Generation of G-to-A and C-to-U Changes in HIV-1 Transcripts by RNA Editing

K. Bourara, S. Litvak, A. Araya*

RNA editing involves posttranscriptional alterations of messenger RNA (mRNA) sequences modifying the information content encoded by the genetic message. Here, it is shown that, in chronically infected H9 cells, human immunodeficiency virus–type 1 (HIV-1) mRNAs undergo guanine-to-adenine (G-to-A) and cytosine-to-uracil (C-to-U) changes. G-to-A modification in the untranslated region of exon 1 was present only in spliced HIV-1 mRNAs. The creation of stop codons in HIV-1 mRNAs may function to control the translation of viral proteins, such as viral protein R, that are involved in the regulation of HIV-1 expression and the survival of chronically infected cells.

Retroviruses display a complex pathway of gene expression controlled by viral- and cellular-encoded regulatory proteins with implications in pathogenesis (*1, 2*). At the translational level, overlapping reading frames are resolved by ribosomal frameshift and readthrough (*3*). At the transcriptional level, RNA editing provides yet another mechanism for controlling viral gene expression through posttranscriptional modifications of mRNAs that introduce amino acid changes or initiation or termination codons (*4–9*).

We analyzed viral protein R (*vpr*) transcripts in H9 cells chronically infected with HIV-1 (*10*). *vpr* encodes a 96–amino acid protein whose biochemical function is not yet understood (*11, 12*). Mutations in *vpr* have been observed in T cells chronically infected with HIV-1, suggesting that reduced or abolished expression of VPR may be related to cell survival (*13*).

The HIV-1 genome organization and the splicing pattern of some mRNAs are shown in Fig. 1A. Polyadenylated [poly(A)$^+$] RNA was isolated from chronically HIV-1 Bru–infected H9 cells (*14*). This model was used because it contains a wild-type *vpr* open reading frame (ORF), and the chronically infected state of the cells precludes the generation of modifications by error-prone retrotranscription. Moreover, the choice of an infected clonal cell line is important to ensure that we are observing posttranscriptional events that are not emerging from variant provirus in heterogeneous samples obtained from HIV-1–infected individuals. Viral mRNA was submitted to reverse transcriptase–polymerase chain reaction (RT-PCR) (Fig. 1B), cloned, and sequenced. We detected several base changes in the 5′ noncoding and *vpr* sequences in the viral mRNAs, relative to the genomic sequence (*15*). G-to-A changes were observed at positions 181, 5301, and 5363, and a C-to-U change was observed at position 5396 (Fig. 2A). Partially modified and unmodified RNAs were also observed. We found no modifications in the 5′ trans-activation responsive region of HIV-1 RNA, which has been reported to be edited by double-stranded RNA adenine deaminase after injection into *Xenopus* oocytes (*16*).

To test whether RT-PCR might have generated the observed changes, we used a genomic fragment spanning the viral infectivity factor (*vif*)–*vpr* region (nucleotides 5029 to 5385) to synthesize sense and antisense RNAs (*17*), which were subjected to RT-PCR, cloned, and sequenced. From 20 independent clones analyzed, neither the sense nor the antisense RNAs were modified, showing that neither reverse transcription, DNA synthesis, nor a specific template (hot spots) caused the base changes observed in retroviral mRNAs.

HIV-1 DNA synthesis is characterized by a high mutation rate, particularly in G-to-A changes (*18, 19*). To determine if RNA modifications were the result of an RNA editing event, we analyzed the sequences of (i) viral mRNA transcripts, (ii) integrated proviral DNA, and (iii) encapsidated genomic RNA, using primers 6 and 7. These primers amplify all mRNAs containing the *vif-vpr* ORFs. Four

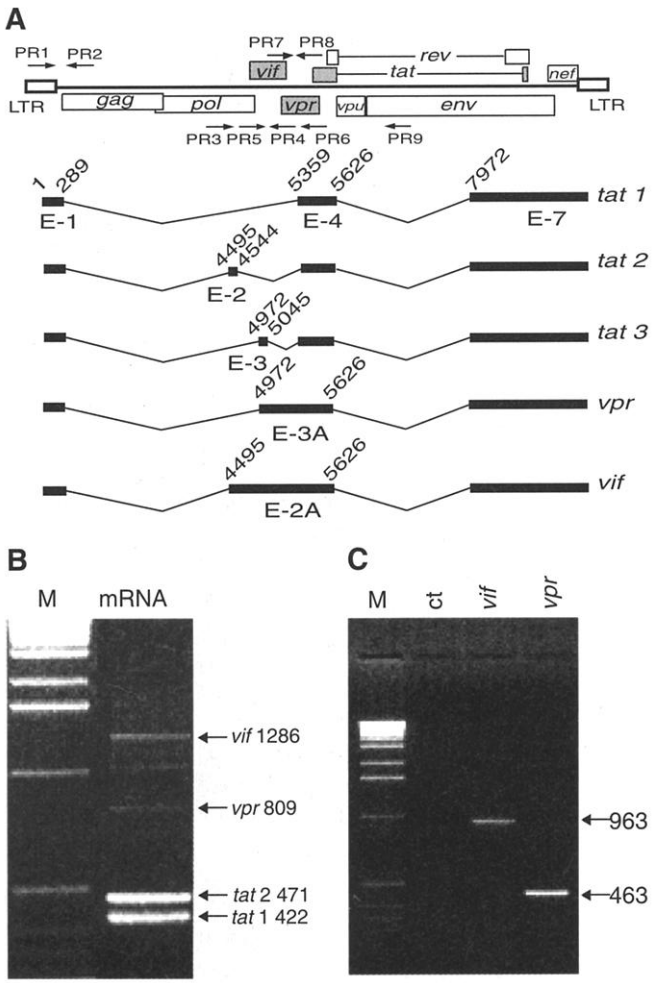

Fig. 1. (**A**) Scheme of the HIV-1 RNA genome. Boxes represent the viral ORFs. Thick lines represent exons (E) involved in the splicing of *tat*, *vpr*, and *vif* mRNAs. The borders of different exons are indicated by genomic numbering. Primers (PR) are indicated by arrows. (**B**) RT-PCR analysis of poly(A)$^+$ mRNA with primers 1 and 6. Numbers indicate the fragment length in base pairs. M, molecular weight markers, 1-kb ladder (Gibco-BRL). (**C**) *vif* (1286 bp) and *vpr* (809 bp) fragments were isolated from agarose gel and amplified with primers 5 and 6 and primers 7 and 6, respectively. ct, control PCR analysis with primers 5 and 8 to detect cross contamination of the *vif* fragment on the *vpr* (809 bp) product.

Laboratoire de Réplication et Expression des Génomes Eucaryotes et Rétroviraux, UMR 5097, CNRS–Université Victor Segalen-Bordeaux 2, and IFR 66 Pathologies Infectieuses, 146 rue Léo Saignat, 33076 Bordeaux Cedex, France.

*To whom correspondence should be addressed. E-mail: Alexandre.Araya@reger.u-bordeaux2.fr

Table 1. RNA editing frequency in mature and unspliced HIV-1 mRNAs. Data are shown for base changes in *vpr*-containing mRNAs after RT-PCR with primers 1 and 6. The number of edited molecules out of the number of analyzed clones is indicated.

Base change	Residue number	*tat*1	*tat*2	*vpr*	*vif*	Unspliced mRNA
G → A	181	8/10	6/7	0/1	5/5	0/18*
G → A	5301	–	–	0/1	5/5	9/10†
G → A	5363	7/8	5/7	0/1	5/5	9/10†
C → U	5396	7/8	5/7	0/1	5/5	9/10†

*RNA editing was analyzed with primers 1 and 2. †RNA editing was analyzed with primers 3 and 9.

Table 2. Base changes in *vif* and *vpr* mRNAs analyzed with primers 5 and 6 and primers 7 and 6, respectively.

Base change	Residue number	*vpr*	*vif*
C → U	5396	5/11	9/10
G → A	181	5/11*	5/7*
G → A	5301	5/11	8/10
G → A	5363	5/11	9/10

*RNA editing was analyzed with primers 1 and 4.

kinds of mRNAs were found: 13 out of 30 RNAs were nonmodified, 4 out of 30 were modified at site 5301, 1 out of 30 was modified at sites 5363 and 5396, and 12 out of 30 were modified at all sites (Fig. 2, A and B). One explanation is that at least four different proviruses are present in the nuclear genome. To test this, we performed PCR experiments on genomic DNA. None of the 11 clones analyzed showed sequence changes, indicating that the observed modifications derive from viral transcripts of one provirus. This conclusion was confirmed by Southern blot analysis of total DNA using *pol*- and *env*-specific probes (*20, 21*).

Next, we analyzed the full-length, unspliced genomic RNA from virions produced by the infected H9 cells to determine if RNA polymerase II (RNA Pol II) generated these changes (*20*). The viral genome is the primary product of RNA Pol II and is indistinguishable from the mRNA assigned to translation, except that the former is encapsidated and the latter is engaged in RNA splicing and translation (*22*). If base changes occur cotranscriptionally, viral particles produced by infected cells should contain either unmodified or modified genomic RNA. None of the 24 independent clones analyzed showed base changes. Thus, RNA editing modifications are restricted to transcripts intended for processing and translation, suggesting that editing allows a distinction between two primary transcript pools. Whether the edited RNA is chosen for translation or precluded from encapsidation remains to be determined.

The changes at positions 5301, 5363, and 5396 are in the *vpr* ORF and affect codons that are highly conserved in various strains of HIV-1 and simian immunodeficiency virus–CPZ. The G-to-A (position 5301) change generated a stop codon, whereas a Gly (GGG) triplet is changed into an Arg (AGG) codon at position 5363 (*21*). Both sites were embedded in a TTGGG motif (underline indicates edited residue), but no consensus was found in surrounding sequences, nor were consistent secondary structures evident in these regions. The C-to-U (position 5396) change, which was always associated with the G-to-A (position 5363) modification, generated a stop codon (UAG) from a Gln (CAG) triplet. C-to-U transitions have been observed in the editing of apolipoprotein B (apoB) mRNA (*9, 23*). The sequence in the vicinity of the C residue in *vpr* mRNA is different from that involved in apoB editing, suggesting that both systems operate differently.

The result obtained with primers 6 and 7 does not discriminate between the different transcripts. In HIV-1, >30 different RNA molecules arise from a full-length transcript through highly regulated splicing events (*22*). We analyzed spliced mRNAs containing the *vpr* ORF using RT-PCR instead of other approaches, such as primer extension. We did so because RT-PCR reveals the status of each editing site in individual alternatively spliced transcript. Moreover, primer extension is not well adapted for sites embedded in TGGG sequences. At least five mRNAs [trans-activator 1 (*tat*1), *tat*2, *tat*3, *vpr*, and *vif*] differing in exon composition can be examined with primers 1 and 6 (Fig. 1B). From 130 clones analyzed, most corresponded to *tat*1 and *tat*2 mRNAs; five clones contained *vif* mRNA, and one corresponded to *vpr* mRNA. No *tat*3 mRNA molecules were detected (Tables 1 and 2).

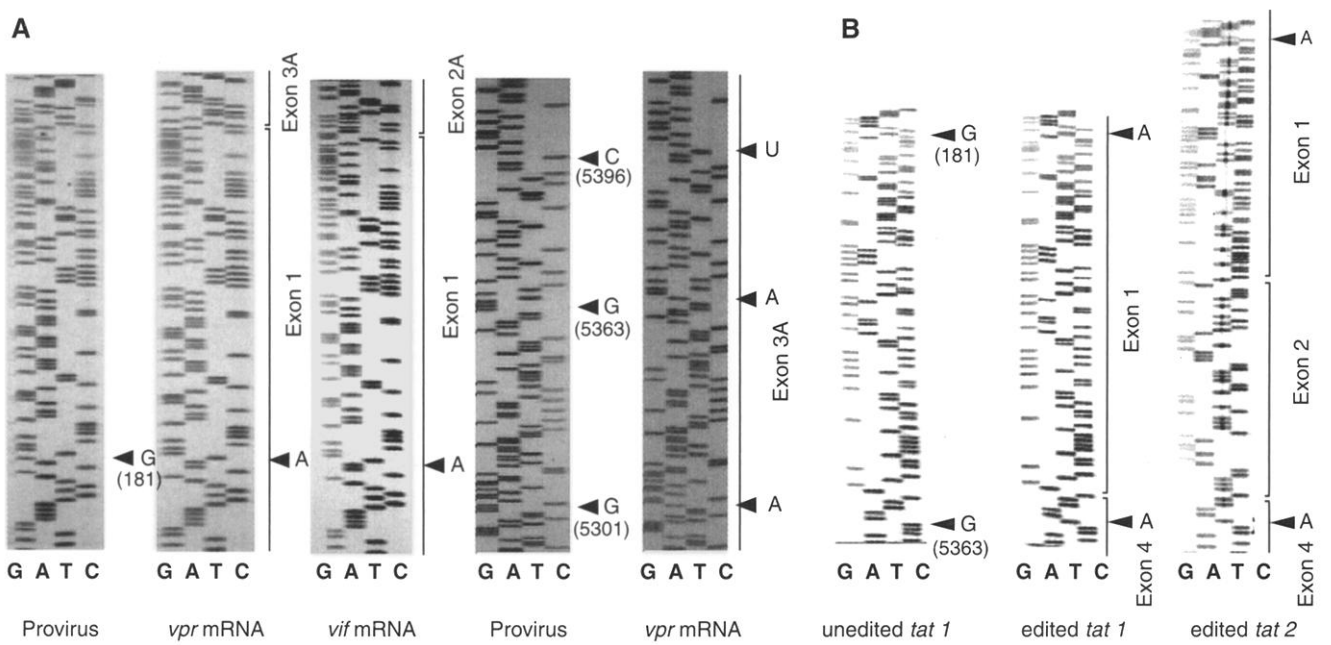

Fig. 2. (**A**) Sequence analysis of *vpr* and *vif* mRNA and the corresponding proviral regions. The exon junctions of the respective mRNAs are indicated. Editing sites are shown by arrowheads. (**B**) Editing sites 181 and 5363 in *tat* mRNAs. *tat*1 and *tat*2 are identified by the exon composition. The antisense strand is shown. All nucleotide positions correspond to the viral genome numbering, irrespective of splicing (*15*).

All *vif* and >90% of *tat*1 and *tat*2 clones showed changes at residues 5301, 5363, and 5396 (Table 1 and Fig. 2, A and B); site 5301 is missing from *tat* mRNAs because exon 4 begins at residue 5359. To increase the number of *vif* and *vpr* clones, we amplified PCR products obtained with primers 1 and 6 by using specific primers (Fig. 1C and Table 2). All three changes were found in 4 out of 11 and 8 out of 10 *vpr* and *vif* clones, respectively, whereas partially edited clones containing the changes 5363 and 5396 were found in 1 out of 11 *vpr* and 1 out of 10 *vif* clones. Unspliced mRNAs, amplified with primers 3 and 9, showed that 9 out of 10 clones carried all three changes (Table 1). The *vpr* ORF in unspliced mRNAs also had stop codons generated by RNA editing, thereby reducing the possibility of producing VPR by ribosomal scanning (*3, 24*).

RNA editing may control alternative splicing of HIV-1 mRNAs by creating the splicing silencing sequence TAGG (*25*) from TGGG in exon 3A of *vpr* mRNA. That could account for the low levels of *vpr* mRNA found in chronically infected H9 cells. The absence of *tat*3 mRNA in the analyzed clones supports this idea because *tat*3and *vpr* mRNAs share the same splicing acceptor site (*26*). Our attempts, using polyclonal antibodies, to detect truncated VPR resulting from newly generated termination codons were unsuccessful. It is unclear if this is due to an inherent instability or a low rate of production. In any case, reduced VPR expression resulting from the creation of stop codons and the decrease of spliced mRNA levels could explain how chronically infected H9 cells, in contrast with other systems, conserve an intact *vpr* ORF (*13*).

Residue G at site 181 in the leader region of exon 1, common to all viral mRNAs, was analyzed in *tat, vif, vpr,* and unspliced mRNAs. The G-to-A transition at site 181 was found with a high frequency in processed mRNAs but was absent from unspliced transcripts (Tables 1 and 2), suggesting that RNA editing of the *vpr* ORF and the leader region occur independently. Site 181 is not embedded in a sequence resembling the target TTGGG present at sites 5301 and 5363. The importance of the change at site 181, located at position +1 of the tRNA primer binding site, is not clear. It has been proposed that nucleotide substitutions in the U5 region are critical for efficient reverse transcription (*27*). RNA editing in this leader sequence would generate an AUG initiation triplet from GUG, upstream from the start codons of *tat*, *vif*, and *vpr* mRNAs, which could affect translation of the corresponding ORFs (*3, 24*). A preliminary search for editing in other mRNAs has not yielded consistently altered sites, except for *vif* mRNA, which presents two G-to-A changes at positions 4800 and 4831 in sequences TAGGG and TTGGG, respectively (*21*). This region is spliced out in all other mature mRNAs.

C-to-U changes in HIV-1 RNA may occur by deamination, as shown in apoB mRNA (*23*) and plant mitochondrial RNA editing (*28*), but G-to-A changes are difficult to explain by a one-step mechanism. The nature of the editing products remains to be determined. Results presented here reflect the Watson-Crick base-pairing rules that operate during DNA synthesis, RT-PCR, and sequencing protocols and do not necessarily describe the chemical identity of the modified residue. It cannot be ruled out that the RNA editing activity generates a residue such as 2,6-diaminopurine from G, which would be able to base pair like an A (*29, 30*). Other mechanisms (transglycosylation and nucleotide exchange, for example) cannot be excluded. Another possibility is that these changes are C-to-U modifications in an antisense RNA; however, using different approaches, we found no evidence for such antisense RNAs in the regions analyzed.

G-to-A and, to a lesser extent, C-to-U modifications are observed in HIV-1 mRNA of chronically infected cells. Because these modifications are specific to mRNAs engaged in viral gene expression and are not observed in either genomic viral RNA or proviral sequences, RNA editing may play a role in the modulation of HIV-1 gene expression. We are currently trying to extend this observation by obtaining clonal cells from infected individuals to ascertain whether the changes described here in a cultured clonal line can be correlated to the HIV life cycle in vivo.

References and Notes

1. J. M. Coffin, S. H. Hughes, H. E. Varmus, Eds., *Retroviruses* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1997), and references therein.
2. S. Litvak, *Retroviral Reverse Transcriptases* (Landes, Austin, TX, 1996), and references therein.
3. S. Schwartz, B. K. Felber, G. N. Pavlakis, *Mol. Cell. Biol.* **12**, 207 (1992).
4. J. P. Jacques, S. Hausmann, D. Kolakofsky, *EMBO J.* **13**, 5496 (1994).
5. J. L. Casey and J. L. Gerin, *J. Virol.* **69**, 7593 (1995).
6. R. Cattaneo, K. Kaelin, K. Baczo, M. A. Billeter, *Cell* **56**, 759 (1989).
7. H. Zheng, T. B. Fu, D. Lazinski, J. Taylor, *J. Virol.* **66**, 4693 (1992).
8. A. Araya, D. Bégu, S. Litvak, *Physiol. Plant.* **91**, 543 (1994).
9. L. M. Powell *et al.*, *Cell* **50**, 831 (1987).
10. Chronically HIV-1 Bru–infected H9 cells were kindly provided by M. Ventura (Bordeaux, France). The viral genotype was described [see (*15*)] (GenBank accession number K02013).
11. D. Trono, *Cell* **82**, 189 (1995).
12. S. Popov *et al.*, *EMBO J.* **17**, 909 (1998).
13. M. E. Rogel, L. I. Wu, M. Emerman, *J. Virol.* **69**, 882 (1995).
14. Poly(A)$^+$ RNA was prepared from chronically infected cells with a Promega kit (PolyATract System 1000). RT-PCR was performed with the appropriate primers (Promega kit). Thirty PCR cycles were performed as follows: 1 min at 94°C, 1 min at 55°C, and 1 min at 72°C in a DNA thermal cycler. No amplification products were obtained when RT was omitted. Control PCR reactions were performed with β-actin mRNA (Promega). The primers (PR) and their genomic localization (in parentheses) were as follows: PR1, GGTCTCTCTGGTTAGACCAG (positions 1 to 20); PR2, CTCTCGCACCCATCTCTCT (330 to 348); PR3, GTAGACATAATAGCAACAGACATAC (4412 to 4436); PR4, CAGCAGTTGTTGCAGAATTC (5325 to 5344); PR5, AAGCTCCTCTGGAAAGGTGA (4529 to 4548); PR6, CAATAGCAAGTGGTACAAGC (5472 to 5491); PR7, AAGCAGGACATAACAAGGTAGG (5029 to 5050); PR8, CCTATTTGCTATGTTGACACC (5365 to 5385); and PR9, GGTACACAGGCATGTGTGGCC (6009 to 6029). PCR products were cloned in a pGEM-T plasmid (Promega), and sequencing was performed with a T7 DNA polymerase kit (Pharmacia). For DNA purification, cells were resuspended in 0.5 ml of lysis buffer [25 mM tris-HCl (pH 8.0), 1% SDS, and 12 mM EDTA] and treated with proteinase K (100 μg/ml) for 1 hour at 37°C. After phenol extraction and ethanol precipitation, total DNA was resuspended in distilled water.
15. S. Wain-Hobson, P. Sonigo, O. Danos, S. Cole, M. Alizon, *Cell* **40**, 9 (1985).
16. L. Sharmeen, B. Bass, N. Sonenberg, H. Weintraub, M. Groudine, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8096 (1991).
17. In vitro transcription was performed with T7 RNA polymerase (Promega) on a Spe I–linearized plasmid, containing 356 base pairs (bp) of the *vif-vpr* region of HIV-1 Bru inserted either in sense or antisense orientation in the pGEM-T vector.
18. M. A. Martinez, M. Sala, J. P. Vartanian, S. Wain-Hobson, *Nucleic Acids Res.* **23**, 2573 (1995).
19. T. Kim, R. A. Mudry, C. A. Rexrode II, V. K. Pathak, *J. Virol.* **70**, 7594 (1996).
20. HIV-1 Bru virus suspensions were prepared as described by M. Ventura *et al.* [*Arch. Virol.* **144**, 513 (1999)]. Virions were resuspended in 1 ml of 4 M guanidine isothiocyanate, 25 mM sodium citrate (pH 7.1), and 2% β-mercaptoethanol and extracted with phenol (pH 4), isoamyl alcohol, and chloroform. Viral RNA was precipitated with ethanol. Control for contaminant DNA was carried out by omitting RT before PCR amplification. Southern blot analysis was performed on total DNA from HIV-1 Bru–infected cells. DNA was digested with Eco RV and Eco RI and hybridized with S1 (positions 2066 to 2471) and S2 (positions 5535 to 6029) probes as described by T. Sambrook, E. F. Fritsch, and T. Maniatis [*Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ed. 2, 1989)]. Both restriction enzymes allow the detection of the flanking regions of the virion linked to the insertion region on the cellular genome. Controls performed with total DNA from uninfected H9 cells gave no signal for both S1 and S2 probes.
21. Supplemental figures are available at www.sciencemag.org/feature/data/1049356.shl.
22. S. Schwartz, B. K. Felber, D. M. Benko, E. M. Fenyö, G. N. Pavlakis, *J. Virol.* **64**, 2519 (1990).
23. B. Teng, C. F. Burant, N. O. Davidson, *Science* **260**, 1816 (1993).
24. M. Kozak, *Cell* **44**, 283 (1986).
25. F. Del Gatto, M. C. Gesnel, R. Breathnach, *Nucleic Acids Res.* **24**, 2017 (1996).
26. L. Damier, L. Domenjoud, C. Branlant, *Biochem. Biophys. Res. Commun.* **237**, 182 (1997).
27. Y. Li, Z. Zhang, J. K. Wakefield, S. M Kang, C. D. Morrow, *J. Virol.* **71**, 6315 (1997).
28. V. Blanc, S. Litvak, A. Araya, *FEBS Lett.* **373**, 56 (1995).
29. R. C. Lord and G. J. Thomas Jr., *Dev. Appl. Spectrosc.* **6**, 179 (1968).
30. C. R. Cantor and P. R. Schimmel, Eds., *Biophysical Chemistry: The Conformation of Biological Macromolecules* (Freeman, San Francisco, 1980), part I, chap. 6.

31. We acknowledge L. Tarrago-Litvak, T. Astier, D. Bégu, and J. C. Farré for helpful discussions. We thank E. Sargos and B. Matusiak for excellent technical assistance. This work was supported by the French Ministère de l'Enseignement Supérieur et de la Recherche and the Conseil Régional d'Aquitaine. K.B. was supported by a predoctoral fellowship from SIDACTION.

31 January 2000; accepted 20 July 2000