Vol. 191, No. 1, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
February 26, 1993 Pages 207-213

EXPRESSION AND ANTIGENECITY OF HUMAN IMMUNODEFICIENCY VIRUS TYPE-1 TRANSMEMBRANE PROTEIN GP41 IN INSECT CELLS

Sung-Liang Yu[1,2], Min-Ji Chou[2], Ming F. Tam[3], Tun-Hou Lee[4], and Wan-Jr Syu[1,*]

[1]Graduate Institute of Microbiology & Immunology, Natl. Yang-Ming Medical College, Taipei, Taiwan

[2]Dept. Medical Technology, Natl. Taiwan University, Taipei, Taiwan

[3]Institute of Molecular Biology, Academica Sinica, Taipei, Taiwan

[4]Department of Cancer Biology, Harvard School of Public Health, Boston

Received January 6, 1993

SUMMARY: The HIV-1 transmembrane protein, gp41, is processed together with the envelope glycoprotein, gp120, from the same precursor, gp160, during the virus maturation. We used a baculovirus expression system to demonstrate that gp41 could be properly expressed without the preceding gp120 sequence. Two constructs with slight differences in the N-terminal region of gp41 were generated: one with a deletion of the first 7 hydrophobic residues of gp41, which have been suggested to be in a region important for membrane fusion and penetration, whereas the second with a complete sequence of gp41 except that a nonconserved leucine was substituted with a glutamine during DNA manipulation. Results from Western blotting with specific antisera confirm the gp41 identity. The sizes of gp41 were sensitive to tunicamycin treatment, indicating that N-linked glycosylation did occur. Further immunoblotting analyses with 90 different serum samples from HIV-1-infected individuals gave similar reaction patterns, suggesting that gp120 as well as the N-terminal region of gp41 are not critical for the expression and antigenecity of gp41. These eucaryotic constructs should provide valuable gp41 sources for detailed characterization of gp41 functions.

Two envelope components of human immunodeficiency virus type 1(HIV-1), glycoproteins gp120 and gp41, are essential for the virus and play major roles in the viral life cycle, particularly during the early stages of infection [1-3]. The coat protein gp120 determines the interaction of virus with CD4 receptors on major targeted T helpers whereas the transmembrane protein gp41 may mediate the immediate post binding events. Both glycoproteins are encoded by the viral *env* gene and proteolytically processed from the same precursor protein, gp160, during transport to the cell surface where viral particles are assembled. After maturation, gp120 remains

*To whom correspondence should be addressed.

0006-291X/93 $4.00

bound to viral membrane via a noncovalent interaction with gp41. This gp120-gp41 interaction appears to be weak so that some gp120 molecules are shed and detected in media [4, 5]. In the cases of truncation expression without gp41, stop codons introduced before the gp41 transmembrane region have rendered *env*-related products secreted directly into media [6, 7]. It is also known that these gp120 variants so obtained retain antigenecity and CD4 binding activity [8, 9]. However, little information is available for gp41 independently expressed in eucaryotic systems.

Many different proteins have been expressed from insect cells with functions equivalent to their nature counterparts [10]. This system takes advantage of the strong polyhedrin promoter of baculovirus and the eukaryotic exocytotic protein synthesis of insect cells. HIV-1 gp160 so expressed are properly glycosylated and retains CD4 binding ability [7, 8], although the oligosaccharide side chains may differ from the native form [7, 9]. However, the proteolytic cleavage of gp160 into gp120 and gp41 in insect cells is slow and inefficient [7, 11, 12]. The lack of efficient cleavage also occured to the processing of baculovirus-expressed hemagglutinin of influenza virus [13]. In this study, we used insect cells to generate gp41 alone and bypassed the difficulty of producing gp41 from its gp160 precursor. Two different constructs of gp41 were prepared and their glycosylation and antigenecity have been examined.

MATERIALS AND METHODS

To independently express gp41 protein in insect cells, our strategy was to locate gp41 sequence immediately after the *env* leader peptide coding sequence (Fig. 1) and left out the entire gp120 coding region. To facilitate the plasmid constructions, we eluted a *Kpn* I/*Xho* I fragment from pBL518E, which was derived from *env*-containing pBWT [14] with a difference of an *Eco*R I site introduced to the junction of sequences coding for gp120 and gp41. This 2546 base pair (bp) fragment was ligated to *Kpn* I/*Bam*H I-digested pSP72 (Promega). A unique *Bam*H I site (at nucleotide 8475 of HXB2 [15] proviral sequence) of the resulted plasmid was then mutated by crossover linker mutagenesis as described [16, 17]: the oligonucleotides used were 5'GATCtTTGGCACTTATCTGG and 5'CCAGAGAAGTGCCAAa, where small letters indicates the designated nucleotide substitutions which result in no amino acid change. The mutated sequence was confirmed by superhelical DNA sequencing. With a similar strategy, a *Bam*H I site was introduced after the unique *Xho* I site, which is 103 bp downstream of the termination codon for gp41. The oligonucleotides used were 5'AGGATTTTGCTATAAGGATCC and 5'TCGAGGATCCTTATAGCAAATTCCTT. The inserted region in the resulting plasmid, pS41B, was sequenced and confirmed accordingly.

To link the *env* leader peptide with gp41, the signal coding region in pBWT was first amplified by polymerase chain reaction (PCR) using the following two primers: 5'GGATCCAGAAGACAGTGGCAATGAG (sense) and 5'CAAGAATTCTGTAGCAC TACAGATCAT (antisense); underlined nucleotides indicate the incorporated 5'-end *Bam*H I and 3'-end *Eco*R I sites, respectively. The amplified product was one-end digested with *Eco*R I and ligated to *Eco*R V/*Eco*R I digested pBluescript II SK+ (Strategene). After white colony selection with the use of isopropyl-β-thio-galactopyranoside and 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside, plasmids were harvested and the entire 123-bp region of PCR product was sequenced and confirmed to be as expected. A *Hinc* II/*Eco*R I fragment containing the123-bp fragment

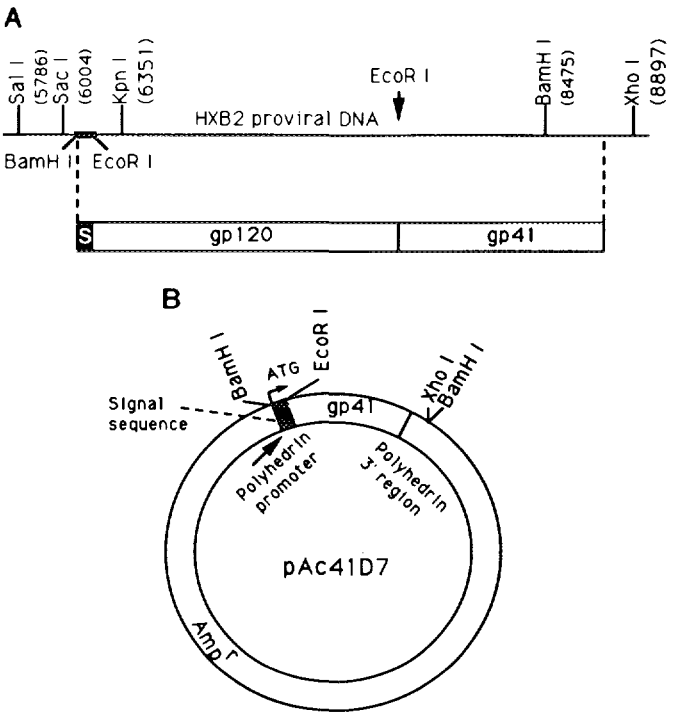


Fig. 1. Diagrams for the construction of baculovirus transfer vector. (A) The relative location of restriction enzyme sites along the HIV-1(HXB2) sequence. The arrow indicates the location of an EcoR I site created by site-directed mutagenesis in the pBL518E (see Text). The thick line flanked by BamH I and EcoR I denotes the region where the signal sequence was amplified by polymerase chain reaction. The hatched region labeled with S represents the *env* signal peptide. (B) Physical map of transfer vector pAc41D7. The *env* signal sequence immediately followed by gp41 coding sequence is driven by the polyhedrin promoter in the orientation indicated by the arrow.

was then eluted and ligated with an *Eco*R I/*Eco*R V fragment (3.5 Kb) of pS41B. The resulting plasmid, pB41EB, was confirmed by sequencing to have the 123-bp fragment followed by the gp41 coding region. A *Bam*H I fragment (1.3 Kb) from pB41EB was then inserted into the baculovirus transfer vector pAcYM1 [18] and the resulted plasmid was named pAc41D7 (Fig. 1). The DNA and the predicted amino acid sequences in the junction between the *env* leader peptide and the N-terminal region of gp41 are shown in Fig. 2A.

Signal sequences, glycosylation sites, sorting signals, and proteolytic cleavage sites of several proteins have apparently been authentically recognized when expressed in insect cells using baculovirus recombinants [10]. Considering the usage of the same system and the construct of pAc41D7, a proper removal of the *env* signal peptide would probably yield gp41 molecules having Thr-Glu replacing the first seven N-terminal hydrophobic residues in the authentic gp41(Fig. 2A). This variant would allow us to evaluate the effect on expression and the antigenic contribution of the very N-terminal hydrophobic residues of gp41.

To obtain a gp41 molecule with a complete sequence, pB41EB was digested with *Eco*R I and linker mutagenesis was done as described above. The oligonucleotides used were 5'GATCGGTAGTGCTGCAGTGGGAATAGGAGCTC and 5' AATTGAGCTCCTATTCCCACTGCAGCACTACCGATC, where nucleotides underlined

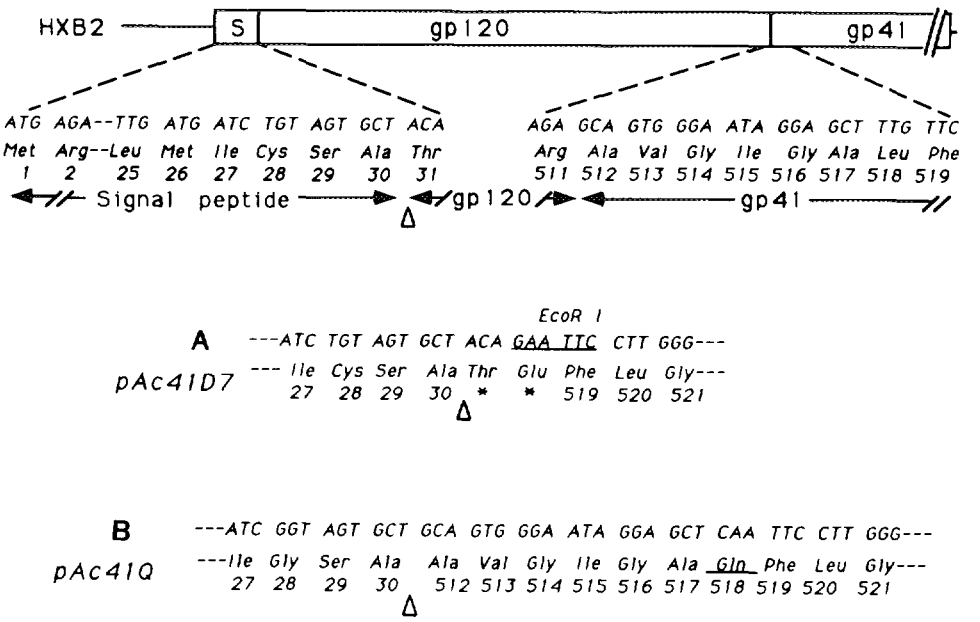


Fig. 2. Sequences immediately follow the *env* signal peptide coding regions in pAc41D7 (A) and pAc41Q (B). Partial *env* sequences around the junctions of signal peptide/gp120 and gp120/gp41 are provided on the top for comparison. The open arrow heads denote the putative cleavage sites used by the signal peptidase. Amino acid residues were numbered according to the positions in *env* open reading frame of the parental HXB2 [15]. Two residues replacing the AVGIGAL segment are asterisked. The Gln underlined in pAc41Q sequence illustrates the residue different from that of the parental gp41 in this position.

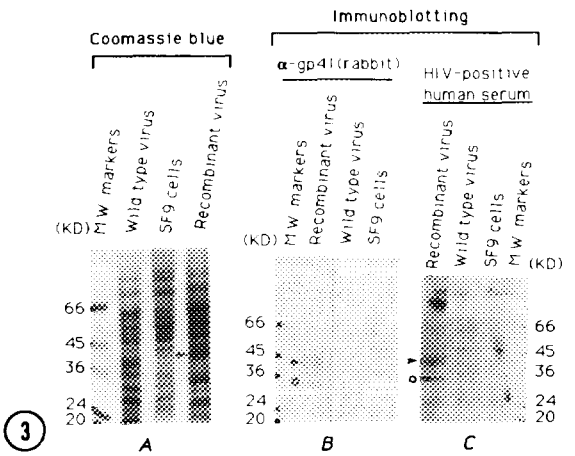
were *Pst* I and *Sac* I sites, respectively, which readily facilitated the screening of desired mutants. By an analogy to the construction of pAc41D7, a transfer plasmid termed pAc41Q was thereafter obtained and confirmed by sequencing. The corresponding sequences in the junction between the *env* leader peptide and the N-terminal region of gp41 are shown in Fig. 2B. This construct has the exact gp41 sequence of HXB2 clone except that a Leu to Gln substitution occurs at the amino acid 518, a position which is only semiconserved for neutral residues [19].

Recombinant baculoviruses were generated basically according to the previously described method [20]. In brief, SF9 cells were cotransfected with 1 μg wild type baculoviral DNA and 3 μg transfer vector DNA. After four-day cultivation, the viruses obtained in the media were limitingly diluted and used to infect fresh SF9 cells in 96-well plates. Cells were microscopically examined for the absence of polyhedrin occlusion bodies and DNA extracted 10 days post infection were dotted onto nylon membranes and hybridized with probes derived from random priming of the entire 1.3 Kb *Bam*H I inserts containing the gp41 sequences. By doing so, recombinant viruses from wells that gave signals positive for dot-blotting hybridization and negative for presence of wild type virus were obtained.

RESULTS AND DISCUSSION

Expression of gp41D7 and gp41Q in SF9 cells was first examined for extra protein bands on 10% polyacrylamide gel by direct staining with Coomassie blue.

Typical results with gp41Q were shown in Fig 3A. A broad band centering at approximate 41 Kd in mass was observed in the lysates of cells infected with recombinant virus, but not in that derived from parental virus infection or from SF9 cells (Fig. 3A). This protein with gp41-related identity was demonstrated by Western blotting. A rabbit antiserum raised against fragment of HIV-1 gp41 (Syu, W.J. unpublished results) bound to this protein. In contrast, no protein with the same size was detected in the lysate of the parental baculovirus-infected cells or that of cells without infection (Fig. 3B). The fact that recombinant gp41 was expressed was substantiated by the specific reaction with HIV-1+ human sera (Fig. 3C). Interestingly, along with the detection of gp41Q molecules, a minor protein with a molecular weight estimated to be 33 Kd was specifically and reproducibly detected by the rabbit anti-gp41 and HIV-1+ human sera. This protein may represent the product of alternative initiation from the gp41 reading sequence or the degradative gp41, or simply be the nonglycosylated form of gp41Q molecules. Similar results were observed with the gp41D7 expressed (data not shown).



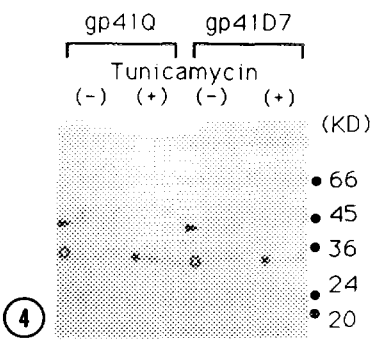


Fig. 3. SF9-expressed gp41Q in crude cell extracts. (A) Total cellular proteins in SDS sample buffer extracts were separated by SDS-PAGE (10% gel) and stained with Coomassie blue. (B)Proteins on gels were transferred onto nitrocellulose membrane, marked with Amido Black 10B, and reacted with a gp41-specific rabbit antiserum. After additional reactions with biotinylated goat anti-rabbit IgG and strepavidin-horseradish peroxidase, blots were developed with 4-chloro-1-naphthol [27]. (C) Blots were processed as in B except that primary antibodies are from an HIV-1+ human serum and secondary antibodies are biotinylated goat anti-human IgG. The primary antibodies used were 1 to 100 diluted. Arrow heads indicate the stained recombinant gp41 whereas open circles label the gp41-related minor protein.

Fig. 4. Western blotting of lysates from cells infected with recombinant viruses in the presence of tunicamycin. Tunicamycin (20 μg/ml) was added to media four hours after infection. Cells were harvested 2 days post infection. Western blotting with the anti-gp41 rabbit serum was performed as described in the Fig. 3 legend. Arrow heads mark the recombinant gp41; open circles label the minor gp41 products; asterisked bands indicate the single gp41-related molecules after the N-linked glycosylation was blocked by tunicamycin.

To better evaluate the natural antigenecity of gp41D7 and gp41Q in infected human individuals, Western blotting were performed with 90 different HIV-1 positive and 20 HIV-1 negative human serum samples. While HIV-1 negatives showed no reaction, all the tested HIV^+ sera reacted with gp41D7 and gp41Q with patterns similar to that in Fig. 3C (data not shown). Also, the minor 33 Kd product described above was all positively detected. Peculiarly, a broad band with approximate 120-160 Kd in mass seen in Fig. 3C was often observed when cell lysates prepared for gp41Q, but not for gp41D7, were reacted with HIV^+ human sera. We are currently investigating whether this band may represent aggregating complexes of intact gp41.

To establish that gp41D7 and gp41Q are glycosylated and that the minor 33 Kd protein could be derived from nonglycosylated gp41, we infected SF9 cells with recombinant viruses in the presence of tunicamycin, which blocks the de novo N-linked glycosylation [21]. Results from Western blotting (Fig. 4) with cell lysates indicate that both gp41Q and gp41D7 are sensitive to tunicamycin action. A shift from 41 Kd to 33 Kd was observed, suggesting that the normally occurring N-linked glycosylation of gp41 was inhibited. The resulted band overlapping with the 33 Kd protein did, however, support the notion that the minor product may represent the nonglycosylated gp41.

In summary, we have expressed gp41 molecules in SF9 insect cells without the coexpression of gp120. In the production of gp160 [7, 11, 12], the signal peptide encoded by HIV *env* apparently was properly recognized in the insect cell system. As shown above, this *env* signal sequence also directed most of the gp41 synthesis to the Golgi N-linked glycosylation pathway. However, like other proteins heterologously expressed, it remains to be tested whether our human viral signal peptide was removed at the exact site by the putative insect signal peptidase. The broad gp41 band detected in SDS PAGE may reflect the glycosylation properties of glycoproteins [6, 22]. This notion was supported by that the sharp protein bands with decreased molecular weight were detected after glycosylation was inhibited (Fig. 4).

The role of gp41 in the viral infectivity and pathogenesis has been suggested [6, 23-25] and the relationship between structure and function of gp41 has been correlated to some extent. Experiments with linker insertion reveals that regions in the the N-terminal half of gp41 may be involved in association with gp120 [26]. The very N-terminal end of gp41 is a stretch of hydrophobic and nonpolar amino acid residues. By an analogy to other viruses, this region was proposed to be the fusogenic domain [6, 24]. Our construct of gp41D7 has a hydrophilic N-terminal dipeptide, Thr-Glu, replacing the continuous seven nonpolar residues. Apparently, no significant difference was observed in terms of expression, glycosylation, and antigenecity between this variant and that of the full length construct, gp41Q. With these eucaryote-expressed gp41 molecules, it becomes possible to reconstitute the viral envelope for further examination of the ill-defined functions of gp41.

ACKNOWLEDGMENTS

We thank N. Wang, K.W. Leou Y.Y. Ching, L.S. Chang and Dr. J.F. Lo for helpful discussion and Drs. S.P. Lin and H.J. Lin for valuable reagents. This work was supported in part by grants NSC80-0412-B010-037 and NSC 80-0412-B010-060 from National Science Council, Taiwan, R.O.C. and an award from Medical Research and Advancement Foundation in Memory of Dr. Chi-Shuen Tsou.

REFERENCES

1. Willey, R. L., Bonifacino, J.S., Potts, B.J., Matin, M.A., and Klausner, R.D. (1988) Proc. Natl. Acad. Sci. USA **85**, 9580-9584.
2. Fennie,C., and Lasky, L. A. (1989) J. Virol. **63**, 639-646.
3. Ozel, M., Pauli, G., and Gelderblom, H. R. (1988) Arch. Virol. **100**, 255-266.
4. Gelderblom, H. R., Hausmann, E. H. S., Ozel, M., Pauli, G., and Koch, M. A. (1987) Virology **156**, 171-176.
5. Schneider, J., Kaaden, O., Copeland, T. D., Oroszlan, S., and Hunsmann, G. (1986) J. Gen. Virol.**67**, 22533-2538.
6. Berman, P. W., Nunes, W. M., and Haffar, O. K. (1988) J. Virol. **62**, 3135-3142.
7. Wells, D. E., and Compans, R. W. (1990) Virology **176**, 575-586.
8. Berman, P. W., Riddle, L., Nakamura, G., Haffar, O. K., Nunes, W. M., Skehel, P.,Byrn, R., Groopman, J., Mathews, T., and Gregory, T. (1989) J. Virol. **63**, 3489-3498.
9. Rusche, J. R., Lynn, D. L., Robert-Duroff, M., Langlois, A. L., Lyerly, H. K., Carson, H., Krohn, K., Rankia. A., Gallo, R. C., Bolognesi, D. P., Putney, S. D., and Matthews, T. J. (1987) Proc. Natl. Acad. Sci. USA **84**, 6924-6928.
10. Luckow, V. A., and Summers, M. D. (1988) Bio/Technology **6**, 47-55.
11. Cochran, M. A., Ericson, B. L., Knell, J. D., and Smith, G. E. (1987) In " Vaccines 87 " (R. M. Chanock, R.A. Lerner, F. Brown, and H. Ginsberg, Eds.), pp. 384-388. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
12. Hu, S., Kosowski, S. G., and Schaff, K. F. (1987) J. Virol. **61**, 3617-3620.
13. Kuroda, K., Hauser, C., Rott, R. Klenk, H. D., and Doerfler, W. (1986) EMBO J. **5**, 1359-1365.
14. Syu, W. J., Lee, W. R., Essex, M., and Lee, T. H. (1991) J. Virol. **65**, 6349-6352.
15. Fisher, A. G., Collati, Ratner, L., Gallo, R. C., and Wong-Staal, F. (1985) Nature (London) **316**, 262-265.
16. Luo, L., Li,Y., and Kang, C. Y. (1990) Virology **179**, 874-880.
17. Sung, W. L., Zahab, D. M., MacDonald, C. A., and Tam, C. S. Gene (1986) **47**, 261-267.
18. Chang, L. H., Fan, J. Y., Liu, L. F., Tsai, S. P., and Tam, M. F. Biochem. J. (1992) **281**, 545-551.
19. Myers, G., Rabson, A. B., Josephs, S. F., Smith, T. F., Berzofsky, J.A.and Wong-Staal, F. (1989) In "Human Retroviruses and AIDS 1989: A Complication and Analysis for Nucleic Acid and Amino Acid Sequences" Los Alamos National Laboratory, Los Alamos, N. Mex.
20. Smith, G. E., Ju, G., Ericson, B. L., Moschera, J., Lahm, H. W., Chizzonite, R., and Summers, M. D. (1985) Proc. Natl. Acad. Sci. USA **82**, 8404-8408.
21. Allan, J. S., Coligan, J. E., Barin, F., McLane, M. F., Sodroski, J. G., Rosen, C. A., Haseltine, W. A., Lee, T. H., and Essex, M. (1985) Science **228**, 1091-1094.
22. Kuroda, K., Geyer, R., Doerfler, W., and Klenk, H. D. (1990) Virology, **174**, 418-429.
23. Cohen, D. I., Tani, Y., Tian, H., Boone, E., Samelson, L. E., and Lane, H. C. (1992) Science **256**, 542-545.
24. Gallaher, W. R. (1987) Cell **50**, 327-328.
25. Dubay, J. W., Roberts, S.J. Hahn, B.H., and Hunter, E. (1992) J. Virol. **66**, 6616-6625 .
26. 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.
27. Syu, W. J., and Kahan, L. (1987) J. Immunol. Methods **103**, 247-252.