Virus Genes
DOI 10.1007/s11262-007-0150-8

ORIGINAL ARTICLE

Functional and structural characterisation of *Ag*MNPV *ie1*

Marcos Fabián Bilen · Marcela Gabriela Pilloff · Mariano Nicolás Belaich · Vanina Gabriela Da Ros · Julio Carlyle Rodrigues · Bergmann Morais Ribeiro · Víctor Romanowski · Mario Enrique Lozano · Pablo Daniel Ghiringhelli

Received: 30 August 2006 / Accepted: 20 November 2006

Abstract We have located and cloned the *Anticarsia gemmatalis* multicapsid nucleopolyhedrovirus isolate 2D (*Ag*MNPV-2D) genomic DNA fragment containing the *immediate early 1* ORF and its flanking regions. Computer assisted analysis of the complete *ie1 locus* nucleotide sequence information was used to locate regulatory signals in the upstream region and conserved nucleotide and amino acid sequences. Comparative studies led to the identification of several characteristic protein motifs and to the conclusion that *Ag*MNPV-2D is more closely related to *Choristoneura fumiferana* defective NPV than to other Group I nucleopolyhedrovirus. We have also shown that the *Ag*MNPV IE1 protein was able to transactivate an early *Autographa californica* MNPV promoter and its own promoter in transient expression assays. In order to investigate the biological functionality of the *ie1* promoter, the *ie1* upstream activating region (UAR) was molecularly dissected and cloned upstream of the *E. coli lac*Z ORF. The results obtained, after transfection of UFL-AG-286 insect cells, leading us to find that the –492 and –357 versions contains sequence motifs important for the level of the *lac*Z reporter gene expression.

Keywords Baculovirus · *Anticarsia gemmatalis* · *Ag*MNPV · *ie1* · Immediate early genes

Introduction

Baculoviruses, a diverse family of viruses with large double-stranded, circular, covalently closed, DNA genomes, are pathogenic for arthropods, particularly insects of the order *Lepidoptera* [1]. They have attracted wide attention not only for their use as vectors for the expression of recombinant proteins in insect cells but also for their potential use as biological control agents, since they are the

Electronic supplementary material The online version of this article (doi: 10.1007/s11262-007-0150-8) contains supplementary material, which is available to authorized users.

The GenBank accession number of the sequence reported in this paper is AF368905.

M. F. Bilen · M. G. Pilloff · M. N. Belaich · V. G. Da Ros · V. Romanowski · M. E. Lozano · P. D. Ghiringhelli (✉)
Laboratorio de Ingeniería Genética y Biología Celular y Molecular (LIGBCM), Departamento de Ciencia y Tecnología, Centro de Estudios e Investigaciones, Universidad Nacional de Quilmes, Roque Saenz Peña 352, 1876 Bernal, Buenos Aires, Argentina
e-mail: pdg@unq.edu.ar

J. C. Rodrigues · B. M. Ribeiro
Departamento de Biología Celular, Laboratorio de Virologia e Microscopia Eletrônica, Universidade de Brasilia-UnB, Brasilia, DF CEP 70910-900, Brazil

B. M. Ribeiro
e-mail: bergmann@unb.br

V. Romanowski
Instituto de Bioquímica y Biología Molecular (IBBM), Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Calles 49 y 115, 1900 La Plata, Buenos Aires, Argentina
e-mail: victor@biol.unlp.edu.ar

Present Address:
V. G. Da Ros
Instituto de Biología y Medicina Experimental (IBYME), Buenos Aires, Argentina

most prominent group of viruses known to infect insect populations [2].

To date, 18 genomes of nucleopolyhedrovirus infecting *Lepidoptera* have been completely sequenced; eight corresponding to group I NPVs and ten corresponding to group II NPVs (Table 1).

Productive baculovirus infection requires the proper expression of early and late genes that are scattered throughout the genome. Current evidence indicates that baculovirus early gene expression is dependent upon multiple *cis*-acting regulatory elements and the host transcription machinery. These genes are transcribed by host RNA polymerase II and the *cis*-acting elements regulate early promoter activity using host- and virus-encoded factors that are involved in increasing gene transcription [19–23]. One of the most important virus-encoded transcription factors is the transactivator protein IE1 [24].

Autographa californica MNPV (*Ac*MNPV) 67 kDa IE1 is a multifunctional protein which transregulates the expression of several viral genes under the control of early and delayed-early promoters [24–26] and has been implicated as one of the six viral proteins essential for DNA replication in transient replication assays [27]. Wild-type IE1 is synthesised early in infection and is maintained through the late phases [28]. The baculovirus early INR also matches the proposed consensus (A/C/T)CA(G/T)T for arthropod INR (initiator) elements [29]. In group I NPVs, the transcription of *ie1* is initiated from the CA dinucleotide within the CAGT motif [23, 25, 30].

The *Ac*MNPV *ie1* promoter contains a functional TATA element 30 bp upstream from the RNA start site, and a downstream activating region (DAR), which extends from nucleotides +11 to +24 from INR [22]. Additional regulatory elements can be found in the upstream activating region (UAR), such as GC-box, GATA and CGT motifs [31].

The *Ac*MNPV transregulator protein contains an N-terminal domain required for transactivation of viral promoters and a C-terminal domain essential for DNA–protein interaction and protein dimerisation [31, 32].

An isolate of a nucleopolyhedrovirus of *Anticarsia gemmatalis* (*Ag*MNPV) is currently used in Brazil to control the velvetbean caterpillar *A. gemmatalis*, which is a serious defoliator of soybean in Argentina, Brazil and other countries [33]. *Ag*MNPV is today the most widely used biopesticide in more than two million hectares of soybean (Moscardi, personal communication). The improvement of its biopesticide performance will demand a thorough knowledge of the viral genome.

In this article we report the sequence of the *Ag*MNPV-2D *ie1 locus* and a comparative analysis with *ie1 loci* of other nucleopolyhedroviruses. We analysed the promoter

Table 1 Nucleopolyhedrovirus sequences

Genus	Group	Name	Abbreviation	Accession number	Ref.
Nucleopolyhedrovirus	I	*Autographa californica* MNPV	AcMNPV	L22858	[3]
		Bombyx mori NPV	BmNPV	L33180	[4]
		Choristoneura fumiferana DEFECTIVE NPV	CfDEFNPV	AY327402	[5]
		Choristoneura fumiferana MNPV	CfMNPV	NC_004778	[6]
		Epiphyas postvittana NPV	EppoNPV	NC_003083	Hyink et al. (unpublished)
		Hyphantria cunea NPV	HycuNPV	AP009046	[7]
		Orgya pseudotsugata MNPV	OpMNPV	U75930	[8]
		Rachiplusia ou MNPV	RoMNPV	NC_004323	[9]
	II	*Adoxophyes honmai* NPV	AdhoNPV	NC_004690	[10]
		Agrotis segetum NPV	AgseNPV	NC_007921	[11]
		Chrysodeixis chalcites MNPV	ChchMNPV	NC_007151	[12]
		Helicoverpa armigera NPV	HaNPV	NC_002654	[13]
		Helicoverpa zea SNPV	HzSNPV	NC_003349	Chen et al. (unpublished)
		Lymantria dispar MNPV	LdMNPV	AF081810	[14]
		Mamestra configurata NPV	MacoNPV	NC_003529	[15]
		Spodoptera exigua MNPV	SeMNPV	AF169823	[16]
		Spodoptera litura NPV	SpltNPV	NC_003102	[17]
		Trichoplusia ni SNPV	TnSNPV	NC_007383	[18]

MNPV: multiple nuclear polyhedrovirus; NPV: nuclear polyhedrovirus; SNPV: single nuclear polyhedrovirus. The accession numbers are from National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov).

activity of the 5′ untranslated region (5′ UTR) of *Ag*MNPV *ie1 locus* using different length versions of the 5′ UTR cloned upstream of the *E. coli LacZ* ORF, in transient expression assays. In addition, we demonstrated the transactivating activity of IE1 protein on its own promoter and over the heterologous *Ac*MNPV *etl* promoter.

Materials and methods

Cell manipulation and virus

Anticarsia gemmatalis UFL-AG-286 and *Spodoptera frugiperda* IPLB-SF21 (SF-21) cells were grown at 28°C in TC-100 (GIBCO – Life Technologies) containing 10% foetal bovine serum (Bioser, Buenos Aires, Argentina) plus 40 mg/l gentamicin. Viruses were obtained by infection of insect cell monolayers. Supernatants of UFL-AG-286 cells infected with budded virus of *Ag*MNPV-2D [34] were used as *inocula*. The virus titres were assessed by plaque assay on UFL-AG-286 monolayers.

For transfections, a total of 1×0^6 SF-21 or UFL-AG-286 cells were seeded in 60 mm plate and incubated at 28°C for 1 h. Six microlitres (6 µl) of CellFECTIN™ (Gibco Life Technologies) were diluted in 100 µl of serum-free TC100, mixed with the DNA (1–2 µg plasmid DNA diluted in 100 µl of serum-free TC100), incubated at room temperature for 40 min and diluted with 800 µl of serum-free TC100. After removing the culture medium, the DNA/CellFECTIN mix was added to the cell monolayers and incubated at 28°C for 5 h without agitation. Subsequently, 2 ml of TC100 with 10% FBS were added and the cells incubated at 28°C. Another set of transfected cells were infected with *Ag*MNPV-2D (MOI 5) immediately after transfection, incubated for 1 h and then 2 ml of TC100 with 10% FBS were added.

Chloramphenicol-acetyl-transferase expression assay

Twenty-four hours after transfection with pETLCAT [26, 35], the medium was removed, the cells suspended in 1× phosphate buffered saline (PBS) and collected by centrifugation (1 min at 2,500*g*). Cells were then resuspended in 100 µl of 0.25 M Tris (pH 7.5) and disrupted by three cycles of 5 min freezing in liquid nitrogen and thawing for 5 min at 37°C. The cellular debris was centrifuged (10,000*g* for 10 min) and the supernatant transferred to a new tube. The chloramphenicol-acetyl-transferase activity (CAT) assay was performed as described previously. In brief, 2 µl of [^{14}C]-chloramphenicol (Amersham Pharmacia Biotech) were added to 20 µl of the cell free supernatant, diluted in 150 µl of 0.25 M Tris (pH 7,5) plus 4 mM acetyl-CoA, and incubated at 37°C for 1 h. The acetylated and non-acetylated forms of the chloramphenicol were extracted with 1 ml of ethyl acetate and vacuum-dried (DNA SpeedVac-Savant) for 45 min. The dried pellet was suspended in 20 µl of ethyl acetate and spotted on to a Sybron SIL G/UV254 silica plate (Brinkmann); chloramphenicol and its acetylated forms were separated by thin layer chromatography. The silica plate was dried at room temperature and the [^{14}C]-chloramphenicol containing spots were visualised by autoradiography after a 24 h exposure.

Identification, cloning and sequencing of the *ie1* ORF and its flanking regions

The plasmid pBSIE1HC containing a 3.1 kbp *Hind* III-*Cla* I (94.7–96.9 m.u.) fragment of *Ac*MNPV DNA that comprises the *ie1* ORF and promoter [36] was used as a heterologous probe to detect the *Ag*MNPV *ie1* gene. We labelled the gel purified 3.1 kbp insert with alkaline phosphatase (AlkaPhos Direct Labelling, Amersham, UK) and used it as a probe to search for the *Ag*MNPV *ie1* gene. *Ag*MNPV DNA was digested in separate reactions with restriction endonucleases *Hind* III and *Cla* I, electrophoresed in a 0.5% agarose gel and transferred onto a nylon membrane (Hybond N, Amersham, UK).

A hybridisation-positive 4.5 kb *Bgl* II fragment (*Bgl* II F), detected with a chemiluminescent reagent (CDP-Start kit, Amersham, UK), was selected, purified from the gel and cloned into the *Bam*H I site of pBluescript II-SK+ (pBS, *Stratagene*, La Jolla, CA). The clone, designated pBF, was mapped with restriction endonucleases. Based on these results a genomic fragment *Bgl* II/*Xho* I was obtained and cloned in pBS-SK+ to generate pBF1 (see Fig. 1 in supplementary data). The insert of pBF1 was divided into three fragments that were cloned in pZErO-2 (*Invitrogen*, Carlsbad, CA, USA) and designated pZF2, pZF3 and pZF4 (see Fig. 1 in supplementary data). The sequence of the complete *Bgl* II-*Xho* I region cloned in the pBF1 was determined. Plasmid templates for sequencing were prepared by a protocol recommended by Applied Biosystems Inc. (Foster City, CA), which involved SDS-alkaline lysis followed by polyethylene glycol precipitation. Sequencing reactions were performed using the Taq DyeDeoxy Terminator Cycle Sequencing kit (Applied Biosystems Inc., Foster City, CA) and analysed on an ABI 373A Automated DNA Sequencer.

Construction of recombinant plasmids

The plasmid pBF1 was digested completely with *Xba* I and partially with *Bsp*H I. The gel purified 4,840 bp fragment was modified using the Klenow fragment and then religated

to generate pBF1Δp plasmid. The structure was confirmed by restriction analysis and PCR.

In order to eliminate the prokaryotic *gpt* promoter found upstream of the *lac*Z ORF in the pCH110 (Amersham Pharmacia Biotech, UK), a *lac*Z 5′-end fragment from pCH110 was amplified using LacZ5 and LacZ3 primers (Table 2). The agarose-purified PCR fragment was digested with *Kpn* I and *Hind* III and ligated to the pCH110 plasmid previously linearised with *Kpn* I and *Hind* III. The resulting plasmid was designated pCH110Δgpt.

The different reporter plasmids were constructed as follows (outlined in Fig. 6A). The *ie1* downstream region was amplified by PCR (template: pBF1) using the ie1-down and the T7-prom primers (Table 2). The gel purified PCR fragment was digested with *Sal* I and *Xho* I and inserted into the same sites of the pBS-SK+ plasmid. The resulting plasmid was named pAgIE1-Down.

To generate recombinant plasmids containing the *lac*Z ORF under the control of 5′ and 3′ non-coding regions from *ie1*, plasmid pAgIE1-Down was digested with *Sal* I, filled-in using the *Klenow* fragment, and digested with *Hind* III. A 3.7 kbp fragment, from pCH110Δgpt, encompassing the *E. coli lac*Z ORF, was generated by digestion with *BamH* I, fill-in using the *Klenow* fragment, and digestion with *Hind* III, and then ligated into modified pAgIE1-down to generate pAg492Δp.

For the *ie1* upstream region, five different size versions—492, 357, 257, 157, 154—were obtained by PCR amplification of pBF1. The different versions were obtained using the ie1-prom-H primer and M13 reverse, 357, 257, 157 and 154 primers (Table 2). The agarose-purified PCR fragments were digested with *Xba* I and *Hind* III and inserted into the pAg492Δp plasmid, previously linearised with *Xba* I and *Hind* III. The resulting plasmids were designated pAg492ZD, pAg357ZD, pAg257ZD, pAg157ZD and pAg154ZD.

Two reporter plasmids were used to evaluate the IE1-mediated transactivation in transfection experiments: pETLCAT plasmid [26, 35] containing the CAT ORF under the control of the *etl* early promoter of *Ac*MNPV and pETLZD comprising the *E. coli lac*Z ORF under the control of the same *Ac*MNPV promoter.

The IE1 responsive promoter *etl* from *Ac*MNPV was obtained by PCR using etl-Ac-5′ and the etl-Ac-3′ primers (Table 2). The agarose-purified PCR fragment of the *etl* promoter was digested with *Xba* I and *Hind* III and ligated to the pAg492Δp plasmid previously linearised with *Xba* I and *Hind* III to generate the pETLZD plasmid, placing the *lac*Z ORF under the control of *Ac*MNPV *etl* promoter. These plasmids were used in transient expression experiments.

Time course of β-galactosidase expression

UFL-AG cells (1×10^6) seeded in 35 mm Petri dishes were transfected with identical number of molecules (~1×10^{11} molecules) of pAg492ZD, pAg357ZD, pAg257ZD, pAg157ZD, pAg154ZD or pAg492Δp plasmid DNAs using CellFECTIN™ (Gibco Life Technologies), according to the manufacturer's protocol.

Samples were collected at 0, 4, 7, 10, 14, 18, 27 and 34 h post-transfection. Three dishes per time point were used, as well as one dish for a mock-transfected control. Cell monolayers were harvested by scraping them into the supernatant medium and recovered by centrifugation (1 min at 2,500*g*). The pellet was rinsed with PBS and assayed for β-galactosidase expression as described

Table 2 Primers

Name	Sequence[a]	Restriction site	Template	Position
LacZ5	5′-cccaagcttcat**ATG**AGCGAAAAATACATCG-3′	*Hind* III	pCH110	1–19[b]
LacZ3	5′-TTCGCCATTCAGGCTGCGCAACTG-3′	–	pCH110	357–334[b]
ie1-down	5′-ggccgtcgacTTGAAGATGCTAAGCAGTTGT-3′	*Sal* I	pBF1	2138–2158[c]
357	5′-cccgtctagaatcCGCACCACCACAAAACCC-3′	*Xba* I	pBF1	136–163[c]
257	5′-cccgtctagaatcGCCGCGCAGCGCCGATGG-3′	*Xba* I	pBF1	236–253[c]
157	5′-ccgtctagaatcTTTGAATAAATAAACGATAACGCC-3′	*Xba* I	pBF1	336–359[c]
154	5′-cccgtctagaAATAAATAAACGATAACGCCATTGGT-3′	*Xba* I	pBF1	340–365[c]
ie1-prom-H	5′-ccccaagcttGATTGTCGGTGAGCGTTGCGCGT-3′	*Hind* III	pBF1	492–470[c]
5′ ext	5′-cgcgttgaacagcgcgtgac-3′	–	pBF1	564–545[c]
Etl-Ac-5′	5′-gctctagaGATAAAATTGAGCGCAAATATGTGG-3′	*Xba* I	AcMNPV	–
Etl-Ac-3′	5′-cccaagcttTTAGCAGTGATTCTAATTGCAGC-3′	*Hind* III	AcMNPV	–
Poly-dT	5′-ccggatcctctagagcggccgcttttttttttttttttv-3′	*Hind* III, *Xba* I, *Not* I	cDNA	–

[a] The non-ie1 sequences are in lowercase and the restriction site is underlined

[b] *Lac*Z orf relative position. ATG is in bold

[c] pBF1 insert sequence relative position (see Fig. 2A)

previously [37]. The enzyme activity was measured spectrophotometrically after 1 h of incubation at 37°C using *o*-nitrophenylgalactoside (ONPG) as substrate.

Transactivation

Co-transfection with 1 μg of each version of the pAg plasmid (pAg492ZD, pAg357ZD, pAg257ZD, pAg157ZD or pAg154ZD) and 0.5 μg pBF1 plasmid were used to evaluate the self-regulation of *ie1* promoter. Co-transfection with 1 μg of pETLZD and 0.5 μg of pBF1 plasmids, were used to evaluate IE1 transactivate activity. In other assay, the cells transfected with each version of pAg plasmid were post-infected with *Ag*MNPV (MOI 5). In all experiments β-galactosidase activity was measured 24 h post-transfection.

5′-Rapid amplification of cDNA ends (5′-RACE)

Three hours post-infection extraction of mRNA was performed according to protocol of kit poly ATtract (Promega).

cDNA was synthesised with AMV RT using the 5′ ext primer, incubated to 42°C, 1 h. The sample was treated with RNAse H and T1 during 30 min. The cDNA was precipitated and incubated with dATP and terminal transferase during 30 min at 37°C.

PCR with oligo-dT and 5′ ext primers was carried out at 92°C 20 s, 35°C 20 s, 72°C 30 s for 10 cycles and 92°C 20 s, 50°C 20 s, 72°C 30 s for 25 cycles. The PCR amplification product was cloned in pZErO-2 plasmid and sequenced with Sp6 and T7 universal primers.

Computer assisted analyses

The relative similarities were calculated using the ClustalX [38, 39] consensus symbols (*, : and .) as input sequence, in an overlapping windows-based strategy. The sum of assigned values for each residue in each window was divided by the window width and allotted to the central position to generate the plots (Ghiringhelli, unpublished). For nucleic acids calculation, we used a windows width of 11 residues and assigned arbitrary values of +1 for identical (*) and –1 for non-identical residues. For amino acids calculation, we used a windows width of 15 residues and assigned arbitrary values of +1 for identical (*), +0.5 and +0.25 for different degrees of residue conservation [40] and –1 for non-identical residues.

Sequence logos were computed using multiple alignment of all virus sequences included in this study. Logos were computed using Alpro and Makelogo [41] and generated using a WEB server (http://www.bio.cam.ac/seqlogo).

The secondary structures of IE1 proteins were predicted using the JPred2 server (http://jura.ebi.ac.uk:8888, [42, 43]). Briefly, this automated server performs a multiple alignment between the query and sequences in the database. Then, it predicts the secondary structure as a consensus of several independent algorithms; PHD [44], NNSSP [45], DSC [46], PREDATOR [47].

Putative post-translational modification sites were searched using the PROSITE [48] database at the Expasy Proteomics WEB server (http://www.expasy.ch).

Average charge and hydrophobicity profiles were calculated using an overlapping windows-based method as indicated above for the sequence similarity analysis. The sum of assigned values for each residue in each window was divided by the window width and allotted to the central position to generate the plots (Ghiringhelli, unpublished). For charge profiles calculation, we used a windows width of 15 residues and assigned values of: +1 for K and R, +0.5 for H and –1 for D and E. For hydrophobicity profiles calculation, the window width was 7 residues and assigned values of: 1.00 for A, 0.17 for C, –3.00 for D, –2.60 for E, 2.50 for F, 0.67 for G, –1.70 for H, 3.10 for I, –4.60 for K, 2.20 for L, 1.10 for M, –2.70 for N, –0.29 for P, –2.90 for Q, –7.50 for R, –1.10 for S, –0.75 for T, 2.30 for V, 1.50 for W and 0.08 for Y. In order to facilitate the comparison of graphs, gaps were included in each individual graph on the basis of amino acid multiple alignments. The consensus plots were calculated taking the minimum and maximum values of all individual graphs at each position, plotting the average value and assigning half of the difference between maximum and minimum as dispersion.

Results

Transient expression assays

Insect cells (UFL-AG-286 and IPLB-SF-21) were transfected with *Ag*MNPV DNA and a reporter plasmid (pETLCAT). Passarelli and Miller [26], using transient assays with this plasmid, showed that the *Ac*MNPV IE1 protein is necessary and sufficient for the transactivation of this promoter. No activity has been observed when this plasmid was transfected alone into insect cells. We observed the transactivation of the *Ac*MNPV *etl* promoter when we co-transfected the pETLCAT and the *Ag*MNPV genome into IPLB-SF-21 and UFL-AG-286 cells (Fig. 1A) demonstrating that the *Ag*MNPV genome has probably a homologous IE1 protein.

Identification, cloning and sequence of the *ie1* ORF and its flanking regions

We detected the *ie1* gene into the *Bgl* II F genomic fragment of *Ag*MNPV by a Southern blot experiment using a

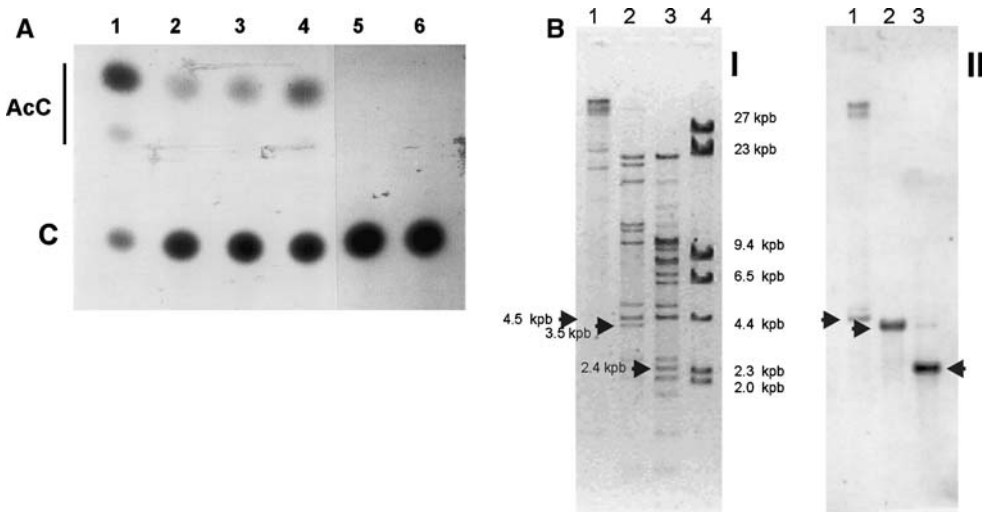

Fig. 1 (**A**) Transient CAT expression assay of insect cells 24 h post-transfection with 2 µg of pETLCAT plasmid and 1 µg of genomic DNA from *Ac*MNPV (lane 1) or *Ag*MNPV (lane 2) in Sf21 cells and AcMNPV (lane 3) or *Ag*MNPV (lane 4) in UFL-AG-286 cells. pETLCAT alone in Sf21 cells (lane 5) and in UFL-AG-286 cells (lane 6). AcC: acetylated chloramphenicol; C: chloramphenicol. (**B**) (I) Digested DNAs resolved in a 0.5% agarose gel. 1. *Ag*MNPV genome digested with *Bgl* II. 2. *Ag*MNPV genome digested with *Xho* I. 3. *Ag*MNPV genome digested with *Bgl* II and *Xho* I. 4. λ phage DNA digested with *Hind* III. (II) Southern blot. The previous gel was transferred to nitrocellulose membrane and hybridised with a heterologous ^{32}P probe derived from the plasmid pBSIE1HC (see Materials and Methods). Lane 1. *Ag*MNPV genome digested with *Bgl* II. Lane 2. *Ag*MNPV genome digested with *Xho* I. Lane 3. *Ag*MNPV genome digested with *Bgl* II and *Xho* I. (**C**) Plasmid containing the *Bgl*II-*Xho*I fragment. *Ag*MNPV genome was digested with *Bgl* II and *Xho* I restriction endonucleases and the resulting fragment of 2,445 bp was cloned in pBS-SK+ to generate pBF1 plasmid. The corresponding subclones pZF2 (*Xba* I/*Hind* III), pZF3 (*Hind* III/*Hinc* II) and pZF4 (*Hinc* II/*Xho* I) were fully sequenced

heterologous probe (Fig. 1B). This 4.5 kbp fragment was cloned into pBS-SK+ to generate pBF. In order to assess the integrity of *ie1 locus* in the recombinant plasmid, we carried out a transient expression experiment through co-transfection with pBF and pETLCAT plasmids. Twenty-four hours post-transfection, the cell extracts were analysed for CAT activity, showing that the insert of pBF plasmid is capable of transactivating the *etl* promoter from *Ac*MNPV (data not shown). The insert of pBF was mapped and a fragment of the 2.4 kpb from a *Bgl* II/*Xho* I digestion was subcloned in pBS-SK plasmid to generate pBF1 (see Fig. 1 in supplementary data). Using the transactivation of the *etl* promoter, we found that the *ie1* open reading frame and its flanking regions was contained in pBF1. The insert of pBF1 was subcloned into the pZErO plasmid (pZF2, pZF3 and pZF4, see Fig. 1 in supplementary data) and sequenced. The pBF1 2,445 bp insert (Fig. 2A) (GB accession no. AF368905), contains a 1,707 bp ORF flanked by 492 bp and 246 bp of 5′ UTR and 3′ UTR regions, respectively. This ORF codes for a predicted 568 amino acids long polypeptide with a calculated molecular weight of 65.7 kDa.

A comparative analysis with all other available nucleopolyhedrovirus (NPVs) *ie1* sequences shows that the *Ag*MNPV *ie1* ORF nucleotide sequence and the deduced amino acid sequence are more similar to those of Group I NPVs than to the corresponding gene of Group II [49] (see Tables 1–3 in supplementary data).

The *Ag*MNPV 5′ UTR contains an INR element (CAGT) located 44 bp upstream from the ATG. 5′ RACE experiments demonstrated that the *ie1* mRNA begins at the G of this INR element (Fig. 2B). The upstream activating region (UAR) has various putative transcriptional regulatory elements, such as several GC-boxes, one GATA motif, five CGT motifs, one TATA-box and one CAAT-like motif (Fig. 2A), similar to those found in other NPVs *ie1* UAR region (Fig. 3). The downstream activating region (DAR) has one CACNG motif similar to that of *Ac*MNPV *ie1* promoter [31].

Sequence analysis of NPV IE1 proteins

A computer-assisted study of the IE1 secondary structure predicted that 43.3% of the residues are involved in the formation of 15 α-helices and 17 β-strands (Fig. 4). An amphipathicity analysis performed by construction of helical wheels of the last two C-terminal α-helices showed that both have a more hydrophobic face (positive values in the helical wheels of Fig. 4, see materials and methods), probably involved in protein–protein interaction [50].

A relative similarity plot of all Group I NPV IE1 proteins (Fig. 5A) revealed that the regions with predicted α-helices or β-strands have the highest degree of sequence similarity. This finding is probably related with structural requirements for the biological protein function.

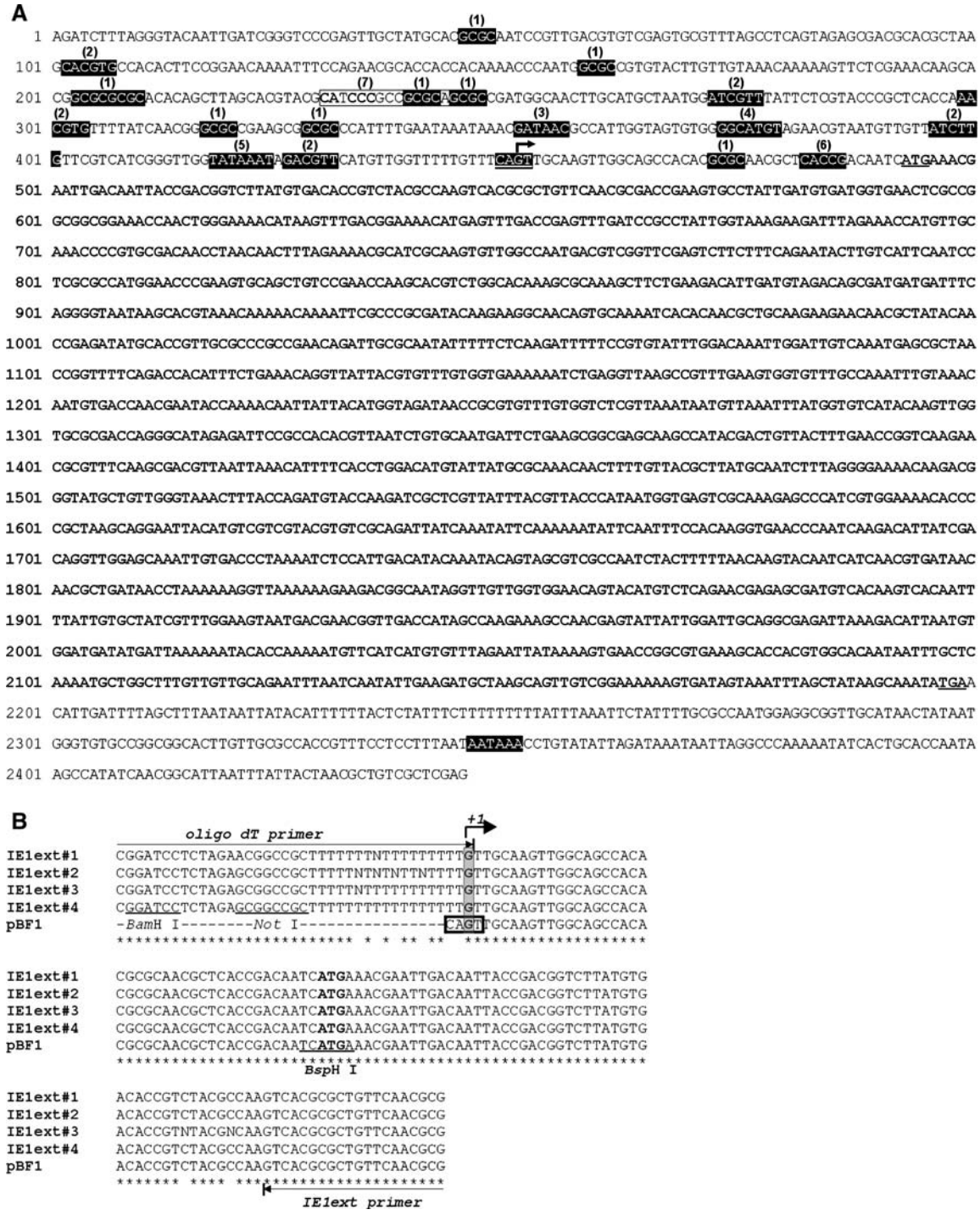

Fig. 2 (**A**) Nucleotide sequence of the *Ag*MNPV *ie1 locus*. The 2,445 nt long fragment inserted in pBF1 comprises a 1,707 nt long ORF (bold characters) and untranslated flanking regions upstream (492 nt) and downstream (246 nt). A consensus INR motif (CAGT) is shaded and underlined, the *arrow* indicates the transcription start point. The ATG start codon and the TGA stop codon for IE1 translation, are underlined. Putative transcriptional regulatory motifs in the UAR and DAR regions are shaded and numbered: (1) GC box, (2) CGT-like motif, (3) GATA motif, (4) CAAT-like motif, (5) TATA-box, (6) CACNG motif (for comparison with other baculoviruses see Fig. 3). Boxed and numbered (7) is indicated a motif that could bind Sp family-like proteins, this is composed by the residues in bold and the loose CG boxes. A putative polyadenylation signal (*shaded*) is located 147 nt downstream the TGA stop codon. (**B**) 5′ RACE-PCR assay. The sequences of four independent RACE clones were aligned with the *ie1* gene sequence obtained from the pBF1 plasmid to show the consensus for the 5′ end mRNA sequence. The transcription start site (+1) is indicated with an arrow, the conserved CAGT motif, present in the pBF1-derived sequence, is boxed and the ATG start codon is indicated in *bold*. *Bam*HI, *Not*I and *Bsp*HI sites are underlined

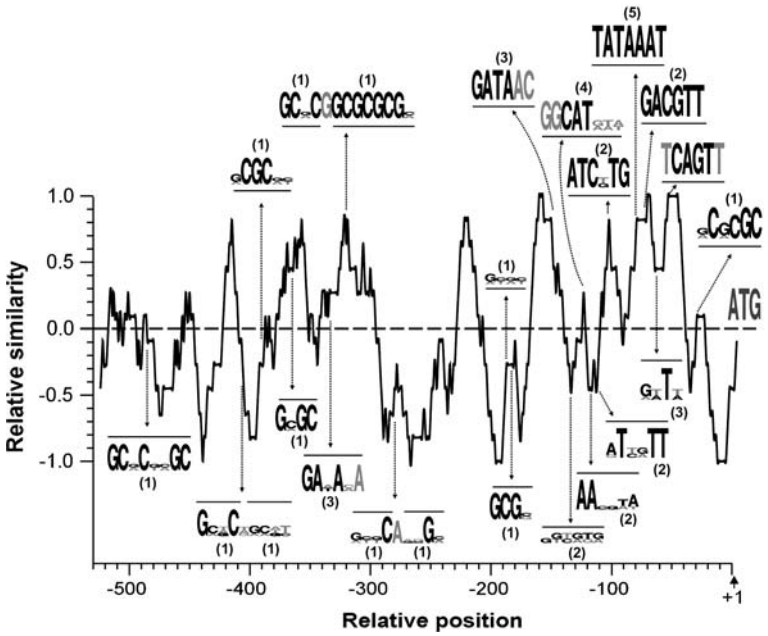

Fig. 3 Relative similarity of *ie1* promoter region and conservation of transcriptional motifs in Group I nucleopolyhedroviruses. The relative similarity was calculated as described in materials and methods. The *x*-axis is numbered relative to ATG start codon (A was considered +1). Sequence conservation of the different transcriptional motifs is depicted as sequence logos [41]. The relative height of each letter within a stack is proportional to their frequencies in all analysed Group I nucleopolyhedrovirus *ie1* promoters. The nucleotides involved in the transcriptional motifs are indicated in black, other residues are indicated in light grey. ATG codon is also indicated (*dark grey*). Motifs are identified by numbers as in Fig. 2A

A comparative analysis of the charge distribution along each sequence of Group I NPVs shows profiles with an overall high degree of similarity. The N-terminal region comprised between residues 8 and 120 has a consistent negative local net charge ranging between –9.5 (*Op*MNPV) and –16.3 (*Bm*NPV) (Fig. 5B). It has been shown that this N-terminal domain of *Ac*MNPV IE1 is required for transactivation of viral promoters [31, 32].

A similar analysis performed on hydrophobicity distribution indicates that the first third of the IE1 proteins has the highest diversity (Fig. 5C). Conversely, the region comprising the α14helix has the highest degree of profile similarity. This C-terminal domain has been identified as essential for DNA-protein complex formation and protein dimerisation in AcMNPV [31, 32].

Expression of *lac*Z driven by different versions of the *Ag*MNPV *ie1* promoter

In order to evaluate the promoter activity of the *ie1* 5′ UTR, a series of recombinant plasmids carrying the *E. coli lac*Z ORF under the control of different length versions of the UTR were used (Fig. 6A). The kinetics and levels of *β*-galactosidase expression from each recombinant plasmid were determined using transient expression assays in UFL-AG 286 cells. The highest level of *β*-galactosidase activity obtained with pAg492ZD, 34 h post-transfection, was nearly 14-fold higher than that achieved with the plasmid pAg154ZD (Fig. 6B). The levels of *β*-galactosidase activity with pAg154ZD were similar to the basal activity obtained with the plasmid pAcDZ1, in which the *Drosophila hsp70* promoter drives the transcription of the *lac*Z ORF, when the heat shock induction was omitted (data not shown). The basal level of expression driven by the *Drosophila hsp*70 promoter is derived from the interaction with the generalised machinery of cellular transcription factors. As expected, the level of ONPG chemical hydrolysis obtained in the negative control (transfection with the plasmid pAg492Δp in which the promoter in front *lac*Z ORF has been deleted) was very low. Furthermore, we analysed the kinetics of gene expression driven by the five promoter versions in transfected cells. The *β*-galactosidase activity was detectable after 4 h post-transfection in all cases (Fig. 6B). Transfection with pAg154ZD plasmid resulted in a *β*-galactosidase activity *plateau* between 20 and 27 h post-transfection. However, in the case of the transfection with pAg492ZD, the enzyme activity continued to rise up to the end of the time course experiment (34 h post-transfection). The last result was completely consistent with the kinetics of *ie1* mRNA synthesis in *Ag*MNPV infected cells when analysed by Northern blot (Fig. 6C).

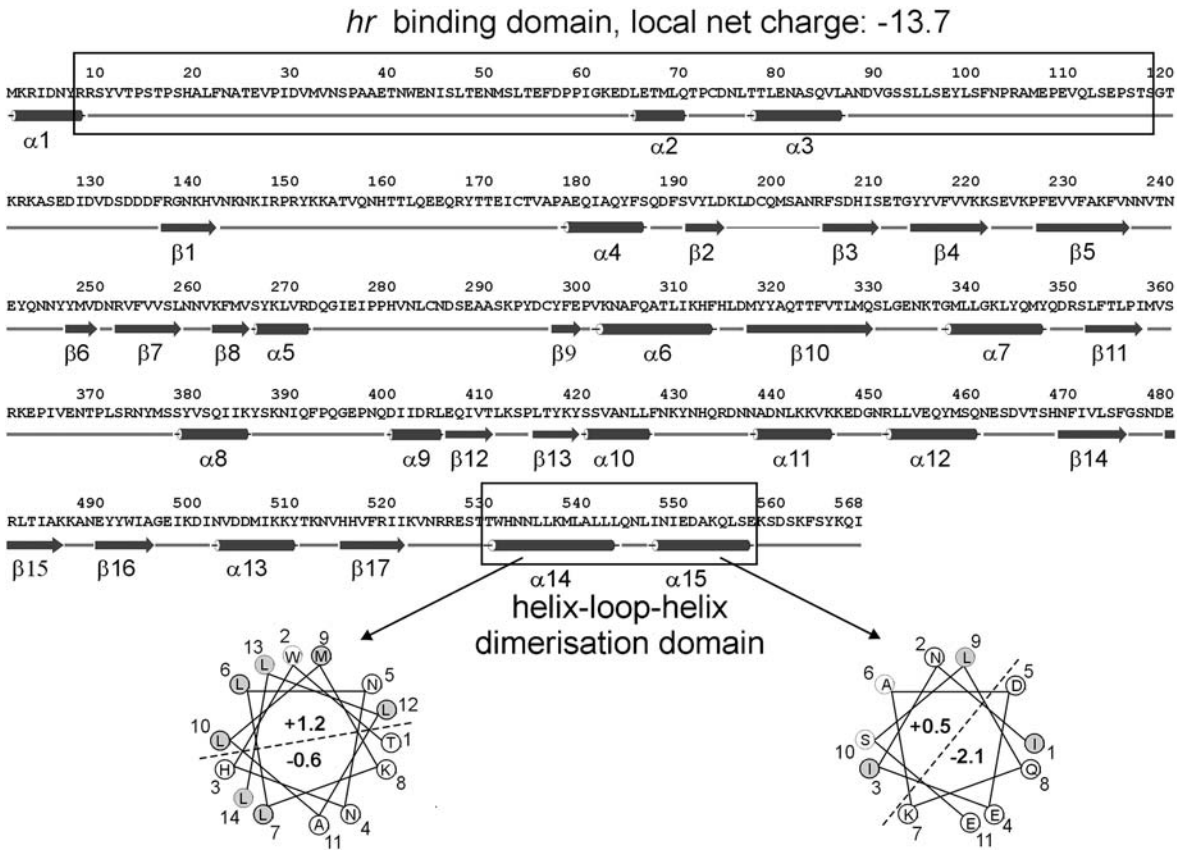

Fig. 4 *Ag*MNPV predicted amino acid sequence and secondary structure prediction. The consensus secondary structure (calculated with the JPred2 server: http://jura.ebi.ac.uk:8888/) was predicted for three alternative states: helix, sheet and loop. Helices are shown as cylinders, sheets as filled arrows and loops as grey lines. The two best-defined domains, N-terminal acidic and dimerisation domains (helix-loop-helix) are boxed. The helical wheel representation of the two last α helices are shown at the bottom. Numbers inside the wheels indicate the average hydrophobicity of each face of the wheel. Positive values correspond to higher hydrophobicity

IE1 regulation

To evaluate the regulatory activity of IE1 on its own promoter, UFL-AG 286 cells were transfected with each version of reporter plasmid used in the previous section, co-transfected with pBF1 plasmid or infected with *Ag*MNPV-2D, and *β*-galactosidase activity was measured 24 hrs post-transfection. Infected samples allow us to evaluate the regulatory activity in a viral gene expression context. The *β*-galactosidase activity increased in the presence of IE1 or after viral infection, in all cases (Fig. 7A, B). We observed that the relative increment of expression was significant when comparing pAg257ZD and pAg357ZD (Fig. 7C), indicating the presence of an important binding region between nucleotides 257 and 357 from the start codon, implicated in the interaction that leads to a higher transcription level.

Trans-regulation

We constructed a recombinant plasmid, pETLZD, comprising the *E. coli lac*Z ORF under the control of the IE1 responsive promoter from *Ac*MNPV *etl* gene, to verify the transactivation by *Ag*MNPV IE1 expressed from the pBF1 plasmid. As shown in Fig. 7D, only the co-transfection experiment performed using pETLZD plus pBF1, yielded levels of *β*-galactosidase higher than the basal activity obtained in untransfected cells.

Discussion

We have identified the presence of the *ie1* gene of *Ag*MNPV in a functional assay followed by cloning and sequencing of its ORF and flanking regions. Furthermore using a 5′ RACE experiment, we have determined the 5′ end of the *ie1* mRNA, which maps at the G of the INR element (Fig. 2B). Current evidence suggests that transcription of baculoviral immediate-early genes is predominantly initiated from a CA dinucleotide into the INR CAGT motif [22, 51]. This RNA start motif appears to be highly conserved among early baculovirus genes. In particular, most article described that early mRNAs begins at

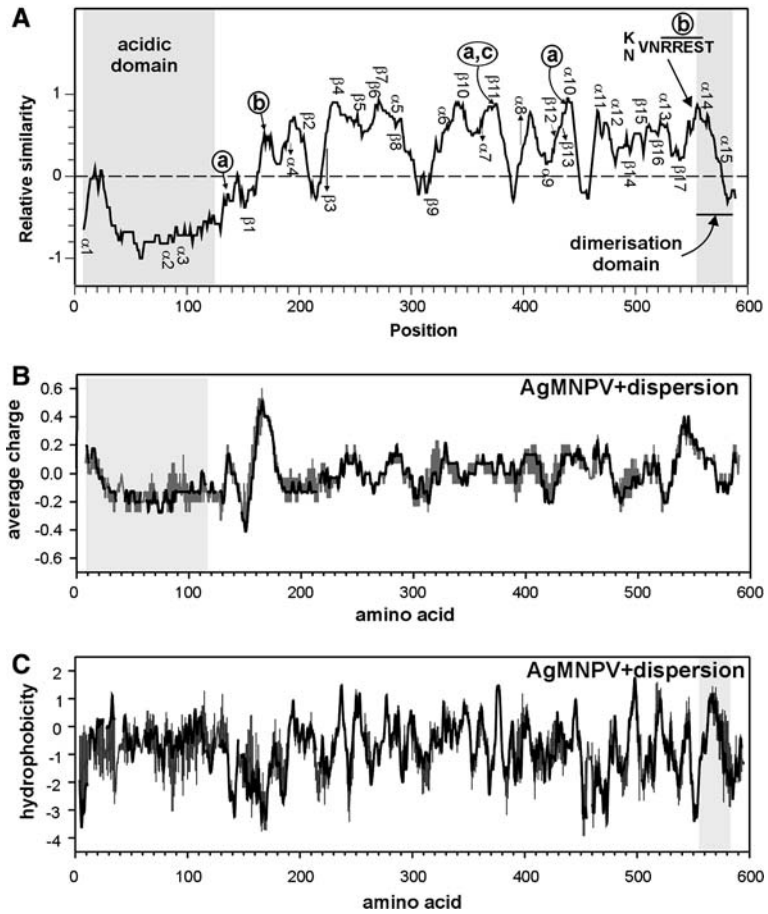

Fig. 5 (**A**) Group I Nucleopolyhedrovirus relative similarity of IE1 proteins. The relative similarities were calculated as described in the "Materials and methods" section. The identified secondary structures are indicated and numbered according to Fig. 4. The *shaded* areas correspond to the acidic and dimerization domains, respectively. Lowercase letters identify putative phosphorylation sites conserved in all sequences in a multiple alignment. (a) Protein kinase C phosphorylation site. (b) cAMP-o cGMP-dependent protein kinase phosphorylation site. (c) Casein kinase II phosphorylation site. (**B**) Average charge profile of different Group I nucleopolyhedroviruses, calculated using a 15-residue-wide window (*black line*). Gaps were included in the individual data on the basis of amino acid multiple alignments. The vertical bars (*dark grey*) show the dispersion at each position. The acidic domain is indicated as a *grey shaded* area. (**C**) Hydrophobicity profiles of different Group I nucleopolyhedroviruses, calculated using a 7-residue-wide window (black line). Gaps were included in the individual graphs in function of an amino acid multiple alignment. The vertical bars (*dark grey*) show the dispersion at each position. The *grey shaded* area corresponds to the dimerisation domain

the A base of the INR sequence (*Ac*MNPV, *Op*MNPV) [19, 23, 51]. However, in one baculovirus from group II (*Ld*MNPV) none of the 5′ ends of different transcripts mapped into an early transcriptional initiator (INR) conserved sequence [52]. In addition, other three baculovirus of group II (*Adho*NPV, *Agse*NPV and *Splt*NPV) lacks of a canonical INR in the –492 promoter region.

Biological activity

It is known that the IE1 protein can modulate the expression of different baculovirus genes, mainly members of the immediate and delayed-early classes [25, 32]. In addition to this, it has been demonstrated that *Ac*MNPV IE1 protein can transactivate cellular promoters augmenting the level of expression in the absence of other viral products [53].

At this point, the mode of action of *Ag*MNPV IE1 and its molecular basis can be inferred only from the functional properties of other IE1 proteins [25, 32, 54, 55]. In this context, we have demonstrated the ability of *Ag*MNPV IE1 protein to transactivate its own promoter and the heterologous *Ac*MNPV IE1-dependent promoters (*etl* promoter).

Sequence analysis of IE1 protein and Acidic domain

Based on the sequence conservation of several baculovirus species [23, 30, 56], we have identified different putative structural features in the IE1 protein reminiscent of DNA binding transcription factors. These are mainly represented by the presence of a helix-loop-helix protein dimerisation domain, near the C-terminus of IE1 [57] and by an acidic domain, similar to that found in some transcription

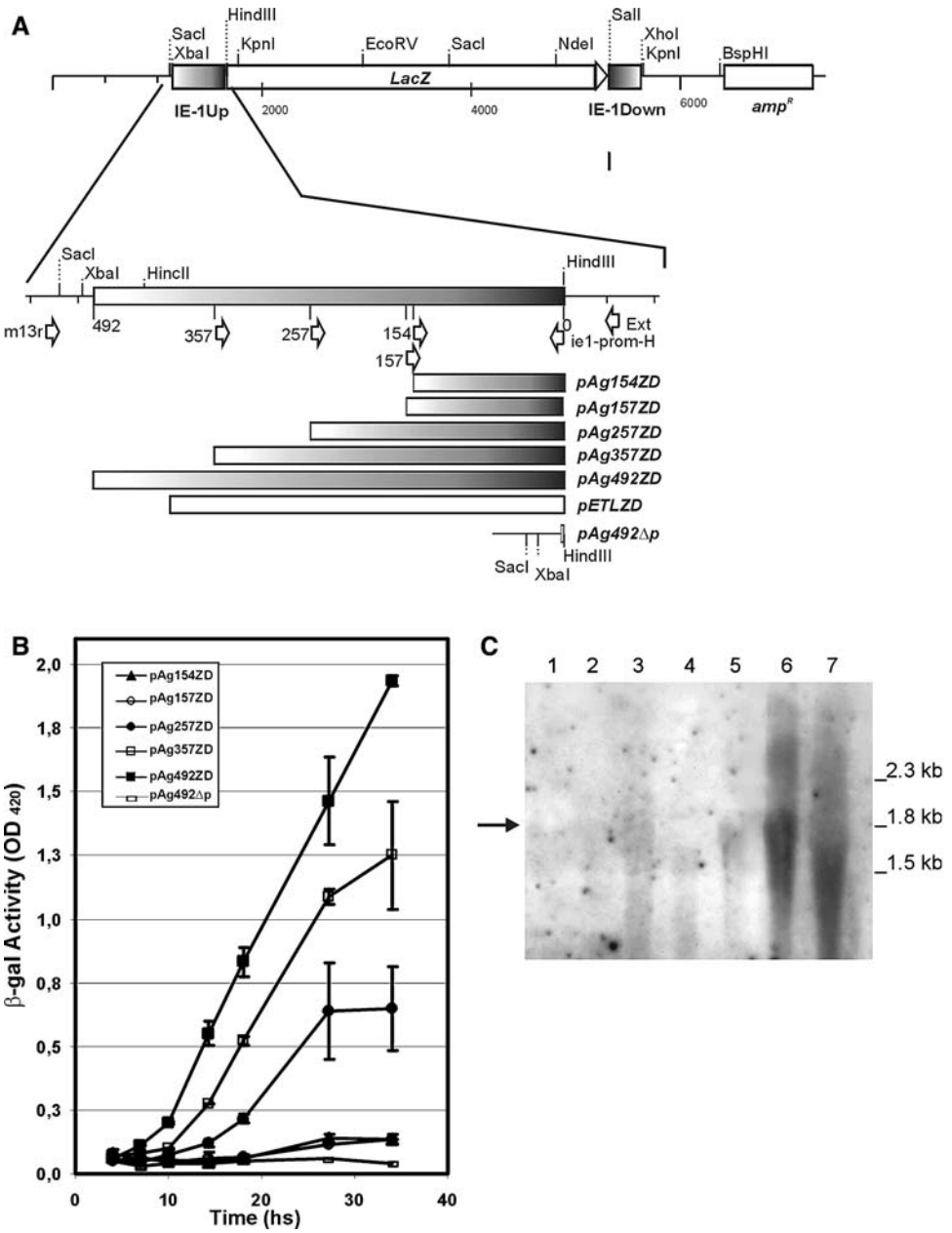

Fig. 6 *Ag*MNPV *ie1* promoter activity. (**A**) Map of recombinant plasmids carrying the *E. coli lac*Z ORF under the control of different versions of *Ag*MNPV *ie1* (492, 357, 257, 157, 154 and 492Δp) promoter or AcMNPV *etl* promoter. Primers used to generate the different promoter constructs are indicated by arrows underneath the scheme of the expanded *ie1* region. (**B**) Time course of *β*-galactosidase expression, driven by all versions of *Ag*MNPV *ie1* promoter. The plasmid pAg492Δp was used as negative control. (**C**) Northern blot assay. Total RNA obtained from infected and non infected UFL-AG 286 cells was analysed by denaturing agarose gel electrophoresis, transferred to a nylon membrane and hybridised with a labelled probe containing *ie1* sequences derived from the pBF1 plasmid (see methods). Negative control (lane 1) contains RNA from uninfected cells. Samples from infected cells were obtained at 1 h (lane 2), 2 h (lane 3), 3 h (lane 4), 5 h (lane 5), 24 h (lane 6) and 48 h (lane 7), post-infection with *Ag*MNPV. *Arrow* highlights the position of *ie1* transcript

regulatory proteins, near their N-terminus [58–61]. The protein dimerisation domain is preceded by a cluster of eight highly conserved residues that include either two or three basic amino acids (K/NVNRREST), and have been proposed to constitute a DNA binding domain [53].

On the other hand, it has been demonstrated that the acidic domain of *Ac*MNPV IE1 and *Op*MNPV IE1 has a crucial role for the transactivation of transcription, i.e. this activity is abolished if the acidic domain is deleted [32, 62]. This domain belongs to a more general family of transcriptional activation domains, since it can be functionally replaced by acidic domains of other viruses, such as acidic domain of *Ac*MNPV IE1 or the acidic domain from the herpes simplex virus transactivator VP16 [63]. Despite the fact that the N-terminal domains of all the IE1 protein sequences available to date have the lowest sequence similarity (Fig. 5A), there is a high similarity in their charge profiles (Fig. 5B), which apparently is its most important feature. Electrophoretic mobility shift assays have shown that the *Ac*MNPV IE1 will bind either directly or indirectly to the *Ac*MNPV enhancer *hr* elements [50, 54]. Other experiments have shown that either the *Op*MNPV IE1 or *Ac*MNPV IE1 could activate the *Ac*MNPV enhancer elements in a sequence-dependent manner indicating that these IE1 proteins are interchangeable [23]. In addition, it has been demonstrated that IE1 can activate transcription in an indirect manner via interaction with other proteins such as cellular transcription

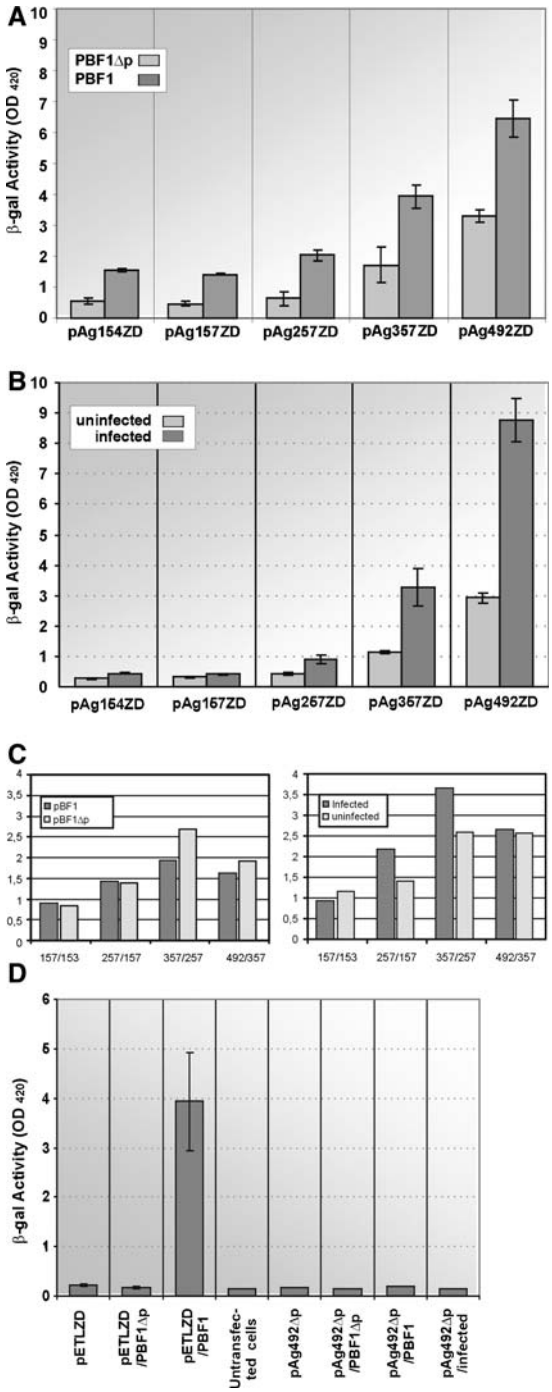

Fig. 7 (**A**) Levels of *β*-galactosidase expression, 24 h after co-transfection of UFL-AG 286 cells with the plasmids containing the different *ie1* promoter versions (154, 157, 257, 357 and 492; see Fig. 6) and pBF1 (IE1, *dark grey bars*) or pBF1Δp (pBF1-derived plasmid comprising the complete *Ag*MNPV *ie1* ORF lacking the *ie1* promoter; negative control: *light grey bars*). (**B**) Transfection with the indicated plasmids and post-infection with *Ag*MNPV: *dark grey bars*; mock infected controls: *light grey bars*. (**C**) Ratios between *β*-galactosidase expression levels under the control of the *ie1* promoter versions as assayed in **A** (left panel) and **B** (right panel). (**D**) *β*-Galactosidase expression levels under the control of the AcMNPV *etl* promoter with and without *Ag*MNPV IE1 protein supplied by co-transfection with different plasmid constructs or by *Ag*MNPV infection. pAgETLZD: transfected cells with pAgETLZD plasmid that contain *lac*Z ORF under the control of *Ac*MNPV *etl* promoter. pAgETLZD/pBF1: co-transfected cells with pAgETLZD and pBF1 plasmid comprising the complete *Ag*MNPV *ie1* ORF. *Negative controls*: pAgETLZD/pBF1Δp: co-transfected cells with pAgETLZD and a pBF1Δp. pAg492Δp/pBF1: co-transfected cells with a pAg492Δp lacking promoter, and pBF1 plasmid. pAg492Δp: cells transfected with pAg492Δp. pAg492Δp/infected: cells transfected with pAg492Δp and post-infected with *Ag*MNPV (MOI 5)

factors and/or RNApol II [19, 64], and through the construction of chimerical IE1 proteins [63], that the acidic domain is sufficient to activate both, homologous or heterologous early promoters.

IE1 phosphorylation

Stewart and Theilmann [28] in *Op*MNPV and Choi and Guarino [65, 66] and Slack and Blissard [62] in *Ac*MNPV have shown that IE1 expressed in insect cells is phosphorylated in multiple sites. Interestingly, the pattern of phosphorylated bands changes with the progress of infection. Nothing is known about the kinases responsible for the phosphorylation of IE1 and their targets within this protein. Nonetheless, all group I NPV IE1 proteins contain several phosphorylation consensus sites for the general cellular kinases, such as cAMP- or cGMP-dependent protein kinase, protein kinase C, casein kinase II, tyrosine kinase, etc. Six of these phosphorylation targets are conserved in type and position in a multiple alignment of IE1 proteins (Fig. 5A). Therefore, we speculate which either cellular or viral kinases (e.g. *Ac*MNPV PK-1), and phosphatases play an important role in the modulation of the IE1 activity. For example, in *Ac*MNPV, different phosphorylation patterns are obtained when IE1 is synthesised in the presence or in the absence of an oligonucleotide containing the direct repeat sequence of the *hr5* site.

Helix-loop-helix domain (dimerisation domain)

Deletion and insertion experiments in other group I NPVs, that disrupted the two α-helices potentially involved in the protein-protein dimerisation abolished the oligomerisation of IE1 proteins [32, 50]. The association of these defective IE1 variant polypeptides with a functional IE1 restores the binding capacity to *hr* sequences and supports the idea that oligomerisation is essential for DNA binding. By analogy to the helix-loop-helix motifs of other transcriptional activators [67], we predict that putative IE1 α-helices, α14 and α15, would have an equivalent function and that the most hydrophobic faces of the predicted C-terminal α-helices would be essential for the protein–protein interaction (Fig. 4).

Promoter analysis

We have shown the transcriptional activity of *Ag*MNPV *ie1* promoter by constructing recombinant plasmids carrying the *lac*Z ORF under the control of different versions of the *ie1* 5′ UTR, in complete absence of other viral factors acting in *cis* or *trans.*

The results obtained with the small UAR version (154 pb) suggest that it contain a minimum promoter sequence sufficient for the recognition by host transcription factors. However, it is clear that additional sequences are needed for the achievement of maximum promoter activity. Interestingly, the "small UAR version" corresponds to the region of highest similarity within group I nucleopolyhedrovirus *ie1* promoters, containing several fully conserved transcriptional motifs (one copy of each, GATA, CAAT-like, TATA-box; two copies of CGT-like and CAGT-INR). In addition, the TATA-box and CAGT-INR are separated by a stretch of 27 nucleotides, 24 of which are identical in group I nucleopolyhedroviruses (Fig. 3) [51].

Furthermore, the steady increase in expression level yields obtained with the other longer promoter versions suggests the existence of one or more *cis* elements in the distal UAR region that could bind different host transcription factors that ensure a high level of expression. Interestingly, the more distal CGT-like motif ($_{-391}$CACGTG$_{-386}$) has been described as a binding motif for transcription factors belonging to the bZIP family [68]. Otherwise, the sequence $_{-263}$CATCCCGCCGCGC$_{-251}$ resembles the binding motif of the Sp-like transcription factors founded in SF-9 cells [69]. Both motifs only exist in the 492 promoter version.

On the other hand we observed that the presence of IE1, increase the transcription activity from all promoter versions. The highest relative increment was found into the region between –257 and –357 pb; this region is highly conserved in baculovirus of group I (Fig. 3). Probably, this jump in the transcriptional activity is due to that in the construction –257 the binding motif for Sp-like proteins is disrupted.

Further studies are underway to dissect new functional sequences that are relevant for the *ie1* promoter activity.

Acknowledgements This work was supported by grants from Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT, FONCyT), and Universidad Nacional de Quilmes (UNQ) to PDG; Centro Argentino Brasileño de Biotecnología (CABBIO) to VR, PDG and BMR. MFB and MGP hold a fellowship from CONICET. PDG, VR and MEL are Research Career Members of CONICET, Argentina.

References

1. M.H.V. van Regenmortel, C.M. Fauquet, D.H.L. Bishop, E.B. Carstens, M.K. Estes, S.M. Lemon, J. Maniloff, M.A. Mayo, D.J. McGeoch, C.R. Pringle, R.B. Wickner, Virus Taxonomy, The Classification and Nomenclature of Viruses. In *The Seventh Report of the International Committee on Taxonomy of Viruses* (Academic Press, San Diego, 2000)
2. L. Miller, J. Invertebr. Pathol. **65**, 211–216 (1995)
3. M.D. Ayres, S.C. Howard, J. Kuzio, M. Lopez-Ferber, R.D. Possee Virology **202**, 586–605 (1994)
4. S. Gomi, K. Majima, S. Maeda J. Gen. Virol. **80**, 1323–1337 (1999)
5. H.A. Lauzón, P.B. Jamieson, P.J. Krell, B.M. Arif J. Gen. Virol. **86**, 945–961 (2005)
6. J.G. de Jong, H.A. Lauzon, C. Dominy, A. Poloumienko, E.B. Carstens, B.M. Arif, P.J.. Krell, J. Gen. Virol. **86**, 929–943 (2005)
7. M. Ikeda, M. Shikata, N. Shirata, S. Chaeychomsri, M. Kobayashi, J. Gen. Virol. **87**, 2549–2562 (2006)
8. C.H. Ahrens, R.L. Russell, C.J. Funk, J.T. Evans, S.H. Harwood, G.F. Rohrmann, Virology **229**, 381–399 (1997)
9. R.L. Harrison, B.C. Bonning, J. Gen. Virol. **84**, 1827–1842 (2003)
10. M. Nakai, C. Goto, W. Kang, M. Shikata, T. Luque, Y. Kunimi, Virology 316, 171–183 (2003)
11. A.K. Jakubowska, S.A. Peters, J. Ziemnicka, J.M. Vlak, M.M. van Oers, J. Gen. Virol. **87**, 537–551 (2006)
12. M.M. van Oers, M.H. Abma-Henkens, E.A. Herniou, J.C. de Groot, S. Peters, J.M. Vlak, J. Gen. Virol. **86**, 2069–2080 (2005)
13. X. Chen, W.F. Ijkel, R. Tarchini, X. Sun, H. Sandbrink, H. Wang, S. Peters, D. Zuidema, R.K. Lankhorst, J.M. Vlak, Z. Hu, J. Gen. Virol. **82**, 241–257 (2001)
14. J. Kuzio, M.N. Pearson, S.H. Harwood, C.J. Funk, J.T. Evans, J.M. Slavicek, G.F. Rohrmann, Virology **253**, 17–34 (1999)
15. Q. Li, C. Donly, L. Li, L.G. Willis, D.A. Theilmann, M. Erlandson, Virology **294**, 106–121 (2002)
16. W.F IJkel, E.A. van Strien, J.G. Heldens, R. Broer, D. Zuidema, R.W. Goldbach, J.M. Vlak, J. Gen. Virol. **80**, 3289–3304 (1999)
17. Y. Pang, J. Yu, L. Wang, X. Hu, W. Bao, G. Li, C. Chen, H. Han, S. Hu, H. Yang, Virology **287**, 391–404 (2001)
18. W. Wang, N. Leat, B. Fielding, S. Davison, Virus Genes **23**, 53–62 (2001)
19. G.W. Blissard, P.H. Kogan, R. Wei, G.F. Rohrmann, Virology **190**, 783–793 (1992)
20. J.A. Dickson, P.D. Friesen, J. Virol. **65**, 4006–401 (1991)
21. R.A. Krappa, Behn-A. Krappa, F. Jahnel, W. Doerfler, D. Knebel-Mörsdorf, J. Virol. **66**, 3494–3503 (1992)
22. S.S. Pullen, P.D. Friesen, J. Virol. **69**, 156–165 (1995)
23. D.A. Theilmann, S. Stewart, Virology **180**, 492–508 (1991)
24. D.D. Carson, M.D. Summers, L.A. Guarino, J. Virol. **65**, 945–951 (1991)
25. F.R. Kovacs, L.A. Guarino, M.D. Summers, J. Virol. **65**, 5281–5288 (1991)
26. A.I. Passarelli, L.K. Miller, J. Virol. **67**, 2149–2158 (1993)
27. M. Kool, C.H. Ahrens, R.W. Goldbach, G.F. Rohrmann, J.M. Vlak, Proc. Natl. Acad. Sci. U.S.A. **91**, 11212–11216 (1994)
28. S. Stewart, D.A. Theilmann, J. Gen. Virol. 74, 1819–1826 (1993)
29. L. Cherbas, P. Cherbas, Insect Biochem. Mol. Biol. **23**, 81–90 (1993)
30. G.E. Chisholm, D.J. Henner, J. Virol. **62**, 3193–3200 (1988)
31. P.D. Friesen, in *The Baculoviruses*, ed. by L.K. Miller (Plenum Press, New York, 1997) pp. 141–170
32. F.R. Kovacs, J. Chio, L.A. Guarino, M.D. Summers, J. Virol. **66**, 7429–7437 (1992)
33. F. Moscardi, Annu. Rev. Entomol. **44**, 257–289 (1999)
34. D.W. Johnson, J.E. Maruniak, J. Gen. Virol. **70**, 1877–1883 (1989)
35. A.M. Crawford, L.K. Miller, J. Virol. **62**, 2773–2781 (1988)
36. B.M. Ribeiro, K Hutchinson, L.K. Miller, J. Virol. 68, 1075–1084 (1994)

37. J.H. Miller, in *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory Press, New York 1977)
38. J.D. Thompson, D.G. Higgins, T.J. Gibson, W. Clustal, Nucleic Acids Res. **22**, 4673–4680 (1997)
39. J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, Nucleic Acids Res. 25, 4876–4882 (1997)
40. M.O. Dayhoff, R.M. Schwartz, B.C. Orcutt, in *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 3, ed. by M.O. Dayhoff (NBRF, Washington, 1978) p. 345
41. T.D. Schneider, R.M. Stephens, Nucleic Acids Res. **18**, 6097–6100 (1990)
42. J.A. Cuff, G.J. Barton, Proteins **34**, 508–519 (1999)
43. J.A. Cuff, M.E. Clamp, A.S. Siddiqui, M. Finlay, G.J. Barton, Bioinformatics **14**, 892–893 (1998)
44. B Rost, C Sander, J. Mol. Biol. **232**, 584–599 (1993)
45. A.A. Salamov, V.V. Solovyev, J. Mol. Biol. **247**, 11–15 (1995)
46. R.D. King, M.J.E. Sternberg, Protein Sci. **5**, 2298–2310 (1996)
47. D Frishman, P Argos, Proteins **27**, 329–335 (1997)
48. K. Hofmann, P. Bucher, L. Falquet, A. Bairoch, Nucleic Acids Res. **27**, 215–219 (1999)
49. D.M. Bulach, C.A. Kumar, A. Zaia, B. Liang, D.E. Tribe, J. Invertebr. Pathol. **73**, 59–73 (1999)
50. S.M. Rodems, S.S. Pullen, P.D. Friesen, J. Virol. **71**, 9270–9277 (1997)
51. S.S. Pullen, P.D. Friesen, J. Virol. **69**, 3575–3583 (1995)
52. M.N. Pearson, G.F. Rohrmann, Virology **235**, 153–65 (1997)
53. M. Lu, R.R. Johnson, K. Iatrou, Virology **218**, 103–113 (1996)
54. L.A. Guarino, W Dong, J. Virol. **65**, 3676–3680 (1991)
55. L.A. Guarino, W. Dong, Virology **200**, 328–335 (1994)
56. R. Huybrechts, L. Guarino, M.V. Brussel, V. Vulsteke, Biochim. Biophys. Acta 1129, 328–330 (1992)
57. R.L. Davis, P.F. Cheng, A.B. Lassar, H. Weintraub, Cell **60**, 733–746 (1990)
58. D. Chattopadhyay, T. Raha, D. Chattopadhyay, Virology **239**, 11–19 (1997)
59. Colberg-A.M. Poley, L. Huang, V.E. Soltero, A.C. Iskenderian, R.F. Schumacher, D.G. Anders, Virology **246**, 400–408 (1998)
60. S.T. Triezenberg, R.C. Kingabury, S.L. McKnight, Genes Dev. **2**, 718–729 (1988)
61. H Zhang, Al-H.O. Barazi, A.M. Colberg-Poley, Virology **223**, 292–302 (1996)
62. J.M. Slack, G.W. Blissard, J. Virol. 71, 9579–9587 (1997)
63. I.J. Forsythe, C.E. Shippam, L.G. Willis, S. Stewart, T. Grigliatti, D.A. Theilmann, Virology **252**, 65–81 (1998)
64. G.W. Blissard, G.F. Rohrmann, J. Virol. **65**, 5820–5827 (1991)
65. J. Choi, L. Guarino, J. Virol. **69**, 4548–4551 (1995)
66. J. Choi, L. Guarino, Virology **209**, 99–107 (1995)
67. C. Murre, P.S. McCaw, D. Baltimore, Cell **56**, 777–783 (1989)
68. P.H. Kogan, G.W. Blissard, J. Virol. **68**, 813–822 (1994)
69. A. Ramachandran, A. Jain, P. Arora, M.D. Bashyam, U. Chatterjee, S. Ghosh, V.K. Parnaik, S.E. Hasnain, J. Biol. Chem. **276**, 23440–23449 (2001)