**, 3123 (1978).
2. P. A. Marx *et al.*, *Science* **223**, 1083 (1984); D. H. Maul *et al.*, *Am. J. Vet. Res.*, in press.
3. K. Stromberg *et al.*, *Science* **224**, 289 (1984).
4. M. D. Daniel *et al.*, *ibid.* **223**, 602 (1984).
5. M. Bryant *et al.*, *Hematol. Oncol.* **3**, 187 (1985); G. Heidecker, in preparation.
6. D. H. Maul *et al.*, *Vet. Immunol. Immunopathol.* **8**, 201 (1985); K. G. Osborn *et al.*, *Am. J. Pathol.* **114**, 99 (1984).
7. P. J. Kanki *et al.*, *Science* **228**, 1199 (1985); M. D. Daniel *et al.*, *ibid.*, p. 1201; N. L. Letvin *et al.*, *ibid.* **230**, 71 (1985).
8. N. Lerche *et al.*, in preparation.
9. N. E. Varmus, *Science* **216**, 812 (1982).
10. H. Temin, *Cell* **27**, 1 (1981).
11. L. A. Donehower *et al.*, *J. Virol.* **34**, 226 (1980).
12. K. Shimotohno *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3101 (1985).
13. M. Seiki *et al.*, *ibid.* **80**, 3618 (1983).
14. M. A. Muesing *et al.*, *Nature (London)* **313**, 450 (1985).
15. S. Wain-Hobson *et al.*, *Cell* **40**, 9 (1985).
16. L. Ratner *et al.*, *Nature (London)* **313**, 227 (1985).
17. R. Sanchez-Pescador *et al.*, *Science* **227**, 484 (1985).
18. H. R. Chen and W. C. Barker, *Nucleic Acids Res.* **12**, 1767 (1984).
19. P. Sonigo *et al.*, *Cell* **42**, 369 (1985).
20. R. Breathnach and P. Chambon, *Annu. Rev. Biochem.* **50**, 349 (1981).
21. T. Yamamoto, B. deCrombrugge, I. Pastan, *Cell* **22**, 788 (1980).
22. N. J. Proudfoot and G. G. Brownlee, *Nature (London)* **252**, 359 (1974).
23. H. Weiher, M. König, P. Gruss, *Science* **219**, 626 (1983).
24. W. S. Dynan and R. Tjian, *Nature (London)* **316**, 774 (1985).
25. M. Kozak, *Nucleic Acids Res.* **12**, 857 (1983).
26. C. Dickson *et al.*, in *RNA Tumor Viruses*, R. A. Weiss, N. M. Teich, H. E. Varmus, J. M. Coffin, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982), p. 513.
27. L. E. Henderson *et al.*, *J. Virol.* **55**, 778 (1985).
28. J. Bradac and E. Hunter, *Virology* **138**, 260 (1984).
29. T. D. Copeland *et al.*, *ibid.* **133**, 137 (1984).

30. N. Sagata *et al.*, *FEBS Lett.* **178**, 79 (1984).
31. H. Toh *et al.*, *EMBO J.* **4**, 1267 (1985).
32. N. Sagata *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4741 (1984).
33. D. Derse *et al.*, *Virology* **141**, 162 (1985).
34. M. Ono *et al.*, *J. Virol.* **55**, 387 (1985).
35. H. Toh *et al.*, *Nature (London)* **305**, 827 (1983).
36. I.-M. Chiu, R. Callahan, S. R. Tronick, J. Schlom, S. A. Aaronson, *Science* **223**, 364 (1984).
37. P. J. Hippenmeyer and D. P. Grandgenett, *Virology* **137**, 358 (1984).
38. D. Schwartz *et al.*, *Cell* **32**, 853 (1983).

39. M. Bryant *et al.*, *Virology*, in press.
40. R. Thayer *et al.*, unpublished information.
41. J. Messing and J. Viera, *Gene* **19**, 269 (1982).
42. D. Shinnick *et al.*, *Nature (London)* **293**, 543 (1981).
43. We thank J. D. Kluge, V. Sharpe, S. Sweet, D. Topping, A. Thomas, and C. Sarason for technical assistance. Supported in part by NIH grants AI20573, CA37467, RR00169, and by special appropriation from the State of California.

21 October 1985; accepted 21 January 1986

Inhibition of Vasopressin Action by Atrial Natriuretic Factor

MARK A. DILLINGHAM AND ROBERT J. ANDERSON

Atrial natriuretic factor results in diuresis in animals and humans, perhaps because atrial natriuretic factor increases renal blood flow. The possibility that this diuresis is due to direct inhibition of renal tubular epithelial water transport was examined in rabbit collecting tubules perfused *in vitro*. Atriopeptin III inhibition of the hydraulic conductivity response to the hormone arginine vasopressin but not to either 3'5'-cyclic adenosine monophosphate or forskolin was found. These results suggest that atriopeptin III acts proximal to cyclic adenosine monophosphate formation to directly affect vasopressin-stimulated water transport in the mammalian nephron. They also suggest a potential role for inhibition by atrial natriuretic factor of the renal response to arginine vasopressin as a contributor to a diuretic state.

MAMMALIAN ATRIA CONTAIN SECRETORY granules. In response to atrial distension, these granules release a group of closely related 21- to 26-amino-acid peptides (1-3). Collectively referred to as atrial natriuretic factor (ANF), these peptides exert potent vasodilatory and diuretic effects in animals and humans (1-7). The mechanism of the diuretic action of ANF has not yet been established (1-3). Evidence obtained *in vivo* suggests that ANF increases the glomerular filtration rate and that the filtered load of salt and water is responsible for the diuresis (2-4). However, ANF reduces systemic arterial pressure, has a variable effect on renal blood flow and vascular resistance, and increases solute excretion (1-7). The multiplicity of systemic and intrarenal effects exerted by ANF *in vivo* renders determination of a direct renal epithelial cell effect on salt and water transport difficult. Whether tubular effects also contribute to the diuresis of ANF is unclear. To date, there is no evidence that ANF directly inhibits intact renal tubular epithelial cell salt and water reabsorption (1-3, 7). Our studies were designed to determine whether ANF exerts a direct effect on water transport in renal collecting tubules.

Individual rabbit renal cortical collecting tubules were obtained by microdissection and perfused *in vitro* by slight modifications of the method developed by Burg and oth-

ers (8-10). Tubules were bathed in a solution of NaCl, 115; MgSO₄, 1.2; CaCl₂, 1.0; KCl, 5.0; sodium acetate, 10; NaH₂PO₄, 1.2; NaHCO₃, 25; and dextrose, 5.5 (all in millimoles per liter). Bath fluid of pH 7.40 and 25°C was completely changed every 3 to 4 minutes (10). The composition of the perfusion fluid was the same as that of the bathing fluid except that the final concentration of NaCl was reduced to 50 mmol/liter. Perfusion fluid also contained sufficient [¹⁴C]inulin (New England Nuclear) to result in collected fluid count per minute at least 10- to 15-fold above background. The tubule was visually inspected at 1- to 3-minute intervals throughout the study. Hydraulic conductivity (measured in cm atm⁻¹ sec⁻¹ × 10⁻⁷) was calculated from the formula derived by Al-Zahid *et al.* (11).

Collecting tubules were allowed to equilibrate at 25°C for 4 hours (10). Tubules were perfused at 10 to 12 nl/min by adjusting hydrostatic pressure of the fluid entering the perfusion pipette. Tubular length was comparable in all groups of studies. In all studies each tubule served as its own control. In each tubule four or five collections were obtained for measurement of hydraulic con-

Department of Medicine, University of Colorado Health Sciences Center, and Veterans Administration Medical Center, Denver, CO 80262.