

# Truncated variants of gp120 bind CD4 with high affinity and suggest a minimum CD4 binding region

Stuart R. Pollard<sup>1,2</sup>, Margaret D. Rosa<sup>1</sup>,  
Joseph J. Rosa<sup>1</sup> and Don C. Wiley<sup>2,3</sup>

<sup>1</sup>Biogen Inc., 14 Cambridge Center, Cambridge, MA 021242,  
<sup>2</sup>Department of Biochemistry and Molecular Biology and <sup>3</sup>Howard  
Hughes Medical Institute, Harvard University, 7 Divinity Avenue,  
Cambridge, MA 02138, USA

Communicated by J.J. Skehel

**The envelope glycoprotein, gp120, of human immunodeficiency virus type 1 (HIV-1) binds the cellular protein CD4 with high affinity. By deletion we show that 62 N- and 20 C-terminal residues along with the V1, V2 and V3 variable regions of gp 120 are unnecessary for CD4 binding. A 287 residue variant (ENV59), missing those 197 amino acids, binds to CD4 with high affinity. A polyclonal antibody failed to efficiently precipitate ENV59 which is consistent with the loss of immunodominant antigenic structures in the regions deleted. This suggests that ENV59 may have potential as an immunogen, able to elicit antibodies against more conserved regions of gp120. Additionally, complementing co-expressed gp120 fragments as well as a circularly permuted molecule bind CD4, and suggest either that the molecular termini are adjacent in the folded structure, or that an N-terminal region folds into the structure unconstrained by its method of attachment to the rest of the molecule.**

**Key words:** CD4 binding/complementation/deletion mutants/gp120/permutated variants

## Introduction

Human immunodeficiency virus type 1 (HIV-1) is the etiological agent of acquired immune deficiency syndrome (AIDS) (Barré-Sinoussi, *et al.*, 1983; Popovic *et al.*, 1984). The viral surface glycoprotein gp120, produced by cleavage of a precursor gp160 (Leis *et al.*, 1988), initiates infection by binding the cell surface receptor CD4 with high affinity. This interaction allows efficient infection of CD4-positive cells and directs a major cell tropism associated with HIV-1 infection (Dalglish *et al.*, 1985; Klatzmann *et al.*, 1985; McDougal *et al.*, 1986). Indeed, soluble recombinant forms of CD4 have been shown to neutralize virus *in vitro* and are being assessed as potential antivirals (Smith *et al.*, 1987; Deen *et al.*, 1988; Fisher *et al.*, 1988; Hussey *et al.*, 1988; Traunacker *et al.*, 1988).

Mutant gp120 molecules engineered *in vitro* indicate that CD4 binding is sensitive to mutations in the C-terminal half of gp120 (Kowalski *et al.*, 1987; Lasky *et al.*, 1987; Linsley *et al.*, 1988; Cordonnier *et al.*, 1989a,b; Olshevsky *et al.*, 1990). Additionally, the epitopes of monoclonal antibodies that specifically block CD4 binding map to the C-terminal C4 domain of gp120 (Lasky *et al.*, 1987; Sun *et al.*, 1989). Many other mutations, particularly in the N-terminal half

of the molecule, do not appear to affect CD4 binding. The lack of significant CD4 binding of mutants truncated at the N-terminus, however (Dowbenko *et al.*, 1988), suggests that the binding domain either includes N-terminal regions (Olshevsky *et al.*, 1990; Syu *et al.*, 1990), or that the N-terminus is required for the correct folding of the molecule. Further evidence that correctly folded gp120 is required for CD4 binding derives from the demonstration that several monoclonal antibodies that bind only to 'conformational epitopes' in gp120 block CD4 binding (Ho *et al.*, 1991; Posner *et al.*, 1991).

Here, we describe the CD4 binding properties of a number of deletion mutants of gp120 which were engineered *in vitro* and expressed transiently in COS7 cells. We show that 62 N- and 20 C-terminal residues along with V1, V2 and V3 variable sequence regions of gp120 are unnecessary for CD4 binding. ENV59, a 287 residue variant missing those 197 amino acids and which therefore combines all of the allowed deletions above, binds to CD4 with high affinity. The deleted regions include dominant antigenic structures. Consistent with this, ENV59 failed to precipitate with an anti-gp120 polyclonal antiserum. This result suggests that ENV59, which retains CD4 binding, may have potential as an immunogen. Additionally, complementing co-expressed gp120 fragments and a circularly permuted molecule can bind CD4. These results suggest that an N-terminal region, required for CD4 binding, folds into the structure unconstrained by its method of attachment to the rest of the molecule. The cyclic permutability may additionally, or alternatively, suggest that the molecular termini are adjacent in the folded structure.

## Results

### *N- and C-terminal residues are not required for CD4 binding*

Previous reports indicate that while a deletion of 30 amino acids from the mature N-terminus does not affect CD4 binding (Lasky *et al.*, 1987), deletions of 164 N-terminal and 44 C-terminal amino acids abolish binding (Dowbenko *et al.*, 1988; Linsley *et al.*, 1988, respectively). To determine the limits of N- and C-terminal truncations that are allowed before the expressed protein no longer binds CD4, 17 mutants were constructed using the plasmid pCAS and the polymerase chain reaction (see Materials and methods), and expressed in COS7 cells. Expression and CD4 binding (Table I) was then determined by co-precipitations of metabolically labeled protein from cell supernatants with a polyclonal antibody and with soluble CD4, respectively. Eighty-two amino acids, 62 from the N-terminus (construct N93/C491; Figure 1A) and 20 from the C-terminus (N82/C491; Figure 1B) of mature gp120 could be deleted without altering the observed CD4 binding. The removal of two additional amino acids at the N-terminus (N95/C491; Figure 1A) or one more amino acid at the C-terminus

(N82/C490; Figure 1B) results in a reduction of CD4 binding. Immune precipitations of the cell supernatants using a polyclonal anti-gp120 antibody (in excess) were quantitatively similar, and therefore suggested that there were similar amounts of the mutant proteins in each of the transfected cell supernatants (not shown). Even allowing for the possibility of variations in the antigenicity of the mutant proteins, the observed reduction in the relative amount of

protein in CD4 co-precipitates suggests loss of CD4 binding activity. The specificity of binding of the gp120 variants in the CD4 co-precipitations was verified by competition with excess soluble CD4. Non-specifically precipitated material was observed in all experiments, both in the stacking gel (hatched region in Figure 1) and in the resolving gel (~68 kDa protein). However, these materials were not competed away with soluble CD4 and also appeared in control precipitations from cell supernatants which were transfected with carrier DNA only. They were therefore considered artifactual and not relevant to the specific CD4 binding exhibited by the truncated gp120 variants.

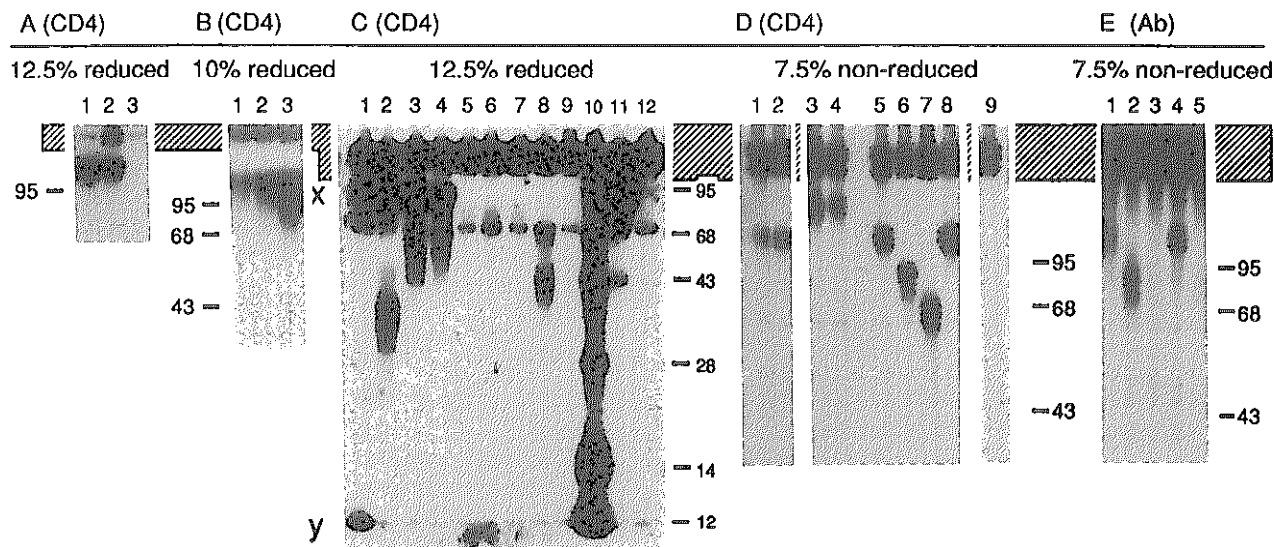
**Table 1.** CD4 binding of N- and C-terminally truncated gp120 mutants

Construct <sup>a</sup>	CD4 binding	Construct	CD4 binding
N60/C502	+	N173/C502	-
N82/C502	+	N173/C495	-
N60/C495	+	N100/C502	-
N82/C495	+	N60/C172	-
N60/C491	+	N60/C488	-
N88/C491	+	N82/C488	-
N91/C491	-	N82/C488	±
N93/C491	+	N82/C490	±
		N95/C491	-

<sup>a</sup>The number refers to the N- (N) or C-terminal (C) amino acid included in the mutant protein and corresponds to the HXBG2 HIV-1 gp120 sequence (Genbank) and includes the signal (Fisher *et al.*, 1985). The + or - refers to whether the mutant bound CD4 (+) or not (-). NB. Mature gp120 comprises 482 amino acids, i.e. residues 30-511.

#### Some gp120 fragments can complement each other

Although many terminally truncated variants of gp120 were not able to bind CD4 when expressed alone, we have observed that in some instances the co-expression of N- and C-terminal fragments results in molecules that bind CD4 (Figure 1C). We first observed this when an N-terminal fragment consisting of residues 60-172 and a C-terminal fragment representing residues 173-502 were co-expressed. Neither exhibits significant CD4 binding when expressed alone, but when co-expressed they complement each other and are both co-precipitated with CD4 (not shown). The C-terminal fragment, 173-502, corresponds to a V8 proteolytic fragment previously reported to retain indepen-



**Fig. 1.** Co-precipitations of gp120 variants. Shown are several examples of data for the results summarized in Figures 2-5. A-D and E are autoradiographs of SDS polyacrylamide gels of labeled proteins precipitated from cell supernatants by CD4 or polyclonal antibody, respectively. The percent acrylamide and whether proteins were reduced or not is indicated. The hatched region indicates non-specifically precipitated material which remains in the stacking gel. **Gel A:** N-terminal truncations. Lanes 1, 2 and 3 correspond, respectively, to constructions N91/C491 and N93/C491 which bind CD4, and N95/C491 which does not (numbers after N and C refer to N- and C-terminal gp120 amino acids in the expressed protein). **Gel B:** C-terminal truncations. Lanes 1, 2 and 3 correspond to constructions N82/C489 and N82/C490 which show reduced CD4 binding, and N88/C491 which binds CD4 as well as controls. **Gel C:** Co-expression experiments (summarized in Figure 2), e.g. lane 1 shows expt 'a' from Figure 2 demonstrating that fragments 119-502 and 60-118 (x and y in C, respectively) both precipitate with CD4 when co-expressed. Other experiments indicated are as follows (lane in Figure 1: experiments from Figure 2)—1: 1, 2: b. 3: c, 4: d, 5: f, 6: g, 7: h, 8: k, 9: n. Lane 10: molecular weight markers; lane 11: N30/C502 (a CD4 binding control) lane 12: no-DNA control. **Gel D:** Lanes 1-4, covalent association of complementing fragments. experiments (from Figure 2) p, a, c and d correspond to lanes 1, 2, 3 and 4, respectively. Non-reducing conditions show covalent association of the fragments in experiments p and a (lanes 1 and 2) where a protein with a molecular weight approximating to gp120 results (lane 8; N30/C502). For experiments c and d (lanes 3 and 4, a protein of higher molecular weight than gp120 results (lane 8) because of the sequence overlap in the fragments (Figure 2c and d). Lanes 5-9: CD4 precipitations of lane 5: HTT.1 (a circularly permuted molecule, Figure 4); lane 6: HTT.1.V1/V2 (a deletion mutant of HTT.1, Figure 4); lane 7: ENV59 (Figures 4 and 5); lane 8: N30/C502 (CD4 binding control); lane 9: no-DNA control. ENV59 (lane 7) binds CD4 as well as the gp120 (lane 8). **Gel E:** Immune precipitations of lane 1: HTT.1; lane 2: HTT.1.V1/V2; lane 3: ENV59; lane 4: N30/C502; lane 5: no-DNA control. This gel was overexposed to illustrate the poor immune precipitation of ENV59 by the polyclonal antibody.

dent CD4 binding (Nygren *et al.*, 1988). Explanations for the apparent discrepancy with the data of Nygren *et al.* which include evidence that suggests the V8 proteolytic fragment does not in fact retain independent CD4 binding, have been presented elsewhere (Pollard *et al.*, 1991). However, the complementation that occurs between this fragment (residues 173–502) and an N-terminal fragment (residues 60–172) does suggest that N-terminal residues are at least required for the correct folding of gp120, if not directly for CD4 binding itself. Figure 2 summarizes a series of other co-expression experiments using other truncated variants which do not bind to CD4 when expressed alone. Some examples of these experiments are shown in Figure 1C [complementation with purified protein fragments has been observed previously with ribonuclease S and staphylococcal nuclease (Anfinsen, 1973)]. In all cases where complementation was observed, comparison of reducing and non-reducing SDS gels indicated that the complementing fragments always became covalently associated via disulfide bonds (see Figure 1D, lanes 1–4). The disulfide structure of gp120 (Leonard

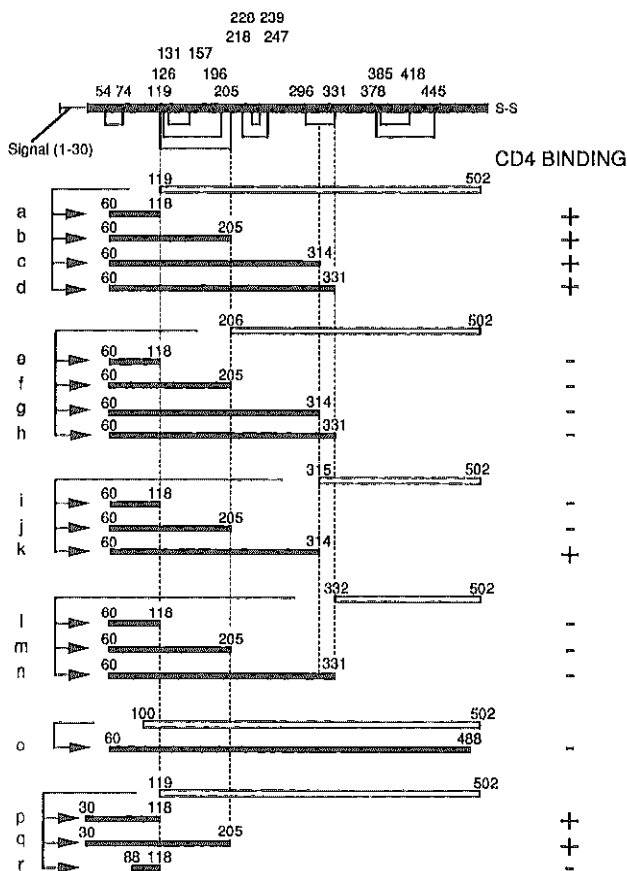


Fig. 2. Co-expression of N- and C-terminal fragments of gp120 sometimes results in complementation. At the top is a diagrammatic representation of the disulfide structure of gp120 (Leonard *et al.*, 1990). Disulfide pairs are indicated below the line. Numbers at the top refer to the 18 cysteines in gp120. Below, indicated diagrammatically, are a series of co-expression experiments performed with a set of N- and C-terminal constructions which produce proteins unable to bind CD4 when expressed alone. Each experiment involved the co-expression of two fragments which are represented by an open and a closed bar. The numbers refer to the terminal gp120 amino acid which is included in the expressed protein. Dashed lines serve to illustrate the relative positions of cysteine residues. The left-hand column indicates CD4 binding: +, co-precipitates with CD4; no strong precipitable band observed.

*et al.*, 1990) indicates that for this to occur in most cases in Figure 2, at least one incorrect disulfide must be made between the complementing fragments relative to the normal gp120 molecule and not affect CD4 binding. In the cases tested, complementation was not observed if the sum of the two fragments contained an internal 'deletion' of gp120 amino acids. Complementation also appears to be sensitive to the relative position of the N- and C-terminus of the two fragments. The failure of fragments 60–331 and 332–502 to complement contrasts with the complementation observed for fragments 50–314 and 315–501 (Figure 2n and k). Similarly fragment 206–502 fails to complement with any of the N-terminal fragments able to complement fragment 119–502, even when no internal 'deletions' would result (cf. Figure 2e–h to 2a–d). The reasons for these different results are unclear, but may be related to the particular cysteines present in each fragment. For example, if covalent association via disulfides is required for complementation to occur, it is possible that for some fragments (i.e. those which fail to complement) intramolecular disulfide bonds are more favorable than intermolecular bonds. Comparison of CD4 precipitations with precipitations with polyclonal antibody gives some indication of the proportion of active CD4 binding molecules (CD4 precipitation) in the total pool (antibody precipitation). When compared to an intact gp120 control, such comparisons suggest that complementation is a relatively efficient process. In the cases looked at, the active CD4 binding molecules constituted the major proportion (~50%) of the molecules expressed. When complementation was observed, non-reducing gels of the material precipitated by the polyclonal antibody indicated that not all fragments become covalently associated via disulfide bonds. These non-disulfide-linked fragments constituted <50% of the molecules expressed and did not bind CD4.

All of the N-terminal fragments used in the initial experiments began at residue 60 and therefore contained an unpaired cysteine at position 74. Unlike fragment 60–118 (Figure 2a), fragment 88–118 (Figure 2r) did not complement fragment 119–502 despite containing all the gp120 residues required for CD4 binding (see above). These results suggest that, in these cases, the complementing region occurs between residues 60 and 88. As no complementation has been observed without covalent disulfide association, the presence of cysteine 74 between residues 60 and 88 suggests, but does not prove it to be the complementing factor in the region. In these experiments, we were unable to detect any protein of appropriate size by immune precipitation that would correspond to that expected to be produced by the construct N88/C118. However, neither of two independent clones of N88/C118 was able to complement the C-terminal fragment N119/C502, and nucleotide sequence analysis confirmed that the constructions were correct. It therefore seems reasonable to assume that this fragment is expressed and suggests that there are no antibodies in the polyclonal antibody to this small region of the gp120 molecule which can precipitate the expressed protein.

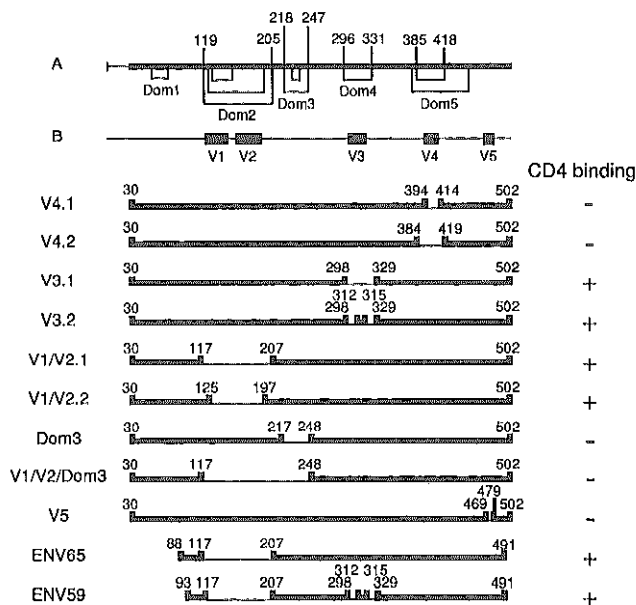
Complementation with formation of inter-fragment disulfides was also observed between fragment 119–502 and the N-terminal fragments 30–118 and 30–205, neither of which contain an odd cysteine (i.e. they contain the 54 and 74 disulfide pair; Figure 2p and q). This indicates that even when all normal pairs of cysteines are present in each fragment, a novel disulfide can arise between fragments, producing a molecule with CD4 binding activity.

### CD4 binding can tolerate some internal deletions of gp120 sequences

In the covalently linked complementing fragments described above which contain novel disulfide bonds, the local structure around the incorrect disulfide must be quite different from that in the normal gp120 molecule. Furthermore, when two fragments with large sequence overlap become associated covalently, such as 60–331 and 119–502 (Figure 2d), large structurally aberrant regions must exist and yet do not destroy CD4 binding. It therefore seemed reasonable to investigate internal deletions of the linear sequence of gp120 that correspond to the likely location of these aberrant regions.

A series of deletion mutants were constructed using the polymerase chain reaction (PCR) technique referred to as 'splicing by overlap extension' (Horton *et al.*, 1989). In general, the mutants were engineered so as not to leave unpaired cysteines (Leonard *et al.*, 1990) and to delete sequences corresponding to regions of hypervariable sequence (Modrow *et al.*, 1987), i.e. the V1, V2, V3, V4 and V5 regions or, in one case, a region within domain 3 of the disulfide structure (Leonard *et al.*, 1990) to which antibodies bind but do not block CD4 binding (Ho *et al.*, 1988) (see Figure 3). Sometimes the deletions removed paired cysteines in a region as for deletions V4.2, V1/V2.1, V1/V2.2, Dom3 and V1/V2/Dom3. Other deletions removed the sequences bounded by a disulfide, but retained the cysteines, e.g. V4.1, V3.1 and V3.2.

All deletions reported here which remove regions V4, V5 and the disulfide domain 3 destroyed CD4 binding. However, deletions of the entire V1/V2 region (deletions V1/V2.1 and V1/V2.2) and the V3 region (V3.1 and V3.2)



**Fig. 3.** Deletion mutants of gp120. **Bar A** shows the disulfide structure of gp120 (Leonard *et al.*, 1990). Cysteines relevant to deleted regions are numbered. The disulfide domains are indicated, Dom1–5. **Bar B** shows the position of the five hypervariable regions of gp120 as filled boxes (Modrow *et al.*, 1987). Below are deletion mutants constructed and expressed in COS-7 cells. Numbers refer to the terminal gp120 amino acids which are included in the expressed protein. Thin lines indicate deleted regions and thick lines refer to amino acids included in the mutant protein. The right-hand column indicates CD4 binding: +, specific co-precipitation with CD4; -, no strong precipitable band observed.

did not result in any observable loss in CD4 binding. The immune precipitation for deletion V3.1, which binds CD4, indicated that it was susceptible to degradation. However, deletion V3.2 which is similar but contains an arginine-proline-glycine-arginine sequence between the cysteines, binds CD4 and is stable (not shown). The original construction intended for the four residue sequence to be glycine-proline-glycine-arginine, a possible  $\beta$ -turn motif, and which corresponds to a conserved sequence in the V3 loop (LaRosa *et al.*, 1990). However, nucleotide sequence analysis of V3.2 revealed a missense mutation, probably resulting from the PCR manipulation, which changes the first glycine or arginine. This sequence, however, retains the potential to form a  $\beta$ -turn structure, suggesting this stabilizes the protein and that the substitution is inconsequential, at least for CD4 binding. The mutants (V1/V2.1 and V1/V2.2) which remove the V1/V2 regions, but still bind CD4, occur in the area where sequence overlap was tolerated in the complementation experiments above, and are probably therefore regions where aberrant structures can be accommodated in the molecule without inhibiting CD4 binding.

The N-terminal, C-terminal and internal deletions that did not affect CD4 binding were then combined to produce a molecule, ENV59 (Figure 3). This variant protein, comprising only 59% of gp120 sequences, is missing the N- and C-terminal truncations of 62 and 20 residues, respectively, the V1, V2 and V3 deletions, four of the nine disulfide bonds and eight of the 24 N-linked glycosylation sites. The mutant ENV59 is efficiently precipitated with CD4 (Figure 1D, lane 7) but was not efficiently precipitated by the polyclonal antibody that was used (Figure 1E, lane 3). Although the polyclonal antibody used was raised against a disparately glycosylated gp120 produced in insect cells, the weak immune precipitation is consistent with the loss in ENV59 of dominant antigenic structures in the N- and C-terminal regions, and in the V1, V2 and V3 variable regions of the molecule.

### ENV59 binds to CD4 with high affinity

As a consequence of the inability to precipitate the ENV59 mutant with the polyclonal antibody, we are unable to determine the relative expression of ENV59 to normal gp120. The various gp120 mutants described here, of the size of ENV59 or larger, do not exhibit significantly different expression levels (see Figure 1E). Therefore we reasoned that unless the mutant ENV59 construct expresses at significantly higher levels than the wild-type control, the similar quantities of protein precipitated by CD4 for both (Figure 1D, lanes 7 and 8), suggests that ENV59 binds CD4 with an affinity similar to that of wild-type gp120. However, to ascertain the relative binding affinity of ENV59 more definitively, a series of CD4 precipitations were tried, for both ENV59 and a wild-type gp120 control (construct N30/C502), using a range of CD4 concentrations (200 pM–200 nM). In these experiments, the CD4 co-precipitations were achieved using free soluble CD4 and an anti-CD4 monoclonal antibody which is not cross-reactive with the gp120 binding site and can precipitate gp120/CD4 complexes. This method was employed as a wider range of CD4 concentrations could be tested than was practically possible with the CD4–Sepharose beads. Co-precipitated material was resolved by SDS–PAGE and the resultant autoradiographs analyzed by scanning densitometry. The

densitometric analyses of these titrated CD4 co-precipitations resulted in the binding curves shown in Figure 4. Both curves clearly plateau, indicating that all the protein able to bind CD4 has co-precipitated. After correction for the expected differences in the specific activities of ENV59 relative to wild-type gp120, the curves essentially superimpose, indicating that the expression levels are similar, as expected above. In addition, half-maximal binding for both occurs in the low nanomolar range, which is in agreement with previous data (4 nM; Lasky *et al.*, 1987) and demonstrates that when all the independently allowed deletions are

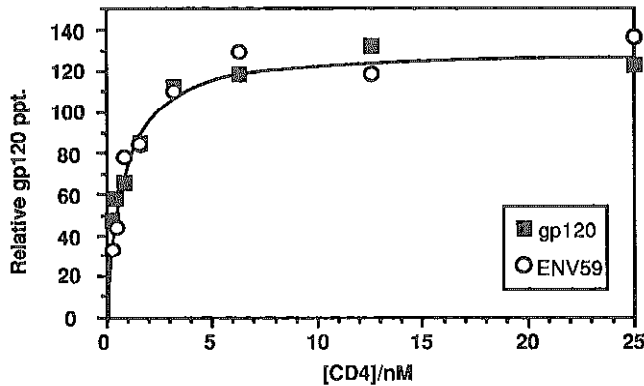


Fig. 4. CD4 binding curves for the mutant ENV59 and wild-type gp120 (residues 30–502). Co-precipitations from metabolically labeled cell supernatants were performed with soluble CD4 at different CD4 concentrations (200 nM–200 pM). The gp120/ENV59–CD4 complexes were precipitated with a CD4 monoclonal antibody (MAbE7) and protein A–Sepharose beads. Precipitated material was resolved on an SDS polyacrylamide gel and the autoradiograph was analyzed by scanning densitometry. Densitometric read-outs were corrected for the expected difference in the specific activity between gp120 (■) and ENV59 (○) (ENV59 values multiplied by 1.8). At 200 nM CD4, the densitometric readings of precipitated bands indicated that for both the mutant ENV59 and the control gp120 a plateau of maximal precipitation had been achieved (the experiment is carried out in antibody excess so that in the region of this plateau the precipitating agents CD4 and antibodies/protein A are in excess). Therefore, half-maximal binding for ENV59 and wild-type gp120 is in the low nanomolar range, which is consistent with previous results (Lasky *et al.*, 1987).

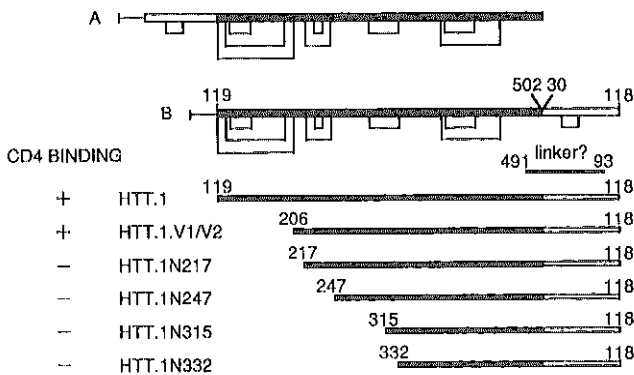


Fig. 5. Circularly permuted gp120 molecules. A shows the normal disulfide structure of gp120 as in Figure 2. B shows the presumed disulfide structure of a circularly permuted molecule where N-terminal residues 30–118 are moved to the C-terminus adjacent to residue 502. The left-hand column indicates CD4 binding: +, co-precipitates with CD4; -, no strong precipitable band observed. The names of the constructions are referred to on the left. Numbers refer to the terminal gp120 amino acids which are included in the expressed protein.

combined in ENV59, there appears to be little or no loss of CD4 binding affinity.

#### Circularly permuted variant binds CD4

The 25 amino acid sequence between residues 93 and 117, at least some of which the deletion experiments indicated to be required for CD4 binding (see Figure 6), appears able to tolerate being linked to the rest of gp120 in several ways and still produce a variant with CD4 binding activity. The complementation experiments indicate that it may be linked by an incorrect disulfide involving cysteine 74 (Figure 2A, e.g. fragment 60–118 and 119–502), while the deletion

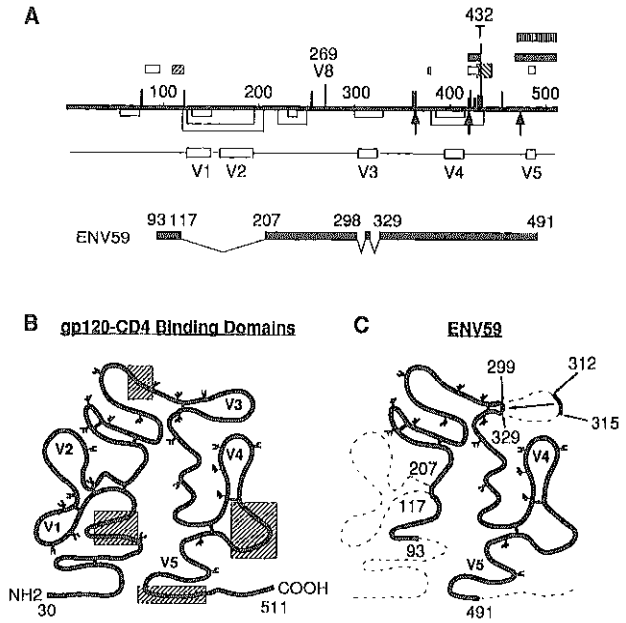


Fig. 6. CD4 binding regions of gp120. (A) shows the disulfide structure of gp120 (Leonard *et al.*, 1990). Numbers above line A are amino acid number, unlabeled vertical lines indicate point mutations shown to affect CD4 binding. ▮ (Olshevksy *et al.*, 1990); ▩ (Cordonnier *et al.*, 1989a); ▨ (Lasky *et al.*, 1987). Boxes above line A refer to deletions which destroy CD4 binding. □ (Cordonnier *et al.*, 1989b); ▤ (Syu *et al.*, 1990); ▥ (Lasky *et al.*, 1987); ▦ (Cordonnier *et al.*, 1989a); ▧ (Linsley *et al.*, 1988); ▨ (Kowalski *et al.*, 1987). Arrows below the line indicate insertion mutations which affect CD4 binding (Kowalski *et al.*, 1987). V8-269 and T-432 refer to the V8 protease and trypsin sites which are in regions important for CD4 binding (Pollard *et al.*, 1991). The line below shows the hypervariable domains (open boxes). ENV59 is represented as a linear thick line with the deletions represented by thin lines. The sequences retained appear consistent with mutant gp120 molecules previously reported. (B) The disulfide structure of gp120 is shown as diagrammed by Leonard *et al.* (1990). The amino acid backbone is represented as a thick black line with the N- and C-termini indicated, along with the residue number (i.e. 30 and 511, respectively). The disulfides are illustrated with crossbars (nine in total) and the 24 N-linked glycosylation sites are represented with the symbols ♣ or ♠. The locations of the five hypervariable sequence domains are indicated (V1–V5). The hatched boxes highlight regions which site-directed mutants indicate to be important for CD4 binding (Kowalski *et al.*, 1987; Lasky *et al.*, 1987; Linsley *et al.*, 1988; Cordonnier *et al.*, 1989a, b; Olshevksy *et al.*, 1990; Syu *et al.*, 1990). (C) The deletion mutant ENV59. The residues retained in the mutant ENV59 are shown as a thick black line in the context of the disulfide structure shown in A. Deleted sequences are indicated with a thin dotted line contiguous with the thick line. The variable sequence domains which are retained (V4 and V5), as well as the 16 N-linked glycosylation sites also retained in ENV59, are indicated. The numbers refer to the deleted sequences and are the residues retained in ENV59.

mutant V1/V2.1 demonstrates that it can be placed directly adjacent to residue 207 and still bind CD4 (Figure 3). To investigate further the independence of these 25 amino acids with respect to the rest of the structure, we genetically fused them to the C-terminus of gp120 rather than their usual place at the N-terminus. This circularly permuted, head-to-tail gp120 (HTT.1) is able to bind CD4 (Figure 1D, lane 5). The permuted molecule (HTT.1) corresponds to a gp120 where the N-terminal residues 30–118 are placed at the C-terminus adjacent to residue 502 (Figure 5). It is possible that two such molecules could complement each other to create CD4 binding in a way analogous to that observed in the co-expression of fragments 30–118 and 119–502 (see Figure 2p). However, this is unlikely as the HTT.1 molecule does not form the covalent dimers expected if it were analogous to the complementation experiments above (Figure 1D, lane 5). Additionally, an N-terminal deletion of this molecule (HTT.1.V1/V2), akin to the deletion V1/V2.1 of Figure 3, results in another circularly permuted molecule that binds CD4 (Figure 4 and Figure 1D, lane 6). This deletion of the N-terminus of HTT.1 up to residue 206 further suggests this not to be a complementation phenomenon because fragment 206–502 could not be complemented by any N-terminal fragment (see Figure 2). Additionally, this molecule is missing the full set of gp120 amino acids which appear to be a prerequisite for the complementation observed above. Deletion of N-terminal residues beyond 206 in HTT.1 abolished CD4 binding (Figure 4), a result that appears consistent with the inability to delete those residues from gp120, as described above (see Figure 3). The terminal deletion results presented above (Figure 1, footnote 18) would suggest that residues 491–502 and 30–93 in HTT.1 are not required directly for CD4 binding, but may be functioning as a spacer or linker (see Figure 4).

## Discussion

The results of the deletion, complementation and other expression experiments presented here indicate that 41% of gp120 residues, i.e. 197 amino acids which include four disulfide bonds and one-third of the N-linked glycosylation sites, are not required for CD4 binding. Consistent with these results, some of the regions identified here as not required for binding have been implicated in other functions of gp120. For example, monoclonal antibodies to the V3 variable region have been shown to inhibit membrane fusion and not to affect CD4 binding (Linsley *et al.*, 1988; Skinner *et al.*, 1988). Additionally, site-directed mutants have indicated that both N- and C-terminal residues of gp120 are important for gp41 association, but not for CD4 binding (Helseth *et al.*, 1991; Ivey-Hoyle *et al.*, 1991).

In every one of the 54 variant gp120 molecules presented, a soluble protein could be precipitated (with either antibody or CD4) from the supernatant of the expressing cell, suggesting that the protein had been secreted and had presumably lost its signal sequence. In the positive cases of CD4 binding, specificity of binding to CD4 was shown by competition with excess soluble CD4. Binding data from mutant gp120 molecules, previously reported, which have suggested regions of gp120 important for CD4 binding appear consistent with the sequences which remain in the CD4 binding variants presented here (summarized in Figure

6a,b). Regions previously identified to be required for binding are also included in the 287 residue variant molecule, ENV59, which retains high-affinity CD4 binding activity (Figure 6). A series of co-precipitations performed at a range of CD4 concentrations, for both ENV59 and gp120, indicate that for both, half-maximal binding occurs in the low nanomolar range. These values are consistent with published data (Lasky *et al.*, 1987) and therefore indicate that ENV59 binds CD4 with an affinity at least approaching that of wild-type gp120. Furthermore, complementation, deletion and, remarkably, cyclic permutation of the sequence indicate that a short  $\leq 25$  amino acid (93–117) N-terminal sequence apparently required for CD4 binding, or perhaps for correct gp120 folding, appears to be critical, somewhat independent of its method of linkage to the rest of the molecule. The circular permutability may indicate that the N- and C-termini of gp120 are in close proximity to each other in the folded molecule. [The proximity of N- and C-termini enabled functional circularly permuted variants of bovine pancreatic trypsin inhibitor (BPTI) and ribosyl anthranilate isomerase to be produced (Goldenberg and Creighton, 1983; Luger *et al.*, 1989)]. The failure of the polyclonal antibody used here to precipitate ENV59 is consistent with the absence of the antigenic structures in the N- and C-termini, and in the V1, V2 and V3 regions. The use of ENV59 as an immunogen may therefore elicit the production of antibodies against more conserved regions of gp120 which include the CD4 binding domain. Such antigenicity, if exhibited, might be of therapeutic or prophylactic value in the treatment or prevention of HIV-1 infection.

## Materials and methods

### Plasmid construction

Genetic truncations were engineered using the PCR and introduced into a plasmid pCAS.ENV. Plasmid pCAS.ENV is derived from plasmid pBG381 (Fisher *et al.*, 1988) and contains a gene encoding gp120 from the HXB2 strain of HIV-1 [HXBCG2, Genbank (Fisher *et al.*, 1985)], flanked by unique restriction sites *Mlu*I and *Bsr*II. The gene is engineered to have the CD4 signal sequence at the N-terminus of expressed protein (Fisher *et al.*, 1988). The method of construction results in a stop codon at the C-terminus of the gp120 specific coding sequences. In addition, all mutant proteins expressed using this vector contain the signal sequence and two additional amino acids (Arg Thr; corresponds to the *Mlu*I site) at the N-terminus of the mature secreted protein, in addition to the gp120 sequence. Internal deletion mutants were constructed using the PCR technique referred to as 'splicing by overlap extension' (Horton *et al.*, 1989). In all constructs reported here, cleavage of the signal peptide was demonstrated by immune precipitation of the secreted gp120 variant proteins from the cell supernatant. The plasmid also includes the 3' RNA processing site from the hepatitis B virus genome which enables *rev*-independent expression of gp120 (Emerman *et al.*, 1989). High template concentrations and limited cycles of amplification were used (10 cycles) for all constructions using PCR. This was done to limit the chances of errors caused by the *taq* polymerase (Horton *et al.*, 1989). Independent duplicate clones were also tested for verification of some results.

### Expression of truncated gp120 mutants

All experiments were performed in COS-7 cells in transient expression assays. Plasmid DNA was introduced into the cells by electroporation. Typically, 20  $\mu$ g each of supercoiled plasmid was used with 380  $\mu$ g of salmon sperm carrier DNA. Electroporation was performed using a Biorad gene pulser<sup>TM</sup> with a capacitance extender at 960 mFD, 280 V. The cells were seeded into 6-well dishes to give a confluent monolayer 48 h post-electroporation. Cell proteins were metabolically labeled for 4 h, 48 h post-electroporation, using 1 ml culture medium (DMEM) without methionine, supplemented with [<sup>35</sup>S]methionine at 100 mCi/ml (NEN, DuPont).

### Immune and CD4 co-precipitations

Expression was determined by immune precipitations of material in transfected cell supernatants with a rabbit polyclonal antibody against gp120

and protein A-Sepharose beads (Sigma). The rabbit polyclonal antibody was raised against a recombinant gp120 (HIV-1 strain HXB2), affinity purified from *Spodoptera frugiperda* insect cells infected with a recombinant baculovirus (*Autographa californica*). CD4 binding was determined by precipitation with a soluble form of CD4 (Fisher *et al.*, 1988) cross-linked to Sepharose beads, made using methods described by the manufacturers (Sigma). For each precipitation, 10 µl of a 50% suspension of beads was used. Precipitations were incubated overnight at 4°C on a rocking platform. After three washes with 1 ml phosphate-buffered saline (PBS), bound material was eluted from the beads using 2× gel loading buffer and resolved by SDS-PAGE. The specificity of CD4 precipitations was verified by competition with excess free soluble CD4.

#### Relative affinity measurements

COS7 cells transfected as above with constructs ENV59 or N30/C502 (a wild-type gp120 control) were seeded into T150 flasks to be confluent 48 h post-electroporation. Expressed proteins were metabolically labeled for 4 h in 10 ml culture medium (DMEM) without cysteine, supplemented with [<sup>35</sup>S]cysteine at 100 mCi/ml (NEN, DuPont). Aliquots (0.5 ml) of labeled transfected cell culture medium were used in a series of CD4 precipitations, using a range of CD4 concentrations (200 nM–200 pM), for both ENV59 and a wild-type gp120 control (construct N30/C502). For these experiments, soluble CD4 rather than CD4 Sepharose beads was used. The gp120/ENV59-CD4 complexes were precipitated with a CD4 monoclonal antibody (MAb3-7; kindly supplied by Biogen Inc.) and protein A-Sepharose beads. The binding site of the monoclonal antibody MAb3-7 maps to CD4 domains 3 and 4, and is therefore not cross-reactive with the gp120 binding site and does not inhibit gp120 binding. This method was employed, rather than the CD4-Sepharose beads, as a wider range of CD4 concentrations could be tested than was practically possible with the beads. Co-precipitated material was resolved by SDS-PAGE and the resultant autoradiographs analyzed using a scanning densitometer (Molecular Dynamics, ImageQuant™ version 3.0). The densitometric read-outs for the wild-type gp120 control relative to the mutant ENV59 were corrected for the expected difference in specific activity, i.e. 18 cysteines in gp120 compared to only 10 cysteines in the mutant ENV59.

#### Acknowledgements

The gp120 HXB2 gp120 cDNA was a gift from Drs R.C. Gallo and F. Wong-Staal. We wish to thank Gary Jaworski for oligonucleotide synthesis and purification, and Vicki Sato for allowing S.R. Pollard to carry out these experiments at Biogen, Inc. S.R. Pollard is supported by NIH grant # GM 39589 to D.C. Wiley.

#### References

- Anfinsen, C.B. (1973) *Science*, **181**, 223–230.  
 Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler, B.C., Vezinet, B.F., Rouzioux, C., Rozenbaum, W. and Montagnier, L. (1983) *Science*, **220**, 868–871.  
 Cordonnier, A., Montagnier, L. and Emerman, M. (1989a) *Nature*, **340**, 571–574.  
 Cordonnier, A., Rivière, Y., Montagnier, L. and Emerman, M. (1989b) *J. Virol.*, **63**, 4464–4468.  
 Dalglish, A.G., Beverley, P.C.L., Clapham, P.R., Crawford, D.H., Greaves, M.F. and Weiss, R.A. (1985) *Nature*, **312**, 763–767.  
 Deen, K.C., McDougal, J.S., Inacker, R., Folena, W.G., Arthos, J., Rosenberg, J., Maddon, P.J., Axel, R. and Sweet, R.W. (1988) *Nature*, **331**, 82–84.  
 Dowbenko, D., Nakamura, G., Fennie, C., Shimasaki, C., Riddle, L., Harris, R., Gregory, T. and Lasky, L. (1988) *J. Virol.*, **62**, 4703–4711.  
 Emerman, M., Vazeux, R. and Peden, K. (1989) *Cell*, **57**, 1155–1165.  
 Fisher, A.G., Collati, E., Ratner, L., Gallo, R.C. and Wong-Staal, F. (1985) *Nature*, **316**, 262–265.  
 Fisher, R.A., Bertonis, J.M., Meier, W., Johnson, V.A., Costopoulos, D.S., Liu, T., Tizard, T., Walker, B.D., Hirsch, M.S., Schooley, R.T. and Flavel, R.A. (1988) *Nature*, **331**, 76–78.  
 Goldenberg, D.P. and Creighton, T.E. (1983) *J. Mol. Biol.*, **165**, 407–413.  
 Helseth, E., Olshevsky, U., Furman, C. and Sodroski, J. (1991) *J. Virol.*, **65**, 2119–2123.  
 Ho, D.D., Kaplan, J.C., Rackauskas, I.E. and Gurney, M.E. (1988) *Science*, **239**, 1021–1023.  
 Ho, D.D., McKeating, J.A., Li, X.-L., Moudgil, T., Daar, E.S., Sun, N.-S. and Robinson, J.E. (1991) *J. Virol.*, **65**, 489–493.  
 Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K. and Pease, L.R. (1989) *Gene*, **77**, 61–68.

- Hussey, R.E., Richardson, N.E., Kowalski, M., Brown, N.R., Chang, H.C., Siliciano, R.F., Dorfman, T., Walker, B., Sodroski, J. and Reinherz, E.L. (1988) *Nature*, **331**, 78–81.  
 Ivey-Hoyle, M., Clark, R.K. and Rosenberg, M. (1991) *J. Virol.*, **65**, 2682–2685.  
 Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.C. and Montagnier, L. (1985) *Nature*, **312**, 767–768.  
 Kowalski, M., Potz, J., Basiripour, L., Dorfman, T., Goh, W.C., Terwilliger, E., Dayton, A., Rosen, C., Haseltine, W. and Sodroski, J. (1987) *Science*, **237**, 1351–1355.  
 LaRosa, L.G., Davide, J.P., Weinhold, K., Waterbury, J.A., Profy, A.T., Lewis, J.A., Langlois, A.J., Dreesman, G.R., Boswell, R.N. and Shaduck, P. (1990) *Science*, **249**, 932–935.  
 Lasky, L.A., Nakamura, G., Smith, D.H., Fennie, C., Shimasaki, C., Patzer, E., Berman, P., Gregory, T. and Capon, D.J. (1987) *Cell*, **50**, 975–985.  
 Leis, J., Baltimore, D., Bishop, J.M., Coffin, J., Fleissner, E., Goff, S.P., Oroszlan, S., Robinson, H., Skalka, A.M., Temin, H.M. and Vogt, V. (1988) *J. Virol.*, **62**, 1808–1809.  
 Leonard, C.K., Spellman, M.W., Riddle, L., Harris, R.J., Thomas, J.N. and Gregory, T.J. (1990) *J. Biol. Chem.*, **265**, 10373–10382.  
 Linsley, P.S., Ledbetter, J.A., Kinney, T.E. and Hu, S.L. (1988) *J. Virol.*, **62**, 3695–3702.  
 Luger, K., Hommel, U., Herold, M., Hofsteenge, J. and Kirschner, K. (1989) *Science*, **243**, 206–210.  
 McDougal, J.S., Kennedy, M.S., Sleigh, J.M., Cort, S.P., Mawle, A. and Nicholson, J.K. (1986) *Science*, **231**, 382–385.  
 Modrow, S., Hahn, B.H., Shaw, G.M., Gallo, R.C., Wong, S.F. and Wolf, H. (1987) *J. Virol.*, **61**, 570–578.  
 Nygren, A., Bergman, T., Matthews, T., Jornvall, H. and Wigzell, H. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6543–6546.  
 Olshevsky, U., Helseth, E., Furman, C., Li, J., Haseltine, W. and Sodroski, J. (1990) *J. Virol.*, **64**, 5701–5707.  
 Pollard, S.R., Meier, W., Chow, P., Rosa, J.J. and Wiley, D.C. (1991) *Proc. Natl. Acad. Sci. USA*, in press.  
 Popovic, M., Sarngadharan, M.G., Read, E. and Gallo, R.C. (1984) *Science*, **224**, 497–500.  
 Posner, M.R., Hideshima, T., Cannon, T., Mukherjee, M., Mayer, K.H. and Byrn, R.A. (1991) *J. Immunol.*, **146**, 4325–4332.  
 Skinner, M.A., Langlois, A.J., McDanal, C.B., McDougal, J.S., Bolognesi, D.P. and Matthews, T.J. (1988) *J. Virol.*, **62**, 4195–4200.  
 Smith, D.H., Byrn, R.A., Marsters, S.A., Gregory, T., Groopman, J.E. and Capon, D.J. (1987) *Science*, **238**, 1704–1707.  
 Sun, N.-C., Ho, D.D., Sun, C.R.Y., Liou, R.-S., Gordon, W., Fung, M.S.C., Li, X.-L., Ting, R.C., Lee, T.-H., Chang, N.T. and Chang, T.-W. (1989) *J. Virol.*, **63**, 3579–3585.  
 Syu, W.J., Huang, J.H., Essex, M. and Lee, T.H. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 3695–3699.  
 Trautnecker, A., Luke, W. and Karjalainen, K. (1988) *Nature*, **331**, 84–86.

Received on October 4, 1991; revised on November 18, 1991

