

647 190

P/00/001
Section 29
P/00/008
Section 29(1)
Regulation 3.1(2)

AUSTRALIA
Patents Act 1990

PATENT REQUEST/NOTICE OF ENTITLEMENT : STANDARD PATENT

We THE WELLCOME FOUNDATION LIMITED
of Unicorn House, 160 Euston Road, LONDON NW1 2BP, ENGLAND

being the Applicant and Nominated Person, request the grant of a patent for an invention entitled EXPRESSION SYSTEM which is described in an accompanying standard complete specification.

Convention priority is claimed from the following basic application:

Basic Applicant	Application Number	Application Date	Country	Country Code
The Wellcome Foundation Limited	9023111.9	24 October 1990	United Kingdom	GB

Timothy Charles Peakman, Martin John Page and Ian George Charles are the actual inventors of the invention.

The actual inventors made the invention for and on behalf of the person nominated for the grant of the patent

The basic application was the first application made in a Convention country in respect of the invention the subject of this request.

DATED this 23rd day of October 1991

THE WELLCOME FOUNDATION LIMITED
By their Patent Attorney

GRIFFITH HACK & CO

6025182 23/10/91



AU9186072

(12) PATENT ABRIDGMENT (11) Document No. AU-B-86072/91
(19) AUSTRALIAN PATENT OFFICE (10) Acceptance No. 647190

- (54) Title
EXPRESSION SYSTEM
- International Patent Classification(s)
(51)⁵ C12N 015/86 C12N 015/11
- (21) Application No. : 86072/91 (22) Application Date : 23.10.91
- (30) Priority Data
- (31) Number (32) Date (33) Country
9023111 24.12.90 GB UNITED KINGDOM
- (43) Publication Date : 30.04.92
- (44) Publication Date of Accepted Application : 17.03.94
- (71) Applicant(s)
THE WELLCOME FOUNDATION LIMITED
- (72) Inventor(s)
TIMOTHY CHARLES PEAKMAN; MARTIN JOHN PAGE; IAN GEORGE CHARLES
- (74) Attorney or Agent
GRIFFITH HACK & CO., GPO Box 4164, SYDNEY NSW 2001
- (57) Claim

8. A recombinant baculovirus having a promoter capable of directing expression of a heterologous gene having a coding sequence in insect cells infected with the recombinant baculovirus, the promoter being operably linked to the heterologous gene and the coding sequence being preceded by an untranslated leader of sequence T A A T C A T C C A C A G G A G A C T T T C T G.

13. A process for preparing a recombinant baculovirus, which process comprises:

(a) providing a recombinant baculovirus transfer vector having a promoter capable of directing expression of a heterologous gene in insect cells infected with the recombinant baculovirus, the promoter being operably linked to the heterologous gene and the coding sequence of the heterologous gene being preceded by an untranslated leader of sequence T A A T C A T C C A C A G G A G A C T T T C T G; and

(b) co-transfecting the cells with the recombinant baculovirus transfer vector and with intact wild-type baculovirus DNA.

26. A baculovirus transfer vector comprising

(i) a restriction site into which a heterologous gene may be cloned downstream of the sequence encoding the N-terminus of the polyhedrin gene;

(11) AU-B-86072/91
(10) 647190

-2-

(ii) another triplet of nucleotides in place of the natural ATG translation start codon of the polyhedrin coding sequence;

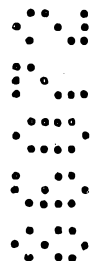
(iii) from up to the first 24 to up to the first 50 bases of the polyhedrin coding sequence; and

(iv) downstream of (iii) but upstream of (i), an untranslated leader of sequence T A A T C A T C C A C A G G A G A C T T T C T G.

AUSTRALIA
Patents Act 1990

647 190

ORIGINAL
COMPLETE SPECIFICATION
STANDARD PATENT



Invention Title: **EXPRESSION SYSTEM**



The following statement is a full description of this invention, including
the best method of performing it known to us:



GH&CO REF: 15431-BO:GJH:RK

EXPRESSION SYSTEM

This invention relates to the preparation of proteins by way of a baculovirus expression system.

The baculovirus vector system has been used to express both procaryotic and eucaryotic genes. The expression of these genes is controlled by the polyhedrin promoter of a baculovirus, in particular of Autographa californica nuclear polyhedrosis virus (AcNPV). The foreign genes are expressed in insect cells in culture which are infected with recombinant baculovirus incorporating the foreign gene.

In order to obtain a recombinant baculovirus, a baculovirus transfer vector is used. This vector includes the polyhedrin promoter. Baculovirus DNA flanking the polyhedrin gene is provided. The rest of the vector is typically made up of DNA from a bacterial plasmid. A foreign gene is inserted into the transfer vector downstream of the polyhedrin promoter such that its expression is controlled by this promoter.

The transfer vector containing the foreign gene and a baculovirus then co-transfect insect cells susceptible to baculovirus infection. Typically Spodoptera frugiperda cell cultures are utilised. Homologous recombination then occurs involving the viral DNA in the transfer vector upstream and downstream of the foreign gene and the corresponding DNA of the baculovirus. The foreign gene and its polyhedrin promoter is therefore transferred to the baculovirus. The recombinant baculovirus is cultured to express the foreign gene.

There are two types of baculovirus transfer vectors. First, there are transfer vectors which include a restriction site into which a foreign gene may be cloned downstream of the polyhedrin ATG start codon for translation. Such transfer vectors result in fusion

proteins in which the product of the foreign gene is fused to a N-terminal portion of the polyhedrin peptide.

Second there are transfer vectors in which the 5' untranslated leader of the polyhedrin gene ends before the natural ATG translation codon for polyhedrin and a restriction site is then provided. Such a transfer vector is pAc373 in which the 5' untranslated leader ends 8 bases upstream of the natural polyhedrin ATG. Genes cloned into the restriction site are expressed as mature proteins if they contain an ATG followed by an open reading frame coding for the desired product.

In EP-A-0340359 we described and claimed a baculovirus transfer vector which incorporates a restriction site into which a foreign gene may be cloned a short distance downstream of the N-terminus of the polyhedrin gene and in which another triplet of nucleotides is present in place of the natural ATG translation start codon for the polyhedrin gene. A preferred such vector is pAc36C. Increased levels of expression of a protein could be achieved by use of this type of transfer vector.

We have now surprisingly found that expression levels can be further improved by cloning a prokaryotic sequence in front of the coding sequence of a heterologous gene it is wished to express. The baculovirus expression system is not a prokaryotic expression system. Rather, expression occurs in insect cells which are classified as higher eukaryotic. Use of the prokaryotic leader sequence of the invention surprisingly results in higher levels of protein expression than if a Kozak sequence is employed instead. A Kozak sequence is a sequence believed optimal for eukaryotic expression.

Accordingly, the present invention provides a process for preparing a protein, which process comprises culturing, under such conditions that the protein is obtained, insect cells infected with a recombinant baculovirus having a promoter capable of directing expression in the cells of a



heterologous gene having a coding sequence encoding the protein, the promoter being operably linked to the heterologous gene and the coding sequence being preceded by an untranslated leader of sequence T A A T C A T C C A C A G G A G A C T T T C
5 T G.

The invention further provides a recombinant baculovirus having a promoter capable of directing expression of a heterologous gene having coding a sequence in insect cells
10 infected with the recombinant baculovirus, the promoter being operably linked to the heterologous gene and the coding sequence being preceded by an untranslated leader of sequence T A A T C A T C C A C A G G A G A C T T T C T G.

15 The leader of the invention may be extended at its 5' end, for example by up to 12 or up to 6 further nucleotides. Additional nucleotides may have been provided therefore to enable the leader sequence to be cloned into a restriction site of a baculovirus transfer vector. The additional nucleotides
20 may therefore represent the sequence of a restriction site or restriction sites. A leader sequence extended at its 5' end to accommodate a Bam HI site is:
A T C C T A A T C A T C C A C A G G A G A C T T T C T G.

25 This sequence may be provided with a G nucleotide at its 5' end. A sequence incorporating both a Hind III site and a Bam HI site is:
Hind III Bam HI
A G C T T G G A T C C T A A T C A T C C A C A G G A G A C T T T
30 C T G.

Typically, the leader sequence is fused to the translation initiation codon of the coding sequence of the heterologous gene it is wished to express. In such
35 circumstances, therefore, the 3' end of the leader sequence is fused directly to the ATG translation start codon of the coding sequence. By "heterologous gene" is meant a gene which encodes a polypeptide not ordinarily produced by the cells infected with the recombinant baculovirus. Preferably, the heterologous



gene is not a baculovirus gene or an insect cell gene.

The coding sequence of the heterologous gene commences with a translation initiation codon. The gene may be
5 a procaryotic or eukaryotic gene. It may be a synthetic gene. The coding sequence encodes a selected polypeptide. The selected polypeptide may be a physiologically active polypeptide, for example, a polypeptide active in the human body. The polypeptide may be capable of raising antibody. The
10 polypeptide may be tetanus toxin fragment C, stromelysin, pertussis major surface antigen or the HIV integrase protein. Other polypeptides which may be encoded by the heterologous gene include:
hepatitis B virus surface antigen;
15 hepatitis B virus core antigen;
growth factors such as epidermal growth factor or transforming growth factor;
interferons,
polypeptides derived from the envelope of an AIDS-associated
20 retrovirus;
polypeptides derived from an organism of the genus Plasmodium;
polypeptides derived from the Epstein-Barr virus;
the heavy chain of a human immunoglobulin molecule;
the light chain of a human immunoglobulin molecule;
25 the heavy chain of a mouse immunoglobulin molecule;
the light chain of a mouse immunoglobulin molecule;
hybrid immunoglobulin molecules e.g. molecules wherein the variable region of the hybrid immunoglobulin molecule is derived from a non-human, such as a mouse, variable region
30 sequence and the constant region of the hybrid immunoglobulin molecule is derived from a human constant region or wherein the complementarily determining regions (CDRs) of the hybrid immunoglobulin molecule are derived from a non-human immunoglobulin such as a murine immunoglobulin and the
35 remaining sequences of the hybrid immunoblogulin molecule are derived from a human immunoglobulin; and fusion polypeptides which comprise a hybrid immunoglobulin molecule contiguous with or connected through an amino acid sequence to an enzyme molecule.



The promoter is typically a polyhedrin promoter. The baculovirus is preferably Autographa californica. The insect cells which are infected with the recombinant baculovirus are suitably cells of Trichoplusia ni, Spodoptera frugiperda, Heliothis zea or Manduca sexta. The cells are preferably cells of a Spodoptera frugiperda cell line.

The cells which are infected with the recombinant baculovirus are typically cultured according to standard procedures (Summers, M.D. and Smith, G.E., 1987. A manual of methods for baculovirus vectors and insect cell culture procedures, Texas Agricultural Experimental Station, Bulletin No. 1555). The desired protein is expressed. It may then be isolated and purified.

The recombinant baculovirus may be prepared by a process comprising:

- (a) providing a recombinant baculovirus transfer vector having a promoter capable of directing expression of a heterologous gene in insect cells infected with the recombinant baculovirus, the promoter being operably linked to the heterologous gene and the coding sequence of the heterologous gene being preceded by an untranslated leader of sequence T A A T C A T C C A C A G G A G A C T T T C T G; and
- (b) co-transfecting the cells with the recombinant baculovirus transfer vector and with intact wild type baculovirus DNA.

Cells of Trichoplusia ni, Spodoptera frugiperda, Heliothis zea or Manduca sexta may be transfected. Preferably the cells are cells of a Spodoptera frugiperda cell line. The recombinant baculovirus transfer vector typically comprises DNA from the baculovirus Autographa californica flanking the promoter, leader sequence and foreign gene. The wild-type baculovirus DNA which also transfects the cells is then the genomic DNA of Autographa californica.

Following homologous recombination, the promoter, leader



sequence, foreign gene and flanking viral DNA from the recombinant transfer vector are transferred to the wild-type baculovirus DNA. Recombinant baculovirus is screened for, for example by plaque assay. The recombinant baculovirus can be
5 purified.

The recombinant baculovirus transfer vector which is employed in step (a) also forms part of the invention. Such a transfer vector may be prepared by cloning a DNA sequence
10 consisting essentially of a coding sequence and the leader sequence of the invention into a baculovirus transfer vector. The DNA sequence is cloned into the transfer vector at an appropriate restriction site so that it is under transcriptional control of the promoter that will direct
15 expression of the gene. The transfer vector is preferably pAc36C, in which case the DNA sequence is cloned into the Bam HI site in the vector.

Alternatively, the recombinant baculovirus transfer
20 vector may be prepared from a transfer vector which incorporates the leader sequence. A coding sequence of a heterologous gene may be cloned into the transfer vector at a restriction site immediately downstream of the leader sequence. A preferred transfer vector of this type comprises:
25 (i) a restriction site into which a heterologous gene may be cloned downstream of the sequence encoding the N-terminus of the polyhedrin gene;
(ii) another triplet of nucleotides in place of the natural ATG translation start codon of the polyhedrin
30 coding sequence;
(iii) from up to the first 24 to up to the first 50 bases of the polyhedrin coding sequence; and
(iv) downstream of (iii) but upstream of (i), an untranslated leader of sequence T A A T C A T C C A C A G G A
35 G A C T T T C T G.

Preferably from up to the first 27 to up to the first 39 bases of the polyhedrin coding sequence are present. Up to the first 33 bases may be present. These bases may be



followed by a linker sequence of up to 10 bases, for example of up to 5 bases, before the leader sequence. Alternatively, no linker sequence may be present. The leader sequence is typically immediately followed by the restriction site (i).

5

The DNA sequence comprising the leader sequence and a coding sequence of a heterologous gene is typically prepared by constructing the leader sequence and fusing the leader sequence to the coding sequence. A heterologous coding sequence for use in the invention may be prepared by synthesis, for example by annealing overlapping oligonucleotides. Alternatively, it may be a cloned or a cDNA sequence.

The following Examples illustrate the invention. In the accompanying drawings:

Figure 1 shows the construction of p36C-TETtac2;

Figure 2 shows (a) the construction of transfer vectors p36CST1 and p36CSTM2 and (b) the oligonucleotide used for oligonucleotide directed mutagenesis on the M13 derivative of pKSST2;

Figure 3 shows the structure of p36CST1;

25

Figure 4 shows the activity of cell lysates infected with v36CSTM2 on days 1, 2, 3 and 4 post-infection in the presence (■) or absence (□) of tissue metalloproteinase inhibitor (TIMP), C denoting control, W denoting wild-type baculovirus, DPM denoting disintegrations per minute and ENZ denoting a control aliquot of pure enzyme.

30

Figure 5 shows the sequences of the oligonucleotides used in the construction of p36CP40 and p36CP40p1;

35

Figure 6 shows the results of sodium dodecyl



sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) of cell infected with vP40, v36CP40pl and v36P60pl (lanes 3 to 5), whilst lanes 1 and 7 contain protein molecular weight markers and lanes 2 and 6: uninfected cells;

Figure 7 shows the sequences of the nucleotides used in the construction of pAc36CHIVK and pAc36CHIVP;

Figure 8 shows SDS-PAGE analysis of lysates of insect cells expressing the integrase protein from HIV-1, in which Lane 1; Lysate of cells infected with vAc36CHIV/K. Lane 2; Lysate of cells infected with vAc36CHIV/P. lane 3; Lysate of cells infected with wild type virus. Lane 4; Lysate of uninfected cells. Lane 5; Protein molecular weight markers (The arrow indicates the HIV integrase band); and

Figure 9 is a restriction map of the 1.2 kb EcoRI-BamHI fragment employed in Example 2, in which the translation start codon is underlined.

EXAMPLE 1: Expression of tetanus toxin fragment C

1. Materials and Methods

Bacterial strains and recombinant DNA techniques

E. coli strain TG1 and plasmid pTETtac2 have been described previously (Makoff et al, Biotechnology 7, 1989, 1043-1046; WO 90/15871). DNA manipulations were performed as described by Maniatis et al ("Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Lab., Cold Spring Harbor, New York, US, 1982).

Baculovirus procedures

Baculovirus recombinants were generated by cotransfection of 1µg AcNPV DNA with 2µg recombinant plasmid DNA (prepared by the method of Birnboim and Doly, Nucl. Acids Res 7, 1979, 1513) into Spodoptera frugiperda cells by calcium phosphate precipitation. After 4 hours, the cells were rinsed once, re-fed with TC100 medium (Flow Labs) and cultured for 3 days at 28°C. Supernatants were

then plaque titrated and screened visually for inclusion negative plaques. Stocks of S. frugiperda cells were maintained in spinner flasks (Techne) at room temperature. Plaque assays were performed by seeding S. frugiperda cells into 35mm petri dishes at 10^6 cells/dish, adhering for 10-15 minutes at room temperature, and then infecting with 100 μ g virus/dish. After one hour at room temperature virus inoculum was removed by aspiration and monolayers overlaid with 1% agarose (2 volumes TC100 plus 1 volume 3% Seaplaque LGT agarose). When set, a further 1ml of TC100 was overlaid on the agarose. After 3 days at 28°C the TC100 was removed and plaques visualised by staining for 1 hour with 1ml/dish of 10% neutral red (BDH) in PBS. Plaque purified clones were expanded in Techne spinner flasks at 28°C.

Immunisation procedures

Mice were immunised with recombinant fragment C in alhydrogel and were tested for immunity to tetanus by toxin challenge as described previously (Fairweather et al, Infect. Immun. 55, 1987, 2541-2545).

Affinity chromatography and ELISA

Fragment C was purified from S. frugiperda extracts by affinity chromatography using monoclonal antibody TT08 (Sheppard et al, Infect. Immun. 43, 710-714, 1984) bound to cyanogen bromide activated Sepharose (Trade Mark) 4B. Fragment C was eluted with 0.1M sodium citrate pH 3.0 and was neutralised by the addition of an equal volume of 0.1M sodium phosphate. Fragment C was assayed by an enzyme-linked immunosorbent assay (ELISA) as described previously (Makoff et al, 1989).

2. RESULTS AND DISCUSSION

Cloning of fragment C into baculovirus

The structural gene for fragment C is present on pTETtac2, an E. coli expression plasmid which directs the synthesis of fragment C at 3-4% total cell protein (Makoff et al, 1989). As fragment C is a cleavage product of tetanus toxin, pTETtac2 directs the synthesis of mature fragment C with a methionine residue at the N-terminus. A 1.4kb BglII-BamHI fragment, containing the entire fragment C structural gene preceded by the sequence

AGATCTTAATCATCCACAGGAGACTTTCTGATG, was cloned into the unique BamHI site of p36C which is a high level baculovirus expression vector (Page, Nucl. Acids Res. 17, 454, 1989).

The resulting plasmid, p36C-TeTtac2 (see Figure 1) was cotransfected with baculovirus AcNPV DNA into S. frugiperda cells by calcium phosphate precipitation. Cells containing recombinant viruses were identified by screening for inclusion negative plaques, and after plaque purification several recombinants were screened for the production of fragment C by SDS-PAGE and Western blotting (data not shown).

One recombinant, BVFC1, was studied further. S. frugiperda cells were infected with BVFC1 at a multiplicity of infection of 0.1. Cells were taken at daily intervals for analysis by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Fragment C was synthesised by BVFC1 infected cells at days 2-4, with the maximum amount being produced on day 3. Fragment C was clearly one of the major polypeptides produced by the cells, at a level only slightly lower than polyhedrin.

S. frugiperda cells were grown in spinner flask culture and were infected with BVFC1 virus at an m.o.i. of 10. After 4 days, cells were harvested, lysed by sonication and a soluble extract was obtained by low and



high speed centrifugation. The cell proteins were then analysed by SDS-PAGE and Western blotting. A band of 5 was one of the major proteins in the cell extract. Its identity as fragment C was confirmed by Western blotting against anti-fragment C sera.

Very high expression of fragment C was thus obtained (about 400µg/ml). The fragment C was purified from the cell extract by affinity chromatography using monoclonal antibody TT08. Fragment C purified in this way was judged to be 80% pure as determined by SDS-PAGE.

Immunogenicity of fragment C expressed by baculovirus

Fragment C from C. tetani or E. coli is able to protect mice from challenge with tetanus toxin. In order to determine whether fragment C produced from baculovirus retained this activity, mice were immunised with the soluble extract of S. frugiperda cells which had been infected with baculovirus BVFC1. Preliminary results indicated that such preparations had immunising activity (data not shown). This test was repeated with immunopurified fragment C followed by challenge with tetanus toxin. The results are shown in Table 1.

Table 1 Immunisation of mice with recombinant fragment C
expressed from Baculovirus

<u>Treatment</u>	<u>Amount(μg)^a</u>	<u>Survivors^b</u>	
		<u>One dose</u>	<u>Two doses</u>
Frag C	1	10	10
Frag C	0.25	10	10
Frag C	0.06	0	10
Frag C	0.01	0	0
PBS		0	0
Frag C ^c		10	10

^aAntigen in alhydrogel was injected into BALB/c mice subcutaneously. The second dose as given 4 weeks after the primary injection.

^bMice (n=10 per group) were challenged with 50 LD₅₀s of tetanus toxin 4 weeks after the last injection.

^cDerived from tetanus toxin.

Mice immunised with one dose of 0.25 μ g or with two doses of 0.06 μ g were fully protected against tetanus challenge, indicating that the fragment C expressed by baculovirus was biologically active in this assay. The levels of fragment C necessary for immunity are similar to those required of fragment C derived from C. tetani or E. coli (Makoff et al, 1989) showing comparable activity of the baculovirus expressed material.

EXAMPLE 2: Expression of the Stromelysin protein

1. General

Expression of fragment C in Example 1 had not been compared with expression of fragment C using a sequence believed optional for eukaryotic translation (a Kozak sequence). The baculovirus expression system, in which expression occurs in insect cells, is classed as higher eukaryotic. The translation rules would be expected to conform to the Kozak rules determined for mammalian genes, although there is no experimental evidence that Kozak rules apply. The leader of the invention employed in Example 1 is very sub-optimal with respect to the consensus Kozak sequence, though.

Experiments were therefore conducted to compare the leader sequence of the invention with Kozak sub-optimal, Kozak optimal and Kozak perfect sequences. The first set of experiments involved expression of the stromelysin protein. Two constructs were made and assayed. The first construct was a recombinant baculovirus termed v36ST1 which was sub-optimal around the translation initiation codon ATG for the Kozak consensus sequence:

v36ST1 5' GATCCGAATTC ATG T 3'

The second construct was a recombinant baculovirus termed v36STM2 which contained the leader sequence of the

invention fused to the translation initiation codon of the stromelysin gene. The infected cells were assayed by determining protease activity. The protease activity of cells infected with v36ST1 was indistinguishable from background activity (about 400 units). The activity of uninfected cells and cells infected with wild-type virus was similarly low (200-300 units). In contrast, cells infected with v36STM2 had a protease activity of about 8000 units, i.e. a stimulation of 20-fold. In more detail:

2. Construction of transfer vectors p36CST1 and p36CSTM2

The transfer vectors, p36CST1 and p36CSTM2, containing the engineered stromelysin gene were derived from pAc36C (Page, 1989). The entire stromelysin gene (Whitham *et al*, Biochem. J. 240, 913-916, 1986; EMBL database, code HSSTROMR) was synthesised as a 1.2kb EcoRI - BamHI fragment (Figure 9), which was cloned into the pUC18 plasmid. This fragment was purified by gel electrophoresis and electroelution, ligated to one or other of the two oligonucleotides shown in Figure 2a and then cloned into the BamHI site of Bluescript KS. The resulting plasmids were designated pKSST1 (Figure 2a(i)) and pKSST2 (Figure 2a(ii)).

Plasmid pKSST1 was digested with BamHI and the 1.2kb fragment purified by gel electrophoresis followed by gel electroelution and cloned in the correct orientation into the baculovirus transfer vector pAc36C. This plasmid was called p36CST1 (Figure 3).

3. Construction of p36CSTM2 - mutagenesis of pM13ST2

The 1.2kb BamHI fragment from pKSST2 was purified by gel electrophoresis followed by gel electroelution and cloned into the BamHI site of the replicative form in the antisense orientation of the M13 mp19. This plasmid was called pM13ST2. However, in order to regenerate exactly the sequence of the prokaryotic leader it was necessary to

delete the sequence AATTC positioned before the translation initiation codon of the stromelysin gene. This was done using the Amersham International oligonucleotide mutagenesis kit, and single strand M13 template derived from pM13ST2. The oligonucleotide used is shown in Figure 2b. The sequence position around the translation start of the mutated form, pM13STM2, was confirmed by DNA sequencing. The 1.2kb BamH1 fragment from pM13ST2M was purified from the replicative form of the plasmid and cloned in the correct orientation into the baculovirus transfer vector pAc36C. The resulting plasmid was called pAc36CSTM2.

4. Generation of recombinant virus

Spodoptera frugiperda (Sf21) cells (Vaughen et al, In vitro, 13 213-217, 1977) were grown in suspension cultures at 27°C in TC100 medium (Flow Laboratories), supplemented with 10% foetal bovine serum (Gibco) and 5µg/ml gentamycin. AcNPV was propagated in Sf21 cells grown at 27°C in suspension cultures. Recombinant viruses containing the heterologous gene sequences were generated by co-transfection of a 20:1 molar ratio of the recombinant transfer vectors p36CST1 and p36CSTM2 with AcNPV DNA purified from extracellular virus (ECV) cultures in Sf21 cells. The method of purification of AcNPV DNA was as described (Summers and Smith, 1987, "A Manual of methods for Baculovirus vectors and Insect cell culture procedure"), except that cultures were infected at 0.1 pfu per cell and harvested after 6 days and the sucrose gradient step was omitted. Co-transfection was as described (Summers and Smith 1987, Method 1).

The resultant ECV was screened for recombinants by standard agarose overlay plaque assay (Brown and Faulkner, 1977, J.Gen.Virol, 36, 361-364) using serial dilutions of the transfection culture supernatant. Following neutral red staining, potential recombinant plaques were selected



visually by screening for the absence of polyhedra, picked into 0.5ml of culture medium and purified by repeated plaque assay. Recombinant viruses were scaled up to high titre ($> 10^8$ pfu/ml) by infection of monolayer cultures of Sf21 cells at 27°C. The resulting high titre viruses were called v36CST1 and v36CSTM2.

5. Analysis of proteins from cells infected with the recombinant viruses

The structure of the recombinant viruses described in section 4 was confirmed by Southern blotting. Polyhedrin negative, insert positive plaques were picked into 0.5 mls of culture medium and allowed to diffuse for at least 1 hr. A 100µl aliquot of this was used to infect 2×10^6 Sf21 cells in a 35 mm tissue culture dish, which was incubated at 27°C for 1 hr. The inoculum was removed and replaced with 1 ml of fresh culture medium and the plates incubated at 27°C for three days. The cells were then washed in phosphate buffered saline solution (PBS) and removed.

6. Protease activity of baculovirus expressed stromelysin

The proteins expressed from the recombinant viruses were assayed as follows. Infected cells were lysed with 1% Nonidet P40 (NP40) and 150mM NaCl and the membranous material was removed by centrifugation. The remaining lysate was incubated with tritiated casein (200µg in 100µl Tris buffer - Sigma). The reaction was incubated at 37°C for 2 hours and terminated by the addition of 50µl of a 30% of trichloroacetic acid (TCA). The activity of the stromelysin was determined by measuring the TCA precipitated radioactivity. Figure 4 shows the results of this assay on v36STM2. The values obtained for v36ST1 were virtually indistinguishable from background and are not included.

Surprisingly, presence of tissue metalloproteinase inhibitor (TIMP), an inhibitor of stromelysin, did not

EXAMPLE 3: Expression of Pertussis major surface antigen

Two constructions were initially expressed in baculovirus. The first, called vP60p1, encoded the pertussis major surface antigen (molecular weight = 60kDa) and was expressed in p36C and contained the leader of the present invention. Expression of this protein in insect cells infected with vP60p1 was very efficient (400-500µg/ml). The second construction, vP40, was a truncated form of the major surface antigen. The sequences upstream of this gene were Kozak optimal, though not Kozak perfect, as follows:

Those bases starred (*) in the Kozak consensus are more important in translation. Those in brackets are less important. Expression of the 40kDa protein in cells infected with vP40 was still efficient (about 100µg/ml) but 4-5 fold lower than expression of P60 from vP60pl infected cells.

The experiment that was then done was to make a recombinant virus expressing the 40kDa protein preceded by the same prokaryotic leader as in vp60pl. Expression of the 60kDa protein and the 40kDa protein from the two constructions was determined. Expression of the 40kDa protein was 3-4 fold higher from vp40pl than from vp40 and

approached the levels of protein expressed from vP60pl.

2. Construction of p36CP60pl, p36CP40 and p36CP40pl

The plasmids p36CP40 and p36CP40pl were made by fusing the pertussis major surface antigen (PMSA) gene body lacking its N-terminus to synthetic oligonucleotides containing various translation leaders and the N-terminus of the PMSA. The sequence of oligonucleotides used in the construction of the two plasmids is shown in Figure 5.

(i) Construction of p36CP60pl

The 1.8kb BglII-BamHI fragment from pPertac 7, containing the PMSA coding sequence fused to the leader of the invention, was purified by gel electrophoresis and gel electroelution and cloned into the Bam HI site of the baculovirus transfer vector pAc36C (Page, 1989). The resulting plasmid was called p36CP60pl.

(ii) Construction of pAc36CP40 and pAc36CP40pl

The 4.6kbp BglII - SacI fragment from pPertac 9 (derived from pPertac 7, Makoff et al.) was purified by electrophoresis followed by electroelution. An oligonucleotide pair containing a Kozak optimal sequence was cloned into this vector as a BglII-SacI fragment. After sequencing, the 1.1 kbp BglII - BamHI fragment was cloned in the correct orientation into the Bam HI site of pAc36C. The resulting plasmid was called pAc36CP40.

Plasmid pAc36CP40pl was constructed in two stages. The 1.1kbp BglII - BamHI fragment of pPertac 7 (Makoff et al., Biotechnology 8, 1990, 1030-1033) was purified and cloned into the BamHI site of pUC9 in the antisense orientation. This plasmid was called pUC69-6. A pair of oligonucleotides, containing the leader of the invention and the N-terminal sequence of the 40kDa protein, was cloned as a BamHI - HindIII fragment into the replication form of M13 mp19. The sequence of the oligonucleotide was confirmed by DNA sequencing. The BamHI-HindIII insert from the replicative form of this construction was purified and cloned between the BamHI and HindIII sites of the



pUC9 plasmid. The resulting plasmid was designated pUC9-3.

The pUC9-3 plasmid was digested with Sac I and Hind III and the 3.0kb fragment, containing the leader of the invention, the N-terminal sequence of the 40 kDa protein and the pUC9 vector body, was isolated. The 1.1kbp Sac I - Hind III fragment, containing the C-terminal portion of the 40kDa protein, was purified and cloned into the Sac I - Hind III digested pUC9-3. The resulting plasmid was designated pUC 9-1. This plasmid contained the entire coding sequence for the 40kDa truncated PMSA fused to the leader and flanked by two Bam HI sites. The 1.2kb Bam HI fragment from pUC 9-1 was purified and cloned into the Bam HI sites of pAc36C. The resulting plasmid was designated pAc36CP40pl.

3. Generation of recombinant virus and analysis of proteins from cells infected with the recombinant viruses

This was performed as described in Example 2. The resulting high titre viruses were called v36CP60pl, v36CP40 and v36CP40pl. Infected cells were lysed with 5% SDS and the products analysed by SDS/polyacrylamide electrophoresis on 10% gels followed by Coomassie blue staining. The results are shown in Figure 6.

The identity and antigenicity of these products was confirmed by Western blotting using a polyclonal rabbit antisera (not shown). The 60kDa product (lane 5) actually migrated as a 69kDa band. The reason for this anomalous migration is unclear. The predicted molecular weight as determined from the amino acid sequence is 60kDa. The 60kDa product shows some breakdown to the 40kDa product which cross-reacts with the antibody. Both the 40kDa and 60kDa proteins are stable as determined by pulse-chase experiments (data not shown). The levels of expression were as indicated above, demonstrating the superiority of the use of the leader sequence of the invention.

EXAMPLE 4: Expression of the human immunodeficiency virus (HIV) integrase protein

1. General

Two recombinant baculoviruses were prepared and assayed. The first, v36HIVK, contained the integrase gene body fused to a perfect Kozak leader sequence. The second, v36HIVP, contained the integrase gene body fused to the leader sequence of the invention. The level of expression of the integrase protein by v36HIVP was 4-5 fold better than the level of expression of the integrase protein by v36HIVK.

2. Construction of plasmids pAc36CHIVK and pAc36CHIVP

10 The fragment of DNA containing the integrase gene in pAc36CHIVK and pAc36CHIVP was derived from the plasmid pINpUC19. This construction contained the HIV integrase protein coding sequence on a 1.6kbp Kpn I - Hind III fragment (Sherman and Fyfe, PNAS USA, 87, 5119-5123, 1990). A Nde I
15 site had been introduced by site directed mutagenesis immediately adjacent to the N-terminal phenylalanine of the HIV integrase protein. Plasmids pAc36CHIVP and pAc36CHIVK were constructed by purifying the 898bp Nde I fragment from pINpUC19 by gel electrophoresis followed by gel electroelution. The
20 fragment was then ligated to one of the pairs of oligonucleotides shown in Figure 7 and cloned into the Bam HI site of pAc36C. The construction containing the leader of the invention fused to the ATG was called pAc36CHIVP (Figure 7a) whilst that containing the perfect Kozak consensus sequence was
25 called pAc36CHIKK.

3. Generation of recombinant virus and analysis of proteins from cells infected with the recombinant viruses

This was performed as described in Example 2. The resulting high titre viruses were called v36HIVK and v36HIVP.
30 Infected cells (2×10^6 , 72 hpi) were lysed with 100 μ l of 150mM Tris, 5% SDS, 25% glycerol and 0.1% w/v bromophenol blue. Samples of these lysates were analysed on 10% SDS-PAGE gels (Figure 8). Cells infected with v36CHIVP expressed the HIV integrase protein at around 5 fold higher levels than cells
35 infected with v36HIVK.



THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A process for preparing a protein, which process comprises culturing, under such conditions that the protein is obtained, insect cells infected with a recombinant
5 baculovirus having a promoter capable of directing expression in the cells of a heterologous gene having a coding sequence encoding the protein, the promoter being operably linked to the heterologous gene and the coding sequence being preceded by an untranslated leader of sequence T A A T C A T C C A C A G G
10 A G A C T T T C T G.
2. A process according to claim 1, wherein the leader is immediately preceded by the sequence A T C C.
3. A process according to claim 1 or 2, wherein the leader is fused to the translation initiation codon of the
15 coding sequence.
4. A process according to any one of the preceding claims, wherein the coding sequence encodes tetanus toxin fragment C, stromelysin, pertussis major surface antigen or the HIV integrase protein.
- 20 5. A process according to any one of the preceding claims, wherein the promoter is the polyhedrin promoter.
6. A process according to any one of the preceding claims, wherein the insect cells are cells of a
25 Spodoptera frugiperda cell line.
7. A process according to any one of the preceding claims, further comprising isolating and purifying the protein.
8. A recombinant baculovirus having a promoter
30 capable of directing expression of a heterologous gene having a coding sequence in insect cells infected with the recombinant baculovirus, the promoter being operably linked to the heterologous gene and the coding sequence being preceded by an untranslated leader of sequence T A A T C A T C C A C A G G A G
35 A C T T T C T G.
9. A recombinant baculovirus according to claim 8, wherein the leader is immediately preceded by the sequence A T C C.
10. A recombinant baculovirus according to claim



8 or 9, wherein the leader is fused to the translation initiation codon of the coding sequence.

11. A recombinant baculovirus according to any one of claims 8 to 10, wherein the coding sequence encodes
5 tetanus toxin fragment C, stromelysin, pertussis major surface antigen or the HIV integrase protein.

12. A recombinant baculovirus according to any one of claims 8 to 11, wherein the promoter is a polyhedrin promoter.

10 13. A process for preparing a recombinant baculovirus, which process comprises:

(a) providing a recombinant baculovirus transfer vector having a promoter capable of directing expression of a heterologous gene in insect cells infected with the recombinant
15 baculovirus, the promoter being operably linked to the heterologous gene and the coding sequence of the heterologous gene being preceded by an untranslated leader of sequence T A A T C A T C C A C A G G A G A C T T T C T G; and

(b) co-transfecting the cells with the
20 recombinant baculovirus transfer vector and with intact wild-type baculovirus DNA.

14. A process according to claim 13, wherein the leader is immediately preceded by the sequence A T C C.

15. A process according to claim 13 or 14,
25 wherein the leader is fused to the translation initiation codon of the coding sequence.

16. A process according to any one of claims 13 to 15, wherein the coding sequence encodes tetanus toxin fragment C, stromelysin, pertussis major surface antigen or the
30 HIV integrase protein.

17. A process according to any one of claims 13 to 16, wherein the promoter is the polyhedrin promoter.

18. A process according to any one of claims 13 to 17, wherein the insect cells are cells of a Spodoptera
35 frugiperda cell line.

19. A recombinant baculovirus transfer vector as defined in any one of claims 13 to 17.

20. A DNA sequence consisting essentially of a coding sequence preceded by an untranslated leader of sequence

T A A T C A T C C A C A G G A G A C T T T C T G.

21. A DNA sequence according to claim 20, wherein the leader is immediately preceded by the sequence A T C C.

22. A DNA sequence according to claim 20 or 21,
5 wherein the leader is fused to the translation initiation codon of the coding sequence.

23. A DNA sequence according to any one of claims 20 to 22, wherein the coding sequence encodes tetanus toxin fragment C, stromelysin, pertussis major surface antigen or the
10 HIV integrase protein.

24. A process for preparing a recombinant baculovirus transfer vector, which process comprises cloning a DNA sequence as claimed in any one of claims 20 to 23 into a baculovirus transfer vector.

25. A process according to claim 24, wherein the transfer vector is pA36C.

26. A baculovirus transfer vector comprising
(i) a restriction site into which a heterologous gene may be cloned downstream of the sequence encoding the N-
20 terminus of the polyhedrin gene;

(ii) another triplet of nucleotides in place of the natural ATG translation start codon of the polyhedrin coding sequence;

(iii) from up to the first 24 to up to the first 50
25 bases of the polyhedrin coding sequence; and

(iv) downstream of (iii) but upstream of (i), an untranslated leader of sequence T A A T C A T C C A C A G G A G A C T T T C T G.

27. A process for preparing a recombinant
30 baculovirus transfer vector, which process comprises cloning a heterologous gene into the restriction site (i) of a baculovirus transfer vector as claimed in claim 26.

28. A process for preparing a protein, which process is substantially as described hereinbefore in Example 2
35 with reference to recombinant baculovirus v36STM2, Example 3 with reference to recombinant baculovirus v36CP60pl or v36CP40pl, or Example 4 with reference to recombinant baculovirus v36CIntP.

29. A recombinant baculovirus substantially as

described hereinbefore in Example 2 with reference to recombinant baculovirus v36STM2, Example 3 with reference to recombinant baculovirus v36CP60pl or v36CP40pl, or Example 4 with reference to recombinant baculovirus v36CHIVP.

5 30. A process for preparing a recombinant baculovirus, which process is substantially as described in Example 1 with reference to recombinant baculovirus BVFC1, Example 2 with reference to recombinant baculovirus v36STM2, Example 3 with reference to recombinant baculovirus v36CP60pl
10 or v36CP40pl, or Example 4 with reference to recombinant baculovirus v36HIVP.

 31. A baculovirus transfer vector substantially as described in Example 1 with reference to p36C-TETtac2, Example 2 with reference to p36CSTM2, Example 3 with reference
15 to p36CP60pl or pAc36CP40pl, or Example 4 with reference to pAc36CHIVP.

Dated this 23rd day of December 1993

THE WELLCOME FOUNDATION LIMITED
By their Patent Attorney
GRIFFITH HACK & CO.



ABSTRACT
EXPRESSION SYSTEM

A process for the preparation of a protein, which process comprises culturing under such conditions that the
5 protein is obtained cells infected with a recombinant baculovirus having a promoter, capable of directing the expression of a gene in the cells, operably linked to a heterologous gene which encodes the protein and which is preceded by a leader sequence comprising:

10 T A A T C A T C C A C A G G A G A C T T T C T G.

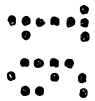
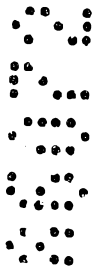


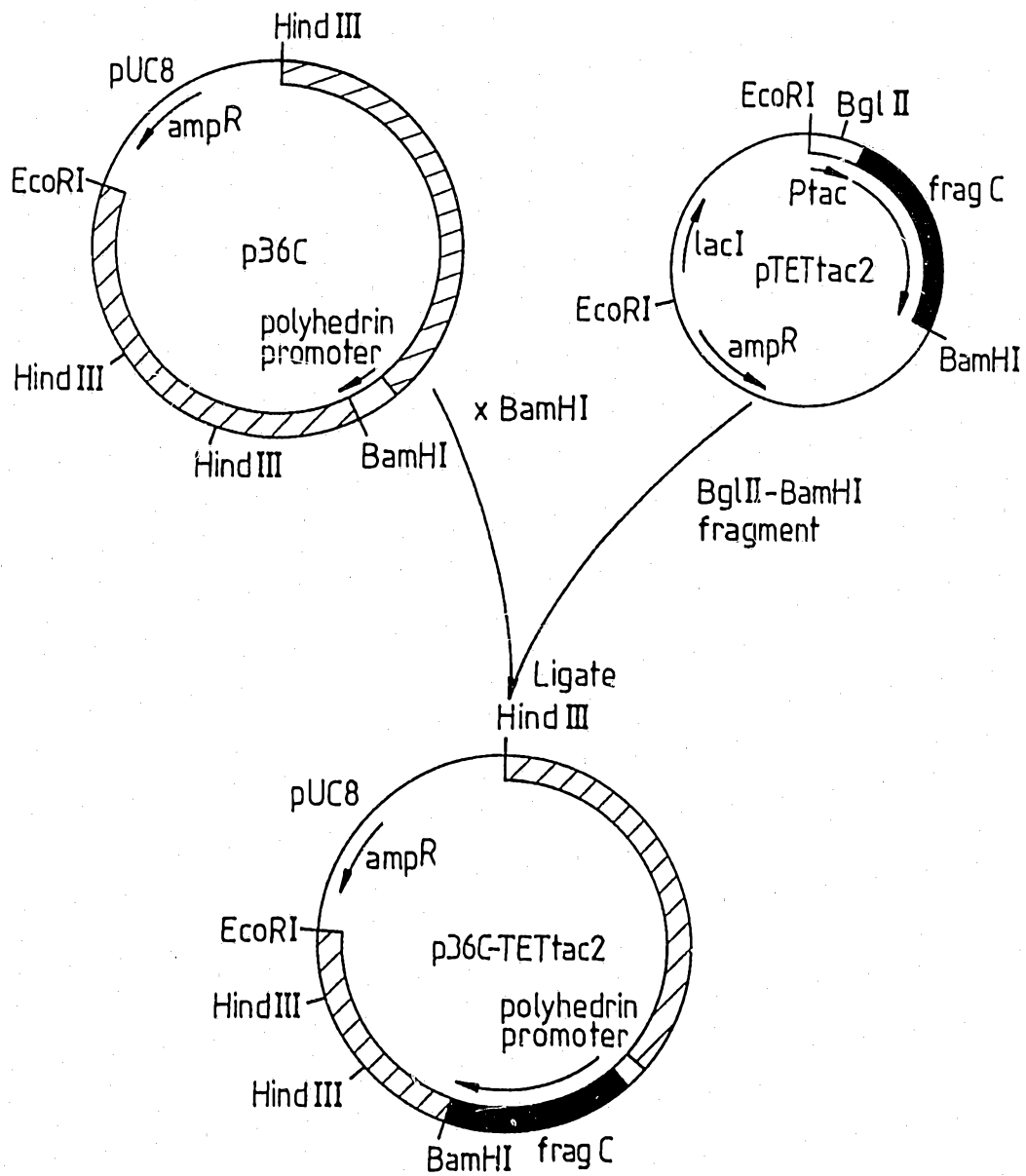
Fig.1.

Fig. 2.

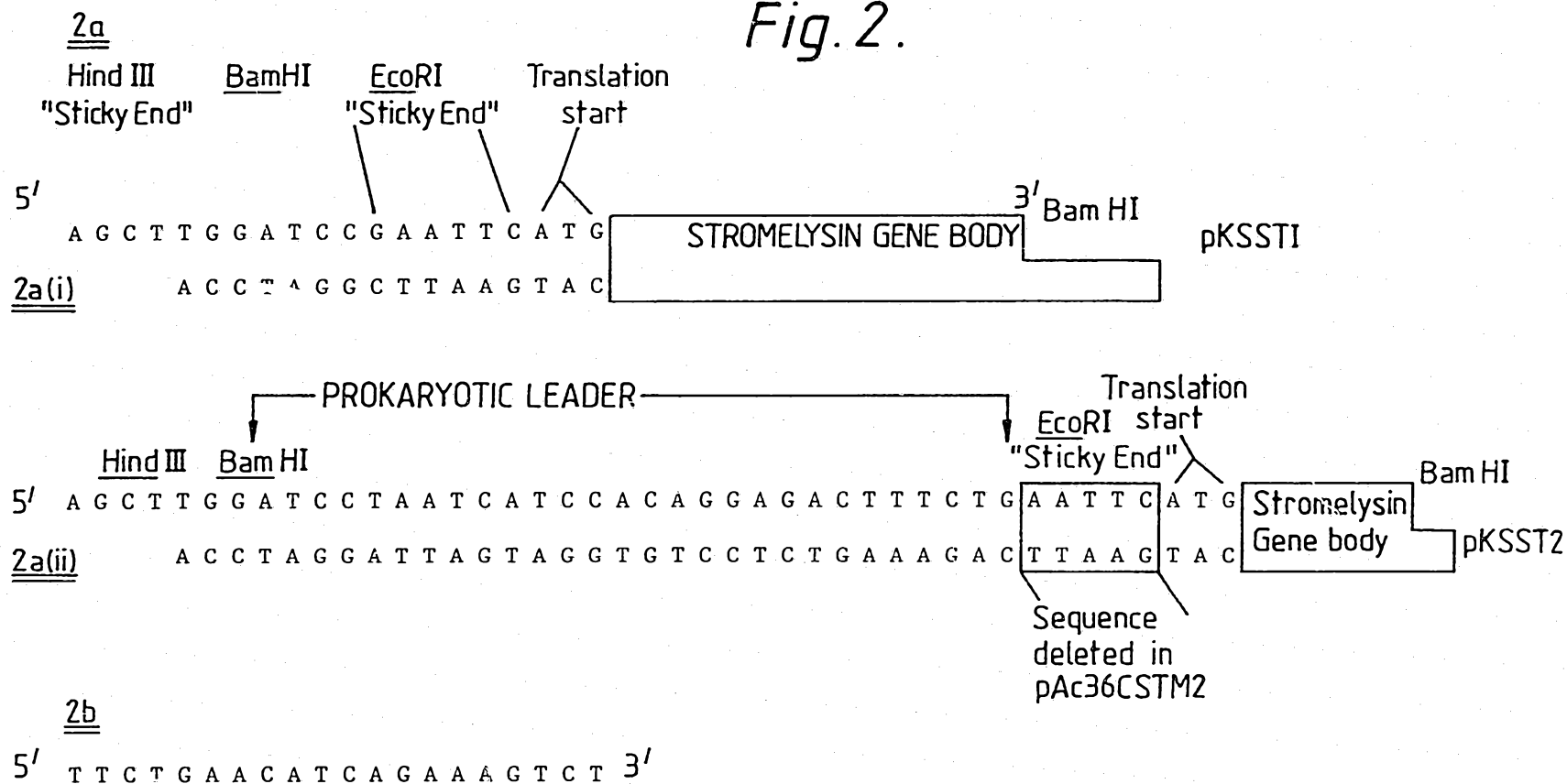


Fig.3.

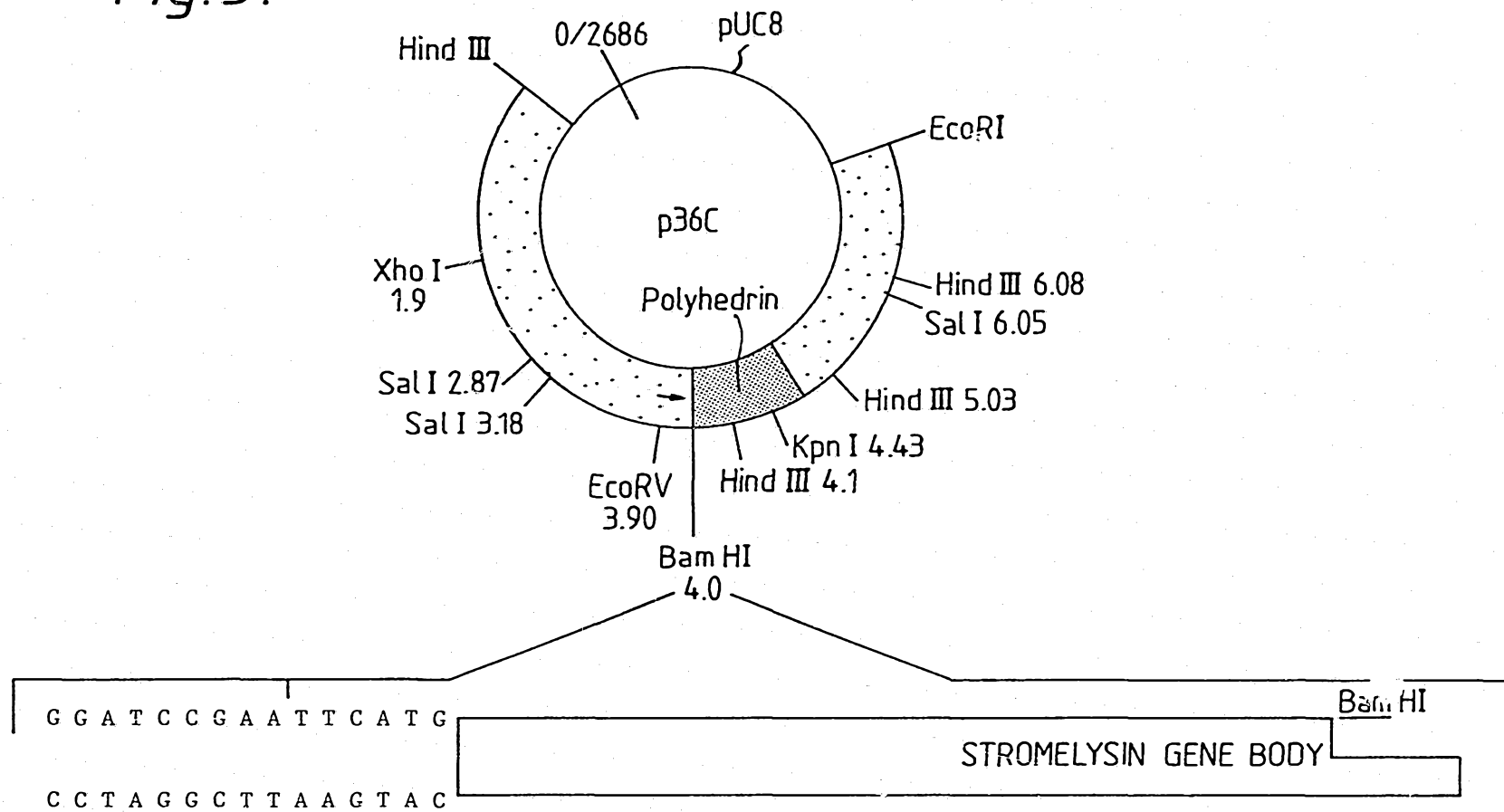


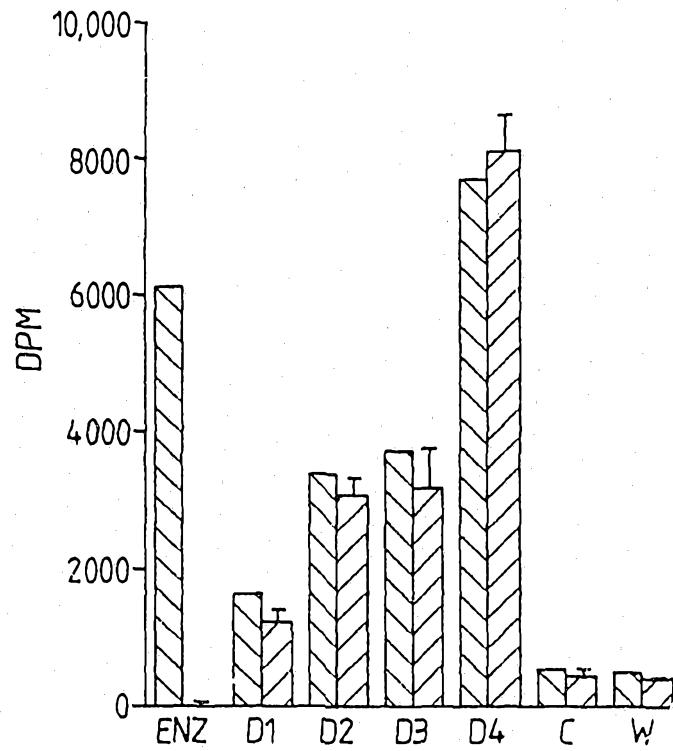
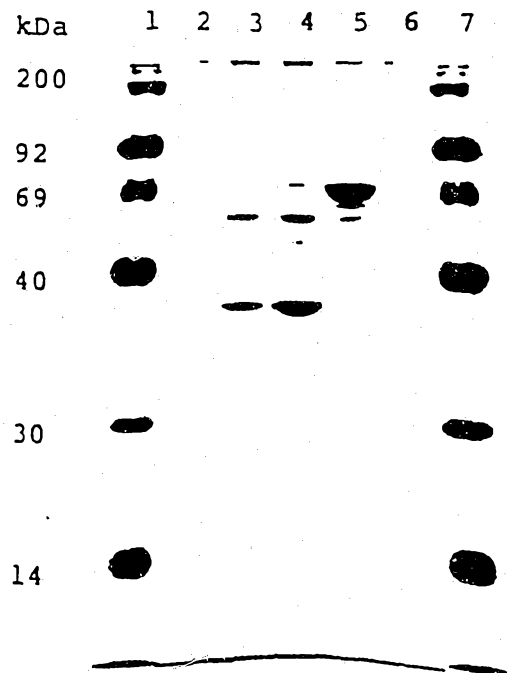
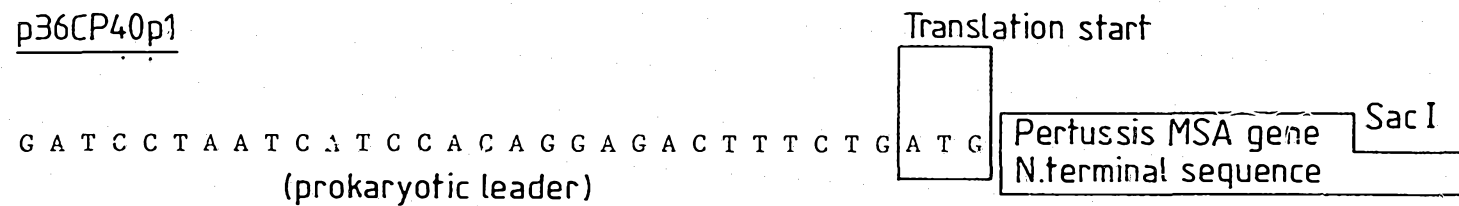
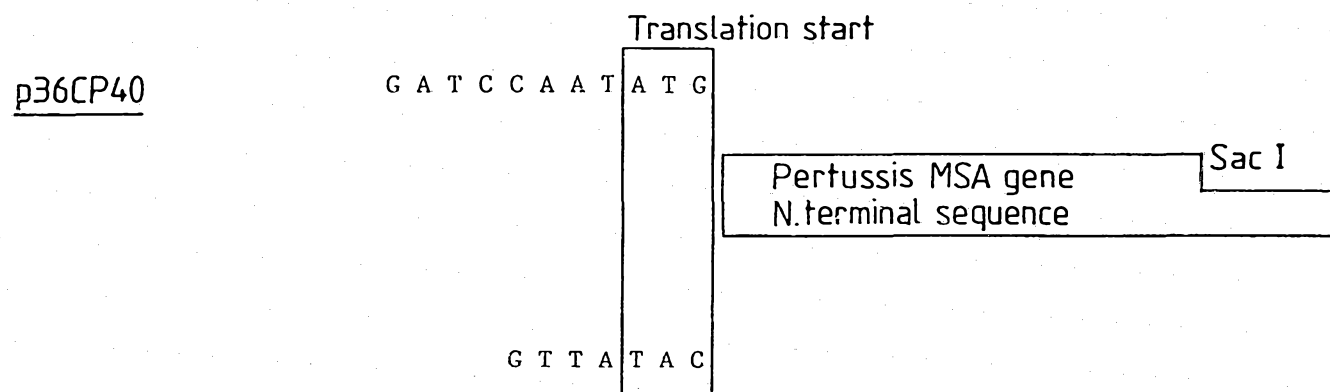
Fig. 4.*Fig. 6.*

Fig. 5.



1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 26

1. *Journal of the American Medical Association*, 1997; 278: 1039-1044.

1. 2. 3.

100

Fig. 8.

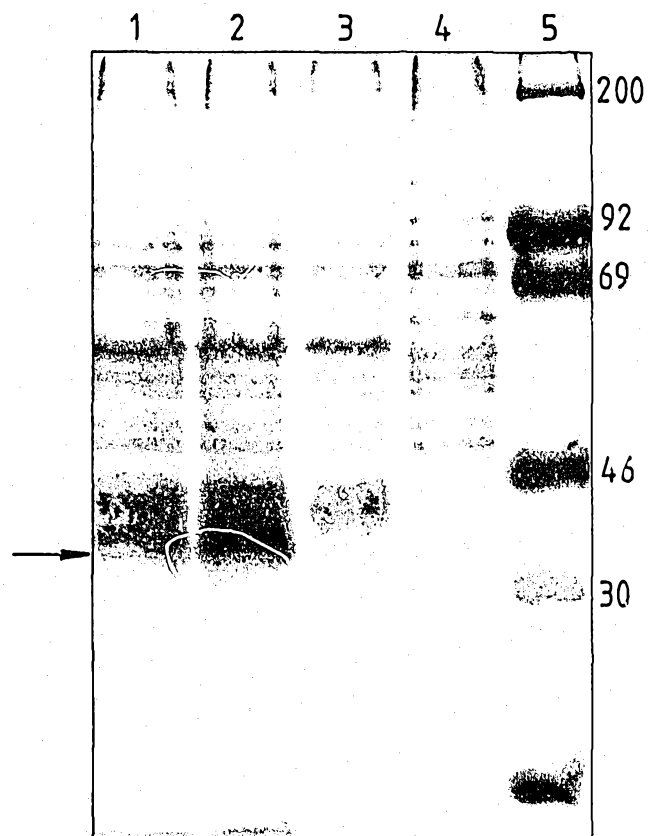


Fig. 9.

10	20	30	40	50	60
GAATTCATGT	TCAGAACCTT	TCCTGGCATC	CCGAAGTGGA	GGAAAACCCA	CCTTACATAC
E N		B SF	M		
C L		S FO	N		
O A		T AK	L		
1 3		1 11	1		

70	80	90	100	110	120
AGGATTGTGA	ATTATACACC	AGATTTGCCA	AAAGATGCTG	TTGATTCTGC	TGTTGAGAAA
			S	H	A
			F	I	L
			A	N	U
			1	1	1

130	140	150	160	170	180
GCTCTGAAAG	TCTGGGAAGA	GGTGACTCCA	CTCACATTCT	CCAGGCTGTA	TGAAGGAGAG
	M I' MP	HM M		B	M
	N B LL	PA L		S	N
	L O YE	HE Y		T	L
	1 2 11	13 1		1	1

190	200	210	220	230	240
GCTGATATAA	TGATCTCTTT	TGCAGTTAGA	GAACATGGAG	ACTTTTACCC	TTTTGATGGA
	S		BN		A
	A		SL		V
	U		MA		A
	A		13		2

250	260	270	280	290	300
CCTGGAAATG	TTTTGGCCCA	TGCCTATGCC	CCTGGGCCAG	GGATTAATGG	AGATGCCCCAC
B	S N		S E SH S	AM	S
S	A L		E C AA E	SS	F
T	U A		C O UE C	EE	A
1	6 3		1 1 63 1	11	1

310	320	330	340	350	360
TTTGATGATG	ATGAACAATG	GACAAAGGAT	ACAACAGGGA	CCAATTTATT	TCTCGTTGCT
			NA	B	F
			LV	B	N
			AA	V	U
			42	1	1

370	380	390	400	410	420
GCTCATGAAA	TTGGCCACTC	CCTGGGTCTC	TTTCACTCAG	CCAACACTGA	AGCTTTGATG
BN	MH P	S S	B	D	T
SL	SA F	E E	S	D	T
PA	CE L	C C	M	E	H
13	13 1	1 1	1	1	2
					3 1
					1

Fig. 9(Cont).

430	440	450	460	470	480
TACCCACTCT	ATCACTGACT	CACAGACCTG	ACTCGGTTCC	GCCTGTCTCA	AGATGATATA
		MP	M	N	B
		LL	L	L	S
		YE	Y	A	M
		11	1	4	1
490	500	510	520	530	540
AATGGCATTC	AGTCCCCTCTA	TGGACCTCCC	CCTGACTCCC	CTGAGACCCC	CCTGGTACCC
B	M	AD M	MP	M	B
S	N	VR N	LL	L	S
M	L	AD L	YE	Y	M
1	1	21 1	11	1	1
550	560	570	580	590	600
ACGGAACCTG	TCCCTCCAGA	ACCTGGGACG	CCAGCCAACT	GTGATCCTGC	TTTGTCCCTT
N	M	PA	S	B	
L	N	FL	E	B	
A	L	LW	C	I	
4	1	11	1	1	
610	620	630	640	650	660
GATGCTGTCA	GCACTCTGAG	GGGAGAAATC	CTGATCTTTA	AAGACAGGCA	CTTTTGGCGC
S	M	D	S	DM	H
F	N	D	A	RS	I
A	L	E	U	AE	N
1	1	1	A	11	1
670	680	690	700	710	720
AAATCCCTCA	GGAAGCTTGA	ACCTGAATTG	CATTTGATCT	CTTCATTTTG	GCCATCTCTT
MD	HA		S	EM	EM
ND	IL		A	AB	AB
LE	NU		U	RO	RO
11	31		A	12	12
730	740	750	760	770	780
CCTTCAGGCG	TGGATGCCGC	ATATGAAGTT	ACTAGCAAGG	ACCTCGTTTT	CATTTTTTAAA
	FS	F	N	M	DM
	OF	N	D	A	RS
	KA	U	E	E	AE
	11	1	1	12	11
790	800	810	820	830	840
GGAAATCAAT	TCTGGGCCAT	CAGAGGAAAT	GAGGTACGAG	CTGGATACCC	AAGAGGCATC
	SHM	M	R	A	SF
	AAN	N	S	L	FO
	UEL	L	A	U	AK
	631	1	1	1	11

10/10

Fig.9(Cont).

850	860	870	880	890	900
CACACCCTAG	GTTTCCCTCC	AACCGTGAGG	AAAATCGATG	CAGCCATTTC	TGATAAGGAA
AM	M M		CT S B		
VA	N N		LA F B		
RE	L L		AQ A V		
21	1 1		11 1 1		
910	920	930	940	950	960
AAGAACAAAA	CATATTTCTT	TGTAGAGGAC	AAATACTGGA	GATTTGATGA	GAAGAGAAAT
	M		B	E	M
	N		S	A	B
	L		R	R	O
	1		1	1	2
970	980	990	1000	1010	1020
TCCATGGAGC	CAGGCTTTCC	CAAGCAAATA	GCTGAAGACT	TTCCAGGGAT	TGACTCAAAG
NN N B		A X		S MP	M
CL L S		L M		E LL	L
OA A T		U N		C YE	Y
13 4 1		1 1		1 11	1
1030	1040	1050	1060	1070	1080
ATTGATGCTG	TTTTTGAAGA	ATTGGGGTTC	TTTTIATTTCT	TTACTGGATC	TTCACAGTTG
S	M			B AS M	
F	B			S LA B	
A	O			R WU O	
1	2			1 1A 2	
1090	1100	1110	1120	1130	1140
GAGTTTGACC	CAAATGCAAA	GAAAGTGACA	CACACTTTGA	AGAGTAACAC	CTGGCTTAAT
		M	E M	M PA	M
		A	A B	A VL	S
		E	R O	E UU	E
		3	1 2	3 21	1
1150	1160				
TGTTGATAAC	TGCAGGATCC				
P A	BB				
S L	AI				
T W	MN				
1 1	11				