ELSEVIER

Virus Research 100 (2004) 179–189

Research

www.elsevier.com/locate/virusres

Characterization and phylogenetic analysis of the chitinase gene from the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus

Hualin Wang [a,b,d], Dong Wu [a,b], Fei Deng [a,b], Huiyin Peng [a], Xinwen Chen [a], Hilary Lauzon [c], Basil M. Arif [c], Johannes A. Jehle [d], Zhihong Hu [a,*]

[a] *Joint-Laboratory of Invertebrate Virology and Key Laboratory of Molecular Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, PR China*

[b] *Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430071, PR China*

[c] *Laboratory for Molecular Virology, Great Lakes Forestry Research Centre, Sault Ste. Marie, Ont., Canada P6A 5M7*

[d] *Section Biotechnological Crop Protection, State Education and Research Center for Agriculture, Viticulture and Horticulture, 67435 Neustadt/Wstr., Germany*

Received 2 September 2003; received in revised form 19 November 2003; accepted 19 November 2003

Abstract

A putative chitinase gene was identified within the fragment *Eco*RI-K of the *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus (HearNPV, also called HaSNPV) genome. The open reading frame (ORF) contains 1713 nucleotides (nt) and encodes a protein of 570 amino acids (aa) with a predicted molecular weight of 63.6 kDa. Transcription started at about 18 h post infection (p.i.) and the protein was first detected at 20 h p.i. The times of transcription and expression are characteristic of a late baculovirus gene. 5′ and 3′ RACE indicated that transcription was initiated from the adenine residue located at −246 nt upstream from the ATG start site and the poly (A) tail was added at 267 nt downstream from the stop codon. This is the first report on the molecular characterization of a *chitinase* from a single nucleocapsid NPV. The phylogeny of baculoviral chitinase genes were extensively examined in comparison with chitinases derived from bacteria, fungi, nematode, actinomycetes, viruses, insects and mammals. Neighbor-joining and most parsimony analyses showed that the baculoviral chitinases were clustered exclusively within γ-proteobacteria. Our results strongly suggest that baculoviruses acquired their chitinase genes from bacteria.

Keywords: HaSNPV; Chitinase; Characterization; Phylogenetic analysis; Secondary structure analysis

1. Introduction

The family Baculoviridae is of rod-shaped viruses with circular, supercoiled, double-stranded DNA genomes of a size ranging from about 80–180 kbp. Baculoviruses are insect specific pathogens particularly for *Lepidoptera*, *Hymenoptera* and *Diptera* (Moscardi, 1999). Generally, baculoviruses are host-specific and cause infected larvae to liquefy shortly after the onset of mortality (Volkman and Keddie, 1990). The liquefaction of cadavers is probably pivotal to the release of progeny virus and its dissemination into the environment. Two baculoviral gene products, chitinase and V-cathepsin, are considered responsible for the liquefaction process (Hawtin et al., 1997; Hom and Volkman, 2000; Ohkawa et al., 1994). V-cathepsin, a protease, is believed to be responsible for the degradation of the proteinaceous components of cadavers, whereas the chitinase is essential in liquefying infected cadavers. Genes encoding chitinases are present in most of the so far completely sequenced baculoviral genomes, except that of *Plutella xylostella* granulovirus (PxGV; Hashimoto et al., 2000). The first chitinase gene in baculoviridae was identified from the *Autographa californica* (Ac) MNPV. It was extensively studied as a prototype and was found to have endo- and exo-chitinase activities in a wide pH range from 3.0 to 10.0. In AcMNPV-infected Sf21 cells, the expressed chitinase is located in the endoplasmic reticulum (Hawtin et al., 1995,1997).

Chitinases are ubiquitous in nature and have been isolated from bacteria, fungi, nematodes, plants, insects and some vertebrates (Fukamizo, 2000; Koga et al., 1999; Raghaven et al., 1994). Usually, chitinases are grouped into distinct classes according to the divergence of amino acid sequences.

* Corresponding author. Tel.: +86-27-87197180; fax: +86-27-87198072.
E-mail address: huzh@pentium.whiov.ac.cn (Z. Hu).

0168-1702/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2003.11.015

In plants, there are at least three classes of chitinases, while in the bacterium *Serratia marcescens*, chitinases have been grouped into at least six classes (Shinshi et al., 1990). According to this classification, baculovirus chitinases were unambiguously classified as chitinase A (chiA). However, the classification of chitinases based on the divergence of sequences was questioned since it caused confusion and was not efficient (Henrissat, 1999). A simple system was suggested to nomenclature chitinases as family 18 chitinases and family 19 chitinases (chi18 and chi19, respectively) depending on their catalytic active site (Henrissat, 1999). Chi18 are the most widely present chitinases found in procaryotes, eucaryotes and viruses. The baculoviral chitinases identified so far are all belonging to family 18 chitinases. It was previously suggested that the AcMNPV chitinase is closely related to the chitinase of *S. marcescens* (Hawtin et al., 1995). Further phylogenetic analyses, however, could not confirm the proposed lateral transfer from bacteria to baculoviruses (Kang et al., 1998).

The cotton bollworm, *Helicoverpa armigera*, is a world-wide pest causing serious damage to cotton, tobacco, pepper and tomato plants (Fitt, 1989). The HaSNPV has been extensively used for the control of *H. armigera* since the first isolation of this pathogen in Hubei province of China. The genome of this virus has been completely sequenced (Chen et al., 2001) and several of its genes such as the *egt* (Chen et al., 1997), *lef*-2 (Chen et al., 1999), *p6.9* (Wang et al., 2001), Ha122 (Long et al., 2003) and Ha94 (Fang et al., 2003) were characterized. In this paper, we describe the molecular properties of the HaSNPV chitinase gene including temporal transcriptional regulation and immunological properties. Additionally, we performed extensive phylogentic analyses including secondary structure comparison in order to shed more light into the evolutionary relationship between baculoviral and other chitinases. Accession number AF243498.

2. Materials and methods

2.1. Virus and insect cells

HaSNPV was initially isolated from diseased *H. armigera* larvae collected in Hubei province of China and propagated in the fourth instar larvae (Zhang et al., 1981; Sun and Zhang, 1994). The insect cell line Hz2e5 was maintained at 27 °C in Grace's medium supplemented with 10% fatal bovine serum (Gibco-BRL) and was infected with budded virus obtained from the hemolymph of diseased larvae. Titers of BVs were determined by the end-point dilution method and expressed as $TCID_{50}$ units/ml (O'Reilly et al., 1992).

2.2. Cloning and sequencing of the HaSNPV chitinase gene

The putative chitinase gene was previously identified in the fragment *Eco*RI-K of the HaSNPV genome (Peng et al., 1998). *Eco*RI-K was cloned into the pTZ19R vector (Gibco-BRL) and denoted as plasmid pCXW125 (Chen et al., 2000). The chitinase ORF and its flanking regions were sequenced in both directions by primer walking (Sanger et al., 1977).

2.3. 3′-RACE and 5′-RACE

3′RACE was conducted by using the 3′-full RACE core set (TaKaRa) as recommended by manufacturer. The oligo(dT)-3 sites adaptor was used for the synthesis of first strand cDNA by incubation with AMV reverse transcriptase and 1 μg of total RNA isolated 72 h post infection (p.i.) for 1 h at 50 °C. The synthesized first strand cDNA was used as template in the following PCR reaction with the three sites adaptor primer and a *chiA* specific forward primer S1: 5′-TTTGTCGAATGGGGCGTTTAC-3′ (493–513 nucleotides (nt) in chitinase ORF). The PCR products were separated with agarose gel and purified using Glassmilk (BioStar). The purified products were cloned into pGEM-T easy vector (Promega) and sequenced with M13 primers.

5′RACE was performed by using 5′-full RACE core set (TaKaRa). The first strand cDNA was synthesized using a 5′-end phosphorated *chiA* specific reverse transcription primer (5′-pGTCAACACAAATTCACGCACG-3′; reverse to 911–891 nt) and AMV reverse transcriptase by incubation for 1 h with 1 μg of total RNA at 50 °C. The RNAs in the mixture was degraded by incubation with RNase H for 1 h at 37 °C. The released first strand cDNA was ligated to circular cDNA by incubation with T4 DNA ligase at room temperature for 3 h. The junction fragment was amplified with two *chitinase* specific primers A1 (5′-GTTTGACTAGCCGAACAGCCG-3′; reverse to 377–357 nt) and S1. A nest-PCR was followed with two specific internal primers A2 (5′-GCCTTGTTCGCCCGACCATAC-3′; reverse to 213–193 nt) and S2 (5′-TTAGCGCATGGAACGAACCCT-3′; 728–748 nt). The nested PCR products were purified from Agarose gel using Glassmilk, cloned into pGEM-T easy vector (Promega) and sequenced with M13 primers.

2.4. Transcription time course

Hz2e5 cells were infected with HaSNPV at a multiplicity of infection (m.o.i.) of 10. Total mRNA, collected at 0, 6, 8, 12, 14, 16, 18, 20, 24, 48, 72 h p.i., was isolated by using Trizol reagent (Gibco-BRL) in a one-step procedure and blotted onto Hybond-N cellulose membrane (Ausubel et al., 1996; Chomczynski and Sacchi, 1987). A chitinase gene specific probe was generated by PCR on pCXW125 template and two primers (forward primer: 5′-ATGAATAATTATTGTTTGTATTTG-3′, reverse primer: 5′-TTCCTATACTAGTAGTCCATTA-3′) located at the start and stop codons, respectively. The product was purified by the Glassmilk method (BioStar), labeled by nick-translation

using [α-^{32}P] dATP, and hybridized to the total mRNA blot (Church and Gilbert, 1984).

2.5. Western blot analysis

Hz2e5 cells were infected with HaSNPV at an m.o.i. of 10 and the protein samples were prepared at 0, 18, 20, 24, 36, 48, 72, 96 h p.i. They were separated on 12% SDS–PAGE gel and transferred onto a PVDF membrane by using a Semi-Dry Transfer Cell (Bio-Rad). The membrane was blocked in 2% fat-free milk at 4 °C overnight. Immuno-detection was performed by using the primary specific antibody against HaSNPV chitinase (Wu et al., 2002) and the secondary antibody against rabbit IgG conjugated with alkaline phosphatase. BCIP and NBT (Roche) detection system were performed as recommended by the manufacturer.

2.6. Computer assisted analysis

The sequence data were compiled and analyzed using DNASTAR software. The deduced amino acid sequence of *chitinase* was explored using the PROSITE network service for putative functional motifs searching (Bairoch et al., 1997). Homologues in GenBank/EMBL databanks were explored by using PSI-BLAST searching tool (Altschul et al., 1992). The sequence alignment was performed using Clustal X software (Thompson et al., 1997). GeneDoc software (version 1.1.1004) was used for similarity shading and scoring of alignment. PAUP program (Version 4.0 beta 10) was used for generating the phylogenetic trees using the neighbor-joining method (NJ), most parsimony method and bootstrap analysis. Phylogenetic tree was visualized using Treeview program. Secondary structures were predicted using network service PHD (profile fed neural network systems from Heidelberg; Rost and Sander, 1993, 1994). The results were imported into Bioedit referring to the corresponding amino acid sequence alignment. Per cent identity of secondary structures were calculated using GeneDoc software. All alignments discussed in the text are available to be downloaded from http:\www.slfa.de\biotech\chitinase.

3. Results

3.1. Sequence analysis

The putative chitinase gene was previously located at different restriction enzyme fragments of the HaSNPV genome by hybridization with CfMNPV chitinase probe (Peng et al., 1998). The positive *Eco*RI-K fragment was cloned and sequenced by primer walking. The sequence was identical to the ORF41 of the complete HaSNPV genomic sequence (Chen et al., 2001). In the HaSNPV genome, the v-cathepsin and chitinase genes are not adjacent to each other as usually found in other baculovirus genomes. The chitinase ORF has a size of 1713 nt, potentially encoding a protein with a predicted molecular weight of 63.6 kDa. The amino acid (aa) sequence was deduced using DNAstar software. It has a size of 570 aa, which is similar to that of known baculoviral chitinases ranging from 551 to 588 aa. The functional motifs of HaSNPV chitinase were explored by using the protein motif searching tool PROSITE. A putative tyrosine kinase phosphorylation motif RASGDLISY (located at aa 507–515), a typical chitinase family 18 active motif FDGIDIDWE (aa 308–317) and six *N*-glycosylation sites were identified. A trans-membrane signal MNNYCLYLFAFSILILHSYA was found at the N-terminus. A characteristic endoplasmic reticulum (ER) targeting sequence, HNEL, was identified at the C-terminus (Fig. 1; Lewis and Pelham, 1990; Pelham, 1990).

3.2. Transcriptional and expressional analysis

The baculoviral consensus late transcription initiation motif ATAAG was located at 29 nt upstream of the putative translation start codon ATG. Two TATA boxes were found at position −33 and −21 nt upstream from the ATG. The typical poly(A) signal, AATAAA, was not present downstream of the chitinase ORF. In order to determine the transcriptional start and stop sites of HaSNPV chitinase, 5′RACE and 3′RACE were performed with total RNA sample isolated at 72 h.p.i. For 5′RACE, three clones were sent for sequencing. The results of two clones were same, indicating that the transcription of HaSNPV *chitinase* was initiated from the nucleotide A located at −246 nt upstream from the start code ATG (Fig. 1). The other clone failed to give any sequencing signal. It is surprising that our results indicate HaSNPV chitinase initiates from TAGT (Fig. 1) instead of the common baculoviral late transcription motif TAAG. The transcriptional initiation site will have to be confirmed by further experiments such as primer extension. For 3′RACE, five clones were sent for sequencing. Four of the five sequences indicated that the poly(A) tail was added at 267 nt downstream from the stop codon and it is 14 nt downstream from the last A of an AATATA sequence (Fig. 1).

The temporal transcription regulation of HaSNPV chitinase gene in vitro was determined by using Northern blot analysis. An HaSNPV *chitinase* specific ^{32}P labelled probe was hybridized to total RNA isolated at 0–72 h p.i. Northern blots showed that a specific transcript with a size of about 2,300 nt was detected by 18 h p.i. (Fig. 2). Thereafter the transcripts accumulated until 72 h p.i. The size of the transcript was in agreement with the size predicted by the results of 5′RACE and 3′RACE. The transcription profile was indicated as a typical pattern of a late baculovirus gene.

In order to investigate the expression profile of the native HaSNPV chitinase in Hz2e5 cells, Western blot analysis was performed on total cellular proteins prepared at 0–96 h p.i. A product of about 66 kDa was detected as early as 20 h p.i. The product increased only slightly with time until 96 h p.i. (Fig. 3). The results indicated that the HaSNPV chitinase

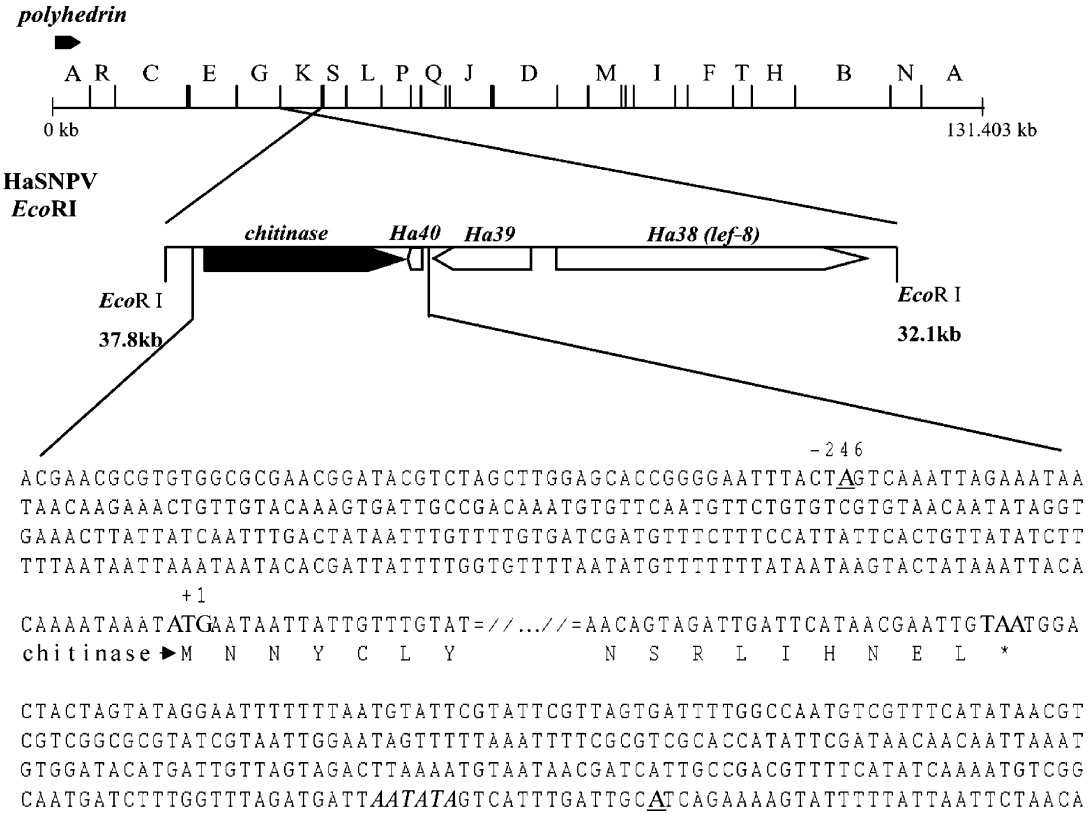

Fig. 1. Location of the chitinase gene on the *Eco*RI map of HaSNPV genome. The position and direction of transcription of *polyhedrin*, *Ha40*, *Ha39*, *Ha38* and *chitinase* (*Ha41*) are shown. Part of the nucleotide sequence and its flanking regions are shown. The predicted amino acids sequence was indicated by one letter code. The transcription initiation and termination nucleotides revealed by 5′ and 3′RACE were bolded and underlined. The putative poly (A) signal AATATA was bolded and italicized.

is a late gene. Its expression pattern was similar to that of AcMNPV (Hawtin et al., 1995).

3.3. Phylogenetic analysis

An extensive study was conducted to elucidate the phylogeny of chitinases derived from various species. The deduced amino acid sequence from HaSNPV *chitinase* was explored using PSI-Blast service against the GenBank. A total of 489 chitinase sequences (all hits with an e value lower than $6e^{-7}$) were downloaded and aligned using Clustal W program. From this alignment a neighbor-joining (NJ) tree was constructed (data not shown) to examine the distances among these sequences. In order to condense the

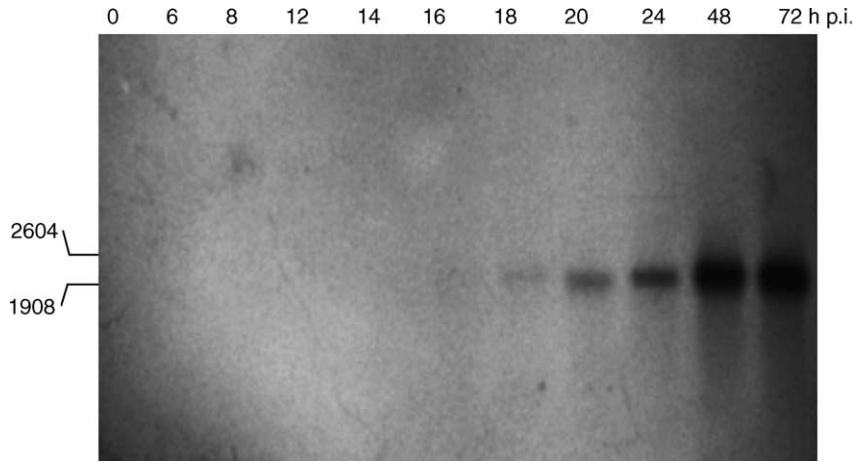

Fig. 2. Northern blot analysis of total mRNA samples extracted at 0, 6, 8, 12, 14, 16, 18, 20, 24, 48 and 72 h p.i. Size of RNA ladder, 2604 and 1909 nt (Promega) are given to the left. The size of the chitinase transcript is about 2300 nt.

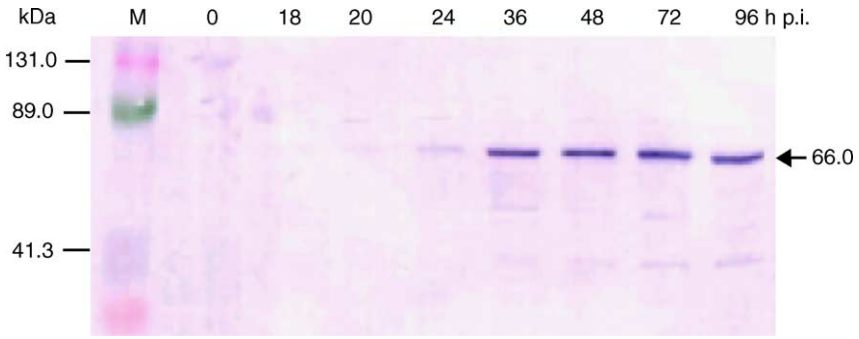

Fig. 3. Expression of HaSNPV chitinase in infected Hz2e5 cells. Samples were prepared at 0, 18, 20, 24, 36, 48, 72 and 96 h p.i., blotted onto PVDF membrane and reacted with anti chitinase antibody. Protein markers are in lane M. The size of the detected chitinase is indicated by an arrow.

Table 1
Chitinase sequences used in this analysis

Taxonomy	Name/source	Abbreviation	Accession
Viruses			
Baculoviridae			
Nucleopolyhedrovirus			
	Helicoverpa armigera SNPV	HaSNPV	AF243498
	Autographa californica MNPV	AcMNPV	AAA66756
	Bombyx mori NPV	BmNPV	NP_047523
	Hyphantria cunea NPV	HycuNPV	AAD31762
	Lymantria dispar NPV	LdMNPV	NP_047707
	Spodoptera exigua NPV	SeMNPV	AAF33549
	Spodoptera litura NPV	SpliNPV	NP_258310
Granulovirus			
	Cydia pomonella GV	CpGV	NP-148794
	Xestia c-nigrum GV	XcGV	AAF05217
Chlorovirus			
	Chlorella virus PBCV-1	PBCV-1	T17756
Cellular organisms			
Bacteria			
Proteobacteria			
Gammaproteobacteria			
Vibrionaceae			
	Listtonella anguillarm	*L. anguillarm*	BAA78114
	Vibrio alginolyticus	*V. alginolyticus*	BAB21607
	Vibrio harveyi	*V. harveyi*	AAK11576
Alteromonadaceae			
	Pseudoalteromonas sp. S9	P.sp.S9	AAC79667
	Shewanella baltica	*S. baltica*	AAF21124
Aeromonas			
	Aeromonas punctata	*A. punctata*	CAD58828
Enterobactriaceae			
	Serratia plymuthica	*S. plymuthica*	CAD32933
	Serratia marcescens	*S. marcescens*	A25090
	Sodalis glossinidius	*S. glossinidius*	CAA72201
	Xenorhabdus nematophila	*X. nematophila*	CAC38398
Burkholderiales			
	Burkholderia cepacia	*B. cepacia*	AAK72610
Firmicutes			
	Bacillus halodurans	*B. halodurans*	BAB04635
	Bacillus thuringiensis	*B. thuringiensis*	AAL71886
	Bacillus cereus	*B. cereus*	BAB16891

Table 1 (*Continued*)

Taxonomy	Name/source	Abbreviation	Accession
Actinomycetales			
	Cellulomonas uda	*C. uda*	AAG27061
	Streptomyces coelicolor A3(2)	*S. coelicolor*	CAB61662
Eukaryota			
Magnoliophta			
	Arabidopsis thaliana	*A. thaloana*	T04756
Fungi			
Ascomycota			
	Ajellomyces capsulatus	*A. capsulatus*	AAF80370
	Coniothyrium minitans	*C. minitans*	AAG00504
	Grifola umbellata	*G. umbellata*	AAO42981
	Hypocrea rufa	*H. rufa*	AAF19616
	Hypocrea virens	*H. virens*	AAL78812
	Trichoderma asperellum	*T. asperellum*	AAF19624
Bilateria			
Pseudocoelomata			
	Nematoda microfilarial chitinase	Nematoda	AAC47324
Coelomata			
Deuterostomi (vertebrata)			
	Bos taurus	*B. taurus*	NP_777124
	Danio rerio	*D. rerio*	AAH46004
	Homo sapiens (glycoprotein 39)	*H. sapiens*1	AAH38354
	Homo sapiens (chondrocyte protein 39)	*H. sapiens*2	NP_003991
	Mus musculus	M. mus	AAG60018
Protostomia (insecta)			
	Araneus ventricocus	*A. ventricocus*	AAN39100
	Braconid wasp (*Chelonus sp.*)	wasp	A53918
	Bombyx mori	*B. mori*	BAB20017
	Choristoneura fumiferana	*C. fumiferana*	AAM43792
	Drosophila melanogaster	*D. melanogaster*	AAF45396
	Glossina morsitans	*G. morsitans*	AAL65401
	Spodoptera litura	*S. litura*	BAB12678

number of sequences and to exclude identical sequences, 48 chitinase sequences representing distinct clades were selected from this tree (Table 1). An NJ and most parsimony (MP) tree was then inferred and tree topology was analyzed using bootstrapping (1000 replicates). The chitinase of *A. thaliana* (e value $= 6e^{-7}$) was used as an outgroup in order to root the tree. The analysis summarized in Fig. 4 show that the chitinase sequences were clustered into two well supported subgroups corresponding to animal (insecta and vertebrata) and non-animal chitinases (fungi, bacteria, viruses). The baculovirus chitinases clustered within a monophyletic clade, which was well supported by bootstrapping (NJ: 81% and MP: 93%). The topology within the baculovirus clade was consistent with the most parsimonious tree generated with baculovirus chitinases only (data not shown). The baculoviral chitinases were clustered within the group of γ-proteobacteria (bootstrap support NJ: 100% and MP: 100%). The chitinases of *S. marcescens*, *B. cepia*, *S. plymuthica* and *A. punctata* appeared to be more closely related to baculovirus chitinases than to other γ-protebacteria, indicating that these bacterial and the baculovirus chitinases had a common origin.

3.4. Analysis of secondary structures and amino acid sequences

Seven chitinases (HaSNPV, AcMNPV, CpGV, *S. marcescens*, P.sp.S9, *C. minitans* and *B. mori*) representing distinct clades were selected to analyze the relatedness of their secondary structure. The deduced amino acid sequences were aligned and their predicted secondary structures were introduced into the alignment (Fig. 5). A putative secretary signal sequence composed of an alpha helix and iscommon to chitinases, was found at the N-terminus and denoted as region I (Ayres et al., 1994; Thomas et al., 1998). A putative ChiN or fibronectin type III domain (Perrakis et al., 1994, 1997) composed of beta-sheet rich structure was identified and denoted as region II. The family 18 chitin hydrolytic domain was denoted as region III. The high similarity of the secondary structures of the three baculovirus chitinases (HaSNPV, AcMNPV and CpGV) and two bacterial chitinases (*S. marcescens* and P.sp.S9) is note worthy (Fig. 5). In comparison with the secondary structures of γ-proteobacterial and baculoviral chitinases, the insect chitinases (*B. mori*) and fungal chitinases (*C. minitans*) lack region II (Fig. 5). However, significant high

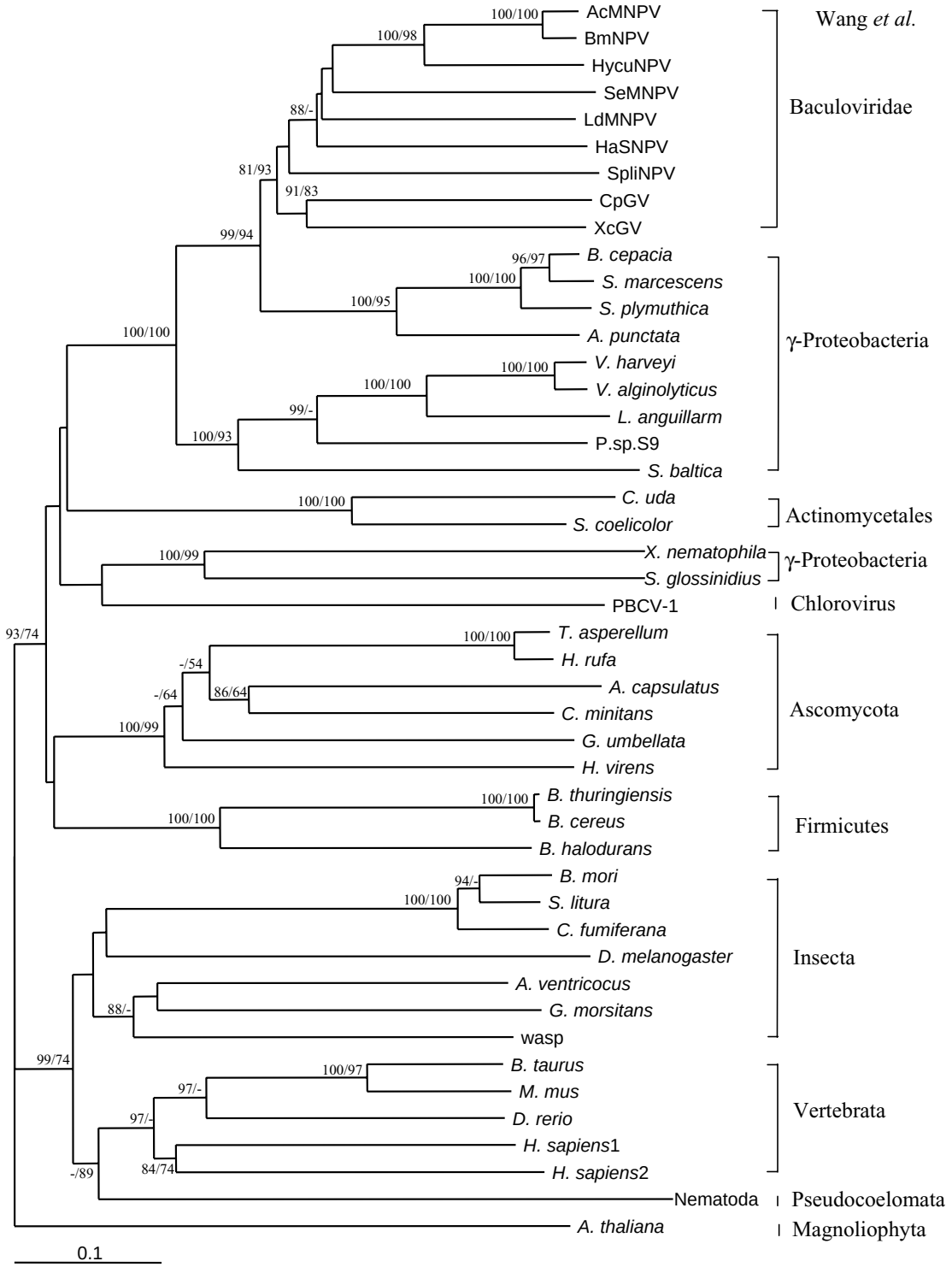

Fig. 4. A neighbor-joining tree derived from 48 chitinases representing various families of bacteria, fungi, nematodes, insects, mammals and viruses. Bootstrap values (1000 replicates, nodes supported with more than 50%) are given for neighbor-joining and most parsimony analyses on the branch lines (NJ/MP). The bootstrap frequency lower than 50 per cent in MP or lower than 80 per cent in NJ is indicated as dashes ('-') and not shown.

secondary structure similarity among these seven chitinases was found in region III, the family 18 hydrolytic active domain. This indicated that they could have similar catalytic function and substrate. Conserved amino acid residues among prokaryotic and baculoviral chitinases (sequences with e value higher than $1e^{-27}$) were explored using Bioedit software. Three amino acid residues, Asp309, Asp312 and Glu316, which constituted the family 18 active site and were demonstrated to play a central roles in the hydrolysis of chitin (Papanikolau et al., 2001), were highly conserved.

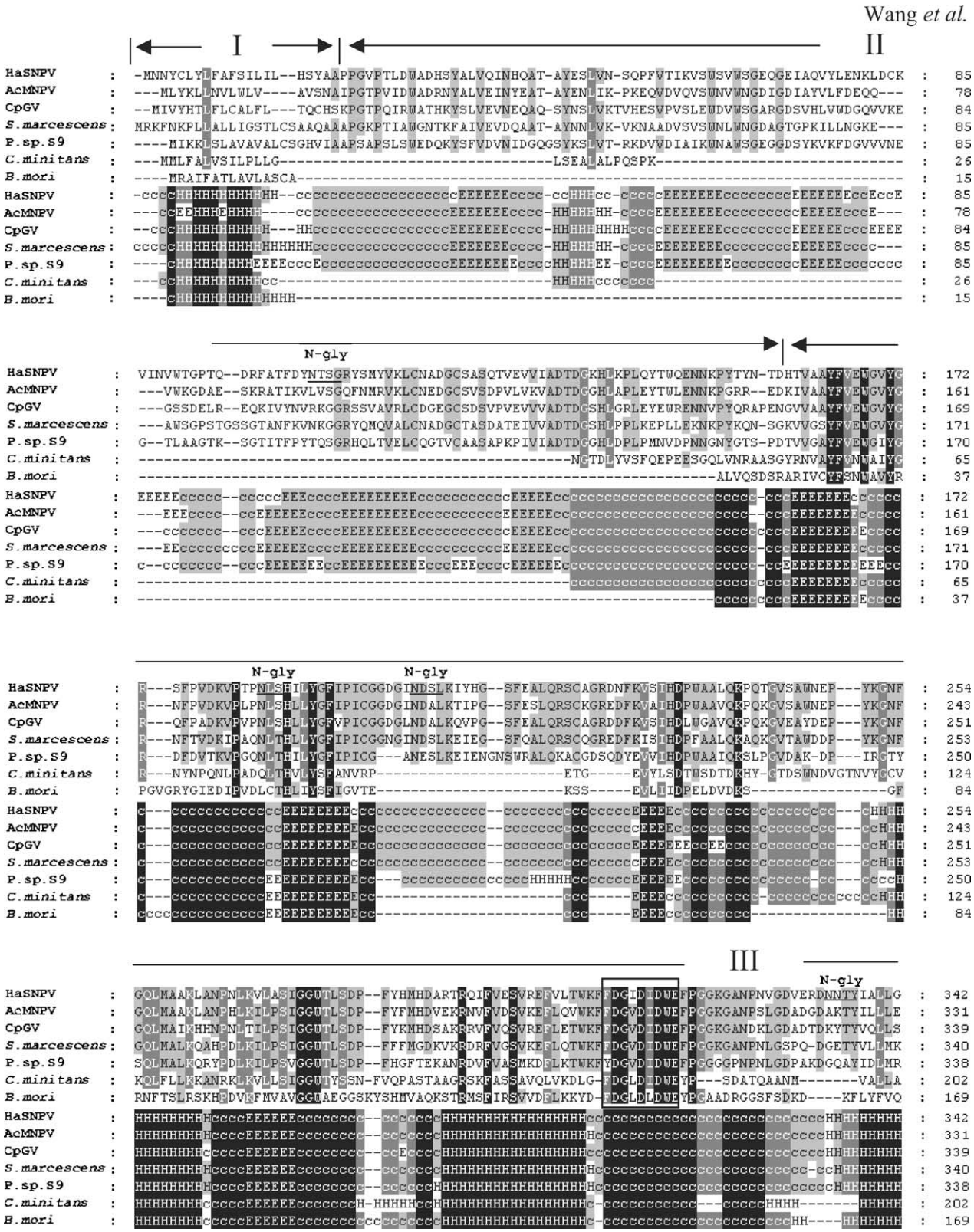

Fig. 5. Alignment of the amino acid sequences and the predicted secondary structures of nine selected chitinases. The alignment was refined using GeneDoc software (version 1.1.004). Three shading levels were set: black for 100% identity, dark gray for 80% identity and light gray for 60%. Regions were indicated with Roman numbers. The family 18 chitin hydrolytic site was boxed. The six *N*-glycosylation sites found in HaSNPV chitinase are indicated by line and marked as *N*-gly. The secondary structure units alpha-helixes (H), beta-strands (E) and coils (c) were represented below the sequences. The gaps are indicated as dashes ('-').

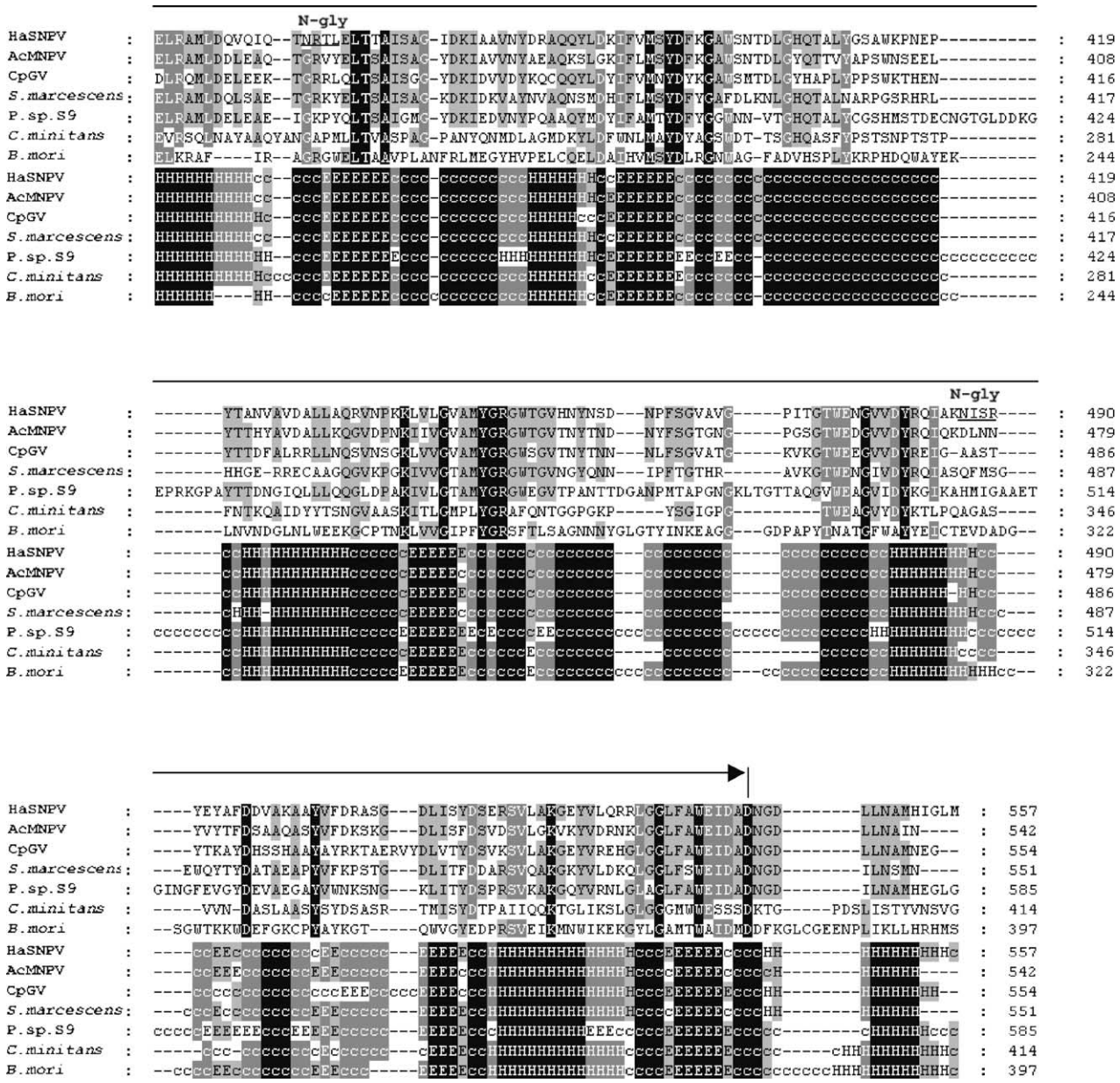

Fig. 5. (*Continued*).

Table 2
Pairwise amino acid sequences and secondary structure identity (%) among seven chitinases

	HaSNPV	AcMNPV	CpGV	*S. marcescens*	P.sp.S9	*C.minitans*	*B. mori*
HaSNPV	–	92	90	92	77	64	59
AcMNPV	63	–	90	94	78	63	58
CpGV	56	59	–	89	76	63	57
S. marcescens	53	58	50	–	77	62	58
P.sp.S9	44	44	43	45	–	56	55
C. minitans	21	21	20	19	19	–	74
B. mori	15	15	16	16	16	19	–

The percentage amino acid identity was shown below the diagonal and the percentage identity of predicted secondary structure was shown above the diagonal.

The similarity of the chitinases was evaluated by calculating their identities of amino acid positions and the secondary structures (Table 2). This confirmed the high similarity between baculoviral and γ-proteobacterial chitinases.

4. Discussion

H. armigera is the natural host of HaSNPV and infected larvae liquefy shortly after death suggesting an active chitinase gene. This study elucidates the transcriptional and expression properties of the HaSNPV chitinase gene. The data confirm earlier findings that chitinase is the product of a late baculovirus gene (Hawtin et al., 1995). The predicted amino acid sequence of HaSNPV chitinase has 54–91% identity to other baculovirus chitinases (data not shown). The presence of the motif HNEL at the C-terminus suggested that the chitinase might accumulates within the ER before released (Thomas et al., 1998). Although the mechanism for the chitinase secretion from the ER to the cytoplasm is still unclear, there is evidence that this motif is important in the retention and stability of the enzyme within the ER vesicles (Frigerio et al., 2001; Saville et al., 2002; Wandelt et al., 1992).

Recent phylogenetic analyses to elucidate the relationship between baculoviral and bacterial chitinases were not always in agreement. Hawtin et al. (1995) noted that the AcMNPV chitinase was closely related to *S. marcescens* chitinase and suggested a horizontal transfer of this gene from bacteria to baculovirses. In contrast, this finding could not be corroborated by Kang et al. (1998) who used a slightly larger data set. We have performed a comprehensive phylogenetic analyses starting with nearly 500 chitinase sequences deposited in GenBank. From the initial analysis, 48 chitinase sequences which represent distant clades, were selected, realigned and subjected to NJ and MP analyses (Fig. 4). Both of the independent methods clearly indicated a monophyletic relationship of the baculovirus chitinases to the chitinases of a sub-branch of γ-proteobacteria, including that of *S. marcescens*, *B. cepacia*, *S. plymuthica* and *A. punctata*. This conclusion was well supported by bootstrap analyses. This relationship strongly supports the hypothesis that baculovirus chitinase is most likely to be of a bacterial origin and was acquired by gene transfer. The separation distance of baculovirus chitinases from those of insects leaves only a minimal likelihood to the alternative assumption of horizontal transfer from the host insect to baculoviruses.

It is well documented that chitinase gene phylogeny only partially follows the 16S rRNA gene phylogeny, which is the base of the current bacterial phylogeny. Lateral transfer of chitinases among bacteria is assumed to be the reason for this deviation (Cottrell et al., 2000). Although the deviation in chitinase gene phylogeny from 16S rRNA gene phylogeny impairs the interpretation of chitinase phylogeny trees, our conclusion that baculovirus chitinases most likely derived from *Serratia*-like γ-proteobacteria is not hampered by observation.

Acknowledgements

H. Wang is a recipient of a DAAD Sandwich Ph.D. fellowship. This research was supported partly by NSFC (39980001, 30025003 and 30070034), Chinese Academy of Sciences (Kscx2-1-02, Kscx2-SW-301-09), MOST (J00-A-003) and the 863 projects of China (101-06-10-01, 2001AA214031).

References

Altschul, S.F., Gish, W., Miller, W., Meyers, E.W., Lipman, D.J., 1992. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Ausubel, F., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., 1996. Current Protocols in Molecular Biology, vol 3. John Wiley & Sons Inc.

Ayres, M.D., Howard, S.C., Kuzio, J., Lopez Ferber, M., Possee, R.D., 1994. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. Virology 202, 586–605.

Bairoch, A., Bucher, P., Hofmann, K., 1997. The PROSITE database, its status in 1997. Nucl. Acids Res. 25, 217–221.

Chen, X., Hu, Z., Jehle, J.A., Zhang, Y., Vlak, J.M., 1997. Analysis of the ecdysteroid UDP-glucosyltransferase gene of *Heliothis armigera* single-nucleocapsid baculovirus. Virus Genes 15, 219–225.

Chen, X.W., Ijkel, W.F.J., Dominy, C., Zanotto, P.M.D., Hashimoto, Y., Faktor, O., Hayakawa, T., Wang, C.H., Prekumar, A., Mathavan, S., Krell, P.J., Hu, Z.H., Vlak, J.M., 1999. Identification, sequence analysis and phylogeny of the *lef*-2 gene of *Helicoverpa armigera* single nucleocapsid baculovirus. Virus Res. 65, 21–32.

Chen, X.W., Li, M., Sun, X.L., Arif, B.M., Hu, Z.H., Vlak, J.M., 2000. Genomic organization of *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus. Arch. Virol. 145, 2539–2555.

Chen, X.W., Ijkel, W.F.J., Tarchini, R., Sun, X.L., Sandbrink, H., Wang, H.L., Peters, S., Zuidema, D., Lankhorst, R.K., Vlak, J.M., Hu, Z.H., 2001. The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus genome. J. Gen. Virol. 82, 241–257.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Biochemistry 162, 156–159.

Church, G.M., Gilbert, W., 1984. Genomic sequencing. Proc. Natl. Acad. Sci. U.S.A. 81, 1991–1995.

Cottrell, M.T., Wood, D.N., Yu, L., Kirchman, D.L., 2000. Selected chitinase genes in cultured and uncultured marine bacteria in the α- and γ-subclasses of the proteobacteria. Appl. Environ. Microbiol. 66, 1195–1201.

Fang, M., Wang, H., Wang, H., Yuan, L., Chen, X., Vlak, J.M., Hu, Z., 2003. The open reading frame 94 of *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus encodes a novel occlusion-derived virion protein: ODV-EC43. J. Gen. Virol. 84, 3021–3027.

Fitt, G.P., 1989. The ecology of *Heliothis* species in relation to agroecosystems. Ann. Rev. Entomol. 34, 17–52.

Frigerio, L., Pastres, A., Prada, A., Vitale, A., 2001. Influence of KDEL on the fate of trimeric or assembly-defective phaseolin: selective use of an alternative route to vacuoles. Plant Cell 13, 1109–1126.

Fukamizo, T., 2000. Chitinolytic enzymes: catalysis, substrate binding and their application. Curr. Protein Pept. Sci. 1, 105–124.

Hashimoto, Y., Hayakawa, T., Ueno, Y., Fujita, T., Sano, Y., Matsumoto, T., 2000. Sequence analysis of the *Plutella xylostella* granulovirus genome. Virology 275, 358–372.

Hawtin, R.E., Arnold, K., Ayres, M.D., Zanotto, P.M.A., Howard, S.C., Gooday, G.W., Chappell, L.C., Kitts, P.A., King, L.A., Possee, R.D., 1995. Identification and preliminary characterization of a chitinase gene in the *Autographa californica* nuclear polyhedrosis virus genome. Virology 212, 673–685.

Hawtin, R.E., Zarkowska, T., Arnold, K., Thomas, C.J., Gooday, G.W., King, L.A., Kuzio, J.A., Possee, R.D., 1997. Liquefaction of *Autographa californica* nucleopolyhedrovirus-infected insects is dependent on the integrity of virus-encoded chitinase and cathepsin genes. Virology 238, 243–253.

Henrissat, B., 1999. Classification of chitinases modules. In: Jolles, P., Muzarelli, R.A.A. (Eds.), Chitin and Chitinases. Birkhäuser Verlag, Basel.

Hom, L.G., Volkman, L.E., 2000. *Autographa californica* M nucleopolyhedrovirus chiA is required for processing of V-CATH. Virology 277, 178–183.

Kang, W., Tristem, M., Maeda, S., Crook, N.E., O'Reilly, D.R., 1998. Identification and characterization of the *Cydia pomonella* granulovirus cathepsin and chitinase genes. J. Gen. Virol. 79, 2283–2292.

Koga, D., Mitsutomi, M., Kono, M., Matsumiya, M., 1999. Biochemistry of chitinases. EXS 87, 111–123.

Lewis, M.J., Pelham, H.R., 1990. A human homologue of the yeast HDEL receptor. Nature 348, 162–163.

Long, G., Chen, X., Peters, D., Vlak, J.M., Hu, Z., 2003. Open reading frame 122 of *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus encodes a novel structural protein of occlusion-derived virions. J. Gen. Virol. 84, 115–121.

Moscardi, F., 1999. Assessment of the application of baculoviruses for control of Lepidoptera. Ann. Rev. Entomol. 44, 257–289.

Ohkawa, T., Majima, K., Maeda, S., 1994. A cysteine protease encoded by the baculovirus *Bombyx mori* nuclear polyhedrosis virus. J. Virol. 68, 6619–6625.

O'Reilly, D.K., Miller, L.K., Luckow, V.A., 1992. Baculovirus Expression Vectors. A Laboratory Manual. W.H. Freeman and Company, New York, pp. 129–132.

Papanikolau, Y., Prag, G., Tavlas, G., Vorgias, C.E., Oppenheim, A.B., Petratos, K., 2001. High resolution structural analyses of mutant chitinase A complexes with substrates provide new insight into the mechanism of catalysis. Biochemistry 40, 11338–11343.

Pelham, H.R., 1990. The retention signal for soluble proteins of the endoplasmic reticulum. Trends Biochem. Sci. 15, 483–486.

Peng, H.Y., Li, X., Zhang, S.M., Arif, B.M., Chen, X., Hu, Z., 1998. Localization and cloning of the chitinase gene of *Heliothis armigera* single nucleocapsid nucleopolyhedrovirus. Virol. Sin. 13, 139–143.

Perrakis, A., Tews, I., Dauter, Z., Oppenheim, A.B., Chet, I., Wilson, K.S., Vorgias, C.E., 1994. Crystal structure of a bacterial chitinase at 2.3 Å resolution. Structure 2, 1169–1180.

Perrakis, A., Ouzounis, C., Wilson, K.S., 1997. Evolution of immunoglobulin-like modules in chitinases: their structural flexibility and functional implications. Fold Des. 2, 291–294.

Raghaven, N., Freedmann, D.O., Fitzgerald, P.C., Unnasch, T.R., Otteson, E.A., Nutman, T.B., 1994. Cloning and characterization of a potentially protective chitinase-like recombinant antigen from *Wuchereria bancrofti*. Infection and Immunity 62, 1901–1908.

Rost, B., Sander, C., 1993. Prediction of protein structure at better than 70% accuracy. J. Mol. Biol. 232, 584–599.

Rost, B., Sander, C., 1994. Combining evolutionary information and neural networks to predict protein secondary structure. Proteins 19, 55–72.

Sanger, F., Nichlen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. 74, 5463–5467.

Saville, G.P., Thomas, C.J., Possee, R.D., King, L.A., 2002. Partial redistribution of the *Autographa californica* nucleopolyhedrovirus chitinase in virus-infected cells accompanies mutation of the carboxy-terminal KDEL ER-retention motif. J. Gen. Virol. 83, 685–694.

Shinshi, H., Neuhas, J.M., Ryals, J., Meins, F.J., 1990. Structure of a tobacco endochitinase gene: evidence that different chitinase genes can arise by transposition of sequences encoding a cysteine-rich domain. Plant Mol. Biol. 14, 357–368.

Sun, X., Zhang, G., 1994. A comparison of four wild isolates of *Heliothis nuclear* polyhedrosis virus. Virol. Sin. 9, 309–318.

Thomas, C.J., Brown, H.L., Hawes, C.R., Lee, B.Y., Min, M.K., King, L.A., Possee, R.D., 1998. Localization of a baculovirus-induced chitinase in the insect cell endoplasmic reticulum. J. Virol. 72, 10207–10212.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl. Acids Res. 25, 4876–4882.

Volkman, L.E., Keddie, B.A., 1990. Nuclear polyhedrosis virus pathogenesis. Semin. Virol. 1, 249–256.

Wandelt, C.I., Khan, M.R., Craig, S., Schroeder, H.E., Spencer, D., Higgins, T.J., 1992. Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. Plant J. 2, 181–192.

Wang, H.L., Chen, X.W., Wang, H.Z., Arif, B.M., Vlak, J.M., Hu, Z.H., 2001. Nucleotide sequence and transcriptional analysis of a putative basic DNA-binding protein of *Helicoverpa armigera* nucleopolyhedrovirus. Virus Genes 22, 113–120.

Wu, D., Wang, H., Deng, F., Peng, H., Chen, X., Hu, Z., 2002. Molecular cloning and expression of the chitinase gene from *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus in *E. coli*. Virol. Sin. 17, 331–335.

Zhang, G., Zhang, Y., Ge, L., Shan, Z., 1981. The production and application of the nuclear polyhedrosis virus of *Heliothis armigera* (HUBNER) in biological control. Acta Phytophylacica Sin. 8, 235–240.