FEBS Letters 414 (1997) 50–54 FEBS 19136

Probing the structure of HIV-1 Rev by protein footprinting of multiple monoclonal antibody-binding sites

Torben Heick Jensen[a], Allan Jensen[a], Anne Marie Szilvay[b], Jørgen Kjems[a,*]

[a]*Department of Molecular and Structural Biology, University of Aarhus, C.F. Møllers Allé Building 130, DK-8000 Aarhus C, Denmark*
[b]*Centre for Research in Virology, The Gade Institute, Department of Microbiology and Immunology, University of Bergen, Bergen, Norway*

Received 14 July 1997

Abstract Human immunodeficiency virus type 1 (HIV-1) Rev is a small RNA-binding protein which is essential for viral replication. To investigate the structure of Rev we have mapped the binding sites of a panel of monoclonal antibodies (mAb) by protein footprinting and identified a mAb protecting amino acids within both the N- and C-terminal parts of Rev. Our mapping results support a previously proposed structure (Auer et al., *Biochemistry*, 33 (1994) 2988–2996) predicting that a helix–loop–helix motif in Rev brings the termini of the protein into proximity. Furthermore, we demonstrate that the binding sites mapped by protein footprinting are in agreement with conventional epitope mapping results and that this technique provides an advantageous strategy for mapping discontinuous sites.

Key words: HIV; Rev; Protein footprinting; Epitope mapping; Conformational epitope

1. Introduction

Human immunodeficiency virus type 1 (HIV-1) encodes a regulatory protein Rev which is required for expression of viral structural proteins. It binds to an RNA element, the Rev response element (RRE), and up-regulates the cytoplasmic appearance of incompletely spliced viral transcripts, thereby inducing a shift in viral protein synthesis from regulatory to structural proteins. Rev is 116 amino acids (aa) long and based on mutational analysis; it has been divided into functional domains essential for biological activity (Fig. 1A). These include an arginine-rich RNA-binding/nuclear localization domain (aa 34–50) [1–4] encompassed by less well-defined sequences necessary for multimerization of the Rev protein [2,5–7], and a leucine-rich nuclear export signal (NES, aa 78–83) [8,9]. Whereas the functional implications of the interaction between the cellular protein p32 and the RNA-binding domain remain unclear [10,11], the transport is probably mediated by one or multiple cellular protein(s) interacting with the NES. Thus far two types of factors have been described: the nucleoporin-like protein Rab/hRip/Rip1p [12–14] and the eukaryotic initiation factor 5A (eIF-5A) [15].

Determination of the 3-dimensional structure of full-length Rev by NMR and X-ray crystallography has been impeded by the strong tendency of the protein to aggregate at higher concentrations [16] and so far NMR has only solved the structure of a peptide spanning the RNA-binding domain of Rev [17]. The information available on the full-length Rev structure derives mainly from circular dichroism (CD) experiments [18–22] and from a footprinting assay involving partial proteolysis of the protein [23]. CD spectra have estimated that full-length Rev protein contains approximately 50% α-helix, and an analysis of fragments of Rev suggests that the majority of the N-terminal part of the protein (aa 8–66) is α-helical in solution, whereas the C-terminal part of the protein is structurally more flexible [18]. A peptide spanning the RNA-binding domain of Rev (aa 34–50) also confines to an α-helical conformation [24], a feature recently confirmed by NMR [17]. Furthermore, partial proteolysis studies suggest that the RNA-binding domain is also α-helical in the context of the full-length protein [23]. A model suggested by Auer et al. implies, that the N-terminal α-helical is interrupted by a proline-rich loop (aa 27–31) and that the flanking α-helical segments interact through conserved hydrophobic residues forming a helix–loop–helix motif in Rev [18,25].

Using a protein footprinting approach, which involves limited proteolysis of N- or C-terminally radiolabeled Rev protein in either the free state or bound to a panel of monoclonal antibodies (mAb), we here report a different line of experimental evidence supporting such a structure.

2. Materials and methods

2.1. Protein expression vectors

The pGEX–GTH–Rev expression vector, producing C-terminally radiolabeled GST-tagged Rev protein, have been described previously [26]. In order to make Rev protein, containing His-tag and HMK site at opposite ends, two parent vectors were constructed. pET–His–H was obtained by annealing the two synthetic oligonucleotides 5′-

Table 1
Comparison of Rev epitopes mapped by either peptide interaction analysis or protein footprinting

mAb	Peptide interaction[a]	Prot. footprinting[b]
4G9	–[c]	aa 5– 15
8E7, 9G2	aa 70– 84	aa 65– 85
	aa 75– 88	
1G7	aa 91–105	aa 95–105
	aa 96–110	
1G10	aa 91–105	aa 10– 20
	aa 96–110	aa 95–105

[a]Eptope-mapping results from peptide interaction studies (see Section 2). Rev peptides interacting with the indicated mAbs are numbered according to the Rev sequence.
[b]Areas in Rev protected from proteolytic cleavage by indicated mAbs. Endoproteinases used: bromelain and proteinase K (Figs. 2 and 3), and Asp-N, trypsin, Glu-C, elastase and chymotrypsin (data not shown). n.d., not detected.
[c]No peptides were found to bind 4G9.

*Corresponding author. Fax: (45) 8619-6500.
E-mail: kjems@biobase.dk

Abbreviations: HIV-1, human immunodeficiency virus type 1; aa, amino acid(s); NMR, nuclear magnetic resonance; mAb, monoclonal antibody; HMK, heart muscle kinase; GST, glutathione *S*-transferase; NES, nuclear export signal

PII S0014-5793(97)00988-5

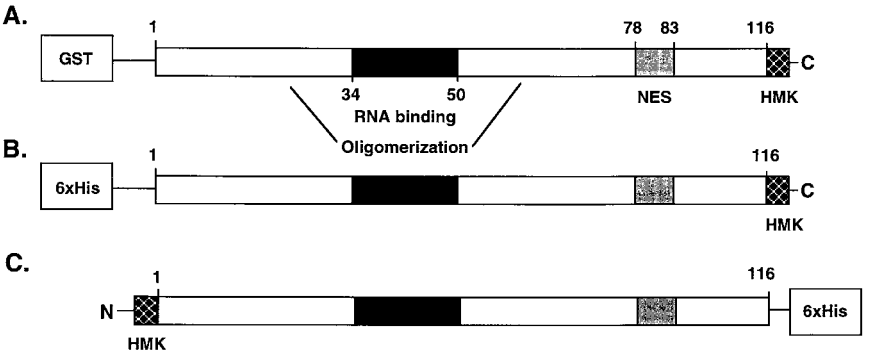

Fig. 1. Schematic representation of functional domains in the Rev fusion constructs used in this study. See text for details. C-terminally ^{32}P-labeled Rev with either a GST (A), or a 6×His (B) purification tag positioned at the N-terminus, and N-terminally ^{32}P-labeled Rev with a C-terminal 6×His tag (C). NES, nuclear export signal; HMK, heart muscle kinase phosphorylation site.

CATGCATCACCATCACCATCACGGATCCCCGGGAATTCGT-CGTGCATCTGTTTGA and 5′-GATCTCAAACAGATGCACGA-CGAATTCCCGGGGATCCTGGATGGTGATGGTGATG, and inserting the double-stranded fragment into a *Nco*I+*Bam*HI digest of pET11dΔ*Eco*RI [26]. pET–H–His was created likewise from the oligonucleotides: 5′-CATGCGTCGTGCATCTGTTGGATCCCCGG-GAATTCATCACCATCACCATCACTGA and 5′-GATCTCAGTG-ATGGTGATGGTGATGAATTCCCGGGGATCCAACAGATGC-ACGACG. To obtain pET–His–Rev–H and pET–H–Rev–His a *Bcl*I–*Eco*RI cleaved Rev PCR fragment was generated as described previously [26] and inserted into the *Bam*HI and *Eco*RI sites of pET–His–H and pET–H–His, respectively.

2.2. Production of radiolabeled proteins

GST-tagged Rev protein was expressed from pGEX–GTH–Rev and radiolabeled as described [26]. To obtain His-tagged radiolabeled Rev proteins, pET–His–Rev–H and pET–H–Rev–His were expressed in *E. coli* strain BL21(DE3), containing an isopropyl-β-D-thiogalactopyra-

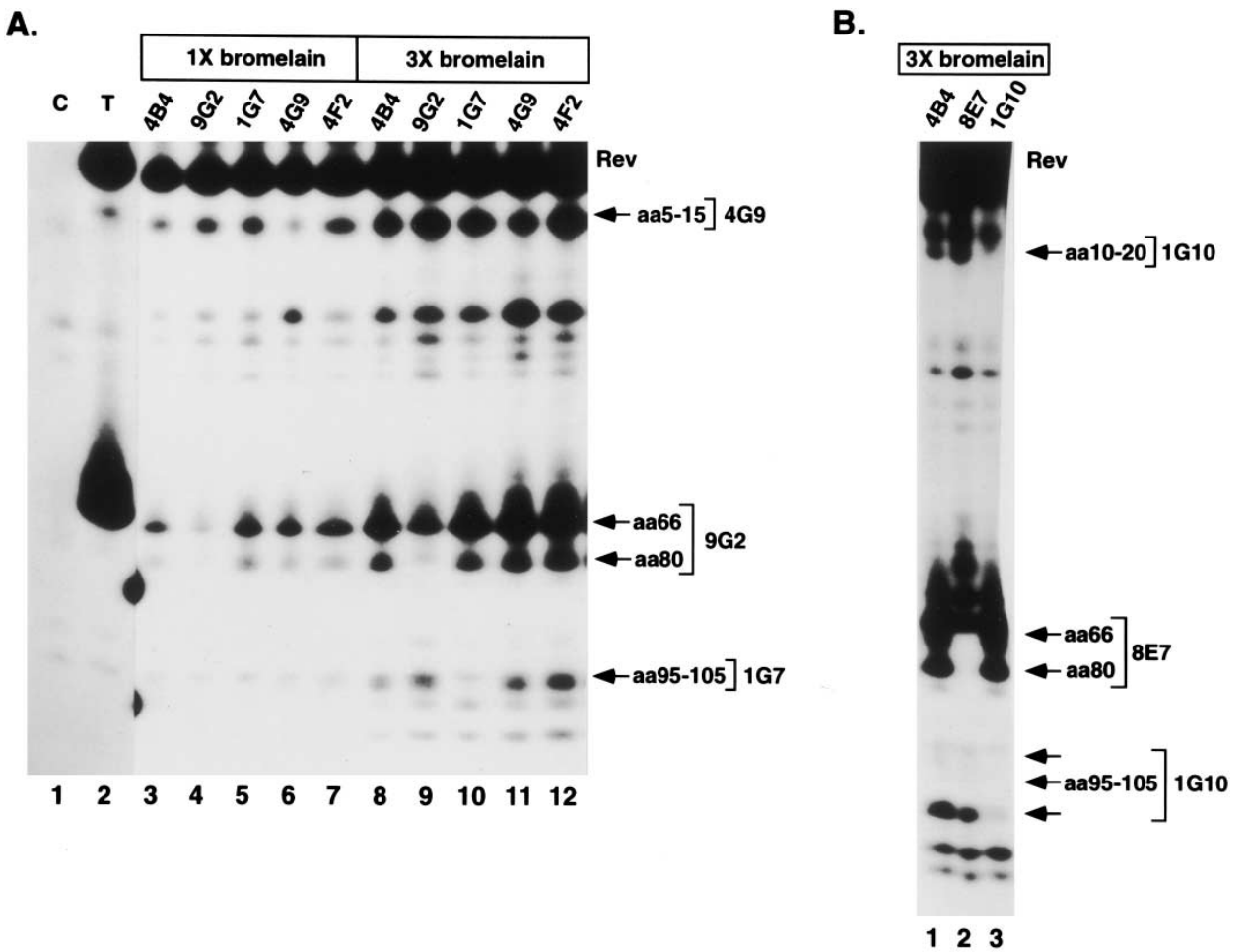

Fig. 2. Autoradiograms of the mAb protection mapping of GST–Rev protein using bromelain. The Rev specific mAbs used were: (A) 9G2, 1G7, 4G9, 4F2; and (B) 8E7 and 1G10 as indicated above the lanes. Horizontal arrows to the right indicate the position of the footprints produced by the denoted mAbs as compared to the control. Where just a single aa in the Rev sequence is given, the exact position of the cleavage site is known. C and T denote control lanes with un- and thrombin-digested GST–Rev, respectively. The thrombin cleavage site in Rev has previously been mapped to aa 66 [23]. Only the portion of the gels covering Rev sequences is shown.

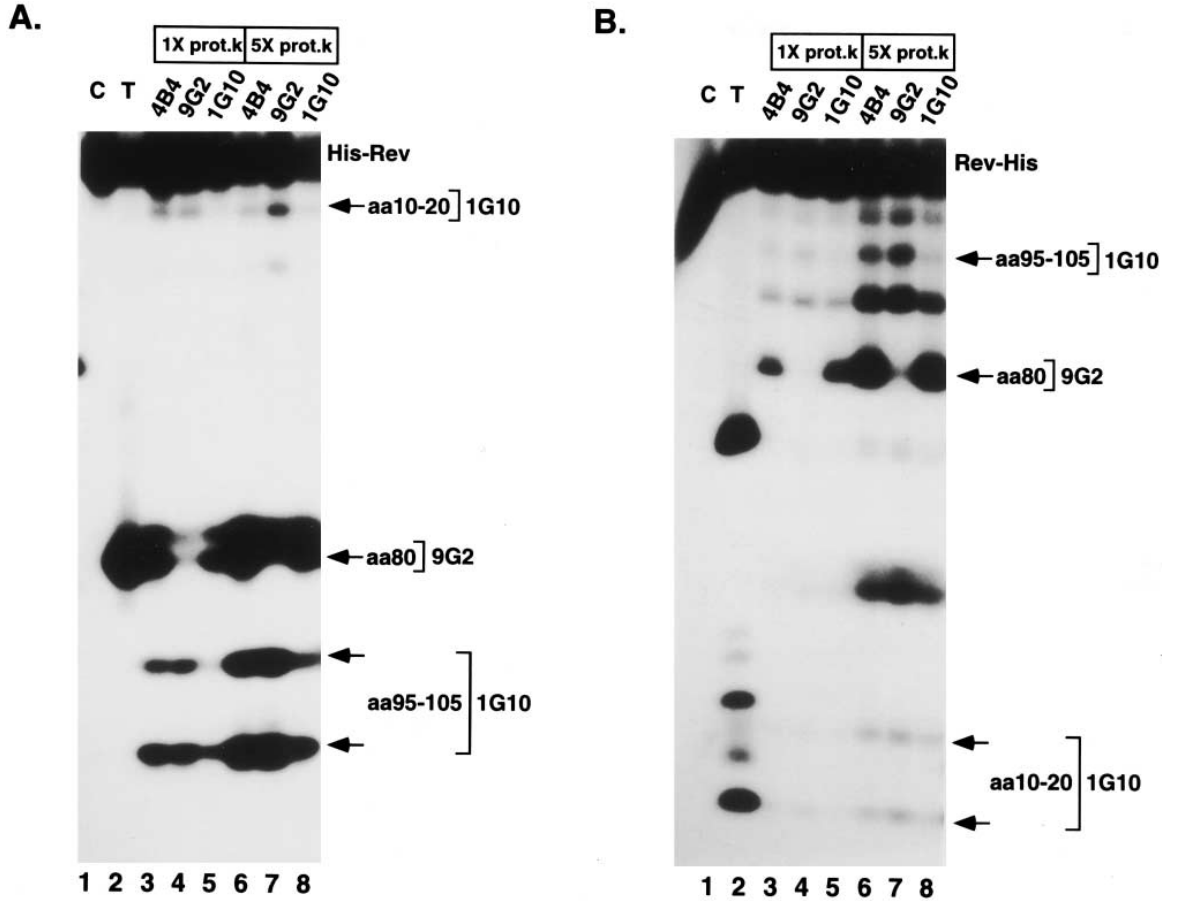

Fig. 3. mAb protection mapping of His-tagged Rev. Autoradiograms of proteinase K digested His–Rev (A), or Rev–His (B) proteins incubated with Tat mAb 4B4 (lanes 3 and 6) or Rev mAbs 9G2 (lanes 4 and 7) and 1G10 (lanes 5 and 8). Nomenclature is as in Fig. 2. As mentioned in the text the 9G2 protected bands migrate abnormally in (A). The exact reason for this discrepancy is not known but small C-terminally peptide fragments created by certain endoproteinases (e.g. Asp-N and proteinase K) has previously been noted to migrate somewhat slower in polyacrylamide gels compared to fragments of the same theoretical size created by other endoproteinases [10,23].

noside inducible *T7* RNA polymerase gene [27]. Proteins were purified under denaturing conditions and renatured as described previously [10]. Dialysed renatured protein was radiolabeled in solution with the HMK enzyme as described [26] and re-incubated with 20 μl of Ni^{2+}-agarose to facilitate 5 washing steps in 500 μl of washing buffer (20 mM HEPES, pH 7.9, 200 mM NaCl, 20% glycerol and 10 mM β-mercaptoethanol) to remove the HMK enzyme and unincorporated nucleotides. Radiolabeled protein was eluted from the Ni^{2+}-agarose in 100 μl of Rev elution buffer (500 mM NaCl, 200 mM imidazole, 50 mM HEPES/KOH, pH 6.0).

2.3. Monoclonal antibodies

mAb 4B4 and mAbs 4G9, 4F2, 9G2, 8E7, 1G7, 1G10 were produced by immunising mice with *E. coli* produced wild-type Tat and Rev protein, respectively. mAbs were purified using mAbTrapG (Pharmacia). Epitope mapping, by binding experiments using overlapping peptides derived from the Rev sequence, has previously been done [28], and the results are summarized in Table 1. mAb 4B4 binds to a fragment of the Tat protein comprising aa 49–86 [29].

2.4. Protein footprinting experiments

All protein footprinting experiments were done as described previously [23] except only 1 ng of both GST- and His-tagged Rev were used. Proteinase concentrations in the final reaction mixtures were: bromelain, 5 pg/μl (1×) and 15 pg/μl (3×) for GST–Rev experiments and 50 pg/μl for His–Rev and Rev–His experiments; proteinase K 0.8 pg/μl (1×) and 4 pg/μl (5×). Elastase and chymotrypsin concentrations were 10–30 pg/μl, and Asp-N, thrombin, trypsin and Glu-C were used as described [23].

3. Results and discussion

3.1. Protein footprinting radiolabeled Rev with mAbs

The topography of Rev was probed by mapping the binding sites of various Rev-specific mAbs using protein footprinting technology. To ease the amino acid identification, the protein had subsequently been ^{32}P-labeled with the heart muscle kinase enzyme (HMK) at the C-terminal HMK site (Fig. 1A). Assembled GST–Rev/mAb complexes were partially cleaved with the following proteinases: Asp-N, trypsin, Glu-C, elastase, chymotrypsin, proteinase K (data not shown), and bromelain (data shown in Fig. 2). As a control for specificity an irrelevant mAb, directed against the HIV-1 Tat protein, was used. Partial proteolysis of GST–Rev with a variety of sequence specific proteinases has previously been performed and the bands observed could readily be assigned to specific amino acids [23]. The positions of the cleavages originating from unspecific proteinases were estimated by comparison to co-electrophoresed markers. Three different mAbs induced contiguous footprints on Rev when using bromelain: (i) 4G9 protected the N-terminal part of Rev around aa 5–15 (Fig. 2A, lanes 6 and 11), (ii) 9G2 and 8E7 gave footprints at aa positions 65–85 (Fig. 2A, lanes 4 and 9, and Fig. 2B, lane 2) and (iii) 1G7 protected cleavages of sequences around

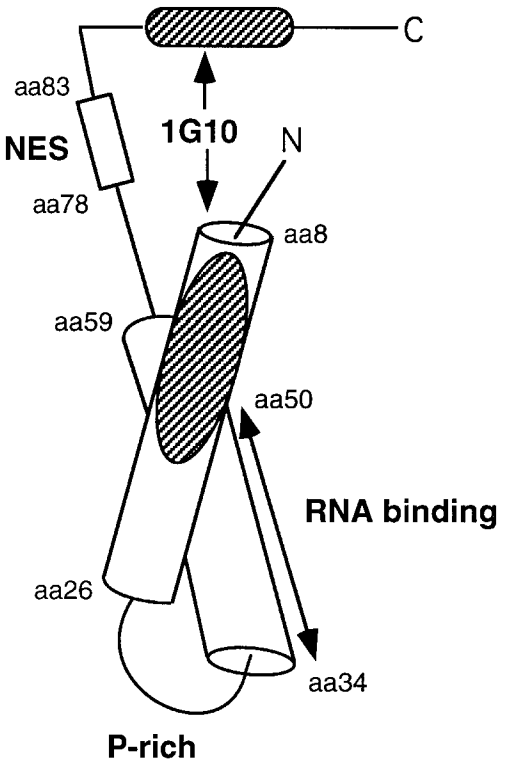

Fig. 4. Putative model of the Rev structure based on this report and the study by Auer et al. [18]The protection of 1G10 is indicated by cross-hatched areas on the model. Cylinders represent α-helices and the C-terminal part of the protein, including the NES, is drawn schematically.

positions 95–105. For mAb 4F2 no effects were seen with the proteinases used.

Interestingly, a discontinuous footprint was obtained with mAb 1G10 yielding protection of cleavages both around aa 10–20 and aa 95–105 (Fig. 2B, compare lanes 1 and 3). The footprint was surprising as previous epitope mapping, using a binding experiment with overlapping Rev peptides, did not find that amino acids in the N-terminal part of Rev were involved in 1G10 binding [28]. The epitope mapping results obtained by peptide interaction analysis and protein footprinting are summarized in Table 1.

3.2. *The conformational epitope is also present in His-tagged Rev*

To test whether the discontinuous protection observed with mAb 1G10 could be reproduced on a Rev protein lacking the GST-tag, we constructed protein expression vectors producing Rev protein containing a 6×His-tag at either the N- or C-terminus (His–Rev and Rev–His, respectively) and a HMK site at the terminus distal to the His-tag (Fig. 1B,C). The advantage of using both N- and C-terminally labeled proteins in protein footprinting experiments is the ability to distinguish primary cleavages (the initial cleavage) from secondary ones (cleavages depending on an initial cleavage event). Purified His-tagged Rev proteins were incubated with 1G10 and partially cleaved with proteinase K (Fig. 3A,B, lanes 5 and 8). Control reactions were performed with Tat mAb 4B4 and Rev mAb 9G2 (Fig. 3A,B, lanes 3, 4, 6 and 7). As seen for GST–Rev, 1G10 also abolished cleavages at positions around aa 10–20 and aa 95–105 in both types of His-tagged Rev proteins, which lead us to conclude that the effects seen are not artefacts due to secondary cleavage events or topological constraints imposed on Rev by the purification tags. 9G2 inhibited proteinase K cleavage at an expected position around aa 80, although in the case of His–Rev the protected bands comigrate with the thrombin cleaved marker, indicating a slightly more N-terminal position at aa 66 (see legend to Fig. 3 for discussion).

From the footprinting analysis we predict that 1G10 recognizes a conformational epitope in Rev comprised of amino acids within positions 10–20 and 95–105, whereas the remaining mAb-binding sites correlate with the epitopes resolved by a peptide-binding procedure (Table 1). Thus, in contrast to conventional peptide mapping strategies the protein footprinting technique seems to be able to yield additional information concerning epitopes consisting of non-linear antigenic determinants.

3.3. *Structural model for Rev*

The discontinuous footprint observed with mAb 1G10 suggests that the termini of Rev are in close proximity. Interestingly, this feature is also favored in the helix–loop–helix model proposed by Auer et al. [18] (Fig. 4). The protein footprinting data cannot resolve whether N- and C-terminal residues are in actual molecular contact since sterical hindrance to the endoproteinases by the mAb could also account for the effects seen. 1G10 recognizes residues more C-terminal than 8E7 and 9G2. Since the latter epitope covers amino acids in the NES of the Rev protein, it is possible that an interaction of the termini plays an important role in presenting the NES to putative cellular partners. More experimental data are required to fully verify this model, but it should prove useful for experimental design.

Acknowledgements: We thank Rita Rosendahl for technical assistence. T.H.J. and A.J. were supported by grants from Aarhus University, and the work was supported by the National Research Council and the Karen Elise Jensen Foundation.

References

[1] Malim, M.H., Bohnlein, S., Hauber, J. and Cullen, B.R. (1989) Cell 58, 205–214.
[2] Olsen, H.S., Cochrane, A.W., Dillon, P.J., Nalin, C.M. and Rosen, C.A. (1990) Genes Dev. 4, 1357–1364.
[3] Perkins, A., Cochrane, A.W., Ruben, S.M. and Rosen, C.A. (1989) J. Acquired Immune Defic. Syndr. 2, 256–263.
[4] Hope, T.J., McDonald, D., Huang, X.J., Low, J. and Parslow, T.G. (1990) J. Virol. 64, 5360–5366.
[5] Bohnlein, E., Berger, J. and Hauber, J. (1991) J. Virol. 65, 7051–7055.
[6] Malim, M.H. and Cullen, B.R. (1991) Cell 65, 241–248.
[7] Zapp, M.L., Hope, T.J., Parslow, T.G. and Green, M.R. (1991) Proc. Natl. Acad. Sci. USA 88, 7734–7738.
[8] Meyer, B.E. and Malim, M.H. (1994) Genes Dev. 8, 1538–1547.
[9] Szilvay, A.M., Brokstad, K.A., Kopperud, R., Haukenes, G. and Kalland, K.H. (1995) J. Virol. 69, 3315–3323.
[10] Tange, T.O., Jensen, T.H. and Kjems, J. (1996) J. Biol. Chem. 271, 10066–10072.
[11] Luo, Y., Yu, H. and Peterlin, B.M. (1994) J. Virol. 68, 3850–3856.
[12] Stutz, F., Neville, M. and Rosbash, M. (1995) Cell 82, 495–506.
[13] Fritz, C.C., Zapp, M.L. and Green, M.R. (1995) Nature 376, 530–533.
[14] Bogerd, H.P., Fridell, R.A., Madore, S. and Cullen, B.R. (1995) Cell 82, 485–494.
[15] Ruhl, M., Himmelspach, M., Bahr, G.M., Hammerschmid, F.,

Jaksche, H., Wollf, B., Aschauer, H., Farrington, G.K., Probst, H., Bevec, D. and Hauber, J. (1993) J. Cell Biol. 123, 1309–1320.
[16] Heaphy, S., Finch, J.T., Gait, M.J., Karn, J. and Singh, M. (1991) Proc. Natl. Acad. Sci. USA 88, 7366–7370.
[17] Battiste, J.L., Mao, H., Rao, S., Tan, R., Muhandiram, D.R., Kay, L.E., Frankel, A.D. and Williamson, J.R. (1996) Science 273, 1547–1551.
[18] Auer, M., Gremlich, H.U., Seifert, J.M., Daly, T.J., Parslow, T.G., Casari, G. and Gstach, H. (1994) Biochemistry 33, 2988–2996.
[19] Daly, T.J., Rusche, J.R., Maione, T.E. and Frankel, A.D. (1990) Biochemistry 29, 9791–9795.
[20] Nalin, C.M. et al. (1990) Proc. Natl. Acad. Sci. USA 87, 7593–7597.
[21] Tan, R., Chen, L., Buettner, J.A., Hudson, D. and Frankel, A.D. (1993) Cell 73, 1031–1040.
[22] Wingfield, P.T., Stahl, S.J., Payton, M.A., Venkatesan, S., Misra, M. and Steven, A.C. (1991) Biochemistry 30, 7527–7534.
[23] Jensen, T.H., Leffers, H. and Kjems, J. (1995) J. Biol. Chem. 270, 13777–13784.
[24] Tan, R. and Frankel, A.D. (1994) Biochemistry 33, 14579–14585.
[25] Thomas, S.L., Hauber, J. and Casari, G. (1997) Prot. Eng. 10, 103–107.
[26] Jensen, T.H., Jensen, A. and Kjems, J. (1995) Gene 162, 235–237.
[27] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Methods Enzymol. 185, 60–89.
[28] Kalland, K.H., Szilvay, A.M., Langhoff, E. and Haukenes, G. (1994) J. Virol. 68, 1475–1485.
[29] Valvatne, H., Szilvay, A.M. and Helland, D.E. (1996) AIDS Res. Hum. Retroviruses 12, 611–619.