第17卷 3期
2002年8月

中 国 病 毒 学
VIROLOGICA SINICA

17(3):234－238
August 2002

棉铃虫核多角体病毒 *lef*－3 基因的序列分析*

张小霞，梁振普，牛国栋，张忠信**，丁清泉
（中国科学院武汉病毒研究所，湖北武汉 430071）

Sequence Analysis of the *lef*-3 Gene of *Helicoverpa armigera* Single-nucleocapsid Nucleopolyhedrovirus *

ZHANG Xiao-xia, LIANG Zhen-pu, NIU Guo-dong, ZHANG Zhong-xin**, DING Qing-quan
(*Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan* 430071, *China*)

Abstract: An *lef*-3 gene of HaSNPV was identified. It was localized in *Bam*HⅠ-F fragment of the genome. Nucleotide sequencing showed the open reading frame was 1140 nt, encoded 379 amino acids with a predicted size of 44kD. A "TATA" box was found in 40～43nt upstream of the transcriptional start codon ATG. A typical poly(A) signal, AATAAA, was found in the downstream of the transcriptional stop codon. HaSPV LEF-3 protein also contains the conserved motif, which exists in nearly 20 kinds of SSBs, so HaSNPV LEF-3 protein is a probable SSB protein.

Key words: *lef*-3; Sequence analysis; Motif SSB

摘要：在棉铃虫核多角体病毒(*Helicoverpa armigera* Single-nucleocapsid Nucleopolyhedrovirus, HaSNPV)基因组中发现了 *lef*-3 基因，该基因位于病毒基因组的 *Bam*HⅠ-F 片段上，全长 1140bp，编码 379 个氨基酸，预计蛋白质分子量为 44kD，启动子区具有 TATAbox，位于起始密码子 ATG 上游 40～43nt 处，在终止密码子下游具典型的 poly(A)终止信号 AATAAA。并且在它的氨基酸序列的 N－端也具有一个类似 SSB 保守序列的基元结构。

关键词：*lef*-3；序列分析；基元结构；SSB
中图分类号：S433　　**文章标识码**：A　　**文章编号**：1003-5125(2002)03-0234-05

病毒 DNA 复制是其生命周期中的最主要环节，病毒复制和转录的调控研究是病毒分子生物学研究的重要内容。杆状病毒基因组呈双链环状 DNA，分子大小在 90kb～160kb 之间[1]，是一类大分子 DNA 病毒。近年来，杆状病毒作为表达载体已广泛应用，但其 DNA 复制转录调控机理却研究的相对较少。1993 年，Li[2]等从 AcNPV 中鉴定出一个病毒晚期基因表达调控因子 LEF-3(late expression factor 3)，该基因是病毒早期基因，它可转录出 2.0kb 的 mRNA，，基因产物是一个 44.5kD 的蛋白质。LEF－3 通过反式作用调控病毒晚期和极晚期基因的表达。1995 年，Hang[3]等进一步证明，*lef*-3 编码的是一种单链 DNA 结合蛋白(single-stranded DNA-binding protein, SSB)。SSB 蛋白仅与单链 DNA 结合，没有核苷酸序列特异性，它能通过结合单链 DNA 维持双链 DNA 复制叉形态等方式调控 DNA 的复制，进而调控晚期基因和极晚期基因的转录和表达。在其它杆状病毒研究中，OpNPV *lef*-3 基因也编码一个 SSB 蛋白[4]，它与 AcNPV *lef*-3 基因的同源性为 39%。AcNPV 和 OpNPV LEF－3 的 N 端均含有一段保守氨基酸序列，该序列中的重要氨基酸基元与 SSB 蛋白一致，这一保守序列可能在 SSB 的功能上具有重要意义。本文报道 HaNPV *lef*-3 基因的序列测定及其编码蛋白

收稿日期：2001-12-10，修回日期：2002-01-16
* 基金项目：中国科学院农办重大课题(NK 十五-B-07-04)；国家 863 项目(2001AA246014)
作者简介：张小霞(1973-)，女，河南省籍，硕士，研究方向为分子病毒学。
** 通讯作者：张忠信(1957-)，男，山西省籍，副研究员，研究方向为病毒学。Correspondence author.

与 SSB 的同源性比较,以便为 *lef*-3 的功能研究及杆状病毒 DNA 复制调控研究奠定基础。

1 材料与方法

1.1 材料

1.1.1 棉铃虫核型多角体病毒由本实验室保存。

1.1.2 菌种 DH5a,载体 pUC19 为本实验室保存。

1.1.3 分子生物学试剂 T_4 连接酶、溶菌酶、去磷酸化酶、RNase、λDNA/*Hin*d Ⅲ分子量标准均购自 TaKaRa;蛋白酶 K 购自 Merk 公司; *Bam*HⅠ购自 GiBCO;柱离心式胶回收试剂盒、一次性胶切割器均为上海化舜生物工程公司产品;质粒小量抽提试剂盒购自 Promega;其他的分析纯级生化试剂分别购自 TaKaRa、Promega、华美等生物公司。

1.1.4 培养基、各种缓冲液及碱解液:LB 液体培养基和固体培养基,SOC 液体培养基以及各种缓冲液的配方参见文献[5],碱解液的配方参照传统方法稍作改进。

1.2 方法

1.2.1 病毒多角体的扩增和纯化 取 4℃冰箱保存的多角体以 2×10^7PIB/mL 的浓度,喂食感染健康的棉铃虫 4 龄初幼虫,4－6d 后收集典型的病毒致死虫,研磨匀浆,加水稀释后三层纱布过滤,1 000r/min离心 1min,取上清,4 000r/min 离心 30min,用水充分悬浮沉淀,如此差速离心数次(一般为 3 次)后,沉淀用 ddH_2O 充分悬浮,混匀后计数,分装于 1.5mL 的离心管,4℃保存备用。

1.2.2 病毒 DNA 的制备 取纯化的多角体悬液,加入溶菌酶(10mg/mL),37℃保温 1－2h,用水洗 2－3 遍,再加入 Proteinase K(20mg/mL)5μL, 37℃保温 30min,加入 500 μL 碱解液,37℃保温 30min(溶液变清),再加入 10%SDS 至终浓度为 1%,37℃保温 30min,碱解产物经平衡酚抽提 2 次,氯仿:异戊醇(24:1)抽提一次后,用 ddH_2O 透析 48 h,纯化的 DNA 置 4℃保存备用,紫外分光光度计测其浓度。

1.2.3 HaNPV *Bam*HⅠ-F 片段的分离回收

取 9μL 病毒 DNA(约 0.5－1μg),加 0.5μL *Bam*HⅠ, 37℃消化 2～3h,酶消化液在 0.75%琼酯糖凝胶上电泳 6～10h,用柱离心式胶回收试剂盒回收 F 片段,具体操作如下:每 100mg 胶块加入 300μL 溶胶液,50℃,10min 使胶彻底溶化,加入 1/3 体积的异丙醇,50℃放置 1min。将溶好的液体移入回收柱,高速离心(12 000～13 000r/min)1min,倒掉收集管中的液体,加入 500μL 洗涤液,静止 1min,离心 1min。将回收柱移到一干净的离心管中,加入 30μL 洗脱液,静止 1～5min,离心 1min,即得到纯化的核酸。纯化后的核酸在 0.7% 琼脂糖凝胶中电泳,检查其回收率。

1.2.4 克隆载体的制备及连接反应、感受态细胞的制备 用 *Bam*HⅠ消化 pUC19 载体并去磷酸化,将上述回收的 *Bam*HⅠ-F 片段与处理过的载体相连,连接反应参照文献进行[5],感受态细胞的制备参考文献进行[5]。

1.2.5 转化 取 200μL 感受态细胞,加入 5μL 连接产物,具体转化过程参考文献进行[5]。

1.2.6 质粒的提取 利用小量质粒抽提试剂盒提取质粒 DNA,具体过程参考产品说明书进行。

1.2.7 质粒的鉴定与测序 *Bam*HⅠ消化质粒 DNA,0.7%琼脂糖凝胶电泳,根据酶切片段大小判断是否为阳性克隆子。序列测定由上海生工公司根据 pUC19 上的通用引物完成。

1.2.8 DNA 序列分析及其编码氨基酸的计算机分析 序列用 FASTA 和 BLAST 软件与 GenBank 数据库进行比较分析,运用 DNAClub、EditSeq 软件进行氨基酸序列的分析,利用 MEGALIGN、CLUSTALX、GeneDoc 进行同源性的比较分析。

2 结果与讨论

2.1 重组质粒的酶切鉴定

将回收的 *Bam*HⅠ-F(7.827kb)片段(图略)与经过同样酶切处理过的 pUC19(2.69kb)载体连接,转化,筛选到重组质粒 pUC-lef3,用 *Bam*HⅠ酶切鉴定,结果显示可切出一 7.827kb 大小的片段[6]。

2.2 *lef*-3 基因序列分析

对插入片段进行 DNA 序列测定,比较分析结果见图 1, *lef*-3 确实位于 HaSNPV 基因组的 *Bam*HⅠ-F 片段,该基因全长 1 140bp,编码 379 个氨基酸,预计蛋白质分子量为 44kD,启动子区具有 TATA box,位于起始密码子 ATG 上游 40～43nt 处,在终止密码子下游具典型的 poly(A)终止信号 AATAAA。与 AcNPV 和 OpNPV 的 LEF－3 一样,

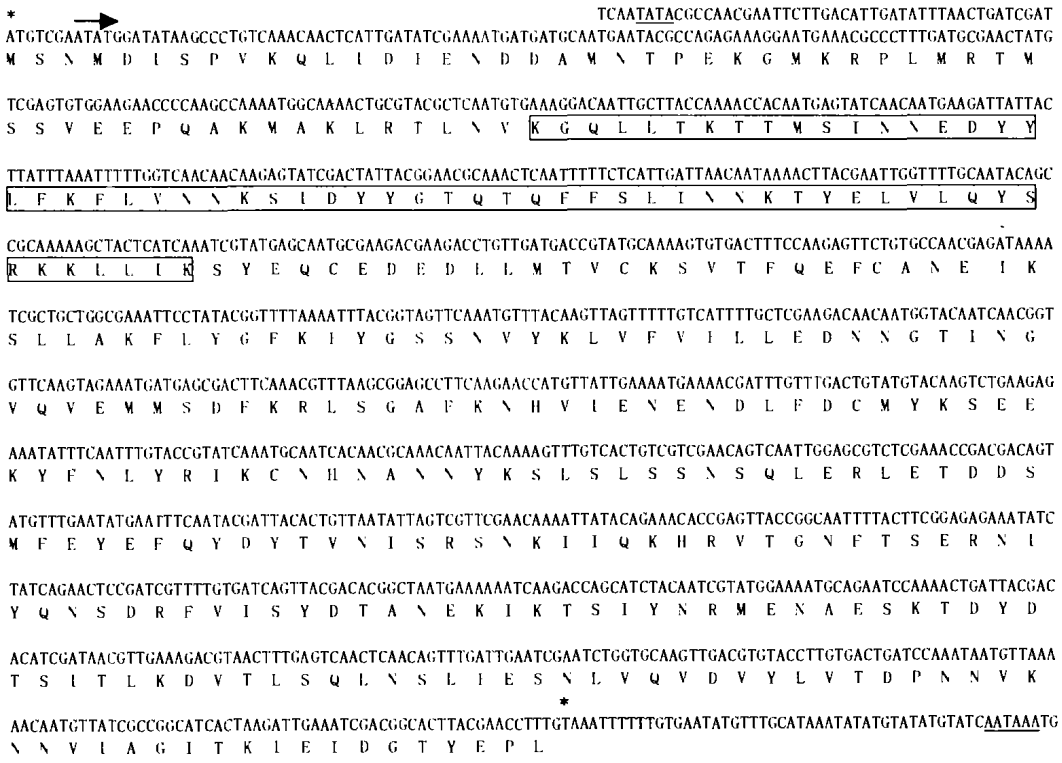

图1 棉铃虫核多角体病毒 *lef*-3 基因的核苷酸序列与推导的氨基酸序列

Fig.1 The *lef*-3 nucleotide sequence and puptative amino acid sequence of HaNPV

The gene's direction of transcription is indicated by arrow. Putative transcription signal TATA box, for early transcription and AATAAA for polyadenylation are underlined. The transcriptional start codon and stop codon are indicated by * above the ATG and TAA The conserved pattern of amino acids similar to SSB motif, in LEF-3 amino-terminal regions, is framed.

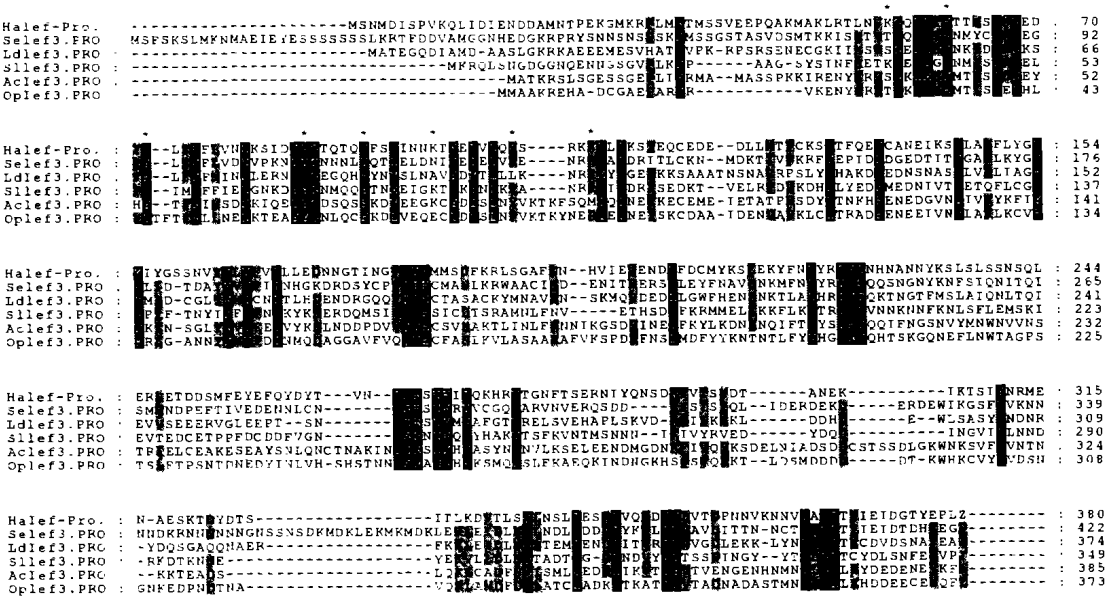

图2 杆状病毒 LEF-3 基因氨基酸序列排列比较

组成类似 SSB 基元的保守碱性氨基酸和芳香族氨基酸均以 * 标出，黑色代表 100% 同源性(identity)，灰色代表 80% 同源性(identity)。

Fig.2 Aligment of amino acids of baculviral LEF-3

The conserved aromatic and basic amino acids, which constitute SSB motif, are indicated by * above the sequence; Two levels of shading were set, black for 100% identity and gray for 80%.

在 HaSNPV LEF－3 的 N 端也具有一个类似 SSB 的保守基元结构，K－X5－K－X09－Y－X13－Y－X6－F－X6－K－X7－Y－X3－K，该结构位于 LEF－3 的第 55 位～111 位氨基酸，共 57 个氨基酸残基(图 1)。将 HaNPV LEF-3 与来源于杆状病毒 AcNPV、OpNPV、LdNPV、SeNPV、SlNPV 的 LEF－3 相比较，结果显示，在整个氨基酸序列上 HaNPV LEF-3 与其它五种 LEF－3 同源性不高，分别为(Identity)19%、20%、21%、24% 和 23%(图略)，但是它们编码的类似 SSB 的 K－X03-K－X08-10－Y－X14－Y－X06－F－X00－K－X13－F/Y－X03－K 结构(X 代表任何氨基酸)保守性很强(图 2)。

将棉铃虫核型多角体病毒 LEF－3 N－端的类似 SSB 保守基元的氨基酸序列与 18 种来源不同的 SSBs 以及 AcNPV LEF-3 和 OpNPV LEF-3 的保守基元相比，结果如图 3 所示。整个基元结构有 8 个保守性氨基酸位点(CCV、HzIE-1 缺少 5 位保守性氨基酸，HCMV 缺少 6 位保守性氨基酸)，我们分别称之为 1、2、3、4、5、6、7、8 位氨基酸，这 8 个位点的氨基酸分别由碱性的精氨酸(R)、赖氨酸(K)，芳香族的苯丙氨酸(F)、酪氨酸(Y)以及含芳香环的色氨酸(W)组成。从图 3 中我们可以看出，1、2 位是碱性的 R 或 K，3、4、5 位是芳香族的 F 或 Y(SSB3 位是含芳香环的 W，CCV 和 HzIE－1 缺少 5 位芳香族氨基酸)，6 位是碱性 R 或 K(HCMV 缺少 6 位芳香族氨基酸)，7 位是芳香族的 F、Y 或含芳香环的 W，8 位是碱性的 R 或 K。HaNPV LEF-3 具有与上述完全相同的结构特征，各保守区氨基酸分别为：1、2 位是 K，3、4 位是 Y，5 位是 F，6 位是 K，7 位是 Y，8 位是 K，因此可以推测 HaNPV *lef*-3 基因很有可能编码 SSB 蛋白。

SSB 在很多复制系统中都是必备的成员之一[7,8]，例如在疱疹病毒[9]和腺病毒中[10,11]。

SSBs 有助于在正好低于正常 Tm 值的温度下使超螺旋 DNA 分子在复制起始点解开，并且 SSB 能维持复制叉处 DNA 的单链状态，也可能影响其它参与复制过程的蛋白质[8]。Li 等人[2]首次证明了 *lef*-3 基因对杆状病毒晚期基因(capsid and polyhedrin)的表达是必需的，杆状病毒中，晚期基因的表达完全依赖于 DNA 的复制，LEF－3 可能直接介于晚期基因的转录过程。目前，人们已在至少三个病毒系统中发现，SSB 明显作用于 DNA 的复制和对晚期基因的表达具有明显的促进作用，腺病毒的组

		1	2	3	4	5	6	7	8	
GP5	16	R-X04-R-X04-Y-X07-Y-X06-Y-X04-K-X26-F-X06-R								80
GP32	67	K-X03-K-X12-Y-X14-Y-X06-Y-X03-K-X04-Y-X07-K								123
SSB	43	K-X05-K-X04-W-X05-F-X-9-Y-X02-K-X04-W-X07-R								86
PIKE	16	R-X05-K-X04-Y-X07-Y-X06-Y-X19-K-X12-F-X06-R								81
AD-5	410	R-X03-K-X03-F-X36-Y-X13-F-----K-X16-W-X11-R								499
HSV-1	803	K-X03-R-X09-F-X04-F-X04-F-X03-K-X10-W-X06-R								849
VZV	801	R-X03-R-X09-F-X04-F-X04-F-X01-R-X12-W-X06-R								847
EHV	808	R-X03-K-X10-F-X04-Y-X04-F-X01-R-X12-W-X06-R								855
EBV	741	K-X03-R-X12-F-X03-Y-X04-F-X03-K-X06-W-X06-R								771
CCV	758	R-X04-R-X22-F-X12-F-X04---X05-R-X03-F-X07-R								821
HCMV	810	R-X05-R-X09-F-X02-Y-X06-F-X01---X09-W-X09-K								857
MCMV	791	R-X04-K-X04-F-X02-F-X06-F-X01-R-X08-W-X09-R								831
AcIE-1	455	K-X03-K-X10-Y-X12-F-X04-F-X04-R-X11-W-X05-K								511
BmIE-1	459	K-X03-K-X10-Y-X12-F-X04-F-X04-R-X11-W-X05-K								516
CfIE-1	433	K-X03-K-X10-Y-X12-F-X04-F-X06-R-X11-W-X05-K								491
HzIE-1	432	K-X02-R-X10-Y-X12-F-X04---X05-R-X11-W-X05-K								491
OpIE-1	529	K-X03-K-X10-Y-X12-F-X04-F-X06-R-X11-W-X05-K								584
SeIE-1	580	K-X02-K-X10-Y-X07-F-X04-F-X09-R-X11-W-X06-K								635
AcLEF-3	39	K-X03-K-X08-Y-X13-Y-X06-F-----K-X13-Y-X03-K								93
OpLEF-3	30	K-X03-K-X10-Y-X13-Y-X06-F-----K-X13-Y-X03-K								86
HaLEF-3	55	K-X05-K-X10-Y-X13-Y-X06-F-X06-K-X07-Y-X03-K								111

图 3　HaNPV LEF-3 与 20 种 SSB 的保守性基元的氨基酸序列比较

X 代表任意氨基酸，数字表示 X 的数目。K、R、Y、F、W 分别代表赖氨酸、精氨酸、酪氨酸、苯丙氨酸和色氨酸。

Fig. 3　A comparison between HaNPV LEF-3 and 20 kinds of putative single- strand-DNA-binding amino acid sequence motifs

The spacer regions between conserved aromatic and basic amino acids are designated by X with a number indicating the spacer residues present. Numbers at the margins of the sequence indicate the location of the motif in the amino acid sequence. The alignment from the gene 5 protein of bacteriophage fd(GP5); gene 32 protein of bacteriophage T4 (GP32); SSB of *E. coli*; DNA-binding proteins from bacteriophage ike (PIKE); HSV-1; Varicella-zoster virus (VZV); Epsstein-Barr virus (EBV); adenovirus type 5 (AD-5); equine herps virus (EHV); human cytomegalovirus (HCMV); mouse cytomegalovirus (MCMV); channel catfish virus (CCV). The baculovirus IE-1 sequences are from AcMNPV (AcIE-1); BmMNPV (BmIE-1); OpMNPV (OpIE-1); CfMNPV (CfIE-1); SeMNPV (SeIE-1); HzSNPV (HzIE-1). The baculovirus LEF-3 sequences are from AcMNPV (AcLEF-3); OpMNPV(OpLEF-3); HaSNPV(HaLEF-3).

成，中间被数目不等的其它氨基酸残基隔 DNA 结合蛋白，单纯疱疹病毒的 ICP8 蛋白，T_4 噬菌体的基因 32 蛋白，分别直接作用于相关晚期基因的转录[12～14]。并且在许多 SSBs 中都有一个保守性基元结构，由保守的芳香族氨基酸和碱性氨基酸开[15,16]，在 AcNPV LEF-3 和 OpNPV LEF-3 的 N 端有类似 SSB 的保守性氨基酸序列基元结构(K-X_{03}-K-X_{08}-X_{10}-Y-X_{14}-Y-X_{06}-F-X_{00}-K-X_{13}-F/Y-X_{03}-K)[4,16]，此基元大约 80 个氨基酸，在 LEF-3 的总序列中显示很高的氨基酸同源性(similarity 为 82%，identity 为 56%)，推测此保守区很可能就是 LEF-3

的功能区。从对 AcNPV、OpNPV LEF-3 的研究结果可看出，LEF－3 对 DNA 复制的重要性可能归于它具有 SSB 功能的行为。棉铃虫核多角体病毒 LEF－3 具有与近 20 种不同来源的 SSB 蛋白完全相同的保守基元结构，因此基因的表达产物很有可能是一种单链 DNA 结合蛋白，即 SSB，关于蛋白的纯化和结合试验还在本实验室进行。

参考文献

[1] Blissatd G W, Rohrmann G F. Baculovirus diversity and molecular biology[J]. Annu Rev Entomol, 1990, 35:127－155.

[2] Li Y H, Passarelli A L, Miller L K. *et al*. Identification, sequence, and transcriptional mapping of *lef*-3, a baculovirus gene involved in late and very late gene expression[J]. J Virol, 1993, 67:5260－5268.

[3] Hang X, Dong W, Guarino L A. The *lef*-3 gene of autographa californica nuclear polyhedrosis virus encodes a single－stranded DNA－binding protein[J]. J Virol, 1995, 69:3924－3928.

[4] Ahrens C H, Carlson C, Rohrmann G F. Identification, sequence, and transcriptional analysis of *lef*-3, a gene essential for Orgyia pseudotsugata baculovirus DNA replication[J]. Virology, 1995, 210:372－382.

[5] 金冬雁. 分子克隆实验指南[M]. 第二版，北京：科学出版社，1992.

[6] 张小霞，张忠信，丁清泉. 棉铃虫核多角体病毒 *iap*2 基因的克隆和序列分析[J]. 中国病毒学，2001，16:338－343.

[7] Chase J W, Williams K R. Single－stranded DNA binding proteins required for DNA replication[J]. Ann Rel Biochem, 1986, 55:103－136.

[8] Kornberg A, Baker T A. DNA Replication[M]. 2nd, New York: W. H. Freeman, 1992.

[9] Weller S K, Lee K J, Sabourin D J, *et al*. Genetic analysis of temperature-sensitive mutants which define the gene for the major herpes simplex virus type 1 DNA-binding protein[J]. J Virol, 1983, 45:354－366.

[10] Challberg M D, Kelly T J. Animal virus DAN replication[J]. Annu Rev Biochem, 1989, 58:671－717.

[11] Stillman B. Initiation of eukaryotic DNA replication in vitro[J]. Annu Rev Cell Biol, 1989, 5:197－245.

[12] Chang L S, Shenk T. The adenovirus DNA-binding protein stimulates the rate of transcription directed by adenovirus and adeno-associated virus promoters[J]. J Virol, 1990, 64:2103－2109.

[13] Gao M, Knipe D M. Potential role for herpes simplex virus ICP8 DNA replication protein in stimulation of late gene expression[J]. J Virol, 1991, 65:2666－2675.

[14] Gauss P K B, Mcpheeters D S, Nelson M A, *et al*. Zinc(Ⅱ) and the single-stranded DNA binding protein of bacteriophage T4 [J]. Proc Natl Acad Sci USA, 1987, 84:8515－8519.

[15] Kool M, Ahrens C, Goldbach R W, *et al*. Identification of genes involved in DNA replication of the Autographa californica baculovirus[J]. Proc Natl Acad Sci USA, 1994, 91:11212－11216.

[16] Wang Y, Hall J D. Characterization of a major DNA binding domain in the Herpes Simplex Virus Type 1 DNA－binding protein (ICP8)[J]. J Virol, 1990, 64:2082－2089.