*Review Article*

# γ-Herpesviruses and cellular signaling in AIDS-associated malignancies

**Kohji Noguchi,[2] Hidesuke Fukazawa, Yuko Murakami, Naoko Takahashi,[3] Satoshi Yamagoe and Yoshimasa Uehara[1,4]**

Department of Bioactive Molecules, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

(Received March 17, 2007/Revised April 24, 2007/Accepted April 30, 2007/Online publication July 12, 2007)

**γ-Herpesviruses, Epstein–Barr virus (EBV/HHV-4) and Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8), are involved in human carcinogenesis, particularly in immunocompromised patients. Virus-associated malignancies are becoming of significant concern for the mortality of long-lived immunocompromised patients, and therefore, research of advanced strategies for AIDS-related malignancies is an important field in cancer chemotherapy. Detailed understanding of the EBV and KSHV lifecycle and related cancers at the molecular level is required for novel strategies of molecular-targeted cancer chemotherapy. The present review gives a simple outline of the functional interactions between KSHV- and EBV-viral gene products and host cell deregulated signaling pathways as possible targets of chemotherapy against AIDS-related malignancies. (*Cancer Sci* 2007; 98: 1288–1296)**

## Herpesviruses and AIDS-related malignancy

Herpesviruses are double-strand DNA viruses and at least eight of them are identified as human pathogens: human herpes simplex virus 1 and 2 (HSV/HHV-1 and -2), varicella-zoster virus (VZV/HHV-3), Epstein–Barr virus (EBV/HHV-4), human cytomegalovirus (HCMV/HHV-5), human herpesviruses 6 and 7 (HHV-6 and -7), and Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8).[1,2] The herpesviruses are divided into three subgroups, α- β- and γ-herpesvirus. EBV and KSHV are similar γ-herpesviruses and associated AIDS-related malignancies and their genomic DNA encodes various homologous gene products with oncogenic potential.[1,2] γ-Herpesviruses are primarily lymphotropic but some show a lytic replication cycle in epithelial and fibroblast cells. EBV and KSHV display two similar alternative phases of infection, latent and lytic. Upon induction of the lytic cycle, lytic viral gene expression initiates lytic viral DNA replication and virion production. In contrast, a restricted set of viral genes is expressed and viral genomes are maintained as circular double-stranded DNA in latently infected malignant cells (Table 1).[2]

Kaposi's sarcoma and AIDS-related lymphoma are the most frequent AIDS-defining malignancies. KSHV was discovered from the Kaposi's sarcoma lesion of an AIDS patient in 1994,[3] and has been linked to this sarcoma and certain lymphoproliferative disorders. Histologically, Kaposi's sarcoma lesions are characterized by neoangiogenesis and proliferative spindle-shaped cells admixed with aberrant endothelial and inflammatory cells.[4] KSHV is detected in spindle-shaped cells that are likely the malignant cells in such lesion.[5] In advanced Kaposi's sarcoma lesions, latently expressed KSHV gene products are believed to promote cell growth, to block apoptosis and host immune response, and to induce neoangiogenesis.[2] KSHV-infection is the only risk factor that is essential for Kaposi's sarcoma development, and the paracrine model is hypothesized for Kaposi's lesion in which KSHV induces a growth-factor rich microenvironment that supports the proliferation of both neighboring and KSHV-infected cells.[6] Additionally, stimulated by this virus-induced microenvironment, the inflammatory cells and cytokines including interleukin (IL)-6, basic fibroblast growth factor (bFGF), vascular endothelial growth factor (VEGF) and the up-regulated matrix metalloproteinases (MMP), activity of endothelial cells probably contribute to profound neoangiogenesis. In addition, HIV may itself contribute to Kaposi's sarcoma tumorigenesis.[7] The Tat protein of HIV appears to stimulate endothelial cell migration, protect KSHV-infected cells from apoptosis, promote spindle-shaped cell growth, and enhance KSHV transmission.[8,9]

More than 50% of AIDS-related lymphoma cases are also associated with γ-herpesviruses, EBV and KSHV. EBV infects B cells as a persistent infection. EBV latent genes activate resting B cells and induce complete transformation, but the host immune system suppresses EBV-induced malignancies.[10] In healthy individuals, immunosurveillance systems including cytotoxic T-cell responses against EBV latent antigen suppress the expansion of these lymphoblasts.[11] However, EBV has strategies to escape from an immunosurveillance system in a host.[12,13] Moreover, HIV-infection increases the incidence of lymphoma, because such virus-induced lymphoblasts have a good opportunity to expand under an immunocompromised condition. KSHV is also associated with another B-cell lymphoproliferative disorder called multicentric Castleman's disease, in which lymphadenopathy with polyclonal hyperimmunoglobulinemia, and high levels of IL-6 in serum are characteristics and there is often HIV co-infection.[14]

Many cellular oncogenic and anti-oncogenic molecules are involved in the mechanisms for malignant cell proliferation and cell survival.[15] Deregulated cell growth signaling, deregulated cell cycle control, dysfunction of tumor suppressor genes, and resistance to apoptotic cell death contribute to malignant phenotypes.[16] Intriguingly, infection of KSHV and EBV has been found to manipulate host cell mechanism, and pathogenic viral gene products also deregulate cellular signaling to promote cell growth and survival. The molecular and epidemiological linkage between AIDS-related malignancies and KSHV- and EBV-infection suggests that these viral gene products are potential targets for molecular-targeted chemotherapy for AIDS-related malignancies. This review describes an outline of the molecular aspects for KSHV- and EBV-associated cell signaling, latency, and malignancy as targets for molecular-targeted chemotherapy of AIDS-associated malignancies.

[1]To whom correspondence should be addressed. E-mail: yuehara@iwate-med.ac.jp
[2]Current address: Department of Chemotherapy, Kyoritsu University of Pharmacy, Tokyo, Japan.
[3]Current address: Division of Hematopoiesis, Center for AIDS Research, Kumamoto University, Kumamoto, Japan.
[4]Current address: Iwate Medical University School of Pharmacy, Iwate, Japan.



doi: 10.1111/j.1349-7006.2007.00555.x

**Table 1. Viral genes expressed in latently infected cells**

| | Malignancies | Viral products |
|---|---|---|
| Epstein–Barr virus | | |
| Type I latency | Burkitt's lymphoma | EBNA1, EBER, BART |
| Type II latency | Hodgkin's lymphoma | EBNA1, LMP-1,-2, |
| | Nasopharyngeal carcinoma | EBER, BART |
| | Natural killer/T-cell lymphoma and AIDS-related non-Hodgkin's lymphoma | |
| Type III latency | AIDS-related proliferative disorders, post-transplant lymphoproliferative disorders, and lymphoblastoid cell lines | EBNA1,-2,-3 A, -3B,-3C,-LP, LMP-1,-2, EBER, BART |
| Kaposi's sarcoma-associated herpesvirus | | |
| | Kaposi's sarcoma | LANA-1, k-Cyclin, vFLIP, vIRF1, Kaposin |
| | Primary effusion lymphomas | LANA-1, k-Cyclin, vFLIP, vIRF3, Kaposin, vIL-6 (~40%) |
| | Multicentric Castleman's disease | LANA-1, k-Cyclin, vFLIP, vIL-6 (~40%) |

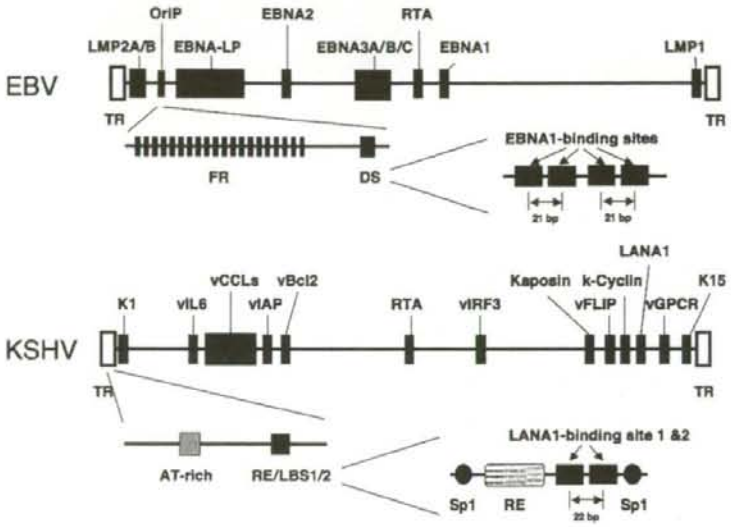


**Fig. 1.** Schematic drawing of the organization of Epstein–Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV) linear genomes. Each viral gene-coding region is flanked by multiple terminal repeat (TR) units. Different viral transcription patterns are observed in each latency types I, II, and III, and different EBNA are encoded by individual mRNA generated by different splicing of the long transcript. The open reading frames and latent replication origins discussed here are indicated., LMP, latent membrane protein.

## Latent replication and maintenance of viral genomes

EBV- and KSHV-associated malignancies involve the viruses' latent cycle. During latency, the viral genome DNA is maintained as an episome, and behaves like a host cellular chromosome. The episomal viral genomes of EBV (~165 kbp) and KSHV (~165 kbp) (Fig. 1) are replicated semiconservatively once per cell cycle and partitioned faithfully to daughter cells during mitosis, and viral factors (*trans*- and *cis*-) and host cell mechanisms are employed to keep these viral latencies in infected cells.[17,18] All EBV- and KSHV-associated malignancies express EBNA1 and LANA-1 proteins, respectively, and these two viral antigens are required for latent replication and maintenance of viral genome.

Latent replication of the EBV genome requires only two viral factors, EBNA1 antigen acting in *trans*, and the latent DNA replication origin, the OriP region in *cis*.[19] OriP consists of two major *cis* elements, namely the family of repeats (FR) and the dyad symmetry (DS) region (Fig. 1). The EBNA1 protein (641 amino acids) consists of three parts, the *N*-terminus (1–391), the joining region (392–458), and the *C*-terminus (459–641). The *N*-terminus contains two regions that can link two independent EBNA1-binding DNA elements, DS and FR, and can form a loop chromatin structure in a viral episome.[20] The two linking regions are separated by Gly-Ala rich repeats, and these repeats contribute to the escape from the host cell immune surveillance system by preventing EBNA1-peptide presentation on the major histocompatibility complex (MHC) class I complex.[12,13] The *C*-terminus of EBNA1 functions as a DNA-binding and EBNA1-dimerization domain, which is essential for OriP-dependent DNA replication.[21] EBNA1 does not have helicase or ATPase activity, but EBNA1-binding to the DS region appears to bend DNA at binding sites,[22] which may stimulate a locally unique chromatin structure and an assembly of components for DNA replication.

The DS region contains four EBNA1-binding sites consisting of two pairs of 21 bp-spacing, and EBNA1-binding on DS induces DNA replication licensed by host cell machineries. EBNA1-binding on DS appears to initiate an assembly of host cell pre-replication complexes, including the Origin recognition complex (ORC) and the Minichromosome maintenance (MCM) complex.[23] Functions of ORC1 and ORC2 are important in facilitating OriP-plasmid DNA replication,[24,25] and, conversely, the licensing factor Geminin negatively controls OriP-plasmid replication during cell cycle.[24] Interestingly, viral OriP-mediated

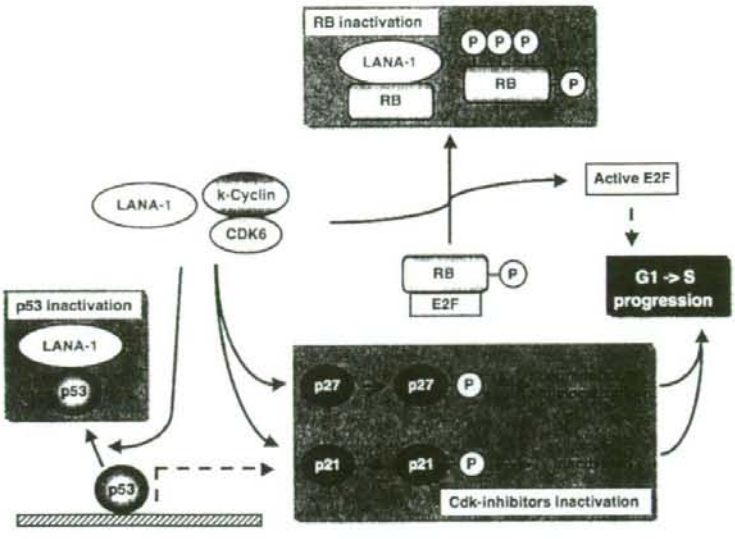


**Fig. 2.** Cell cycle regulatory proteins and viral gene products. Kaposi's sarcoma-associated herpesvirus (KSHV) latent antigen LANA-1 inactivates RB and p53, and KSHV k-cyclin/CDK complex inactivates Cdk-inhibitors and RB by phosphorylation (P). These viral product expressions promote G1-S progression of the host cell. Viral products are indicated by red font.

plasmid replication is more severely affected by ORC1/2 activity changes than host cell DNA replication and proliferation. ORC1 mutations or ORC2 reduction significantly impairs OriP-dependent plasmid replication and maintenance without inhibiting host cell proliferation. Within the DS region, other host cell factors also assemble to help OriP-plasmid replication and maintenance. The DS region contains three nonamer sites (TTAGGGTTA), which resemble telomeric repeats. TRF2 and TRF2-interacting tankyrase assemble on the DS region cooperatively with EBNA1, and these factors contribute to EBNA1-dependent OriP-plasmid replication/maintenance.[26]

The FR consists of a 20-member family of 30-bp repeats, and each member has one 18-bp element with high affinity to EBNA1. EBNA1-binding to FR has various effects. Multiple EBNA1-binding to FR is critical for OriP-mediated maintenance/partitioning of EBV episomal chromosome.[27] Moreover, the EBNA1-binding at FR enhances transcriptional activity of the latency promoters, called BamHIC and LMP1 promoter, and contributes to five latent gene expressions during latent proliferation of infected cells.[28]

Another γ-herpesvirus, KSHV, shows latent infection on some tumors, and most KSHV-infected cells in PEL and Kaposi's sarcoma lesions express latent genes for LANA-1, k-cyclin, vFLIP, and vIRF3.[2] The KSHV genome also has viral latent replication origin for persistent infection. The *trans*-acting factor is KSHV latency antigen LANA-1, and the *cis*-acting is the terminal repeat (TR) region consisting of two LANA-1-binding sites (LBS1/2) and a replication element (RE)[18,29–31] (Fig. 1). LANA-1 is a functional counterpart of EBNA1, and the minimum replication origin in the TR structurally related to EBV DS confers LANA-1-dependent replication origin activity on episomal plasmid DNA. LANA-1 protein (1162 amino acids) comprises three regions, and is required for stable KSHV episomal maintenance in infected cells. The *N*-terminal region of LANA-1 tethers LANA-1 to chromatin and mitotic chromosomes during cell cycle by binding with nucleosomal histone H2A-H2B.[32] Specifically, a peptide of LANA-1 residues at positions 5–13 binds to an acidic patch of nucleosomal H2A-H2B. The central region of LANA-1 contains three different acidic blocks, and the *C*-terminus (996–1139) encodes a DNA-binding domain that binds to an 18-bp imperfect palindrome in the LBS1/2.[18] LANA-1 binds cooperatively to LBS1 and LBS2 with high and low affinity, respectively,[30] and this structural organization is similar to half of the DS element in OriP. Functions of the LANA-1 *N*- and *C*-terminal regions are involved in LANA-1-mediated bridging between plasmid DNA and mitotic chromosome and are required for efficient TR-containing plasmid DNA replication and faithful segregation of episomes during cell division.[33]

As in the case of EBNA1-OriP, LANA-1-dependent KSHV genome replication depends on host cell machineries, and LANA-1-bound TR is associated with host cell DNA replication components. ORC and MCM complexes assemble at or near the LBS of the TR in the KSHV-latently infected cells.[29,34,35] ORC2 is able to interact with the LANA-1 *C*-terminus at least *in vitro*, and the LANA-1-DNA complex is likely to facilitate ORC recruitment at TR *in vivo*. Furthermore, various cellular proteins are assembled on the TR region in KSHV-infected cells.[36]

LANA-1 also contributes to persistent latency by repressing the lytic replication cycle. The immediately early gene product RTA, KSHV ORF50, plays a critical role in switching viral replication from the latent to the lytic phase by transactivating lytic viral replication gene expression through interaction with RBP-Jκ.[37] LANA-1 physically interacts with RTA and inhibits RTA's promoter activity. Such effects by LANA-1 decrease RTA transactivation activity and contribute to the maintenance of viral latency.

Overall, latently expressed major antigens EBNA1 and LANA-1 have pivotal roles in persistent infection and malignancy development caused by EBV and KSHV. Thus, functions of these antigens are attractive targets for molecular therapy/chemoprevention of latently infected cell malignancy.

## Cell cycle deregulation by viral gene products

Some KSHV encoding genes have the potential ability to modulate the cell cycle control mechanism (Fig. 2). k-Cyclin (ORF72) is a homolog of cyclin D2 (32% identity and 54% similarity).[38] The predominant partner for k-cyclin is CDK6, and this viral CDK complex phosphorylates RB protein like cellular CDK complexes.[38] Moreover, k-cyclin/CDK6 shows broader substrate specificity than cellular cyclinD/CDK6. k-Cyclin/CDK6 can phosphorylate CDK2 substrates, including ORC1,


doi: 10.1111/j.1349-7006.2007.00555.x

CDC6, p27Kip1, histone H1, and others.[39] In contrast to their cellular homolog, viral CDK complexes are constitutively active in KSHV-infected cells.[40] k-Cyclin lacking degron is more stable than other cellular cyclins, and viral CDK complexes do not require activation by the upstream kinase, CAK.[41] Moreover, viral CDK complexes antagonize the cellular CDK inhibitor, p27Kip1 and p21Cip1. Viral CDK complexes appear to inactivate p27Kip1 by phosphorylation-mediated degradation and mislocalization.[42–44] Viral CDK6 complex associates with and phosphorylates p21Cip1 on serine 130, and this phosphorylated p21Cip1 is not able to inhibit CDK2.[45] Thus, owing to the unusual dual substrate preference of both CDK2 and CDK6, and escaping from functional CDK-inhibitors,[46] viral CDK complexes promote S-phase entry efficiently. Moreover, LANA-1 also has roles in G1/S transition. LANA-1 expression is co-regulated with k-cyclin expression at latency locus by the major latency promoter.[47] LANA-1 is reported to inactivate RB and p53 like other oncogenic viral proteins,[48,49] and to enhance E2F activity for S-phase entry. These LANA-1 functions during latency modulate host cell growth machinery in multiple ways and contribute to B-cell hyperplasia.[50,51] Another viral protein, KSHV vIRF1, is able to interact with both p53 and ATM to inhibit p53 transcriptional activity.[52]

## Modulation of cellular signaling pathway (PI3K-Akt, NF-κB, MAPK, Wnt, Notch)

During EBV- and KSHV-associated malignancies, different types of latent gene expression have been described (Table 1). Some EBV and KSHV viral products have the ability to manipulate the host cellular signaling pathway during latent and lytic phases (Fig. 3). Such deregulated signaling by viral products contributes to malignant cell proliferation, survival and viral expansion. These manipulated cell signalings are as follows.

**NF-κB and MAPK.** KSHV vFLIP encoded by ORF71, related to cellular FLIP, was initially considered to be an apoptosis inhibitor like its cellular homolog, but its anti-apoptotic activity has recently been shown to be primarily associated with activation of the nuclear factor-κB (NF-κB) pathway (Fig. 3a). vFLIP is expressed in latently infected Kaposi's sarcoma spindle-shaped and PEL cells, and TRAF2-mediated association of vFLIP with IKK-γ causes constitutive activation of the NF-κB pathway.[53,54] EBV latent membrane protein 1 (LMP1) also activates the NF-κB pathway (Fig. 3a).[55] The LMP1 *C*-terminal cytoplasmic tail contains two distinct functional domains, *C*-terminal activation regions 1 and 2 (CTAR1 and 2). Each CTAR1-TRAF2 and CTAR2-TRADD complex links to the IKK–NIK–NF-κB pathway.[56,57] These unscheduled activations of the NF-κB pathway by viral products are required for virus-induced cellular transformation and cell survival. KSHV vFLIP-TRAF and EBV LMP1-TRAF complexes also activate another signaling pathway, the JNK-pathway that induces AP-1 activation to support cell survival.[58,59] In addition, LMP1-mediated activations of NF-κB and/or mitogen-activated protein kinase (MAPK) pathways contribute to VEGF production.[60] KSHV LANA-1 also has an ability to function as a co-activator of c-Jun to activate an AP-1 responsive element (Fig. 3a).[61]

EBV latent membrane protein 2 A (LMP2A) has an *N*-terminal cytoplasmic region containing eight tyrosine residues, two of which (Tyr-74 and -85) function as the immunoreceptor tyrosine-based activation domain (ITAM) motif (Fig. 3a). The phosphorylated ITAM region recruits Src-type tyrosine kinases, and this ITAM signalsome-mediated ERK-activation also stimulates the c-Jun/AP-1 pathway.[62] Analogously, KSHV K1 has a *C*-terminal cytoplasmic tail bearing an ITAM motif (Fig. 3b),[63] which is highly conserved among different K1 subtypes and is similar to that of EBV LMP2A. K1 is expressed in Kaposi's sarcoma lesions, but is not correlated with latency. However, K1 has the ability to immortalize and extend the life span of primary human umbilical vein endothelial cells in culture.[64] K1 homodimerization leads to the formation of signalsome, including Syk and phospholipase C, that activates downstream AP-1 and NF-AT.[65] KSHV ORF K15, a lytic gene product, is a positional and functional homolog of EBV LMP2A. K15 analogously interacts with cellular proteins, TRAF and Src kinases, and activates AP-1, NF-κB, JNK and ERK (Fig. 3b).[66]

In addition, non-protein products from EBV, EBER1 and 2 are shown to affect host cell signaling.[67] EBER1 and 2 are non-translated viral small RNA abundantly expressed in EBV latently infected cells. EBER activate the NF-κB and IRF3 pathways through association with RIG-I (Fig. 3a),[68] and EBER-mediated signaling is thought to affect cytokine expressions in EBV-infected cells.[69] These activations of NF-κB and AP-1 by viral products lead to expression of cellular IL-6, IL-8, bFGF, VEGF and others, the expression of which contributes to host cell survival and malignancy.

**PI3K-Akt.** EBV LMP1 is able to activate the PI3K/Akt pathway in rodent fibroblast through the CTAR1 domain (Fig. 3a), and PI3K inhibitor suppresses CTAR1-induced focus formation and anchorage-independent growth of rodent fibroblast cells.[70] EBV LMP2A activates the RAS/PI3K/Akt pathway (Fig. 3a), and LMP2A-ITAM-dependent Akt activation contributes to suppression of pro-apoptotic activity of forkhead transcription factor (FHKR) and cell survival of B cells from LMP2A-transgenic mice.[71,72] LMP2A-induced activation of the PI3K-Akt axis leads to mTOR-activation.[73] KSHV LANA-1 functions to promote VEGF/KDR signaling. LANA-1 binds to Daxx, and this interaction interferes with Daxx-mediated repression of Ets-dependent VEGF receptor expression (Fig. 3a).[74]

During the lytic cycle, the cytoplasmic tail of KSHV K1 has the ability to activate the PI3K/Akt pathway (Fig. 3b), and K1 lytic expression and activation of PI3K/Akt in B-cell appears to inactivate the FHKR family members and promote cell survival by preventing apoptosis.[75] The KSHV G-protein coupled receptor (vGPCR) encoded by ORF74 has been revealed to be an attractive therapeutic target for KSHV, although only a very small fraction of cells in advanced Kaposi's sarcoma lesions express it.[76,77] The expression of vGPCR, which appears to activate multiple pathways including those of MAPK, NF-κB, and Akt (Fig. 3b), leads to endothelial cell immortalization and tumorigenicity, and induces an angiogenic phenotype mediated by VEGF/KDR autocrine/paracrine signaling.[78,79] vGPCR in PEL cells activates the PI3K/Akt pathway via the heterotrimeric GTP-binding protein family.[80] VEGF/KDR autocrine/paracrine activation also promotes PI3K/Akt-mTOR signaling, and pharmacological inhibitions against this signaling axis effectively block vGPCR-induced endothelial cell proliferation.[77,79] EBV also encodes a lytic G-protein-coupled receptor protein, BILF1, that stimulates endogenous GTP-binding protein,[81] although the biological significance of BILF1 on EBV pathogenesis remains to be determined.

**Wnt.** A recent advance in γ-herpesvirus studies has uncovered another important oncogenic pathway manipulated by viral gene products. Both EBV and KSHV modulate the Wnt-signaling pathway during development of malignancy.[82] Deregulated Wnt-signaling (accumulation of β-catenin and activation of TCF) is often responsible for many types of human cancer development.[83] EBV LMP2A-mediated PI3K/Akt activation leads to phosphorylation and suppression of GSK3β in epithelial cells (Fig. 3a).[72] GSK3β is a negative regulator of Wnt signaling, and its inactivation by LMP2A results in stabilization of β-catenin and activation of TCF transcriptional factors. ITAM and phosphotyrosine motifs of LMP2A are required for accumulation and nuclear translocation of β-catenin. In addition, EBV LMP1 can

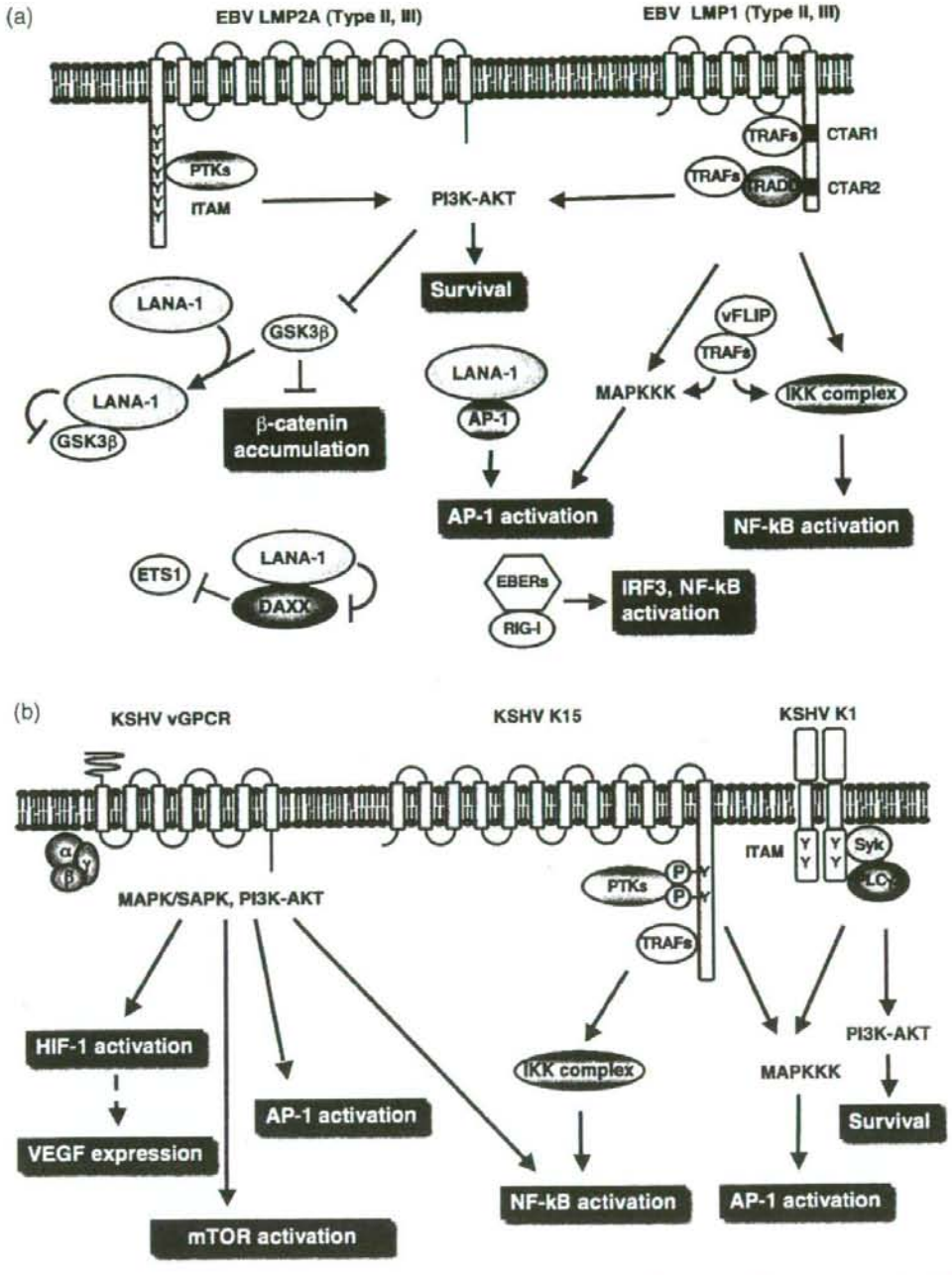


**Fig. 3.** Cell signaling manipulated by Epstein–Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV) viral products. (a) Latent phase. KSHV latent antigen LANA-1, vFLIP and EBV non-coding-RNA EBER are expressed in latently infected cells. EBV latent membrane protein (LMP)2A and LMP1 are expressed in type II and III latency. (b) Lytic phase. The KSHV G-protein coupled receptor (vGPCR), K15, and K1 induce multiple signals. The tyrosine residues (Y) in the cytoplasmic tails of K15 and K1 mediate signal initiation. Complex cell signaling manipulated by viral products is an advantage to virus-infected cell survival. Viral products are indicated by red font.



doi: 10.1111/j.1349-7006.2007.00555.x

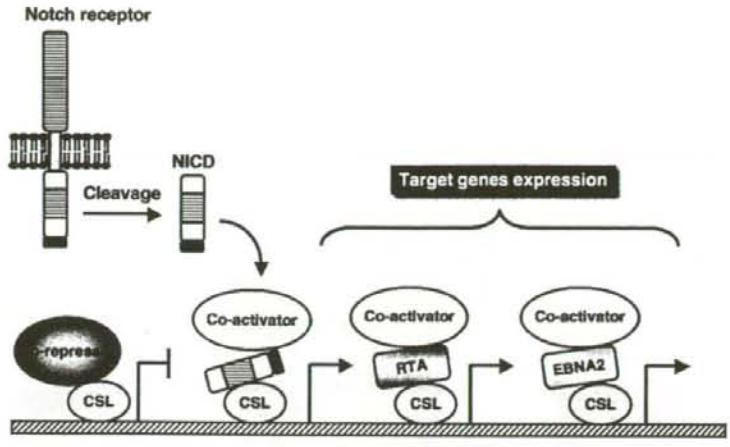


**Fig. 4.** Epstein–Barr virus (EBV) EBNA2 and Kaposi's sarcoma-associated herpesvirus (KSHV) RTA promote Notch-signaling. Ligand-receptor interaction initiate Notch signaling, and Notch receptor cleaved product, Notch intracellular domain (NICD) activates CSL-initiated transcription of target genes by recruiting transcriptional co-activators. Expression of EBNA2 induces ligand-independent CSL activation, and RTA expression activates Notch-signaling by stimulation of CSL DNA-binding. Although RTA causes reactivation of the KSHV lytic cycle, and both EBNA2 and RTA promote Notch-signaling, EBNA2 does not reactivate the KSHV lytic cycle. Viral products are indicated by red font.

up-regulate β-catenin in B lymphoma cells by inhibiting the Siah-1-dependent ubiquitination and degradation of β-catenin.[84] Stable expression of β-catenin in Kaposi's sarcoma lesion and PEL cell lines is also observed,[85] and β-catenin deregulation is suspected to play a role in KSHV-tumorigenesis. Fujimuro *et al.* found that LANA-1 causes GSK3β nuclear localization through interaction via *C*-terminal axin-like and *N*-terminal domains of LANA-1.[85] LANA-1-binding depletes cytoplasmic GSK3β, and thus subsequently allows accumulation of β-catenin to promote S-phase entry (Fig. 3a).[86,87]

**Notch.** Notch receptor and its neighboring cell ligand pathway regulate cell proliferation, differentiation, and apoptosis during development. Notch signaling is involved in physiological angiogenesis, and disruption of the Notch pathway has been shown in multiple types of tumors.[88] Ligand-binding activates receptor proteolysis and induces nuclear translocation of the cleaved intracellular domain, Notch intracellular domain (NICD), which binds to DNA-binding protein CSL/CBF-1/RBP-Jκ to convert CSL-complex to a positive transcriptional complex from a co-repressor complex. EBV EBNA2 has been shown to be partially interchangeable with NICD for transactivation of CSL/CBF-1/RBP-Jκ target genes such as anti-apoptotic bfl-1, and this modulation may contribute to EBV-mediated-transformation (Fig. 4).[89–91]

Modulation of the Notch pathway is involved in KSHV viral reactivation from latency. KSHV lytic replication may contribute to the spread of virus and progression of Kaposi's sarcoma pathogenesis, and RTA, a transactivator of the lytic gene expression program, is a switch molecule from latency to lytic replication cycle.[92] RTA binds to and transactivates CSL-complex, and RTA-targeted gene activation requires CSL (Fig. 4).[93] In addition, RTA-CSL interaction contributes to KSHV latency establishment. During the early stage of infection, RTA activates LANA-1 expression by using CSL,[37,94] and LANA-1 expression conversely represses RTA expression by down-regulation of its promoter.[95] This feedback mechanism seems to function during KSHV latency establishment. Recent analysis has also showed that NICD expression was elevated in latent Kaposi's sarcoma lesions and PEL cells.[96] Increased NICD is potentially able to reactivate KSHV from latency by activating the CSL complex on the RTA promoter. However, the above studies suggest that LANA-1, predominantly expressed in latently infected cells, represses NICD-mediated transactivation on the RTA promoter. Therefore, Notch signaling is regulating the balance between latency and the lytic cycle of KSHV.

## Apoptotic pathway deregulated by viral gene products

Viral infection sometimes causes host cell apoptosis by activating endogenous machinery during the lytic cycle. Therefore, escaping from apoptosis is advantageous for viral expansion and survival. γ-Herpesviruses, KSHV and EBV, encode anti-apoptotic genes to guard the host cell during the lytic cycle[97] (Fig. 5). The KSHV gene ORF16-encoded protein (vBcl-2) is highly homologous to Bcl-2, particularly at the BH1 and BH2 domains, and shows functionally anti-apoptotic activity.[98] vBcl-2 is expressed in spindle-shaped cells at late stages of the Kaposi's sarcoma lesions, and the anti-apoptotic effect of KSHV vBcl-2 seems to contribute to Kaposi's sarcoma progression. Similarly, EBV also encodes a Bcl-2 homolog, BHRF1, an early lytic cycle protein.[99] The function of BHRF1 resembles anti-apoptotic Bcl-2 in some types of cells, and contributes to both the initial evasion of apoptosis during early infection and the establishment of latency for cellular transformation. Another type of anti-apoptotic mode is provided by viral products related to IAP (vIAP), encoded by KSHV ORF K7.[100] ORF K7 is structurally related to survivin ΔEX3, a splice variant of survivin that inhibits apoptosis. The vIAP BIR domain interacts with active caspase-3, and the vIAP BH2 domain binds to Bcl-2. Thus, vIAP promotes anti-apoptotic activity of Bcl-2 against active caspase-3.

In addition, LANA-1 interaction with p53 leads to inactivation of p53-induced cell death during latency, which is considered to increase the chance for viral persistence and KSHV-tumorigenesis.[48] p53-induced apoptosis is also antagonized by a latent antigen of EBV, EBNA1. USP7/HAUSP is a key de-ubiquitination enzyme regulating the p53–Mdm2 pathway, and stabilizes p53 by preventing p53 protein lowering.[101] Competitive binding of EBNA1 to USP7/HAUSP disrupts p53 stabilization control balance, and protects cells from p53-induced apoptosis by destabilizing p53 protein.[102]

## Concluding remarks

Numerous studies on cancer research have shown that multiple genetic alterations cause malignancies, and here γ-herpesviruses, EBV and KSHV, are considered to be strong causes of AIDS-related malignancies. Although some viral gene products do not themselves have oncogenic activity, they contribute to virus survival and ensure persistent infection of the host cells. Molecular studies on these viral lifecycles have provided us knowledge of molecular bases for viral-associated malignancies. At present, these viral gene products from γ-herpesviruses are

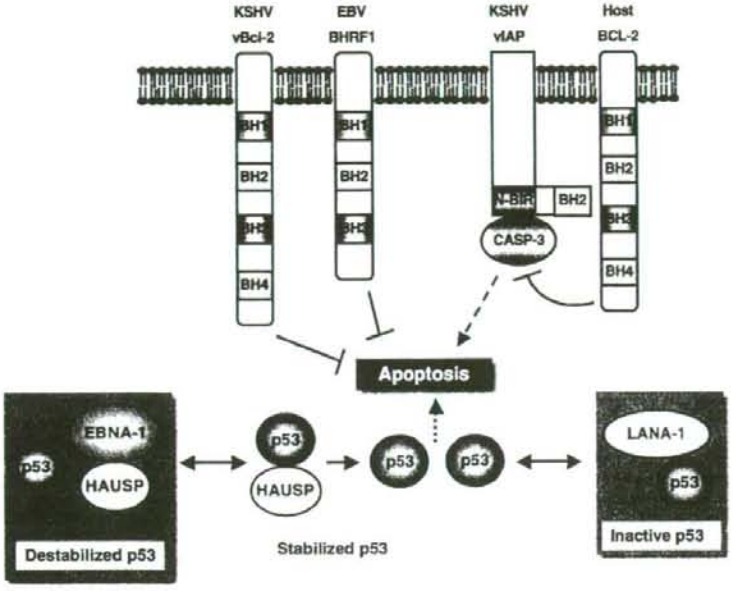


Fig. 5. Anti-apoptotic effects by viral products. Epstein–Barr virus (EBV) BHRF1 and Kaposi's sarcoma-associated herpesvirus (KSHV) vBcl-2 are functionally related to host cell anti-apoptotic Bcl-2 family containing Bcl-2-homology region 1–4 (BH1-4). KSHV vIAP bridges active caspase-3 (CASP-3) and host cell Bcl-2 through baculovirus IAP repeat domain (N-BIR) to inhibit apoptosis-execution by caspase-3. Latent antigens, EBV EBNA1 and KSHV LANA-1 suppress p53 activity and prevent p53-induced apoptosis execution. Viral products are indicated by red font.

found to manipulate host cell signaling to initiate cell transformation and contribute to the progression and development of some types of cancer. Therefore, all these viral proteins modulating cell signaling are recognized as selective targets for chemotherapy/chemoprevention of EBV and KSHV-associated malignancies. Admittedly, there are many unsolved questions in AIDS-related malignancies, and such questions must be investigated to understand virus-associated malignancies. Further extensive research is crucial but holds promise for the development of selective molecular-targeted therapy for AIDS-associated malignancies.

## Acknowledgments

We are grateful to Drs Harutaka Katano and Shizuko Harada (National Institute of Infectious Diseases), and Masahiro Fujimuro (Hokkaido University) for valuable comments on this manuscript. We apologize that we could not cite the excellent work of many investigators owing to space limitations. This work was supported by a grant-in-aid for Cancer Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

1 Boshoff C, Weiss R. AIDS-related malignancies. *Nat Rev Cancer* 2002; **2**: 373–82.
2 Damania B. Oncogenic gamma-herpesviruses: comparison of viral proteins involved in tumorigenesis. *Nat Rev Microbiol* 2004; **2**: 656–68.
3 Chang Y, Cesarman E, Pessin MS *et al.* Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 1994; **266**: 1865–9.
4 Katano H, Sato Y, Kurata T *et al.* High expression of HHV-8-encoded ORF73 protein in spindle-shaped cells of Kaposi's sarcoma. *Am J Pathol* 1999; **155**: 47–52.
5 Boshoff C, Schulz TF, Kennedy MM *et al.* Kaposi's sarcoma-associated herpesvirus infects endothelial and spindle cells. *Nat Med* 1995; **1**: 1274–8.
6 Hayward GS. Initiation of angiogenic Kaposi's sarcoma lesions. *Cancer Cell* 2003; **3**: 1–3.
7 Pati S, Foulke JS Jr, Barabitskaya O *et al.* Human herpesvirus 8-encoded vGPCR activates nuclear factor of activated T cells and collaborates with human immunodeficiency virus type 1 Tat. *J Virol* 2003; **77**: 5759–73.
8 Barillari G, Sgadari C, Palladino C *et al.* Inflammatory cytokines synergize with the HIV-1 Tat protein to promote angiogenesis and Kaposi's sarcoma via induction of basic fibroblast growth factor and the alpha v beta 3 integrin. *J Immunol* 1999; **163**: 1929–35.
9 Aoki Y, Tosato G. HIV-1 Tat enhances Kaposi sarcoma-associated herpesvirus (KSHV) infectivity. *Blood* 2004; **104**: 810–14.
10 Thorley-Lawson DA. Epstein-Barr virus: exploiting the immune system. *Nat Rev Immunol* 2001; **1**: 75–82.
11 Voo KS, Fu T, Wang HY *et al.* Evidence for the presentation of major histocompatibility complex class I-restricted Epstein-Barr virus nuclear antigen 1 peptides to CD8 + T lymphocytes. *J Exp Med* 2004; **199**: 459–70.
12 Levitskaya J, Coram M, Levitsky V *et al.* Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen-1. *Nature* 1995; **375**: 685–8.
13 Yin Y, Manoury B, Fahraeus R. Self-inhibition of synthesis and antigen presentation by Epstein-Barr virus-encoded EBNA1. *Science* 2003; **301**: 1371–4.
14 Soulier J, Grollet L, Oksenhendler E *et al.* Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease. *Blood* 1995; **86**: 1276–80.
15 Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med* 2004; **10**: 789–99.
16 Tsuruo T, Naito M, Tomida A *et al.* Molecular targeting therapy of cancer: drug resistance, apoptosis and survival signal. *Cancer Sci* 2003; **94**: 15–21.
17 Yates JL, Warren N, Sugden B. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. *Nature* 1985; **313**: 812–15.
18 Hu J, Garber AC, Renne R. The latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus supports latent DNA replication in dividing cells. *J Virol* 2002; **76**: 11 677–87.
19 Wang J, Sugden B. Origins of bidirectional replication of Epstein-Barr virus: models for understanding mammalian origins of DNA synthesis. *J Cell Biochem* 2005; **94**: 247–56.
20 Goldsmith K, Bendell L, Frappier L. Identification of EBNA1 amino acid sequences required for the interaction of the functional elements of the Epstein-Barr virus latent origin of DNA replication. *J Virol* 1993; **67**: 3418–26.
21 Bochkarev A, Barwell JA, Pfuetzner RA *et al.* Crystal structure of the DNA-binding domain of the Epstein-Barr virus origin-binding protein, EBNA1, bound to DNA. *Cell* 1996; **84**: 791–800.
22 Bashaw JM, Yates JL. Replication from oriP of Epstein-Barr virus requires exact spacing of two bound dimers of EBNA1 which bend DNA. *J Virol* 2001; **75**: 10 603–11.



doi: 10.1111/j.1349-7006.2007.00555.x

23 Schepers A, Ritzi M, Bousset K *et al.* Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus. *EMBO J* 2001; **20**: 4588–602.
24 Dhar SK, Yoshida K, Machida Y *et al.* Replication from oriP of Epstein-Barr virus requires human ORC and is inhibited by geminin. *Cell* 2001; **106**: 287–96.
25 Noguchi K, Vassilev A, Ghosh S *et al.* The BAH domain facilitates the ability of human Orc1 protein to activate replication origins in vivo. *EMBO J* 2006; **25**: 5372–82.
26 Deng Z, Lezina L, Chen CJ *et al.* Telomeric proteins regulate episomal maintenance of Epstein-Barr virus origin of plasmid replication. *Mol Cell* 2002; **9**: 493–503.
27 Rawlins DR, Milman G, Hayward SD *et al.* Sequence-specific DNA binding of the Epstein-Barr virus nuclear antigen (EBNA-1) to clustered sites in the plasmid maintenance region. *Cell* 1985; **42**: 859–68.
28 Reisman D, Sugden B. trans activation of an Epstein-Barr viral transcriptional enhancer by the Epstein-Barr viral nuclear antigen 1. *Mol Cell Biol* 1986; **6**: 3838–46.
29 Lim C, Sohn H, Lee D *et al.* Functional dissection of latency-associated nuclear antigen 1 of Kaposi's sarcoma-associated herpesvirus involved in latent DNA replication and transcription of terminal repeats of the viral genome. *J Virol* 2002; **76**: 10 320–31.
30 Hu J, Renne R. Characterization of the minimal replicator of Kaposi's sarcoma-associated herpesvirus latent origin. *J Virol* 2005; **79**: 2637–42.
31 Verma SC, Choudhuri T, Robertson ES. The minimal replicator element of KSHV terminal repeat supports replication in a semi-conservative and cell cycle dependent manner. *J Virol* 2007; **81**: 3402–13.
32 Barbera AJ, Chodaparambil JV, Kelley-Clarke B *et al.* The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science* 2006; **311**: 856–61.
33 Komatsu T, Ballestas ME, Barbera AJ *et al.* KSHV LANA1 binds DNA as an oligomer and residues N-terminal to the oligomerization domain are essential for DNA binding, replication, and episome persistence. *Virology* 2004; **319**: 225–36.
34 Stedman W, Deng Z, Lu F *et al.* ORC, MCM, and histone hyperacetylation at the Kaposi's sarcoma-associated herpesvirus latent replication origin. *J Virol* 2004; **78**: 12 566–75.
35 Verma SC, Choudhuri T, Kaul R *et al.* Latency-associated nuclear antigen (LANA) of Kaposi's sarcoma-associated herpesvirus interacts with origin recognition complexes at the LANA binding sequence within the terminal repeats. *J Virol* 2006; **80**: 2243–56.
36 Ohsaki E, Ueda K, Sakakibara S *et al.* Poly (ADP-ribose) polymerase 1 binds to Kaposi's sarcoma-associated herpesvirus (KSHV) terminal repeat sequence and modulates KSHV replication in latency. *J Virol* 2004; **78**: 9936–46.
37 Lan K, Kuppers DA, Verma SC *et al.* Induction of Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen by the lytic transactivator RTA. a novel mechanism for establishment of latency. *J Virol* 2005; **79**: 7453–65.
38 Li M, Lee H, Yoon DW *et al.* Kaposi's sarcoma-associated herpesvirus encodes a functional cyclin. *J Virol* 1997; **71**: 1984–91.
39 Verschuren EW, Jones N, Evan GI. The cell cycle and how it is steered by Kaposi's sarcoma-associated herpesvirus cyclin. *J General Virol* 2004; **85**: 1347–61.
40 Van Dross R, Yao S, Asad S *et al.* Constitutively active K-cyclin/cdk6 kinase in Kaposi sarcoma-associated herpesvirus-infected cells. *J Natl Cancer Inst* 2005; **97**: 656–66.
41 Kaldis P, Ojala PM, Tong L *et al.* CAK-independent activation of CDK6 by a viral cyclin. *Mol Biol Cell* 2001; **12**: 3987–99.
42 Ellis M, Chew YP, Fallis L *et al.* Degradation of p27 (Kip) cdk inhibitor triggered by Kaposi's sarcoma virus cyclin-cdk6 complex. *EMBO J* 1999; **18**: 644–53.
43 Mann DJ, Child ES, Swanton C *et al.* Modulation of p27 (Kip1) levels by the cyclin encoded by Kaposi's sarcoma-associated herpesvirus. *EMBO J* 1999; **18**: 654–63.
44 Sarek G, Jarviluoma A, Ojala PM. KSHV viral cyclin inactivates p27KIP1 through Ser10 and Thr187 phosphorylation in proliferating primary effusion lymphomas. *Blood* 2006; **107**: 725–32.
45 Jarviluoma A, Child ES, Sarek G *et al.* Phosphorylation of the cyclin-dependent kinase inhibitor p21Cip1 on serine 130 is essential for viral cyclin-mediated bypass of a p21Cip1-imposed G1 arrest. *Mol Cell Biol* 2006; **26**: 2430–40.
46 Swanton C, Mann DJ, Fleckenstein B *et al.* Herpes viral cyclin/Cdk6 complexes evade inhibition by CDK inhibitor proteins. *Nature* 1997; **390**: 184–7.
47 Talbot SJ, Weiss RA, Kellam P *et al.* Transcriptional analysis of human herpesvirus-8 open reading frames 71, 72, 73, K14, and 74 in a primary effusion lymphoma cell line. *Virology* 1999; **257**: 84–94.
48 Friborg J Jr, Kong W, Hottiger MO *et al.* p53 inhibition by the LANA protein of KSHV protects against cell death. *Nature* 1999; **402**: 889–94.
49 Radkov SA, Kellam P, Boshoff C. The latent nuclear antigen of Kaposi sarcoma-associated herpesvirus targets the retinoblastoma-E2F pathway and with the oncogene Hras transforms primary rat cells. *Nat Med* 2000; **6**: 1121–7.
50 Katano H, Sato Y, Kurata T *et al.* Expression and localization of human herpesvirus 8-encoded proteins in primary effusion lymphoma, Kaposi's sarcoma, and multicentric Castleman's disease. *Virology* 2000; **269**: 335–44.
51 Fakhari FD, Jeong JH, Kanan Y *et al.* The latency-associated nuclear antigen of Kaposi sarcoma-associated herpesvirus induces B cell hyperplasia and lymphoma. *J Clin Invest* 2006; **116**: 735–42.
52 Shin YC, Nakamura H, Liang X *et al.* Inhibition of the ATM/p53 signal transduction pathway by Kaposi's sarcoma-associated herpesvirus interferon regulatory factor 1. *J Virol* 2006; **80**: 2257–66.
53 Liu L, Eby MT, Rathore N *et al.* The human herpes virus 8-encoded viral FLICE inhibitory protein physically associates with and persistently activates the Ikappa B kinase complex. *J Biol Chem* 2002; **277**: 13 745–51.
54 Field N, Low W, Daniels M *et al.* KSHV vFLIP binds to IKK-gamma to activate IKK. *J Cell Sci* 2003; **116**: 3721–8.
55 Mosialos G, Birkenbach M, Yalamanchili R *et al.* The Epstein-Barr virus transforming protein LMP1 engages signaling proteins for the tumor necrosis factor receptor family. *Cell* 1995; **80**: 389–99.
56 Izumi KM, Kieff ED. The Epstein-Barr virus oncogene product latent membrane protein 1 engages the tumor necrosis factor receptor-associated death domain protein to mediate B lymphocyte growth transformation and activate NF-kappaB. *Proc Natl Acad Sci USA* 1997; **94**: 12 592–7.
57 Sylla BS, Hung SC, Davidson DM *et al.* Epstein-Barr virus-transforming protein latent infection membrane protein 1 activates transcription factor NF-kappaB through a pathway that includes the NF-kappaB-inducing kinase and the IkappaB kinases IKKalpha and IKKbeta. *Proc Natl Acad Sci USA* 1998; **95**: 10 106–11.
58 Kieser A, Kilger E, Gires O *et al.* Epstein-Barr virus latent membrane protein-1 triggers AP-1 activity via the c-Jun N-terminal kinase cascade. *EMBO J* 1997; **16**: 6478–85.
59 An J, Sun Y, Sun R *et al.* Kaposi's sarcoma-associated herpesvirus encoded vFLIP induces cellular IL-6 expression: the role of the NF-kappaB and JNK/AP1 pathways. *Oncogene* 2003; **22**: 3371–85.
60 Murono S, Inoue H, Tanabe T *et al.* Induction of cyclooxygenase-2 by Epstein-Barr virus latent membrane protein 1 is involved in vascular endothelial growth factor production in nasopharyngeal carcinoma cells. *Proc Natl Acad Sci USA* 2001; **98**: 6905–10.
61 An J, Sun Y, Rettig MB. Transcriptional coactivation of c-Jun by the KSHV-encoded LANA. *Blood* 2004; **103**: 222–8.
62 Chen SY, Lu J, Shih YC *et al.* Epstein-Barr virus latent membrane protein 2A regulates c-Jun protein through extracellular signal-regulated kinase. *J Virol* 2002; **76**: 9556–61.
63 Lanier LL. Viral immunoreceptor tyrosine-based activation motif (ITAM) - mediated signaling in cell transformation and cancer. *Trends Cell Biol* 2006; **16**: 388–90.
64 Wang L, Dittmer DP, Tomlinson CC *et al.* Immortalization of primary endothelial cells by the K1 protein of Kaposi's sarcoma-associated herpesvirus. *Cancer Res* 2006; **66**: 3658–66.
65 Lagunoff M, Majeti R, Weiss A *et al.* Deregulated signal transduction by the K1 gene product of Kaposi's sarcoma-associated herpesvirus. *Proc Natl Acad Sci USA* 1999; **96**: 5704–9.
66 Brinkmann MM, Glenn M, Rainbow L *et al.* Activation of mitogen-activated protein kinase and NF-kappaB pathways by a Kaposi's sarcoma-associated herpesvirus K15 membrane protein. *J Virol* 2003; **77**: 9346–58.
67 Komano J, Maruo S, Kurozumi K *et al.* Oncogenic role of Epstein-Barr virus-encoded RNAs in Burkitt's lymphoma cell line Akata. *J Virol* 1999; **73**: 9827–31.
68 Samanta M, Iwakiri D, Kanda T *et al.* EB virus-encoded RNAs are recognized by RIG-I and activate signaling to induce type I IFN. *EMBO J* 2006; **25**: 4207–14.
69 Iwakiri D, Eizuru Y, Tokunaga M *et al.* Autocrine growth of Epstein-Barr virus-positive gastric carcinoma cells mediated by an Epstein-Barr virus-encoded small RNA. *Cancer Res* 2003; **63**: 7062–7.
70 Mainou BA, Everly DN Jr, Raab-Traub N. Epstein-Barr virus latent membrane protein 1 CTAR1 mediates rodent and human fibroblast transformation through activation of PI3K. *Oncogene* 2005; **24**: 6917–24.
71 Swart R, Ruf IK, Sample J *et al.* Latent membrane protein 2A-mediated effects on the phosphatidylinositol 3-Kinase/Akt pathway. *J Virol* 2000; **74**: 10 838–45.
72 Morrison JA, Klingelhutz AJ, Raab-Traub N. Epstein-Barr virus latent membrane protein 2A activates beta-catenin signaling in epithelial cells. *J Virol* 2003; **77**: 12 276–84.
73 Moody CA, Scott RS, Amirghahari N *et al.* Modulation of the cell growth regulator mTOR by Epstein-Barr virus-encoded LMP2A. *J Virol* 2005; **79**: 5499–506.
74 Murakami Y, Yamagoe S, Noguchi K *et al.* 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. *J Biol Chem* 2006; **281**: 28 113–21.
75 Tomlinson CC, Damania B. The K1 protein of Kaposi's sarcoma-associated herpesvirus activates the Akt signaling pathway. *J Virol* 2004; **78**: 1918–27.
76 Montaner S, Sodhi A, Ramsdell AK *et al.* The Kaposi's sarcoma-associated herpesvirus G protein-coupled receptor as a therapeutic target for the treatment of Kaposi's sarcoma. *Cancer Res* 2006; **66**: 168–74.
77 Sodhi A, Chaisuparat R, Hu J *et al.* The TSC2/mTOR pathway drives endothelial cell transformation induced by the Kaposi's sarcoma-associated herpesvirus G protein-coupled receptor. *Cancer Cell* 2006; **10**: 133–43.
78 Montaner S, Sodhi A, Molinolo A *et al.* Endothelial infection with KSHV genes in vivo reveals that vGPCR initiates Kaposi's sarcomagenesis and can promote the tumorigenic potential of viral latent genes. *Cancer Cell* 2003; **3**: 23–36.
79 Bais C, Van Geelen A, Eroles P *et al.* Kaposi's sarcoma associated herpesvirus G protein-coupled receptor immortalizes human endothelial cells by activation of the VEGF receptor-2/KDR. *Cancer Cell* 2003; **3**: 131–43.
80 Cannon M, Philpott NJ, Cesarman E. The Kaposi's sarcoma-associated herpesvirus G protein-coupled receptor has broad signaling effects in primary effusion lymphoma cells. *J Virol* 2003; **77**: 57–67.
81 Paulsen SJ, Rosenkilde MM, Eugen-Olsen J *et al.* Epstein-Barr virus-encoded BILF1 is a constitutively active G protein-coupled receptor. *J Virol* 2005; **79**: 536–46.
82 Fujimuro M, Hayward SD. Manipulation of glycogen-synthase kinase-3 activity in KSHV-associated cancers. *J Mol Med* 2004; **82**: 223–31.
83 Kikuchi A. Tumor formation by genetic mutations in the components of the Wnt signaling pathway. *Cancer Sci* 2003; **94**: 225–9.
84 Jang KL, Shackelford J, Seo SY *et al.* Up-regulation of beta-catenin by a viral oncogene correlates with inhibition of the seven in absentia homolog 1 in B lymphoma cells. *Proc Natl Acad Sci USA* 2005; **102**: 18 431–6.
85 Fujimuro M, Wu FY, ApRhys C *et al.* A novel viral mechanism for dysregulation of beta-catenin in Kaposi's sarcoma-associated herpesvirus latency. *Nat Med* 2003; **9**: 300–6.
86 Fujimuro M, Liu J, Zhu J *et al.* Regulation of the interaction between glycogen synthase kinase 3 and the Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen. *J Virol* 2005; **79**: 10 429–41.
87 Fujimuro M, Hayward SD. The latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus manipulates the activity of glycogen synthase kinase-3beta. *J Virol* 2003; **77**: 8019–30.
88 Rehman AO, Wang CY. Notch signaling in the regulation of tumor angiogenesis. *Trends Cell Biol* 2006; **16**: 293–300.
89 Hsieh JJ, Hayward SD. Masking of the CBF1/RBPJ kappa transcriptional repression domain by Epstein-Barr virus EBNA2. *Science* 1995; **268**: 560–3.
90 Henkel T, Ling PD, Hayward SD *et al.* Mediation of Epstein-Barr virus EBNA2 transactivation by recombination signal-binding protein J kappa. *Science* 1994; **265**: 92–5.
91 Pegman PM, Smith SM, D'Souza BN *et al.* Epstein-Barr virus nuclear antigen 2 trans-activates the cellular antiapoptotic bfl-1 gene by a CBF1/RBPJ kappa-dependent pathway. *J Virol* 2006; **80**: 8133–44.
92 Lukac DM, Renne R, Kirshner JR *et al.* Reactivation of Kaposi's sarcoma-associated herpesvirus infection from latency by expression of the ORF 50 transactivator, a homolog of the EBV R protein. *Virology* 1998; **252**: 304–12.
93 Liang Y, Chang J, Lynch SJ *et al.* The lytic switch protein of KSHV activates gene expression via functional interaction with RBP-Jkappa (CSL), the target of the Notch signaling pathway. *Genes Dev* 2002; **16**: 1977–89.
94 Lan K, Kuppers DA, Robertson ES. Kaposi's sarcoma-associated herpesvirus reactivation is regulated by interaction of latency-associated nuclear antigen with recombination signal sequence-binding protein Jkappa, the major downstream effector of the Notch signaling pathway. *J Virol* 2005; **79**: 3468–78.
95 Lan K, Kuppers DA, Verma SC *et al.* Kaposi's sarcoma-associated herpesvirus-encoded latency-associated nuclear antigen inhibits lytic replication by targeting Rta: a potential mechanism for virus-mediated control of latency. *J Virol* 2004; **78**: 6585–94.
96 Lan K, Murakami M, Choudhuri T *et al.* Intracellular-activated Notch1 can reactivate Kaposi's sarcoma-associated herpesvirus from latency. *Virology* 2006; **351**: 393–403.
97 Cuconati A, White E. Viral homologs of BCL-2: role of apoptosis in the regulation of virus infection. *Genes Dev* 2002; **16**: 2465–78.
98 Sarid R, Sato T, Bohenzky RA *et al.* Kaposi's sarcoma-associated herpesvirus encodes a functional bcl-2 homologue. *Nat Med* 1997; **3**: 293–8.
99 Henderson S, Huen D, Rowe M *et al.* Epstein-Barr virus-coded BHRF1 protein, a viral homologue of Bcl-2, protects human B cells from programmed cell death. *Proc Natl Acad Sci USA* 1993; **90**: 8479–83.
100 Wang HW, Sharp TV, Koumi A *et al.* Characterization of an anti-apoptotic glycoprotein encoded by Kaposi's sarcoma-associated herpesvirus which resembles a spliced variant of human survivin. *EMBO J* 2002; **21**: 2602–15.
101 Li M, Brooks CL, Kon N *et al.* A dynamic role of HAUSP in the p53-Mdm2 pathway. *Mol Cell* 2004; **13**: 879–86.
102 Saridakis V, Sheng Y, Sarkari F *et al.* Structure of the p53 binding domain of HAUSP/USP7 bound to Epstein-Barr nuclear antigen 1 implications for EBV-mediated immortalization. *Mol Cell* 2005; **18**: 25–36.



doi: 10.1111/j.1349-7006.2007.00555.x

〔ウイルス　第56巻　第1号，pp.9-16，2006〕

総　説

# 2. 酵母を用いた動物ウイルスの研究

森川　裕子
北里大学北里生命科学研究所

大腸菌が原核細胞のモデルであるように，酵母は真核細胞のモデルとしてしばしば用いられる．古くから酵母は詳細に解析され，近年そのゲノムも完全解読された．高等真核細胞との類似性が高くかつ分子遺伝学や分子生物学の解析手法が応用できることから，酵母は真核細胞のモデル細胞として多用されている．こうした酵母細胞で動物ウイルスの複製が再現できることが近年明らかとなってきた．本稿では，酵母の遺伝子発現と酵母における動物ウイルスの複製を概説する．特筆すべきは，酵母における完全なウイルス複製過程の成立と酵母を用いたウイルス様粒子ワクチンの製造である．酵母を用いた近年の研究は，ウイルス複製に関与する宿主因子や宿主機構の同定を中心に展開されており，それらは酵母遺伝学の利用や酵母の全遺伝子破壊株ライブラリを用いた網羅的解析によって推進されている．

## 1. はじめに

酵母 Yeast は単細胞性の fungi であり，厳密にいうと「出芽・分裂によって増殖し，性接合を行う ascomycetous あるいは basidiomycetous fungi」の微生物細胞である．出芽酵母の代表として Saccharomyces cerevisiae（パン酵母）が，分裂酵母の代表として Schizosaccharomyces pombe がよく知られている．酵母は大腸菌のように取り扱いが簡単な細胞であるにもかかわらず，高等真核細胞との類似性が高くかつ分子生物学・細胞生物学の解析手法が応用できることから，真核細胞のモデル細胞として多用されている．特筆すべきは遺伝子破壊や変異導入などの分子遺伝学（特に，逆遺伝学）である．また，酵母 Two-Hybrid Assay は結合する蛋白質をライブラリから探索する方法としてよく知られているところである．

出芽酵母には内在性 RNA ウイルス（L-A ウイルスやレトロエレメント Ty1-5）が存在することが古くから知られている．これらは細胞外感染経路をもたず，接合などの細胞質混合で伝達される．酵母の細胞外からのウイルス感染を最初に報告したのはタバコモザイクウイルスを用いてである[5]．下等真核細胞である酵母は細胞壁があるという点では植物細胞であり，植物ウイルスの報告があるのはなるほどと思われる．しかし近年，動物ウイルスの増殖も酵母で再現できることが明らかとなってきた．本稿では，酵母が動物ウイルスの研究にどう利用されているのか，またどう利用できるのか，実例と利用の利点・注意点などを概説する．

連絡先
〒108-8641 東京都港区白金5丁目9番1号
北里大学北里生命科学研究所
TEL: 03-5791-6129
E-mail: morikawa@lisci.kitasato-u.ac.jp

## 2. 酵母における遺伝子発現

ウイルス研究者にとって，高等真核細胞はウイルスを増殖させるための宿主細胞であり，大腸菌はウイルス遺伝子をクローニングするための細胞である．しかし，酵母には馴染みが薄いと思う．酵母におけるウイルス遺伝子の発現を紹介する前に，酵母自体の遺伝子発現について少し解説する．

### 1）酵母の遺伝子発現系

#### ①DNA 複製

出芽酵母は DNA 複製開始部位の構造がわかっている唯一の真核生物である．酵母細胞内で複製開始点となる DNA 配列を自律複製配列(Autonomously Replicating Sequence: ARS)と呼び，5'-(T/A)TTTA(C/T)(A/G)TTT(T/A)-3'-という共通配列が見出されている．ただし，分裂酵母の ARS とは一般的に交換できない．この ARS に加え，分極に関与するセントロメア CEN と染色体安定性に関与するテロメア TEL が染色体機能として必要である．酵母における DNA 複製機構の解析から同様の複製機構がヒト細胞で

も存在すると推測されている.

②転写

酵母から動物細胞まで,DNA依存性RNAポリメラーゼの基本的な構造や機能は共通であり,酵母RNAポリメラーゼはPol A, B, C(動物細胞のPol I, II, IIIに相当)と呼ばれている.mRNAの5'末端には$m^7$GpppNのキャップ構造が付加され,モノシストロニックである.

動物細胞に比べ,出芽酵母の遺伝子にはイントロン-エクソン構造が少ない(約3%).スプライシング機構は動物細胞の場合とほとんど同じであるが,スプライシングシグナルの配列がやや異なる.分裂酵母では遺伝子の3-4割がイントロンをもつ.

③翻訳

原核生物では翻訳開始点の上流にリボソーム結合部位が存在するのに対し,酵母を含め真核生物ではそれが認められない.また酵母では,転写開始点と翻訳開始コドンAUGの間の距離は長くても短くても翻訳レベルは変わらない.しかしながら,mRNAの2次構造(特に,Gと多く含む配列)は翻訳を強く阻害する.高等真核細胞では翻訳開始コドンAUGの周辺にはKozakのコンセンサス配列CCACCAUGGCが認められるが,酵母ではこれがAAAAAAAAUGUCUである[13].ただし,変異させても翻訳効率の低下はわずかである[3].酵母のリボソームは一度翻訳を終結すると翻訳を再開できないといわれている.

2) 酵母におけるウイルス遺伝子の発現

①ウイルス遺伝子の複製・転写

一般的にDNAウイルスの遺伝子の複製は,ウイルスゲノム長が短いウイルス(パポーバ科やパルボ科)では宿主細胞酵素系への依存性が高く,宿主のDNAポリメラーゼを利用する.転写はポックス科を除いてすべてのDNAウイルスが宿主細胞のRNA Pol IIに依存している.出芽酵母におけるDNAウイルスの複製・転写は,最も宿主細胞依存性が高いと思われるパポーバ科のパピローマウイルスを用いて成功している[2, 22, 49, 50].酵母細胞内で安定して複製するには自律複製配列ARSなどが必要であるが,パピローマウイルスDNAのL1-LCR領域に酵母細胞内での自律複製に関与する配列[22]つまり弱いながらもARSとの塩基配列相同性のある部位が見出されている[49].この酵母におけるウイルスmRNAの転写を直接証明した報告はないが,驚くべきことに,細胞壁を消化した酵母スフェロプラスト細胞にパピローマウイルス粒子を取り込ませると,ウイルスゲノムDNAがパッケージングされた感染性ウイルス粒子が回収されることから[50],酵母細胞で完全なウイルス複製過程が進行すると思われる.

出芽酵母における(+)鎖RNAウイルスの複製・転写は,この分野の第1人者であるP. Ahlquistらが植物ウイルスのブロモモザイクウイルス(BMV)を用いて精力的に解析している.このウイルスの蛋白質1a(RNA helicaseとキャッピング機能をもつ)と蛋白質2a(RNA依存性RNA合成酵素)を共発現させた酵母細胞では,これら蛋白質がウイルスゲノムRNAと複合体を形成し,転写・複製がおこることが明らかになっている[12, 19, 20, 39, 46].さらに,酵母細胞における完全なウイルス複製過程がノダ科の昆虫ウイルスを用いて示されており,感染性ウイルス粒子が回収されている[36, 37, 38].これに対し,(-)鎖RNAウイルスの複製・転写ついて詳細な報告はないが,水泡性口炎ウイルス(VSV)粒子を取り込ませた酵母スフェロプラスト細胞ではウイルス蛋白の発現が観察されている.ただし,その著者らはウイルスゲノムが複製されたとはあまり考えていないようである[27].

②ウイルスmRNAの翻訳

上記のように,酵母で感染性ウイルスが産生されることから,間接的ではあるが,多くのウイルスmRNAは正確に翻訳されると考えられる.気をつけねばならない点はInternal Ribosome Entry Site(IRES)である.上述したように,酵母では転写開始点と翻訳開始点の距離はさほど問題ではないが,mRNAの2次構造は翻訳に致命的である.そのせいなのか不明だが,酵母においてはヒトC型肝炎ウイルス(HCV)IRESの翻訳成功例[17, 40]とポリオウイルスIRESの翻訳不成功例[6]が報告されている.後者の不成功例では,酵母細胞からLa自己抗原のIRESへの結合を阻害する60塩基長RNAが見出されている[7, 8, 9].筆者らもヒト免疫不全ウイルス1型(HIV-1)mRNAが酵母で翻訳されないことを見出している.ウイルスmRNAと酵母のRNA helicaseとの相性も指摘されている[33].

### 3. 酵母を用いた蛋白質発現とその応用

酵母は真核細胞のモデルとして代用されるだけでなく,蛋白質産生の細胞としても利用されてきた.異種蛋白質発現のための酵母の遺伝子操作については多くの総説があると思うのでそれを参考にしてほしい.入門書としては柴垣と水本による総説「酵母発現系とその実際」がわかりやすい[42].

蛋白質発現のための細胞として,大腸菌や高等真核細胞など様々な細胞が用いられる.**表1**にそれらの特徴をまとめた.酵母も大腸菌と同じで操作は簡単であり,寒天プレートで形質転換コロニーを形成させ,これを液体培地で大量に振盪培養する.ただし,酵母の生育は大腸菌より遅いので各ステップはovernightではなく2日はかかる.酵母の欠点として,①分厚い細胞壁をもつので細胞粉砕が困難である,②蛋白質分解酵素の活性が高く蛋白質を精製しづらい,③さほど発現量が高くないなどが挙げられる.昆虫細胞や哺乳類細胞での発現は,transfectionによる一過性発現(大量発現に不向き)か,transfectantの細胞株を樹立する(時間がかかる)かであり,そうでなければウイル

表1 蛋白質発現系の比較

| | 大腸菌 | 酵母 | 昆虫細胞 | 哺乳類細胞 |
|---|---|---|---|---|
| 培養形式 | 浮遊 | 浮遊 | 浮遊 | 接着 |
| 細胞倍加時間 | 30分 | 90-180分 | 24時間 | 24時間以上 |
| 蛋白質発現量 | 大 | 小 | 中 | 極小 |
| 蛋白質の翻訳後修飾 | | | | |
| 糖鎖付加 | なし | マンナン型 | 高マンノース型 | 複合型 |
| リン酸化 | His, Asp | Ser, Thr, Tyr | Ser, Thr, Tyr | Ser, Thr, Tyr |
| アシル化 | なし | あり | あり | あり |

スベクターに因らねばならないという難点があり，選択に迷うところである．

大腸菌とは異なる利点として，酵母では翻訳後修飾がおこることが挙げられる．ただし，酵母における糖鎖付加は問題で，その付加シグナル配列は高等真核細胞と同じであるが，マンナンと呼ばれる巨大なマンノースポリマー構造が形成される．例えば，HIV-1のエンベロープ蛋白gp120（高等真核細胞では120kDaの蛋白質として生成されこの分子量の約半分を糖鎖が占める）を酵母で発現させると600kDa以上になり抗体で認識されない[14]．しかし，糖鎖付加部位が少ない場合は活性あるエンベロープ蛋白となるようで，VSV G蛋白を発現させた酵母スフェロプラスト細胞では膜融合による多核巨細胞が形成される[27]．酵母における糖鎖付加のもう1つの問題は，複合型の糖鎖付加がおこらないことである（ただし，この点は昆虫細胞で発現させた場合も同じで，昆虫細胞でも複合型の糖鎖付加はおこらない）．こうした問題点を解決する目的で，糖鎖改変酵母株の分子育種がさかんに行われている．すなわち，マンナン型糖鎖ができないように酵母染色体遺伝子（OCH1, MNN1, MNN4遺伝子など）を破壊し，さらに複合型糖鎖が生成されるようヒト遺伝子（Mannosidase I, IIを含む7つの遺伝子）を導入した酵母株の作出が試みられている[4, 21, 32]．ヒト遺伝子をもたせた酵母株はまだ市販されていないが，酵母の糖鎖遺伝子破壊株はAmerican Type Culture Collection（ATCC）から入手できる．

酵母の蛋白質発現系が大変有用であると認識されたのは，酵母を用いたヒトB型肝炎ウイルス（HBV）ワクチンの製造によってである．当時はヒト感染血清から危険を冒しながらHBV粒子を精製しそれを不活化してワクチンとしていた．しかし，P. Valenzuelaらは HBVの表面抗原であるHBsAgを酵母で発現させたところ，球状のHBsAg粒子が形成されていることを見出した[31, 45]．さらに，その細胞破砕液から精製したウイルス様粒子（VLP）はヒト感染血清由来不活化ワクチンと同等の力価をもつことが明らかにされ[29]，現行のHBVワクチンとなっている．酵母細胞におけるVLP形成はパピローマウイルス[15, 16]やBMV[23]でも報告されている．また近年，HIV-1でも示され[41]，そのVLPの免疫原性が検討されている[43]．

### 3. 酵母における細胞生物学

1）細胞の構造と機能

酵母は核と細胞質が核膜で分離された真核細胞であるが，①有糸分裂の時も核膜（ラミンの裏打ち構造はない）が消失しないこと，②細胞は厚い細胞壁に覆われていることが，動物細胞との決定的な違いである．しかし，その他の細胞内構造はよく似ている．粗面小胞体は核膜と連続したものと，網状のネットワークが形質膜近くに存在する．ゴルジ装置はシス・ミディアム・トランスに分かれた層状構造として認められる．初期エンドソームや後期エンドソームの存在も知られており，液胞は動物細胞のリソソームに相当する．酵母オルガネラマーカーの抗体はMolecular Probesから市販されている．

2）酵母細胞におけるウイルス蛋白質の動態

ウイルスの吸着・侵入過程の解析に酵母を積極的に用いた研究は少ない．酵母には細胞壁があるからである．しかしながら，酵母は細胞壁を除去しても活発にエンドサイトーシスがおこる細胞であり，それによると思われる酵母スフェロプラスト細胞へのウイルス感染成立が，パピローマウイルス[49, 50]とVSV[26, 28]で報告されている．特にVSVでは，その感染成立が温度依存性・低pH依存性の膜融合によることが明らかにされている[28]．

ウイルス蛋白質の細胞内輸送や局在は，その蛋白質がエンベロープ蛋白か，キャプシド蛋白か，複製・転写酵素かによって様々だが，それぞれ酵母でも解析されている．例えば，一般的に(+)鎖RNAウイルスのRNA依存性RNA合

成酵素は細胞内の膜小器官に結合して複製・転写複合体を形成するが，どの膜小器官に targeting するかは個々のウイルスによって異なっている．この膜小器官に対する特異性は酵母細胞でも再現され，かつこの複製・転写複合体を人為的に本来とは異なる膜小器官に retargeting させても活性ある複製・転写複合体として維持されることから，この膜小器官特異性はゲノムの複製・転写に必要な現象ではなく，その他の段階(ゲノムのパッケージングやアッセンブリー)に必要な現象だと考えられるようになってきた[25,30]．

脂質二重膜をもたないウイルスでは，ウイルス粒子は細胞質内で形成され細胞破壊によって放出される．これに対し，脂質二重膜をもつウイルスでは，キャプシド蛋白が小胞体やゴルジ体あるいは形質膜に targeting し粒子を形成しながら出芽する．筆者らはこのようなウイルスキャプシド蛋白の膜 targeting を酵母で解析し，HIV-1 のキャプシド蛋白が形質膜に targeting し，細胞壁を除去すると VLP が出芽することを報告している[41]．

## 4. 酵母遺伝学の利用

酵母をウイルスの宿主細胞として使用する，その真価は酵母の分子遺伝学（特に，逆遺伝学）の利用にあるといっても過言でないと思う．ウイルスは細胞内寄生生物であり，その増殖過程において様々な宿主因子を利用する．こうした宿主因子を同定する目的で，ウイルスの研究に染色体遺伝子を破壊させた酵母株を宿主として用いる実験である．

### 1) 遺伝子破壊株の利用

酵母は相同組換えの頻度が非常に高い生物であり，酵母の染色体遺伝子をクローニングしその遺伝子を改変（欠損や点変異を導入）して再び酵母の染色体に戻すと，遺伝子改変の酵母が作出できる．自分が調べたい宿主因子がある程度予想できるならば，こうした遺伝子改変酵母株を分与してもらうか（大抵もらえる）あるいは作成して実験に使ってみると良い．筆者らも細胞内輸送に関連する遺伝子破壊株を分与してもらい実験に使っている．

### 2) 突然変異株の分離

しかし，調べたい宿主因子が予想できない場合にはもっと積極的なアプローチが必要となる．すなわち，エチルメタン硫酸やニトロソグアニジンあるいは紫外線照射によって突然変異を誘発し，ウイルス増殖を許容しなくなった酵母細胞を分離する試みである．知恵を絞らねばならないのはその突然変異株の選択方法である．最も理想的なのは，ウイルス増殖を許容しなくなった突然変異株だけが生育してくるといった条件設定であるが，そんなに都合良くいかない．ウイルス遺伝子あるいはウイルス蛋白を薬剤耐性や栄養要求性のマーカーで標識し，それを指標に突然変異株を分離するのが現実的である．ただし，こうした変異誘発で分離された突然変異株は，その表現型が単一遺伝子の変異によるものか，戻し交配と四分子解析が必要となることが多い．前者は突然変異株と親株の backcross でまだ簡単だが，後者は慣れないと難しい．次に，その遺伝子を同定する．酵母では突然変異株が得られた場合，その遺伝子を特定するのは比較的容易である．酵母の遺伝子ライブラリを導入してその突然変異を相補するのである．すなわち，その突然変異株に親株のゲノム DNA ライブラリを導入し突然変異の表現型が打ち消されたものを分離するか（その突然変異が劣性の場合）（図1），逆に，親株に突然変異株

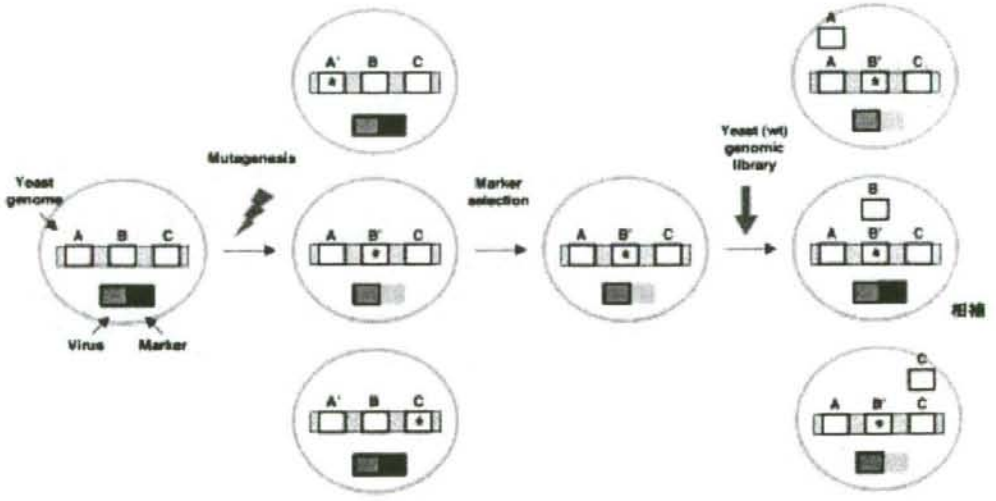


図1 酵母相補性試験による遺伝子の同定

のゲノムDNAライブラリを導入し突然変異の表現型が出現したものを分離する（その突然変異が優性の場合）．こうした解析方法により，BMVではウイルスmRNAの安定性に関与する宿主因子Sm蛋白[10,18]や転写・複製に関する分子シャペロンDNA J[43]が同定されている[1]．

3）全遺伝子破壊株ライブラリを用いた網羅的解析

もっと全遺伝子を網羅的に調べたいという要求が出てくるかもしれない．これを可能にするのがここに述べる全遺伝子破壊株ライブラリの利用である．

出芽酵母のゲノム解析は1988年より国際共同プロジェクトとして開始され1996年に全塩基配列データが公開された．出芽酵母のゲノム長はリボソームDNAの反復配列の長さによって幅があるものの，12-16 Mbpで真核生物の中で最も短い．出芽酵母のゲノムは16本の染色体から構成されており，全ORF数は6000以上でそれぞれに規則的なID番号がついている．出芽酵母のゲノムのデータベースとしてhttp://mips.gsf.de/genre/proj/yeast/（欧州マックスプランク研究所）とhttp://www.yeastgenome.org/（米国スタンフォード大学）がある．一方，分裂酵母のゲノムは14 Mbpであり，3本の染色体から成り立っている．代表的なデータベースとしてhttp://www.sanger.ac.uk/Projects/S_pombe/（英国サンガーセンター）がある．

これらのポストゲノム研究はすべての遺伝子の体系的機能解析研究であり，出芽酵母ではEuropean Functional Analysis Network（EUROFAN）プロジェクトにより全遺伝子破壊株コレクションが作成された．この遺伝子破壊株作製には，PCR法による遺伝子破壊アレルの作製をベースとしながら，全遺伝子共通配列（約20 bp）と個別識別のための分子バーコード（約20 bp）の挿入などの技術を駆使して作製された[11,48]（図2）．この破壊株ライブラリはゲノム解析によって明らかにされた約6000個の遺伝子ORFをそれぞれ薬剤耐性遺伝子で破壊した1倍体あるいは2倍体のセット（96穴フォーマット）で，Invitrogenやフナコシを通じて入手できる．この破壊株ライブラリをpooledで用いた場合には，遺伝子同定には蛍光標識分子バーコードをプローブとしたDNAマイクロアレイが使用できる[48]．こうした遺伝子破壊株ライブラリを用いて，細胞形態[34]やヒト疾患[47]に関与する遺伝子のhigh-throughputな解析が始まっている．2003年にはこの酵母遺伝子破壊株ライブラ

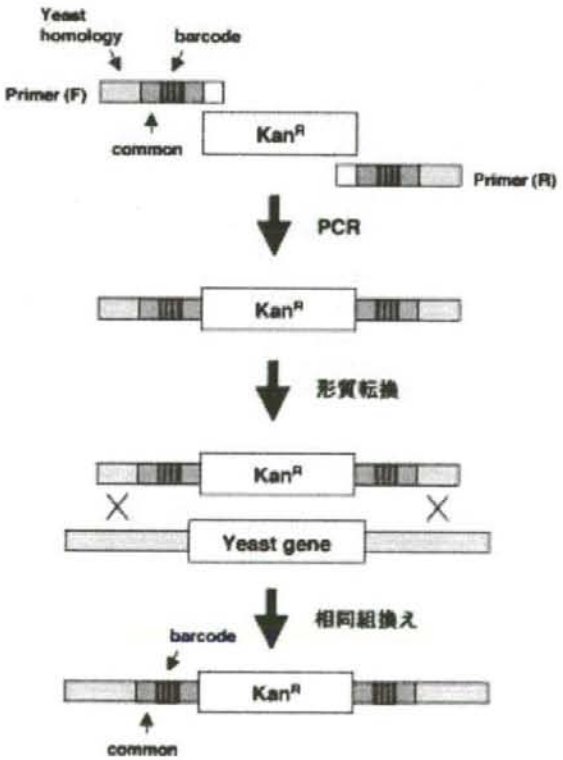


図2　分子バーコードをもつ酵母遺伝子破壊株の作出法

リ（約4500株）を用いてBMV RNAの複製・転写に関与する宿主因子の網羅的解析が行われ，約100個の宿主因子が発表された[24]．

4）全遺伝子破壊株ライブラリVS siRNAライブラリ

上述のように，酵母は真核生物における遺伝学研究を担ってきた．それは酵母では相同組換え頻度が高くgene targetingが容易だったからである．哺乳類細胞における相同組換え頻度は残念ながら今なお低く，遺伝子破壊株ライブラリを作成できるまでに至っていない．しかし近年，RNA干渉法によるmRNAのknock-down法が確立され，この技術は哺乳類細胞における標的宿主因子のdepletionを可能にした．また，網羅的解析を目的としたsiRNAライブラリの構築も開始された．それを用いて，VSVによるクラスリン依存性エンドサイトーシスとSV40によるラフト/カベオラ依存性エンドサイトーシスに関与してくるキナーゼ群が解析されたところ，調べた590個のキナーゼのうち208個がこれらのウイルス感染によるエンドサイトーシスに関与していることが明らかとなった[35]．酵母の遺伝子破壊株かsiRNAか，どちらが良いか筆者にはまだ結論が出せない．

## 5. おわりに

筆者はフツウのウイルス研究者であり，酵母学者ではない．ただ，当時ラボに来てくれた櫻木小百合博士（現，阪大微研）と「酵母からHIV-1の粒子を出芽させる」といった実験を組んだことから酵母も扱うようになった．そうした経験から，ウイルス研究者が酵母に対して疑問に思うことや不安に思うことはできる限り書いたつもりであるが，書き漏らしはあると思うし，酵母学としては体系的に書けていないと思う．また，酵母Two-Hybrid Assayとその細胞質内Assayバージョンである酵母CytoTrap Two-Hybrid Assayの利用については紙面（時間か？）の関係から書けなかった．これらの点はどうぞお許しいただければと願う．足りない点などご指摘いただければ幸いである．

## 文 献


1）Ahlquist P, Noueiry AO, Lee WM, Kushner DB, Dye BT.：Host factors in positive-strand RNA virus genome replication. J Virol 77: 8181-8186, 2003.
2）Angeletti PC, Kim K, Fernandes FJ, Lambert PF.：Stable replication of papillomavirus genomes in Saccharomyces cerevisiae. J Virol 76: 3350-3358, 2002.
3）Baim SB, Sherman F.：mRNA structures influencing translation in the yeast Saccharomyces cerevisiae. Mol Cell Biol 8: 1591-1601, 1988.
4）Chiba Y, Sakuraba H, Kotani M, Kase R, Kobayashi K, Takeuchi M, Ogasawara S, Maruyama Y, Nakajima T, Takaoka Y, Jigami Y.：Production in yeast of alpha-galactosidase A, a lysosomal enzyme applicable to enzyme replacement therapy for Fabry disease. Glycobiol 12: 821-828, 2002.
5）Coutts RH, Cocking EC, Kassanis B.：Infection of protoplasts from yeast with tobacco mosaic virus. Nature 240: 466-467, 1972.
6）Coward P, Dasgupta A.：Yeast cells are incapable of translating RNAs containing the poliovirus 5' untranslated region: evidence for a translational inhibitor. J Virol 66: 286-295, 1992.
7）Das S, Coward P, Dasgupta, A.：A small yeast RNA selectively inhibits internal initiation of translation programmed by poliovirus RNA: specific interaction with cellular proteins that bind to the viral 5'-untranslated region. J Virol 68: 7200-7211, 1994.
8）Das S, Kenan DJ, Bocskai D, Keene JD, Dasgupta A.：Sequences within a small yeast RNA required for inhibition of internal initiation of translation: interaction with La and other cellular proteins influences its inhibitory activity. J Virol 70: 1624-1632, 1996.
9）Das S, Ott M, Yamane A, Tsai W, Gromeier M, Lahser F, Gupta S, Dasgupta A.：A small yeast RNA blocks hepatitis C virus internal ribosome entry site (HCV IRES)-mediated translation and inhibits replication of a chimeric poliovirus under translational control of the HCV IRES element. J Virol 72: 5638-5647, 1998.
10）Diez J, Ishikawa M, Kaido M, Ahlquist P.：Identification and characterization of a host protein required for efficient template selection in viral RNA replication. Proc Natl Acad Sci USA 97: 3913-8, 2000.
11）Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre, B et al.：Functional profiling of the Saccharomyces cerevisiae genome. Nature 418: 387-391, 2002.
12）Gopinath K, Dragnea B, Kao C.：Interaction between Brome mosaic virus proteins and RNAs: effects on RNA replication, protein expression, and RNA stability. J Virol 79: 14222-14234, 2005.
13）Hamilton R, Watanabe CK, de Boer HA.：Compilation and comparison of the sequence context around the AUG startcodons in Saccharomyces cerevisiae mRNAs. Nucleic Acids Res 15: 3581-3593, 1987.
14）Hitzeman RA, Chen CY, Dowbenko DJ, Renz ME, Liu C, Pai R, Simpson NJ, Kohr WJ, Singh A, Chisholm V.：Use of heterologous and homologous signal sequences for secretion of heterologous proteins from yeast. Methods Enzymol 185: 421-440, 1990.
15）Hofmann KJ, Cook JC, Joyce JG, Brown DR, Schultz LD, George HA, Rosolowsky M, Fife KH, Jansen KU.：Sequence determination of human papillomavirus type 6a and assembly of virus-like particles in Saccharomyces cerevisiae. Virology 209: 506-518, 1995.
16）Hofmann KJ, Neeper MP, Markus HZ, Brown DR, Muller M, Jansen KU.：Sequence conservation within the major capsid protein of human papillomavirus (HPV) type 18 and formation of HPV-18 virus-like particles in Saccharomyces cerevisiae. J Gen Virol 77: 465-468, 1996.
17）Iizuka N, Najita L, Franzusoff A, Sarnow P.：Cap-dependent and cap-independent translation by internal initiation of mRNAs in cell extracts prepared from Saccharomyces cerevisiae. Mol Cell Biol 14: 7322-

# Proteomics Reveals *N*-Linked Glycoprotein Diversity in *Caenorhabditis elegans* and Suggests an Atypical Translocation Mechanism for Integral Membrane Proteins*

**Hiroyuki Kaji‡§, Jun-ichi Kamiie‡, Hirotaka Kawakami‡, Kazuki Kido‡, Yoshio Yamauchi‡, Takashi Shinkawa‡, Masato Taoka‡, Nobuhiro Takahashi¶‖, and Toshiaki Isobe‡‖**

**Protein glycosylation is one of the most common post-translational modifications in eukaryotes and affects various aspects of protein structure and function. To facilitate studies of protein glycosylation, we paired glycosylation site-specific stable isotope tagging of lectin affinity-captured *N*-linked glycopeptides with mass spectrometry and determined 1,465 *N*-glycosylated sites on 829 proteins expressed in *Caenorhabditis elegans*. The analysis shows the diversity of protein glycosylation in eukaryotes in terms of glycosylation sites and oligosaccharide structures attached to polypeptide chains and suggests the substrate specificity of oligosaccharyltransferase, a single multienzyme complex in *C. elegans* that incorporates an oligosaccharide moiety en bloc to newly synthesized polypeptides. In addition, topological analysis of 257 *N*-glycosylated proteins containing a putative single transmembrane segment that were identified based on the relative positions of glycosylation sites and transmembrane segments suggests that an atypical non-cotranslational mechanism translocates large N-terminal segments from the cytosol to the endoplasmic reticulum lumen in the absence of signal sequence function.** ***Molecular & Cellular Proteomics 6:2100–2109, 2007.***

Protein post-translational modifications (PTMs)[1] such as proteolysis or addition of a chemical group to one or more amino acid residues may change the properties of a protein. A large body of evidence suggests that PTMs are critical for various cell regulatory and signaling processes (1), and thus the analysis of the status of PTMs on proteins is a major objective of proteomics research. The development of certain new technologies for mapping PTMs on a proteomic scale has begun to yield fruitful results (2–8). Because all PTMs accompany changes in molecular mass of a protein, MS-based analysis is often selected for large scale PTM analyses.

Among the approximately 200 different known PTMs (9), protein glycosylation is one of the most common in eukaryotes: on average there are potential targets in more than half of the genes encoded in eukaryotic genomes (10). Protein glycosylation plays a role in protein folding, subcellular localization, turnover, activity, protein-protein interactions, etc. and contributes significantly to physiology as evidenced by the growing number of human diseases with defects in glycoconjugate assembly and processing (11, 12). Thus, the analysis of protein glycosylation is important for both basic biology and clinical applications, including the discovery of protein biomarkers for diagnosis and drug discovery. Previous studies show that protein glycosylation is quite diverse because the oligosaccharide structure may vary widely between different proteins. In addition, a single protein can be glycosylated at multiple sites, and subsequent processing may differentially or partially modify an oligosaccharide attached at each site. These factors generate the observed complexity of glycoprotein structure and cause difficulties in characterizing protein glycosylation on a proteomic scale. At present, little is known about the final structure of most glycoproteins; however, the specific structure of each oligosaccharide and the rate of the modification(s) are often critical to individual glycoprotein function, and defects in these processes may cause disease (13). Thus, the mechanisms by which protein glycosylation is regulated remain a challenging problem for proteomics research.

Currently two methods allow large scale glycoprotein analysis directly from a complex biological mixture, and both methods utilize MS-based shotgun technology but differ in

From the ‡Department of Chemistry, Graduate School of Science and Engineering, Tokyo Metropolitan University, Minamiosawa 1-1, Hachioji, Tokyo 192-0397, Japan, ¶Department of Applied Biological Science, Tokyo University of Agriculture and Technology, Saiwai-cho 3-5-8, Fuchu, Tokyo 183-8509, Japan, and ‖Core Research for Evolutional Science & Technology (CREST), Japan Science and Technology Agency, Honmachi 4-1-8, Kawaguchi, Saitama 332-0012, Japan

Received, October 12, 2006, and in revised form, April 20, 2007
Published, MCP Papers in Press, August 30, 2007, DOI 10.1074/mcp.M600392-MCP200

[1] The abbreviations used are: PTM, post-translational modification; PNGase, peptide-*N*-glycanase; IGOT, isotope-coded glycosylation site-specific tagging; Con A, concanavalin A; WGA, wheat germ agglutinin; GaL6, galectin 6; ER, endoplasmic reticulum; 2D, two-dimensional; OST, oligosaccharyltransferase; SRP, signal recognition particle; LDL, low density lipoprotein.

the way glycopeptides are collected. One of the methods captures glycopeptides, regardless of the glycan structure, on a solid support by chemical coupling between the *cis*-diol group of the glycan and hydrazide on the support, and then *N*-linked glycopeptides are released specifically from the support by peptide-*N*-glycanase (PNGase) digestion (14, 15). Another method captures a subset of glycopeptides by lectin affinity chromatography (16–18). The type of glycopeptides captured by this method depends on the specificity of the lectin used; however, comprehensive analysis of glycoproteins can be achieved by using multiple lectin columns with distinct binding specificity (*e.g.* non-reducing end oligosaccharides). This approach, termed isotope-coded glycosylation site-specific tagging (IGOT), includes a step to remove the glycan moiety of glycopeptides with PNGase in $^{18}O$-labeled water (16). When the enzyme releases *N*-linked glycans in $H_2{}^{18}O$, the glycosylated Asn residue (in the consensus tripeptide sequence for *N*-linked glycosylation, Asn-Xaa-(Ser/Thr) where Xaa is any amino acid except Pro) is converted to Asp with concomitant incorporation of $^{18}O$ from water (19). This PNGase-mediated incorporation of the $^{18}O$-tag distinguishes glycosylated peptides from non-glycosylated peptides that have non-enzymatically deamidated Asp residues. The conversion of Asn to Asp via $^{18}O$ incorporation in the glycosylation consensus sequence strongly indicates that the peptide was formerly *N*-glycosylated.

In this study, we paired IGOT with automated multidimensional liquid chromatography-MS technology and identified 1,465 *N*-glycosylated sites on 829 proteins expressed in *Caenorhabditis elegans*. We report here the diversity of protein glycosylation and the specificity of the oligosaccharyltransferase of *C. elegans* that incorporates an oligosaccharide moiety en bloc into nascent polypeptide chains. Based on the analysis of the relative positions of *N*-glycosylation sites and putative transmembrane segments of 257 potential integral membrane glycoproteins identified in this study, we also suggest that an atypical, non-cotranslational mechanism determines the topology of integral membrane glycoproteins.

## EXPERIMENTAL PROCEDURES

*Preparation of a Galectin 6 Column*—The coding sequence of the *C. elegans* galectin 6 (GaL6) cDNA (provided by Dr. Hirabayashi, National Institute of Advanced Industrial Science and Technology (AIST), Ibaraki, Japan) was inserted into the *Escherichia coli* expression vector pET and introduced into *E. coli* BL21(DE3)pLysS (20). The transformant was cultured in M9CA medium containing 0.2 mg/ml ampicillin at 37 °C, and gene expression was induced with 1 mM isopropyl 1-thio-β-D-galactopyranoside at a midlog phase of growth ($A_{600}$ = 0.6~0.8). After further cultivation for 3 h, *E. coli* cells were lysed by sonication at 4 °C in 50 mM sodium phosphate buffer, pH 7.5, and centrifuged at 10,000 × *g* for 30 min. The supernatant was then applied to an asialofetuin column (Toyopearl 650M, 2.5-cm inner diameter × 5 cm) equilibrated with 50 mM sodium phosphate buffer, pH 7.5, at a flow rate of 0.5 ml/min. After washing the column with the equilibration buffer, the adsorbed GaL6 was eluted with the same buffer containing 0.2 M lactose. The purified GaL6 (20 mg) was immobilized on TSK-GEL Tresyl-5PW (2 ml; TOSOH) according to the protocol provided by the supplier and was packed into a 4.6-mm-inner diameter × 10-cm column.

*Preparation of Tryptic Digests of Soluble and Insoluble Protein Fractions of C. elegans*—*C. elegans* strain N2 was cultured in liquid medium at 20 °C as described previously (16, 21). A mixed growth phase culture of the worm (5–20 g, wet weight) was lysed by sonication in 5 volumes of TBS (50 mM Tris-HCl, pH 7.5, 150 mM NaCl) containing a protease inhibitor mixture (Sigma), and the homogenate was centrifuged at 1,000 × *g* for 10 min at 4 °C to remove cell debris. The soluble extract was then centrifuged at 100,000 × *g* for 30 min at 4 °C to separate the soluble and insoluble protein fractions. Each fraction was solubilized in 7 M guanidine HCl in 0.5 M Tris-HCl, pH 8.6, containing 50 mM EDTA, and the proteins were reduced with dithiothreitol and *S*-carbamoylmethylated with iodoacetamide (22). The *S*-carbamoylmethylated proteins were dialyzed against 10 mM HEPES-NaOH, pH 7.5, and digested with $N^{\alpha}$-tosylphenylalanyl chloromethyl ketone-treated trypsin (Pierce) at an enzyme:substrate ratio of 1:50 at 37 °C. After 18 h, an aliquot of protease inhibitor mixture (Sigma) was added to the mixture to stop digestion and to protect the lectin columns.

*Preparation of Lectin-bound Glycopeptides*—In our earlier attempts, we prepared an *N*-glycosylated protein fraction by lectin affinity chromatography of *C. elegans* crude extract and then obtained *N*-glycosylated peptides from a tryptic digest of the glycosylated protein fraction by a second round of lectin affinity chromatography (16). In this study, however, we modified the procedure to more efficiently identify the integral membrane glycoproteins; the crude protein extract was first digested with trypsin after *S*-carbamoylmethylation in 7 M guanidine HCl, and then the *N*-glycosylated peptides were recovered by lectin affinity chromatography. To increase the purity of glycopeptides, we incorporated an additional "hydrophilic interaction" chromatography step (23) before PNGase-mediated $^{18}O$ labeling (described later).

To collect *N*-glycosylated peptides, the tryptic digests of soluble and insoluble protein fractions of *C. elegans* were subjected to affinity chromatography on three lectin columns, concanavalin A (Con A) (LA-Con A; 4.6-mm inner diameter × 15 cm; Seikagaku Corp., Tokyo, Japan), wheat germ agglutinin (WGA) (LA-WGA; 4.6-mm inner diameter × 15 cm; Seikagaku Corp.), or GaL6 (4.6-mm inner diameter × 10 cm). Approximately 50–200 mg of peptide mixture was applied to each column equilibrated with 10 mM HEPES-NaOH, pH 7.5. After washing the column with the equilibration buffer, adsorbed glycopeptides were recovered by elution with the buffer containing a cognate sugar: 0.2 M α-methyl mannopyranoside for the Con A column, 0.2 M *N*-acetyl-D-glucosamine (GlcNAc) for the WGA column, or 0.2 M lactose for the GaL6 column. To maximize the recovery of glycosylated peptides, the flow-through fraction of the first chromatography was applied again to the same lectin column, and the chromatography was repeated as described above. The glycopeptide fractions from individual lectin columns of the first and second rounds of chromatography were combined for subsequent steps.

*Purification of Glycopeptides by Hydrophilic Interaction Chromatography*—The *N*-glycosylated peptide mixture recovered by lectin affinity chromatography (10–20 ml containing 200–500 μg of peptides) was added to an equal volume of ethanol (EtOH) and 4 volumes of 1-butanol (BuOH) and was applied immediately to a Sepharose CL-4B column (5-mm inner diameter × 50 mm) equilibrated with the solvent $H_2O$:EtOH:BuOH = 1:1:4 (v/v/v). After washing the column with the same solvent, adsorbed glycopeptides were eluted with $H_2O$:EtOH, 1:1 (v/v). The column eluent was monitored at 220 nm, and the recovered glycopeptides were quantitated fluorometrically after reaction with *o*-phthalaldehyde (24).

*PNGase-mediated $^{18}O$ Labeling of Glycopeptides*—*N*-Glycosylated peptides were labeled specifically with $^{18}O$ by IGOT as de-

scribed previously (16). Briefly the sample glycopeptides were dried under vacuum to remove solvent containing $H_2{}^{16}O$ and then redissolved in 0.1 M Tris base prepared in $H_2{}^{18}O$ (≥99 atom % $^{18}O$; Taiyo Nippon Sanso Corp., Tokyo, Japan). The peptide solution was then adjusted to pH 8–9 with a minimal volume of acetic acid, and then PNGase-A (lyophilized; Seikagaku Corp.), dissolved in $H_2{}^{18}O$, was added to a final concentration of 1 milliunit/10 μg of peptide. The reaction was incubated overnight at 37 °C in a sealed polypropylene tube.

*Automated 2D Nano-LC-MS/MS Analysis of $^{18}O$-Tagged Peptides*—The deglycosylated $^{18}O$-tagged peptide mixture (approximately 5–10 μg) was analyzed by automated 2D LC-MS/MS. The instrument used was a miniaturized version of that described previously (25, 26) and was equipped with a first dimensional microscale cation-exchange column (1-mm inner diameter × 50 mm) of Bioassist-S (7-μm particles; TOSOH) and a second-dimensional direct nanoflow spray tip reversed phase column (150-μm inner diameter × 50 mm) of Mightysil-$C_{18}$ (3-μm particles; Kanto Chemicals) connected in tandem through an electric column switching valve and an automated solvent desalting device. The chromatography was performed automatically under the time-dependent control program, and the eluate was directly sprayed into a high resolution Q-TOF hybrid mass spectrometer (Q-TOF Ultima; Waters-Micromass) at a flow rate of 100 nl/min. The spectrometer was operated in a data-dependent MS/MS mode where a full MS scan (1 s, *m/z* 400–1500) was followed by two MS/MS scans (1 s each, *m/z* 100–1500). The two most intensive precursor ions with a charge state (*z*) of +2 or +3 were dynamically selected and subjected to collision-induced dissociation with a collision energy as recommended by the manufacturer and a dynamic exclusion duration of 30 s. The total analysis time for a single 2D nano-LC-MS/MS process was 24 h.

*Protein Identification by Database Search*—The large volume of MS/MS data generated by the 2D nano-LC-MS/MS analysis was converted to text files using MassLynx software (version 4.0, Micromass). The peak list files were then created with smoothing by the Savitzky-Golay method (window channels, ±3) using the same software and processed by the Mascot algorithm (version 1.9, Matrix Science, Ltd.) to assign peptides on the *C. elegans* Wormpep 124 protein sequence database (22,259 entries, www.sanger.ac.uk/Projects/C_elegans/WORMBASE/current/wormpep.shtml). The database search was performed with the parameters as described previously (16, 21) except that we defined a custom modification, "deamidation with $^{18}O$ (asparagine + 3 Da)," for the deamidation of Asn incorporating $^{18}O$. We first screened the candidate peptides with probability-based Mowse scores that exceeded their thresholds ($p < 0.05$) and with MS/MS signals for y- or b-ions >3; finally we selected "identified peptides" that contained one or more aspartic acid tagged with $^{18}O$ atoms on the basis of their MS/MS spectra. If a prospective "identified peptide" did not contain the consensus tripeptide sequence for *N*-linked glycosylation (Asn-Xaa-(Ser/Thr)), the data were eliminated regardless of the match score. The resulting dataset was finally evaluated by in-house software STEM (27) to remove unreliable Mascot peptide identifications and redundant assignments and to integrate the results with key parameters of the experiment.

*Characterization of the Identified Glycoproteins*—The transmembrane segment and the signal peptide of proteins were predicted by SignalP 3.0 (28) and/or ConPredII (29) bioinformatics tools.

## RESULTS AND DISCUSSION

### *Identification of N-Glycosylated Proteins Expressed in C. elegans*

Because the glycosylation reaction takes place within the lumen of the endoplasmic reticulum (ER), *N*-linked glycoproteins should also have a signal sequence and/or a transmembrane segment as discussed later. We identified these two structural elements in ~25% of the 22,500 genes predicted from the genome sequence of *C. elegans* (Wormpep), suggesting that there are ~6,000 potential targets for *N*-linked glycosylation. To catalog *N*-glycosylated proteins expressed in *C. elegans* and to study details of protein glycosylation, we used IGOT coupled with MS-based proteomics. To increase the coverage, we used three types of lectin columns with different binding specificity for the oligosaccharide attached to the polypeptide chain; thus, the columns contained immobilized Con A, WGA, and GaL6 (20), which are specific for the non-reducing end of Man, GlcNAc, and Gal, respectively. In addition, the lectin affinity chromatography was performed with tryptic peptide mixtures derived from soluble and insoluble protein fractions of *C. elegans* crude extract (see "Experimental Procedures"). The glycopeptide mixtures were further purified by hydrophilic interaction chromatography on Sepharose CL-4B, subjected to IGOT (*i.e.* *N*-glycanase-mediated $^{18}O$ labeling), and analyzed by automated 2D nano-LC-MS/MS shotgun technology to identify $^{18}O$-labeled formerly *N*-glycosylated peptides. To maximize the number of identifications, the shotgun analysis was repeated three times for each peptide mixture prepared by Con A, WGA, and GaL6 affinity chromatography of the soluble/insoluble fractions. Supplemental Table 1 lists all the candidate glycosylated peptides in *C. elegans* identified in this study and all their MS/MS spectra are shown in Supplemental Fig. 1-1 to 1-9.

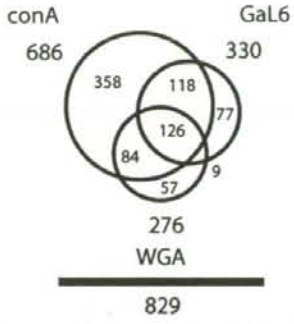


FIG. 1. **Venn diagram of the number of *N*-linked glycoproteins captured on Con A, WGA, and GaL6 columns.** The glycoproteins were identified from the tryptic digests of soluble and insoluble protein fractions of *C. elegans* by IGOT-LC-MS/MS. *Numbers* shown are the numbers of *N*-linked glycoproteins identified by triplicate 2D LC-MS/MS analyses. See text for details.

Supplemental Table 2 lists the *C. elegans* *N*-glycosylated proteins and the number of glycosylation sites identified in this study. We identified 1,204 *N*-glycosylated sites on 686 proteins from Con A-captured glycopeptide mixtures and likewise 474 sites on 276 proteins from WGA- and 382 sites on 330 proteins from GaL6-captured glycopeptide mixtures. After eliminating redundant identifications, we had identified 1,465 *N*-glycosylated sites on 829 unique proteins (Fig. 1 and

Supplemental Table 2). The number of glycosylated sites assigned on each protein ranged from 1 to 24 with an average of 1.5. The glycoproteins we identified were quite diverse in terms of subcellular localization and function, etc., yet many (approximately 50%) were integral membrane proteins such as cell surface receptors, transporters, channels, extracellular matrix proteins, and proteases.

### *Structural Heterogeneity of Oligosaccharides of the Glycoproteins*

Previous studies have shown that most glycans liberated from *C. elegans* membrane proteins contain neutral sugars and have an oligomannose-type structure and that approximately 80% of *N*-linked glycans in *C. elegans* have a nonreducing end mannose that is recognized by Con A (30, 31). These glycans lack sialic acid as the *C. elegans* genome has no sialyltransferase gene, implying that the glycan structure is relatively simple as compared with that of mammalian cells (32). Thus, our lectin affinity analysis of glycopeptides showed that the largest subset of *N*-glycoproteins was identified from the Con A-captured peptide mixtures (Fig. 1); however, the glycopeptides collected by each lectin column overlapped significantly (Supplemental Table 3). For example, 24 glycopeptides assigned for *him-4* (F15G9.4) were identified using the Con A column, whereas some of these glycopeptides were also recovered from the WGA and GaL6 columns, suggesting that the *him-4* product has a highly heterogeneous glycan structure. Of the 1,465 glycosylated sites we determined, 138 sites on 105 proteins were found redundantly in the peptides captured by the three lectin columns, and 317 sites on 228 proteins were found in the peptides captured by two of the lectin columns. Although a subset of those peptides should have hybrid-type glycan structures that would be recognized by multiple lectins, our study implies that most of the worm glycoproteins have complex glycoforms that are typical of eukaryotes. It should be noted that glycan structures may be heterogeneous not only at the protein level but also that each glycosylation site may carry a complex series of *N*-linked glycans if one particular peptide on a single protein is recovered by multiple lectin columns (*e.g.* Supplemental Table 3). Unlike these proteins, however, we also found that many proteins, such as neprilysin (ZK20.6) and integrin $\alpha$ (F54G8.3), have relatively homogeneous glycans because multiple glycopeptides were identified only in the Con A-bound fraction, suggesting that the proteins contain a high mannose-type oligosaccharide chain(s). However, our argument, based on the binding specificity of different types of lectins, should certainly be confirmed by direct structural analysis of the oligosaccharides attached to each site of the polypeptide chain.

### *Amino Acid Residues Close to the Glycosylated Site*

Although we identified ~6,000 potential targets for *N*-glycosylation, not all those proteins were found to be *N*-glycosylated, and as a matter of course, not all Asn residues in the consensus sequences were glycosylated. This suggests that local sequence elements may help determine the specificity of oligosaccharyltransferase (OST), an enzyme responsible for attachment of an oligosaccharide to the newly synthesized polypeptide. We reported previously the frequencies of the amino acid residues around the 400 glycosylated sites on Con A-bound soluble proteins in *C. elegans* (16). In the present study, the analysis was performed for 1,465 unique glycosylated sites on 829 proteins captured on the three types of lectin columns (Table I). Within the consensus tripeptide sequence for *N*-linked glycosylation, Thr occurs at position 3 more than twice as frequently than Ser (819 Thr *versus* 359 Ser). Pro does not occur at positions 2 or 4 with only one exception. We also found that Cys occurs at positions −3 to 6 at 1.5–2.5 times greater frequency than that expected from natural abundance of this residue. However, we could not detect other strong amino acid preferences around the glycosylation sites. This suggests that the nematode OST can introduce an oligosaccharide to almost any Asn-Xaa-(Ser/Thr) (Xaa ≠ Pro) sequence if the nascent polypeptide chain has a properly folded tertiary structure or meets some other criteria. The genome sequence implies that *C. elegans* may have a single OST-translocon (protein-conducting channel) complex in the ER lumen, whereas the yeast *Saccharomyces cerevisiae* has multiple OST-translocon complexes that might have different specificities for other sequences near consensus glycosylation sites (33–35). However, the factors that guide OST-mediated glycosylation remain unknown.

### *Topological Analysis of Integral Membrane Glycoproteins*

*Classification of Integral Membrane Glycoproteins*—Although *N*-linked glycoproteins are found throughout the secretory pathway, including the cell surface, the ER, Golgi, and lysosomes, the initial attachment of *N*-glycans to polypeptides occurs only in the ER lumen. Thus, proteins destined for *N*-glycosylation contain structural elements that serve as signals that target proteins to the ER. The best characterized signal is the signal peptide, a short hydrophobic segment at the N terminus of a nascent polypeptide chain that serves as a recognition site for signal recognition particle (SRP). Upon binding the signal peptide, SRP translocates the translating ribosome to the translocon on the ER membrane. The translocon then inserts the C-terminal portion of the polypeptide chain co-translationally into the ER (36). Subsequently the signal peptide is cleaved off from the nascent polypeptide chain by the ER enzyme signal peptidase, thereby exposing the newly emerged N terminus in the lumen of the ER. Thus, the signal peptide plays a major role in the determination of the transmembrane topology of integral membrane proteins. This mechanism generates the Type I transmembrane proteins that have single transmembrane segments (Table II). Another ER targeting signal is the signal anchor, which, like

TABLE I
*Amino acid occurrences around the N-glycosylated sites*

| Amino acid \ position | -3 | -2 | -1 | 1 (N) | 2 (X) | 3 (S/T) | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|---|---|
| G | 67 (1.07) | 65 (1.04) | 76 (1.21) | 0 | 115 (1.84) | 0 | 61 (0.97) | 62 (0.99) | 56 (0.89) |
| A | 86 (1.17) | 75 (1.02) | 84 (1.14) | 0 | 69 (0.93) | 0 | 73 (0.99) | 70 (0.95) | 70 (0.95) |
| V | 85 (1.17) | 70 (0.96) | 77 (1.05) | 0 | 99 (1.36) | 0 | 102 (1.40) | 81 (1.11) | 69 (0.95) |
| L | 106 (1.04) | 118 (1.16) | 110 (1.08) | 0 | 83 (0.81) | 0 | 102 (1.00) | 88 (0.86) | 98 (0.96) |
| I | 59 (0.81) | 84 (1.15) | 62 (0.85) | 0 | 88 (1.21) | 0 | 96 (1.32) | 99 (1.36) | 74 (1.02) |
| M | 20 (0.65) | 32 (1.04) | 19 (0.62) | 0 | 27 (0.88) | 0 | 24 (0.78) | 18 (0.58) | 17 (0.55) |
| F | 44 (0.76) | 80 (1.39) | 66 (1.15) | 0 | 57 (0.99) | 0 | 56 (0.97) | 59 (1.02) | 57 (0.99) |
| Y | 45 (1.20) | 59 (1.58) | 55 (1.47) | 0 | 55 (1.47) | 0 | 37 (0.99) | 39 (1.04) | 51 (1.36) |
| W | 10 (0.77) | 17 (1.30) | 17 (1.30) | 0 | 11 (0.84) | 0 | 21 (1.61) | 22 (1.68) | 17 (1.30) |
| S | 65 (0.69) | 90 (0.95) | 91 (0.96) | 0 | 102 (1.08) | 359 | 94 (0.99) | 95 (1.00) | 81 (0.86) |
| T | 76 (1.10) | 63 (0.92) | 81 (1.18) | 0 | 70 (1.02) | 819 | 79 (1.15) | 89 (1.29) | 74 (1.07) |
| C | 41 (1.70) | 38 (1.58) | 35 (1.45) | 0 | 61 (2.53) | 0 | 36 (1.49) | 37 (1.54) | 42 (1.74) |
| N | 51 (0.88) | 64 (1.10) | 50 (0.86) | 1178 | 59 (1.02) | 0 | 44 (0.76) | 54 (0.93) | 64 (1.10) |
| Q | 50 (1.04) | 28 (0.58) | 54 (1.12) | 0 | 46 (0.95) | 0 | 40 (0.83) | 27 (0.56) | 36 (0.75) |
| D | 60 (0.97) | 43 (0.69) | 58 (0.93) | 0 | 48 (0.77) | 0 | 79 (1.27) | 62 (1.00) | 78 (1.26) |
| E | 51 (0.67) | 55 (0.72) | 54 (0.71) | 0 | 84 (1.10) | 0 | 90 (1.18) | 80 (1.05) | 112 (1.46) |
| H | 24 (0.88) | 28 (1.02) | 30 (1.10) | 0 | 15 (0.55) | 0 | 21 (0.77) | 20 (0.73) | 23 (0.84) |
| K | 108 (1.42) | 55 (0.72) | 71 (0.93) | 0 | 50 (0.66) | 0 | 58 (0.76) | 49 (0.64) | 56 (0.74) |
| R | 74 (1.21) | 60 (0.98) | 46 (0.75) | 0 | 39 (0.64) | 0 | 64 (1.05) | 40 (0.65) | 52 (0.85) |
| P | 55 (0.96) | 53 (0.92) | 42 (0.73) | 0 | 0 (0.00) | 0 | 1 (0.02) | 87 (1.51) | 51 (0.89) |

The numbers in parentheses are the relative rate estimated from the actual rate (=number of occurrences/1,178) and the expected rate from the natural abundance of each amino acid.

the signal peptide, contains a hydrophobic segment recognized by SRP. The internal signal anchor sequence generates Type II and Type III transmembrane proteins, which differ with regard to the membrane topology of the polypeptide chain, by inserting either the N- or C-terminal transmembrane segment into the ER (37). Unlike the signal peptide, however, the signal anchor sequence is not cleaved by signal peptidase and remains in the mature protein as a transmembrane segment.

To characterize the glycoproteins identified in this study, we searched for ER targeting signals in the amino acid sequences of *C. elegans* *N*-linked glycoproteins. We first searched for signal peptide sequences using SignalP 3.0 (neural network and hidden Markov model methods) (28) and ConPred II (module DetecSig) (29). We accepted the results of prediction only when both softwares gave identical results; otherwise we considered the results ambiguous (Table II) and did not further consider them. The number and position of transmembrane segments were predicted with ConPred II (29). The results, including several representative known glycoproteins, are summarized in Table II. Of the 829 *N*-linked glycoproteins identified in this study, 463 had putative signal peptides, whereas 238 did not; the remaining 128 proteins gave ambiguous results. Among the 463 signal peptide-containing proteins, 181 had a single putative transmembrane segment, 58 had multiple transmembrane segments, and 224 had no predicted transmembrane segment. Transmembrane segments were also found in 193 of 238 proteins that had no detectable signal peptide (76 with single transmembrane segments and 117 with multiple transmembrane segments). We also identified 45 "ovalbumin-like" *N*-linked glycoproteins with no detectable signal peptide or transmembrane sequence. Thus, our predictions suggest that the *N*-glycosylated proteins identified in this study included 224 secretory, 181 Type I, and 76 Type II/III proteins (Supplemental Table 4).

*Analysis of Integral Membrane Glycoproteins Containing a Signal Peptide*—Because the initial protein glycosylation takes place only in the ER lumen and the glycosylated segments do not cross the membrane bilayer, the structural segment around the glycosylated site must face toward the ER lumen or an equivalent topological space (*e.g.* the Golgi lumen). Our topological analysis of membrane glycoproteins was based on the positions of experimentally determined glycosylation sites and putative transmembrane segments on the polypeptide chains. To simplify our analysis, we focused on the 257 glycoproteins containing a single transmembrane segment. The translated polypeptide segment that follows the signal peptide is introduced into the ER lumen through the translocon. If the transmembrane segment is translated and recognized by the translocon, it exits laterally from the channel into the membrane lipid, and the C-terminal portion remains in the cytosol (36). Therefore, for single span transmembrane proteins containing a signal peptide, the N-terminal portion of the transmembrane segment resides in the ER lumen. We assigned 181 of these Type I transmem-