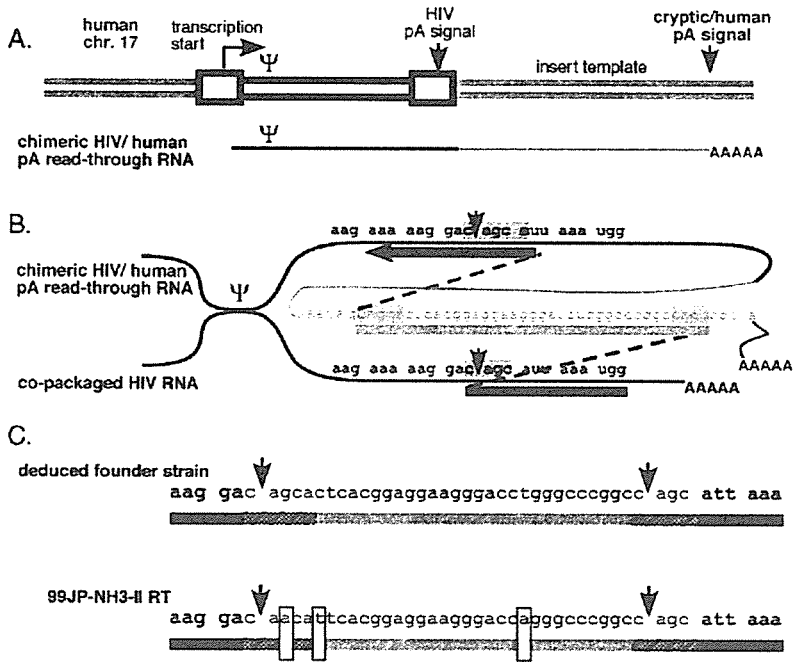

Fig. 2. Splinted recombination model for human sequence insertion. (A) Proviral insertion and read-through RNA. Not to scale. Boxes indicate HIV-1 long terminal repeats; Ψ indicates RNA packaging/dimerization signal. Thick lines are DNA; thin lines are RNA. Black lines are viral sequences; shaded lines are human sequences. Arrows indicate direction of transcription and positions of polyadenylation signals. (B) Template switches postulated to generate 99JP-NH3-II. Shown are two co-packaged RNAs joined at Ψ, and DNA products of template switching (thick and dotted lines: shaded boxes indicate regions of microhomology between viral and human sequences). Note that experimental studies have demonstrated that polyadenylation read-through occurs during the synthesis of about 10% of all HIV RNAs and that even unlinked host sequences can become incorporated in proviruses (Olsen et al., 1990; Sun et al., 2001; An and Telesnitsky, 2004 and references therein). (C) Genetic differences between 99JP-NH3-II and the predicted founder strain. Arrowheads indicate borders of the 33 base insert; hatched sequences were templated by microhomology, boxed residues are sites of mutations evoked to generate 99JP-NH3-II from the deduced founder strain.

insertions are observed in roughly 1% of the viral strains in treated HIV/AIDS patients world-wide (Winters and Merigan, 2005). Although a simple duplication of flanking sequences is sufficient to enhance drug resistance and may be the most frequent cause of β3- β4 insertions, a wide variety of β3- β4 insert sequences – varying in both length and sequence composition – have been observed (Winters and Merigan, 2005).

Identifying likely human origins of the duplication-flanked insert in 99JP-NH3-II was possible because this insert was unusually lengthy. However, alternate means of searching reveal that Genbank contains additional examples of direct repeat-flanked RT insertions, albeit with shorter sequence regions with uncertain origins (Masquelier et al., 2001; van der Hoek et al., 2005) (see Materials and methods). Although these latter inserts are too short for homology searching to implicate specific sequences, it seems reasonable to postulate that their synthesis involved either virus or host bridging templates.

Not surprisingly, when the human genome is used to query all HIV-1 sequences in GenBank, the few isolates that appear genuine and contain long (>50 nt) human inserts are annotated as replication-defective (for example, GenBank accession no. AY561239; see Materials and methods). Among replication-competent HIV-1 isolates, the inserts with strongest virus–human match belong to the class of mutations called AVT codon-rich *env* variable region extensions (Kitrinos et al., 2003). The codon bias differences between these inserts and other HIV-1 sequences, as well as the similarity of these inserts to repetitive human microsatellite sequences, have been described previously (An and Telesnitsky, 2004; Bosch et al., 1994; Kitrinos et al., 2003). However, most HIV AVT-rich extensions identified by BLAST using human query sequences have fewer than 30 contiguous bases of match to the human genome reference sequence. Assessing possible human origins for these AVT-rich inserts is complicated by these inserts' genetic plasticity, the extent of microsatellite variation within the human population, and the relatively high frequency with which AVT-rich extensions are observed: leaving open the possibility that some arose via recombination between viral isolates (Kitrinos et al., 2003; Zhivotovsky et al., 2003). Nonetheless, observations such as RT β3- β4 inserts that lack genetically linked regions of sequence identity (Masquelier et al., 2001) and the *de novo* generation of AVT-rich extensions during virus replication in tissue culture (Kuhmann et al., 2004) suggest that on rare occasions additional instances of short patch human sequence transduction, like that reported here for 99JP-NH3-II, may contribute to HIV-1 genetic variation.

Materials and methods

Sequences and sequence analysis

Analyses shown were performed using default settings in nucleotide-nucleotide BLAST (blastn) via NCBI (http://www.ncbi.nlm.nih.gov). All sequences analyzed in this report were previously deposited in GenBank. Searches described as unrestricted databases in GenBank or "nr" were searches of (all GenBank + RefSeq Nucleotides + EMBL + DDBJ + PDB sequences). Expect values presented in Fig. 1B are those assigned by NCBI blastn using default gap and mismatch penalties.

Additional examples of β3- β4 insert splinted recombinants (GenBank accession nos. AY877315 and AF315271) mentioned in the discussion were identified using blastn by querying Genbank with an artificial sequence comprised of NH3's preinsertion sequence (Fig. 1A) modified to contain a β3- β4 insert of arbitrary length (24 b) while applying reduced mismatch and increased gap penalty values and then by manually screening isolates for strong epidemiologic support and structures predicted by splinted recombination.

Methods and findings from querying all HIV sequences in GenBank with the human genome will be described elsewhere.

Acknowledgments

The authors thank Yuki Naito and James Gergel for their help with the sequence analyses, and David Friedman and Mary Jane Wieland for their useful discussions. This work was supported by NIH no. GM64479 to AT.

References

An, W., Telesnitsky, A., 2002. HIV-1 genetic recombination: experimental approaches and observations. AIDS Rev. 4, 195–212.

An, W., Telesnitsky, A., 2004. Human immunodeficiency virus type 1 transductive recombination can occur frequently and in proportion to polyadenylation signal readthrough. J. Virol. 78, 3419–3428.

Berkhout, B., Grigoriev, A., Bakker, M., Lukashov, V.V., 2002. Codon and amino acid usage in retroviral genomes is consistent with virus-specific nucleotide pressure. AIDS Res. Hum. Retroviruses 18, 133–141.

Bosch, M.L., Andeweg, A.C., Schipper, R., Kenter, M., 1994. Insertion of N-linked glycosylation sites in the variable regions of the human immunodeficiency virus type 1 surface glycoprotein through AAT triplet reiteration. J. Virol. 68, 7566–7569.

Dunn, M.M., Olsen, J.C., Swanstrom, R., 1992. Characterization of unintegrated retroviral DNA with long terminal repeat-associated cell-derived inserts. J. Virol. 66, 5735–5743.

Fang, H., Pincus, S.H., 1995. Unique insertion sequence and pattern of CD4 expression in variants selected with immunotoxins from human immunodeficiency virus type 1-infected T cells. J. Virol. 69, 75–81.

Gao, X.-S., Cheng, J., Zhen, Z., Guo, J., Zhang, L.-Y., Tao, M.-L., 2005. Cloning of hepatitis B virus DNAPTP1 transactivating genes by suppression subtractive hybridization technique. Shijie Huaren Xiaohua Zazhi 13, 2371–2374.

Hacker, J., Kaper, J.B., 2000. Pathogenicity islands and the evolution of microbes. Annu. Rev. Microbiol. 54, 641–679.

Hajjar, A.M., Linial, M.L., 1993. A model system for nonhomologous recombination between retroviral and cellular RNA. J. Virol. 67, 3845–3853.

Harris, R.S., Bishop, K.N., Sheehy, A.M., Craig, H.M., Petersen-Mahrt, S.K., Watt, I.N., Neuberger, M.S., Malim, M.H., 2003. DNA deamination mediates Innate immunity to retroviral infection. Cell 113, 803–809.

Jetzt, A.E., Yu, H., Klarmann, G.J., Ron, Y., Preston, B.D., Dougherty, J.P., 2000. High rate of recombination throughout the human immunodeficiency virus type 1 genome. J. Virol. 74, 1234–1240.

Kitrinos, K.M., Hoffman, N.G., Nelson, J.A., Swanstrom, R., 2003. Turnover of env variable region 1 and 2 genotypes in subjects with late-stage human immunodeficiency virus type 1 infection. J. Virol. 77, 6811–6822.

Kuhmann, S.E., Pugach, P., Kunstman, K.J., Taylor, J., Stanfield, R.L., Snyder, A., Strizki, J.M., Riley, J., Baroudy, B.M., Wilson, I.A., Korber, B.T., Wolinsky, S.M., Moore, J.P., 2004. Genetic and phenotypic analyses of human immunodeficiency virus type 1 escape from a small-molecule CCR5 inhibitor. J. Virol. 78, 2790–2807.

Leitner, T., Albert, J., 1999. The molecular clock of HIV-1 unveiled through analysis of a known transmission history. Proc. Natl. Acad. Sci. U.S.A. 96, 10752–10757.

Lobato, R.L., Kim, E.Y., Kagan, R.M., Merigan, T.C., 2002. Genotypic and phenotypic analysis of a novel 15-base insertion occurring between codons 69 and 70 of HIV type 1 reverse transcriptase. AIDS Res. Hum. Retroviruses 18, 733–736.

Malim, M.H., Emerman, M., 2001. HIV-1 sequence variation: drift, shift, and attenuation. Cell 104 (4), 469–472.

Masquelier, B., Race, E., Tamalet, C., Descamps, D., Izopet, J., Buffet-Janvresse, C., Ruffault, A., Mohammed, A.S., Cottalorda, J., Schmuck, A., Calvez, V., Dam, E., Fleury, H., Brun-Vezinet, F., ANRS AC11 Resistance study group, French Agence Nationale de Recherches sur le SIDA, 2001. Genotypic and phenotypic resistance patterns of human immunodeficiency virus type 1 variants with insertions or deletions in the reverse transcriptase (RT): multicenter study of patients treated with RT inhibitors. Antimicrob. Agents Chemother. 45, 1836–1842.

Mikkelsen, J.G., Pedersen, F.S., 2000. Genetic reassortment and patch repair by recombination in retroviruses. J. Biomed. Sci. 7, 77–99.

Muriaux, D., Rein, A., 2003. Encapsidation and transduction of cellular genes by retroviruses. Front. Biosci. 8, 135–142.

Olsen, J.C., Bova-Hill, C., Grandgenett, D.P., Quinn, T.P., Manfredi, J.P., Swanstrom, R., 1990. Rearrangements in unintegrated retroviral DNA are complex and are the result of multiple genetic determinants. J. Virol. 64 (11), 5475–5484.

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., 2004. Complete sequencing and characterization of 21,243 full-length human cDNAs. Nat. Genet. 36, 40–45.

Parthasarathi, S., Varela-Echavarria, A., Ron, Y., Preston, B.D., Dougherty, J.P., 1995. Genetic rearrangements occurring during a single cycle of murine leukemia virus vector replication: characterization and implications. J. Virol. 69, 7991–8000.

Pathak, V.K., Temin, H.M., 1990. Broad spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: deletions and deletions with insertions. Proc. Natl. Acad. Sci. U.S.A. 87, 6024–6028.

Preston, B.D., Dougherty, J.P., 1996. Mechanisms of retroviral mutation. Trends Microbiol. 4 (1), 16–21.

Pulsinelli, G.A., Temin, H.M., 1991. Characterization of large deletions occurring during a single round of retrovirus vector replication: novel deletion mechanism involving errors in strand transfer. J. Virol. 65, 4786–4797.

Sato, H., Tomita, Y., Ebisawa, K., Hachiya, A., Shibamura, K., Shiino, T., Yang, R., Tatsumi, M., Gushi, K., Umeyama, H., Oka, S., Takebe, Y., Nagai, Y., 2001. Augmentation of human immunodeficiency virus type 1 subtype E (CRF01_AE) multiple-drug resistance by insertion of a foreign 11-amino-acid fragment into the reverse transcriptase. J. Virol. 75 (12), 5604–5613.

Sun, G., O'Neil, P.K., Yu, H., Ron, Y., Preston, B.D., Dougherty, J.P., 2001. Transduction of cellular sequence by a human immunodeficiency virus type 1-derived vector. J. Virol. 75, 11902–11906.

Temin, H.M., 1993. Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. Proc. Natl. Acad. Sci. U.S.A. 90 (15), 6900–6903.

van der Hoek, L., Back, N., Jebbink, M.F., de Ronde, A., Bakker, M., Jurriaans, S., Reiss, P., Parkin, N., Berkhout, B., 2005. Increased mulitnucleoside drug resistance and decreased replicative capacity of a human immunodeficiency virus type 1 variant with an 8-amino-acid insert in the reverse transcriptase. J. Virol. 79, 3536–3543.

Wain-Hobson, S., Renoux-Elbe, C., Vartanian, J.P., Meyerhans, A., 2003. Network analysis of human and simian immunodeficiency virus sequence sets reveals massive recombination resulting in shorter pathways. J. Gen. Virol. 84, 885–895.

Winters, M.A., Merigan, T.C., 2005. Insertions in the human immunodeficiency virus type 1 protease and reverse transcriptase genes: Clinical impact and molecular mechanisms. Antimicrob. Agents Chemother. 49, 2575–2582.

Zhang, J., Temin, H.M., 1993. 3′ junctions of oncogene-virus sequences and the mechanisms for formation of highly oncogenic retroviruses. J. Virol. 67 (4), 1747–1751.

Zhivotovsky, L.S., Rosenberg, N.A., Feldman, M.W., 2003. Features of evolution and expansion of modern humans, inferred from genomewide microsatellite markers. Am. J. Hum. Genet. 72, 1171–1186.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 38, pp. 28113–28121, September 22, 2006

Ets-1-dependent Expression of Vascular Endothelial Growth Factor Receptors Is Activated by Latency-associated Nuclear Antigen of Kaposi's Sarcoma-associated Herpesvirus through Interaction with Daxx*

Received for publication, March 3, 2006, and in revised form, July 19, 2006 Published, JBC Papers in Press, July 20, 2006, DOI 10.1074/jbc.M602026200

Yuko Murakami‡, Satoshi Yamagoe‡[1], Kohji Noguchi‡, Yutaka Takebe§, Naoko Takahashi‡, Yoshimasa Uehara‡, and Hidesuke Fukazawa‡

From the ‡Department of Bioactive Molecules and §Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, Tokyo 162, Japan

Vascular endothelial growth factor (VEGF) and its receptors are highly expressed in Kaposi's sarcoma (KS) lesion and play a key role in angiogenesis. Latency-associated nuclear antigen (LANA) of Kaposi's sarcoma-associated herpesvirus (KSHV/HHV8) has multiple functions related to viral latency and KSHV-induced oncogenesis. In this report, we have identified Daxx as a LANA-binding protein by co-immunoprecipitation analysis of HeLa cells stably expressing LANA. LANA associated with Daxx in a PEL cell line infected with KSHV. LANA and Daxx also bound *in vitro*, suggesting direct interaction. From the results of binding assays, a region containing the Glu/Asp-rich domain within LANA, and a central region including the second paired amphipathic helix within Daxx contributed to the interaction. To address the physiological significance of this interaction, we focused on a Daxx-mediated VEGF receptor gene regulation. We found that Daxx repressed Ets-1-dependent Flt-1/VEGF receptor-1 gene expression, and that LANA inhibited the repression by Daxx in a reporter assay. Analyses of flow cytometry and real-time PCR revealed that expression of VEGF receptor-1 and -2 in LANA-expressing human umbilical vein endothelial cells (HUVECs) significantly increased. Co-immunoprecipitation and immunoblotting experiments suggested that LANA-bound Daxx to inhibit the interaction between Daxx and Ets-1. Chromatin immunoprecipitation assays showed that Daxx associated with VEGF receptor-1 promoter in HUVECs, and that LANA expression reduced this association. These results suggested that LANA contributes to a high expression of VEGF receptors in KS lesion by interfering with the interaction between Daxx and Ets-1.

Kaposi's sarcoma-associated herpesvirus (KSHV[2]/HHV8) has been found to be the pathogen of Kaposi's sarcoma (KS) (1), two B cell malignancies, primary effusion lymphoma (PEL), and multicentric Castleman's disease (MCD) (3). Among over 80 ORFs of KSHV (4), LANA (latency-associated nuclear antigen) is exceptionally highly expressed in KS lesion, PEL, and also in MCD (5) (3), so that LANA is used as a diagnostic marker of KSHV. LANA is reported to be a multifunctional protein that tethers its own viral episomal DNA to host chromosomes in mitosis to segregate KSHV into progeny cells (6) (7) and also binds many host molecules to regulate expression of cellular genes. LANA inhibits p53-induced apoptosis (8), transforms fibroblast by co-transfection with the Ras oncogene (9) and also stabilizes β-catenin to stimulate entry into S phase (10). LANA seems to contribute to pathogenesis of KSHV-associated malignancies through these interactions.

We identified Daxx as a new LANA-interacting host protein. To know the biological significance of the interaction between LANA and Daxx, we focused on Daxx-modulated transcription. Daxx was found initially as a Fas-binding protein to regulate apoptosis (11) and later reported to bind with several nuclear proteins and transcription factors. Daxx was shown to act as a transcriptional repressor of Ets-1 (12), Pax3 (13), Pax5 (14), and p53 (15) through protein-protein interaction. In the case of Ets-1, Daxx repressed Ets-1-dependent expression of matrix metalloproteinase 1 (MMP1) and Bcl-2 (12). Ets-1 belongs to the Ets family of transcriptional factors, and regulates various gene expressions through binding to a unique motif (GGAA) on their promoters. Ets-1 regulates genes related to angiogenesis: Flt-1/VEGF receptor-1, KDR/VEGF receptor-2, and matrix metalloproteinases (MMPs) (16). Ets-1 is specifically expressed in lymphoid tissues, endothelial cells (17), and also in the spindle cells of KS lesion, derived from endothelial origins (18). In KS lesion, angiogenic factors such as VEGF and VEGF receptors were highly expressed (1) (19). Vascular angiogenesis plays an important role in the development and progression of tumors, especially KS (20) (21) (22). We therefore examined the role of LANA in interaction between Daxx

* This work was supported in part by a grant for Research on Health Sciences focusing on Drug Innovation from The Japan Human Sciences Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] To whom correspondence should be addressed: Dept. of Bioactive Molecules, National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo, 162-8640, Japan. Tel.: 81-3-5285-1111; Fax: 81-3-5285-1272; E-mail: syamagoe@nih.go.jp.

[2] The abbreviations used are: KSHV, Kaposi's sarcoma-associated herpesvirus; VEGF, vascular endothelial growth factor; ORF, open reading frame; DTT, dithiothreitol; PMSF, phenylmethylsulfonyl fluoride; HA, hemagglutinin; GST, glutathione *S*-transferase; DAPI, 4′,6-diamidino-2-phenylindole; ChIP, chromatin immunoprecipitation assay; GFP, green fluorescent protein; PAH, paired amphipathic helix; HUVEC, human vascular endothelial cells; LANA, latency-associated nuclear antigen; aa, amino acids; Ets, E26 transformation-specific.

and Ets-1, and show here a possible new function of LANA on the expression of VEGF receptors.

EXPERIMENTAL PROCEDURES

Plasmids—LANA gene was cut from L54 Lambda FIX II vector (NIH AIDS Research & Reference Reagent Program), at the EheI site (123743 and 127293 of KSU75698) and inserted into the EcoRV site of pFLAG-CMV-2 expression vector (Sigma) to make an N-terminal FLAG-tagged LANA expression vector, pFLAG-LANA. For *in vitro* translation/transcription system, the *LANA* gene from pFLAG-LANA was subcloned between the EcoRI and KpnI sites of pBluescript II KS(+) plasmid after correction of the N-terminal 5 bases using synthetic oligonucleotides, 5′-AATTCATCGATGGCGCCCCCGGGA-ATGCG-3′ and 5′-CATTCCCGGGGGCTCTATCGATG-3′ (using EcoRI and BsmI sites), to obtain pBluescript-LANA. A series of C-terminal deletion mutants of LANA (L1 to L4) were constructed from pBluescript-LANA using an exonuclease III/mung bean deletion kit (Toyobo, Tokyo, Japan) according to the manufacturer's instructions. An N-terminal deletion mutant of LANA (L5) was constructed with pBluescript-LANA by cutting the N-terminal region at EcoRI and PstI sites and joining it with synthetic oligonucleotides, 5′-AATTCATC-GATGGAGCCCCTGCA-3′ and 5′-GGGCTCCATCGATG-3′. PFLAG-LANA deletion mutants (pFLAG-LANA-N1 to pFLAG-LANA-C) were constructed with L1–L5 and pFLAG-CMV-2 vector. Full-length and various deletion mutants of the *Daxx* gene were generated by PCR amplification from cDNA of HeLa cells, subcloned into pCR-Blunt II-TOPO plasmid (Invitrogen, Carlsbad, CA), and cloned between the EcoRI and the SalI sites of pcDNA3.1 (−) (Invitrogen) (termed pcDNA-Daxx), pCMV-HA (Clontech Laboratories, Inc. Palo Alto, CA), or pGEX-6P-3 (Amersham Biosciences, Piscataway, NJ). A luciferase reporter plasmid, pFlt-1-luc (containing human Flt-1 promoter −748/+248, D64016) was kindly provided by Dr. K. Morishita (23). Human *ets-1* genes of p51Ets-1 and p42Ets-1 (the full-length Ets-1 and a variant lacking the regulatory domain, exon VII, respectively (24)), were cut from plasmids kindly provided by Dr. R. Li (12), and cloned into pcDNA3.1(+) (Invitrogen). The constructed plasmids were termed pcDNA-p51Ets-1 and pcDNA-p42Ets-1, respectively. For flow cytometric analysis, the full-length *LANA* gene from the pBluescript-LANA was cloned between the SacI and SalI sites of pIRES2-EGFP vector (Clontech), termed pIRES2-LANA-EGFP.

Cell Culture and Transfection—HeLa cells and human embryonic kidney 293T cells were cultured in Dulbecco's modified medium supplemented with 10% bovine fetal serum. A KSHV-infected PEL cell line, BCBL-1 cells (kindly provided by Dr. H. Katano) were cultured in RPMI 1640 with 10% bovine fetal serum. Human umbilical vein endothelial cells (HUVEC) (Clonetics, San Diego, CA) were cultured in EGM-2 medium (Clonetics). Transfection was performed with FuGENE6 (Roche Diagnostics, Indianapolis, IN) for HeLa and 293T cells or by Nucleofecter system (amaxa GmbH, Cologne, Germany) for HUVEC.

Identification of LANA-binding Protein—PFLAG-LANA was transfected into HeLa cells and stable LANA-expressing clones were selected. LANA-expressing cells of six liters were harvested, nuclear extract was prepared as previously described (25), and dialyzed against a buffer containing 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.2 mM EDTA, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 10 mM β-mercaptoethanol. The nuclear extract was adjusted to 150 mM NaCl and 0.1% Tween 20 before absorption into anti-FLAG antibody (M2) affinity gel (Sigma). The gel was washed with 150 mM washing buffer (20 mM Tris-HCl, pH 8.0, 5 mM $MgCl_2$, 150 mM NaCl, 1 mM dithiothreitol (DTT), 10% glycerol, 1 mM PMSF), and eluted with the same buffer containing 200 μg/ml of FLAG peptides. The eluted protein was applied to SDS-PAGE and stained with Coomassie Brilliant Blue or silver stained. The Coomassie-stained band was cut and treated with lysyl endopeptidase. The extracted peptides were purified using HPLC, and analyzed with a Procise 494 HT Protein Sequencing System (Applied Biosystems, Foster City, CA).

Immunoprecipitation and Western Blotting—Cells were harvested and lysed with low salt buffer (10 mM HEPES, pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT containing 0.5% Nonidet P-40, 1 mM PMSF, 25 μg/ml each of antipain, pepstatin, and leupeptin), then centrifuged to collect the nuclei. The nuclear pellet was lysed with nuclear extract buffer (20 mM Tris-HCl, pH 7.9, 5 mM, EDTA, 300 mM NaCl, 1 mM PMSF), and the same volume of distilled water was added. Nuclear extract was subjected to immunoprecipitation either with anti-FLAG antibody (M2), anti-Daxx antibody (sc-7152) (Santa Cruz Biotechnology, Inc., Santa Cruz, CA), or anti-Ets-1 antibody (Santa Cruz Biotechnology, sc-350). The immune complex was washed with 150 mM washing buffer, resolved in Laemmli's sample buffer and applied to SDS-PAGE. Proteins in gels were transferred to PVDF membrane followed by Western blotting using anti-FLAG antibody (M5, Sigma), anti-Daxx antibody (sc-7152), anti-LANA antibody (Advanced Biotechnologies, Columbia, MD), anti-Ets-1 antibody (sc-350), or anti-HA antibody (Sigma).

GST Pull-down Assay—Glutathione *S*-transferase (GST)-Daxx fusion proteins were expressed in *Escherichia coli*, and purified using affinity matrix glutathione-Sepharose beads (Amersham Biosciences). ^{35}S-labeled LANA was made *in vitro* with TnT quick-coupled reticulocyte transcription/translation systems (Promega, Madison, WI). GST-Daxx fusion proteins were bound to glutathione-Sepharose beads and incubated with the translated products containing ^{35}S-labeled LANA in binding buffer (25 mM HEPES, pH 7.6, 50 mM NaCl, 2.5 mM $MgCl_2$, 1 mM DTT, 0.05% Triton X-100, 1 mM PMSF) at room temperature for 15 min. After washing with washing buffer (25 mM HEPES, pH 7.6, 150 mM NaCl, 2.5 mM $MgCl_2$, 1 mM DTT, 0.5% Triton X-100, 1 mM PMSF), proteins adsorbed to the beads were resolved and applied to SDS-PAGE. Proteins in the gel were stained with Coomassie Brilliant Blue, and the radioactivity was detected by BAS-1500 (Fuji Film, Tokyo).

Immunofluorescence Assay—BCBL-1 cells were attached to slide glasses with a cell concentrator (StatSpin, Norwood, MA). HeLa cells were seeded on chamber slides (Lab-Tek, Campbell, CA). The cells were fixed in 4% paraformaldehyde in PBS(−), permeabilized with 0.2% Triton X-100, and incubated with anti-LANA antibody (Advanced Biotechnologies) (diluted to 1:500 or 1:1000) and anti-Daxx antibody (sc-7152) (diluted to

1:100 or 1:200) followed by fluorescent-conjugated second antibody (Alexa Fluor 488 anti-rat IgG and Alexa Fluor 594 anti-rabbit IgG, Molecular Probes, Eugene, OR) (diluted to 1:200 each). Cell nuclei were stained with DAPI in mounting oil (Vectashield with DAPI, Vector Laboratories, Burlingame, CA). Immunostained cells were analyzed by a confocal laser scanning microscope using a Carl Zeiss LSM510 system (Carl Zeiss, Oberkochen, Germany).

Transcriptional Reporter Assay—293T cells (2×10^5 cells/well) grown in 24-well plates were transfected with pFlt-1-luc, pRSV-β-Gal (for transfection efficiency), and the combination of pcDNA-p51Ets-1, pcDNA-p42Ets-1, pcDNA-Daxx, or pFLAG-LANA, with 2 μl of FuGENE6. Total DNA was adjusted to a constant amount (800 ng). Two days after transfection, the cells were lysed and applied to luciferase assay (Toyo Inki, Tokyo) and β-galactosidase enzyme assay system (Promega). Assays were performed in triplicate, and the experiments were repeated three times.

Flow Cytometric Analysis—HUVECs transfected with either pIRES2-LANA-EGFP or pIREAS2-EGFP as control, were subjected to FACS Vantage (Becton Dickinson, Franklin Lakes, NJ) to collect GFP-expressing cells 2 days after transfection. The GFP-expressing cells cultured for 10 days were incubated with anti-Flt-1 or anti-KDR antibody (V4262, V9134, respectively, Sigma) and PE-labeled secondary antibody (R0439, Dako Cytomation, Carpinteria, CA). The cells were analyzed by a FACS Calibur flow cytometer (Becton Dickinson).

Quantitative Real-time RT-PCR—HUVECs were transfected with either pIRES2-LANA-EGFP or pIREAS2-EGFP using Nucleofector system, and sorted with FACS Aria (Becton Dickinson) to collect GFP-expressing cells. Total RNA was extracted with RNeasy Mini kit (Qiagen GmbH, Hilden, Germany), and reverse-transcribed to cDNA with oligo(dT) by Superscript First-strand synthesis system according to the manufacturer's instructions (Invitrogen). The cDNA was applied to Real-Time PCR using SYBR Premix Ex Taq (Takara Bio Co.) with ABI PRISM7000 (Applied Biosystems). PCR was performed at 95 °C for 10 s, followed by 40 cycles of 95 °C for 5 s and 60 °C for 34 s. The primers were designed using software Primer Express (Applied Biosystems). The forward and reverse primers for Flt-1 were 5′-CCCTTATGATGCCAGCAAGTG-3′ and 5′-CCAAAAGCCCCTCTTCCAA-3′, respectively, and primers for KDR were 5′-CACCACTCAAACGCTGACATGTA-3′ and 5′-CCAACTGCCAATACCAGTGGAT-3′. Primers for Daxx were 5′-GCCCTTCACCACTGTCTTAGAGA-3′ and 5′-GAGACGCCTCCATTGAAGGA-3′. As an internal control, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used with primers 5′-GGAGTCAACGGATTTGGTCGTA-3′, and 5′-GGCAACAATATCCACTTTACCAGAGT-3′. The primers for Ets-1 were 5′-CTGCGCCCTGGGTAAAGA-3′ and 5′-CCCATAAGATGTCCCCAACAA-3′. In the case of Ets-1, primers for GAPDH were 5′-CCACCCATGGCAAATTCC-3′ and 5′-TGGGATTTCCATTGATGACAAG-3′. Obtained data were analyzed according to the sequence detector program (Applied Biosystems).

Chromatin Immunoprecipitation (ChIP) Assay—ChIP assays were performed basically using a kit from Upstate Biotechnology (Lake Placid, NY) with some modifications. Cells (5×10^6 cells/assay) were treated with 1% formaldehyde for 5 min for cross-linking, lysed with 400 μl of lysis buffer (10 mM HEPES, pH 7.9, 60 mM KCl, 0.5% Nonidet P-40, 1 mM PMSF, 25 μg/ml each of antipain, pepstatin, and leupeptin), and centrifuged to collect the nuclei. The nuclear pellet was lysed with 200 μl of SDS lysis buffer (50 mM Tris-Hcl, pH 8.1, 10 mM EDTA, 1% SDS), and sonicated 6 times for 30 s each time. Centrifuged supernatants were then diluted with 1.8 ml of ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8.0), and precleared with protein A agarose/salmon sperm DNA. Anti-Ets-1 (Santa Cruz Biotechnology, sc-350), anti-Daxx antibody (Santa Cruz Biotechnology, sc-7152), or rabbit IgG (Sigma) as the negative control was added respectively to the supernatant, and rotated overnight at 4 °C. The mixture was then incubated with protein A-agarose/salmon sperm DNA for 1 h at 4 °C. The protein A-agarose-conjugated complex was washed, and DNA fragments were eluted and prepared according to the manufacturer's instructions. The prepared DNA was resolved in 20 μl of H_2O and 2 μl was used for PCR. Primers used were 5′-GGGACCCCTTGACGTCACCA-3′ (corresponding to −90 to −71 of Flt-1 promoter) and 5′-ACCTCGATGAAGAGCAGCCG-3′ (corresponding to −12 to +8 of Flt-1 promoter). Ex Taq polymerase (Takara Bio Co.) was used, and PCR conditions were 30 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 30 s. The PCR products were analyzed on a 1.8% agarose gel.

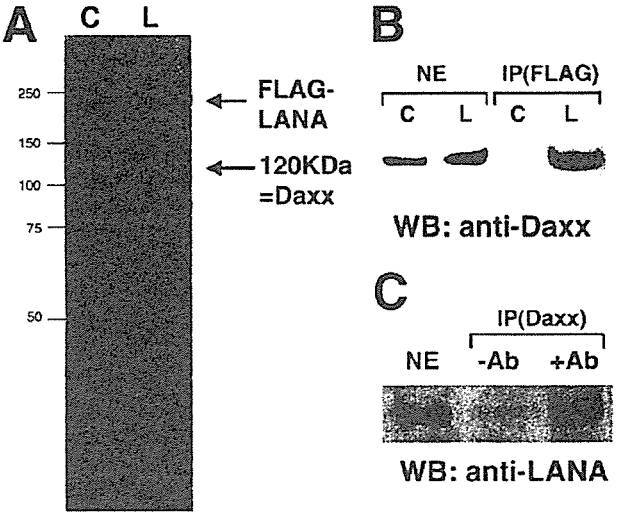

FIGURE 1. **Identification of Daxx as a LANA-binding protein.** *A*, immune complex with anti-FLAG antibody of nuclear extract (*NE*) of HeLa cells was analyzed by SDS-PAGE. Proteins were detected with silver staining. A protein of 120 kDa associated with FLAG-LANA was identified as Daxx. (*C*: control parent HeLa cells, *L*: LANA-expressing HeLa cells). *B*, immune complex with anti-FLAG antibody followed by Western blotting (*WB*) with anti-Daxx antibody. Daxx was detected at about 120 kDa. *C*, immune complex with anti-Daxx antibody followed by Western blotting with anti-LANA antibody using nuclear extract of BCBL-1 cells. A band of about 250 kDa was detected as LANA.

RESULTS

Identification of Daxx as a LANA-binding Protein—To identify host proteins that associate with LANA, we constructed a plasmid expressing FLAG-LANA (N terminus-tagged) to transfect into HeLa cells and established several stable LANA-

expressing clones. We cultured one clone of the LANA-expressing cells and prepared nuclear extract. This extract was incubated with anti-FLAG affinity gel (M2-agarose), followed by elution with FLAG peptides. Eluate was subjected to SDS-PAGE to detect a prominent 120-kDa band (Fig. 1*A*). Although there was a 75-kDa band, which was a nonspecific binding protein commonly found with the antibody. We determined the sequences of the N-terminal 10 residues of the 120-kDa protein, which revealed the protein to be Daxx. To confirm the identification, the nuclear extract (each 500 μg of protein) was immunoprecipitated with anti-FLAG antibody to apply to immunoblotting with anti-Daxx antibody. As shown in Fig. 1*B*, anti-Daxx antibody recognized a band of 120 kDa. These results indicated that Daxx is a cellular binding protein of exogenously expressed LANA in the HeLa cell. To confirm LANA-Daxx interaction in a physiological context, we immunoprecipitated with anti-Daxx antibody from nuclear extracts of BCBL-1 cells, a PEL cell line infected with KSHV. LANA was co-immunoprecipitated with Daxx as well (Fig. 1*C*). This result suggested that LANA formed a complex with Daxx in KSHV-infected cells.

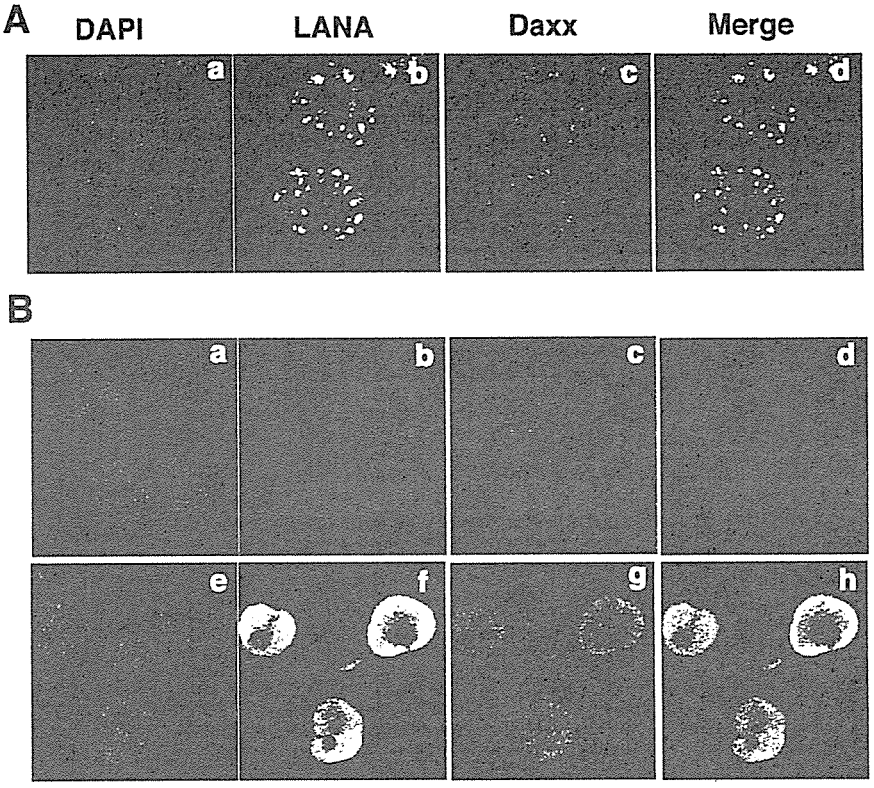

FIGURE 2. **LANA co-localizes with Daxx in BCBL-1 cells and HeLa cells.** Confocal microscopic images of PEL cell line, BCBL-1 cells (*A*), and HeLa cells (*B*, control (*panels a–d*) and LANA-expressing cells (*panels e–h*)). Cells were doubly immunostained with anti-LANA antibody (1:500 for *A*, 1:1000 for *B*) and anti-Daxx antibody (1:100 for *A*, 1:200 for *B*). Images represent cells stained with DAPI (*panels a* and *e*), anti-LANA antibody (*panels b* and *f*), or anti-Daxx antibody (*panels c* and *g*), and merged images of LANA and Daxx staining (*panels d* and *h*).

Colocalization of LANA and Daxx in the Nuclei of KSHV-infected Cell Line BCBL-1—Next we examined the localization of LANA and Daxx in BCBL-1 by immunofluorescence microscopic assay (Fig. 2*A*). LANA gave a characteristic speckled staining pattern in nuclei of the cells (Fig. 2*A*, *panel b*), Daxx also showed some speckles in the nuclei (Fig. 2*A*, *panel c*). The merged image indicated that LANA considerably co-localized with Daxx in the nuclear dots (Fig. 2*A*, *panel d*). We also investigated the localization of Daxx using HeLa cells (Fig. 2*B*). LANA gave fine patchy staining in the nucleus (Fig. 2*B*, *panel f*), which is a typical observation in the absence of KSHV genome (Fig. 2*B*, *panel g*). The parental HeLa cells showed diffused staining of Daxx throughout the cell (Fig. 2*B*, *panel c*). In contrast, Daxx appeared to accumulate in the nuclei of the LANA-expressing cells (used in Fig. 1) (Fig. 2*B*, *panel g*). LANA and Daxx largely localized in the nucleus of the HeLa cells (Fig. 2*B*, *panel h*). We performed biochemical fractionation using three independent clones of LANA-expressing HeLa cells and examined cellular localization of Daxx by Western blotting. The results indicated that the amount of Daxx in the nuclear fraction increased as LANA expression increased, although total amounts of Daxx were comparable in these HeLa clone cells (data not shown).

A Region Containing the Acidic-rich Domain in LANA Is Required for Binding with Daxx—To determine the interacting domain of LANA with Daxx, we constructed a series of LANA deletion mutants (Fig. 3*A*), which were translated *in vitro* and subjected to pull-down assay with GST-Daxx. As shown in Fig. 3*B*, full-length LANA was pulled down with GST-fused full-length Daxx, indicating direct interaction between LANA and Daxx. Three N-terminal mutants of LANA (L1–L3) bound with GST-Daxx, but the shortest N-terminal LANA (aa 1–261) (L4), and C-terminal LANA (aa 496–740) (L5) failed (Fig. 3*B*). We constructed mammalian expression plasmids, LANA-N (aa 1–564), LANA-C (aa 496–1162), LANA-N1 (aa 1–260), LANA-N2 (aa 1–320), LANA-N3 (aa 1–344), and LANA-ΔAD (with aa 322–493 deleted) (Fig. 3*A*). These plasmids were co-transfected with pcDNA-Daxx into 293T cells, and the nuclear extracts were analyzed. Immunoprecipitation with anti-Daxx antibody and Western blotting with anti-FLAG antibody indicated that Daxx formed a complex with full-length LANA and LANA-N, and weakly with LANA-N3, but not with the other LANA fragments (Fig. 3*C*). Taken together, these results suggested that aa 320–344 of LANA, which contains many aspartic acids and glutamic acids, were required for binding with Daxx.

A Central Domain of Daxx Is Required to Interact with LANA—To determine the critical region of Daxx for binding with LANA, a series of GST-fused deletion mutants of Daxx (Fig. 4*A*)

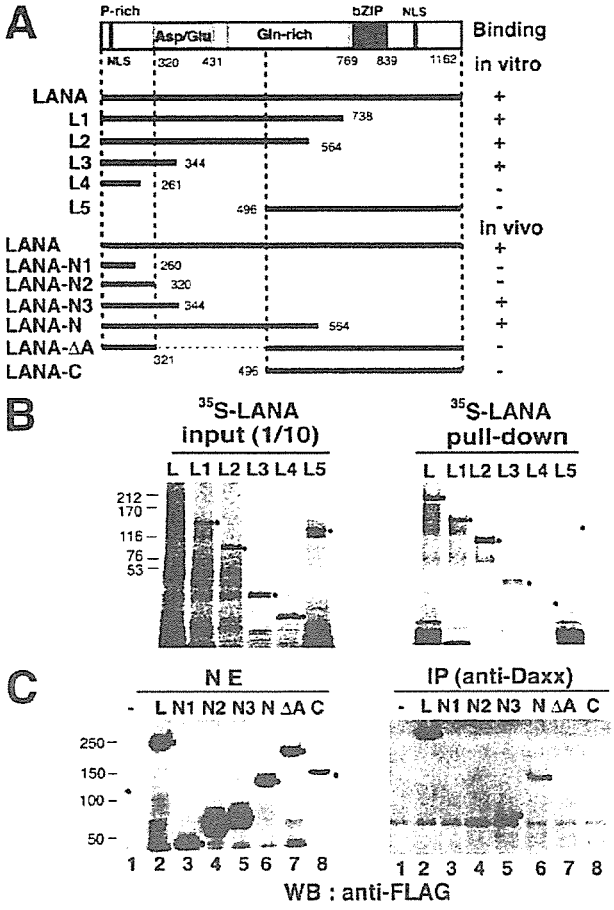

FIGURE 3. **A region containing an acidic-rich domain in LANA is required for binding with Daxx *in vitro* and *in vivo*.** *A*, domain structure of LANA and its deletion mutants. LANA is constituted of domains of proline-rich, acidic-rich, glutamine-rich, and basic leucine-zipper. A series of deletion mutants of LANA and the binding activity *in vitro* and *in vivo* are shown. *B*, result of pull-down assay with GST-fused full-length Daxx of ^{35}S-labeled LANA deletion mutants (*L1–L5*). *C*, co-immunoprecipitation of Daxx and LANA deletion mutants in 293T cells. PFLAG-CMV-2 vector (4.0 μg) (*lane 1*), pFLAG-LANA (4.0 μg) (*lane 2*), pFLAG-LANA-N1 (1.0 μg) (*lane 3*), pFLAG-LANA-N2 (2.0 μg) (*lane 4*), pFLAG-LANA-N3 (2.0 μg) (*lane 5*), pFLAG-LANA-N (2.0 μg) (*lane 6*), pFLAG-LANA-ΔA (4.0 μg) (*lane 7*), or pFLAG-LANA-C (4.0 μg) (*lane 8*) was individually co-transfected with pcDNA-Daxx (1.0 μg) in 60-mm dishes with adjustment of total DNA amount (5.0 μg). The immunoprecipitates (*IP*) with anti-Daxx antibody were followed by immunoblotting (*WB*) with anti-FLAG antibody (M5).

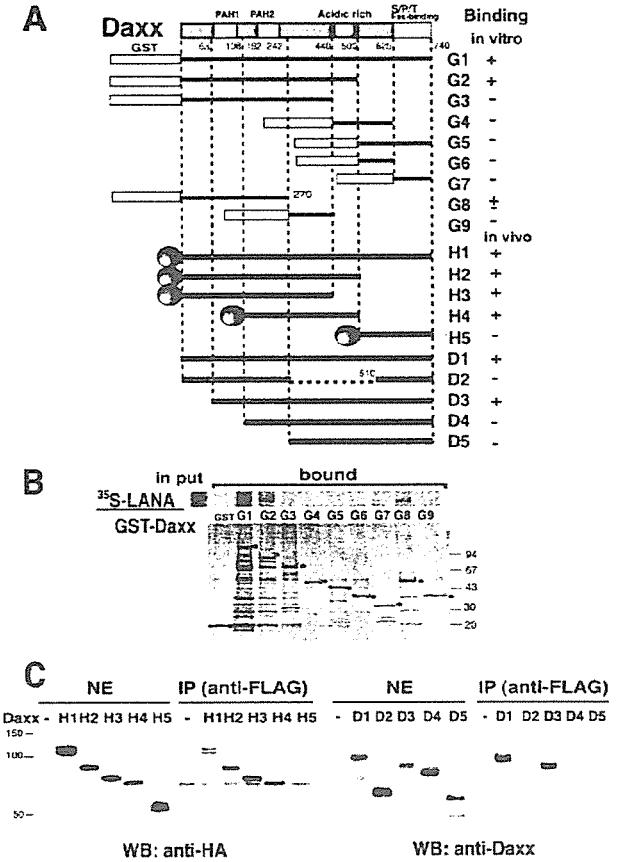

FIGURE 4. *A*, central region containing PAH 2 and acidic-rich domain in Daxx is required to interact with LANA. *A*, domain structure of Daxx and various deletion mutants. Daxx is composed of two PAH and acidic-rich and Ser/Pro/Thr-rich domains. A series of mutants of Daxx and the binding activity *in vitro* and *in vivo* are shown. *B*, purified GST-Daxx variants (*G1–G9*) were applied in *in vitro* pull-down assay with full-length ^{35}S-LANA. *C*, mammalian expression plasmids, pCMV-HA-Daxx-H1 (full-length) (2.0 μg), pCMV-HA-Daxx-H2 (aa 1–500) (2.0 μg), pCMV-HA-Daxx-H3 (aa 1–440) (1.0 μg), pCMV-HA-Daxx-H4 (aa 110–500)(1.0 μg), pCMV-HA-Daxx-H5 (aa 500–740) (1.0 μg) were co-transfected with pFLAG-LANA (1.0 μg) (*left two panels*). PcDNA-Daxx-D1 (full-length) (1.0 μg), pcDNA-Daxx-D2 (deleted aa 271–509) (3.0 μg), pcDNA-Daxx-D3 (aa 63–740)(3.0 μg), pcDNA-Daxx-D4 (aa 111–740)(3.0 μg), and pcDNA-Daxx-D5 (aa 243–740) (2.0 μg) were individually co-transfected with pFLAG-LANA (1.0 μg) (*right two panels*). Immunoprecipitates (*IP*) with anti-FLAG antibody (M2) were followed by Western blotting (*WB*) with anti-HA antibody (*left panels*) or anti-Daxx antibody (*right panels*).

were produced in *E. coli*, and applied to pull-down assay with full-length ^{35}S-labeled LANA. The GST-fused full-length Daxx (G1) and the Daxx-deleted aa 500–740 (G2) bound to LANA, but deleted aa 440–740 (G3) failed (Fig. 4*B*). From the *in vitro* result above, the region of aa 440–500 in Daxx was thought to be critical for the binding. However, GST-fused aa 440–625 of Daxx (G4) did not bind (Fig. 4*B*), nor did any other mutants, although weak binding was observed with GST-fused aa 1–270 of Daxx (G8)(Fig. 4*B*). We constructed a series of deletion mutants of N-terminal HA-tagged Daxx (H2–H5), and co-expressed them with pFLAG-LANA in 293T cells. Immunoprecipitation with anti-FLAG antibody followed by Western blotting with anti-HA antibody showed that all the mutants except H5 bound to LANA (Fig. 4*C*, *left two panels*). The acidic-rich region (aa 440–500) of Daxx was not critical to the binding with LANA in cells, not corresponding with the results *in vitro*. To examine contribution of N terminus of Daxx, a series of mutant Daxx expression vectors with N-terminal deletion (D3–D5) and a deletion mutant without central region aa 271–509 (D2), were constructed and transiently expressed in 293T cells. Experiments of immunoprecipitation with anti-FLAG antibody and Western blotting with anti-Daxx antibody (sc-7152, that recognizes the C terminus of Daxx) showed that D3 bound firmly with LANA, but D4 did very little (Fig. 4*C*, *right two panels*). The first paired amphipathic helix (PAH), aa 63–108 appeared to be of some importance for the binding, although HA-tagged Daxx without PAH1 (H4) bound LANA. These results indicated that a central region aa 63–440 within Daxx,

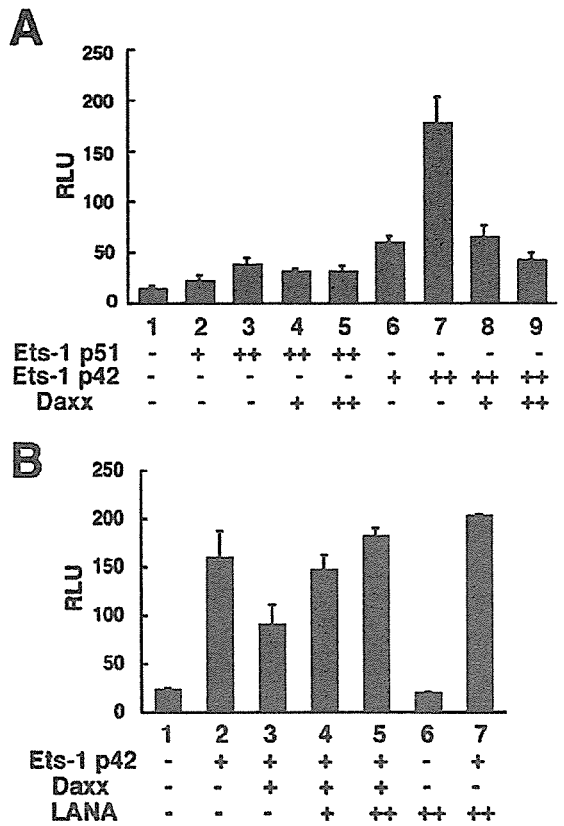

FIGURE 5. **LANA inhibited Daxx-mediated repression on Ets-1-dependent VEGF receptor 1 (Flt-1) gene expression.** *A*, Daxx repressed Ets-1-dependent Flt-1 expression. PcDNA-p51Ets-1 or pcDNA-p42Ets-1 (+; 25 ng, ++; 50 ng) were co-transfected with pcDNA-Daxx (+; 200 ng, ++; 500 ng) and pFlt-luc (100 ng). *B*, LANA counteracts Daxx-mediated repression in Flt-1 expression in the presence or absence of exogenous Daxx. PcDNA-p42Ets-1 (50 ng), pcDNA-Daxx (200 ng), and pFLAG-LANA (0, +; 50, ++; 100 ng, respectively) were co-transfected with pFlt-1-luc (100 ng). The relative luciferase activity (*RLU*) was normalized by β-galactosidase activity. Assays were performed in triplicate, and *error bars* indicate S.D.

containing two paired PAHs and its following 200 aa, was important for the binding with LANA in cells.

LANA Inhibited Daxx-mediated Repression of Ets-1-dependent VEGF Receptor 1 (Flt-1) Gene Expression— To examine the role of Daxx in Kaposi's sarcoma, we focused on Ets-1 transcription factor. It was reported that Daxx interacts with Ets-1 to repress Ets-1-dependent transcriptional activity of MMP-1 and Bcl-2 (12). On the other hand, as a characteristic feature of KS, it is known that VEGF and its receptors, Flt-1 and KDR (VEGF receptor-1 and -2, respectively), are highly expressed in KS (20). There are several Ets-1 motifs in Flt-1 and KDR promoters to regulate the expression (26) (27). We examined the effect of Daxx on Ets-1-dependent Flt-1 expression. We co-transfected a luciferase reporter plasmid pFlt-1-luc driven by Flt-1 promoter, an Ets-1 expression vector, and a Daxx expression vector into 293T cells, to perform luciferase assay. Transcriptional activity on Flt-1 increased depending on the amount of Ets-1 plasmid, although the effect of p51-Ets-1 was quite weak. Daxx evidently repressed Ets-1-dependent activation (Fig. 5*A*). p51 and p42 are two human variants of the Ets-1 molecule. It is reasonable that the activity of p51-Ets-1 is lower than p42-Ets-1 because p42-Ets-1 lacks exon VII, the internal transcriptional regulatory domain (24). This result is similar to the case of MMP-1 and Bcl-2 expression (28). As we observed the repressive activity of Daxx on Ets-1-dependent Flt-1 expression, we examined the effect of LANA on the Daxx-mediated repression with p42Ets-1. Co-transfection with a LANA expression plasmid dose-dependently reactivated the transcriptional activity repressed by exogenous Daxx (Fig. 5*B*, *4* and *5*), although LANA slightly activated it in the absence of exogenous Daxx (Fig. 5*B*, *7*). These results suggested that LANA inhibited the repression via interaction with Daxx.

LANA Activated Expression of VEGF Receptors in Vascular Endothelial Cells—To investigate the possibility that LANA induces Flt-1 in Kaposi's sarcoma lesion, we tried to express LANA in HUVEC, because endothelial cells (ECs) are regarded as the origins of KS lesions. We constructed a plasmid, pIRES2-LANA-GFP, which contains an internal ribosomal entry site (IRES) to express both LANA and GFP from a single mRNA. We transfected pIRES2-LANA-GFP or pIRES2-GFP as control into HUVEC and Flt-1 and KDR expression in GFP-positive cells were analyzed by flow cytometry. Flt-1 of GFP-positive cells in pIRES2-LANA-GFP-transfected cells was significantly increased as compared with that in control cells (Fig. 6*A*, *left*). The number of cells expressing Flt-1 over log intensity 1 (M1) was about 1.9× higher (Fig. 6*A*, *upper, right graph*) than that of control. M1 of KDR also increased 1.4× (Fig. 6*A*, *lower, right graph*). Furthermore, to examine the level of mRNA of the two receptors, we performed real-time PCR with total RNA prepared from the GFP-expressing HUVEC. LANA expression in pIRES2-LANA-GFP-transfected cells was confirmed by using PCR with primers of LANA (data not shown). The relative expressions of Flt-1 and KDR in LANA-expressing cells were 1.4 and 2.0× higher than that of control cells, respectively (Fig. 6*B*). Although there was discrepancy between rise of protein and mRNA, results of both FACS and real-time PCR indicated that LANA induced the two receptors in human endothelial cells. The expression of Ets-1 and Daxx was not altered between LANA-expressing cells and control cells (Fig. 6*B*).

LANA Sequesters Daxx from Ets-1—To resolve the mechanism of the activation of VEGF receptors expression by LANA, we examined the relation of the three molecules, LANA, Daxx, and Ets-1. 293T cells were co-transfected with a constant amount of pcDNA-Daxx and pcDNA-Ets-1, and a variable amount of pFLAG-LANA. Nuclear extracts were prepared and subjected to immunoprecipitation and Western blotting with anti-Ets-1 antibody, anti-Daxx antibody or anti-FLAG antibody. Daxx and Ets-1 were expressed in a fixed amount (Fig. 7*A*, *row a*), *middle and right panel*, respectively) and FLAG-LANA was dose-dependently increased in the nuclear extract (Fig. 7*A*, *row a, left panel*). When we performed immunoprecipitation with anti-FLAG antibody, Daxx was detected in the immune complex in proportion to the amount of LANA (Fig. 7*A*, *row b, middle panel*). On the other hand, we detected no specific interaction between LANA and Ets-1 in the immune complex (Fig. 7*A*, *row b, right panel*). Next, by immunoprecipitation with anti-Daxx antibody, FLAG-LANA was detected in direct proportion to the amount of LANA (Fig. 7*A*, *row c, left panel*). The immune complex also contained Ets-1 in inverse

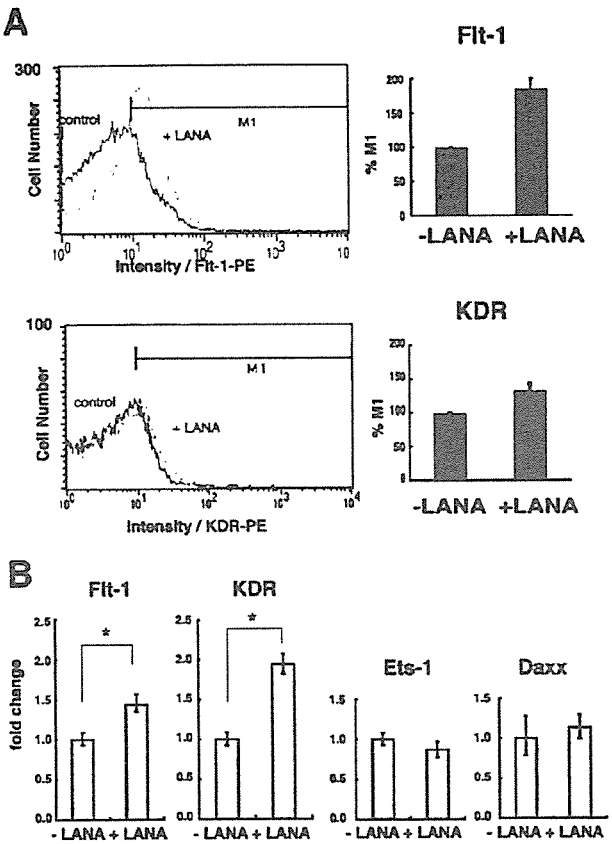

FIGURE 6. **LANA induced VEGF receptors in HUVEC.** *A*, flow cytometric analysis of Flt-1 (*upper graphs*) and KDR (*lower graphs*) expression of control (pIRES2-GFP transfected cells; *black lines*) and LANA-expressing cells (pIRES2-LANA-GFP-transfected cells; *gray lines*). The graphs to the *right* of each indicate percentages of cells that exceed 1 of the relative log intensity (*M1*). Experiments were repeated three times and the M1 values represent means of the three experiments; *error bars* indicate S.D. *B*, real-time PCR analysis of Flt-1, KDR, Ets-1, and Daxx. HUVECs transfected transiently with pIRES2-LANA-GFP or pIRES2-GFP (as a control) were sorted 2 days after transfection. Total RNA extracted from the cells (1 μg) was reverse-transcribed to cDNA (40 μl), and aliquots (0.4 μl) were applied to real-time PCR (20 μl) with each primer (0.4 mM) in triplicate described under "Experimental Procedures." Values represented relative expression of Flt-1, KDR, Ets-1, and Daxx (calculated with threshold cycle number, CT) of LANA-expressing cells compared with that of control cells. Each value was adjusted with CT of internal control (GAPDH). *, p value < 0.02.

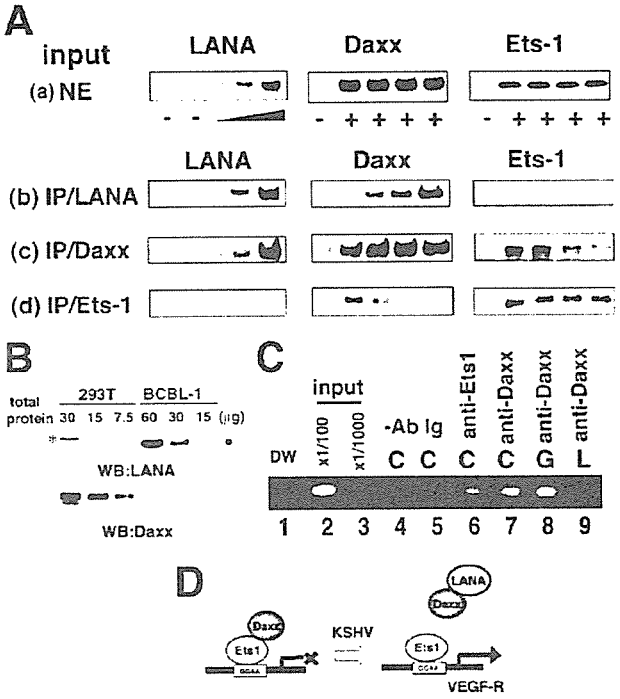

FIGURE 7. **LANA interacted with Daxx to sequester from Ets-1.** *A*, Western blotting analysis of immunoprecipitates with anti-FLAG, anti-Daxx, and anti-Ets-1 antibodies. 293T cells were transfected with a constant amount of pcDNA-Daxx (2 μg) and pcDNA-p42Ets-1 (2 μg), and an increasing amount of pFLAG-LANA (0.25, 0.5, 1.0 μg). Total DNA amounts were adjusted with pFLAG-CMV-2 vector to be 5 μg. Nuclear extracts (*row a*), immune complex using anti-FLAG antibody (*row b*), immune complex using anti-Daxx antibody (*row c*), and immune complex using anti-Ets-1 antibody (*row d*), were subject to Western blotting with anti-FLAG antibody (*left*), anti-Daxx antibody (*middle*) or anti-Ets-1 antibody (*right*). *B*, relative protein amounts of LANA and Daxx in BCBL-1 cells and those of transfected 293T cells. 293T cells were co-transfected with pcDNA-Daxx (2 μg), pcDNA-p42Ets-1 (2 μg), and pFLAG-LANA (1.0 μg). Nuclear extract (30, 15, 7.5 μg of the 293T cells and 60, 30, 15 μg of BCBL-1 cells) were subjected to Western blotting with anti-LANA antibody or anti-Daxx antibody. FLAG-LANA (*) migrated slower than native LANA(●) did. *C*, chromatin immunoprecipitation of Ets-1 and Daxx interaction with Flt-1 promoter in HUVECs. Bands indicate PCR products targeting -90 to $+8$ of Flt-1 promoter. 2 μl of water (*lane 1*), 1/100 and 1/1000 of input (cross-linked and sonicated pre-immunoprecipitation lysate) (*lanes 2* and *3*), eluate from no antibody (*lane 4*), rabbit IgG (2 μg) (*lane 5*), anti-Ets-1 antibody (2 μg) (*lane 6*), and anti-Daxx antibody (2 μg) (*lane 7*) were applied to the PCR reaction, respectively. Eluate from anti-Daxx antibody of LANA-expressing HUVECs (*L*, *lane 9*) and that from the control GFP-expressing HUVECs (*G*, *lane 8*) were subjected to PCR reaction. *D*, possible mechanism for induction of VEGF receptors by LANA. Daxx interacts with Ets-1, and represses Ets-1-dependent expression in the absence of LANA, while LANA sequesters Daxx from Ets-1 to inhibit the interaction between Daxx and Ets-1, resulting in activation of Ets-1-dependent expression of VEGF receptors.

proportion to LANA expression (Fig. 7*A*, *row c*, *right panel*). Consistently, Daxx was detected in the immune complex with anti-Ets-1 antibody in inverse proportion to LANA expression (Fig. 7*A*, *row d*, *middle panel*). LANA was not detected in the immune complex with the anti-Ets-1 antibody (Fig. 7*A*, *row d*, *left panel*), which implies that increasing LANA caused increase of Daxx-LANA interaction, and reduction of Daxx-Ets-1 interaction. These results suggested that LANA sequesters Daxx from Ets-1, which results in inhibition of the interaction between Daxx and Ets-1.

In the experiments above we used transiently transfected 293T cells (Fig. 7*A*). To address whether the transient expression system for LANA-Daxx interaction is physiologically relevant or not, we analyzed relative expression levels of LANA and Daxx proteins using BCBL-1 and the transfected 293T cells. As shown in Fig. 7*B*, the expression level of exogenous LANA protein in 293T cells in the same condition of Fig. 7*A* was similar to that of endogenous LANA in BCBL-1 cells. In contrast, endogenous Daxx expression level is much lower in BCBL-1 cells than in the 293T cells. These data indicated that relative expression ratio of endogenous LANA to Daxx in BCBL-1 cells was much higher than that of LANA-transfected 293T cells.

Daxx associated with Flt-1 promoter and LANA reduced its association in HUVEC. To investigate the possibility that Daxx affects transcriptional activity of Ets-1 for Flt-1 expression in endothelial cells (ECs), we performed ChIP assay using HUVEC. Cross-linked nuclear extract from HUVECs was immunoprecipitated with anti-Ets-1 antibody or anti-Daxx

antibody, and subjected to PCR to amplify a 98-bp fragment. The PCR product is designed to span the fourth ets motif (−54 to −51) that is thought to be indispensable for Flt-1 promoter activity (26). The anti-Daxx antibody precipitated the Flt-1 promoter as well as anti-Ets-1 antibody (Fig. 7*C*, *lanes 6* and *7*). The result indicated that Daxx as well as Ets-1 associated with Flt-1 promoter in ECs. Furthermore, HUVECs transfected with pIRES2-LANA-GFP or pIRES2-GFP as the control were sorted and subjected to ChIP assay with anti-Daxx antibody. The PCR product from the LANA-expressing cells (Fig. 7*C*, *L*, *lane 9*) was lower than that of control GFP-expressing cells (Fig. 7*C*, *G*, *lane 8*), indicating reduction of Daxx association with the promoter of the *Flt-1* gene in ECs.

DISCUSSION

LANA is reported to have multiple functions in KS lesion. It interacts with many host cellular molecules: p53 (8), pRb (9), ATF4/CREB2 (29), CBP (30), c-Jun (31), RING3 (32), mSin3A (33), HP-1 (33), Dek (34), GSK-3b (10) and so on. In the present study, we identified Daxx as a new member of LANA-binding proteins. Daxx was prominently detected in our immunoaffinity system, but this system also detected previously reported LANA-interacting protein such as RING3 by Western blotting (data not shown). We showed the interaction between the two proteins *in vivo* (Fig. 1) and *in vitro* (Figs. 3 and 4), which indicates that Daxx and LANA directly bound to each other. Fluorescent immunostaining assay showed co-localization of LANA and Daxx in BCBL-1 cells, supporting LANA-Daxx interaction in cells (Fig. 2).

Daxx is reported to bind many cellular molecules, indicating its involvement in multiple cellular processes. Although Daxx could interact with proteins of cytoplasm or membrane, it also interacted with some transcription factors and localized sometimes in the nuclear matrix structure, PML NBs (promyelocytic leukemia nuclear bodies). PML NBs are thought to provide platforms for transcription regulation, DNA repair, apoptosis, DNA replication, RNA transport, and many viruses target PML NBs to pirate host functions (reviewed by Everett, Ref. 35). Ets-1 associates with a PML NBs protein, Sp100 (36). Therefore, it might be a strategy of KSHV that LANA targets Daxx of PML NBs to modulate the cellular function(s) of Ets-1.

Although most LANA-binding proteins are reported to interact through the C or N terminus of LANA, the critical domain for binding with Daxx seemed to be a central region, aa 321–344 of LANA (Fig. 3). The aa 320–431 of LANA consists mainly of aspartic acid and glutamic acid. It is reported that a transcriptional co-activator, CBP interacts through this acidic-rich region of LANA (30). This domain may have some roles in transcriptional regulation. On the other hand, although most Daxx-binding proteins interact around the C terminus of Daxx, a central domain containing the PAHs and the following region of Daxx appeared to be important for binding with LANA *in vivo* (Fig. 4). There was a discrepancy between *in vitro* and *in vivo* binding. Protein modification may be one possibility explaining *in vivo* binding activity. It is reported that Daxx is modified by hyperphosphorylation (13) and sumoylation (37). Because the sumoylation sites of Daxx are reported to be Lys^{630} and Lys^{631}, it is unlikely to affect the interaction. The hyperphosphorylation site on Daxx has not been identified, but it is possible to be related to the binding. There may be other possibilities, for example, constructive interference by fused GST protein. PAH is a characteristic domain that is involved in transcriptional co-repressors such as mSin3 (38). It is interesting that mSin3A binds to aa 1–340 of LANA (28). There is a report that acetylated histone H4 interacts through PAH1 within Daxx, but no report that any other host molecule binds through this region of Daxx. As Daxx interacts with Ets-1 through the C-terminal region of Daxx (12), there may be no direct competition for Daxx between LANA and Ets-1.

Based on the interaction between LANA and Daxx (Figs. 1–4), we found that LANA induced VEGF receptors in ECs (Fig. 6) in accordance with the results of reporter assays (Fig. 5). Although expression level changes were not consistent for Flt-1 and KDR in protein (Fig. 6*A*) and mRNA (Fig. 6*B*), it may be caused by time point difference. This is the first report of the function of LANA in angiogenesis. It is reported that KSHV ORF74 (viral G-protein coupled receptor, v-GPCR) contributes to expression of VEGF receptors (39). Because ORF74 is expressed in the viral lytic infection cycle, it is unlikely that ORF74 is the only gene of KHSV that induces angiogenesis in KS. It is likely that some other factors such as VEGF and hypoxia-inducible factor (HIF) additionally affect on expression of these receptors in KS (40) (41).

As to the mechanism of activation of the receptor expression by LANA, we propose a hypothesis that LANA sequesters Daxx from Ets-1 (Fig. 7*D*), based on the results of co-immunoprecipitation (Fig. 7*A*) and ChIP assay (Fig. 7*C*). LANA slightly activated Ets-1 dependent Flt-1 expression without exogenous Daxx in the reporter assay (Fig. 5*B*). It is thought that LANA sequestered endogenous Daxx. However it is possible that LANA activates Flt-1 expression through an unidentified mechanism(s). At least LANA did not activate Flt-1 expression through up-regulation of Ets-1 expression (Fig. 6*B*). In human Flt-1 promoter, there are five Ets motifs and a CRE (cAMP response element). It is reported that co-existence of the fourth Ets motif, and the CRE is necessary for Flt-1 expression (26). LANA is reported to modulate the expression of a reporter plasmid with CRE, but the effect of LANA on CRE is repression (29). There is no CRE in the promoter of KDR.

Given that LANA induces VEGF receptors in KS lesion, we propose this hypothesis: Daxx binds Ets-1 to repress expression of VEGF receptors in normal ECs, while in KSHV-infected cells, LANA binds to Daxx to inhibit Daxx-Ets-1 interaction, resulting in the activation of Ets-1-dependent VEGF receptors. Furthermore, LANA-Daxx interaction might contribute to not only VEGF receptor gene expression but also to other Daxx-mediated gene regulation related to the pathogenesis of KS, PEL, and MCD malignancy.

Acknowledgments—We thank Dr. Kaoru Morishita (Daiichi Pharmaceutical Co., Ltd., Tokyo), and Dr. Runzhao Li (Medical University of South Carolina) for kindly providing plasmids. We thank Dr. Harutaka Katano (Department of Pathology, National Institute of Infectious Diseases) for providing BCBL-1 cells and useful advice, Dr. Kazuo Suzuki (Department of Bioactive Molecules, National Institute of Infectious Diseases) for useful discussion, and Eri Watanabe, Junko Kondo, and Yuki Hashimoto for their technical assistance.

REFERENCES

1. Flore, O., Rafii, S., Ely, S., O'Leary, J. J., Hyjek, E. M., and Cesarman, E. (1998) *Nature* **394,** 588–592
2. Moore, P. S., and Chang, Y. (1998) *Am. J. Epidemiol.* **147,** 217–221
3. Dupin, N., Fisher, C., Kellam, P., Ariad, S., Tulliez, M., Franck, N., van Marck, E., Salmon, D., Gorin, I., Escande, J. P., Weiss, R. A., Alitalo, K., and Boshoff, C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4546–4551
4. Russo, J. J., Bohenzky, R. A., Chien, M. C., Chen, J., Yan, M., Maddalena, D., Parry, J. P., Peruzzi, D., Edelman, I. S., Chang, Y., and Moore, P. S. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 14862–14867
5. Katano, H., Sato, Y., Kurata, T., Mori, S., and Sata, T. (1999) *Am. J. Pathol.* **155,** 47–52
6. Ballestas, M. E., Chatis, P. A., and Kaye, K. M. (1999) *Science* **284,** 641–644
7. Garber, A. C., Shu, M. A., Hu, J., and Renne, R. (2001) *J. Virol.* **75,** 7882–7892
8. Friborg, J., Jr., Kong, W., Hottiger, M. O., and Nabel, G. J. (1999) *Nature* **402,** 889–894
9. Radkov, S. A., Kellam, P., and Boshoff, C. (2000) *Nat. Med.* **6,** 1121–1127
10. Fujimuro, M., Wu, F. Y., ApRhys, C., Kajumbula, H., Young, D. B., Hayward, G. S., and Hayward, S. D. (2003) *Nat. Med.* **9,** 300–306
11. Yang, X., Khosravi-Far, R., Chang, H. Y., and Baltimore, D. (1997) *Cell* **89,** 1067–1076
12. Li, R., Pei, H., Watson, D. K., and Papas, T. S. (2000) *Oncogene* **19,** 745–753
13. Hollenbach, A. D., Sublett, J. E., McPherson, C. J., and Grosveld, G. (1999) *EMBO J.* **18,** 3702–3711
14. Emelyanov, A. V., Kovac, C. R., Sepulveda, M. A., and Birshtein, B. K. (2002) *J. Biol. Chem.* **277,** 11156–11164
15. Kim, E. J., Park, J. S., and Um, S. J. (2003) *Nucleic Acids Res.* **31,** 5356–5367
16. Oikawa, T., and Yamada, T. (2003) *Gene (Amst.)* **303,** 11–34
17. Sacchi, N., de Klein, A., Showalter, S. D., Bigi, G., and Papas, T. S. (1988) *Leukemia* **2,** 12–18
18. Wernert, N., Raes, M. B., Lassalle, P., Dehouck, M. P., Gosselin, B., Vandenbunder, B., and Stehelin, D. (1992) *Am. J. Pathol.* **140,** 119–127
19. Masood, R., Cai, J., Zheng, T., Smith, D. L., Naidu, Y., and Gill, P. S. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 979–984
20. Masood, R., Cesarman, E., Smith, D. L., Gill, P. S., and Flore, O. (2002) *Am. J. Pathol.* **160,** 23–29
21. Sato, Y., Kanno, S., Oda, N., Abe, M., Ito, M., Shitara, K., and Shibuya, M. (2000) *Ann. New York Acad. Sci.* **902,** 201–205; discussion 205–207
22. Tang, D. G., and Conti, C. J. (2004) *Semin. Thromb. Hemost.* **30,** 109–117
23. Morishita, K., Johnson, D. E., and Williams, L. T. (1995) *J. Biol. Chem.* **270,** 27948–27953
24. Koizumi, S., Fisher, R. J., Fujiwara, S., Jorcyk, C., Bhat, N. K., Seth, A., and Papas, T. S. (1990) *Oncogene* **5,** 675–681
25. Yamagoe, S., Kanno, T., Kanno, Y., Sasaki, S., Siegel, R. M., Lenardo, M. J., Humphrey, G., Wang, Y., Nakatani, Y., Howard, B. H., Ozato, K., Xu, R. H., Peng, Y., Fan, J., Yan, D., Princler, G., Sredni, D., and Kung, H. F. (2003) *Mol. Cell. Biol.* **23,** 1025–1033
26. Wakiya, K., Begue, A., Stehelin, D., and Shibuya, M. (1996) *J. Biol. Chem.* **271,** 30823–30828
27. Kappel, A., Schlaeger, T. M., Flamme, I., Orkin, S. H., Risau, W., and Breier, G. (2000) *Blood* **96,** 3078–3085
28. Krithivas, A., Young, D. B., Liao, G., Greene, D., and Hayward, S. D. (2000) *J. Gen. Virol.* **74,** 9637–9645
29. Lim, C., Sohn, H., Gwack, Y., and Choe, J. (2000) *J. Gen. Virol.* **81,** 2645–2652
30. Lim, C., Gwack, Y., Hwang, S., Kim, S., and Choe, J. (2001) *J. Biol. Chem.* **276,** 31016–31022
31. An, J., Sun, Y., and Rettig, M. B. (2004) *Blood* **103,** 222–228
32. Platt, G. M., Simpson, G. R., Mittnacht, S., and Schulz, T. F. (1999) *J. Virol.* **73,** 9789–9795
33. Lim, C., Lee, D., Seo, T., Choi, C., and Choe, J. (2003) *J. Biol. Chem.* **278,** 7397–7405
34. Krithivas, A., Fujimuro, M., Weidner, M., Young, D. B., and Hayward, S. D. (2002) *J. Virol.* **76,** 11596–11604
35. Everett, R. D. (2001) *Oncogene* **20,** 7266–7273
36. Wasylyk, C., Schlumberger, S. E., Criqui-Filipe, P., and Wasylyk, B. (2002) *Mol. Cell. Biol.* **22,** 2687–2702
37. Jang, M. S., Ryu, S. W., and Kim, E. (2002) *Biochem. Biophys. Res. Commun.* **295,** 495–500
38. Spronk, C. A., Tessari, M., Kaan, A. M., Jansen, J. F., Vermeulen, M., Stunnenberg, H. G., and Vuister, G. W. (2000) *Nat. Struct. Biol.* **7,** 1100–1104
39. Bais, C., VanGeelen, A., Eroles, P., Mutlu, A., Chiozzini, C., Dias, S., Silverstein, R. L., Rafii, S., and Mesri, E. A. (2003) *Cancer Cell* **3,** 131–143
40. Chen, Z., Fisher, R. J., Riggs, C. W., Rhim, J. S., and Lautenberger, J. A. (1997) *Cancer Res.* **57,** 2013–2019
41. Elvert, G., Kappel, A., Heidenreich, R., Englmeier, U., Lanz, S., Acker, T., Rauter, M., Plate, K., Sieweke, M., Breier, G., and Flamme, I. (2003) *J. Biol. Chem.* **278,** 7520–7530

Identification of a Novel Circulating Recombinant Form (CRF33_01B) Disseminating Widely Among Various Risk Populations in Kuala Lumpur, Malaysia

Kok Keng Tee, MMedSc,† Xiao-Jie Li, MD, PhD,* Kyoko Nohtomi, BSc,* Kee Peng Ng, MBBS, PhD,‡ Adeeba Kamarulzaman, MBBS, FRACP,† and Yutaka Takebe, MD, PhD**

Summary: A molecular epidemiological investigation was conducted among various risk populations (n = 184) in Kuala Lumpur, Malaysia, in 2003 to 2005, on the basis of nucleotide sequences of protease and reverse transcriptase regions. In addition to circulating HIV-1 strains, including CRF01_AE (57.1%), subtype B (20.1%), and subtype C (0.5%), we detected a candidate with a new circulating recombinant form (CRF). We determined four near-full-length nucleotide sequences with identical subtype structure from epidemiologically unlinked individuals of different risk and ethnic groups. In this chimera, two short subtype B segments were inserted into the *gag*-RT region in a backbone of CRF01_AE. The recombinant structure was distinct from previously identified CRF15_01B in Thailand. In agreement with the current HIV nomenclature system, this constitutes a novel CRF (CRF33_01B). The overall prevalence of CRF33_01B is 19.0% (35/184). Although the prevalence of CRF33_01B is particularly high among injecting drug users (42.0%, 21/50), it is also detected in a substantial proportion of homo-/bisexual males (18.8%, 3/16) and heterosexuals (9.8%, 9/92). Moreover, unique recombinant forms composed of CRF01_AE and subtype B that have a significant structural relationship with CRF33_01B were detected in 1.6% (3/184) of study subjects, suggesting an ongoing recombination process in Malaysia. This new CRF seems to be bridging viral transmission between different risk populations in this country.

Key Words: HIV-1 circulating recombinant form (CRF), unique recombinant form (URF), CRF01_AE, subtype B, CRF33_01B, Malaysia, Asia

(*J Acquir Immune Defic Syndr* 2006;43:523–529)

Received for publication May 10, 2006; accepted August 7, 2006.

From the *Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, Shinjuku-ku, Tokyo, Japan; †Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia; and ‡Department of Medical Microbiology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia.

Funded by grants from the Ministry of Health, Labour and Welfare, the Ministry of Education, Science and Technology, and the Japanese Foundation for AIDS Prevention (JFAP).

Reprints: Yutaka Takebe, MD, PhD, Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan (e-mail: takebe@nih.go.jp).

Since the first case of AIDS was reported in Malaysia in 1986,[1] a total of 67,438 HIV infections had been identified nationwide by the Ministry of Health by June 2005. Although the adult HIV prevalence rate remains low (0.5%:0.2%–1.5%), elements exist that could cause the epidemic to erupt suddenly. In 2004 alone, an average of 18 new cases of HIV infections was reported daily. The Joint United Nations Program on HIV/AIDS (UNAIDS) estimated that approximately 69,000 people in Malaysia were living with HIV in 2005.[2] The main route of HIV transmission in Malaysia is through injecting drug use, which accounts for 75% of all reported HIV cases (Fig. 1, inset). The growing proportion of HIV cases attributed to sexual transmission (17% in 2002 compared with 11.5% in 2001 and 7% in 1995; 46% increase between 2001 and 2002), suggesting that HIV is spreading in the general population.[3,4]

HIV-1 exhibits a tremendous genetic diversity that is driven by high rates of mutation (3×30^{-5} sites/genome/replication cycle) and recombination (2–3 crossovers/replication cycle), coupled with high viral turnovers ($>10^9$ per day) and persistent nature of infections (300 replication cycles/year).[5] By these mechanisms, HIV-1 group M, which is largely responsible for the global pandemic, has diversified into 11 subtypes and sub-subtypes (A1, A2, B, C, D, F1, F2, G, H, J, and K) and various types of recombinants (http://hiv-web.lanl.gov/). HIV-1 recombinants with epidemic spread are known as circulating recombinant forms (CRFs). Thirty-two CRFs are currently recognized.[6] Four CRFs have been reported so far in Asia: CRF01_AE and CRF15_01B in Thailand and CRF07_BC and CRF08_BC in China. To define a CRF, at least three epidemiologically unlinked HIV-1 sequences with identical mosaic structures should be characterized, at least two of them in near-full-length genomes (>8 kb).[7] In addition to CRFs, various types of unique recombinant forms (URFs) that were detected in a single individual or a single epidemiologically linked cluster have been identified in the region, where the multiple lineages of HIV-1 strains cocirculate in the same population.

In the early phase of the Thai epidemic, two HIV-1 strains—CRF01_AE and subtype B′ (Thai variant of subtype B)—were circulating relatively independently among different risk populations. CRF01_AE was distributed among persons at risk of sexual exposure, while subtype B′ was distributed mainly among injecting drug users (IDUs).[8,9] However, by 1999 it was reported that CRF01_AE accounted for >50% of new infections among IDUs.[10–12] Cocirculation of CRF01_AE

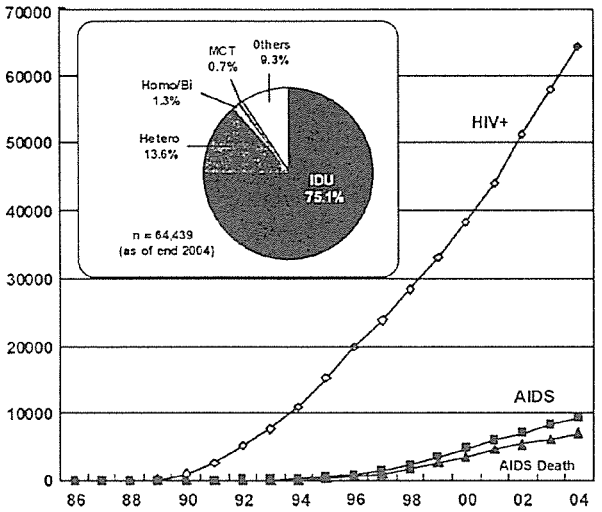

FIGURE 1. Status and distribution of HIV-1 infection in different risk categories in Malaysia. Cumulative numbers of HIV-1 infections, AIDS cases, and AIDS-related deaths reported by the Ministry of Health in Malaysia (1986–December 2004). Inset, Distribution of HIV-1 infections in different risk categories as of the end of 2004 (n = 64,439). Abbreviations: IDU, injecting drug users; Hetero, heterosexuals; Homo/Bi, male homo-/bisexual; MCT, mother-to-child transmission.

and subtype B′ in Thailand led to the generation of various forms of CRF01_AE/B′ recombinants,[13–18] including CRF15_01B.[19] A similar molecular epidemiological trend has been observed in Malaysia. Studies conducted in 1992 to 1997 showed that CRF01_AE and subtype B′ were prevalent among 81% of heterosexuals and 55% to 92% of IDUs, respectively.[20–22] However, more recent studies based on partial protease (Pro) and reverse transcriptase (RT) sequence data suggested the emergence of new forms of CRF01_AE/B recombinants in Malaysia.[23,24]

For the present study, we characterized near-full-length nucleotide sequences of CRF01_AE/B recombinants from eight epidemiologically unlinked individuals to define a novel CRF and other recombinant forms emerging in Malaysia and discuss herein their epidemiological and biological implications.

MATERIALS AND METHODS

Study Subjects and Specimens

All subjects (n = 184) were recruited between July 2003 and August 2005 in the HIV clinic of the University Malaya Medical Center (UMMC), Kuala Lumpur. This study was approved by the UMMC Medical Ethics Committee. Information on the patients' clinical and epidemiological backgrounds was collected from the HIV patient-management database. Study subjects included 50 IDUs, 92 heterosexuals, 14 male homosexuals, 2 male bisexuals, 4 persons who were infected from their mothers (mother-to-child transmission, MCT), and 22 cases with unknown risk factors. Subjects consisted of 150 adult males with an age range of 19 to 67 years old (mean: 38.2 ± 8.9 years old) and 31 adult females with an age range of 26 to 53 years old (mean: 36.0 ± 7.9 years old). Three MCT adolescents (2 males, 1 female) ranging in age from 5 to 9 years old (mean: 7.7 ± 2.3 years old) were also included. Specimens from these 184 patients were serologically determined to be HIV-1 positive. No HIV-2 infections were detected. HIV-1 genotypes were screened on the basis of the Pro and RT regions using plasma HIV-1 RNA, as described previously.[23,24]

Viral Isolation, Near-Full-Length DNA Amplification, and Sequencing

Peripheral-blood mononuclear cells (PBMCs) from selected patients were separated on Ficoll-Hypaque density gradient centrifugation (Amersham Biosciences AB, Uppsala, Sweden). For virus isolation, PBMCs from HIV-1–positive individuals were cocultured with phytohemagglutinin-stimulated (1 μg/ml) $CD8^+$ T-cell–depleted PBMCs (Miltenyi Biotec GmbH, Bergisch Gladbach, Germany) from HIV-negative healthy donors in RPMI 1640 containing 10% fetal calf serum and interleukin-2 (20 U/mL). Virus production was detected by a virion-associated RT assay as described previously.[25] HIV-infected PBMCs were harvested, and proviral DNA was isolated with guanidine detergent (Invitrogen, Carlsbad, CA). Near-full-length viral DNA was amplified with primers pbsA (5′-AGT GGC GCC CGA ACA GG-3′; nucleotide positions relative to HXB2: 634–650) and 9KU5B (5′-GGT CTG AGG GAT CTC TAG TTA CCA G-3′; nucleotide positions relative to HXB2: 9666–9690) by the Expand Long Template PCR System (Roche Diagnostic GmbH, Penzberg, Germany), gel-purified, and TA-cloned with the pCR-XL-TOPO vector (Invitrogen).[26,27] Positive clones were selected, and the near-full-length genome of ~9.1 kb was directly sequenced with an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA) using the primer-walking method. For plasma samples, HIV-1 RNA was reverse-transcribed and nested-PCR–amplified to produce 10 overlapping fragments; directly sequenced; and then assembled to generate the near-full-length genomes. Primers were carefully designed to obtain sufficient numbers of overlapping nucleotides to ensure that recombinants were not generated by assembling the sequence fragments derived from different HIV-1 subtypes that had been amplified from an individual (primer sets and PCR parameters are available upon request).

Phylogenetic Analysis

Nucleotide sequences were aligned manually using Se-Al, version 1.0,[28] with HIV-1 reference subtypes and CRFs from the HIV database (http://hiv-web.lanl.gov/). Phylogenetic trees were constructed by the neighbor-joining method[29] based on the Kimura two-parameter model with a transition-to-transversion ratio of 2.0.[30] The reliability of the branching orders was tested by bootstrap analysis of 100 replicates. Bootscanning and informative-site analyses were performed using SimPlot, version 3.5,[31] with a sliding window of 250 nucleotides overlapped by 50 nucleotides to define the recombinant structure. The origin of each segment was analyzed by subregion neighbor-joining tree analysis.

RESULTS

HIV-1 Genotype Distribution in Malaysia

HIV-1 genotypes circulating among various risk populations (n = 184) were determined by phylogenetic and recombination breakpoint analyses based on the nucleotide sequences of the Pro and RT regions. Distribution of HIV-1 genotypes in Malaysia is currently as follows: CRF01_AE (105/184, 57.1%), subtype B′ (37/184, 20.1%), and subtype C (1/184, 0.5%) (Fig. 2). A substantial proportion (41/184, 22.3%) of specimens showed unique subtype structure composed of CRF01_AE and subtype B.[23,24] The majority of them (35/184, 19.0%) harbored the identical subtype structure, and a small proportion of specimens (6/184, 3.3%) displayed different profiles of recombinant structure. We designate these two groups of Malaysian HIV-1 recombinant strains as the major recombinant form (major RF) and the minor RF, respectively. The prevalence of the major RF was highest among IDUs (42.0%, 21/50), but it was also found among other risk populations: male homo-/bisexuals (18.8%, 3/16) and heterosexuals (9.8%, 9/92; Fig. 2).

Near-Full-Length Nucleotide Sequence Analysis Detects New CRF in Malaysia

To characterize the detailed subtype structure of new HIV-1 recombinant strains in Malaysia, we determined near-full-length nucleotide sequences of HIV-1 strains from eight epidemiologically unlinked individuals who are infected with

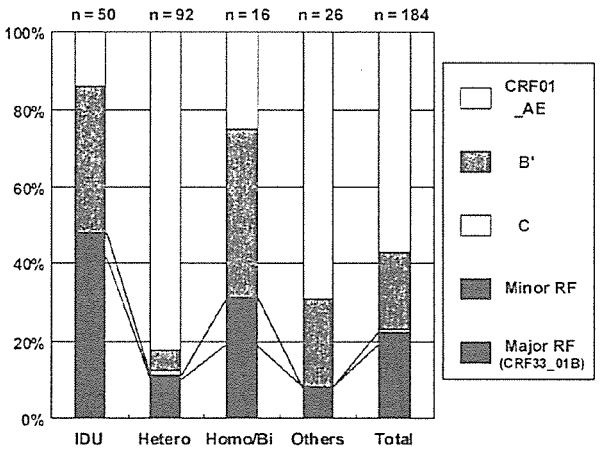

FIGURE 2. Distribution of HIV-1 genotypes in different risk groups in Malaysia. Genotypes were determined by phylogenetic and recombination breakpoint analyses of the nucleotide sequences in protease and reverse transcriptase regions (see text). Bars indicate the relative frequency (%) of HIV-1 genotypes in each risk group shown at the bottom. "n" indicates the number of specimens analyzed in each risk group. "Others" include 22 cases with unknown risk factors and 4 cases of MCT. Abbreviations: B′, HIV-1 subtype B′ (Thai variant of subtype B); C, subtype C; major RF, major recombinant forms (CRF33_01B); minor RF, minor recombinant forms (unique recombinant forms composed of CRF01_AE and subtype B).

either major RFs (n = 4) (05MYKL007.1, 05MYKL015.2, 05MYKL031.1, and 05MYKL045.1) or minor RFs (n = 4) (04MYKL016.1, 03MYKL018.1, 04MYKL019.1, and 05MYKL043.1) (Table 1). The bootscanning plot of near-full-length nucleotide sequences revealed that four major RFs indeed displayed the identical recombinant structure composed of CRF01_AE and subtype B (Fig. 3A). In this chimera, two short subtype B segments are inserted into a backbone of CRF01_AE. The profile of the subtype structure of Malaysian major RFs is distinct from that of CRF15_01B previously reported in Thailand,[19] where most of the *env* region belongs to subtype B in a backbone of CRF01_AE (Figs. 3B, C). In major RFs, a total of four recombination breakpoints (sites I–IV in Fig. 3C) are found to be clustered in a 0.9-kb region in *gag*-RT gene. The subregion tree analysis demonstrated that CRF01_AE segments (regions 1 and 5 in Fig. 3B) in major RFs belonged to the cluster of Thailand CRF01_AE, indicating that these segments indeed originated from CRF01_AE of Thai origin. However, region 3 (121 bp) was too short to discern its origin from either African or Thailand CRF01_AEs. Similarly, we were not able to discern whether subtype B segments (regions 2 and 4) originated from subtype B′ because subtypes B and B′ are phylogenetically indistinguishable in the Pro-RT region (Fig. 3B).

On the other hand, minor RFs (n = 4) showed a different degree of structural relatedness with major RFs. In particular, site II in Pro region is shared between major RFs and three minor RFs (1.6%, 3/184) (04MYKL016.1, 04MYKL019.1, and 05MYKL043.1; Fig. 3C). Similarly, the recombination breakpoints in *gag* p7 (site I) and RT (site IV) regions in major RF are shared with 04MYKL016.1 and 05MYKL043.1, respectively. 03MYKL018.1 is the most distantly related recombinant that shows no obvious structural relatedness with other RFs, while 05MYKL043.1 is most closely related with major RFs, sharing two recombination breakpoints (sites II and IV) but harboring slightly longer subtype B segments in *gag*-Pro and Pro-RT regions than major RFs (Fig. 3C).

Consistent with the results of recombination breakpoint analyses, the phylogenetic tree analysis of near-full-length nucleotide sequences showed that four major RFs (05MYKL007.1, 05MYKL015.2, 05MYKL031.1, and 05MYKL045.1) and the most closely related minor RF (05MYKL043.1) formed a monophyletic cluster with high bootstrap values (93%) (Fig. 3D). In contrast, three other minor RFs (04MYKL016.1, 03MYKL018.1, and 04MYKL019.1) are located outside any of the CRF01_AE-related sequences, including CRF15_01B (Fig. 3D).

Characterization of the Recombination Breakpoints in Malaysian RFs

To search for the possible sequence signatures near the subtype boundaries in Malaysian RFs, we mapped out all four recombination breakpoints (sites I–IV in Fig. 3C) estimated by informative site analysis (Fig. 4). It is noted that the homopolymeric nucleotide tracts (polyadenine ($A_{\geq 4}$), polyguanine ($G_{\geq 3}$), polycytosine ($C_{\geq 4}$), or polythymine ($T_{\geq 4}$)), known to pause in vitro reverse transcription and promote template switch,[32,33] are observed within or adjacent to all four recombination breakpoints. In particular, a characteristic

TABLE 1. Epidemiological and Clinical Background of Study Subjects Harboring Novel CRF01_AE/B Recombinants Emerging in Malaysia

Sample ID	Sex	Age (y)	Ethnicity	Risk Factor*	Date of First Positive HIV-1 Test (mo/yr)	Date of Collection (mo/yr)	Treatment Status†
05MYKL007.1	Male	32	Chinese	Homo	11/04	6/05	ARV-naive
05MYKL015.2	Male	33	Malay	Homo	4/04	8/05	ARV-naive
05MYKL031.1	Male	44	Indian	Hetero	1/05	8/05	3TC d4T NVP
05MYKL045.1	Male	43	Malay	Hetero	7/05	8/05	AZT 3TC EFV
04MYKL016.1	Female	30	Malay	Hetero	9/02	4/04	ARV-naive
03MYKL018.1	Male	32	Chinese	Bisexual	1/97	12/03	ARV-naive
04MYKL019.1	Male	49	Chinese	Hetero + IDU	2/04	6/04	ARV-naive
05MYKL043.1	Male	32	Malay	Hetero + IDU	8/03	8/05	AZT 3TC EFV

Sample ID	CD4 Count (cells/mm^3)	Viral Load (copies/mL)	Specimen Type	HIV-1 Genotype‡	Accession No.
05MYKL007.1	389	14,300	Isolate	CRF33_01B	DQ366659
05MYKL015.2	310	9,900,000	Isolate	CRF33_01B	DQ366660
05MYKL031.1	3	188,000	Isolate	CRF33_01B	DQ366661
05MYKL045.1	121	4300	Isolate	CRF33_01B	DQ366662
04MYKL016.1	401	3290	Plasma	Minor RF	DQ366663
03MYKL018.1	74	328,000	Plasma	Minor RF	DQ366664
04MYKL019.1	78	67,100	Plasma	Minor RF	DQ366665
05MYKL043.1	375	<50	Isolate	Minor RF	DQ366666

*Homo indicates male homosexual; Hetero, heterosexual; IDU, injecting drug user.
†ARV indicates antiretroviral; 3TC, lamivudine; d4T, didanosine; NVP, nevirapine; AZT, zidovudine; EFV, efavirenz.
‡Minor RF refers to the unique recombinant form composed of CRF01_AE and subtype B that showed a different degree of structural relatedness with CRF33_01B (Fig. 3).

homopolymeric nucleotide tract, the A_6GA_6 sequence, was found within site IV (in the RT gene). Similarly, noncontiguous sets of homopolymeric tracts, $A_5/G_5/T_4$ sequence, were observed within site II (in Pro gene). It is also noted that sites I and III are located adjacent to the boundaries between *gag* p7 and p1 and between Pro and RT, respectively (Fig. 4).

DISCUSSIONS

In the present study, we identified a novel CRF (CRF33_01B) composed of CRF01_AE and subtype B in Malaysia. Four sets of near-full-length nucleotide sequences with identical recombinant structure were determined from epidemiologically unlinked individuals with different risk factors and ethnicities in Malaysia (Table 1). This new chimera displays the recombinant structure distinct from any other CRFs reported to date, including CRF15_01B identified in Thailand.[19] The prevalence of CRF33_01B was particularly high among IDUs (42.0%, 21/50), compared with other risk populations (18.8% (3/16) in male homo-/bisexuals; 9.8% (9/92) in heterosexuals; Fig. 2). This may suggest that CRF33_01B first emerged among IDUs and subsequently spread to other risk populations.[23,24]

In addition to this new CRF33_01B, we detected six minor RFs consisting of CRF01_AE and subtype B. Among them, we determined the near-full-length nucleotide sequences of four minor RFs (Table 1). Most of them (except 03MYKL018.1) seem to share the recombination breakpoints with those of CRF33_01B (Fig. 3C). This suggests that they are closely related relatives and are most likely to be secondary recombinants derived from CRF33_01B that were generated by subsequent crossover(s) with either CRF01_AE or subtype B strain circulating in Malaysia.

The sequence characteristics found in or near the recombination breakpoints in CRF33_01B are suggestive of the following traits: (i) The homopolymeric tracts, known to pause in vitro reverse transcription and promote template switch (retroviral recombination),[32,33] are detected within or adjacent to all four recombination breakpoints in CRF33_01B (sites I–IV in Fig. 3C). Particularly, site II recombination breakpoint in Pro gene that contains a characteristic set of noncontiguous homopolymeric tracts, $A_5/G_5/T_4$ sequence (Fig. 4), was shared with three out of four minor RFs (05MYKL043.1, 04MYKL016.1, and 04MYKL019.1; Fig. 3C). Interestingly, this recombination breakpoint (site II) is shared with some of CRF01_AE/B recombinants reported in Thailand (unpublished data). It is thus tempting to speculate that site II may be one of the preferred sites for in vivo recombination. (ii) Some recombination breakpoints are found near the boundaries of functional domains of HIV-1 proteins: sites I and III are located adjacent to the boundaries between *gag* p7 and p1 and between Pro and RT, respectively (Fig. 4). This may reflect the selection pressure to maintain the functional integrity of HIV-1 proteins in recombination processes. Extensive dissemination of this new CRF over CRF01_AE and subtype B particularly among IDUs in Malaysia may suggest the presence of yet undefined selective advantages over parental strains.

Finally, we observed a unique parallelism of the transition in molecular epidemiological features of HIV-1

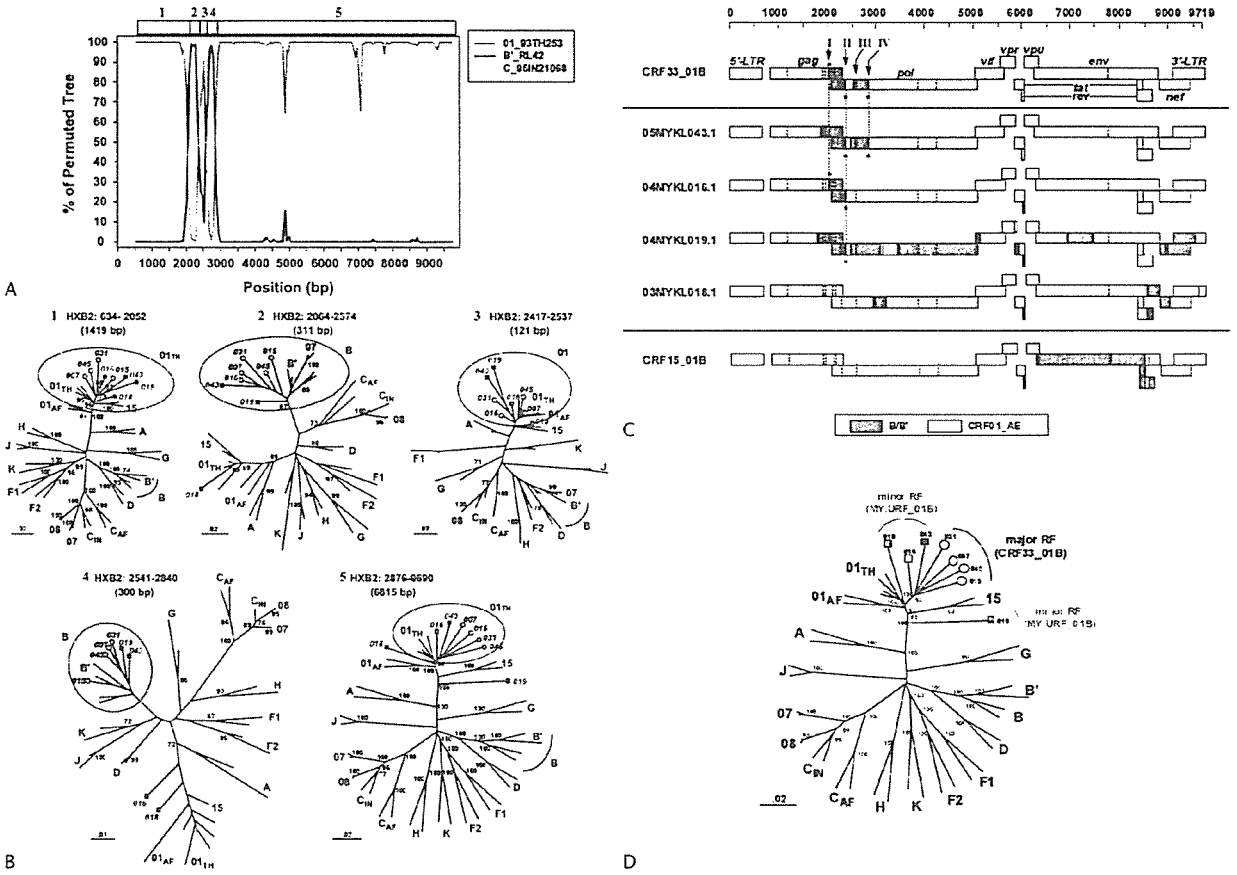

FIGURE 3. Characterization of newly emerging HIV-1 recombinant forms (RFs) in Malaysia. A, Bootscanning analysis of major RF (CRF33_01B). CRF01.93TH253 and B′. RL42 are used as the putative parental subtypes. Bootstrap values of 100 replicates were plotted for a window of 250 bp moving in increments of 50 bp along near-full-length sequences. B, Subregion trees in the segments (1–5) defined by informative site analysis of CRF33_01B are illustrated. Nucleotide positions of each segment relative to HXB2 and the size of segments (in parentheses) are indicated. Bootstrap values (≥70%) are shown. Major RFs (CRF33_01B) (n = 4: 05MYKL007.1, *007*; 05MYKL015.2, *015*; 05MYKL031.1, *031*; 05MYKL045.1, *045*). Minor RFs (n = 4: 03MYKL018.1 *018*; 04MYKL016.1 *016*; 04MYKL019.1 *019*; 05MYKL043.1, *043*). Abbreviations of strain codes for minor RFs are shown in gray. C, Schematic representation of subtype structures of novel HIV-1 recombinants. Recombination breakpoints (sites I–IV) are indicated. Dotted vertical lines indicate the breakpoints (marked with asterisks) shared among Malaysian RFs. D, Neighbor-joining tree of eight near-full-length nucleotide sequences of novel Malaysian recombinants with reference sequences of HIV-1 group M subtypes and sub-subtypes (http://hiv-web.lanl.gov/). Selected HIV-1 genotypes of geographical importance are included: B′, Thai variant of subtype B; C_{IN}, Indian subtype C; C_{AF}, African subtype C; 01_{TH}, Thailand CRF01_AE; 01_{AF}, African CRF01_AE; 07, CRF07_BC; 08, CRF08_BC; 15, CRF15_01B, minor RF, minor recombinant form composed of CRF01_AE and subtype B. Branching orders were assessed by 100 bootstrap analyses, and values ≥70% are shown at the corresponding nodes of the tree. Scale bar represents 2% genetic distance (0.02 substitution/site).

epidemics between Thailand and Malaysia. In the early phase of the Malaysian epidemic, CRF01_AE and subtype B′ were circulating relatively independently among heterosexuals and IDUs, respectively, similar to the early stage of the Thai epidemic.[8,9] However, Tovanabutra and others recently reported that in a high-risk cohort in northern Thailand, CRF01_AE/B recombinants began to be identified in 2001 at 8.3% of incident cases and increased to 57.1% in 2002.[18] Due to the lack of availability of archival HIV specimens in the current study, we were not able to pinpoint the timing of emergence of CRF33_01B. However, the relatively long branch length of the CRF33_01B cluster (interstrain genetic diversity of 6.06 ± 0.50% (n = 4, isolates collected in 2005) compared with the interstrain genetic diversity of 3.10 ± 0.55% for Thailand CRF01_AE cluster (n = 4, isolates collected in 1990–1997) in near-full-length genomes) (Fig. 3D) suggests that the timing of the origin of CRF33_01B could be dated back to mid-1990s (unpublished data).

In summary, we report herein on a novel CRF that is circulating widely among various risk groups in Malaysia.

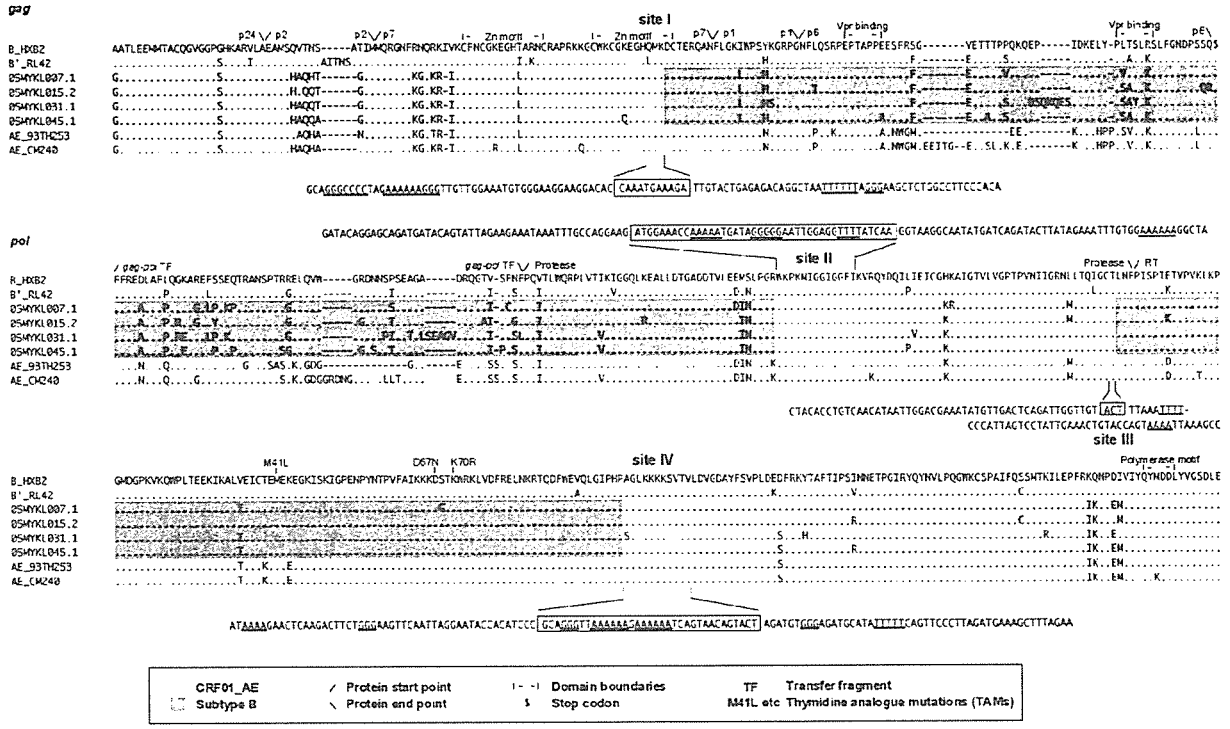

FIGURE 4. Sequence signatures in the vicinity of the recombination breakpoints in CRF33_01B. Amino acid sequence alignment of major recombinant form (RF) (CRF33_01B, n = 4) with subtype B (HXB2), B′ (RL42), and CRF01_AE (93TH253 and CM240) reference strains in *gag-pol* regions is shown. Four recombination breakpoints are clustered in this region. Periods (.) indicate the sequence identity with HXB2, and dashes (–) denote the gaps in the alignment. The CRF01_AE and subtype B regions are marked with different shadows. Unshaded areas are the presumed recombination breakpoints estimated by informative site analysis. Nucleotide sequences (consensus of four CRF33_01B isolates) within (boxed) and adjacent to (50-nucleotide region) the estimated recombination breakpoints (sites I–IV) are shown. Sequence landmarks in the Pro-RT region, including the boundaries of the domains of HIV-1 proteins, the conserved motifs, and the positions of major drug resistance mutations, are indicated at the top of the alignment.

Wide distribution of CRF33_01B involving all major ethnic and risk groups provides evidence for extensive bridging of HIV-1 transmission between different risk groups in Malaysia.

ACKNOWLEDGMENTS

We thank Kouichi Watanabe, Toshinari Onogi, Clarence Sim, and Margaret Lau for assistance; Brian Foley and Thomas Leitner for valuable advice on assignment of the new CRF; and Tim Mastro and Sai Kit Lam for critical reading of the manuscript. We also thank the patients for their participation in this study.

REFERENCES

1. Goh KL, Chua CT, Chiew IS, et al. The acquired immune deficiency syndrome: a report of the first case in Malaysia. *Med J Malaysia.* 1987;42: 58–60.
2. Joint United Nations Program on HIV/AIDS (UNAIDS). Available at: http://www.unaids.org. Accessed September 17, 2006.
3. Ministry of Health. *HIV/AIDS Report*, 2004.
4. UNAIDS. *AIDS Epidemic Update: December 2005*, 2005.
5. Malim MH, Emerman M. HIV-1 sequence variation: drift, shift, and attenuation. *Cell.* 2001;104:469–472.
6. Leitner T, Korber B, Daniels M, et al. HIV-1 subtype and circulating recombinant form (CRF) references sequences. In: Leitner T, Foley B, Hahn B, et al, eds. *HIV Sequence Compendium 2005.* Report no. LA-UR 06-0680. Los Alamos, NM: Theoretical Biology and Biophysics Group, Los Alamos National Library; 2005:41–48.
7. Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. *Science.* 2000;288:55–57.
8. Ou CY, Takebe Y, Weniger BG, et al. Independent introduction of two major HIV-1 genotypes into distinct high-risk populations in Thailand. *Lancet.* 1993;341:1171–1174.
9. Weniger BG, Takebe Y, Ou CY, et al. The molecular epidemiology of HIV in Asia. *AIDS.* 1994;8(Suppl 2):S13–S28.
10. Wasi C, Herring B, Raktham S, et al. Determination of HIV-1 subtypes in injecting drug users in Bangkok, Thailand, using peptide-binding enzyme immunoassay and heteroduplex mobility assay: evidence of increasing infection with HIV-1 subtype E. *AIDS.* 1995;9:843–849.
11. Vanichseni S, Kitayaporn D, Mastro TD, et al. Continued high HIV-1 incidence in a vaccine trial preparatory cohort of injection drug users in Bangkok, Thailand. *AIDS.* 2001;15:397–405.
12. Hudgens MG, Longini IM Jr, Vanichseni S, et al. Subtype-specific transmission probabilities for human immunodeficiency virus type 1 among injecting drug users in Bangkok, Thailand. *Am J Epidemiol.* 2002; 155:159–168.

13. Tovanabutra S, Polonis V, De Souza M, et al. First CRF01_AE/B recombinant of HIV-1 is found in Thailand. *AIDS*. 2001;15:1063–1065.
14. Viputtijul K, de Souza M, Trichavaroj R, et al. Heterosexually acquired CRF01_AE/B recombinant HIV type 1 found in Thailand. *AIDS Res Hum Retroviruses*. 2002;18:1235–1237.
15. Nguyen L, Hu DJ, Choopanya K, et al. Genetic analysis of incident HIV-1 strains among injection drug users in Bangkok: evidence for multiple transmission clusters during a period of high incidence. *J Acquir Immune Defic Syndr*. 2002;30:248–256.
16. Ramos A, Nguyen L, Hu DJ, et al. New HIV type 1 CRF01_AE/B recombinants displaying unique distribution of breakpoints from incident infections among injecting drug users in Thailand. *AIDS Res Hum Retroviruses*. 2003;19:667–674.
17. Swanson P, Devare SG, Hackett J Jr. Full-length sequence analysis of HIV-1 isolate CM237: a CRF01_AE/B intersubtype recombinant from Thailand. *AIDS Res Hum Retroviruses*. 2003;19:707–712.
18. Tovanabutra S, Beyrer C, Sakkhachornphop S, et al. The changing molecular epidemiology of HIV type 1 among northern Thai drug users, 1999 to 2002. *AIDS Res Hum Retroviruses*. 2004;20:465–475.
19. Tovanabutra S, Watanaveeradej V, Viputtikul K, et al. A new circulating recombinant form, CRF15_01B, reinforces the linkage between IDU and heterosexual epidemics in Thailand. *AIDS Res Hum Retroviruses*. 2003;19:561–567.
20. Brown TM, Robbins KE, Sinniah M, et al. HIV type 1 subtypes in Malaysia include B, C, and E. *AIDS Res Hum Retroviruses*. 1996;12:1655–1657.
21. Beyrer C, Vancott TC, Peng NK, et al. HIV type 1 subtypes in Malaysia, determined with serologic assays: 1992–1996. *AIDS Res Hum Retroviruses*. 1998;14:1687–1691.
22. Saraswathy TS, Ng KP, Sinniah M. Human immunodeficiency virus type 1 subtypes among Malaysian intravenous drug users. *Southeast Asian J Trop Med Public Health*. 2000;31:283–286.
23. Tee KK, Pon CK, Kamarulzaman A, et al. Emergence of HIV-1 CRF01_AE/B unique recombinant forms in Kuala Lumpur, Malaysia. *AIDS*. 2005;19:119–126.
24. Tee KK, Saw TL, Pon CK, et al. The evolving molecular epidemiology of HIV type 1 among injecting drug users (IDUs) in Malaysia. *AIDS Res Hum Retroviruses*. 2005;21:1046–1050.
25. Kato K, Sato H, Takebe Y. Role of naturally occurring basic amino acid substitutions in the human immunodeficiency virus type 1 subtype E envelope V3 loop on viral coreceptor usage and cell tropism. *J Virol*. 1999;73:5520–5526.
26. Salminen MO, Koch C, Sanders-Buell E, et al. Recovery of virtually full-length HIV-1 provirus of diverse subtypes from primary virus cultures using the polymerase chain reaction. *Virology*. 1995;213:80–86.
27. Yang R, Kusagawa S, Zhang C, et al. Identification and characterization of a new class of human immunodeficiency virus type 1 recombinants comprised of two circulating recombinant forms, CRF07_BC and CRF08_BC, in China. *J Virol*. 2003;77:685–695.
28. Rambaut A. *Se-Al (Sequence Alignment Editor)*, version 1.0, alpha 1. Oxford, England: University of Oxford, Department of Zoology; 1996.
29. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4:406–425.
30. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol*. 1980;16:111–120.
31. Lole KS, Bollinger RC, Paranjape RS, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol*. 1999;73:152–160.
32. Klarmann GJ, Schauber CA, Preston BD. Template-directed pausing of DNA synthesis by HIV-1 reverse transcriptase during polymerization of HIV-1 sequences in vitro. *J Biol Chem*. 1993;268:9793–9802.
33. Quinones-Mateu ME, Gao Y, Ball SC, et al. In vitro intersubtype recombinants of human immunodeficiency virus type 1: comparison to recent and circulating in vivo recombinant forms. *J Virol*. 2002;76:9600–9613.

Separate elements are required for ligand-dependent and -independent internalization of metastatic potentiator CXCR4

Yuko Futahashi,[1] Jun Komano,[1,4] Emiko Urano,[1] Toru Aoki,[1,2] Makiko Hamatake,[1] Kosuke Miyauchi,[1] Takeshi Yoshida,[3] Yoshio Koyanagi,[3] Zene Matsuda[1] and Naoki Yamamoto[1,2]

[1]AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku, Tokyo 162-8640; [2]Department of Molecular Virology, Tokyo Medical and Dental University, 1-5-45, Yushima, Bunkyo-ku, Tokyo 113-8519; [3]Laboratory of Viral Pathogenesis, Institute for Virus Research, Kyoto University, 53 Shougoin-kawahara machi, Sakyou-ku, Kyoto 606-8507, Japan

(Received September 17, 2006/Revised November 3, 2006/Accepted November 11, 2006/Online publication January 19, 2007)

The C-terminal cytoplasmic domain of the metastatic potentiator CXCR4 regulates its function and spatiotemporal expression. However, little is known about the mechanism underlying constitutive internalization of CXCR4 compared to internalization mediated by its ligand, stromal cell-derived factor-1 alpha (SDF-1α)/CXCL12. We established a system to analyze the role of the CXCR4 cytoplasmic tail in steady-state internalization using the NP2 cell line, which lacks endogenous CXCR4 and SDF-1α. Deleting more than six amino acids from the C-terminus dramatically reduced constitutive internalization of CXCR4. Alanine substitution mutations revealed that three of those amino acids Ser^{344} Glu^{345} Ser^{346} are essential for efficient steady-state internalization of CXCR4. Mutating Glu^{345} to Asp did not disrupt internalization, suggesting that the steady-state internalization motif is S(E/D)S. When responses to SDF-1α were tested, cells expressing CXCR4 mutants lacking the C-terminal 10, 14, 22, 31 or 44 amino acids did not show downregulation of cell surface CXCR4 or the cell migration induced by SDF-1α. Interestingly, however, we identified two mutants, one with E344A mutation and the other lacking the C-terminal 17 amino acids, that were defective in constitutive internalization but competent in ligand-promoted internalization and cell migration. These data demonstrate that ligand-dependent and -independent internalization is genetically separable and that, between amino acids 336 and 342, there is a negative regulatory element for ligand-promoted internalization. Potential involvement of this novel motif in cancer metastasis and other CXCR4-associated disorders such as warts, hypogammaglobulinemia, infections and myelokathexis (WHIM) syndrome is discussed. (*Cancer Sci* 2007; 98: 373–379)

The chemokine receptor CXCR4 is a class-A G protein-coupled receptor (GPCR; reviewed in [1,2]) and its natural ligand is stromal cell-derived factor-1 alpha (SDF-1α)/CXCL12. CXCR4 also serves as the receptor for HIV type 1 (HIV-1). Many cell types express CXCR4, including peripheral blood lymphocytes, monocytes-macrophages, thymocytes, dendritic cells, endothelial cells, epithelium-derived tumor cells, microglial cells, neurons and hematopoietic stem cells. CXCR4 plays multiple biological roles from promoting development of neuronal networks to regulating migration of leukocytes, cerebellar granule cells and hematopoietic stem cells.[3–8] Analysis of knockout mice indicates that the CXCR4/SDF-1α system is essential for maintenance of hematopoiesis and intestinal vascularization.[9,10]

The CXCR4/SDF-1α system also functions in pathological processes, including autoimmune diseases, cancer progression and metastasis, and AIDS caused by HIV-1. Recently, metastasis of breast cancer cells was found to be regulated by the CXCR4/SDF-1α axis.[5] Similarly, other studies have found that metastasis of other malignancies was controlled by the CXCR4/SDF-1α system, including colon carcinoma[11] non-small cell lung cancer[12] and prostate cancer.[13] These observations suggest that the CXCR4/SDF-1α axis is a potential target for metastatic cancer therapy.

Warts, hypogammaglobulinemia, infections and myelokathexis (WHIM) syndrome is a rare combined immunodeficiency characterized by an unusual form of neutropenia. It is reported that the CXCR4 cytoplasmic tail is mutated and often truncated in WHIM syndrome.[14] Thus, determining the biochemical activity of the CXCR4 cytoplasmic tail should facilitate understanding of the pathogenesis of WHIM syndrome as well as suggest ways to control cancer metastasis.

Following SDF-1α binding, CXCR4 is activated, triggering multiple signaling cascades via Gα or β-arrestin 2 (reviewed in[15]). To desensitize activated CXCR4, the G protein-coupled receptor kinase (GRK) is recruited and phosphorylates serine residues on the CXCR4 cytoplasmic tail, thereby inactivating Gα-mediated signal. Simultaneously, CXCR4 is internalized in a clathrin-dependent manner. β-arrestin 2 competes with Gα for CXCR4 binding and can initiate signal transduction independent from Gα. β-arrestin 2 can also induce clathrin-dependent CXCR4 endocytosis. Thus, cell surface levels of CXCR4 transiently decrease after agonist binding but, several hours later, surface levels of CXCR4 return to normal. Most internalized CXCR4 is transported to lysosomes and degraded, but some internalized CXCR4 is recycled. It is reported that amino acids within the cytoplasmic tail are required for agonist-dependent endocytosis of CXCR4.[16–18]

By contrast, it is unclear how steady-state cell surface levels of CXCR4 are maintained in the absence of SDF-1α. Although cell surface levels of CXCR4 could be regulated at the transcriptional level, it is likely that primary regulation occurs post-translationally. Given that the cell surface levels of CXCR4 are positively correlated with cancer cells' ability to metastasize,[5,19] understanding the post-translational behavior of CXCR4 is likely to shed light on metastatic processes. Historically, cells expressing endogenous CXCR4 have been used for analysis of CXCR4 trafficking. However, as is the case with many G protein-coupled receptors (GPCR), CXCR4 trafficking is influenced by spontaneous oligomerization in the absence of ligand.[20–22] Thus, previous observations might not correctly model phenotypes seen in CXCR4 mutants.

In the present study, we analyzed the contribution of the cytoplasmic tail to the post-translational trafficking of CXCR4 in a cell line lacking both endogenous CXCR4 and SDF-1α. Using genetic approaches, we identified two amino acid motifs within the CXCR4 cytoplasmic tail; one that positively regulates

[4]To whom correspondence should be addressed. E-mail:ajkomano@nih.go.jp

doi: 10.1111/j.1349-7006.2007.00388.x