JOURNAL OF VIROLOGY, Sept. 2002, p. 9024–9034 Vol. 76, No. 18
0022-538X/02/$04.00+0 DOI: 10.1128/JVI.76.18.9024–9034.2002

Analysis of the Complete Genome Sequence of the Hz-1 Virus Suggests that It Is Related to Members of the *Baculoviridae*

Chia-Hsiung Cheng,[1] Su-Mei Liu,[2] Teh-Yuan Chow,[2] Yu-Yun Hsiao,[3] Dan-Ping Wang,[3] Jiann-Jang Huang,[2] and Hong-Hwa Chen[1,3]*

Institute of Biotechnology[1] and Department of Biology,[3] National Cheng Kung University, Tainan 701, and Institute of Botany, Academic Sinica, Nankang, Taipei 105,[2] Taiwan

Received 10 February 2002/Accepted 6 June 2002

We report the complete sequence of a large rod-shaped DNA virus, called the Hz-1 virus. This virus persistently infects the *Heliothis zea* cell lines. The Hz-1 virus has a double-stranded circular DNA genome of 228,089 bp encoding 154 open reading frames (ORFs) and also expresses a persistence-associated transcript 1, PAT1. The G+C content of the Hz-1 virus genome is 41.8%, with a gene density of one gene per 1.47 kb. Sequence analysis revealed that a 9.6-kb region at 43.6 to 47.8 map units harbors five cellular genes encoding proteins with homology to dUTP pyrophosphatase, matrix metalloproteinase, deoxynucleoside kinase, glycine hydroxymethyltransferase, and ribonucleotide reductase large subunit. Other cellular homologs were also detected dispersed in the viral genome. Several baculovirus homologs were detected in the Hz-1 virus genome. These include PxOrf-70, PxOrf-29, AcOrf-81, AcOrf-96, AcOrf-22, VLF-1, RNA polymerase LEF-8 (orf50), and two structural proteins, p74 and p91. The Hz-1 virus p74 homolog shows high structural conservation with a double transmembrane domain at its C terminus. Phylogenetic analysis of the p74 revealed that the Hz-1 virus is evolutionarily distant from the baculoviruses. Another distinctive feature of the Hz-1 virus genome is a gene that is involved in insect development. However, the remainder of the ORFs (81%) encoded proteins that bear no homology to any known proteins. In conclusion, the sequence differences between the Hz-1 virus and the baculoviruses outnumber the similarities and suggest that the Hz-1 virus may form a new family of viruses distantly related to the *Baculoviridae*.

The Hz-1 virus is a nonoccluded, rod-shaped, enveloped virus with a particle size of 414 ± 30 nm (11). It contains a circular double-stranded DNA genome of 228 kb with a molecular mass of 131 × 10^6 Da to 140 × 10^6 Da (13, 34). Hz-1 virus was classified as a member of subgroup C of the *Baculoviridae* and is a *Nudibaculoviridae* species. However, low sequence homology between Hz-1 virus and other members of the *Baculoviridae* was observed. It has been shown that the DNA homology of the Hz-1 virus shares 3% homology with *Heliothis armigera* granulovirus (HearGV) and 0.1 to 1% homology with several nuclear polyhedrosis viruses (NPVs) and *Plodia interpunctella* granulovirus (PiGV) by Southern blot hybridization (57). Originally classified as a member of the *Baculoviridae*, the Hz-1 virus is currently unclassified due to its lack of an occlusion body, as well as its low DNA homology with other baculoviruses (59).

The Hz-1 virus was originally identified as a persistent viral infection in the IMC-Hz-1 cell line, which was isolated from the adult ovarian tissues of *Heliothis zea* (26, 44, 50). Induction of persistent Hz-1 virus was observed when the IMC–Hz-1 cells were transfected with *H. zea* NPV (HzNPV) DNA. This resulted in the cytopathic effect typical of a NPV but without occlusion body formation (35). Activation of the persistent Hz-1 virus in IMC–Hz-1 cells was further confirmed by infection with the homologous virus (Hz-1 virus) or the heterologous viruses HzNPV, *Spodoptera frugiperda* MNPV (Sf*M*NPV), *Spodoptera litura* NPV (SpliNPV), and HearGV, as well as heat- and UV-inactivated *Trichoplusia ni* SNPV (TnSNPV) (39). The Hz-1 virus can establish persistent infections in other insect cell lines, including IPLB-Ld652 (*Lymantria dispar*), IPLB-Hz-1075 (*Helicoverpa zea*), IPLB-SF-21 (*Spodoptera frugiperda*), and TN368 (*Tricoplusia ni*) cells (32). It is the first insect virus to be reported that establishes both productive and persistent infections in insect cells (26, 39).

Temporal viral gene expression during persistent infection by Hz-1 virus has been reported (11). There are >100 genes expressed during a productive infection, with transcript sizes ranging from 0.8 to 9.5 kb (14). So far, two late genes of Hz-1 virus, p34 and p51, have been identified as expressed at 8 h postinfection during a productive Hz-1 virus infection (28, 29). In contrast, transcription of the Hz-1 viral genome is mostly silent during persistent infection, with only persistence-associated transcript 1 (PAT1) being expressed (14). The results of in situ hybridization suggest that PAT1 is localized in the nucleus. In addition, it is not associated with polysomes, and in vitro-translated products are not detected, suggesting that PAT1 is a noncoding nuclear RNA (15).

Because of the distinct temporal viral gene expression pattern in a persistent Hz-1 virus infection, the study of Hz-1 virus gene expression regulation may facilitate understanding of the switch between both productive and persistent infections, which is common in a number of viruses pathogenic for humans and animals.

We describe here the complete genome sequence of the Hz-1 virus. This analysis revealed that it has a genome with a 41.8% G+C content and contains 228,089 bp encoding 154

* Corresponding author. Mailing address: College of Sciences, Institute of Biotechnology, National Cheng Kung University, 1 University Rd., Tainan 701, Taiwan. Phone: 886-6-275-7575, ext. 65521. Fax: 886-6-276-6505. E-mail: hhchen@mail.ncku.edu.tw.

open reading frames (ORFs), in addition to the PAT1, with a gene density of one gene per 1.47 kb. Twenty-nine predicted Hz-1 ORFs show significant relatedness to the baculovirus and cellular genes. However, the remaining ORFs (81%) bear no detectable homology to any known proteins, suggesting that the Hz-1 virus is evolutionarily distant from the *Baculoviridae*.

MATERIALS AND METHODS

Virus and viral DNA. The Hz-1 virus was serially diluted and then used to infect TN368 insect culture cells derived from the ovarian tissues of *Tricoplusia ni* larvae (32). The infected cells were incubated at 26°C with TNM-FH medium supplemented with 8% heat-inactivated fetal bovine serum (Gibco-BRL, Gaithersburg, Md.) for 48 to 72 h for plaque production. Hz-1 virus from single plaques was selected and scaled up in TN368 cells with a multiplicity of infection (MOI) of 0.01 PFU/cell at 26°C for 24 to 48 h. The viral titer was determined by determining the 50% tissue culture infective dose(s) (57). To extract the viral DNA, the viral suspensions were centrifuged at 800 × *g* (RT600D; Sorvall, Newton, Conn.) for 10 min. The supernatant was added on top of a 30% sucrose and centrifuged at 52,714 × *g* (SW28 rotor; Optima LE-80K Ultracentrifuge; Beckman, Fullerton, Calif.) for 30 min. The pellet was resuspended with 1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) buffer and then added on top of a 40 to 65% sucrose gradient and centrifuged at 52,714 × *g* as described above for 1 h. The Hz-1 virus particles were banded at a 62 to 63% position. After collection into a new tube, the viral particles were suspended in 1× SSC, centrifuged twice at 52,714 × *g* for 30 min, and then resuspended in 1× extraction buffer (0.1 M Tris-Cl [pH 7.6], 2.5% sodium dodecyl sulfate, 0.1 M EDTA). The viral particles were digested with proteinase K (40 μg/ml; Sigma, St. Louis, Mo.) for 1 h at 50°C, and this step was followed by digestion with 50 μg of proteinase K for another 12 h. Viral genomic DNA was extracted with phenol-chloroform and precipitated with alcohol. After centrifugation, viral DNA pellets were dissolved in TE buffer (10 mM Tris [pH 8.0], 0.5 mM EDTA).

Hz-1 virus DNA cloning and sequence determination. The Hz-1 virus was sequenced to sevenfold genomic coverage by a shotgun approach. The viral DNA was sheared by nebulization into fragments with an average size of 2,000 bp (HydroShear; GeneMachines, San Carlos, Calif.). DNA fragments were size fractionated by gel electrophoresis and cloned into the *Sma*I site of pBluescript SK(+/−) (Promega, Madison, Wis.). After transformation into *Escherichia coli* DH10B competent cells (Gibco-BRL), 3,000 recombinant colonies were picked randomly. DNA templates for sequencing were isolated by using Multiscreen kits (Millipore, Bedford, Mass.). Sequencing was performed by using the ABI Prism Big Dye Terminator Cycle Sequencing Ready reaction kit with FS AmpliTaq DNA polymerase (Perkin-Elmer, Palo Alto, Calif.) and analyzed on an ABI 377 DNA Analyzer. Shotgun sequences were base called by using the PHRED basecaller and assembled with the PHRAP assembler (22, 23). PHRAP-assembled data were stored in the Sun workstation assembly database by using the Sun workstation interface (7). The Sun workstation interface and its features were then used for editing and completing the sequence. Consensus calculations with a quality cutoff value of 40 were performed by using the Sun workstation with a probabilistic consensus algorithm based on expected error rates for output by PHRED. The PCR products bridging the ends of existing contiguous fragments were sequenced, and primers were designed to walk with genomic DNA to fill the remaining gaps in the sequence.

DNA sequence analysis. Genomic DNA composition, structure, repeats, and restriction enzyme patterns were analyzed by using the University of Wisconsin Genetics Computer Group programs (20) and Sequencher v.4.1.2 (GeneCodes, Ann Arbor, Mich.). ORFs encoding more than 50 amino acids (150 bp) were considered protein encoding and hence were designated putative genes. The maximal 154 ORFs were predicted and analyzed from the Artemis program (http://www.sanger.ac.uk/Software/Artemis/) (54) and ORF finder and the BLAST programs of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The overlap between any two ORFs was set to a maximum of 25 amino acids. Otherwise, the largest ORF similar to the previously described ORF was selected. DNA and protein comparisons with entries in the sequence databases were performed by using the FASTA and BLAST programs (3, 50). Multiple sequence alignments were performed with the PileUp and Gap computer programs (version 10.0; Genetics Computer Group, Madison, Wis.) with the gap creation and extension penalties set to 9 and 2, respectively (20). The percent identity indicated the percentage of identical residues between two complete sequences. DNA repeats were identified by using the Miropeats computer program of EBI (European Bioinformatics Institute) (www.ebi.ac.uk/~jparsons/packages/miropets). The presence of transmembrane (TM) domain and signal peptide (SP) in the putative ORFs were predicted by using the transmembrane hidden Markov model (TMHMM) (http://www.cbs.dtu.dk/services/TMHMM) and SignalP (http://www.cbs.dtu.dk/services/SignalP) programs of the Center for Biological Sequence Analysis, Biocentrum-DTU, Technical University of Denmark, Lyngby, Denmark. Comparison between Hz-1 virus and NPV-granulovirus genomes were performed by using the pairwise program of prFLAG (http://flag.itri.org.tw/~cflag) from the Biomedical Engineering Center, Industrial Technology Research Center, Hsinchu, Taiwan. Phylogenetic analysis was performed by using the clustering method, UPGMA, where pairwise distance estimations were based on the proportional distance.

Nucleotide sequence accession number. The complete Hz-1 virus sequence can be obtained from GenBank (accession no. AF451898).

RESULTS AND DISCUSSION

General features of the Hz-1 virus genome. The Hz-1 virus genome was assembled into a contiguous sequence of 228,689 bp in good agreement with previous predictions of 228-kb based on the restriction enzyme fragment analysis and physical mapping (14). The G+C content of Hz-1 virus was 41.8%. A total of 154 ORFs, defined as methionine-initiated ORFs encoding more than 50 amino acids and with a minimal overlap with other ORFs, were present in the Hz-1 virus genome (Table 1 and Fig. 1). Together with the noncoding PAT1 (indicated by a bold arrow in Fig. 1), 155 genes were detected in the Hz-1 virus genome. The gene density of Hz-1 virus genome was one gene per 1.47 kb, much larger than that for seven NPVs and two GVs (0.87 to 0.99 kb and 0.84 to 0.99 kb, respectively) (Table 2). The ORFs were distributed evenly along the genome: 45% of them were clockwise, and 55% were anticlockwise (Fig. 2). The first ORF, Hz1V001, was defined as the first ORF present in the A fragment of *Xho*I-digested Hz-1 virus genome (13). The locations, orientations, sizes, and BLAST results of the ORFs are shown in Table 1. Predicted ORFs represent 68.98% coding density, with a mean ORF length of 1,015 nucleotides. The prediction of the viral capsid/coat proteins of the putative ORFs with the characteristics of the TM domain and SP was carried out with the Hidden-Markov model and SignalP software. The predicted results of ORFs with TM domain and/or SP are presented in Table 1.

A total of 24 Hz-1V ORFs possessed a consensus early promoter motif (a TATA box followed by a CAG/TT motif located 20 to 25 bp downstream) within 180 bp (41) of the initiation codon (Table 1). Of these, five ORFs also possessed a late gene promoter motif, which may allow transcription of theses genes during both early and late stages of infection. A total of 45 Hz-1V ORFs possessed a consensus late gene promoter motif within 160 bp of the initiation codon. One Hz-1V ORF was found that contained a CGTGC motif that has also been identified as an early promoter consensus sequence-transcription initiation site (41). A 9-bp sequence of TTATAG TAT was identified at the upstream regulatory regions of both p34 and p51 late genes of Hz-1 virus (28, 29). A TTATAGTAT motif was found within 200 bp of the initiation codon of three Hz-1V ORFs, including p34 and p51. Of the Hz-1V ORFs, 80 did not possess consensus late or early promoter sequences. It is possible that they may be transcribed from unique early promoters or from consensus late gene promoters other than these sequences.

Repeated regions of Hz-1 virus. The baculovirus contains homologous regions (HRs) containing AT-rich sequences of

TABLE 1. Characterization of putative genes of Hz-1 virus[a]

ORF	Direction[b]	Position		Length[c]		Promoter	BLASTP genes[d]	Source	BLASTP score	Identity (%)
		Start	End	nt	aa					
Hz1V001	F	620	3955	3,336	1,112	L				
Hz1V002	F	4333	4503	171	57		TM, SP			
Hz1V003	R	6218	7921	1,704	568	E	Histidine kinase	*Bacillus halodurans*	151	21
Hz1V004	R	8454	11378	2,925	975	E				
Hz1V005	F	11704	11922	219	73	L	TM, SP			
Hz1V006	R	12779	12931	153	51	E				
Hz1V007	R	13415	14209	795	265		DHFR protein kinase cyclic AMP-	*Heliothis virescens*	443	49
Hz1V008	R	14380	15144	765	255		dependent, catalytic chain 1	*Leishmania major*	213	28
Hz1V009	F	15274	16005	732	244					
	R	16008	18980	2,973		E	PAT1			
Hz1V010	F	16094	16873	780	260		PxOrf-70	*Plutella xylostella* GV	126	24
Hz1V011	F	17883	19931	2,049	683	L	p74	*Spodoptera litura* NPV	402	22
Hz1V012	F	20214	21089	876	292	L				
Hz1V013	F	21412	22146	735	245	L				
Hz1V014	F	22321	22491	171	57	L	TM			
Hz1V015	F	23171	23407	237	79					
Hz1V016	R	23510	23749	240	80					
Hz1V017	F	23954	24190	237	79					
Hz1V018	F	25151	25309	159	53	E				
Hz1V019	R	25312	25479	168	56					
Hz1V020	F	25691	25846	156	52	L				
Hz1V021	F	26150	26383	234	78	L	TM, SP			
Hz1V022	R	26865	27041	177	59	L				
Hz1V023	R	27208	27342	135	45					
Hz1V024	R	27641	27877	237	79	L				
Hz1V025	R	28180	28419	240	80	L	TM			
Hz1V026	F	30177	30386	210	70	E				
Hz1V027	F	30486	30686	201	67	L				
Hz1V028	F	34670	35740	1,071	357					
Hz1V029	R	35912	36850	939	313					
Hz1V030	F	36914	37780	867	289		TM, SP			
Hz1V031	R	38553	41990	3,438	1,146	E				
Hz1V032	R	42056	42457	402	134					
Hz1V033	R	42556	43026	471	157		AcOrf-81	*Autographa californica* NPV	130	25
Hz1V034	F	43141	43785	645	215		TM, SP			
Hz1V035	F	43954	44127	174	58	L	TM, SP			
Hz1V036	R	44312	48547	4,236	1,412		DNA ligase III	*Homo sapiens*	736	44
Hz1V037	R	52106	54937	2,832	944	E	Probable methyl transferase	*Autographa californica* NPV	311	35
Hz1V038	F	55352	55543	192	64					
Hz1V039	F	55655	55813	159	53					
Hz1V040	R	57238	57432	195	65					
Hz1V041	R	57601	58140	540	180					
Hz1V042	R	58165	58476	312	104					
Hz1V043	R	58804	59868	1,065	355	E				
Hz1V044	R	60007	60282	276	92		TM			
Hz1V045	R	60346	60546	201	67					
Hz1V046	R	60720	63215	2,496	832	L	vp91, viral capsid-associated	*Autographa californica* NPV	131	25
Hz1V047	F	63587	65263	1,677	559		protein (AcOrf-83)			
Hz1V048	R	65317	65502	186	62	L				
Hz1V049	R	65726	66694	969	323	E2				
Hz1V050	R	67339	67734	396	132	L	TM, SP			
Hz1V051	R	68053	69963	1,911	637	E,L				
Hz1V052	F	70210	74184	3,975	1,325	L				
Hz1V053	F	74656	74820	165	55	L				
Hz1V054	R	75184	75849	666	222					
Hz1V055	R	76083	77789	1,707	569	L	SP			
Hz1V056	R	78048	78431	384	128					
Hz1V057	R	78718	79641	924	308	L				
Hz1V058	F	80420	82591	2,172	724					
Hz1V059	R	83062	84294	1,233	411					
Hz1V060	F	84435	87065	2,631	877					
Hz1V061	F	87339	87593	255	85	L				
Hz1V062	R	88323	92270	3,948	1,316	E,L				
Hz1V063	R	92396	92950	555	185	L	TM			
Hz1V064	R	93091	94242	1,152	384	HL	p51	*Macaca mulatta* rhadinovirus 17577	95	21
Hz1V065	R	94271	94444	174	58	L	TM, SP			
Hz1V066	F	94610	94948	339	113					
Hz1V067	F	95097	96485	1,389	463	L				
Hz1V068	F	96560	98518	1,959	653					
Hz1V069	F	99357	100409	1,053	351		dUTP pyrophosphatase	*Homo sapiens*	386	58

Continued on facing page

TABLE 1—*Continued*

ORF	Direction[b]	Position		Length[c]		Promoter	BLASTP genes[d]	Source	BLASTP score	Identity (%)
		Start	End	nt	aa					
Hz1V070	R	101531	103909	2,379	793		Matrix metalloproteinase	*Mus musculus*	464	29
Hz1V071	R	104121	104945	825	275		Deoxynucleoside kinase	*Drosophila melanogaster*	397	41
Hz1V072	F	106014	107339	1,326	442		Glycine hydroxymethyltransferase	*Drosophila melanogaster*	1434	61
Hz1V073	R	107960	108961	1,002	334		rr2	Equine herpesvirus 1	1150	67
Hz1V074	R	109072	109536	465	155		TM			
Hz1V075	F	109742	113350	3,609	1203	E	DNA-directed RNA polymerase	Yeast (*Kluyveromyces marxianus* var. *lactis*)	116	26
Hz1V076	F	113737	115065	1,329	443		TM	Plasmid pSKL		
Hz1V077	R	115274	116017	744	248					
Hz1V078	R	116225	116599	375	125	L				
Hz1V079	R	116876	117640	765	255	E, L, HL	p34	Hz-1 insect virus	1308	100
Hz1V080	R	117679	117840	162	54	L				
Hz1V081	F	117933	118313	381	127	E, HL	TM, SP			
Hz1V082	F	118605	118793	189	63	L	TM, SP			
Hz1V083	R	122431	123837	1,407	469	L				
Hz1V084	F	124009	124212	204	68					
Hz1V085	R	124410	124694	285	95	L				
Hz1V086	R	124758	124931	174	58		SP			
Hz1V087	F	124951	125394	444	148					
Hz1V088	R	125907	126548	642	214		PxOrf-29	*Plutella xylostella* GV	96	29
Hz1V089	R	127093	127938	846	282	L				
Hz1V090	F	128075	131866	3,792	1,264	L	RNA polymerase LEF-8 (Orf-50)	*Bombyx mori* NPV	101	24
Hz1V091	F	132319	132498	180	60					
Hz1V092	F	132556	132921	366	122					
Hz1V093	R	133290	134036	747	249					
Hz1V094	F	134193	134555	363	121					
Hz1V095	R	134950	137769	2,820	940	L	rr1	Equine herpesvirus 1	2137	53
Hz1V096	R	138005	139477	1,473	491	E				
Hz1V097	F	139781	140824	1,044	348					
Hz1V098	F	141233	143899	2,667	889					
Hz1V099	R	145139	145900	762	254					
Hz1V100	R	145986	146882	897	299					
Hz1V101	F	147162	147887	726	242	L				
Hz1V102	R	148545	148715	171	57	E				
Hz1V103	R	148816	150930	2,115	705		AcOrf-96	*Xestia c-nigrum* GV	126	23
Hz1V104	F	150932	155689	4,758	1,586					
Hz1V105	F	156572	156727	156	52					
Hz1V106	R	156774	157796	1,023	341	L				
Hz1V107	F	158399	162130	3,732	1,244					
Hz1V108	R	162963	163127	165	55					
Hz1V109	R	163181	164056	876	292	E,L	*ts*	*Mus musculus*	1064	70
Hz1V110	F	164264	164440	177	59	L	TM			
Hz1V111	R	164454	165461	1,008	336					
Hz1V112	F	165694	166974	1,281	427	E				
Hz1V113	R	168405	168560	156	52					
Hz1V114	F	168579	168740	162	54		TM, SP			
Hz1V115	R	169103	170377	1,275	425	L				
Hz1V116	R	170418	170615	198	66					
Hz1V117	F	171120	171602	483	161	E, L	SP			
Hz1V118	R	172592	173560	969	323					
Hz1V119	F	173602	176073	2,472	824		TM, SP			
Hz1V120	F	176098	176286	189	63					
Hz1V121	F	176883	177884	1,002	334	E	VLF-1	*Spodoptera exigua* NPV	84	56
Hz1V122	R	178281	180080	1,800	600					
Hz1V123	R	180164	181306	1,143	381		AcOrf-22	*Bombyx mori* NPV	208	26
Hz1V124	R	181949	182278	330	110	L	TM, SP			
Hz1V125	R	182341	182769	429	143	L				
Hz1V126	R	183078	184616	1,539	513		CG4526 gene product (alt1)	*Drosophila melanogaster*	1140	48
Hz1V127	R	184654	185178	525	175					
Hz1V128	F	185280	187370	2,091	697	E				
Hz1V129	F	188310	188741	432	144					
Hz1V130	R	189943	190356	414	138	L				
Hz1V131	F	190540	193950	3,411	1,137	L	DNA polymerase I	*Aeropyrum pernix*	124	23
Hz1V132	F	194777	195013	237	79	L	TM, SP			
Hz1V133	F	195093	195281	189	63	L				
Hz1V134	R	196267	197592	1,326	442	E				
Hz1V135	R	197755	198303	549	183	E	Inhibitor of apoptosis protein	*Orgyia pseudotsugata* NPV	309	35
Hz1V136	R	199229	200053	825	275					
Hz1V137	R	202089	203438	1,350	450	E				
Hz1V138	R	203887	204453	567	189		IAP	*Buzura suppressaria* NPV	290	37

Continued on following page

TABLE 1—*Continued*

ORF	Direction[b]	Position		Length[c]		Promoter	BLASTP genes[d]	Source	BLASTP score	Identity (%)
		Start	End	nt	aa					
Hz1V139	R	204781	204945	165	55	L				
Hz1V140	F	205208	207643	2,436	812	E,L				
Hz1V141	F	207769	210198	2,430	810					
Hz1V142	F	210245	210493	249	83					
Hz1V143	F	211082	211276	195	65	E	SP			
Hz1V144	F	211278	212285	1,008	336					
Hz1V145	F	213258	215354	2,097	699		Carboxylesterase	*Aphis gossypii*	226	32
Hz1V146	R	215729	215887	159	53	L				
Hz1V147	R	216178	216375	198	66		TM, SP			
Hz1V148	R	216788	217228	441	147					
Hz1V149	R	217623	217796	174	58					
Hz1V150	F	217831	218781	951	317					
Hz1V151	R	220142	220846	705	235	L				
Hz1V152	F	221294	221458	165	55	E	MP			
Hz1V153	R	221867	222625	759	253					
Hz1V154	R	222659	227176	4,518	1,506					

[a] Nucleotides in the Hz-1 virus genome were numbered sequentially, beginning with the nucleotide in the *Xho*I-A fragment. ORFs in the Hz-1 virus genome over 150 bp in length were designated ORF1 to ORF154, beginning with the first ORF in the *Xho*I-A fragment. PAT1 is the noncoding RNA that is transcribed from the reverse strand between the Hz1 V010 and Hz1V011. The presence of baculovirus early (E and E2) and late (L) promoter elements, located within 300 nucleotides of the ATG, is indicated. E and E2 indicate TATAA sequence with a CA(T/G)T (E) or CGTGC (E2) start site sequence 20 to 40 nucleotides downstream. L indicates the presence of a (A/T/G)TAAG motif. In addition, the promoter scanning was also carried out for detecting the Hz-1 virus late (HL) promoter sequence TTATAGTAT. The position of each ORF defines the A the initiation codon (ATG) and the T of the termination codon (TAA/TAG/TGA) of its encoding strand.
[b] The directions of the transcripts are indicated by F (forward) or R (reverse).
[c] nt, nucleotides; aa, amino acids.
[d] TM, TM domain gene; SP, SP gene.

direct repeats, as well as inverse repeats that serve as the origins for DNA replication (49) and enhancers of early transcription (27). Homologous regions have been identified in all baculovirus genomes sequenced to date, including Ac*M*NPV (5), Op*M*NPV (3), BmNPV (25), Ld*M*NPV (43a), and CuniNPV (2). It has been considered that the presence of HRs is a characteristic feature of the baculoviruses. However, no HRs were identified in the Hz-1 virus genome. Instead, abundant tandem sequence repeats of 21 to 75 bp were identified throughout the Hz-1 virus genome. There are highly repeated sequences identified at positions 57243 to 58675 (R1), 59090 to 59359 (R2), and 105345 to 105885 (R3) of the Hz-1 virus genome (Fig. 1).

Hz-1 virus ORFs encoding products homologous to known proteins. Among the 154 ORFs identified, only 29 have significant homologies to the current database. These include nine involved in nucleic acid metabolism, two involved in DNA replication, three involved in gene regulation, one virion, one nucleocasid protein, and p34 (product of Hz1V079) and p51 (product of Hz1V064), which are the Hz-1 virus late genes as previously reported (28, 29) (Table 1). However, 125 ORFs (81%) have no similarities to any known genes, suggesting that the Hz-1 virus may form a novel class of virus that is distant from the baculoviruses.

Enzymes involved in nucleic acid metabolism. Nine genes—Hz1V003, Hz1V007, Hz1V037, Hz1V043, Hz1V069, Hz1V071, Hz1V073, Hz1V095, and Hz1V109—show extensive homologies with previously identified proteins and therefore may encode the Hz-1 virus homologues of enzymes involved in nucleic acid metabolism. The most significant homology (70% identity over 292 amino acids) was detected between the product of Hz1V109 and the human thymidylate synthase gene (*ts*) (Table 1). Thymidylate synthase catalyzes the metabolism of dUTP to yield the nucleotide precursor of dTMP and is an important step in the de novo pathway of biosynthesis of pyrimidine (12). So far, no *ts* gene has been reported for baculoviruses. However, homologs of the *ts* gene have been identified in the genomes of two other insect viruses, *Chilo* iridescent virus (47) and *Melanoplus sanguinipes* entomopoxvirus (1). Our previous analysis suggested that the Hz-1 virus *ts* gene originated from its lepidopteran host and that the *ts* genes of Hz-1 virus, *Chilo* iridescent virus, and *Melanoplus sanguinipes* entomopoxvirus originated from independent recombination events (16).

The products of Hz1V095 and Hz1V073 show significant homologies to both the large (*rr1*) and small (*rr2*) subunits of ribonucleotide reductase (53% identity to equine herpesvirus *rr1* over 940 amino acids, and 67% identity to equine herpesvirus *rr2* over 334 amino acids, respectively). The *rr1* and *rr2* genes were separated by 25,989 bp. This enzyme is involved in nucleotide metabolism and reduces ribonucleotides into deoxyribonucleotides as immediate precursors of DNA (37). Hz1V069 is homologous to dUTP pyrophosphatase (dUTPase) with a 58% identity to *Homo sapiens* dUTPase over 351 amino acids. dUTPase cleaves the alpha-beta phosphodiester bonds of dUTP to form pyrophosphate and dUMP, preventing incorporation of uracil into DNA and providing the substrate for thymine synthesis (42). dUTPase has been shown to be essential for the replication of a number of DNA viruses (6). Hz1V071 shows a 44% identity to deoxynucleoside kinase of *Drosophila melanogaster* over 275 amino acids. Deoxynucleoside kinases are key enzymes in deoxyribonucleoside salvage (37). Hz1V007 shows a 49% identity to dihydrofolate reductase (DHFR) of *Heliothis virescens* over 265 amino acids. Hz1V037 shows a 35% identity to the rRNA methyltransferase J large subunit of Ac*M*NPV over 944 amino acids. Hz1V003 shows a 21% identity to histidine kinase of *Bacillus halodurans* over 568 amino acids (Table 1).

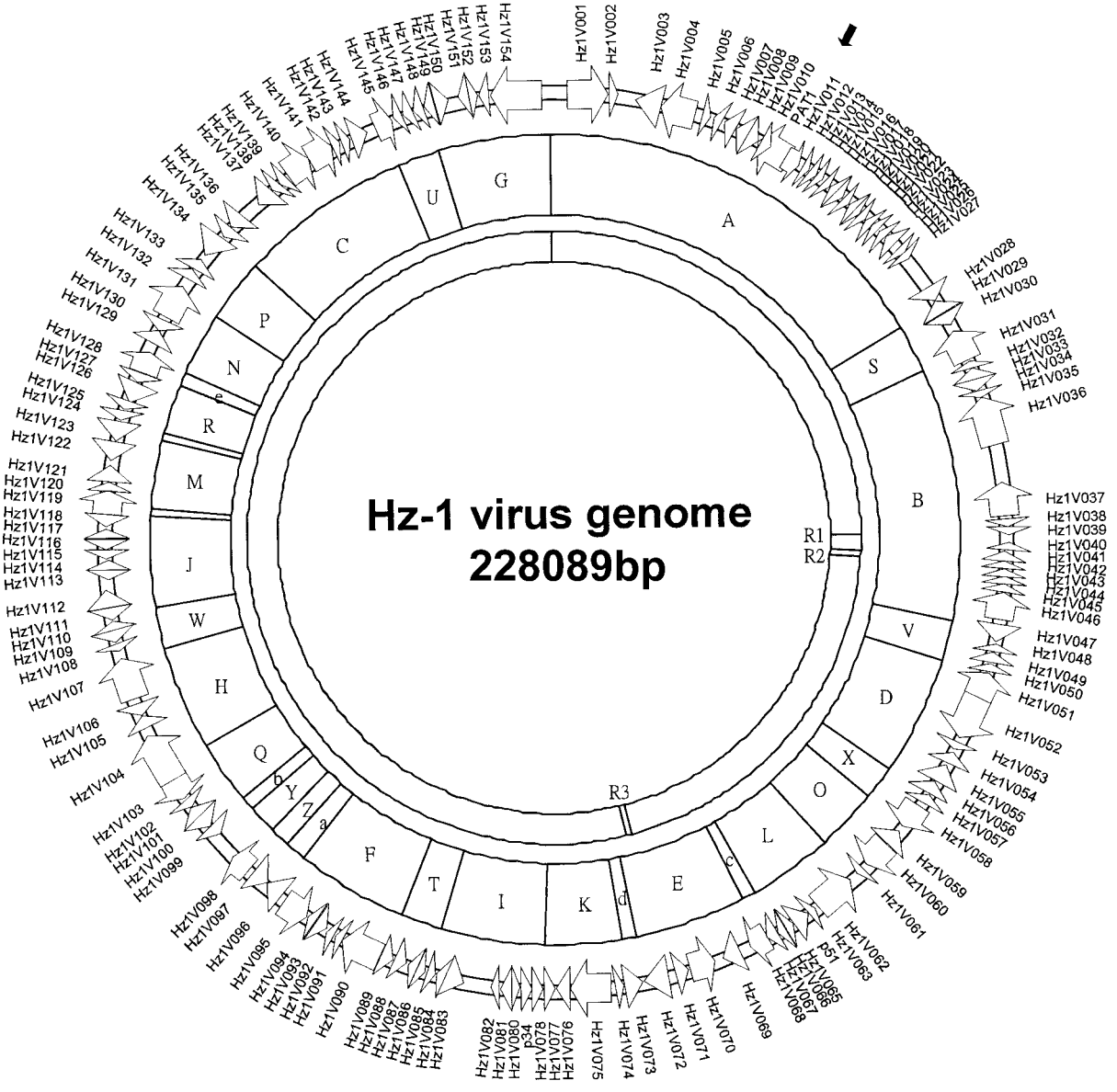

FIG. 1. Circular representation of the Hz-1 virus genome. The arrows indicate the positions (outer ring) of the 154 ORFs and the noncoding PAT1. X, sites of *Xho*I restriction enzymes (inner ring [positions are indicated in parentheses]). The repeated regions (R1, R2, and R3) are indicated in the innermost ring. The position of PAT1 is indicated by a heavy arrow.

The presence of nucleotide metabolism enzymes suggests that the Hz-1 virus may synthesize its nucleotides independently of the host cell machinery. A number of viral genes could be acquired from genome DNA or from another virus infecting a common host (53). Molecular mimicry or genetic piracy, with respect to the utilization of cellular genes captured and modified during the course of viral evolution, has been an area of increasing research with the expansion in virus genome sequencing (19). A common feature of these captured genes is that they are nonessential for virus replication in vitro and that they confer a selective advantage for virus replication, persistence, and spread or in dealing with host cell differentiation and immune defense mechanisms in vivo (9, 10, 47, 55). However, few cases have reported on a viral gene captured from its insect host, although many viruses can infect insects.

TABLE 2. Comparisons of the genome of Hz-1 virus and various baculoviruses

Virus	Genome size	No. of genes	G+C content (%)	Gene density (kb)	No. of HRs	Source or reference
Hz-1	228108	155	41.8	1.47	ND[a]	This study
Ld*M*NPV	161046	163	57.5	0.99	13	43a
Se*M*NPV	135611	139	44.0	0.98	6	36
Ac*M*NPV	133894	154	41.0	0.87	8	5
Op*M*NPV	131990	152	55.0	0.87	5	3
HearNPV	131403	135	39.1	0.97	5	17
BmNPV	128413	136	40.0	0.94	7	25
CuniNPV	108252	109	50.9	0.99	4	2
XecnGV	178733	181	40.7	0.99	8	31
PxGV	100999	120	40.7	0.84	4	30

[a] ND, no homologous regions were detected in the Hz-1 virus genome.

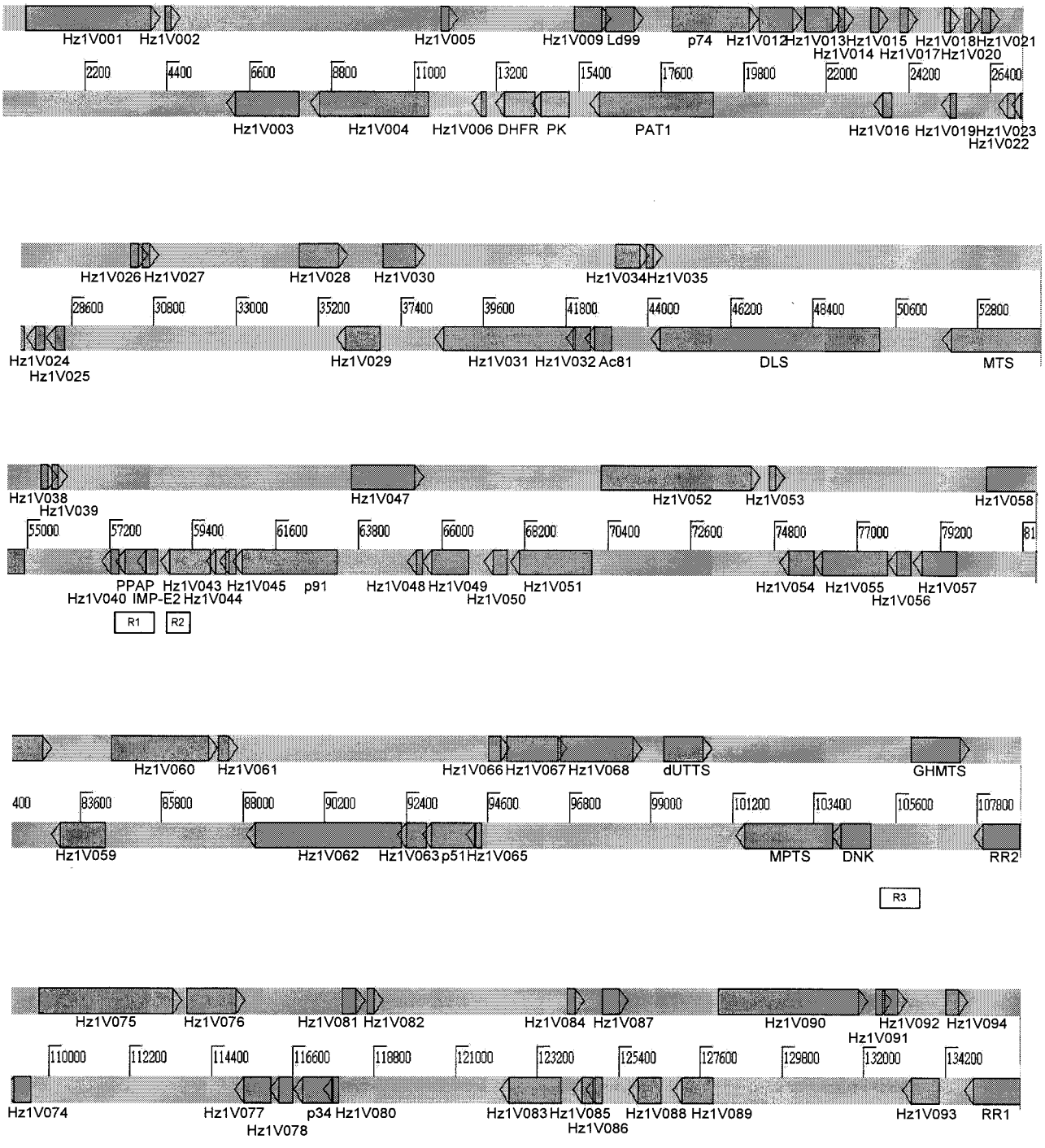

FIG. 2. Layout of the genes and elements in the Hz-1 virus genome. The genome is shown expanded from the representation in Fig. 1. The protein-coding regions and orientations for the recognized genes are as listed in Table 1. The repeated regions (R1, R2, and R3) are indicated.

Proteins involved in DNA replication and transcription. Two genes involved in DNA replication were identified in the Hz-1 virus genome. Hz-1 virus DNA ligase (product of Hz1V036) shows a 44% identity to that of *Homo sapiens* over 1,412 amino acids. The Hz-1 virus DNA polymerase (a product of Hz1V131) was putatively identified by the presence of three highly conserved motifs that are found in most eukaryotic DNA polymerases gene, as well as in some viral polymerases. Among the three homologues to DNA transcription genes, Hz1V075 shows a 26% identity to yeast (*Saccharomyces kluyveri*) plasmid pSKL encoding DNA-directed RNA polymerase (33) over 1,203 amino acids. Hz1V090 shows a 24% identity to DNA-dependent RNA polymerase LEF-8 (orf50) of BmNPV over 1,264 amino acids (Table 1). Hz1V121 shows a high homology to VLF-1 (56% identity to that of Se*M*NPV over 334 amino acids). VLF-1 has been shown to regulate very late gene

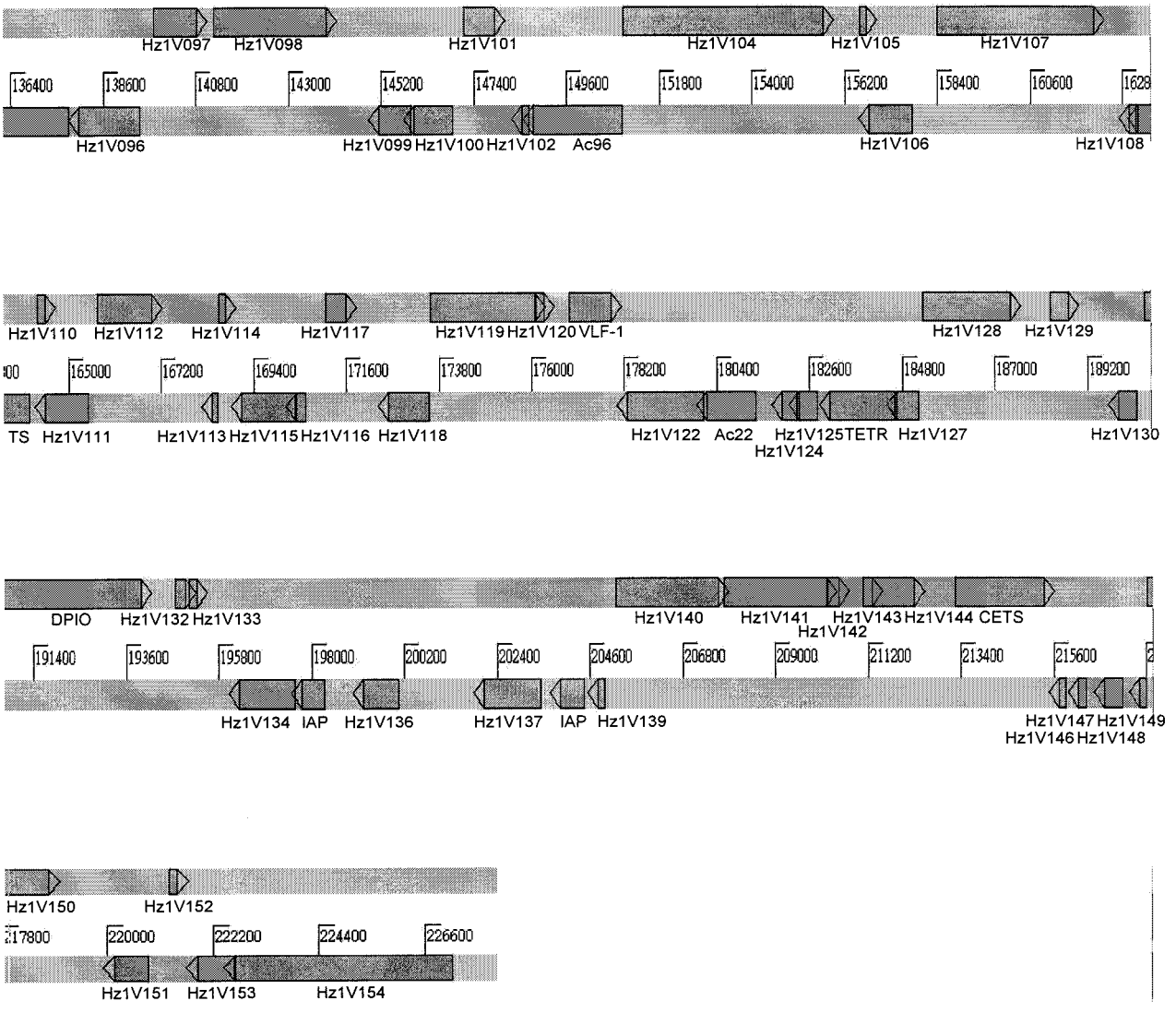

FIG. 2—*Continued.*

expression (45) and also plays a crucial role in the replication of the budded virus form of Ac*M*NPV (60).

Structural proteins. Among the 29 ORFs of Hz-1 virus that show significant homologies to known proteins, two ORFs show significant homologies to two of baculovirus structural proteins: p74 and vp91-capsid associated protein (products of Hz1V011 and Hz1V046, respectively). No homologs of gp64/ld130 groups were identified in the predicted Hz-1V ORFs. Hz1V012 shows a 21 to 23% amino acid identity to p74 of 9 NPVs, 2 GVs, and *Culex nigipalpus* NPV. The amino acid sequences of the p74 gene homologs among the 12 baculoviruses ranged from 31 to 91% identity (56). It has been shown that p74 is localized on the outside of the virion envelope and is important for virus entry into insect midgut cells (24). The N terminus of p74 is exposed on the ODV surface and the C terminus of p74 acts as a TM anchor (24). A remarkable double TM domain at the C terminus of p74 was detected among the 12 baculoviruses and the Hz-1 virus by using the Hidden-Markov model. However, a region with a high amino acid sequence similarity between the Hz-1 virus p74 and that in the baculoviruses was located between amino acids 7 and 579, upstream of the double TM domain (amino acids 624 to 646 and amino acids 673 to 682) (Fig. 3). It has been suggested that the C terminus of p74 directs the protein into the envelope that surrounds occluded virions (56). Consistent with this hypothesis, the deletion of the C terminus of Ac*M*NPV p74 abolishes insect oral infectivity but does not interfere with virus replication in cell cultures (43). No occlusion bodies (OBs) were detected for the Hz-1 virus in cell cultures.

The Hz1V046 product shows a 35% amino acid identity to vp91 capsid-associated protein, a baculovirus capsid-associated protein (5). In both Hz1V046 and the vp91 capsid-associated protein, a region containing chitin-binding peritrophin A domain (amino acids 250 to 305) was revealed by PFAM searches (Fig. 3).

A continuous 9.6-kb locus contains five homologues of cellular genes. A continuous 9.6-kb sequence located at 43.6 to 47.8 map units contained five genes homologous to cellular genes (Hz1V069 to Hz1V073). The Hz1V070 product shows 29% amino acid sequence identity to *Mus musculus* matrix

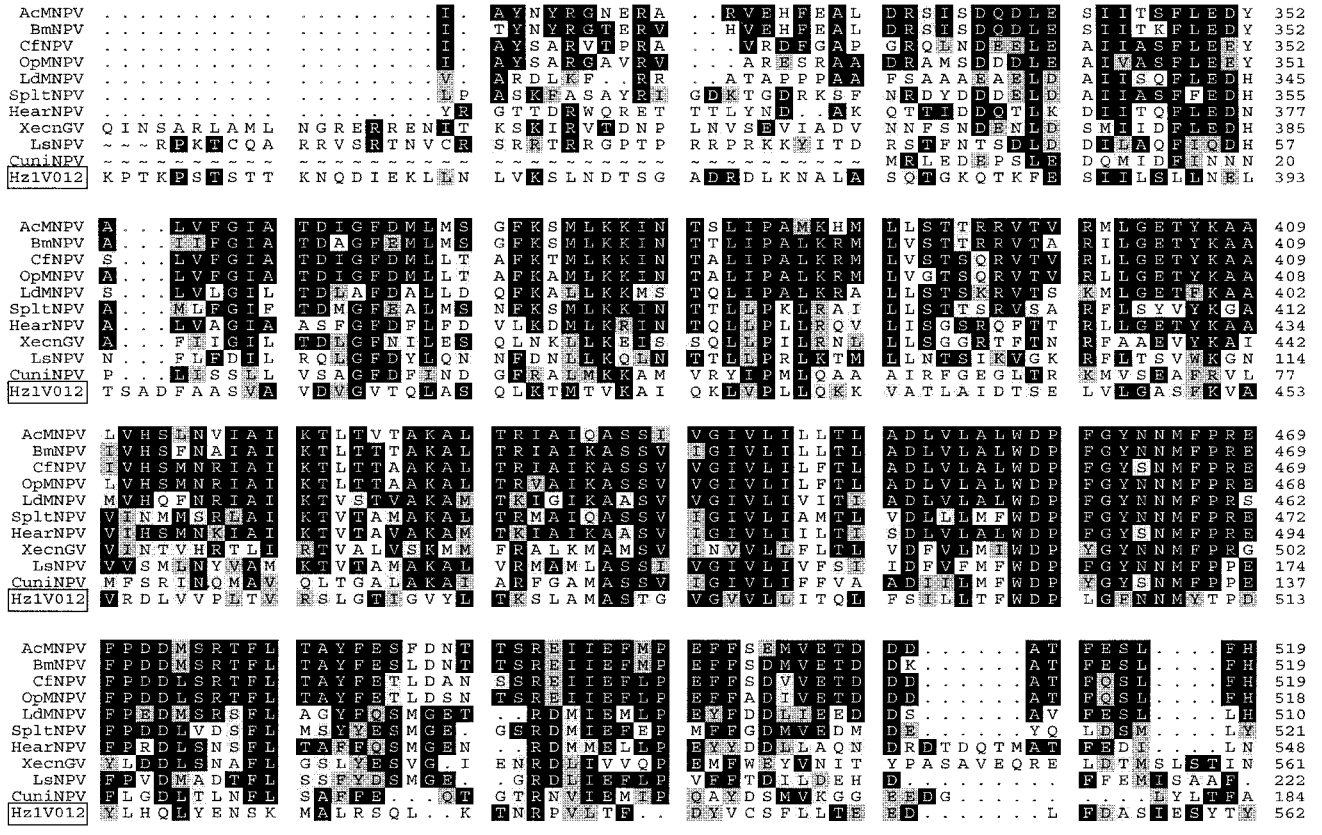

FIG. 3. Multiple amino acid sequence alignment of Hz-1 virus p74 homolog (product of Hz1V011) with the p74 of BmNPV (NP_047536.1), Ac*M*NPV (NP_054168.1), Op*M*NPV (U75930), CfNPV (M97904), Ld*M*NPV (NP_047663.1), HearNPV (NP_075089.1), *Spodoptera litura* NPV (SpltNPV) (AJ01155858), XecnGV (NP_059225.1), Leucania separata NPV (LsNPV) (AB009455), and CuniNPV (AF274288). Identity with consensus is denoted by black box. Similarity with the consensus is denoted by gray shading, differences are indicated by white, and gaps in the alignment are indicated by dots. The positions of the amino acid sequences are indicated on the right.

metalloproteinase (MMP) over 185 amino acids. An MMP gene was also detected in the *Xestia c-nigrum* GV (XecnGV) genome and encodes an ORF of 469 amino acids that has the conserved catalytic domains of human MMP3 (40). MMP homologs have not been reported for other insect viruses. The Hz1V072 product shows a 61% amino acid identity to *Drosophila melanogaster* serine hydroxymethyltransferase (SHMT). It has been suggested that thymidylate synthase and DHFR, along with SHMT, form a metabolic cycle that methylates dUMP to dTMP. The SHMT enzyme catalyzes the THF-dependent reversible conversion of serine to glycine (18).

The acquisition of several continuous cellular proteins has been observed in Kaposi sarcoma-associated herpesvirus (KSHV, also named human herpesvirus 8) (48). A single 13-kb locus in the KSHV genome contains nine ORFs that are homologous to or related to cellular proteins (48). These include a complete *ts* gene, a DHFR gene, four novel cytokine genes, and a *bcl-2* homologue (48).

Other cellular homologs. The Hz1V126 product encodes an ORF of 513 amino acids that shows a 48% amino acid identity to CG4526 of *Drosophila melanogaster* which harbors a sugar transporter domain from amino acids 53 to 508. Interestingly, the Hz1V145 product encodes an ORF of 699 amino acids that shows a 32% amino acid identity to *Drosophila melanogaster* juvenile hormone esterase (JHE) that is involved in insect morphogenesis (52). The Hz1V-JHE contains the conserved carboxylesterase domain from amino acids 3 to 453. No JHE homologues have been detected in other insect viruses. JHE inactivates juvenile hormone, which regulates the outcome of an insect molt, and is an essential enzyme for normal insect development. It has been an attractive targets for biorationally designed, environmentally safe pesticides (8, 21). The expression of JHE by the Hz-1 virus may influence the developmental characteristics, weight gain, and time of mortality of the insects.

Comparison of the Hz-1 virus ORFs with baculovirus ORFs. The complete sequences of five NPVs and two GVs pathogenic for *Lepidoptera*, including Ac*M*NPV, BmNPV, Op*M*NPV, Se*M*NPV, Ld*M*NPV, PlxyGV, and XecnGV, yielded a wealth of information relating to the gene structure and organization of baculoviruses. A comparison of the Hz-1 virus gene content to that of the baculoviruses (31) shows that some Hz-1 virus genes are conserved in baculoviruses, including 11 genes involved in DNA replication, transcription, structural proteins, and the regulation of host metabolism, and five unknown baculovirus homologues (Table 3). However, the most conserved baculovirus gene, *polyhedrin* or *granulin*, was not identified in the Hz-1 virus genome.

Further comparison of the Hz-1 virus genome with the NPVs and GVs genomes was carried out by using the prFLAG

TABLE 3. Baculovirus gene homologs present in the Hz-1 virus genome

Gene type	Conserved genes[a]	Variable genes[b]
Replication genes	*dna-pol*	*rr1*, *rr2*, *dna-lig*, *dutpase*
Transcription-specific genes	*vlf-1*	
Structural protein genes	*p74*, *vp91-capsid*, *pk-1*	
Genes homologs that may alter host metabolism	*iap1*, *iap2*	
Unknown	*AcOrf-22*, *AcOrf-81*, *AcOrf-96*, *PxOrf-29*, *PxOrf-70*	

[a] Conserved genes are identified in all the eight baculoviruses.
[b] Variable genes are identified in some of the eight baculoviruses.

pairwise program. The results showed that no significant similarity was found either between the Hz-1 virus and Ac*M*NPV, HearNPV, PlxyGV, and XecnGV genomes (data not shown). Interestingly, the low similarity between the NPVs and GVs suggests that the Hz-1 virus is at least a different genus from these groups.

Phylogeny of Hz-1 virus. Although the Hz-1 virus has been excluded from the *Baculoviridae*, the data from our sequence analysis indicate a striking similarity, including 16 genes that appear to be shared between the two groups. These include baculovirus homologs that are conserved for DNA replication, gene expression regulation, structural genes, and genes that affect the host metabolism of baculoviruses. However, the lack of homologous regions within the entire Hz-1 virus genome, an overall low level of identity between ORFs, the lack of conserved gene order, the absence of many genes present in all lepidopteran baculovirus genomes, and the lack of polyhedrin/granulin and p10 homologs all suggest that the Hz-1 virus is distantly related to the baculoviruses. Recently, a similar situation has been reported for the *Culex nigripalpus* NPV (CuniNPV), which is a pathogenic baculovirus for dipteran, except that it has four putative *hrs* (2, 46). CuniNPV has globular OBs and is not enveloped (46). The phylogenetic analysis of CuniNPV p74 showed that it is a member of the baculovirus lineage distinct from lepidopteran NPVs and GVs (46). The phylogenetic analysis of Hz-1 virus p74 suggests that it is even further away from the NPVs and GVs than CuniNPV is (Fig. 4). The Hz-1 virus and CuniNPV genomic sequence analyses results suggest that these viruses may have gone through different evolution pathways from those of the NPVs and GVs. However, unlike CuniNPV, neither *hrs* nor OBs were identified in the Hz-1 virus genome. In conclusion, our analysis of the Hz-1 virus genome sequence revealed that the differences outnumber the similarities between the Hz-1 virus and baculoviruses and suggest that the Hz-1 virus may form a different class of virus that is distinct from *Baculoviridae*. Further comparison of the Hz-1 virus with other nudiviruses and baculoviruses should provide important insights into the evolution of both invertebrate and vertebrate viruses.

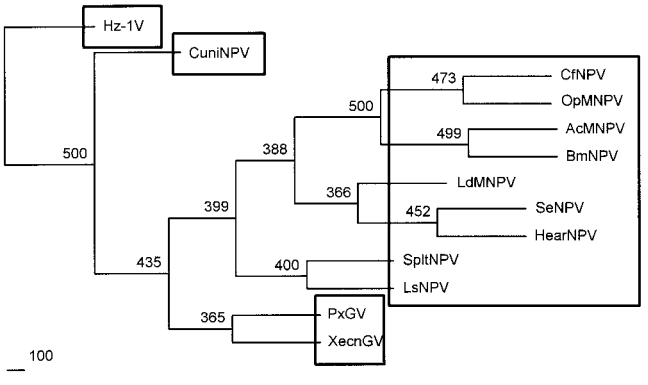

FIG. 4. Phylogenetic analysis of the p74 gene based on the amino acid distances from a variety of organisms. The trees were constructed by using a UPGMA clustering method. Bootstrap values of >50 are shown above the node in 500 replications.

ACKNOWLEDGMENTS

We thank A. Yaw, P. I. Huang, and K. H. Liang of the Biomedical Engineering Center, Industrial Technology Research Institute, Hsinchu, Taiwan, for assistance with prFLAG. We thank H. R. Cheng and T. W. Hsu for excellent technical assistance.

This work was supported by grant NSC89-2311-B-006-003 from the National Science Council of Taiwan.

REFERENCES

1. **Afonso, C. L., E. R. Tulman, Z. Lu, E. Oma, G. F. Kutish, and D. L. Rock.** 1999. The genome of *Melanoplus sanguinipes* entomopoxvirus. J. Virol. **73:** 533–552.
2. **Afonso, C. L., E. R. Tulman, A. Lu, C. A. Balinsky, B. A. Moser, J. J. Becnel, D. L. Rock, and G. F. Kutish.** 2001. Genome sequence of a baculovirus pathogenic for *Culex nigripalpus.* J. Virol. **75:**11157–11165.
3. **Ahrens, C. H., R., Russell, C. J. Funk, J. T. Evans, S. H. Harwood, and G. F. Rohrmann.** 1997. The sequence of the *Orygia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus genome. Virology **229:**381–399.
4. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
5. **Ayres, M. D., S. C. Howard, J. Kuzio, M. Lopez-Ferber, and R. D. Possee.** 1994. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. Virology **202:**586–605.
6. **Baldo, A. M., and M. A. McClure.** 1999. Evolution and horizontal transfer of dUTPase-encoding genes in viruses and their hosts. J. Virol. **73:**7710–7721.
7. **Bonfield, J. K., K. F. Smith, and R. Staden.** 1995. A new DNA sequence assembly program. Nucleic Acids Res. **23:**4992–4999.
8. **Bonning, B. C., P. W. Roelvink, J. M. Vlak, R. D. Possee, and B. D. Hammock.** 1994. Superior expression of juvenile hormone esterase and beta-galactosidase from the basic promoter of *Autographa california* nuclear polyhedrosis virus compared to the p10 protein and polyhedrin promoters. J. Gen. Virol. **75:**1551–1556.
9. **Brazas, R., and D. Ganem.** 1996. A cellular homolog of hepatitis delta antigen: implications for viral replication and evolution. Science **274:**90–94.
10. **Bugert, J. J., and G. Darai.** 2000. Poxvirus homologues of cellular genes. Virus Genes **21**:111–133.
11. **Burand, J. P., C. Y. Kawanishi, and Y. S. Huang.** 1986. *Persistent baculovirus infection*, p. 159–175. *In* R. R. Granados and B. A. Fedderici (ed.), The biology of baculovirus. CRC Press, Inc., Boca Raton, Fla.
12. **Carreras, C. W., and D. V. Santi.** 1995. The catalytic domain and structure of thymidylate synthase. Annu. Rev. Biochem. **64:**721–762.
13. **Chao, Y. C., M. Hamblin, and H. A. Wood.** 1990. Physical map of Hz-1 baculovirus genome from standard and defective interfering particles. J. Gen. Virol. **71:**1265–1270.
14. **Chao, Y. C., H. A. Wood, C. Y. Chang, H. J. Lee, W. C. Shen, and H. T. Lee.** 1992. Different expression of Hz-1 baculovirus genes during productive and persistent infections. J. Virol. **63:**1442–1448.
15. **Chao, Y. C., S. T. Lee, M. C. Chang, H. H. Chen, S. S. Chen, T. Y. Wu, F. H. Liu, E. L. Hsu, and R. F. Hou.** 1998. A 2.9-kilobase noncoding nuclear RNA functions in the establishment of persistent Hz-1 viral infection. J. Virol. **72:**2233–2245.
16. **Chen, H. H., D. J. Tso, W. B. Yeh, H. J. Cheng, and T. F. Wu.** 2001. The thymidylate synthase gene of Hz-1 virus: a gene captured from its *lepidopteran* host. Insect Mol. Biol. **10:**495–503.
17. **Chen, X., M. Li, X. Sun, B. M. Arif, Z. Hu, and J. M. Vlak.** 2000. Genomic organization of *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus. Arch. Virol. **45:**2539–2555.
18. **Chu, E., and C. J. Allegra.** 1996. The role of thymidylateee synthase in cellular regulation. Bioessays **18:**191–198.
19. **Davis-Poynter, N. J., M. Degli-Espsti, and H. E. Farrell.** 1999. Murine cytomegalovirus homologues of cellular immunomodulatory genes. Intervirology **42:**331–341.

20. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
21. **Eldridge, R., D. R. O'Relly, B. D. Hammock, and L. K. Miller.** 1992. Insecticidal properties of genetically engineered baculoviruses expressing an insect juvenile hormone esterase gene. Appl. Environ. Microbiol. **58:**1583–1591.
22. **Ewing, B., and P. Green.** 1998. Base-calling of automated sequencer traces using PHRED. II. Error probabilities. Genome Res. **8:**186–194.
23. **Ewing, B., L. Hillier, and M. C. Wendl, and P. Green.** 1998. Base-calling of automated sequencer traces using PHRED. I. Accuracy assessment. Genome Res. **8:**175–185.
24. **Faulker, P., J. U. Kuzio, G. V. Williams, and J. A. Wilson.** 1997. Analysis of p74, a PDV envelope protein of *Autographa californa* nucleopolyhedrovirus required for occlusion body infectivity in vivo. J. Gen. Virol. **78:**3091–3100.
25. **Gomi, S., K. Majima, and S. Maeda.** 1999. Sequence analysis of the genome of *Bombyx mori* nucleopolyhedrovirus. J. Gen. Virol. **80:**1323–1337.
26. **Granados, R. R., T. Nguyen, and B. Cato.** 1978. An insect cell line persistently infected with a baculovirus-like particle. Intervirology **10:**309–317.
27. **Guarino, L., and M. D. Summers.** 1986. Interspersed homologous DNA of *Autographa californica* nuclear polyhedrosis virus enhances delayed-early gene expression. J. Virol. **60:**215–223.
28. **Guttieri, M. C., and J. P. Burand.** 1996. Nucleotide sequence, temporal expression, and transcriptional mapping of the p34 late gene of the Hz-1 insect virus. Virology **223:**370–375.
29. **Guttieri, M. C., and J. P. Burand.** 2001. Location, nucleotide sequence, and regulation of the p51 late gene of the Hz-1 insect virus: identification of a putative late regulatory element. Virus Genes **23:**17–25.
30. **Hashimoto, Y., T. Hayakawa, Y. Ueno, T. Fugita, Y. Sano, and T. Matsumoto.** 2000. Sequence analysis of the *Plutella xylostella* granulovirus genome. Virology **275:**358–372.
31. **Hayakawa, T., G. F. Rohrmann, and Y. Hashimoto.** 2000. Patterns of genome organization and content in lepidopteran baculoviruses. Virology **278:**1–12.
32. **Hink, W. F.** 1972. Insect tissue culture. Adv. Appl. Microb. **15:**157–214.
33. **Hishinuma, F., and K. Hirai.** 1991. Genome organization of the linear plasmid, pSKL, isolated from *Saccharomyces kluyveri.* Mol. Gen. Genet. **226:**97–106.
34. **Huang, Y. S., M. Hedberg, and C. Y. Kawanish.** 1982. Characterization of the DNA of a nonoccluded baculovirus, Hz-1 V. J. Virol. **43:**174–181.
35. **Ignoffo, C. M., M. Shapior, and W. F. Hink.** 1971. Replication and serial passage of infectious *Heliothis* nuclear polyhedrosis virus in an established line of *Heliothis zea* cells. J. Invertebr. Pathol. **18:**131–134.
36. **Ijkel, W. F., E. A. van Strien, J. G. M. Jeldens, R. Broer, D. Zuiderma, R. W. Goldbach, and J. M. Vlak.** 1999. Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome. J. Gen. Virol. **80:**3289–3304.
37. **Ives, D. H., and S. Ikeda.** 1998. Life on the salvage path: the deoxynucleoside kinase of *Lactobacillus acidophilus* R-26. Prog. Nucleic Acid Res. Mol. Biol. **59:**205–255.
38. **Jordan, A., and P. Reichard.** 1998. Ribonucleotide reductases. Annu. Rev. Biochem. **67:**71–98.
39. **Kelly, D. C., T. Lescott, M. D. Ayres, D. Carey, A. Countts, and K. A. Harrap.** 1981. Induction of a nonoccluded baculovirus persistently infecting *Heliothis zea* cells by *Heliothis armigera* and *Trichopuisa ni* nuclear polyhedrosis virus. Virology **112:**174–180.
40. **Ko, R., K. Okano, and S. Maeda.** 2000. Structural and functional analysis of the *Xestia c-nigrum* granulovirus matrix metalloproteinase. J. Virol. **74:**11240–11246.
41. **Kool, M., and J. M. Vlak.** 1993. The structural and functional organization of the *Autographa californica* nuclear polyhedrosis virus genome. Arch. Virol. **130:**1–16.
42. **Korngberg, A., and T. A. Baker.** 1992. DNA replication, 2nd ed. Freeman, San Francisco, Calif.
43. **Kuzio, J., R. Jacques, and P. Faulkner.** 1989. Identification of p74, a gene essential for virulence of baculovirus occlusion bodies. Virology **173:**759–763.

43a. **Kuzio, J., M. N. Pearson, S. H. Harwood, C. J. Funk, J. T. Evans, J. M. Slavicek, and G. F. Rohrmann.** 1999. Sequence and analysis of the genome of a baculovirus pathogenic *Lymantria dispar*. Virology **253:**17–34.

44. **McIntosh, A. H., and C. M. Ignoffo.** 1981. Establishment of a persistent baculovirus infection in a lepidopteran cell line. J. Invertebr. Pathol. **8:**395–403.
45. **McLachlin, J. R., and Miller, L. K.** 2001. Identification and characterization of *vlf-1*, a baculovirus gene involved in very late gene expression. J. Virol. **68:**7746–7756.
46. **Moser, B. A., J. J. Bencnel, S. E. White, C. Afonos, G. Kutish, S. Shanker, and E. Almira.** 2001. Morphological and molecular evidence that *Culex nigripalpus* baculvoirus is an unusual member of the family *Baculoviridae.* J. Gen. Virol. **82:**283–297.
47. **Muller, K., C. Tidona, A. U. Bahr, and G. Darai.** 1998. Identification of a thymidylate synthase gene within the genome of *Chilo* iridescent virus. Virus Genes **17:**243–258.
48. **Nicholas, J., V. Ruvolo, J. Zong, S. Ciufo, H. G. Guo, M. Reitz, and G. S. Hayward.** 1997. A single 13-kilobase divergent locus in the Kaposi sarcoma-associated herpesvirus (human herpesvirus 8) genome contains nine open reading frames that are homologous to or related to cellular proteins. J. Virol. **71:**1963–1974.
49. **Pearson, M. N., and G. F. Rohrmann.** 1995. *Lymanthia dispar* nuclear polyhedrosis virus homologous regions: characterization of their ability to function as replication origins. J. Virol. **69:**213–321.
50. **Pearson, W. R.** 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol. **183:**63–98.
51. **Ralston, A. L., Y. S. Huang, and C. Y. Kawanish.** 1981. Cell culture studies with the IMC-Hz-1 nonoccluded virus. Virology **115:**33–44.
52. **Roe, R. M., and K. Venkatech.** 1990. Metabolism of juvenile hormones: degradation and titer regulation, p. 125–180. *In* A. P. Gupta (ed.), Morphogenic hormones of arthropods, vol. 1. Rutgers University Press, New Brunswick, N.J.
53. **Rossnick, M. J.** 1997. Mechanisms of plant virus evolution. Annu. Rev. Phytopathol. **354:**191–209.
54. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M.-A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.
55. **Shimizu, N., and T. Gojobori.** 2000. How can human and simian immunodeficiency virused utilize chemokine receptors as their coreceptors? Gene **259:**199–205.
56. **Slack, J. M., E. M. Pougherty, and S. D. Lawrence.** 2001. A study of the *Autographa californica* multiple nucleopolyhedrosis virus ODV envelope protein p74 using a GFP tag. J. Gen. Virol. **82:**2279–2287.
57. **Smith, G. E., and Summers, M. D.** 1982. DNA homology among subgroup A, B, and C baculovirus. Virology **123:**393.
58. **Summers, M. D., and G. E. Smith.** 1987. A manual of methods for baculovirus vectors and insect cell culture procedure, p. 33–35. Department of Entomology Texas Agricultural Experiment Station, Texas A&M University, College Station, Tex.
59. **Volkman, L. E., G. W. Blissard, P. D. Friesen, B. A. Keddie, R. D. Possee, and D. A. Theilmann.** 1995. Virus taxonomy: The classification and nomenclature of viruses, p. 104–113. *In* F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Javis, G. P. Martelli, M. A. Mayo, and M. D. Summers (ed.), The sixth report of the International Committee on Taxonomy of Viruses. Springer-Verlag Wein, Inc., New York, N.Y.
60. **Yang, S., and L. K. Miller.** 1998. Expression and mutational analysis of the baculovirus very late factor 1 (*vlf-1*) gene. Virology **245:**99–109.