

FORM 1

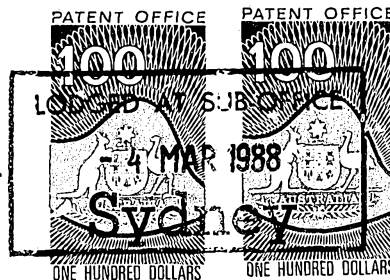
609783

SPRUSON & FERGUSON

COMMONWEALTH OF AUSTRALIA

PATENTS ACT 1952

APPLICATION FOR A STANDARD PATENT



F Hoffmann-La Roche & Co Aktiengesellschaft, of Grenzacherstrasse 124-184, 4002 Basle, SWITZERLAND, hereby apply for the grant of a standard patent for an invention entitled:

Novel Fusion Proteins and their Purification

which is described in the accompanying complete specification.

Details of basic application(s):-

Basic Applic. No: Country:

895/87-2

SWITZERLAND

Application Date

10 March 1987

The address for service is:-

Spruson & Ferguson
Patent Attorneys
Level 33 St Martins Tower
31 Market Street
Sydney New South Wales Australia

DATED this THIRD day of MARCH 1988

F Hoffmann-La Roche & Co Aktiengesellschaft

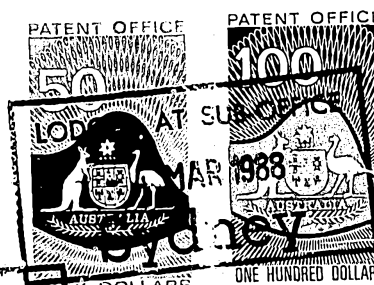
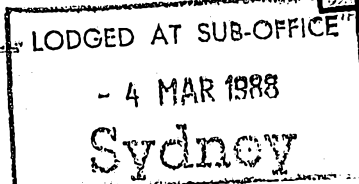
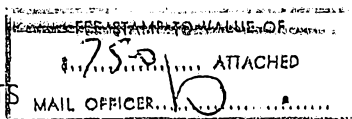
By:

W. J. Anderson

Registered Patent Attorney

TO: THE COMMISSIONER OF PATENTS
OUR REF: 50012
S&F CODE: 55541
APPLICATION ACCEPTED AND AMENDMENTS

ALLOWED 19-2-91
5845/2



COMMONWEALTH OF AUSTRALIA
THE PATENTS ACT 1952
DECLARATION IN SUPPORT OF A
CONVENTION APPLICATION FOR A PATENT

AUSTRALIA
CONVENTION
STANDARD
& PETTY PATENT
DECLARATION

In support of the Convention Application made for a
patent for an invention entitled:

RAM 4100/68

Title of Invention

Novel fusion proteins and their purification

Full name(s) and
address(es) of
Declarant(s)

I Fridolin Klausner

of 187 Baselmattweg, 4123 Allschwil, Switzerland

do solemnly and sincerely declare as follows:—

Full name(s) of
Applicant(s)

1. I am authorised by
F. HOFFMANN-LA ROCHE & CO. Aktiengesellschaft,
of 124-184 Grenzacherstrasse, Basle, Switzerland,

Basic Country(ies)

in Switzerland

Priority Date(s)

on March 10, 1987

Basic Applicant(s)

by ☒ F. HOFFMANN-LA ROCHE & CO., Aktiengesellschaft

☐ the inventor(s) cited in paragraph 3.

Full name(s) and
address(es) of
inventor(s)

3.

- 1) Heinz Döbeli, 22 Lupsingerstrasse,
4417 Ziefen, Switzerland.
- 2) Bernhard Eggimann, 83 Eulerstrasse,
4051 Basle, Switzerland
- 3) Reiner Gentz, 34 Am Hochgericht,
7888 Rheinfelden, Germany
- 4) Erich Hochuli, 25 Kirchackerstrasse,
4411 Arisdorf, Switzerland
- 5) Dietrich Stüber, 9 Bandweg, 7889 Grenzach-Wyhlen,
Germany

(respectively)

is/are the actual inventor(s) of the invention and the facts upon
which the applicant(s) is/are entitled to make the application are
as follows:

Set out how Applicant(s)
derive title from actual
inventor(s) e.g. The
Applicant(s) is/are the
assignee(s) of the
invention from the
inventor(s)

- () the inventor(s) have assigned the invention to
Hoffmann-La Roche Inc., Nutley, USA,
who have re-assigned all their rights for Australia
to the Applicant.
- (x) the Applicant is the assignee of the invention from
the inventor(s).
4. The basic application(s) referred to in paragraph 2 of this
Declaration was/were the first application(s) made in a Convention
country in respect of the invention(s) the subject of the application.

To:

Declared at Basle this 29th day of January, 1988

The Commissioner of Patents,
COMMONWEALTH OF AUSTRALIA

Fridolin Klausner
Signature of Declarant(s)

(12) PATENT ABRIDGMENT (11) Document No. AU-B-12709/88
(19) AUSTRALIAN PATENT OFFICE (10) Acceptance No. 609783

(54) Title
NOVEL FUSION PROTEINS AND THEIR PURIFICATION

International Patent Classification(s)

(51) ⁴	C07K 003/20	A61K 037/26	A61K 037/36	A61K 037/43
	A61K 039/21	C07G 017/00	C07H 021/04	C07K 013/00
	C07K 015/26	C07K 017/08	C12N 001/20	C12N 001/21
	C12N 015/00	C12N 015/70	C12P 001/04	C12P 021/02
	A61K 039/015	A61K 039/385		
(51) ⁵	A61K 037/66	C12N 015/62		

(21) Application No. : 12709/88

(22) Application Date : 04.03.88

(30) Priority Data

(31) Number	(32) Date	(33) Country
895/87	10.03.87	CH SWITZERLAND

(43) Publication Date : 15.09.88

(44) Publication Date of Accepted Application : 09.05.91

(71) Applicant(s)
F HOFFMANN-LA ROCHE & CO AKTIENGESELLSCHAFT

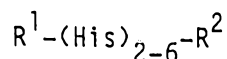
(72) Inventor(s)
HEINZ DOBELI; BERNHARD EGGIMANN; RAINER GENTZ; ERICH HOCHULI; DIETRICH STUBER

(74) Attorney or Agent
SPRUSON & FERGUSON, GPO Box 3898, SYDNEY NSW 2001

(57) Claim

1. Fusion proteins consisting of one or two affinity peptides which contain 2-6 adjacent histidine residues and a biologically active polypeptide or protein covalently linked directly or indirectly to this/these affinity peptide(s).

2. Fusion proteins in accordance with claim 1, wherein the affinity peptide(s) has/have the formula



wherein R^1 represents hydrogen or 1 to 4 amino acids, R^2 represents Q, Q-Ile-Glu-Gly-Arg- or Q-Asp-Asp-Asp-Asp-Lys- wherein Q is a peptide bond, an amino acid or a sequence of 2-30 amino acids and wherein one affinity peptide is covalently linked directly or indirectly to the amino terminal amino acid or to the carboxy terminal amino acid of the biologically active peptide or protein.

10. Genes which code for a fusion protein in accordance with any one of claims 1 to 8.

11. Expression vectors in which a gene in accordance with claim 10 is operatively covalently linked to an expression control sequence.

14. A bacterium transformed with an expression vector in accordance with any one of claims 11 to 13.

18. A process for the purification of a biologically active polypeptide or protein, which process comprises purifying it as a fusion protein in accordance with any one of claims 1 to 8 by means of a process comprising affinity chromatography by NTA resins and then removing the affinity peptide(s) by selective cleavage.

609783

S & F Ref: 50012

FORM 10

COMMONWEALTH OF AUSTRALIA

PATENTS ACT 1952

COMPLETE SPECIFICATION

(ORIGINAL)

This document contains the
amendments made under
Section 49 and is correct for
printing

FOR OFFICE USE:

Class Int Class

Complete Specification Lodged:
Accepted:
Published:

Priority:

Related Art:

Name and Address
of Applicant:

F Hoffmann-La Roche & Co Aktiengesellschaft
Grenzacherstrasse 124-184
4002 Basle
SWITZERLAND

Address for Service: Spruson & Ferguson, Patent Attorneys
Level 33 St Martins Tower, 31 Market Street
Sydney, New South Wales, 2000, Australia

Complete Specification for the invention entitled:

Novel Fusion Proteins and their Purification

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

RAN 4100/68

5

Abstract

10 Fusion proteins as well as a process for their purification by means of metal chelate affinity chromatography on NTA resins.

15

20

25

30

35

RAN 4100/68

5 The possibilities of preparing hybrid genes by gene
technology open up new routes for the working-up of
recombinant proteins. By linking the coding gene sequence
of a desired protein with the coding gene sequence of a
protein fragment having a high affinity for a ligand
10 (affinity peptide) it is possible to purify desired recom-
binant proteins in the form of fusion proteins in one step
using the affinity peptide. By site-directed mutagenesis
it is also possible to introduce specific chemical or
enzymatic cleavage sites at the point of linkage of the
15 affinity peptide and the desired recombinant protein, so
that after the purification of the fusion protein by means
of a suitable affinity resin the desired recombinant
protein can be recovered by chemical or enzymatic cleav-
age. Such purification methods are known, for example,
20 from Science 198, 1056-1063 (1977) (Itakura et al.), Proc.
Natl. Acad. Sci. U.S.A. 80, 6848-6852 (1983) (Germino et
al.), Nucleic Acids Res. 13, 1151-1162 (1985) (Nilsson et
al.), Gene 32, 321-327 (1984) (Smith et al.), as well as
from European Patent Applications publication Nos. 150 126
25 and 184 355.

30 In accordance with the invention it has now been found
that affinity peptides having at least two neighbouring
histidine residues are especially suitable for the purifi-
cation of recombinant proteins by means of metal chelate
affinity chromatography on nitrilotriacetic acid (NTA)
resins. These affinity peptides are distinguished
vis-a-vis the known peptides primarily in that they permit
the problem-free purification of native and denatured
35 proteins by means of NTA resins.

The present invention is therefore concerned with fusion proteins which consist of one or two affinity peptides, which contain neighbouring histidine residues, and a biologically active polypeptide or protein linked directly or indirectly to this/these affinity peptide(s), a process for their preparation by means of recombinant DNA technology as well as a process for their purification by means of metal chelate affinity chromatography on NTA resins. The present invention is also concerned with genes which code for these fusion proteins, expression vectors which contain these genes, microorganisms transformed with these expression vectors as well as a process for the preparation of said genes, expression vectors and transformed microorganisms.

According to a first embodiment of this invention there is provided fusion proteins consisting of one or two affinity peptides which contain 2-6 adjacent histidine residues and a biologically active polypeptide or protein covalently linked directly or indirectly to this/these affinity peptide(s).

According to a second embodiment of this invention there is provided a bacterially prepared fusion protein.

According to a third embodiment of this invention there is provided genes which code for a fusion protein.

According to a fourth embodiment of this invention there is provided expression vectors in which a gene is operatively covalently linked to an expression control sequence.

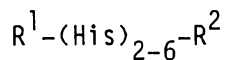
According to a fifth embodiment of this invention there is provided a bacterium transformed with an expression vector.

According to a sixth embodiment of this invention there is provided a process for the preparation of fusion proteins, which process comprises transforming a bacterial host with an expression vector, culturing the transformant under appropriate conditions of growth and isolating said fusion proteins.

According to a seventh embodiment of this invention there is provided a process for the purification of a biologically active polypeptide or protein, which process comprises purifying it as a fusion protein by means of a process comprising affinity chromatography by NTA resins and then removing the affinity peptide(s) by selective cleavage.

According to an eighth embodiment of this invention there is provided vaccines containing a fusion protein and a physiologically compatible carrier material.

The affinity peptides of the fusion proteins in accordance with the invention are defined by the general formula



wherein R^1 represents hydrogen, an amino acid or a sequence of several amino acids, R^2 represents Q, Q-Ile-Glu-Gly-Arg- or Q-Asp-Asp-Asp-Lys- and Q is a peptide bond, an amino acid or a sequence of several, max. 30, amino acids.

Especially preferred affinity peptides of the fusion proteins in accordance with the invention are those with the peptide sequences of the formulae

Met-His-His,
Met-His-His-His,
Met-His-His-His-His,



Met-His-His-His-His-His,
Met-His-His-His-His-His-His,
Met-His-His-Ala-Gly-Ile-Glu-Gly-Arg
and Met-His-His-Ala-Gly-Asp-Asp-Asp-Asp-Lys.

5

The affinity peptides of the fusion proteins in accordance with the invention can be linked directly or indirectly to the biologically active polypeptide or protein. When a single affinity peptide is used, this can
10 be linked either to the amino terminal amino acid or to the carboxy terminal amino acid of the biologically active polypeptide or protein. When two affinity peptides are used, one of them is linked to the amino terminal amino acid of the biologically active polypeptide or protein and
15 the other is linked to the carboxy terminal amino acid of the biologically active polypeptide or protein.

In the case of indirect linking the affinity peptides contain a suitable selective cleavage site via which they
20 are linked to the desired biologically active polypeptide or protein. As suitable selective cleavage sites there preferably come into consideration the amino acid sequences $-(\text{Asp})_n\text{-Lys-}$, wherein n signifies 2, 3 or 4, or $-\text{Ile-Glu-Gly-Arg-}$ which can be specifically recognized
25 by the proteases enterokinase and coagulation factor X_a , respectively. Such affinity peptides can then be cleaved off enzymatically in a manner known per se.

In the case of direct linking the affinity peptides
30 remain linked with the desired biologically active polypeptide or protein, i.e. the affinity peptides can not be cleaved off chemically or enzymatically. This type of linking is advantageous when the activity of the desired polypeptide or protein is not influenced disadvantageously
35 by the presence of the affinity peptide. Such fusion proteins in accordance with the invention can be used for

a number of immunological procedures. They can be used, for example, as reagents for detecting infectious diseases. As they can be admixed with a physiologically compatible carrier material, they can also be used as vaccines in the prevention of diseases.

The term "biologically active polypeptide or protein" which is used in connection with the fusion proteins in accordance with the invention relates to those polypeptides or proteins which themselves are biologically active or to polypeptides or proteins which can be used for the preparation of biologically active polypeptides or proteins.

As biologically active polypeptides or proteins there come into consideration, for example, malaria surface antigens, especially the 5.1 surface antigen, the CS protein and the p190 protein of Plasmodium falciparum, lymphokines, interferons, insulin and insulin precursors, HIV-1 and HIV-2 envelope and structure proteins, growth hormones and growth hormone releasing factors. Especially preferred biologically active polypeptides or proteins of the fusion proteins in accordance with the invention are those having the amino acid sequence of human immune interferon and partial sequences of human immune interferon, especially those having the amino acid sequences of the formulae:

Gln-Asp-Pro-Tyr-Val-Lys-Glu-Ala-Glu-Asn-Leu-Lys-Lys-Tyr-Phe-Asn-Ala-Gly-His-Ser-Asp-Val-Ala-Asp-Asn-Gly-Thr-Leu-Phe-Leu-Gly-Ile-Leu-Lys-Asn-Trp-Lys-Glu-Glu-Ser-Asp-Arg-Lys-Ile-Met-Gln-Ser-Gln-Ile-Val-Ser-Phe-Tyr-Phe-Lys-Leu-Phe-Lys-Asn-Phe-Lys-Asp-Asp-Gln-Ser-Ile-Gln-Lys-Ser-Val-Glu-Thr-Ile-Lys-Glu-Asp-Met-Asn-Val-Lys-Phe-Phe-Asn-Ser-Asn-Lys-Lys-Lys-Arg-Asp-Asp-Phe-Glu-Lys-Leu-Thr-Asn-Tyr-Ser-Val-Thr-Asp-Leu-Asn-Val-Gln-Arg-Lys-Ala-Ile-His-Glu-Leu-Ile-Gln-Val-Met-Ala-Glu-Leu-

Ser-Pro-Ala-Ala-Lys-Thr-Gly-Lys-Arg-Lys-Arg-Ser-Gln-Met-Leu-Phe-Arg-Gly-Arg-Arg-Ala-Ser-Gln,

5 Gln-Asp-Pro-Tyr-Val-Lys-Glu-Ala-Glu-Asn-Leu-Lys-Lys-Tyr-Phe-Asn-Ala-Gly-His-Ser-Asp-Val-Ala-Asp-Asn-Gly-Thr-Leu-Phe-Leu-Gly-Ile-Leu-Lys-Asn-Trp-Lys-Glu-Glu-Ser-Asp-Arg-Lys-Ile-Met-Gln-Ser-Gln-Ile-Val-Ser-Phe-Tyr-Phe-Lys-Leu-Phe-Lys-Asn-Phe-10 Lys-Asp-Asp-Gln-Ser-Ile-Gln-Lys-Ser-Val-Glu-Thr-Ile-Lys-Glu-Asp-Met-Asn-Val-Lys-Phe-Phe-Asn-Ser-Asn-Lys-Lys-Lys-Arg-Asp-Asp-Phe-Glu-Lys-Leu-Thr-Asn-Tyr-Ser-Val-Thr-Asp-Leu-Asn-Val-Gln-Arg-Lys-Ala-Ile-His-Glu-Leu-Ile-Gln-Val-Met-Ala-Glu-Leu-Ser-Pro-Ala-Ala-Lys-Thr-Gly-Lys-Arg-Lys-Arg-Ser-Gln-Met-Leu and 15

Gln-Asn-Pro-Tyr-Val-Lys-Glu-Ala-Glu-Asn-Leu-Lys-Lys-Tyr-Phe-Asn-Ala-Gly-His-Ser-Asp-Val-Ala-Asp-Asn-Gly-Thr-Leu-Phe-Leu-Gly-Ile-Leu-Lys-Asn-Trp-Lys-Glu-Glu-Ser-Asp-Arg-Lys-Ile-Met-Gln-Ser-Gln-Ile-Val-Ser-Phe-Tyr-Phe-Lys-Leu-Phe-Lys-Asn-Phe-20 Lys-Asp-Asp-Gln-Ser-Ile-Gln-Lys-Ser-Val-Glu-Thr-Ile-Lys-Glu-Asp-Met-Asn-Val-Lys-Phe-Phe-Asn-Ser-Asn-Lys-Lys-Lys-Arg-Asp-Asp-Phe-Glu-Lys-Leu-Thr-Asn-Tyr-Ser-Val-Thr-Asp-Leu-Asn-Val-Gln-Arg-Lys-Ala-Ile-His-Glu-Leu-Ile-Gln-Val-Met-Ala-Glu-Leu-Ser-Pro-Ala-Ala-Lys-Thr-Gly-Lys-Arg-Lys-Arg-Ser-Gln-Met-Leu 25

as well as those having the amino acid sequence of mouse dihydrofolate reductase.

30 The preparation of the fusion proteins in accordance with the invention can be effected according to methods of recombinant DNA technology which are described in the literature. Preferably, a nucleotide sequence coding for the desired affinity peptide is firstly synthesized and this is then linked with a nucleotide sequence coding for 35 the desired biologically active polypeptide or protein.

The incorporation of the thus-obtained hybrid gene in expression vectors, for example the plasmids pDS8/RBSII, SphI; pDS5/RBSII,3A+5A; pDS78/RBSII or pDS56/RBSII and other commercial or generally accessible plasmids, is also effected in a manner known per se. In addition, reference can be made to the textbook of Maniatis et al. ("Molecular Cloning", Cold Spring Harbor Laboratory, 1982).

The methods for the expression of the fusion proteins in accordance with the invention are also known per se and are described in detail in the aforementioned textbook. They embrace the following procedures:

- (a) Transformation of a suitable host organism, advantageously *E. coli*, with an expression vector in which an aforementioned hybrid gene is operatively bonded to an expression control sequence;
- (b) cultivation of the thus-obtained host organism under suitable growth conditions; and
- (c) extraction and isolation of the desired fusion protein from the host organism.

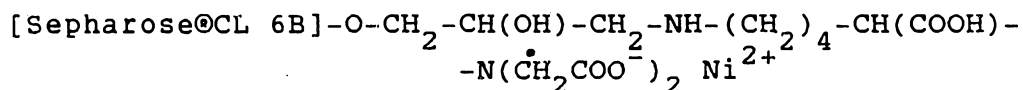
As host organisms there come into consideration gram-negative and gram-positive bacteria, for example *E. coli* and *B. subtilis* strains. *E. coli* strain M15 (see p. 17) is an especially preferred host organism of the present invention. Apart from the above-mentioned *E. coli* strain there can, however, also be used other generally accessible *E. coli* strains, for example *E. coli* 294 (ATCC No. 3144), *E. coli* RR1 (ATCC No. 31343) and *E. coli* W3110 (ATCC No. 27325).

Ideal metal chelate resins for the purification of the fusion proteins in accordance with the invention are nitrilotriacetic acid (NTA) resins of the general formula
Carrier matrix-spacer-NH-(CH₂)_x-CH(COOH)-
-N(CH₂COO⁻)₂ Ni²⁺, wherein x signifies 2, 3 or 4.

As the carrier matrix there come into consideration materials which are used in affinity and gel chromatography, for example cross-linked dextrans, agarose (especially in the form known under the trade mark Sepharose®) or polyacrylamides.

As the spacer there come into consideration the spacer groups which are already known from affinity chromatography, with the groups $-O-CH_2-CH(OH)-CH_2-$ and $-O-CO-$ being preferred.

An especially preferred NTA resin for the purification of the hybrid proteins in accordance with the invention is that of the formula



, the preparation of which can be effected as described in Example 10.

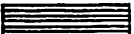
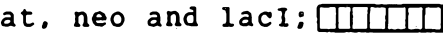
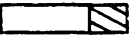
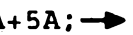
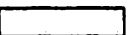
The aforementioned NTA resin can be used batch-wise or in continuously operating columns for the purification of the fusion proteins in accordance with the invention. Prior to the loading with the fusion protein in accordance with the invention, the NTA resin is conveniently equilibrated with an aqueous buffer which itself does not form chelates with nickel, preferably a Tris·HCl buffer, pH 7.5. The equilibration buffer (as well as the elution buffer) can contain a denaturing agent or a detergent, for example guanidine·HCl, urea or Triton. The addition of such a denaturing agent or detergent permits problem-free operations even with fusion proteins in accordance with the invention which are extremely difficultly soluble in aqueous solution. The elution of the fusion proteins in accordance with the invention can be carried out at a

constant pH value or with linear or discontinuously falling pH gradients. The optimal elution conditions depend on the amount and type of impurities which are present, the amount of material to be purified, the column dimensions etc. and are conveniently determined on a case by case basis.

The following Examples illustrate the preparation of fusion proteins in accordance with the invention, their purification by means of metal chelate chromatography as well as the preparation of biologically active polypeptides or proteins by enzymatic cleavage of the purified fusion proteins in accordance with the invention.

These Examples can be understood better when they are read in conjunction with the accompanying Figures. The following abbreviations and symbols appear in these Figures:

B, Bg, E, H, N, Na, Nd, P, S, Sa, Sc, X and Xb denote cleavage sites for the restriction enzymes BamHI, BglII, EcoRI, HindIII, NaeI, NarI, NdeI, PstI, SphI, SalI, ScaI, XhoI and XbaI, respectively.

 represents the promoters of the genes bla, lacI and neo;  represents the ribosomal binding sites of the genes bla, cat, neo and lacI;  represents the terminators t₀ and T₁;  represents the regulatable promoter/operator element P_{N25x/O} or N25OPSN25OP29;  represents the ribosomal binding sites RBSII, SphI and RBSII,3A+5A;  represents the coding regions under control of these ribosomal binding sites;  represents the regions which code for the affinity peptides in accordance with the invention as well as the selective cleavage sites;  represents the regions which code for 2,4 or 6 histidine residues; 

represents the region required for replication (repl.);  represents coding regions for dihydrofolate reductase (dhfr), chloramphenicol acetyltransferase, lac-repressor (lacI), β -lactamase (bla), neomycin phosphotransferase (neo) and the various derivatives of γ -interferon.

Figure 1

Schematic representation of the plasmid pDS8/RBS11, SphI.

Figure 2

Nucleotide sequence of the XhoI/XbaI fragment of the plasmid pDS8/RBS11, SphI. This fragment contains the regulatable promoter/operator element $P_{N25'X/O'}$, the ribosomal binding site RBS11, SphI, the dhfr gene, the terminator t_o , the cat gene as well as the terminator T1. The cleavage sites for restriction enzymes given in Figure 1 are overlined, while the region being under control of RBS11, SphI, which codes for a variant of dihydrofolate reductase, is underlined. In addition, the part of the plasmid pDS8/RBS11, SphI originating from plasmid pRB322 is shown schematically, with the given numbers referring to the sequence of pBR322 (J.G. Sutcliffe, Cold Spring Harbor Symp. Quant. Biol. 43, pp. 77-90 [1979]).

Figure 3

Schematic representation of the plasmid pDS5/RBS11, 3A+5a.

Figure 4

Nucleotide sequence of the plasmid pDS5/RBS11, 3A+5A. The cleavage sites for restriction enzymes given in Figure 3 are overlined, while the region being under control of RBS11, 3A+5A, which codes for a variant of chloramphenicol transferase, is underlined.

Figure 5

Schematic representation of the plasmid pDS78/RBSII.

Figure 6

5 Nucleotide sequence of the plasmid pDS78/RBSII. The cleavage sites for restriction enzymes given in Figure 5 are overlined, while the region being under control of RBSII, which codes for a variant of dihydrofolate reductase, is underlined.

Figure 7

Schematic representation of the plasmid pDS56/RBSII.

Figure 8

15 Nucleotide sequence of the plasmid pDS56/RBSII. The cleavage sites for restriction enzymes given in Figure 7 are overlined, while the region being under control of RBSII is underlined.

Figure 9

Schematic representation of the plasmid pDMI,1.

Figure 10

25 Nucleotide sequence of the plasmid pDMI,1. The cleavage sites for restriction enzymes given in Figure 9 are overlined, while the regions coding for neomycin phosphotransferase (neo) and lac-repressor (lacI) are underlined.

Figure 11

30 Nucleotide sequences of the oligonucleotides which are employed for the construction of the plasmids used in the Examples. In each case two such oligonucleotides were combined and referred to as the adaptor. The cleavage sites for the restriction enzymes NaeI, NarI and BglII are
35 overlined.

Figure 12

5 Schematic representation of the construction and isolation of fragment 1. This fragment was isolated from the plasmid pRC23/IFI-900 and contains the gene for the recombinant human interferon having Cys-Tyr-Cys as the N-terminal amino acids.

10 Figure 13

Schematic representation of the construction of the plasmid pGLS by incorporating fragment 1 in the plasmid pDS8/RBSII,SphI. In the schematic representation of pGLS, (Sc) denotes the position at which fragment 1 has been linked with the plasmid pDS8/RBSII,SphI via the cleavage site for the restriction enzyme ScaI.

15 Figure 14

Schematic representation of the construction and isolation of fragment 2. This fragment codes for a human interferon which is shortened at the C-terminus by 8 amino acids and which is referred to as IFN- γ (-8). In the given nucleotide sequences the corresponding termination codon is underlined.

20 Figure 15

Schematic representation of the construction of the plasmid pIFN- γ (-8) by incorporating fragment 2 in the plasmid pDS8/RBSII,SphI via the cleavage sites for the restriction enzymes EcoRI and HindIII.

25 Figure 16

30 Schematic representation of the construction and isolation of fragment 3. This fragment carries the regulatable promoter/operator element $P_{N25^x/O}$, the ribosomal binding site RBSII,SphI and the adator 3 which codes for the amino acid sequence Met-His-His-Ala-Gly-Ile-Glu-Gly-Arg-Leu-Gly-Ser.

Figure 17

Schematic representation of the construction of the plasmid pDS8/RBSII,SphI-His,His-Xa-BamHI by incorporating fragment 3 in the plasmid pDS8/RBSII,SphI via the cleavage sites for the restriction enzymes XhoI and BamHI.

Figure 18

Schematic representation of the construction and isolation of fragment 4 which was used in the construction of the plasmid pHis,His-Xa-IFN- γ .

Figure 19

Schematic representation of the isolation of fragment 5 which was used in the construction of the plasmids pHis,His-Xa-IFN- γ , pHis,His-Ek-IFN- γ (-8) and pHis,His-Xa-IFN- γ (-8)(Asn).

Figure 20

Schematic representation of the isolation of fragment 6 which was used in the construction of the plasmid pHis,His-Xa-IFN- γ .

Figure 21

Schematic representation of the construction of the plasmid pHis,His-Xa-IFN- γ by linking fragments 4, 5 and 6. The plasmid pHis,His-Xa-IFN- γ codes for a IFN- γ fusion protein having Met-His-His-Ala-Gly-Ile-Glu-Gly-Arg as an additional N-terminal amino acid sequence (His,His-Xa-IFN- γ).

Figure 22

Schematic representation of the construction and isolation of fragment 7 which was used in the construction of the plasmid pHis,His-Ek-IFN- γ (-8).

Figure 23

Schematic representation of the isolation of fragment 8 which was used in the construction of the plasmids pHis,His-Ek-IFN- γ (-8) and pHis,His-Xa-IFN- γ (-8)(Asn).

Figure 24

Schematic representation of the construction of the plasmid pHis-His-Ek-IFN- γ (-8) by linking fragments 5, 7 and 8. The plasmid pHis,His-Ek-IFN- γ (-8) codes for a IFN- γ fusion protein which is shortened by 8 amino acids and which has Met-His-His-Ala-Gly-Asp-Asp-Asp-Lys as an additional N-terminal amino acid sequence (His,His-Ek-IFN- γ (-8)).

Figure 25

Schematic representation of the construction and isolation of fragment 9 which was used in the construction of the plasmid pHis,His-Xa-IFN- γ (-8)(Asn).

Figure 26

Schematic representation of the construction of the plasmid pHis,His-Xa-IFN- γ (-8)(Asn) by linking fragments 5, 8 and 9. This plasmid codes for a IFN- γ fusion protein which is shortened at the C-terminus by 8 amino acids, which is lengthened at the N-terminus by the amino acid sequence Met-His-His-Ala-Gly-Ile-Glu-Gly-Arg and in which additionally at position 2 the amino acid Asp is replaced by the amino acid Asn (His,His-Xa-IFN- γ (-8)-(Asn)).

Figure 27

Schematic representation of the construction and isolation of fragment 10 which was used in the construction of the plasmid p6xHis-DHFR.

Figure 28

Schematic representation of the construction of the plasmid p6xHis-DHFR by linking fragment 10 with the XhoI/BamHI fragment of plasmid pDS78/RBSII containing the replication region. The plasmid p6xHis-DHFR codes for a DHFR fusion protein having 6 histidines at the N-terminus [(His)₆-mDHFR].

Figure 29

Schematic representation of the construction and isolation of fragment 11 which was used in the construction of the plasmid p4xHis-DHFR.

Figure 30

Schematic representation of the construction of the plasmid p4xHis-DHFR by linking fragment 11 with the XhoI/BamHI fragment of plasmid pDS78/RBSII containing the replication region. The plasmid p4xHis-DHFR codes for a DHFR fusion protein having 4 histidines at the N-terminus [(His)₄-mDHFR].

Figure 31

Schematic representation of the construction and isolation of fragment 12 which was used in the construction of the plasmid pRBSII-6xHis.

Figure 32

Schematic representation of the construction of the plasmid pRBSII-6xHis by linking fragment 12 with the XbaI/BamHI fragment of the plasmid pDS56/RBSII containing the replication region.

Figure 33

Schematic representation of the construction and isolation of fragment 13 which was used in the construction of the plasmid pRBSII-4xHis.

Figure 34

Schematic representation of the construction of the plasmid pRBSII-4xHis by linking fragment 13 with the XbaI/BamHI fragment of the plasmid pDS56/RBSII containing the replication region.

Figure 35

Schematic representation of the construction and isolation of fragment 14 which was used in the construction of the plasmid pRBSII-2xHis.

Figure 36

Schematic representation of the construction of the plasmid pRBSII-2xHis by linking fragment 14 with the XbaI/BamHI fragment of the plasmid pDS56/RBSII containing the replication region.

Figure 37

Schematic representation of the construction of the plasmid pDHFR-6xHis by linking the XbaI/BglII fragment of the plasmid pDS78/RBSII containing the replication region with the BglII/XbaI fragment of the plasmid pRBSII/6xHis containing the cat gene. The plasmid pDHFR-6xHis codes for a DHFR fusion protein having 6 histidines at the C-terminus [Met-mDHFR-(His)₆].

Figure 38

Schematic representation of the construction of the plasmid pDHFR-2xHis by linking the XbaI/BglII fragment of the plasmid pDS78/RBSII containing the replication region with the BglII/XbaI fragment of the plasmid pRBSII-2xHis containing the cat gene. The plasmid pDHFR-2xHis codes for a DHFR fusion protein having 2 histidines at the C-terminus [Met-mDHFR-(His)₂].

Figure 39

Schematic representation of the construction of the
5 plasmid p4xHis-DHFR-4xHis by linking the XbaI/BglII
fragment of the plasmid p4xHis-DHFR containing the
replication region with the BglII/XbaI fragment of the
plasmid pRBSII-4xHis containing the cat gene. The plasmid
10 p4xHis-DHFR-4xHis codes for a DHFR fusion protein having
in each case 4 histidines at the N- and at the C-terminus
[(His)₄-mDHFR-(His)₄].

Figure 40

15 Nucleotide sequence of the IFN-γ gene coded from
plasmid pGLS and the amino acid sequence derived therefrom.

Figure 41

20 Nucleotide sequence of the IFN-γ fusion gene coded
from plasmid pHis,His-Xa-IFN-γ and the amino acid
sequence derived therefrom.

Figure 42

25 Nucleotide sequence of the IFN-γ fusion gene coded
from plasmid pHis,His-Ek-IFN-γ(-8) and the amino acid
sequence derived therefrom.

Figure 43

30 Nucleotide sequence of the IFN-γ fusion gene coded
from plasmid pHis,His-Xa-IFN-γ(-8)(Asn) and the amino
acid sequence derived therefrom.

Example 1

35 Description of the plasmids which were used in the
construction of the plasmids pGLS, pHis,His-Xa-IFN-γ,
p-His,His-Ek-IFN-γ(-8), pHis,His-Xa-IFN-γ(-8)(Asn),
p6xHis-DHFR, p4xHis-DHFR-4xHis, pDHFR-2xHis and pDHFR-6xHis

A. Principles

The plasmids pDS8/RBSII,SphI (Figs. 1 and 2), pDS5/
RBSII,3A+5A (Figs. 3 and 4), pDS78/RBSII (Figs. 5 and 6)
5 and pDS56/RBSII (Figs. 7 and 8) were used for the con-
struction of the specified plasmids. E. coli cells trans-
formed with these plasmids have been deposited at the
Deutsche Sammlung von Mikroorganismen in Göttingen (from
November 21, 1987 in Braunschweig) on October 3, 1985
10 [E. coli M15 (pDS5/RBSII,3A+5A; pDMI,1), DSM No.: 3517],
on August 6, 1986 [E. coli M15 (pDS8/ RBSII,SphI; pDMI,1),
DSM No.: 3809], on September 3, 1987 [E. coli M15
(pDS78/RBSII; pDMI,1), DSM No.: 4232], and on December 23,
1987 [E. coli M15 (pDS56/RBSII; pDMI,1), DSM No.: 4330] in
15 accordance with the Budapest Treaty.

The above-mentioned vectors contain the regulatable
promoter/operator element $P_{N25^x/O}$ (Stüber et al., EMBO
J. 3, 3143-3148 [1984]) or N25OPSN25OP29 and the ribosomal
20 binding sites RBSII,SphI, RBSII,3A+5A or RBSII. These
ribosomal binding sites were derived from the ribosomal
binding site of the promoter P_{G25} of the E. coli phage
T5 (R. Gentz, Dissertation, Universität Heidelberg, BRD
[1984]) and were obtained via DNA synthesis. Having regard
25 to the high efficiency of these expression signals the
above-mentioned plasmids can be held stable in E. coli
cells only if the promoter/operator element is repressed
by the binding of a lac repressor to the operator. The lac
repressor is coded in the lacI gene. $P_{N25^x/O}$ and
30 N25OPSN25OP29 can be repressed efficiently only when a
sufficient number of repressor molecules is present in the
cells. Therefore, the lacI^q allele, which contains a
promoter mutant which leads to an increased expression of
the repressor gene, was used. This lacI^q allele is
35 present in the plasmid pDMI,1 (Figs. 9 and 10). This
plasmid carries, in addition to the lac-I gene, the

neo gene, which confers kanamycin resistance to the bacteria and which is used as the selection marker. pDMI,1 is compatible with the aforementioned plasmids. E. coli cells which are transformed with such expression vectors must contain pDMI,1 in order to guarantee that the expression vector is held stable in the cells. An induction of this system is achieved by adding IPTG to the medium at the desired cell density.

B. Plasmid pDS8/RBSII,SphI

The part of pDS8/RBSII,SphI (Figs. 1 and 2), which lies between the restriction cleavage sites for XbaI and XhoI and which contains the replication region as well as the gene for β -lactamase which confers ampicillin resistance to the cells, is derived from the plasmid pBR322 (Bolivar et al., Gene 2, 95-113 [1977]); Sutcliffe, Cold Spring Harbor Symp. Quant. Biol. 43, 77-90 [1979]). The remaining part of the plasmid carries the regulatable promoter/operator element $P_{N25^x/O}$ (Stüber et al., supra) followed by the ribosomal binding site RBSII, SphI, which is part of a EcoRI/BamHI fragment. There follow the dihydrofolate reductase (DHFR) gene of mouse cell line AT-3000 (Chang et al., Nature 275, 617-624 [1978]; Masters et al., Gene 21, 59-63 [1983]), the terminator t_o of E. coli phage lambda (Schwarz et al., Nature 272, 410-414 [1978]), the promoter-free gene of chloramphenicol acetyltransferase (Marcoli et al., FEBS Letters, 110, 11-14 [1980]) and the terminator T1 of the E. coli rrnB operon (Brosius et al., J. Mol. Biol., 148, 107-127 [1981]).

C. Plasmid pDS5/RBSII,3A+5A

The part of pDS5/RBSII,3A+5A (Figs. 3 and 4), which lies between the cleavage sites for the restriction

enzymes XbaI and XhoI and which contains the replication region as well as the gene for β -lactamase which confers ampicillin resistance to the cells, is derived originally from the plasmid pBR322 (Bolivar et al., supra; Sutcliffe, supra). However, the gene for β -lactamase is modified in that the cleavage sites for the restriction enzymes HincII and PstI are eliminated. These alterations in the DNA sequence do not, however, have any effect on the amino acid sequence of the β -lactamase. The remaining part of the plasmid carries the regulatable promoter/operator element $P_{N25}^{X/O}$ (Stüber et al., supra) followed by the ribosomal binding site RBSII,3A+5A, which is part of a EcoRI/BamHI fragment. There follow cleavage sites for the restriction enzymes SalI, PstI and HindIII, the promoter-free gene for chloramphenicol acetyltransferase (Marcoli et al., supra) and the terminator T1 of the E. coli rrnB operon (Brosius et al., supra).

D. Plasmid pDS78/RBSII

The part of pDS78/RBSII (Figs. 5 and 6), which lies between the restriction cleavage sites for XbaI and XhoI and which contains the replication region as well as the gene for β -lactamase which confers ampicillin resistance to the cells, is derived originally from the plasmid pBR322 (Bolivar et al., supra; Sutcliffe, supra). However, the gene for β -lactamase is modified in the manner described for the plasmid pDS5/RBSII,3A+5A. The remaining part of the plasmid carries the regulatable promoter/operator element N25OPSN25OP29 followed by the ribosomal binding site RBSII, which is part of a EcoRI/BamHI fragment. There follow the gene for dihydrofolate reductase of the mouse cell line AT-3000 (Chang et al., supra; Masters et al., supra), which has been altered by introducing a cleavage site for the restriction enzyme BglII immediately prior to the end of the structural gene,

the terminator t_o (Schwarz et al., supra), the promoter-free gene for chloramphenicol acetyltransferase (Marcoli et al., supra) and the terminator T1 (Brosius et al., supra).

5

E. Plasmid pDS56/RBSII

10

The plasmid pDS56/RBSII (Figs. 7 and 8) is very similar to the plasmid pDS5/RBSII,3A+5A. However, in contrast to this, the plasmid pDS56/RBSII contains as expression signals the regulatable promoter/operator element N25OPSN25OP29 and the ribosomal binding site RBSII. In addition, pDS56/RBSII contains the terminator t_o of the E. coli phage lambda (Schwarz et al., supra).

15

F. Plasmid pDMI,1

20

25

The plasmid pDMI,1 (Figs. 9 and 10) carries the gene for neomycin phosphotransferase from the transposon Tn5 (Beck et al., Gene 19, 327-336 [1982]), which confers kanamycin resistance to the E. coli cells, and the lacI gene (Farabough, Nature 274, 765-769 [1978]) with the promoter mutation I^q (Calos, Nature 274, 762-765 [1978]), which codes for the lac repressor. Moreover, the plasmid pDMI,1 contains a region of the plasmid pACYC184 (Chang and Cohen, J. Bacteriol. 134, 1141-1156 [1978]), which contains all information required for the replication and stable transmission to the daughter cells.

30

Example 2

Description of the DNA adaptors which were used in the construction of the various plasmids

35

A. Principles

In order to adapt the ribosomal binding site RBSII,

SphI on the gene for immune interferon (IFN- γ), to shorten this gene, to link IFN- γ and IFN- γ fragments such as e.g. IFN- γ (-8) with an affinity peptide and to express DHFR fusion proteins having at least two neighbouring histidine residues, oligonucleotides were synthesized chemically and, after their working-up, phosphorylated. The nucleotide sequences of the adaptors used are shown as double-stranded DNA sequences in Figure 11.

B. Synthesis and working-up of the oligonucleotides

The oligonucleotides were prepared simultaneously on a multisynthesis apparatus (described in European Patent Application No. 181, published on 21.05.85), with glass of defined pore size (CPG) being used as the carrier material (Kiefer et al., Immuno. Meth. 3, 69-83 [1985]; Sproat et al., Tetrahedr. Lett. 24 5771-5774 [1983]; Adams et al., J. Amer. Chem. Soc., 105, 661-663 [1985]. The lyophilized oligonucleotides were taken up in water and dissolved at 4°C and 1 hour. The DNA concentration was 100 nmoles/ml.

C. Phosphorylation of the oligonucleotides

In each case 150 pmol of the oligonucleotides were incubated in separate batches at 37°C for 20 minutes in 20 μ l of 50 mM Tris HCl, pH 8.5, and 10 mM MgCl₂ with 2 pmol of γ [³²P]-ATP (Amersham, Braunschweig; 5000 Ci/mmol) and 1 unit (U) of T₄ polynucleotide kinase (Gibco-BRL, Basle). Subsequently, 5 nmol of ATP were added and, after a further 20 minutes at 37°C, the reaction was completed by heating to 65°C. The thus-obtained phosphorylated oligonucleotides were used without any further working-up.

Example 3

Construction of the plasmid pGLS

5 A. Principles

For the construction of the plasmid pGLS, the IFN- γ gene was firstly bonded (Fig. 12) with the adaptor 1 (Fig. 11) and isolated. Subsequently the resulting fragment 1
10 was integrated in the plasmid pDS8/RBSII,SphI (Fig. 13).

B. Preparation of fragment 1

4 μ g of the plasmid pRC23/IFI-900 (Europ. Patent
15 Application publication No. 99084, published on 25.1.84)
having a DNA concentration of 400 μ g/ml were digested
with 10 units of the restriction endonuclease NdeI in core
buffer (50 mM Tris HCl, pH 8, 10mM MgCl₂, 50mM NaCl) for
1 hour at 37°C (volume 20 μ l). The sample was subse-
20 quently extracted once with phenol, the remainder of the
phenol was removed with ether and the DNA was finally
precipitated with 66% alcohol and 0.3M K acetate. The
sediment was dried for 2 minutes in a Speed-vac concen-
trator and dissolved in T4 ligase buffer (50mM Tris HCl,
25 pH 7.8, 10mM MgCl₂, 10mM DTT, 500 μ M ATP). 25 pmol of
the phosphorylated adaptor 1 (Fig. 11) were dissolved in
1x ligase buffer and added to this reaction mixture so
that a total volume of 25 μ l was achieved. The ligation
was carried out for 3 hours at 22°C, with 1 μ l of DNA
30 ligase (1 White unit, Boehringer Mannheim) being used. The
ligation was terminated by heating the sample to 65°C for
7 minutes. The DNA was precipitated with alcohol, dried as
described above and then dissolved in 50 μ l of NcoI
digestion buffer (50mM Tris HCl, pH 8, 10mM MgCl₂, 50mM
35 NaCl, 50mM KCl). 10 units of NcoI were added thereto and
the sample was incubated for 1 hour at 37°C. The enzyme

was subsequently inactivated by heating the sample to 65°C for 7 minutes. After a phenol extraction the DNA was precipitated as described above and the sediment was dried. The DNA was dissolved in Klenow buffer (50mM Tris HCl, pH 7.2, 10mM MgSO₄, 100 µM DTT), dATP, dGTP, dCTP and dTTP (final concentration in each case 100 µM) and 1 unit of Klenow enzyme (Boehringer, Mannheim) were added thereto and the sample was held for 1 hour at 22°C. The reaction was terminated by the addition of 2 µl of 0.25M EDTA, the sample was extracted with phenol, the DNA was precipitated with alcohol as described and dissolved in SphI digestion buffer (50mM Tris HCl, pH 7.5, 6mM MgCl₂, 50mM NaCl, 6mM 2-mercaptoethanol). After the addition of 10 units of SphI the sample was incubated for 1 hour at 37°C, the digestion was terminated by heating as described, a phenol extraction was carried out and the DNA was precipitated with alcohol and dried. The DNA sediment was dissolved in 10 µl of sample buffer and the DNA fragments were separated in a 6% polyacrylamide gel (elution buffer; 40mM Tris HCl, 20mM Na acetate, 2mM EDTA, pH 7.8). DNA of phage ΦX (Gibco-BRL, Basle) digested with HaeIII was used as the molecular weight standard. The DNA was stained with ethidium bromide (0.5 µg/ml), made visible with UV light (300 nm wavelength) and the IFN-γ coding band was cut out from the gel with a scalpel. The piece of gel was transferred into a pocket of size 4x4 mm in an agarose gel (0.7% agarose gel, running buffer: 90mM Tris borate, 90mM boric acid, 3mM EDTA, pH 8.3). The pocket was closed with 0.7% liquid agarose in 1xTBE in order to achieve a homogeneous electrical field. A portion of a NA45 membrane (Schleicher and Schüll, Dassel, BRD) was placed in front of the sample and the DNA was electrophoretized onto the membrane (5 minutes, 15 V/cm). After washing with distilled water the membrane was transferred into an Eppendorf test tube which contained 250 µl of 1.5M lithium acetate, 50mM Tris HCl,

pH 8, and 10mM EDTA. The DNA was now eluted for 20 minutes at 65°C. The membrane strip was removed from the test tube and the sample was extracted once with 200 µl of phenol (pH 8). The DNA was precipitated after the addition of 20 µl of 5M lithium acetate and 440 µl of iso-propanol and the sediment was washed with 80% ethanol and subsequently dried. Subsequently, the sediment was dissolved in 10 µl of TE buffer (10mM Tris HCl, pH 7.6, 1mM EDTA). The thus-obtained DNA fragment received the designation fragment 1 (Fig. 12).

C. Preparation of the plasmid pDS8/RBSII,SphI

2 pmol of the plasmid pDS8/RBSII,SphI were firstly cleaved with the restriction enzyme SphI. Thereafter, the resulting linear plasmid DNA was incubated with a limited amount of the restriction enzyme ScaI, whereby the DNA was cleaved only at approximately 50% of the ScaI cleavage sites present. The sample was now extracted with phenol, extracted with ether and the DNA was precipitated as described. The sediment was dried, dissolved in 20 µl of buffer (50mM Tris HCl, pH 8) and 1U of CIP (calf intestinal phosphatase, Boehringer Mannheim) was added. The sample was incubated for 1 hour at 37°C, the enzyme was removed by a phenol extraction and the DNA was precipitated. After the DNA had been dissolved, the ScaI/SphI fragment, which contained part of the cat gene, the terminator T1, the replication region, the bla gene, the promoter N25x/o and the ribosomal binding site RBSII,SphI, was isolated from a 1% agarose gel and transferred electrophoretically to a NA45 membrane as described above. The elution of the DNA, alcohol precipitation and dissolution in 10 µl of TE buffer was also carried out as described above. Approximately 1 pmol of the desired vector fragment was obtained.

D. Assembly of the plasmid pGLS

0.05 pmol of the vector fragment was ligated (s.o.) with 0.05 pmol of insert DNA (fragment 1) in ligase buffer. A control ligation without insert DNA was carried out in parallel thereto. E. coli M15 cells containing plasmid pDMI,1 were pre-prepared for the transformation according to the method of Morrison (Methods Enzymol. 68, 326-331 [1979]). After heating for 7 minutes at 65°C the ligation mixtures were added to 200 µl of these competent cells. The samples were maintained for 30 minutes in ice, then incubated for 2 minutes at 42°C and, after the addition of 0.5 ml of LB medium, incubated for 90 minutes at 37°C. The cells were then plated-out on LB agar plates which contained 100 µg/ml of ampicillin and 25 µg/ml of kanamycin and incubated overnight in an incubator at 37°C. The transformation of the control ligation gave no transformants. The ligation with fragment 1 on the other hand gave about 200 colonies. Individual colonies were picked with a sterile toothpick, transferred into a test tube which contained 10 ml of LB medium with 100 µg/ml of ampicillin and 25 µg/ml of kanamycin and maintained for 12 hours in a shaking incubator. Thereafter, the cells were sedimented and the plasmid DNA was isolated according to the method of Birnboim and Doly (Nucleic Acids Res. 7, 1515-1523 [1979]).

In each case 1 µg of the isolated plasmids was digested with SphI and XbaI in order to test whether a fragment which contains the IFN-γ gene and the terminator T1 is present in these plasmids. All analyzed DNA samples contained the mentioned DNA fragment of about 1 kb. These plasmids received the designation pGLS (Fig. 13).

E. Sequence analysis of the IFN- γ gene integrated in pGLS

In order to detect whether the correct IFN- γ sequence is present in the plasmid pGLS, the double-stranded circular plasmid DNA was sequenced, with a starter sequence (primer) labelled with [γ -³²P]-ATP being used. This starter sequence contains the nucleotides of position 199-218 of the plasmid pDS8/RBSII, SphI and ends 7 nucleotides before the ATG of the SphI cleavage site.

0.3 pmol of the isolated plasmid DNA was precipitated with alcohol, the sediment was washed once with 80% ethanol, dried and finally dissolved in 8 μ l of 1/4 TE buffer. After the addition of 2 pmol of the starter sequence the sample was incubated for 5 minutes at 42°C. The DNA was then sequenced according to the method of Sanger et al. (Proc. Natl. Acad. Sci. USA 74, 5463-6567 [1977]). As a radioactively labelled "primer" was used, all reactions were carried out with unlabelled deoxynucleotide triphosphates. The DNA sequence analysis indicated that the correct IFN- γ sequence had been integrated into the plasmid pGLS (amino acid sequence see Fig. 40).

Example 4

Construction of the plasmid pIFN- γ (-8)

A. Principles

For the construction of the plasmid pIFN- γ (-8), the adaptor 2 (Fig. 11) was firstly linked with the individual HinfI cleavage site in the IFN- γ gene. Because of a translation stop codon in this adaptor, the C-terminal

region of the IFN- γ protein is therefore shortened by 8 amino acids (Fig. 14). The resulting fragment 2 was subsequently integrated into the plasmid pDS8/RBSII,SphI (Fig. 15).

5

B. Preparation of fragment 2

3 pmol of the plasmid pGLS were digested with 15U of HinfI (50 μ l volume, 1 h, 37°C). The restriction enzyme
10 was subsequently inactivated (7 min. at 65°C), the sample was extracted with phenol, extracted with ether, precipitated with K acetate and alcohol and dried. The sediment was dissolved in 50 μ l of ligase buffer. 100 pmol of phosphorylated oligonucleotide (adaptor 2) were mixed with
15 10 μ l of HinfI-cleaved plasmid pGLS and, after the addition of 25 μ l of ligase buffer and 1U of T4-DNA ligase, incubated for 12 hours at 22°C. As described above, the reaction was terminated by heating the sample and the DNA was precipitated. A digestion with the restriction enzymes
20 EcoRI (15U) and HindIII (20U) for 20 hours at 37°C now followed. After heat inactivation, phenol extraction, extraction with ether and alcohol precipitation the sample was dissolved in 10 μ l of sample buffer and the DNA fragments were separated in a 6% polyacrylamide gel. After
25 staining with ethidium bromide the DNA bands were made visible under UV light (300 nm). The band which contained the IFN- γ gene was cut out from the gel with a sterile scalpel and electrophoretized as described above on to a NA45 membrane. The DNA was eluted as described and
30 received the designation fragment 2 (Fig. 14).

C. Preparation of the plasmid pDS8/RBSII,SphI

2 pmol of the plasmid pDS8/RBSII,SphI were digested
35 with 10U of EcoRI and 10U of HindIII for 1 hour at 37°C in a volume of 50 μ l. After heat inactivation of the enzyme

and alcohol precipitation the DNA sediment was dissolved in 10 μ l of sample buffer. After electrophoreses in a 1% agarose gel the EcoRI/HindIII fragment which contained the terminator t_o , the cat gene, the terminator T1, the replication region, the bla gene and the promoter $P_{N25}^{x/o}$ was cut out from the gel and eluted as described above.

D. Assembly of the plasmid pIFN- γ (-8)

10 10 μ l of the isolated EcoRI/Hind III vector fragment and half of the isolated fragment 2 were incubated (20°C, 3 hours) with 1U of T4 ligase. A control ligation without the addition of fragment 2 was carried out in parallel thereto. The ligations were terminated by heating the sample as described.

20 The transformations were carried out according to the method of Morrison (supra), with the E. coli strain M15, which contained the plasmid pDMI.1, being used. The cells were plated-out on LB agar plates which contained 100 μ g/ml of ampicillin and 25 μ g/ml of kanamycin. The plates were held for 15 hours at 37°C in an incubator.

25 No transformants were found on the control plates (= control ligation). The ligation in which the vector DNA and fragment 2 were used gave about 500 colonies. Individual colonies were picked with a sterile toothpick, transferred into 100 ml of LB medium and left to grow as described. The plasmid DNA was isolated according to the method of Birnboim and Doly (supra). In each case 4 μ l of the plasmid DNA, isolated and dissolved in TE buffer, were cleaved as described with in each case 2U of EcoRI and HindIII. A fragment with the desired length of about 35 450 bp could be cut out from all tested plasmids. These plasmids received the designation pIFN- γ (-8) (Fig. 15).

E. Sequence analysis of the plasmid pIFN γ (-8)

The sequence analysis was carried out as described in Example 3. However, as the starter sequence there was used an oligonucleotide which contained the nucleotides of position 928-896 of the plasmid pDS8/RBSII,SphI and which therefore permitted the sequencing of DNA fragments which were integrated in front of the terminator t_0 . The sequence analysis confirmed the desired sequence of the IFN- γ gene which codes for a IFN- γ protein shortened by 8 amino acids (at the carboxyl end).

Example 5.

Construction of the plasmid pDS8/RBSII,SphI-His,His-Xa-BamHI

A. Principles

For the construction of the plasmid pDS8/RBSII,SphI-His,His-Xa-BamHI, the adaptor 3 (Fig. 11), which codes for an affinity peptide which contains two neighbouring histidines and a cleavage site of factor Xa was linked with the ribosomal binding site RBSII,SphI. Subsequently, fragment 3 (Fig. 16) containing the promoter $P_{N25^X/O}$ as well as the said affinity peptide was isolated and integrated into the plasmid pDS8/RBSII,SphI (Fig. 17).

B. Preparation of fragment 3

2 pmol of the plasmid pDS8/RBSII,SphI were cleaved with the restriction enzyme SphI. The reaction was terminated by incubation at 65°C for 7 minutes, the sample was extracted with phenol, extracted with ether and the DNA was precipitated with alcohol and K acetate. The sediment was taken up in 10 μ l of ligase buffer, 25 pmol of the

phosphorylated adaptor 3 (Fig. 11) dissolved in ligase buffer were added and, after the addition of 1U of T4 DNA ligase, incubated for 3 hours at 22°C. The reaction was terminated by heating (7 minutes, 65°C) and, after a phenol extraction and subsequent treatment with ether, the DNA was precipitated with alcohol and K acetate. The sediment was dissolved in 30 µl of buffer, in each case 10U of the restriction enzymes BamHI and XhoI were added and the mixture was incubated at 37°C for 2 hours. Subsequently, 3.5 µl of 10-fold concentrated sample buffer for polyacrylamide gels were added to the sample and the mixture was incubated for 7 minutes at 65°C. The DNA was separated in a 6% polyacrylamide gel and the fragment liberated by XhoI and BamHI was cut out with a scalpel. The DNA was eluted as described and received the designation fragment 3 (Fig. 16).

C. Preparation of the plasmid pDS8/RBSII,SphI

2 pmol of the plasmid pDS8/RBSII,SphI were cleaved with in each case 10U of the restriction enzymes BamHI and XhoI. After heat inactivation of the enzyme the sample was extracted with phenol, extracted with ether and the DNA was precipitated with alcohol and K acetate. The sediment was re-suspended in 50 µl of 50 mM Tris HCl, pH 8. 1U of CIP (see above) was added and the sample was incubated at 37°C for 30 minutes. After heat inactivation of the enzyme the DNA was separated in a 6% polyacrylamide gel after the addition of sample buffer and the plasmid body was eluted from the gel as described.

D. Assembly of the plasmid pDS8/RBSII,SphI-His,His-Xa-BamHI

The above-described fragment 3 was ligated with the vector body (22°C, 2U of T4 DNA ligase, 25 µl of ligase

buffer). A control ligation without the addition of fragment 3 was carried out in parallel thereto. The ligation batches were transformed as described above into the E. coli strain M15 which contained the plasmid pDM1,1 and plated out on LB plates with 100 µg/ml of ampicillin and 25 µg/ml of kanamycin. The transformation of the control ligation gave no transformants. The transformation of the ligation batch with fragment 3 gave about 100 colonies. Individual colonies were grown up in 100 ml of LB medium as described above and the plasmid DNA was isolated according to the method of Birnboim and Doly (supra). All plasmids contained the cleavage sites for NaeI and NarI (see Fig. 17) newly introduced by the adaptor. The sequence analysis of the plasmid DNA was carried out as described above (Example 3,F) and confirmed that the adaptor 3 had been integrated correctly into the vector. These plasmids received the designation pDS8/RBSII,SphI-His,His-Xa-BamHI (Fig. 17).

Example 6

Construction of the plasmid pHis,His-Xa-IFN-γ

A. Principles

For the construction of the plasmid pHis,His-Xa-IFN-γ, the following DNA fragments were isolated and linked with one another (Fig. 21): 1) the IFN-γ gene of the plasmid pGLS (fragment 4, Fig. 18) linked with the adaptor 4 (Fig. 11); 2) the signal unit of the plasmid pDS8/RBSII,SphI-His,His-Xa-BamHI, which contains the promoter P_{N25^x/O}, the ribosomal binding site RBSII, SphI and the region coding for the neighbouring histidines and for the recognition site of factor Xa (fragment 6, Fig. 20); and 3) the replication region with the β-lactamase gene from the plasmid pDS5/RBSII,3A+5A (fragment 5, Fig. 19).

B. Preparation of fragment 4

5 2 pmol of the plasmid pGLS were cleaved with the
restriction enzyme NdeI. After heat inactivation of the
enzyme the sample was extracted with phenol, extracted
with ether and the DNA was precipitated as described. The
sediment was dissolved in 10 µl of ligase buffer.
10 50 pmol of the phosphorylated adaptor 4 (Fig. 11) dis-
solved in ligase buffer were added to the NdeI-cleaved
plasmid pGLS and the sample was incubated (22°C, 3 hours)
with 2U of ligase. After heat inactivation of the ligase
the sample was extracted with phenol, extracted with ether
and the DNA was precipitated as described. The sediment
15 was dissolved and the DNA was cleaved with the restriction
enzymes NarI and XbaI. After the addition of sample
buffer, heating the mixture at 65°C for 7 minutes and
separation of the DNA in a 6% polyacrylamide gel the NarI/
20 XbaI fragment, which contained the IFN-γ gene, was iso-
lated as described. This fragment received the designation
fragment 4 (Fig. 18).

C. Preparation of fragment 5

25 2 pmol of the plasmid pDS5/RBSII,3A+5A were cleaved
with the restriction enzymes XhoI and XbaI. The mixture
was worked-up as described above and the DNA was separated
in a 6% polyacrylamide gel. The fragment which contained
the bla gene and the replication region was isolated from
30 the gel as described. This fragment received the
designation fragment 5 (Fig. 19).

D. Preparation of fragment 6

35 2 pmol of the plasmid pDS8/RBSII,SphI-His,His-Xa-
-BamHI were cleaved with the restriction enzymes XhoI and
NarI. After working-up the sample and gel electrophoreses

there was isolated fragment 6 which contains the promoter $P_{N25^{X/O}}$, the ribosomal binding site RBSII, SphI as well as the region which codes for the neighbouring histidines and the recognition site for factor Xa (Fig. 20).

5

E. Assembly of the plasmid pHis,His-Xa-IFN- γ

10 In each case 0.5 pmol of fragments 4 (Fig. 18), 5 (Fig. 19) and 6 (Fig. 20) were incubated (22°C, 5 hours) in ligase buffer with 2U of T4 DNA ligase. After heat inactivation of the enzyme the batch was transformed as described above into the E. coli strain M15 which contained the plasmid pDMI.1 and the transformation mixture was plated-out on LB agar plates which contained
15 100 μ g/ml of ampicillin and 25 μ g/ml of kanamycin. After incubation at 37°C overnight about 100 transformants were obtained. Individual colonies were grown up in 100 ml of LB medium as described and the plasmids were isolated according to the method of Birnboim and Doly (supra). All
20 plasmids were cleaved with the restriction enzymes XhoI, BamHI and XbaI and the fragments were analyzed in 6% polyacrylamide gels. The restriction enzyme analysis indicated that the plasmids contained the 3 desired fragments. The sequence analyses were carried out as described (Example
25 3F) and showed that the adaptor 4 has been fused correctly with IFN- γ gene. These plasmids received the name pHis,His-Xa-IFN- γ (Fig. 21). The IFN- γ fusion protein (amino acid sequence see Fig. 41) coded from these plasmids received the designation His,His-Xa-IFN- γ .

30

Example 7

Construction of the plasmid pHis,His-Ek-IFN- γ (-8)

35 A. Principles

For the construction of the plasmid pHis,His-Ek-IFN- γ (-8), the following 3 DNA fragments were linked with one another (Fig. 24): 1) a fragment from the plasmid pDS8/RBSII,SphI-His,His-Xa-BamHI containing the promoter $P_{N25^x/O}$, the ribosomal binding site RBSII,SphI and the region coding for the adjacent histidines, which has been lengthened with the aid of adaptor 5 (Fig. 11) by a region coding for the recognition site of enterokinase (EK) (fragment 7, Fig. 22); 2) a fragment from the plasmid pIFN- γ (-8) which contains the gene for IFN- γ (-8) (fragment 8, Fig. 23); and 3) a fragment from the plasmid pDS5/RBSII,3A+5A having the replication region and the β -lactamase gene (fragment 5, Fig. 19). The preparation of the last-named fragment has been described in Example 6.

B. Preparation of fragment 7

4 pmol of the plasmid pDS8/RBSII,SphI-His,His-Xa-BamHI were cleaved with the restriction enzyme NaeI. Subsequently, the enzyme was heat-inactivated, the sample was extracted with phenol, extracted with ether and the DNA was precipitated as described. The sediment was taken up in 50 μ l of TE buffer. 1.5 pmol of the cleaved DNA were incubated in a volume of 200 μ l with 30 pmol of the phosphorylated adaptor 5 (Fig. 11) and 7U of T4 DNA ligase for 14 hours in ligase buffer. After heat inactivation of the enzyme the DNA was cleaved with the restriction enzymes NdeI and XhoI. Subsequently, the enzymes were heat-inactivated, the sample was extracted with phenol, extracted with ether and the DNA was precipitated as described. The sediment was taken up in sample buffer and the mixture was incubated for 7 minutes at 65°C. Thereafter, the DNA was separated in a 6% polyacrylamide gel. 0.2 pmol of the XhoI/NdeI fragment containing the promoter $P_{N25^x/O}$, the ribosomal binding site RBSII,SphI as well as the region coding for the neighbouring histidines and

for the recognition site of enterokinase, was isolated from the gel as described. This DNA fragment received the designation fragment 7 (Fig. 22).

5 C. Preparation of fragment 8

10 0.5 pmol of the plasmid pIFN- γ (-8) was cleaved with the restriction enzymes NdeI and XbaI. Subsequently, the enzyme was heat-inactivated, the sample was extracted with phenol, extracted with ether and the DNA was precipitated as described. The sediment was taken up in sample buffer and the mixture was incubated for 7 minutes at 65°C. The DNA was then separated in a 6% polyacrylamide gel. 0.05 pmol of the NdeI/XbaI fragment containing the IFN- γ (-8) gene, the terminator t_o , the cat gene as well as the terminator T1 was isolated from the gel as described. This DNA fragment received the designation fragment 8 (Fig. 23).

20 D. Assembly of the plasmid pHis₈-His-Ek-IFN- γ (-8)

0.006 pmol of fragment 5 (Fig. 19), 0.02 pmol of fragment 7 (Fig. 22) and 0.005 pmol of fragment 8 (Fig. 23) were incubated in ligase buffer in a volume of 30 μ l with 0.5U of T4 DNA ligase for 3 hours at 15°C. After heat inactivation of the enzyme the batch was transformed as described above into the E. coli strain M15 which contained the plasmid pDMI,1 and the transformation mixture was plated-out on LB plates which contained 100 μ g/ml of ampicillin, 25 μ g/ml of kanamycin and 5 μ g/ml of chloramphenicol. After incubation of the plates for 24 hours at 37°C 2 of the transformants obtained were grown up in 10 ml of LB medium containing 100 μ g/ml of ampicillin, 25 μ g/ml of kanamycin and 5 μ g/ml of chloramphenicol and the plasmids were isolated according to the method of Birnboim and Doly (supra). The plasmids

were analyzed with respect to their size in 0.7% agarose gels and with respect to their composition with the aid of the restriction enzymes HindIII, HinfI, NdeI, SphI, XbaI and XhoI in 6% polyacrylamide gels. Both plasmids contained the desired 3 DNA fragments in the correct orientation to one another. The sequence analyses, which were carried out as described (Example 3,F), indicated that the ribosomal binding site RBSII,SphI with the subsequent elements had been linked correctly with the IFN- γ (-8) gene. These plasmids received the designation pHis,His-Ek-IFN- γ (-8) (Fig. 24). The IFN- γ fusion protein (amino acid sequence, Fig. 42) coded from these plasmids received the designation His,His-Ek-IFN- γ (-8).

Example 8

Construction of the plasmid pHis,His-Xa-IFN- γ (-8)(Asn)

A. Principles

For the construction of the plasmid pHis,His-Xa-IFN- γ (-8)(Asn), the following 3 DNA fragments were linked with one another (Fig. 26): 1) a fragment from the plasmid pIFN- γ (-8) which contained the IFN- γ (-8) gene (fragment 8, Fig. 23, the preparation of which has been described in Example 7); 2) a fragment from the plasmid pDS8/RBSII,SphI-His,His-Xa-BamHI containing the promoter $P_{N25^{x/O}}$, the ribosomal binding site RBSII,SphI as well as the region coding for the neighbouring histidines and for the recognition site of factor Xa, which has been lengthened with the aid of adaptor 6 (Fig. 11) so that, by linkage with the fragment described under 1), the IFN- γ (-8) derivative IFN- γ (-8)(Asn) is coded (fragment 9, Fig. 25); and 3) a fragment from the plasmid pDS5/RBSII, 3A+5A with the replication region and the gene for β -lactamase (fragment 5, Fig. 19, the preparation of which has been described in Example 6).

B. Preparation of fragment 9

2 pmol of the plasmid pDS8/RBSII,SphI-His,His-Xa-BamHI were cleaved with the restriction enzyme NarI. Subsequently, the enzyme was heat-inactivated, the sample was extracted with phenol, extracted with ether and the DNA was precipitated as described. The sediment was taken up in 50 μ l of TE buffer. 1 pmol of the cleaved DNA was incubated in a volume of 150 μ l with 30 pmol of the phosphorylated adaptor 6 (Fig. 11) and 5U of T4 DNA ligase for 14 hours at 20°C in ligase buffer. After heat inactivation of the enzyme the DNA was cleaved with the restriction enzymes NdeI and XhoI. Subsequently, the enzymes were heat-inactivated, the sample was extracted with phenol, extracted with ether and the DNA was precipitated as described. The sediment was taken up in sample buffer and the batch was incubated for 7 minutes at 65°C. The resulting DNA mixture was then separated in a 6% polyacrylamide gel. 0.25 pmol of the XhoI/NdeI fragment with the promoter P_{N25^x/O}, the ribosomal binding site RBSII,SphI and the region which codes for the adjacent histidines, for the recognition site of factor Xa as well as for the amino acid Asn was isolated from the gel as described. This DNA fragment received the designation fragment 9 (Fig. 25).

C. Assembly of the plasmid pHis,His-Xa-IFN- γ (-8)(Asn)

0.006 pmol of fragment 5 (Fig. 19), 0.005 pmol of fragment 8 (Fig. 23) and 0.02 pmol of fragment 9 (Fig. 25) were incubated in 30 μ l of ligase buffer and 0.5 U of T4 DNA ligase for 3 hours at 15°C. After heat inactivation of the enzyme the batch was transformed as described above into the E. coli strain M15 which contained the plasmid pDMI,1 and the transformation mixture was plated-out on LB

plates containing 100 µg/ml of ampicillin and 25 µg/ml of kanamycin. After incubation of the plates at 37°C overnight 2 colonies were grown up as described in 10 ml of LB medium containing 100 µg/ml of ampicillin and 25 µg/ml of kanamycin and the plasmids were isolated according to the method of Birnboim and Doly (supra). The plasmids were analyzed with respect to their size in 0.7% agarose gels and with respect to their composition with the aid of the restriction enzymes HindIII, HinfI, NdeI, SphI, XbaI and XhoI in 6% polyacrylamide gels. One of the two plasmids contained the desired 3 DNA fragments in the correct orientation to one another. The sequence analysis of this plasmid was carried out as described (Example 3,F) and indicated that the ribosomal binding site RBSII, SphI with the following elements had been correctly linked with the region which codes for IFN-γ(-8)(Asn). This plasmid received the designation pHis,His-Xa-IFN-γ(-8)(Asn) (Fig. 26). The IFN-γ fusion protein (amino acid sequence see Fig. 43) coded from this plasmid received the designation His,His-Xa-IFN-γ(-8)(Asn).

Example 9

Construction of the plasmid p6xHis-DHFR

A. Principles

For the construction of the plasmid p6xHis-DHFR, the following DNA fragments were isolated and linked with each other (Fig. 28): 1) the signal unit of the plasmid pDS78/RBSII having the promoter N25OPSN25OP29 as well as the ribosomal binding site RBSII, which has been linked with the adaptor 7 (Fig. 11), (fragment 10, Fig. 27) and 2) the larger of the two XhoI/BamHI fragments of the plasmid pDS78/RBSII (Fig. 28).

B. Preparation of fragment 10

2 pmol of the plasmid pDS78/RBSII were cleaved with the restriction enzyme BamHI. After heat inactivation of the enzyme the sample was extracted with phenol, extracted with ether and the DNA was precipitated as described. The sediment was dissolved in 10 μ l of ligase buffer.

50 pmol of the phosphorylated adaptor 7 (Fig. 27) dissolved in ligase buffer were added to the cleaved plasmid and the sample was incubated with 2U of ligase (22°C, 3 hours). After heat inactivation of the ligase the sample was extracted with phenol, extracted with ether and the DNA was precipitated as described. The sediment was dissolved and the DNA was cleaved with the restriction enzymes XhoI and BamHI. After adding sample buffer, heating the mixture at 65°C for 7 minutes and separating the DNA in a 6% polyacrylamide gel the XhoI/BamHI fragment having the promoter N25OPSN25OP29, the ribosomal binding site RBSII and the region coding for 6 histidines was isolated as described. This fragment received the designation fragment 10 (Fig. 27).

C. Preparation of the BamHI/XhoI fragment of the plasmid pD78/RBSII

2 pmol of the plasmid pDS78/RBSII were cleaved with the restriction enzymes XhoI and BamHI. After working-up the sample and gel electrophoreses there was isolated the fragment which contains the replication region (Fig. 28).

D. Assembly of the plasmid p6xHis-DHFR

In each case 0.1 pmol of the specified fragments were incubated in ligase buffer with 2U of T4 DNA ligase (22°C, 3 hours). After heat inactivation of the enzyme the mixture was transformed as described above into the E. coli

strain M15 which contained the plasmid pDMI,1, the transformation mixture was plated-out on LB agar plates which contained 100 µg/ml of ampicillin and 25 µg/ml of kanamycin and the plates were incubated at 37°C overnight. Individual colonies were grown up in 10 ml of LB medium as described and the plasmids were isolated according to the method of Birnboim and Doly (supra). A restriction analysis with the enzymes XhoI and BamHI indicated that the plasmids contained the 2 desired fragments. The sequence analysis was carried out as described (Example 3,E) and confirmed that the adaptor 7 had been linked correctly with the ribosomal binding site. These plasmids, which code for the DHFR fusion protein (His)₆-mDHFR, received the designation p6xHis-DHFR (Fig. 28).

Example 10

Construction of the plasmid p4xHis-DHFR

The construction of the plasmid p4xHis-DHFR was carried out analogously to the construction of the plasmid p6xHis-DHFR (Example 9), with the following DNA fragments being isolated and linked with each other (Fig. 30): 1) the signal unit of the plasmid pDS78/RBSII having the promoter N25OPSN25OP29 as well as the ribosomal binding site RBSII, which has been linked with the adaptor 8 (Fig. 11), (fragment 11, Fig. 29) and 2) the larger of the two XhoI/BamHI fragments of the plasmid pDS78/RBSII (Fig. 30). The resulting plasmid p4xHis-DHFR codes for the DHFR fusion protein (His)₄-mDHFR.

Example 11

Construction of the plasmid pRBSII-6xHis

A. Principles

5 For the construction of the plasmid pRBSII-6xHis, the following DNA fragments were isolated and linked with each other (Fig. 32): 1) the region from the plasmid pDS56/RBSII having the terminator t_o , the cat gene and the terminator T1, which has been lengthened by the adaptor 9
10 (which codes for 6 histidines), (fragment 12, Fig. 31) and 2) the XbaI/BamHI fragment from the plasmid pDS56/RBSII having the replication region, the bla gene, the promoter N25OPSN25OP29 and the ribosomal binding site RBSII (Fig. 32).

15 B. Preparation of fragment 12

2 pmol of the plasmid pDS56/RBSII were cleaved with the restriction enzyme HindIII. After working-up the sample 50 pmol of the phosphorylated adaptor 9 were added
20 to the cleaved plasmid and the sample was incubated with T4 DNA ligase as described. After working-up the ligation batch the DNA was cleaved with the restriction enzymes BamHI and XbaI and the BamHI/XbaI fragment having the region coding for 6 histidines, the terminator t_o , the cat gene and the terminator T1 was isolated as described.
25 This fragment received the designation fragment 12 (Fig. 31).

30 C. Preparation of the XbaI/BamHI fragment of the plasmid pDS56/RBSII

2 pmol of the plasmid pDS56/RBSII were cleaved with the restriction enzymes XbaI and BamHI and the fragment
35 having the replication region, the bla gene, the promoter N25OPSN25OP29 and the ribosomal binding site RBSII was isolated as described (Fig. 32).

D. Assembly of the plasmid pRBSII-6xHis

In each case 0.1 pmol of the isolated fragments were, as described (Example 9,D), ligated and subsequently transformed into the E. coli strain M15 (pDMI,1). After plating and incubation (Example 9,D) individual colonies were grown up in 10 ml of medium as described and the plasmids were isolated according to the method of Birnboim and Doly (supra). A restriction analysis with the enzymes BamHI and XbaI indicated that the plasmids contained the 2 desired fragments. The sequence analysis (Example 3,E) confirmed that the adaptor 9 had been introduced correctly into the plasmid DNA. These plasmids received the designation pRBSII-6xHis.

Example 12

Construction of the plasmid pRBSII-4xHis

The construction of the plasmid pRBSII-4xHis was carried out analogously to the construction of the plasmid pRBSII-6xHis (Example 11), with the following DNA fragments being isolated and linked with each other (Fig. 34): 1) the region from the plasmid pDS56/RBSII having the terminator t_0 , the cat gene and the terminator T1, which has been lengthened by the adaptor 10 (which codes for 4 histidines), (fragment 13, Fig. 33) and 2) the XbaI/BamI fragment from the plasmid pDS56/RBSII having the replication region, the bla gene, the promoter N25OPSN25OP29 and the ribosomal binding site RBSII (Fig. 34).

Example 13

Construction of the plasmid pRBSII-2xHis

The construction of the plasmid pRBSII-2xHis was carried out analogously to the construction of the plasmid pRBSII-6xHis (Example 11), with the following DNA fragments being isolated and linked with each other (Fig. 36): 1) the region from the plasmid pDS56/RBSII having the terminator t_0 , the cat gene and the terminator T1, which has been lengthened by the adaptor 11 (which codes for 2 histidines), (fragment 14, Fig. 35) and 2) the XbaI/BamHI fragment from the plasmid pDS56/RBSII having the replication region, the bla gene, the promoter N25OPSN25OP29 and the ribosomal binding site RBSII (Fig. 36).

Example 14

Construction of the plasmid pDHFR-6xHis

A. Principles

For the construction of the plasmids pDHFR-6xHis the following DNA fragments were isolated and linked with each other (Fig. 37): 1) the XbaI/BglII fragment from the plasmid pDS78/RBSII having the replication region, the bla gene, the promoter N25OPSN25OP29, the ribosomal binding site RBSII and the dhfr gene and 2) the BglII/XbaI fragment from the plasmid pRBSII-6xHis having the region coding for 6 histidines, the terminator t_0 , the cat gene and the terminator T1. The resulting plasmid pDHFR-6xHis codes for the DHFR fusion protein Met-mDHFR-(His)₆.

B. Preparation of the XbaI/BglII fragment of the plasmid pDS78/RBSII

2 pmol of the plasmid pDS78/RBSII were cleaved with the restriction enzymes XbaI and BglII. After working-up the sample the XbaI/BglII fragment having the replication

region, the bla gene, the promoter N25OPSN25OP29, the ribosomal binding site RBSII and the dhfr gene was isolated as described.

5 C. Preparation of the BglII/XbaI fragment of the plasmid pRBSII-6xHis

2 pmol of the plasmid pRBSII-6xHis were cleaved with the restriction enzymes BglII and XbaI. After working-up
10 the sample the BglII/XbaI fragment having the region coding for 6 histidines, the terminator t_0 , the cat gene and the terminator T1 was isolated as described.

D. Assembly of the plasmid pDHFR-6xHis

15 In each case 0.1 pmol of the isolated fragments were, as described (Example 9,D), ligated and transformed into the E. coli strain M15 (pDMI,1). After plating and incubation (Example 9,D) individual colonies were grown up in
20 10 ml of medium as described and the plasmids were isolated according to the method of Birnboim and Doly (supra). A restriction analysis with the enzymes XbaI and BglII indicated that the two fragments had been linked
25 with one another in the desired manner. These plasmids received the designation pDHFR-6xHis.

Example 15

Construction of the plasmid pDHFR-2xHis

30

The construction of the plasmid pDHFR-2xHis, which codes for the DHFR fusion protein Met-mDHFR-(His)₂, was carried out analogously to the construction of the plasmid pDHFR-6xHis (Example 14), with the following DNA fragments
35 being isolated and linked with each other (Fig. 38): 1) the XbaI/BglII fragment from the plasmid pDS78/RBSII

having the replication region, the bla gene, the promoter N25OPSN25OP29, the ribosomal binding site RBSII and the dhfr gene and 2) the BglII/XbaI fragment from the plasmid pRBSII-2xHis having the region coding for 2 histidines, the terminator t_0 , the cat gene and the terminator T1.

Example 16

Construction of the plasmid p4xHis-DHFR-4xHis

The construction of the plasmid p4xHis-DHFR-4xHis, which codes for the DHFR fusion protein (His)₄-mDHFR-(His)₄, was carried out analogously to the construction of the plasmid pDHFR-6xHis (Example 14), with the following DNA fragments being isolated and linked with each other (Fig. 39): 1) The XbaI/BglII fragment from the plasmid p4xHis-DHFR having the replication region, the bla gene, the promoter N25OPSN25OP29, the ribosomal binding site RBSII, 4xHis and the dhfr gene and 2) the BglII/XbaI fragment from the plasmid pRBSII-4xHis having the region coding for 4 histidines, the terminator t_0 , the cat gene and the terminator T1.

Example 17

Preparation of the NTA resin

41.7 g of bromoacetic acid were dissolved in 150 ml of 2N sodium hydroxide solution and cooled to 0°C. Thereto there was slowly added dropwise at 0°C while stirring a solution of 42 g of N^E-Z-L-lysine in 225 ml of 2N sodium hydroxide solution. After 2 hours the cooling was removed and the mixture was stirred overnight. The reaction mixture was then held at 50°C for 2 hours and 450 ml of 1N hydrochloric acid were subsequently added. After the mixture had been cooled the separated crystals

were filtered off. The product was dissolved in 1N sodium hydroxide solution and again precipitated with the same amount of 1N hydrochloric acid and filtered off. There were obtained 40 g of N-[5-benzyloxycarbonylamino-1-carboxypentyl]-iminodiacetic acid in the form of white crystals, m.p. 172-174°C (dec.), $[\alpha]_D = +9.9^\circ$ (c = 1; 0.1N NaOH).

7.9 g of the lysine derivative obtained were dissolved in 49 ml of 1N sodium hydroxide solution and, after the addition of a spatula tip of 5% Pd/C, hydrogenated at room temperature and normal pressure. The catalyst was filtered off and the filtrate was evaporated. There resulted 6.2 g of N-[5-amino-1-carboxypentyl]-iminodiacetic acid whose structure, $\text{NH}_2(\text{CH}_2)_4\text{-CH}(\text{COOH})\text{-N}(\text{CH}_2\text{COOH})_2$, was confirmed by the NMR spectrum.

100 ml of Sepharose[®] CL-6B (Pharmacia) were washed twice on a glass suction filter with about 600 ml of water and then reacted at 30°C for 4 hours in a 500 ml round flask with 16 ml of 4N sodium hydroxide solution and 8.22 ml of epibromohydrin. The total volume of the reaction mixture was 200 ml. The activated Sepharose was subsequently filtered off, washed neutral with water and transferred back into the reaction vessel. 6.5 g of N-[5-amino-1-carboxypentyl]-iminodiacetic acid were dissolved in 50 ml of water and added to the activated Sepharose together with 10.6 g of solid sodium carbonate. The mixture was stirred slowly at 60°C overnight. The resulting chelate resin with the formula $[\text{Sepharose}^{\text{®}}\text{CL-6B}]\text{-O-CH}_2\text{-CH}(\text{OH})\text{-CH}_2\text{-NH-}(\text{CH}_2)_4\text{-CH}(\text{COOH})\text{-N-}(\text{CH}_2\text{COOH})_2$ (NTA resin) was subsequently washed in a chromatography column in succession with 500 ml of water, 100 ml of aqueous $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$ (2 wt.%), 200 ml of water, 200 ml of 0.2M acetic acid (containing 0.2M NaCl and 0.1 wt./vol.% Tween 20) and 200 ml of water. The nickel ion

concentration in the resulting chelate resin of the formula [Sephacel[®] CL-6B]-O-CH₂-CH(OH)-CH₂-NH-(CH₂)₄-CH(COOH)-N(CH₂COO⁻)₂Ni²⁺ amounted to about 7.1 micromol/ml.

5

Example 18

Metal chelate affinity chromatography with purified IFN-γ

10 A column (diameter 1.6 cm, length = 7.0 cm) was filled with metal-free chelate resin of the formula [Sephacel[®] CL-6B]-O-CH₂-CH(OH)-CH₂-NH-(CH₂)₄-CH(COOH)-N(CH₂COOH)₂ (NTA resin) and the resin was brought into the nickel form by rinsing with a three-fold column volume of 0.1M NiSO₄•5H₂O and subsequently washing
15 with a three-fold column volume of 0.2M acetic acid. The resin was subsequently equilibrated with 0.1M Tris•HCl buffer (pH 7.5) and 0.5M NaCl (throughflow in each case 60 ml/hr.).

20

1 mg of purified IFN-γ (Example 3, amino acid sequence see Fig. 40) was taken up in 3 ml of equilibration buffer and applied to the column. By means of enzyme immunoassay [Gallati, H., J. Clin. Chem. Clin. Biochem.
25 20, 907-914 (1982)] it could be detected that in spite of the two internal protein structural elements Gly-His-Ser and Ile-His-Glu no binding to the NTA column took place.

Example 19

30

Purification of His₂-Xa-IFN-γ by means of NTA resin

E. coli M15 cells containing the plasmids pDMI.1 and pHis₂-Xa-IFN-γ (Example 6) were left to grow in
35 1 litre of LB medium containing 100 µg/ml of ampicillin and 25 µg/ml of kanamycin at 37°C up to an optical

density of $OD_{600} = 0.6$. Then, IPTG was added (final concentration 0.5mM) and the cells were incubated for a further 4 hours. Subsequently, the cells were separated (5 g wet weight) from the culture medium by centrifugation (4000 x g, 10 min, 4°C) and extracted with 15 ml of 7M guanidine•HCl and 0.01M sodium borate (pH 8) (1 hr, 4°C, magnetic stirrer). The thus-obtained crude extract was centrifuged (10,000 x g, 15', 4°C), the supernatant was diluted 10-fold with 0.1M Tris•HCl buffer (pH 7.5) and 0.5M NaCl, again centrifuged (10,000 x g, 15', 4°C) and pumped up on the same NTA column as described in Example 18. The column was subsequently washed with equilibrating buffer until the UV detector (280 nm) again showed the base value. The elution of the His,His-Xa-IFN- γ was effected by lowering the pH value to 5.5. By means of enzyme immunoassay [Gallati, H., supra] it could be detected that this protein was quantitatively absorbed on the NTA column and was eluted only by lowering the pH value. By means of SDS-polyacrylamide gel electrophoresis and RP-18 HPLC it could be detected that the protein obtained was pure His,His-Xa-IFN- γ (purity > 90%). The expected amino-terminal sequence Met-His-His-Ala-Gly-Ile-Glu-Gly-Arg-Gln... was confirmed by Edman degradation.

Example 20

Purification of His,His-Ek-IFN- γ (-8) by means of NTA resin

His,His-Ek-IFN- γ (-8) (Example 7) was expressed in E. coli in an analogous manner to Example 19, extracted and purified via the NTA column. This fusion protein was also bonded to the NTA column at pH 7.5 and was eluted in pure form (purity > 90%) by lowering the pH value to 5.5. The expected sequence Met-His-His-Ala-Gly-Asp-Asp-Asp-Asp-Lys-Gln.... was confirmed by Edman degradation.

Example 21

Purification of His-His-Xa-IFN- γ (-8)(Asn) by means of NTA resin

5 His,His-Xa-IFN- γ (-8)(Asn) (Example 8) was expressed in E. coli in an analogous manner to Example 19, extracted and purified via the NTA column. This protein was also bonded to the NTA column at pH 7.5 and was eluted in pure
10 form (purity > 90%) by lowering the pH value to 5.5.

1 mg of the thus-obtained His-His-Xa-IFN- γ (-8)(Asn) was dialyzed against 0.1M Tris•HCl (pH 7.5), 0.5M NaCl and 1mM CaCl₂. The dialysate (5 ml) was treated with
15 100 μ l (= 1U) of coagulation factor Xa (Boehringer/Mannheim) and incubated for 16 hours at 22°C. The enzymatic degradation of the His,His-affinity peptide was determined by means of SDS-polyacrylamide gel electrophoreses.

20 For the purpose of separating salts, bovine serum albumin (constituent of the commercial factor Xa preparation) and factor Xa, the incubation mixture was firstly dialyzed against water, then lyophilized and subsequently
25 chromatographed on a RP-18 HPLC column (Nucleosil 5C18 column from Brownlee Labs, running agent 0.1% trifluoroacetic acid, gradient with acetonitrile, throughflow 1 ml/min). The resulting purified protein was then freed from solvent and subjected to an Edman degradation. The
30 expected amino-terminal sequence Gln-Asn-Pro-Tyr..... could be confirmed by means of this method.

This experiment shows that the affinity sequence at the NH₂ terminus of the His-His-Xa-IFN- γ (-8)(Asn) can
35 be cleaved off cleanly after the metal chelate affinity chromatography.

Example 22

Purification of (His)₆-mDHFR by means of NTA resin in 6M guanidine•HCl

5 (His)₆-mDHFR (Example 9) was expressed in E. coli in an analogous manner to Example 19. The cells were extracted with 6M guanidine•HCl in 0.1M sodium phosphate buffer (pH 8.0) (5 ml of buffer solution per 1 g of cells,
10 1 hr., 22°C, magnetic stirrer). The thus-produced crude extract was subsequently centrifuged and the supernatant was pumped on to the same NTA column as described in Example 18. The chromatography was carried out analogously to Example 19 with the exception of the buffer solutions
15 used. The buffers used contained in each case 6M guanidine•HCl in 0.1M sodium phosphate buffer having the following pH values: pH 8.0 in order to apply the proteins, pH 6.0 in order to wash-out the non-bound E. coli proteins and pH 4.5 in order to elute the (His)₆-
20 -mDHFR. The eluate obtained was dialyzed against water and subsequently lyophilized. By means of SDS-polyacrylamide gel electrophoresis it could be detected that the protein obtained was pure (His)₆-mDHFR (purity >90%). The expected sequence Met-Arg-Gly-Ser-His-His-His-His-His-
25 -Gly-Ser-Ile-Met... was confirmed by Edman degradation.

Example 23

Purification of (His)₄-mDHFR-(His)₄ by means of NTA resin in 6M guanidine•HCl

(His)₄-mDHFR-(His)₄ (Example 16) was expressed in E. coli in an analogous manner to Example 19, extracted and purified via the NTA column. In place of the stepwise
35 gradients used in Example 22 there was used for the elution a linear pH gradient (pH 8.0 to pH 4.0, 2 hours).

(His)₄-mDHFR-(His)₄ fusion protein was eluted at pH 4.9 and had a purity of at least 90%.

Example 24

5

Purification of Met-mDHFR-(His)₆ by means of NTA resin in 6M urea

Met-mDHFR-(His)₆ (Example 14) was expressed in E. coli in an analogous manner to Example 19. The centrifuged-off cells were extracted with 6M urea in 0.05M sodium phosphate buffer (pH 7.5) (1 g of cells per 10 ml of buffer solution) and ultrasound (10 minutes). After centrifugation of the cell debris the supernatant was applied on to a NTA column (4.5 cm x 2.6 cm) equilibrated with extraction buffer. After washing the column with extraction buffer the Met-mDHFR-(His)₆ fusion protein was eluted with a linear pH gradient of pH 7.5 (extraction buffer) to pH 4.8 (0.05M sodium phosphate buffer containing 6M urea) during 5 hours and at a pump velocity of 18 ml per hour. The fractions which contained protein were analyzed by means of SDS-polyacrylamide gel electrophoresis. 9 mg of Met-mDHFR-(His)₆ fusion protein with a purity of >90% were obtained.

25

Example 25

Purification of Met-mDHFR-(His)₂ by means of NTA resin

Met-mDHFR-(His)₂ (Example 15) was expressed in E. coli in an analogous manner to Example 19. The centrifuged-off cells in 0.05M potassium phosphate buffer (pH 8.0), containing 0.1M potassium chloride and 0.1% Tween 20, were treated with ultrasound in an ice bath during 15 minutes (1 g of cells per 10 ml of buffer solution). Subsequently, the cell debris were centrifuged off

35

and the clear supernatant was placed on a NTA column (4.6 cm x 2.6 cm) equilibrated with extraction buffer. The column was washed with extraction buffer and the Met-mDHFR-(His)₂ fusion protein was eluted with a linear pH gradient of pH 8.0 (extraction buffer) to pH 5.0 (0.05M potassium phosphate buffer containing 0.1M potassium chloride and 0.1% Tween 20) during 10 hours and at a pump velocity of 50 ml per hour. The peak fractions of the eluent were analyzed by means of SDS-polyacrylamide gel electrophoresis. 7 mg of Met-mDHFR-(His)₂ fusion protein with a purity of >85% were obtained.

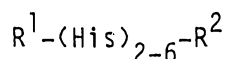
3 mg of the thus-obtained fusion protein Met-mDHFR-(His)₂ were dialyzed against 0.05M Tris•HCl (pH 6.0) at 6°C. The protein solution was then adjusted to pH 9.0 with 0.5M NaOH and incubated at 37°C in the presence of 8.5 units of carboxypeptidase A from bovine pancreas (Serva, Feinbiochemica, Heidelberg BRD). After 0, 15, 30, 90 and 180 minutes samples were removed and analyzed for their histidine content by means of HPLC. After 480 minutes the pH value was lowered to 8.0 and the reaction mixture was pumped on to a NTA column equilibrated with 0.05M potassium phosphate buffer (pH 8). The protein contained in the reaction mixture was detected in the flow-through by means of SDS-polyacrylamide gel electrophoresis. In addition, an amount of histidine residues increasing with time was detected in the samples which had been removed from the protein solution after 15, 30, 90 and 180 minutes.

This experiment shows that the affinity sequence at the carboxyl terminus can be removed cleanly after the purification on NTA resin.

The claims defining the invention are as follows:

1. Fusion proteins consisting of one or two affinity peptides which contain 2-6 adjacent histidine residues and a biologically active polypeptide or protein covalently linked directly or indirectly to this/these affinity peptide(s).

2. Fusion proteins in accordance with claim 1, wherein the affinity peptide(s) has/have the formula



wherein R¹ represents hydrogen or 1 to 4 amino acids, R² represents Q, Q-Ile-Glu-Gly-Arg- or Q-Asp-Asp-Asp-Lys- wherein Q is a peptide bond, an amino acid or a sequence of 2-30 amino acids and wherein one affinity peptide is covalently linked directly or indirectly to the amino terminal amino acid or to the carboxy terminal amino acid of the biologically active peptide or protein.

3. Fusion proteins in accordance with claim 1 or 2, wherein the affinity peptide(s) has/have a peptide sequence of the formula

Met-His-His,

Met-His-His-His,

Met-His-His-His-His,

Met-His-His-His-His,

Met-His-Hi: ' 's-His-His,

Met-His-His Gly-Ile-Glu-Gly-Arg,

Met-His-His- α -Gly-Asp-Asp-Asp-Asp-Lys.

4. Fusion proteins in accordance with any one of claims 1-3, wherein one affinity peptide is covalently linked directly or indirectly to the amino terminal amino acid of the biologically active polypeptide or protein and a further affinity peptide is covalently linked to the carboxy terminal amino acid of the biologically active polypeptide or protein. ✓

5. Fusion proteins in accordance with any one of claims 1 to 4, wherein the affinity peptides complex immobilized nickel ions.

6. Fusion proteins in accordance with any one of claims 1 to 5, wherein the biologically active polypeptide or protein has the amino acid sequence of a human immune interferon or partial sequences thereof or the amino acid sequence of mouse dihydrofolate reductase.



7. A bacterially prepared fusion protein in accordance with any one of claims 1 to 6.

8. A fusion protein in accordance with any one of claims 1 to 6 prepared by E. coli.

9. A fusion protein in accordance with any one of claims 1 to 8 in homogeneous form.

10. Genes which code for a fusion protein in accordance with any one of claims 1 to 8.

11. Expression vectors in which a gene in accordance with claim 10 is operatively covalently linked to an expression control sequence.

12. Expression vectors in accordance with claim 11 which can replicate in a gram-negative bacterium.

13. Expression vectors in accordance with claim 12 which can replicate in E. coli.

14. A bacterium transformed with an expression vector in accordance with any one of claims 11 to 13.

15. An E. coli strain transformed with an expression vector in accordance with any one of claims 11 to 13.

16. E. coli M15 strain transformed with an expression vector in accordance with any one of claims 11 to 13.

17. A process for the preparation of fusion proteins in accordance with any one of claims 1 to 8, which process comprises transforming a bacterial host with an expression vector in accordance with any one of claims 11 to 13, culturing the transformant under appropriate conditions of growth and isolating said fusion proteins.

18. A process for the purification of a biologically active polypeptide or protein, which process comprises purifying it as a fusion protein in accordance with any one of claims 1 to 8 by means of a process comprising affinity chromatography by NTA resins and then removing the affinity peptide(s) by selective cleavage.

19. A process in accordance with claim 18, wherein a protease is used for the cleavage of the affinity peptide(s).

20. A process in accordance with claim 19, wherein the protease is factor Xa.

21. Vaccines containing a fusion protein in accordance with any one of claims 1 to 9 and a physiologically compatible carrier material.

22. Expression vectors in which a gene in accordance with claim 10



is operatively covalently linked to an expression control sequence, which vectors are substantially as herein described with reference to any one of Example 5 and Fig. 17, Example 6 and Fig. 21, Example 7 and Fig. 24, Example 8 and Fig. 26, Example 9 and Fig. 28, Example 10 and Fig. 30, Example 11 and Fig. 32, Example 12 and Fig. 34, Example 13 and Fig. 36, Example 14 and Fig. 37, Example 15 and Fig. 38 or Example 16 and Fig. 39.

23. Fusion proteins consisting of one or two affinity peptides which contain 2-6 adjacent histidine residues and a biologically active polypeptide or protein covalently linked directly or indirectly to this/these affinity peptide(s), which proteins are substantially as herein described with reference to any one of Examples 19 to 25.

24. A process for the preparation of fusion proteins, which process is substantially as herein described with reference to any one of Examples 3 to 16.

25. A process for the purification of fusion proteins, which process is substantially as herein described with reference to any one of Examples 19 to 25.

26. Vaccines containing a fusion protein as defined in claim 23 together with a physiologically compatible carrier material.

DATED this SEVENTH day of FEBRUARY 1991
F Hoffmann-La Roche & Co Aktiengesellschaft

Patent Attorneys for the Applicant
SPRUSON & FERGUSON



Fig. 1

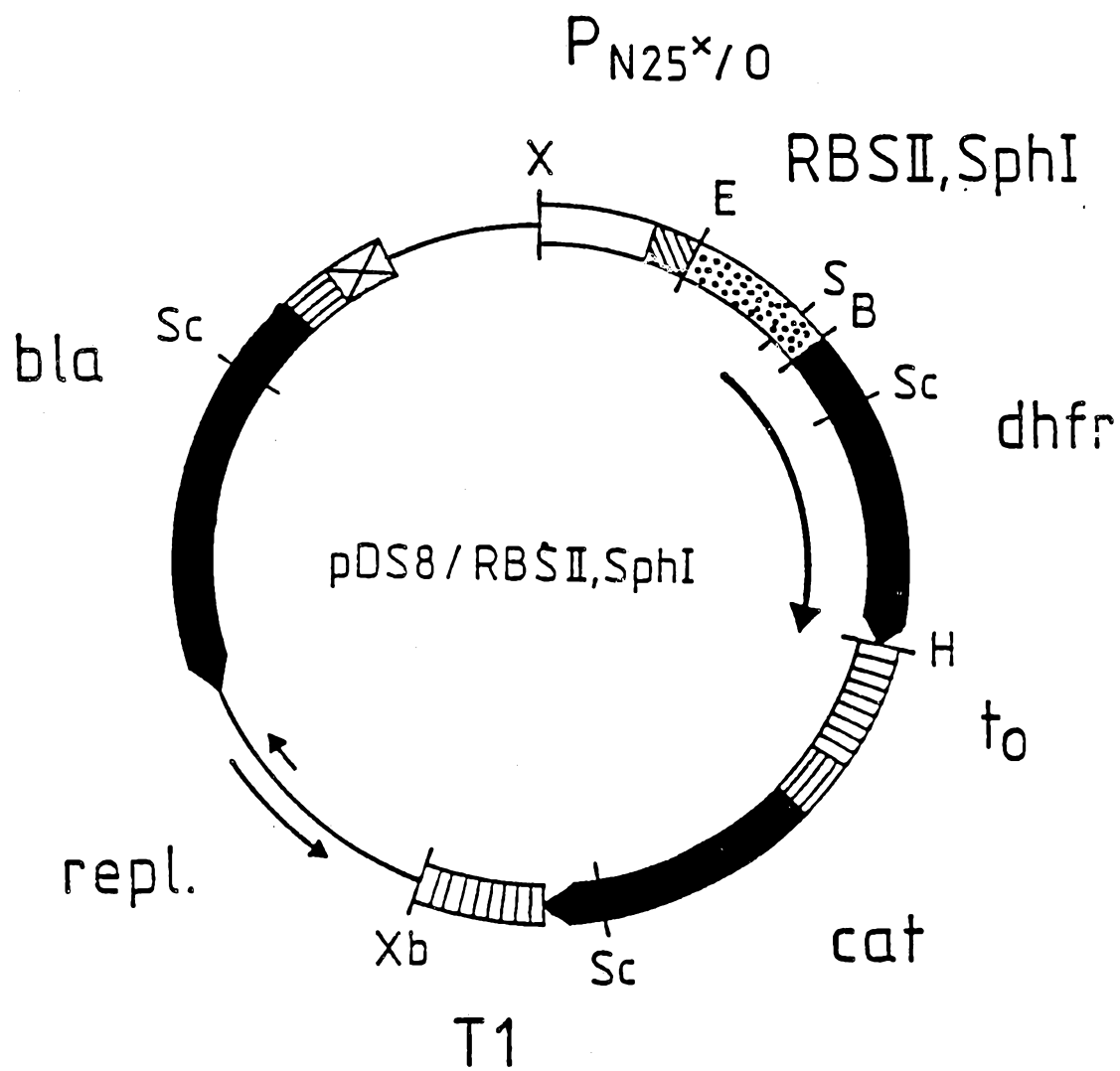


Fig. 2

	10	20	30	40	50
0	<u>CCTCGAGGCT</u>	GGCATCCCTA	ACATATCCGA	ATGGTTACTT	AAACAACGGA
50	GGACTAGCGT	ATCCCTTTCG	ATAGGGTTTG	AGTTAGATAA	AGTATATGCT
100	GAAC TTTCTT	CTTTGCTCAA	AGAATCATAA	AAAATTTATT	TGCTTTCAGG
150	AAAATTTTTC	TGTATAATAG	ATTCAAATTG	TGAGCGGATA	ACAATTTGAA
200	<u>TTCA</u> TTAAAG	AGGAGAAATT	<u>AAGCATGCGA</u>	<u>GGATCCGGCA</u>	TCATGGTTCG
250	<u>ACCATTGAAC</u>	TGCATCGTCG	CCGTGTCCCA	AAATATGGGG	ATTGGCAAGA
300	<u>ACGGAGACCT</u>	ACCCTGGCCT	CCGCTCAGGA	ACGAGTTCAA	<u>GTA</u> CTTCCAA
350	<u>AGAATGACCA</u>	CAACCTCTTC	AGTGGAAGGT	AAACAGAATC	TGGTGATTAT
400	<u>GGGTAGGAAA</u>	ACCTGGTTCT	CCATTCCCTGA	GAAGAATCGA	CCTTTAAAGG
450	<u>ACAGAA</u> TTAA	TATAGTTCTC	AGTAGAGAAC	TCAAAGAACC	ACCACGAGGA
500	<u>GCTC</u> ATTTTC	TTGCCAAAAG	TTTGGATGAT	GCCTTAAGAC	TTATTGAACA
550	<u>ACC</u> GGAATTG	GCAAGTAAAG	TAGACATGGT	TTGGATAGTC	GGAGGCAGTT
600	<u>CTG</u> TTTACCA	GGAAGCCATG	AATCAACCAG	GCCACCTTAG	ACTCTTTGTG
650	<u>ACA</u> AGGATCA	TGCAGGAATT	TGAAAGTGAC	ACGTTTTTCC	CAGAAATTGA
700	<u>TTT</u> GGGGAAA	TATAAACTTC	TCCCAGAATA	CCCAGGCGTC	CTCTCTGAGG
750	<u>TCC</u> AGGAGGA	AAAAGGCATC	AAGTATAAGT	TTGAAGTCTA	CGAGAAGAAA
800	<u>GAC</u> TAACAGG	AAGATGCTTT	CAAGTTCTCT	GCTCCCCTCC	TAAAGCTATG
850	<u>CAT</u> TTTTATA	AGACCATGGG	ACTTTTGCTG	GCTTTAGATC	CGGCCAAGCT
900	<u>TGG</u> ACTCCTG	TTGATAGATC	CAGTAATGAC	CTCAGAACTC	CATCTGGATT
950	TGTTCAGAAC	GCTCGGTTGC	CGCCGGGCGT	TTTTTATTGG	TGAGAATCCA
1000	AGCTAGCTTG	GCGAGATTTT	CAGGAGCTAA	GGAAGCTAAA	ATGGAGAAAA
1050	AAATCACTGG	ATATACCACC	GTTGATATAT	CCCAATGGCA	TCGTAAAGAA
1100	CATTTTGAGG	CATTTTCAGTC	AGTTGCTCAA	TGTACCTATA	ACCAGACCGT
1150	TCAGCTGGAT	ATTACGGCCT	TTTTAAAGAC	CGTAAAGAAA	AATAAGCACA

Fig. 2 (continued)

	10	20	30	40	50
1200	AGTTTTATCC	GGCCTTTATT	CACATTCTTG	CCCGCCTGAT	GAATGCTCAT
1250	CCGGAATTTT	GTATGGCAAT	GAAAGACGGT	GAGCTGGTGA	TATGGGATAG
1300	TGTTACCCCT	TGTTACACCG	TTTTCCATGA	GCAAACCTGAA	ACGTTTTTCAT
1350	CGCTCTGGAG	TGAATACCAC	GACGATTTCC	GGCAGTTTCT	ACACATATAT
1400	TCGCAAGATG	TGGCGTGTTA	CGGTGAAAAC	CTGGCCTATT	TCCCTAAAGG
1450	GTTTATTGAG	AATATGTTTT	TCGTCTCAGC	CAATCCCTGG	GTGAGTTTCA
1500	CCAGTTTTGA	TTTAAACGTG	GCCAATATGG	ACAACTTCTT	CGCCCCCGTT
1550	TTCACCATGG	GCAAATATTA	TACGCAAGGC	GACAAGGTGC	TGATGCCGCT
1600	GGCGATTTCAG	GTTTCATCATG	CCGTCTGTGA	TGGCTTCCAT	GTCGGCAGAA
1650	TGCTTAATGA	ATTACAACAG	TACTGCGATG	AGTGGCAGGG	CGGGGCGTAA
1700	TTTTTTTAAAG	GCAGTTATTG	GTGCCCTTAA	ACGCCTGGGG	TAATGACTCT
1750	CAGCTTGAG	GCATCAAATA	AAACGAAAGG	CTCAGTCGAA	AGACTGGGCC
1800	TTTCGTTTTA	TCTGTTGTTT	GTGGGTGAAC	GCTCTCCTGA	GTAGGACAAA
1850	TCCGCCGCTC	TAGAGC			

2068

————pBR322————A

4358

Fig. 3

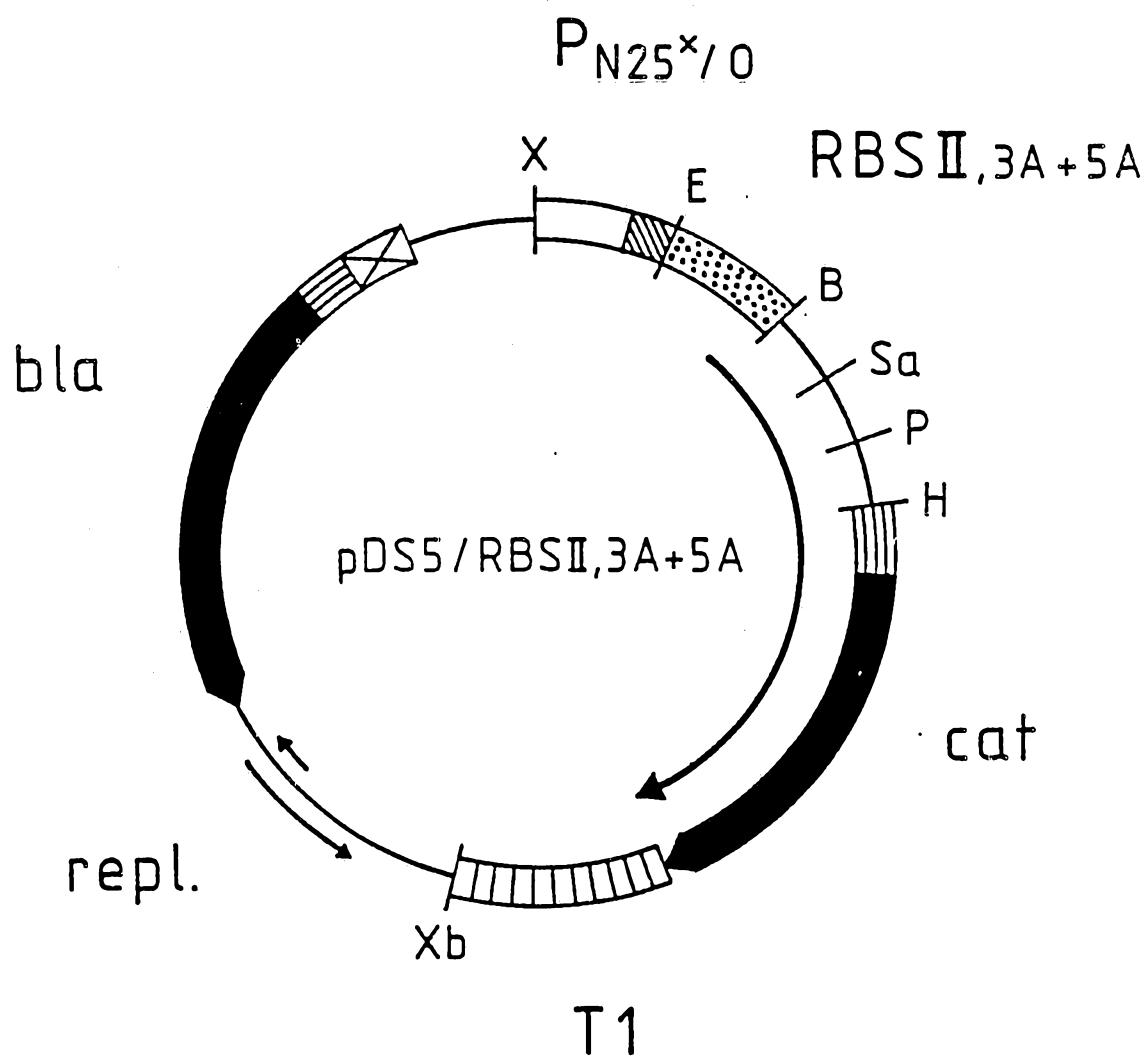


Fig. 4

	10	20	30	40	50
0	<u>CCTCGAGGCT</u>	GGCATCCCTA	ACATATCCGA	ATGGTTACTT	AAACAACGGA
50	GGACTAGCGT	<u>ATCCCTTCGC</u>	ATAGGGTTTG	AGTTAGATAA	AGTATATGCT
100	GAACTTTCTT	CTTTGCTCAA	AGAATCATAA	AAAATTTATT	<u>TGCTTTCAGG</u>
150	AAAATTTTTT	TGTATAATAG	<u>ATTCAAATTG</u>	TGAGCGGATA	ACAATTTGAA
200	<u>TTCAATTAAAG</u>	AGGAGAAATT	AACTATGAGG	<u>GGATCCGTCG</u>	<u>ACCTGCAGCC</u>
250	<u>AAGCTTGGCG</u>	AGATTTTCAG	GAGCTAAGGA	AGCTAAAATG	GAGAAAAAAA
300	<u>TCACTGGATA</u>	<u>TACCACCGTT</u>	GATATATCCC	AATGGCATCG	<u>TAAAGAACAT</u>
350	<u>TTTGAGGCAT</u>	<u>TTCAGTCAGT</u>	<u>TGCTCAATGT</u>	<u>ACCTATAACC</u>	<u>AGACCGTTCA</u>
400	<u>GCTGGATATT</u>	<u>ACGGCCTTTT</u>	<u>TAAAGACCGT</u>	<u>AAAGAAAAAT</u>	<u>AAGCACAAGT</u>
450	<u>TTTATCCGGC</u>	<u>CTTTATTAC</u>	<u>ATTCTTGCCC</u>	<u>GCCTGATGAA</u>	<u>TGCTCATCCG</u>
500	<u>GAATTTGTA</u>	<u>TGGCAATGAA</u>	<u>AGACGGTGAG</u>	<u>CTGGTGATAT</u>	<u>GGGATAGTGT</u>
550	<u>TCACCCTTGT</u>	<u>TACACCGTTT</u>	<u>TCCATGAGCA</u>	<u>AACTGAAACG</u>	<u>TTTTCATCCG</u>
600	<u>TCTGGAGTGA</u>	<u>ATACCACGAC</u>	<u>GATTTCCGGC</u>	<u>AGTTTCTACA</u>	<u>CATATATTCT</u>
650	<u>CAAGATGTGG</u>	<u>CGTGTTACGG</u>	<u>TGAAAACCTG</u>	<u>GCCTATTTCC</u>	<u>CTAAAGGGTT</u>
700	<u>TATTGAGAAT</u>	<u>ATGTTTTTTC</u>	<u>TCTCAGCCAA</u>	<u>TCCCTGGGTG</u>	<u>AGTTTCACCA</u>
750	<u>GTTTTGATTT</u>	<u>AAACGTGGCC</u>	<u>AATATGGACA</u>	<u>ACTTCTTCGC</u>	<u>CCCCGTTTTT</u>
800	<u>ACCATGGGCA</u>	<u>AATATTATAC</u>	<u>GCAAGGCGAC</u>	<u>AAGGTGCTGA</u>	<u>TGCCGCTGGC</u>
850	<u>GATTCAGGTT</u>	<u>CATCATGCCG</u>	<u>TCTGTGATGG</u>	<u>CTTCCATGTC</u>	<u>GGCAGAATGC</u>
900	<u>TTAATGAATT</u>	<u>ACAACAGTAC</u>	<u>TGCGATGAGT</u>	<u>GGCAGGGCGG</u>	<u>GGCGTAATTT</u>
950	<u>TTTTAAGGCA</u>	<u>GTTATTTGGT</u>	<u>CCCTTAAACG</u>	<u>CCTGGGGTAA</u>	<u>TGACTCTCTA</u>
1000	<u>GCTTGAGGCA</u>	<u>TCAAATAAAA</u>	<u>CGAAAGGCTC</u>	<u>AGTCGAAAGA</u>	<u>CTGGGCCTTT</u>
1050	<u>CGTTTTATCT</u>	<u>GTTGTTTGTC</u>	<u>GGTGAACGCT</u>	<u>CTCCTGAGTA</u>	<u>GGACAAATCC</u>
1100	<u>GCCGCTCTAG</u>	<u>AGCTGCCTCG</u>	<u>CGCGTTTTCG</u>	<u>TGATGACGGT</u>	<u>GAAAACCTCT</u>
1150	<u>GACACATGCA</u>	<u>GCTCCCGGAG</u>	<u>ACGGTCACAG</u>	<u>CTTGTCTGTA</u>	<u>AGCGGATGCC</u>

Fig. 4 (continued)

	10	20	30	40	50
1200	GGGAGCAGAC	AAGCCCGTCA	GGGCGCGTCA	GCGGGTGTTG	GCGGGTGTCG
1250	GGGCGCAGCC	ATGACCCAGT	CACGTAGCGA	TAGCGGAGTG	TATACTGGCT
1300	TAACTATGCG	GCATCAGAGC	AGATTGTACT	GAGAGTGCAC	CATATGCGGT
1350	GTGAAATACC	GCACAGATGC	GTAAGGAGAA	AATACCGCAT	CAGGCGCTCT
1400	TCCGCTTCCT	CGCTCACTGA	CTCGCTGCGC	TCGGTTCGTT	GGCTGCGGCG
1450	AGCGGTATCA	GCTCACTCAA	AGGCGGTAAT	ACGGTTATCC	ACAGAATCAG
1500	GGGATAACGC	AGGAAAGAAC	ATGTGAGCAA	AAGGCCAGCA	AAAGGCCAGG
1550	AACCGTAAAA	AGGCCGCGTT	GCTGGCGTTT	TTCCATAGGC	TCCGCCCCCC
1600	TGACGAGCAT	CACAAAAATC	GACGCTCAAG	TCAGAGGTGG	CGAAACCCGA
1650	CAGGACTATA	AAGATACCAG	GCGTTTCCCC	CTGGAAGCTC	CCTCGTGCGC
1700	TCTCCTGTTT	CGACCCTGCC	GCTTACCGGA	TACCTGTCCG	CCTTTCTCCC
1750	TTCGGGAAGC	GTGGCGCTTT	CTCAATGCTC	ACGCTGTAGG	TATCTCAGTT
1800	CGGTGTAGGT	CGTTGCTCC	AAGCTGGGCT	GTGTGCACGA	ACCCCCCGTT
1850	CAGCCCGACC	GCTGCGCCTT	ATCCGGTAAC	TATCGTCTTG	AGTCCAACCC
1900	GGTAAGACAC	GACTTATCGC	CACTGGCAGC	AGCCACTGGT	AACAGGATTA
1950	GCAGAGCGAG	GTATGTAGGC	GGTGCTACAG	AGTTCTTGAA	GTGGTGGCCT
2000	AACTACGGCT	ACACTAGAAG	GACAGTATTT	GGTATCTGCG	CTCTGCTGAA
2050	GCCAGTTACC	TTCGGAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAACAAA
2100	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC
2150	AGAAAAAAG	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA
2200	CGCTCAGTGG	AACGAAAAC	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT
2250	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA	ATTAAAAATG	AAGTTTTAAA
2300	TCAATCTAAA	GTATATATGA	GTAAACTTGG	TCTGACAGTT	ACCAATGCTT
2350	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTTCG	TCATCCATAG
2400	TTGCCTGACT	CCCCGTCTGT	TAGATAACTA	CGATACGGGA	GGGCTTACCA
2450	TCTGGCCCCA	GTGCTGCAAT	GATACCGCGA	GACCCACGCT	CACCGGCTCC
2500	AGATTTATCA	GCAATAAACC	AGCCAGCCGG	AAGGGCCGAG	CGCAGAAGTG
2550	GTCTTGCAAC	TTTATCCGCC	TCCATCCAGT	CTATTAATTG	TTGCCGGGAA

Fig. 4 (continued)

	10	20	30	40	50
2600	GCTAGAGTAA	GTAGTTCGCC	AGTTAATAGT	TTGCGCAACG	TTGTTGCCAT
2650	TGCTACAGGC	ATCGTGGTGT	CACGCTCGTC	GTTTGGTATG	GCTTCATTCA
2700	GCTCCGGTTC	CCAACGATCA	AGGCGAGTTA	CATGATCCCC	CATGTTGTGC
2750	AAAAAAGCGG	TTAGCTCCTT	CGGTCCTCCG	ATCGTTGTCA	GAAGTAAGTT
2800	GGCCGCAGTG	TTATCACTCA	TGGTTATGGC	AGCACTGCAT	AATTCTCTTA
2850	CTGTCATGCC	ATCCGTAAGA	TGCTTTTCTG	TGACTGGTGA	GTA CTCAACC
2900	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT	CTTGCCCGGC
2950	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA
3000	TCATTGGAAG	ACGTTCTTCG	GGGCGAAAAC	TCTCAAGGAT	CTTACCGCTG
3050	TTGAGATCCA	GTTGATGTA	ACCCACTCGT	GCACCGAACT	GATCTTCAGC
3100	ATCTTTTACT	TTCACCAGCG	TTTCTGGGTG	AGCAAAAACA	GGAAGGCAAA
3150	ATGCCGCAAA	AAAGGGAATA	AGGGCGACAC	GGAAATGTTG	AATACTCATA
3200	CTCTTCCTTT	TTCAATATTA	TTGAAGCATT	TATCAGGGTT	ATTGTCTCAT
3250	GAGCGGATAC	ATATTTGAAT	GTATTTAGAA	AAATAAACAA	ATAGGGGTTC
3300	CGCGCACATT	TCCCCGAAAA	GTGCCACCTG	ACGTCTAAGA	AACCATTATT
3350	ATCATGACAT	TAACCTATAA	AAATAGGCGT	ATCACGAGGC	CCTTTCGTCT
3400	TCA				

Fig. 5

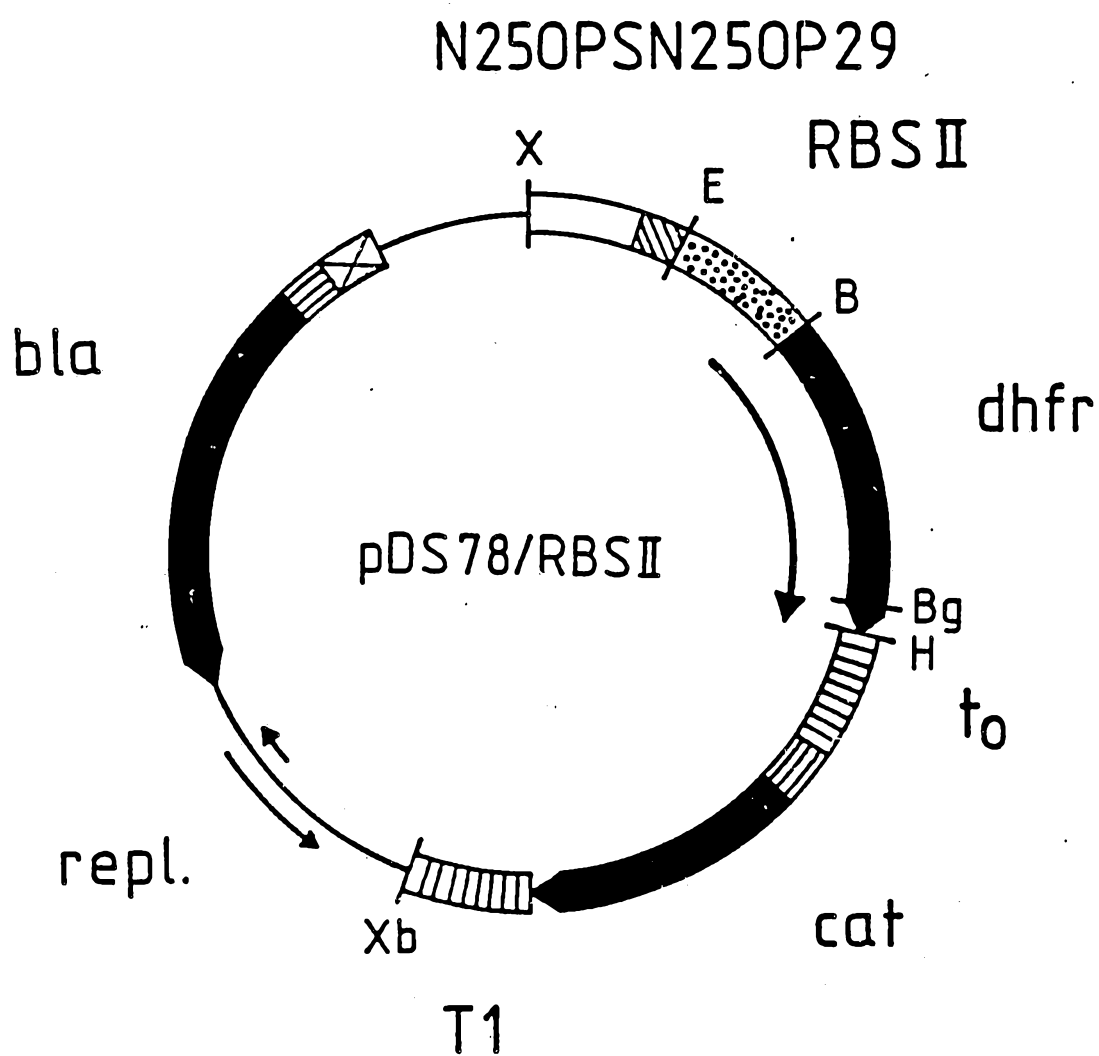


Fig. 6

10 20 30 40 50
1 CTCGAGAAAT | CATAAAAAAT | TTATTTGCTT | TGTGAGCGGA | TAACAATTAT
51 AATAGATTCA ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG
101 AGGAGAAATT AACTATGAGA GGATCCGGCA TCATGGTTCG ACCATTGAAC
151 TGCATCGTCG CCGTGTCCCA AAATATGGGG ATTGGCAAGA ACGGAGACCT
201 ACCCTGGCCT CCGCTCAGGA ACGAGTTCAA GTACTTCCAA AGAATGACCA
251 CAACCTCTTC AGTGGAAGGT AAACAGAATC TGGTGATTAT GGGTAGGAAA
301 ACCTGGTTCT CCATTCCTGA GAAGAATCGA CCTTTAAAGG ACAGAATTAA
351 TATAGTTCTC AGTAGAGAAC TCAAAGAACC ACCACGAGGA GCTCATTTTC
401 TTGCCAAAAG TTTGGATGAT GCCTTAAGAC TTATTGAACA ACCGGAATTG
451 GCAAGTAAAG TAGACATGGT TTGGATAGTC GGAGGCAGTT CTGTTTACCA
501 GGAAGCCATG AATCAACCAG GCCACCTTAG ACTCTTTGTG ACAAGGATCA
551 TGCAGGAATT TGAAAGTGAC ACGTTTTTCC CAGAAATTGA TTTGGGGAAA
601 TATAAACTTC TCCCAGAATA CCCAGGCGTC CTCTCTGAGG TCCAGGAGGA
651 AAAAGGCATC AAGTATAAGT TTGAAGTCTA CGAGAAGAAA GGTTCAGAT
701 CTGTTAACCT ACTTTAACAG GAAGATGCTT TCAAGTTCTC TGCTCCCCTC
751 CTAAAGCTAT GCATTTTTAT AAGACCATGG GACTTTTGCT GGCTTTAGAT
801 CCGGCCAAGC TTGGACTCCT GTTGATAGAT CCAGTAATGA CCTCAGAACT
851 CCATCTGGAT TTGTTCAGAA CGCTCGGTTG CCGCCGGGCG TTTTTTATTG
901 GTGAGAATCC AAGCTAGCTT GGCGAGATTT TCAGGAGCTA AGGAAGCTAA
951 AATGGAGAAA AAAATCACTG GATATACCAC CGTTGATATA TCCCAATGGC
1001 ATCGTAAAGA ACATTTTGAG GCATTTCACT CAGTTGCTCA ATGTACCTAT
1051 AACCAGACCG TTCAGCTGGA TATTACGGCC TTTTAAAGA CCGTAAAGAA
1101 AAATAAGCAC AAGTTTTATC CGGCCTTAT TCACATTCTT GCCCGCCTGA
1151 TGAATGCTCA TCCGGAATTT CGTATGGCAA TGAAAGACGG TGAGCTGGTG

Fig. 6 (continued)

1201 ATATGGGATA GTGTTACCCC TTGTTACACC GTTTTCATG AGCAAACCTGA
1251 AACGTTTTCA TCGCTCTGGA GTGAATACCA CGACGATTTC CGGCAGTTTC
1301 TACACATATA TTCGCAAGAT GTGGCGTGTT ACGGTGAAAA CCTGGCCTAT
1351 TTCCCTAAAG GGTTTATTGA GAATATGTTT TTCGTCTCAG CCAATCCCTG
1401 GGTGAGTTTC ACCAGTTTTG ATTTAAACGT GGCCAATATG GACAACTTCT
1451 TCGCCCCCGT TTTCACCATG GGCAAATATT ATACGCAAGG CGACAAGGTG
1501 CTGATGCCGC TGGCGATTCA GGTTCATCAT GCCGTCTGTG ATGGCTTCCA
1551 TGTCGGCAGA ATGCTTAATG AATTACAACA GTACTGCGAT GAGTGGCAGG
1601 GCGGGGCGTA ATTTTTTTAA GGCAGTTATT GGTGCCCTTA AACGCCTGGG
1651 GTAATGACTC TCTAGCTTGA GGCATCAAAT AAAACGAAAG GCTCAGTCGA
1701 AAGACTGGGC CTTTCGTTTT ATCTGTTGTT TGTCCGTGAA CGCTCTCCTG
1751 AGTAGGACAA ATCCGCCGCT CTAGAGCTGC CTCGCGCGTT TCGGTGATGA
1801 CGGTGAAAAC CTCTGACACA TGCAGCTCCC GGAGACGGTC ACAGCTTGTC
1851 TGTAAGCGGA TGCCGGGAGC AGACAAGCCC GTCAGGGCGC GTCAGCGGGT
1901 GTTGGCGGGT GTCGGGGCGC AGCCATGACC CAGTCACGTA GCGATAACGG
1951 AGTGTATACT GGCTTAACTA TCGGGCATCA GAGCAGATTG TACTGAGAGT
2001 GCACCATATG CGGTGTGAAA TACCGCACAG ATGCGTAAGG AGAAAATACC
2051 GCATCAGGCG CTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC
2101 TGTCGGCTGC GCGGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT
2151 ATCCACAGAA TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC
2201 AGCAAAAGGC CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT
2251 AGGCTCCGCC CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG
2301 GTGGCGAAAC CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCTGGAA
2351 GCTCCCTCGT GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG
2401 TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCAAT GCTCACGCTG
2451 TAGGTATCTC AGTTCGGTGT AGGTCGTTCT CTCCAAGCTG GGCTGTGTGC
2501 ACGAACCCCC CGTTCAGCCC GACCGCTGCG CTTATCCGG TAACTATCGT

Fig. 6 (continued)

2551 CTTGAGTCCA ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC
2601 TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT
2651 TGAAGTGGTG GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC
2701 TGCCTCTGCT TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG
2751 ATCCGGCAAA CAAACCACCG CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC
2801 AGCAGATTAC GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT
2851 TCTACGGGGT CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTTT
2901 GGTCATGAGA TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA
2951 AATGAAGTTT TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC
3001 AGTTACCAAT GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT
3051 TCGTTTATCC ATAGCTGCCT GACTCCCCGT CGTGTAGATA ACTACGATAC
3101 GGGAGGGCTT ACCATCTGGC CCCAGTGCTG CAATGATACC GCGAGACCCA
3151 CGCTCACCGG CTCCAGATTT ATCAGCAATA AACCAGCCAG CCGGAAGGGC
3201 CGAGCGCAGA AGTGGTCCTG CAACTTTATC CGCCTCCATC CAGTCTATTA
3251 ATTGTTGCCG GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA TAGTTTGCGC
3301 AACGTTGTTG CCATTGCTAC AGGCATCGTG GTGTCACGCT CGTCGTTTGG
3351 TATGGCTTCA TTCAGCTCCG GTTCCCAACG ATCAAGGCGA GTTACATGAT
3401 CCCCCATGTT GTGCAAAAAA GCGGTTAGCT CCTTCGGTCC TCCGATCGTT
3451 GTCAGAAGTA AGTTGGCCGC AGTGTTATCA CTCATGGTTA TGGCAGCACT
3501 GCATAATTCT CTTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTGACTG
3551 GTGAGTACTC AACCAAGTCA TTCTGAGAAT AGTGTATGCG GCGACCGAGT
3601 TGCTCTTGCC CGGCGTCAAT ACGGGATAAT ACCGCGCCAC ATAGCAGAAC
3651 TTAAAAGTG CTCATCATTG GAAAACGTTT TTCGGGGCGA AACTCTCAA
3701 GGATCTTACC GCTGTTGAGA TCCAGTTCGA TGTAACCCAC TCGTGCACCC
3751 AACTGATCTT CAGCATCTTT TACTTTCACC AGCGTTTCTG GGTGAGCAAA
3801 AACAGGAAGG CAAAATGCCG CAAAAAGGG AATAAGGGCG ACACGGAAAT
3851 GTTGAATACT CATACTCTTC CTTTTTCAAT ATTATTGAAG CATTTATCAG
3901 GGTATTGTGC TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA

Fig. 6 (continued)

3951 ACAAATAGGG GTTCCGCGCA CATTTCCCCG AAAAGTGCCA CCTGACGTCT
4001 AAGAAACCAT TATTATCATG ACATTAACCT ATAAAAATAG GCGTATCACG
4051 AGGCCCTTTC GTCTTCAC

Fig. 7

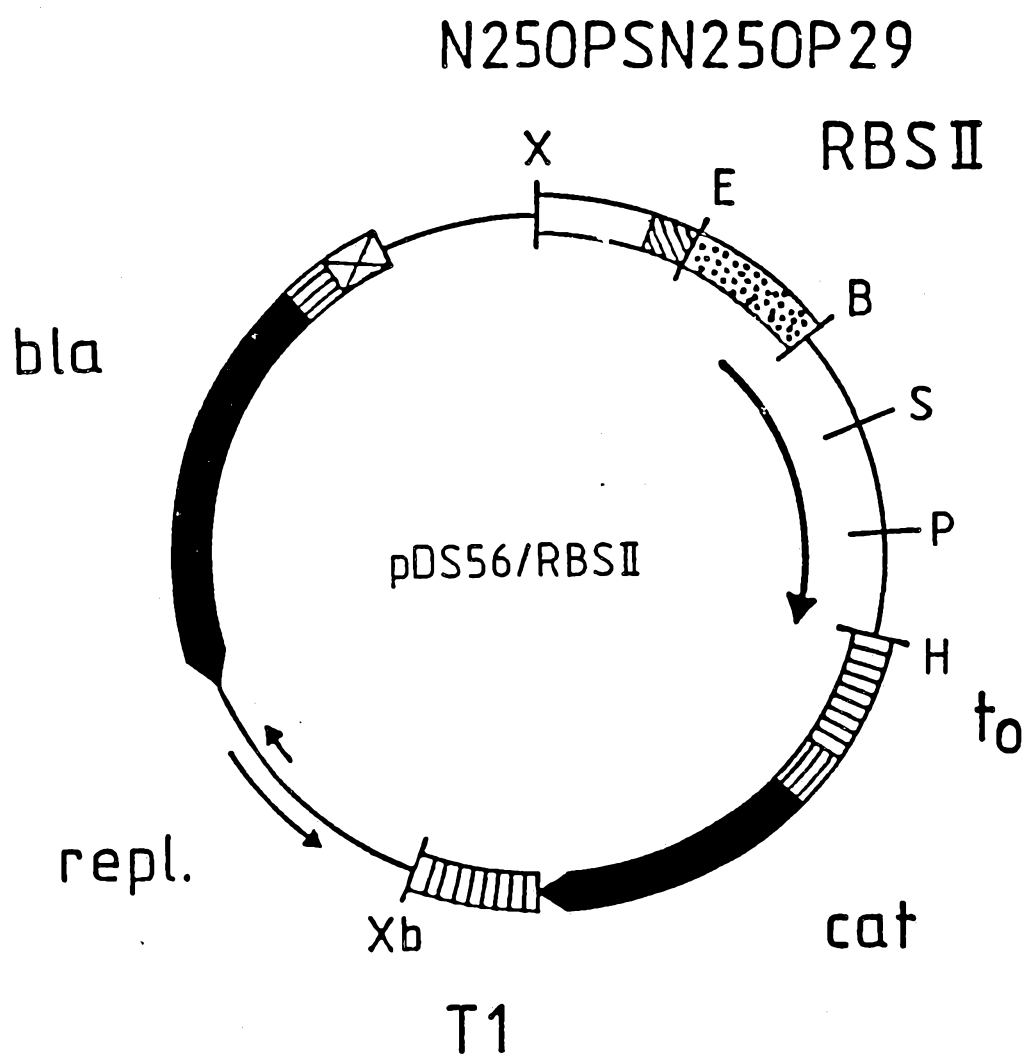


Fig. 8

	10	20	30	40	50
1	<u>CTCGAGAAAT</u>	CATAAAAAAT	TTATTTGCTT	TGTGAGCGGA	TAACAATTAT
51	AATAGATTCA	ATTGTGAGCG	GATAACAATT	TCACACAGAA	<u>TTCATTAAG</u>
101	AGGAGAAATT	AAC <u>TATGAGA</u>	<u>GGATCCGTCG</u>	<u>ACCTGCAGCC</u>	<u>AAGCTTAATT</u>
151	<u>AGCTGAGCTT</u>	GGACTCCTGT	TGATAGATCC	AGTAATGACC	TCAGAACTCC
201	ATCTGGATTT	GTTCAGAACG	CTCGGTTGCC	GCCGGGCGTT	TTTTATTGGT
251	GAGAATCCAA	GCTAGCTTGG	CGAGATTTTC	AGGAGCTAAG	GAAGCTAAAA
301	TGGAGAAAAA	AATCACTGGA	TATACCACCG	TTGATATATC	CCAATGGCAT
351	CGTAAAGAAC	ATTTTGAGGC	ATTTCAGTCA	GTTGCTCAAT	GTACCTATAA
401	CCAGACCGTT	CAGCTGGATA	TTACGGCCTT	TTTAAAGACC	GTAAAGAAAA
451	ATAAGCACAA	GTTTTATCCG	GCCTTTATTC	ACATTCTTGC	CCGCCTGATG
501	AATGCTCATC	CGGAATTTTCG	TATGGCAATG	AAAGACGGTG	AGCTGGTGAT
551	ATGGGATAGT	GTTACCCCTT	GTTACACCGT	TTTCCATGAG	CAAAC <u>TGAAA</u>
601	CGTTTTTCATC	GCTCTGGAGT	GAATACCACG	ACGATTTCCG	GCAGTTTCTA
651	CACATATATT	CGCAAGATGT	GGCGTGTTAC	GGTGAAAACC	TGGCCTATTT
701	CCCTAAAGGG	TTTATTGAGA	ATATGTTTTT	CGTCTCAGCC	AATCCCTGGG
751	TGAGTTTCAC	CAGTTTTTGAT	TTAAACGTGG	CCAATATGGA	CAACTTCTTC
801	GGCCCCGTTT	TCACCATGGG	CAAATATTAT	ACGCAAGGCG	ACAAGGTGCT
851	GATGCCGCTG	GCGATTCAGG	TTCATCATGC	CCTCTGTGAT	GGCTTCCATG
901	TCGGCAGAAT	GCTTAATGAA	TTACAACAGT	ACTGCGATGA	GTGGCAGGGC
951	GGGGCGTAAT	TTTTTTAAGG	CAGTTATTGG	TGCCCTTAAA	CGCCTGGGGT
1001	AATGACTCTC	TAGCTTGAGG	CATCAAATAA	AACGAAAGGC	TCAGTCGAAA
1051	GACTGGGCCT	TTCGTTTTTAT	CTGTTGTTTG	TCGGTGAACG	CTCTCCTGAG
1101	TAGGACAAAT	CCGCCGCTCT	<u>AGAGCTGCCT</u>	CGCGCGTTTC	GGTGATGACG
1151	GTGAAAACCT	CTGACACATG	CAGCTCCCGG	AGACGGTCAC	AGCTTGTCTG

Fig. 8 (continued)

1201 TAAGCGGATG CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT
1251 TGGCGGGTGT CGGGGCGCAG CCATGACCCA GTCACGTAGC GATAGCGGAG
1301 TGTATACTGG CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC
1351 ACCATATGCG GTGTGAAATA CCGCACAGAT GCGTAAGGAG AAAATACCGC
1401 ATCAGGCGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCTG
1451 TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA ATACGGTTAT
1501 CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG
1551 CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG
1601 GCTCCGCCCC CCTGACGAGC ATCACAAAAA TCGACGCTCA AGTCAGAGGT
1651 GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC
1701 TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC
1751 CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC TCACGCTGTA
1801 GGTATCTCAG TTCGGTGTAG GTCGTTGCT CCAAGCTGGG CTGTGTGCAC
1851 GAACCCCCCG TTCAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT
1901 TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG
1951 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG
2001 AAGTGGTGGC CTAAC TACCG CTACACTAGA AGGACAGTAT TTGGTATCTG
2051 CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT
2101 CCGGCAAACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG
2151 CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC
2201 TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG
2251 TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT AAATTAAAAA
2301 TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG
2351 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC
2401 GTTCATCCAT AGCTGCCTGA CTCCCCGTCG TG TAGATAAC TACGATACGG
2451 GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG
2501 CTCACCGGCT CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG

Fig. 8 (continued)

2551 AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA GTCTATTAAT
2601 TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA
2651 CGTTGTTGCC ATTGCTACAG GCATCGTGGT GTCACGCTCG TCGTTTG GTA
2701 TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC
2751 CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCGGTCCTC CGATCGTTGT
2801 CAGAAGTAAG TTGGCCGCAG TGTTATCACT CATGGTTATG GCAGCACTGC
2851 ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT
2901 GAGTACTCAA CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG
2951 CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT AGCAGAACTT
3001 TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG
3051 ATCTTACCGC TGTTGAGATC CAGTTCGATG TAACCCACTC GTGCACCCAA
3101 CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA
3151 CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT
3201 TGAATACTCA TACTCTTCCT TTTTCAATAT TATTGAAGCA TTTATCAGGG
3251 TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC
3301 AAATAGGGGT TCCGCGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA
3351 GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC GTATCACGAG
3401 GCCCTTTCGT CTTCAC

Fig. 9

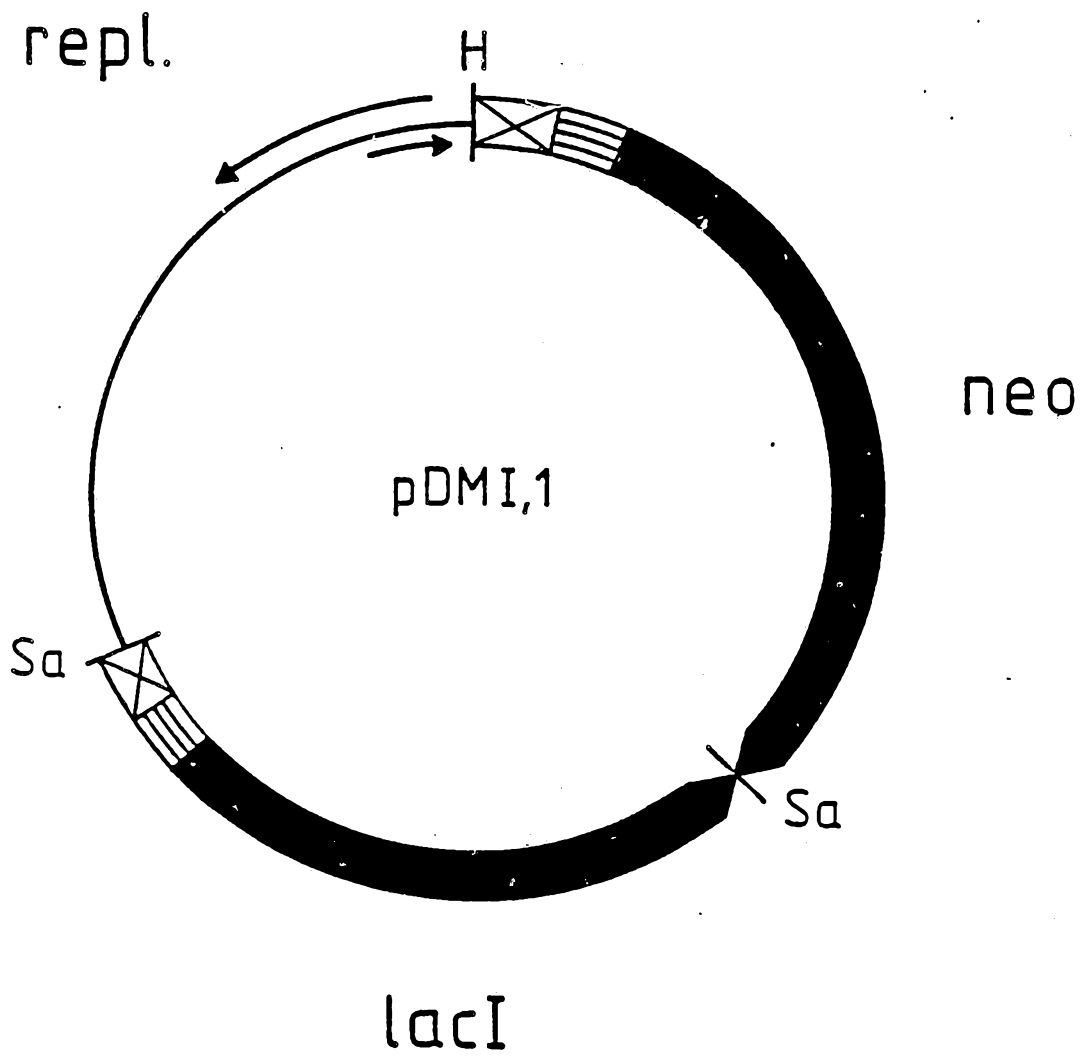


Fig. 10

	10	20	30	40	50
0	<u>AAGCTTCACG</u>	CTGCCGCAAG	CACTCAGGGC	GCAAGGGCTG	CTAAAGGAAG
50	CGGAACACGT	AGAAAGCCAG	TCCGCAGAAA	CGGTGCTGAC	CCCGGATGAA
100	TGTCAGCTAC	TGGGCTATCT	GGACAAGGGA	AAACGCAAGC	GCAAAGAGAA
150	AGCAGGTAGC	TTGCAGTGGG	CTTACATGGC	GATAGCTAGA	CTGGGCGGTT
200	TTATGGACAG	CAAGCGAACC	GGAATTGCCA	GCTGGGGCGC	CCTCTGGTAA
250	GGTTGGGAAG	CCCTGCAAAG	TAAACTGGAT	GGCTTTCTTG	CCGCCAAGGA
300	TCTGATGGCG	CAGGGGATCA	AGATCTGATC	AAGAGACAGG	ATGAGGATCG
350	<u>TTTCGCATGA</u>	<u>TTGAACAAGA</u>	<u>TGGATTGCAC</u>	<u>GCAGGTTCTC</u>	<u>CGGCCGCTTG</u>
400	<u>GGTGGAGAGG</u>	<u>CTATTCGGCT</u>	<u>ATGACTGGGC</u>	<u>ACAACAGACA</u>	<u>ATCGGCTGCT</u>
450	<u>CTGATGCCGC</u>	<u>CGTGTTCCGG</u>	<u>CTGTCAGCGC</u>	<u>AGGGGCGCCC</u>	<u>GGTTCTTTTT</u>
500	<u>GTC AAGACCG</u>	<u>ACCTGTCCGG</u>	<u>TGCCCTGAAT</u>	<u>GAAGTGCAGG</u>	<u>ACGAGGCAGC</u>
550	<u>CGGGCTATCG</u>	<u>TGGCTGGCCA</u>	<u>CGACGGGCGT</u>	<u>TCCTTGCGCA</u>	<u>GCTGTGCTCG</u>
600	<u>ACGTTGTCAC</u>	<u>TGAAGCGGGA</u>	<u>AGGGACTGGC</u>	<u>TGCTATTGGG</u>	<u>CGAAGTGCCG</u>
650	<u>GGGCAGGATC</u>	<u>TCCTGTCATC</u>	<u>TCACCTTGCT</u>	<u>CCTGCCGAGA</u>	<u>AAGTATCCAT</u>
700	<u>CATGGCTGAT</u>	<u>GCAATGCGGC</u>	<u>GGCTGCATAC</u>	<u>GCTTGATCCG</u>	<u>GCTACCTGCC</u>
750	<u>CATTCGACCA</u>	<u>CCAAGCGAAA</u>	<u>CATCGCATCG</u>	<u>AGCGAGCACG</u>	<u>TACTCGGATG</u>
800	<u>GAAGCCGGTC</u>	<u>TTGTGATCA</u>	<u>GGATGATCTG</u>	<u>GACGAAGAGC</u>	<u>ATCAGGGGCT</u>
850	<u>CGCGCCAGCC</u>	<u>GAAGTGTTCG</u>	<u>CCAGGCTCAA</u>	<u>GGCGCGCATG</u>	<u>CCCGACGGCG</u>
900	<u>AGGATCTCGT</u>	<u>CGTGACCCAT</u>	<u>GGCGATGCCT</u>	<u>GCTTGCCGAA</u>	<u>TATCATGGTG</u>
950	<u>GAAAATGGCC</u>	<u>GCTTTTCTGG</u>	<u>ATTCATCGAC</u>	<u>TGTGGCCGGC</u>	<u>TGGGTGTGGC</u>
1000	<u>GGACCGCTAT</u>	<u>CAGGACATAG</u>	<u>CGTTGGCTAC</u>	<u>CCGTGATATT</u>	<u>GCTGAAGAGC</u>
1050	<u>TTGGCGGCCA</u>	<u>ATGGGCTGAC</u>	<u>CGCTTCCTCG</u>	<u>TGCTTTACGG</u>	<u>TATCGCCGCT</u>
1100	<u>CCCGATTCCG</u>	<u>AGCGCATCGC</u>	<u>CTTCTATCGC</u>	<u>CTTCTTGACG</u>	<u>AGTTCTTCTG</u>
1150	<u>AGCGGGACTC</u>	<u>TGGGGTTCGA</u>	<u>AATGACCGAC</u>	<u>CAAGCGACGC</u>	<u>CCAACCTGCC</u>
1200	<u>ATCACGAGAT</u>	<u>TTCGATTCCA</u>	<u>CCGCCGCCTT</u>	<u>CTATGAAAGG</u>	<u>TTGGGCTTCC</u>
1250	<u>GAATCGTTTT</u>	<u>CCGGGACGCC</u>	<u>GGCTGGATGA</u>	<u>TCCTCCAGCG</u>	<u>CGGGGATCTC</u>
1300	<u>ATGCTGGAGT</u>	<u>TCTTCGCCCA</u>	<u>CCCCGGGCTC</u>	<u>GATCCCCCTC</u>	<u>CGAGTTGGTT</u>
1350	<u>CAGCTGCTGC</u>	<u>CTGAGGCTGG</u>	<u>ACGACCTCGC</u>	<u>GGAGTTCTAC</u>	<u>CGGCAGTGCA</u>
1400	<u>AATCCGTCCG</u>	<u>CATCCAGGAA</u>	<u>ACCAGCAGCG</u>	<u>GCTATCCCGC</u>	<u>CATCCATGCC</u>
1450	<u>CCCGAACTGC</u>	<u>AGGAGTGGGG</u>	<u>AGGCACGATG</u>	<u>GCCGCTTTGG</u>	<u>TCGACAATTG</u>
1500	<u>GCGCTAACTT</u>	<u>ACATTAATTG</u>	<u>CGTTGCGCTC</u>	<u>ACTGCCCCGT</u>	<u>TTCCAGTCCG</u>
1550	<u>GAAACCTGTC</u>	<u>GTGCCAGCTG</u>	<u>CATTAATGAA</u>	<u>TGGGCCAACG</u>	<u>CGCGGGGAGA</u>
1600	<u>GGCGGTTTGC</u>	<u>GTATTGGGCG</u>	<u>CCAGGGTGGT</u>	<u>TTTTCTTTTC</u>	<u>ACCAGTGAGA</u>
1650	<u>CGGGCAACAG</u>	<u>CTGATTGCCC</u>	<u>TTACCCGCCT</u>	<u>GGCCCTGAGA</u>	<u>GAGTTGCAGC</u>
1700	<u>AAGCGGTCCA</u>	<u>CGCTGGTTTG</u>	<u>CCCCAGCAGG</u>	<u>CGAAAATCCT</u>	<u>GTTTGATGGT</u>
1750	<u>GGTTAACGGC</u>	<u>GGGATATAAC</u>	<u>ATGAGCTGTC</u>	<u>TTGGGTATCG</u>	<u>TCGTATCCCA</u>
1800	<u>CTACCGAGAT</u>	<u>ATCCGCACCA</u>	<u>ACGCGCAGCC</u>	<u>GGGACTCGGT</u>	<u>AATGGCGCGC</u>
1850	<u>ATTGCGCCCA</u>	<u>GCGCCATCTG</u>	<u>ATCGTTGGCA</u>	<u>ACCAGCATCG</u>	<u>CAGTGGGAAC</u>

Fig. 10 (continued)

	10	20	30	40	50
1900	<u>GATGCCCTCA</u>	<u>TTCAGCATTT</u>	<u>GCATGGTTTG</u>	<u>TTGAAAACCG</u>	<u>GACATGGCAC</u>
1950	<u>TCCAGTCGCC</u>	<u>TTCCCGTTCC</u>	<u>GCTATCGGCT</u>	<u>GAATTTGATT</u>	<u>GCGAGTGAGA</u>
2000	<u>TATTTATGCC</u>	<u>AGCCAGCCAG</u>	<u>ACGCAGACGC</u>	<u>GCCGAGACAG</u>	<u>AACTTAATGG</u>
2050	<u>GCCCCGTAAC</u>	<u>AGCGCGATTT</u>	<u>GCTGGTGACC</u>	<u>CAATGCGACC</u>	<u>AGATGCTCCA</u>
2100	<u>CGCCCAGTCG</u>	<u>CGTACCGTCT</u>	<u>TCATGGGAGA</u>	<u>AAATAATACT</u>	<u>GTTGATGGGT</u>
2150	<u>GTCTGGTCAG</u>	<u>AGACATCAAG</u>	<u>AAATAACGCC</u>	<u>GGAACATTAG</u>	<u>TGCAGGCAGC</u>
2200	<u>TTCCACAGCA</u>	<u>ATGGCATCCT</u>	<u>GGTCATCCAG</u>	<u>CGGATAGTTA</u>	<u>ATGATCAGCC</u>
2250	<u>CACTGACCGG</u>	<u>TTGCCGAGAG</u>	<u>AGATTGTGCA</u>	<u>CCGCCGCTTT</u>	<u>ACAGGCTTCG</u>
2300	<u>ACGCCGCTTC</u>	<u>GTTCTACCAT</u>	<u>CGACACCACC</u>	<u>ACGCTGGCAC</u>	<u>CCAGTTGATC</u>
2350	<u>GGCGCGAGAT</u>	<u>TTAATCGCCG</u>	<u>CGACAATTTG</u>	<u>CGACGGCGCG</u>	<u>TGCAGGGCCA</u>
2400	<u>GACTGGAGGT</u>	<u>GGCAACGCCA</u>	<u>ATCAGCAACG</u>	<u>ACTGTTTGCC</u>	<u>CGCCAGTTGT</u>
2450	<u>TGTGCCACGC</u>	<u>GGTTGGGAAT</u>	<u>GTAATTCAGC</u>	<u>TCCGCCATCG</u>	<u>CCGCTTCCAC</u>
2500	<u>TTTTTCCCGC</u>	<u>GTTTTCGCAG</u>	<u>AAACGTGGCT</u>	<u>GGCCTGGTTC</u>	<u>ACCACGCGGG</u>
2550	<u>AAACGGTCTG</u>	<u>ATAAGAGACA</u>	<u>CCGGCATACT</u>	<u>CTGCGACATC</u>	<u>GTATAACGTT</u>
2600	<u>ACTGGTTTCA</u>	<u>CATTCACCAC</u>	<u>CCTGAATTGA</u>	<u>CTCTCTTCCG</u>	<u>GGCGCTATCA</u>
2650	<u>TGCCATACCG</u>	<u>CGAAAGGTTT</u>	<u>TGCACCATTG</u>	<u>GATGGTGTCA</u>	<u>ACGTAAATGC</u>
2700	<u>ATGCCGCTTC</u>	<u>GCCTTCGCGC</u>	<u>GCGAATTGTC</u>	<u>GACCTGTGCC</u>	<u>CTCCTGTTCA</u>
2750	<u>GCTACTGACG</u>	<u>GGGTGGTGCG</u>	<u>TAACGGCAAA</u>	<u>AGCACCGCCG</u>	<u>GACATCAGCG</u>
2800	<u>CTAGCGGAGT</u>	<u>GTATACTGGC</u>	<u>TTACTATGTT</u>	<u>GGCACTGATG</u>	<u>AGGGTGTCAG</u>
2850	<u>TGAAGTGCTT</u>	<u>CATGTGGCAG</u>	<u>GAGAAAAAAG</u>	<u>GCTGCACCGG</u>	<u>TGCGTCAGCA</u>
2900	<u>GAATATGTGA</u>	<u>TACAGGATAT</u>	<u>ATTCCGCTTC</u>	<u>CTCGCTCACT</u>	<u>GACTCGCTAC</u>
2950	<u>GCTCGGTCTG</u>	<u>TCGACTGCGG</u>	<u>CGAGCGGAAA</u>	<u>TGGCTTACGA</u>	<u>ACGGGGCGGA</u>
3000	<u>GATTTCTCTG</u>	<u>AAGATGCCAG</u>	<u>GAAGATACTT</u>	<u>AACAGGGAAG</u>	<u>TGAGAGGGCC</u>
3050	<u>GCGGCAAAAG</u>	<u>CGTTTTTCCA</u>	<u>TAGGCTCCGC</u>	<u>CCCCCTGACA</u>	<u>AGCATCACGA</u>
3100	<u>AATCTGACGC</u>	<u>TCAAATCAGT</u>	<u>GGTGGCGAAA</u>	<u>CCCGACAGGA</u>	<u>CTATAAAGAT</u>
3150	<u>ACCAGGCGTT</u>	<u>TCCCCTGGCG</u>	<u>GCTCCCTCGT</u>	<u>GCGCTCTCCT</u>	<u>GTTCTTGCCT</u>
3200	<u>TTCCGTTTAC</u>	<u>CGGTGTCAAT</u>	<u>CCGCTGTTAT</u>	<u>GGCCGCGTTT</u>	<u>GTCTCATTCG</u>
3250	<u>ACGCCTGACA</u>	<u>CTCAGTTCCG</u>	<u>GGTAGGCAGT</u>	<u>TCGCTCCAAG</u>	<u>CTGGACTGTA</u>
3300	<u>TGCACGAACC</u>	<u>CCCCGTTTCA</u>	<u>TCCGACCGCT</u>	<u>GCGCCTTATC</u>	<u>CGGTAACATAT</u>
3350	<u>CGTCTTGAGT</u>	<u>CCAACCCGGA</u>	<u>AAGACATGCA</u>	<u>AAAGCACCAC</u>	<u>TGGCAGCAGC</u>
3400	<u>CACTGGTAAT</u>	<u>TGATTTAGAG</u>	<u>GAGTTAGTCT</u>	<u>TGAAGTCATG</u>	<u>CGCCGGTTAA</u>
3450	<u>GGCTAAACTG</u>	<u>AAAGGACAAG</u>	<u>TTTTGGTGAC</u>	<u>TGCGCTCCTC</u>	<u>CAAGCCAGTT</u>
3500	<u>ACCTCGGTTC</u>	<u>AAAGAGTTGG</u>	<u>TAGCTCAGAG</u>	<u>AACCTTCGAA</u>	<u>AAACCGCCCT</u>
3550	<u>GCAAGGCGGT</u>	<u>TTTTTCGTTT</u>	<u>TCAGAGCAAG</u>	<u>AGATTACGCG</u>	<u>CAGACCAAAA</u>
3600	<u>CCATCTCAAG</u>	<u>AAGATCATCT</u>	<u>TATTAATCAG</u>	<u>ATAAAATATT</u>	<u>TCTAGATTTT</u>
3650	<u>AGTGCAATTT</u>	<u>ATCTCTTCAA</u>	<u>ATGTAGCACC</u>	<u>TGAAGTCAGC</u>	<u>CCCATACGAT</u>
3700	<u>ATAAGTTGTT</u>	<u>AATTCTCATG</u>	<u>TTTGACAGCT</u>	<u>TATCATCGAT</u>	

Fig. 11

Adapter 1: (SphI) (NdeI)
 CAGGACCCA
 GTACGTCCTGGGTAT

Adapter 2: (HinfI) (HindIII)
 AGTCAGATGCTGTAGTTAACA
 GTCTACGACATCAATTGTTCTGA

Adapter 3: (SphI) NaeI NarI (BamHI)
 CATCACGCCGGCATCGAAGGGCGCCTTG
 GTACGTAGTGCGGCCGTAGCTTCCCGCGGAACCTAG

Adapter 4: (NarI) (NdeI)
 CGCCAAGATCCA
 GGTTCTAGGTAT

Adapter 5: (NaeI) (NdeI)
 GGCGATGACGATGACAAACAAGATCCA
 CCGCTACTGCTACTGTTTGTTCTAGGTAT

Adapter 6: (NarI) (NdeI)
 CGCCAAAACCCA
 GGTTTTGGGTAT

Fig. 11 (continued)

(BamHI)

Adapter 7: GATCGCATCACCATCACCATCACG
 CGTAGTGGTAGTGGTAGTGCCTAG

(BamHI)

Adapter 8: GATCGCATCACCATCACCATCACG
 CGTAGTGGTAGTGCCTAG

(BamHI) BglII

(HindIII)

Adapter 9: GATCCAGATCTCATCACCATCACCATCACTA
 GTCTAGAGTAGTGGTAGTGGTAGTGATTCGA

(BamHI) BglII

(HindIII)

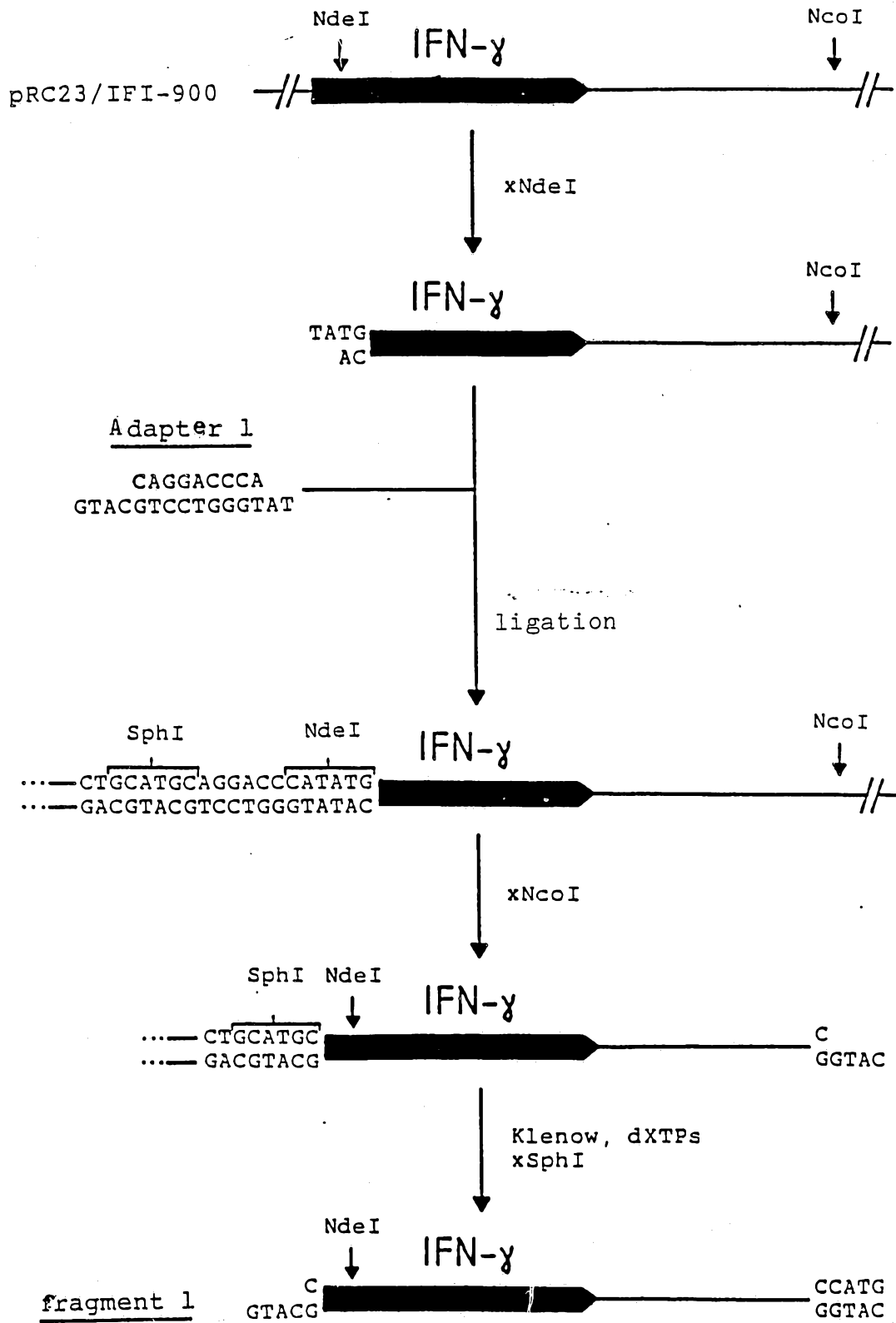
Adapter 10: GATCCAGATCTCATCACCATCACTA
 GTCTAGAGTAGTGGTAGTGATTCGA

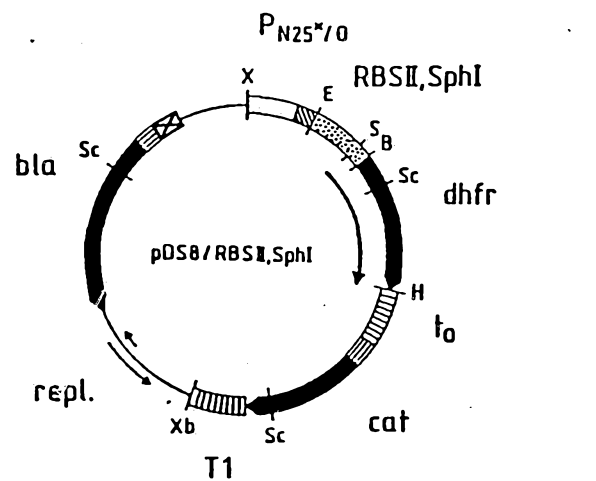
(BamHI) BglII

(HindIII)

Adapter 11: GATCCAGATCTCATCACTA
 GTCTAGAGTAGTGATTCGA

Fig. 12





xScaI, partial
xSphI
fragment isolation

fragment 1



ligation
transformation

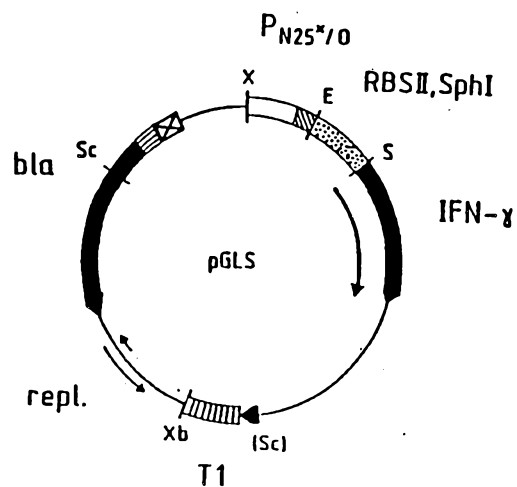


Fig. 13

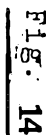


Fig. 15

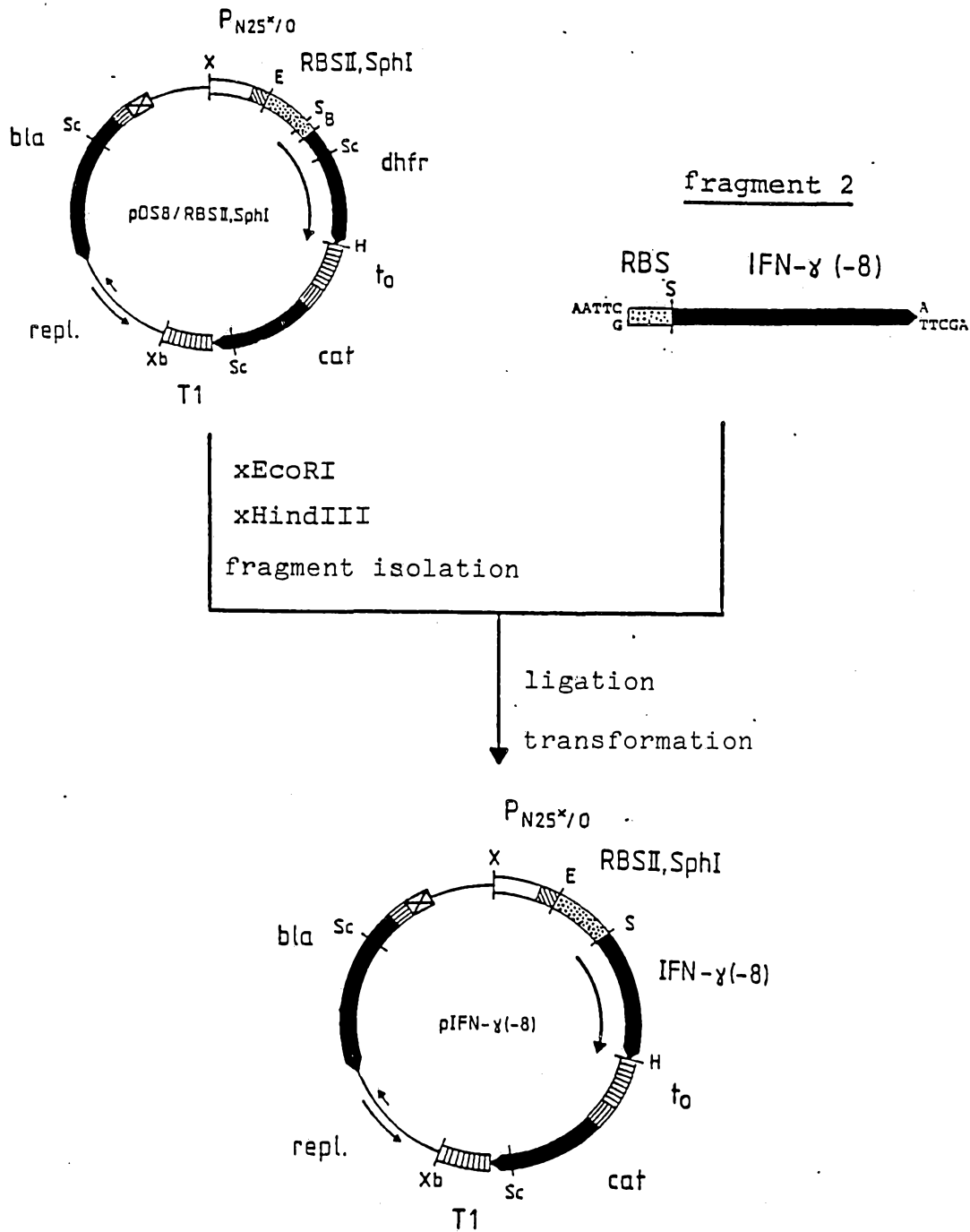
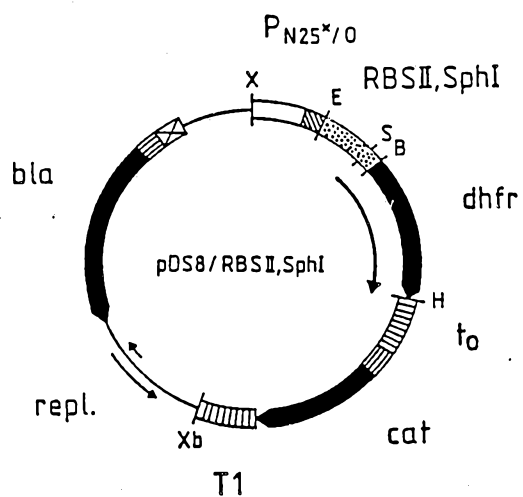


Fig. 16

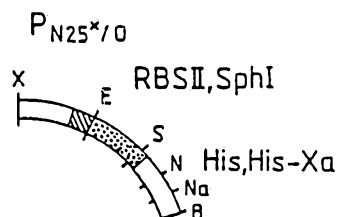


adapter 3

(SphI) NaeI NarI (BamHI)
 CATCACGCCGGCATCGAAGGGCGCCTTG
 GTACGTAGTGCGGCCGTAGCTTCCCGCGGAACCTAG

xSphI

ligation
 xBamHI
 xXhoI
 fragment isolation



fragment 3

Fig. 17

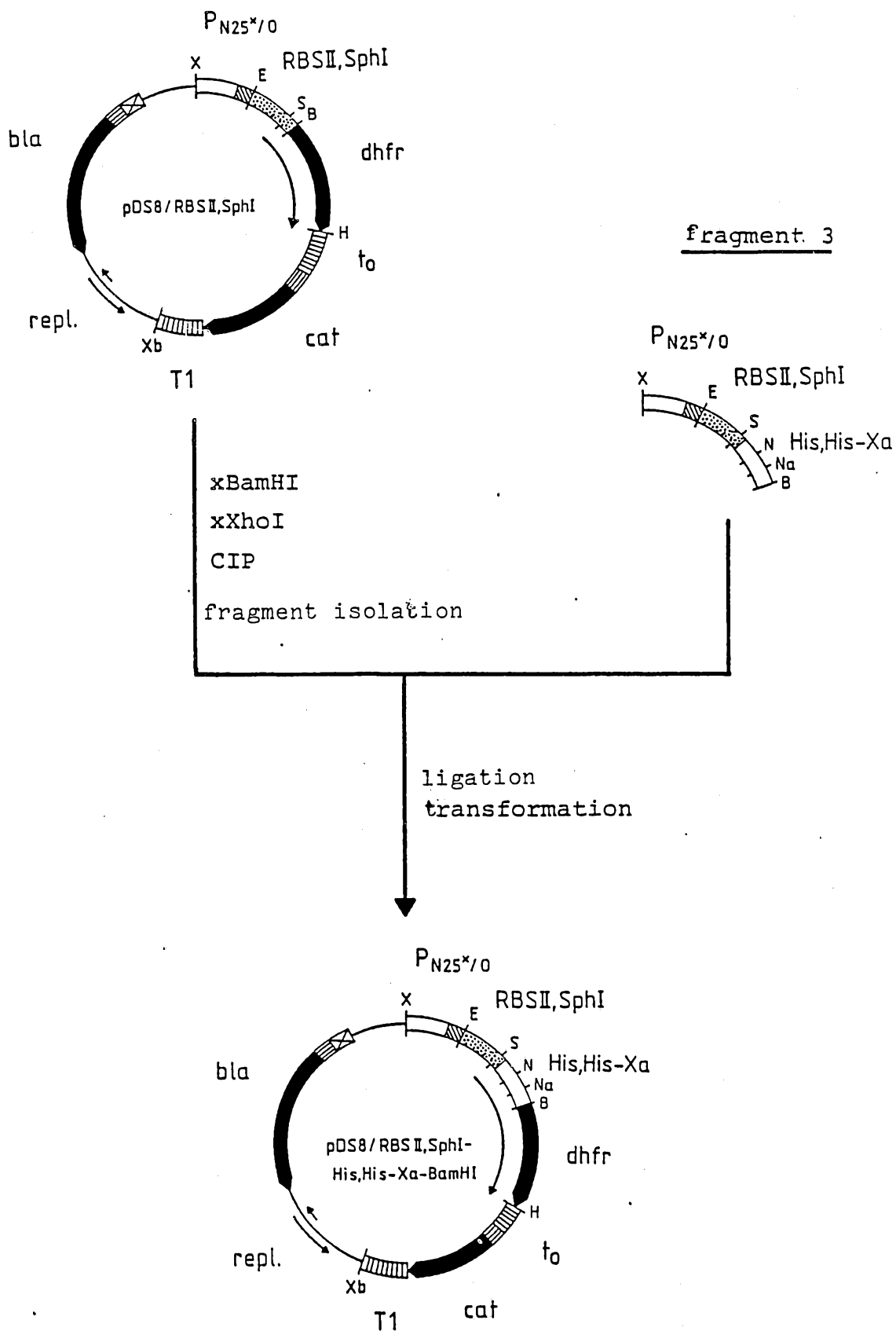


Fig. 18

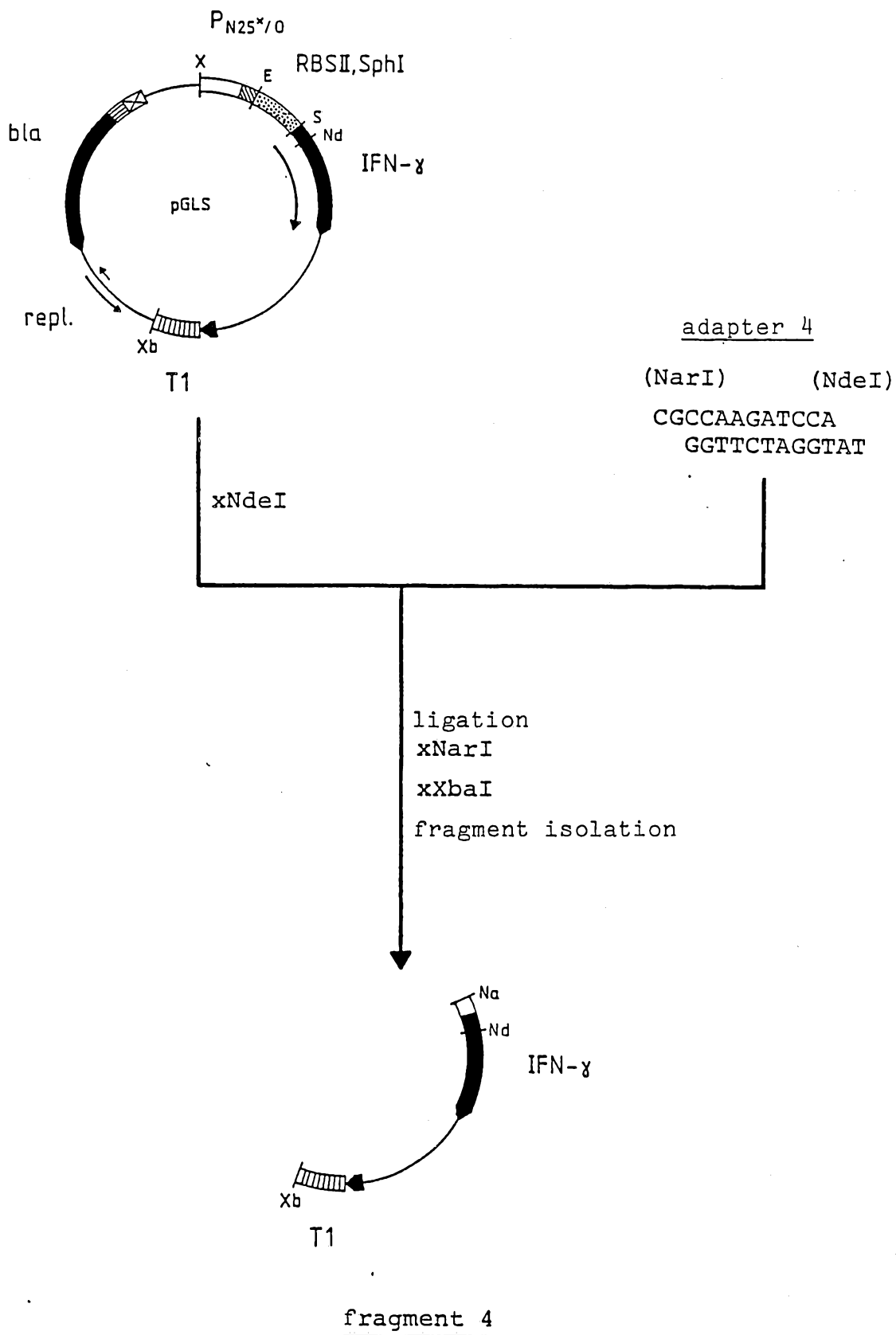


Fig. 19

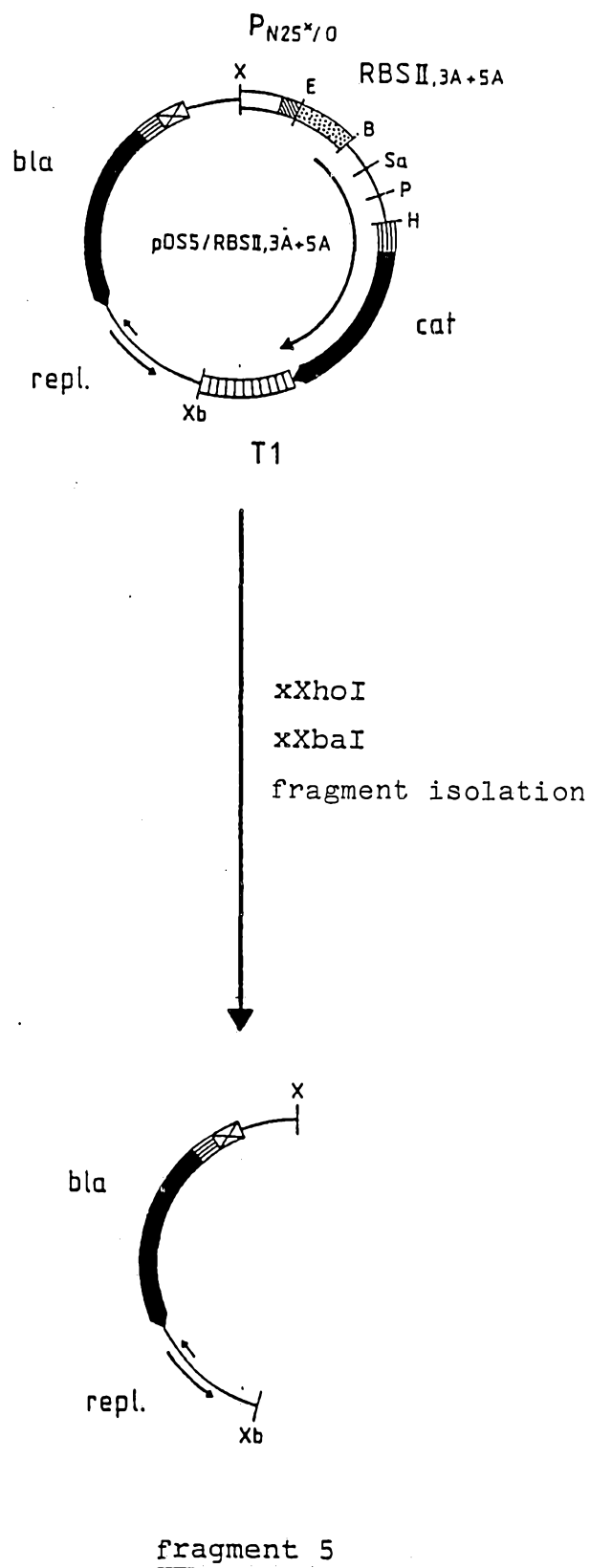
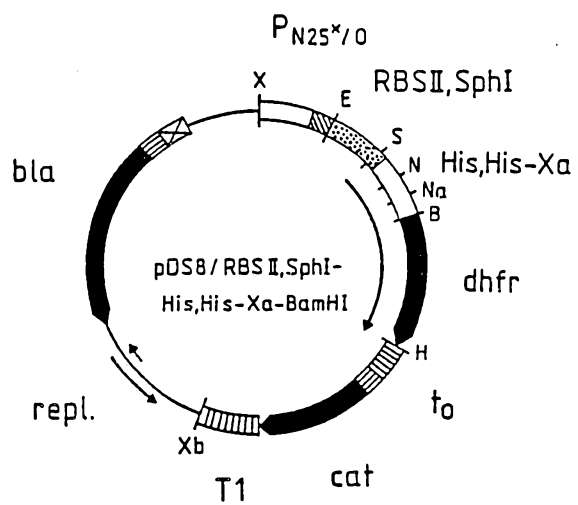
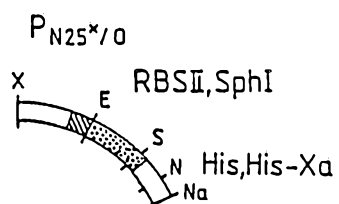


Fig. 20



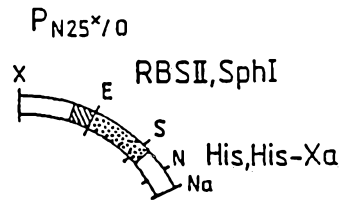
xxhoI
 xNarI
 fragment isolation



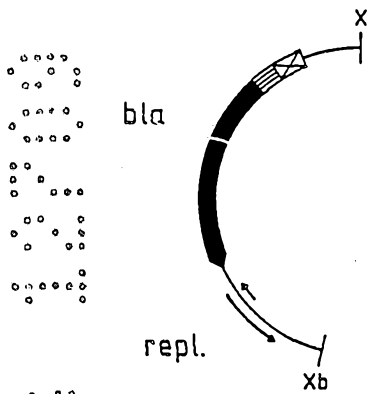
fragment 6

Fig. 21

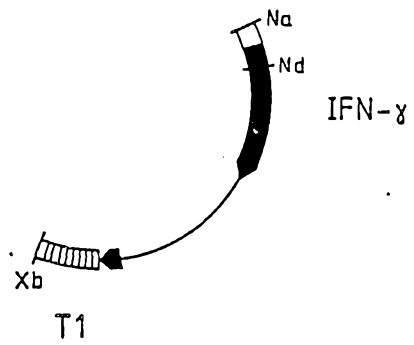
fragment 6



fragment 5



fragment 4



ligation
transformation

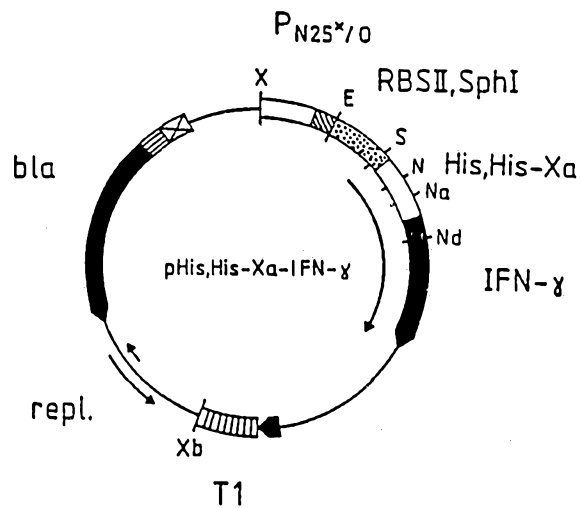
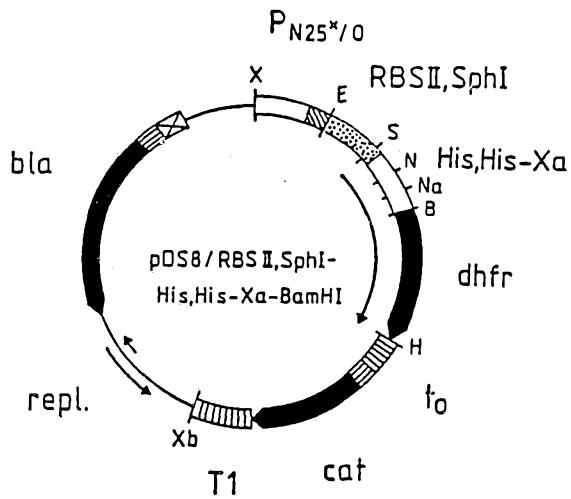


Fig. 22



(NaeI) adapter 5 (NdeI)

GGCGATGACGATGACAAACAAGATCCA

CCGCTACTGCTACTGTTTGTCTAGGTAT

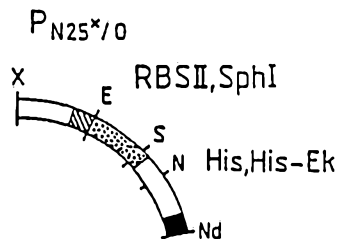
xNaeI

ligation

xxhoI

xNdeI

fragment isolation



fragment 7

Fig. 23

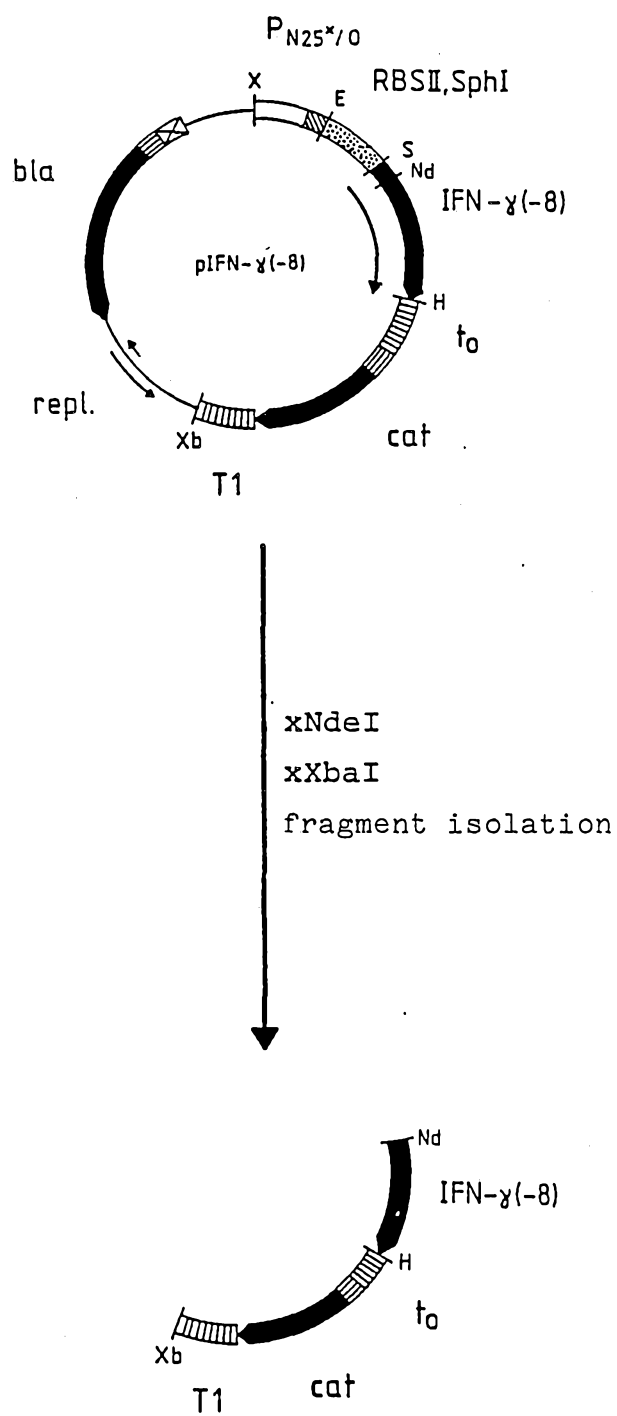


Fig. 24

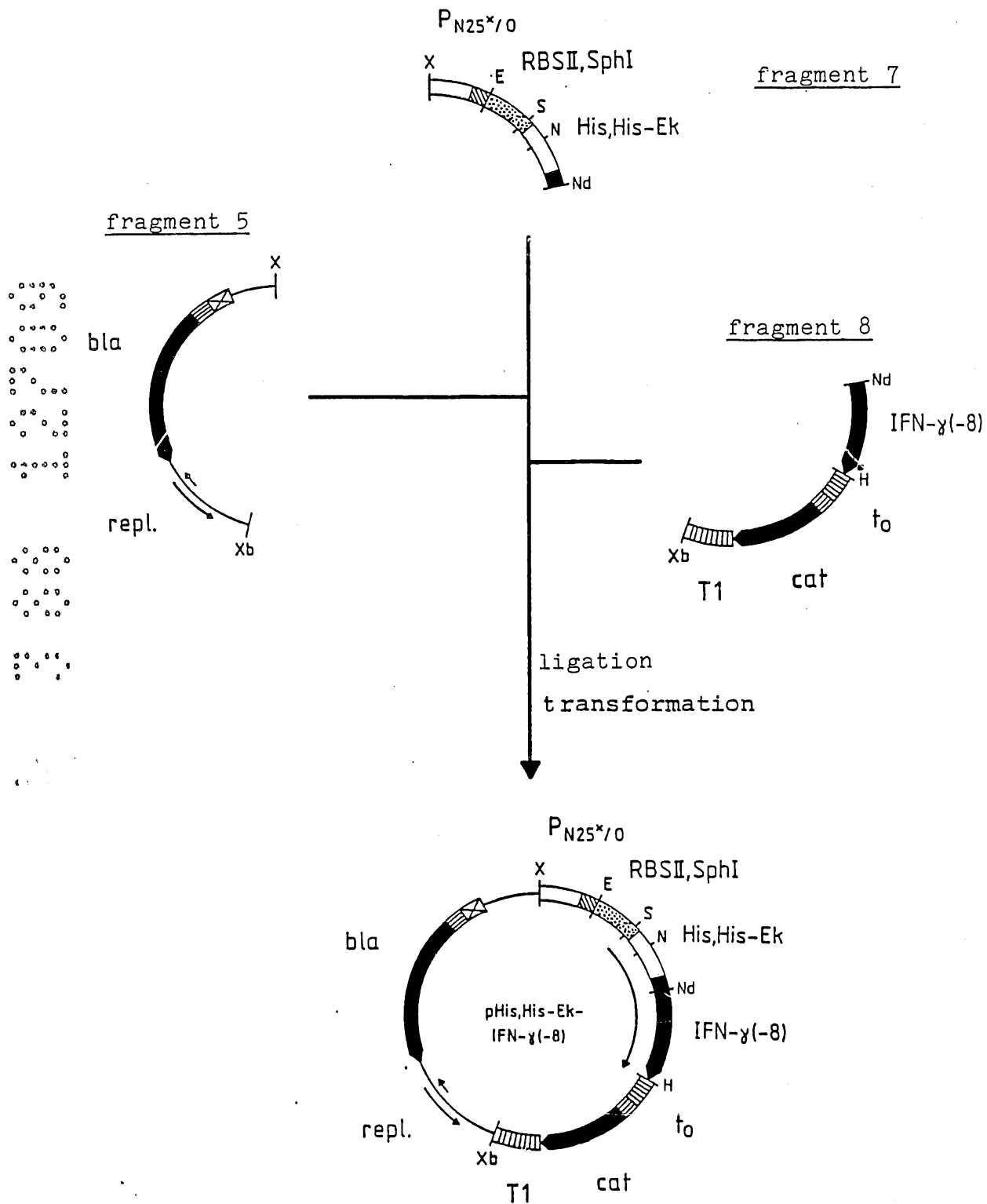
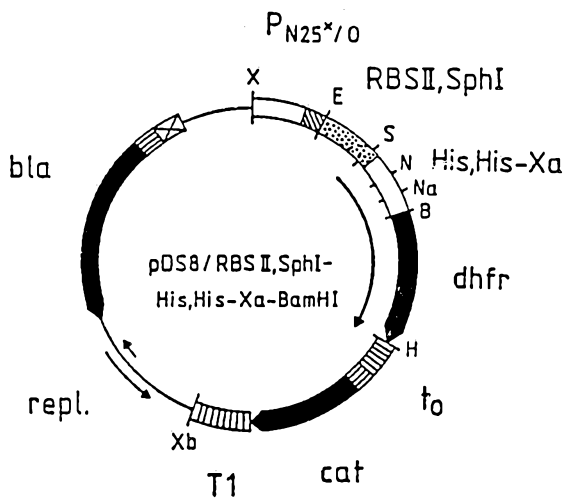


Fig. 25



adapter 6
 (NarI) (NdeI)
 CGCCAAAACCCA
 GGTTTTGGGTAT

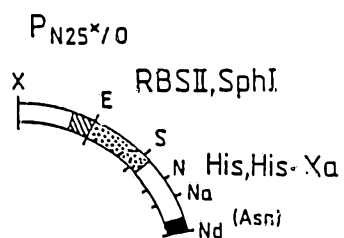
xNarI

ligation

transformation

xXhoI

xNdeI



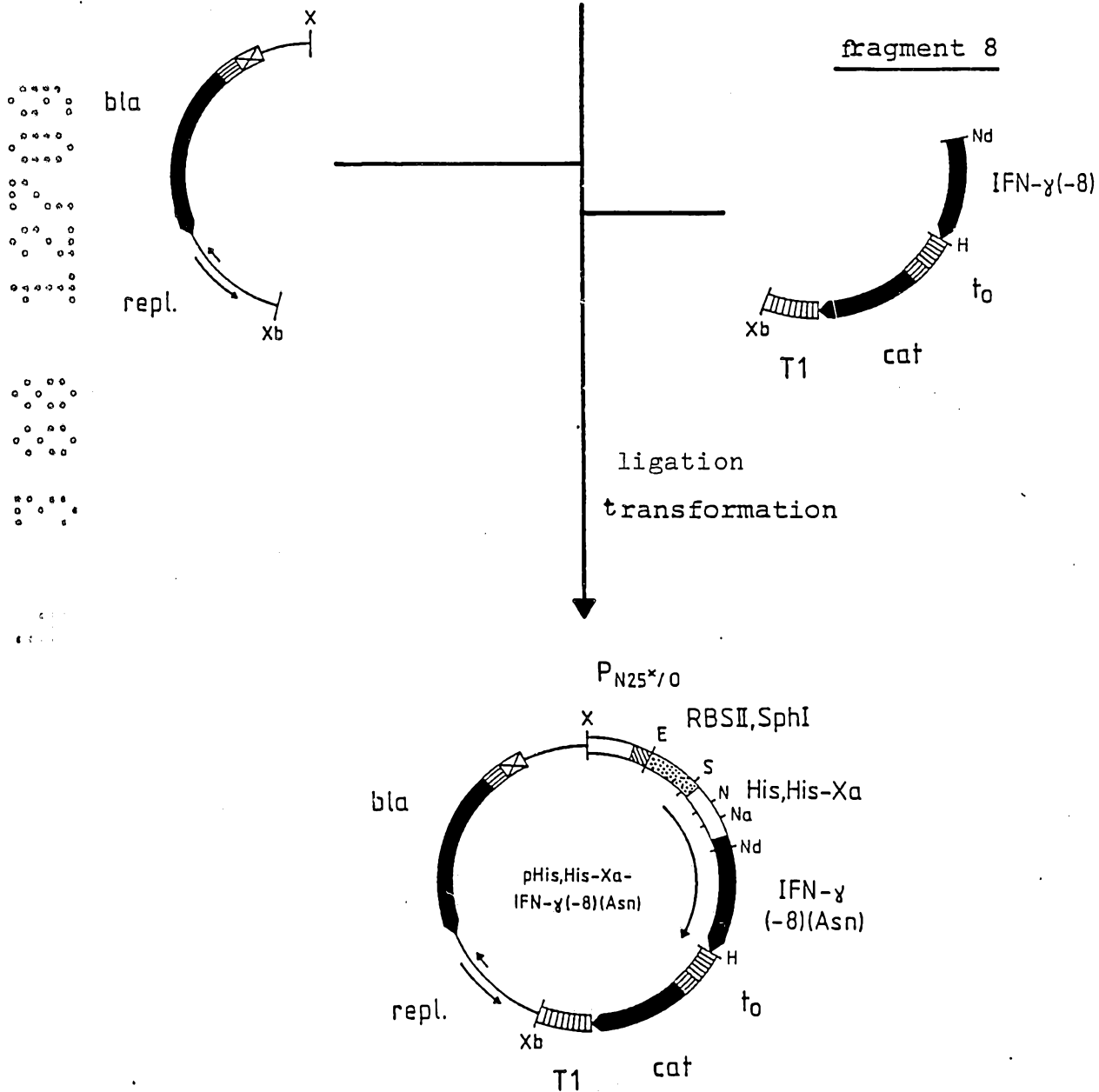
fragment 9

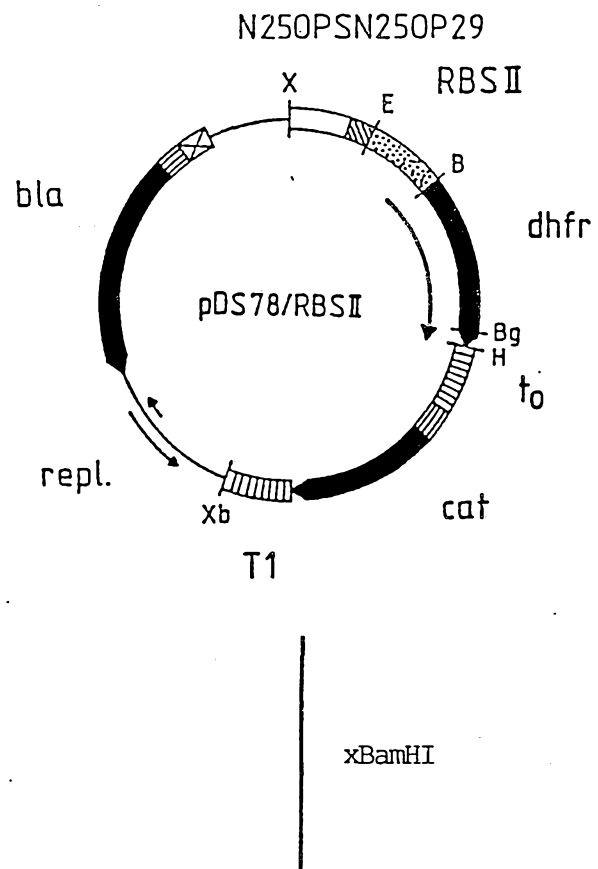
Fig. 26

fragment 9

fragment 5

fragment 8





xBamHI

fragment 10

adapter 7

(BamHI)

GATCGCATCACCATCACCATCACG
CGTAGTGGTAGTGGTAGTGCCTAG

ligation
xXhoI, xBamHI
fragment isolation.

N250PSN250P29 RBSII, 6xHis

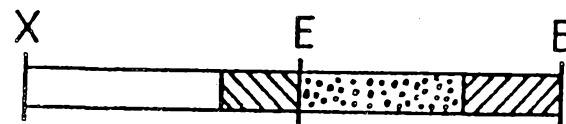
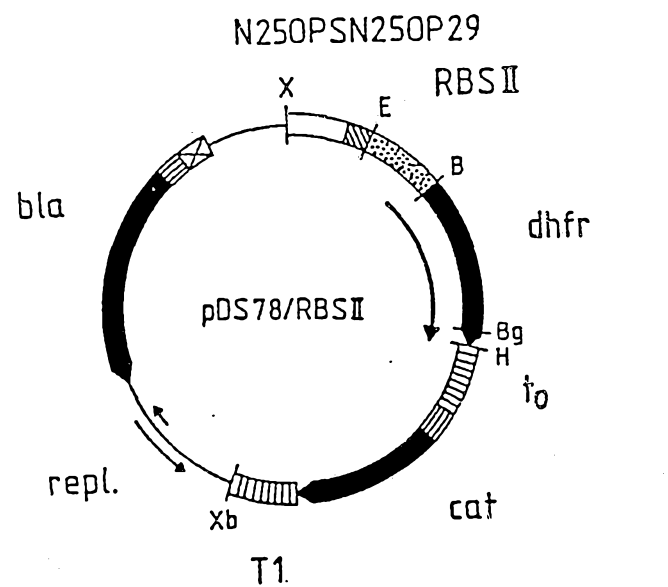


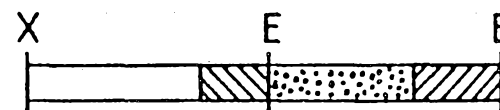
Fig. 27



xXhoI, xBamHI
fragment isolation

fragment 10

N250PSN250P29 RBSII,6xHis



ligation
transformation

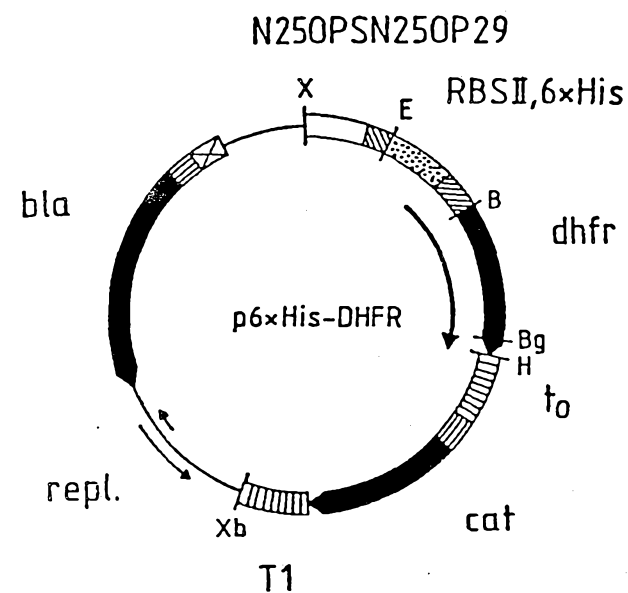
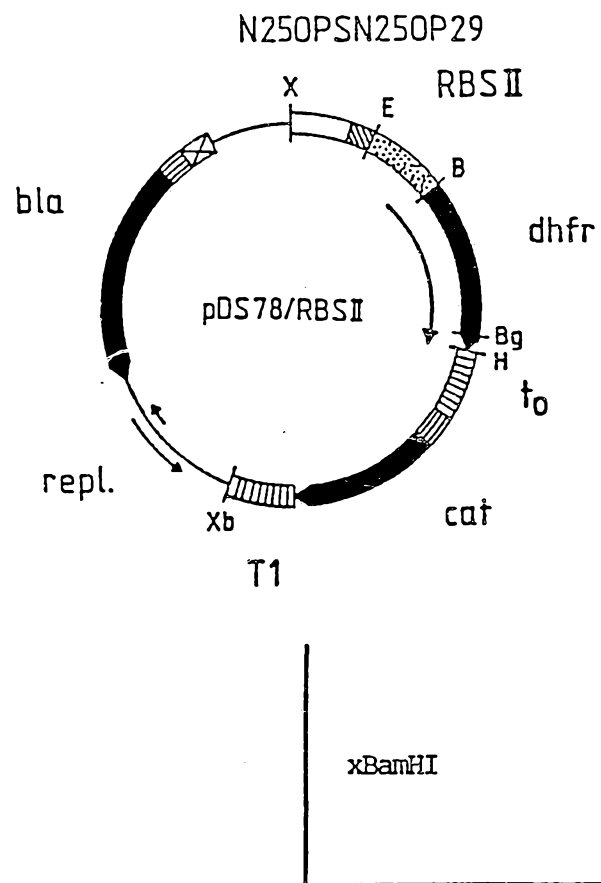


Fig. 28



xBamHI

fragment 11

adapter 8

(BamHI)

GATCGCATCACCATCAG
CGTAGTGGTAGTGCCTAG

ligation
xXhoI, xBamHI
fragment isolation

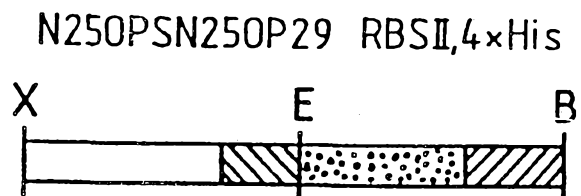
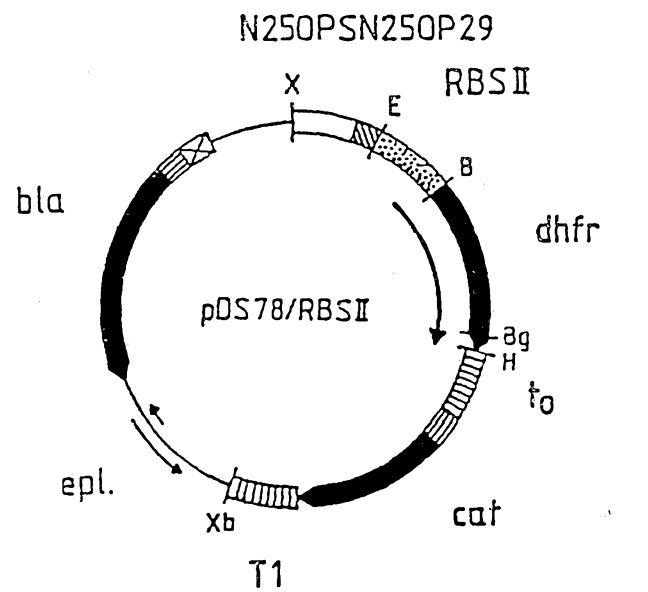


Fig. 29



xXhoI, xBamHI
fragment isolation

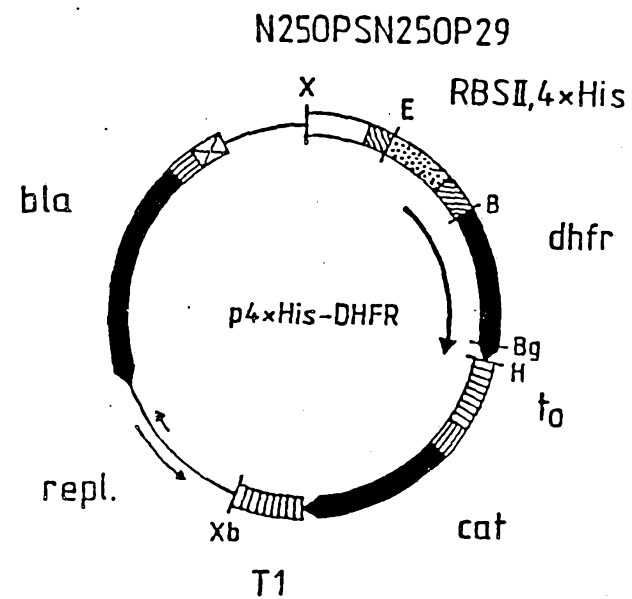
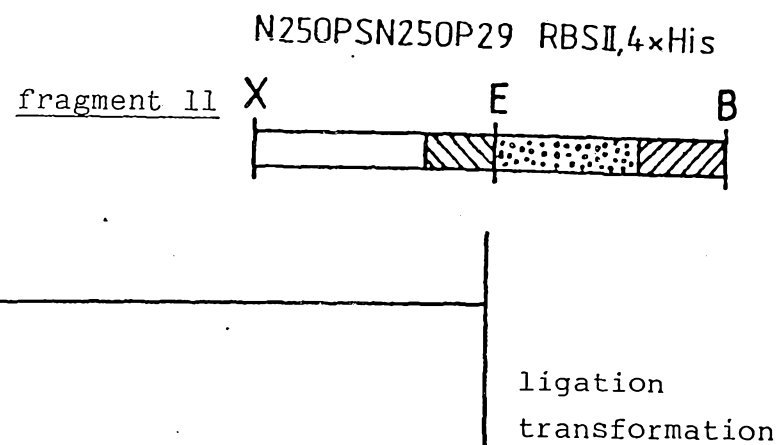
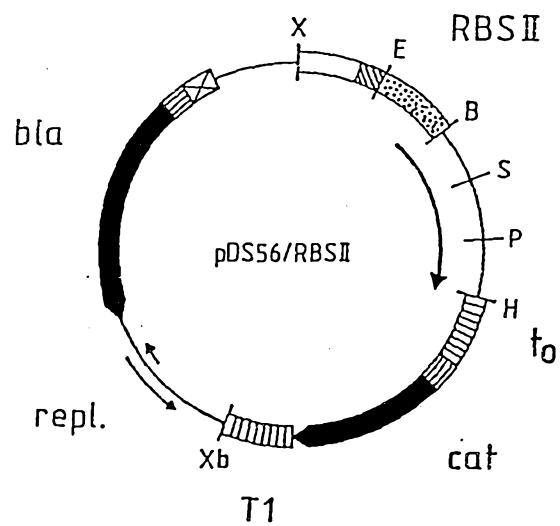


Fig. 30

N250PSN250P29



adapter 9

(BamHI) BglII

(HindIII)

GATCCAGATCTCATCACCATCACCATCACTA
GTCTAGAGTAGTGGTAGTGGTAGTGATTCTGA

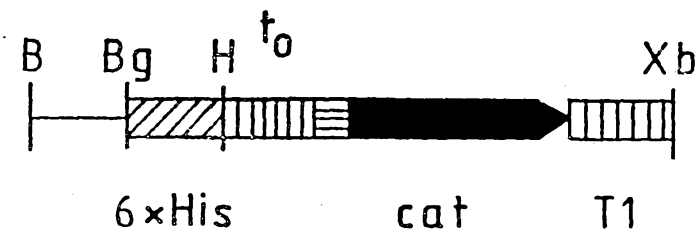
ligation

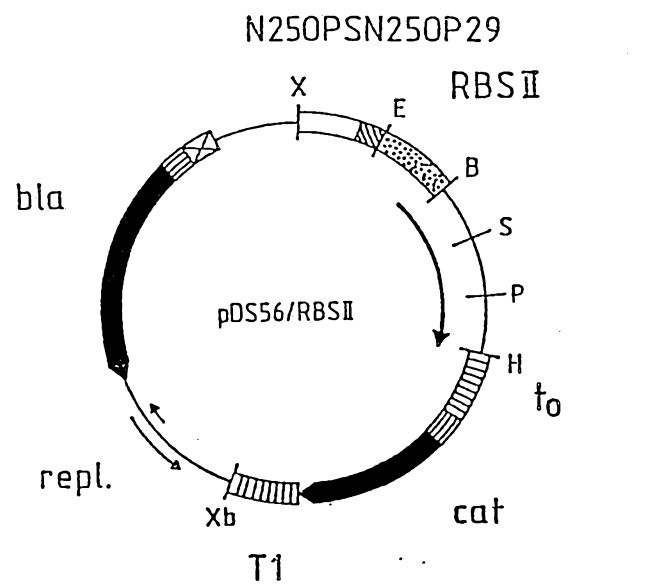
xBamHI, xXbaI

fragment isolation

xHindIII

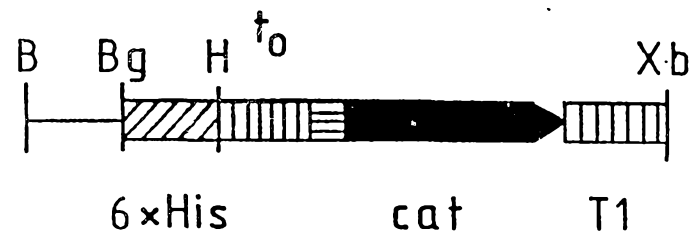
fragment 12

Fig. 31



xBamHI, xXbaI
fragment isolation

fragment 12



ligation
transformation

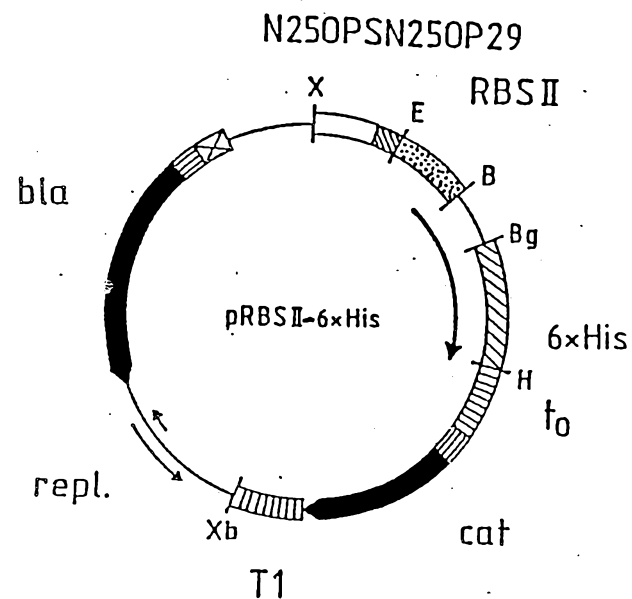
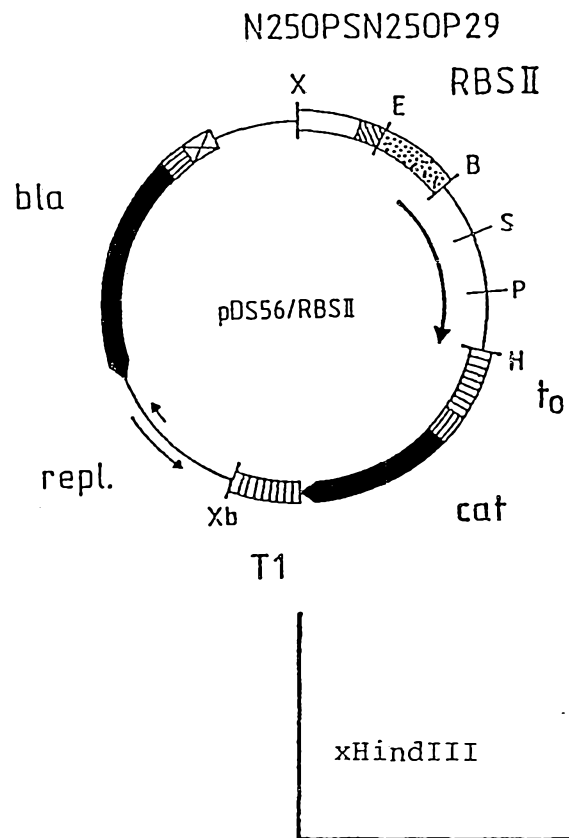


Fig. 32



adapter 10

(BamHI) BglIII

(HindIII)

GATCCAGATCTCATCACCATCACTA
GTCTAGAGTAGTGGTAGTGATTCTGA

ligation

xBamHI, xXbaI

fragment isolation

fragment 13

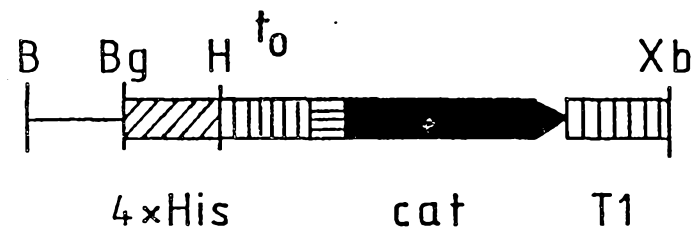


Fig. 33

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

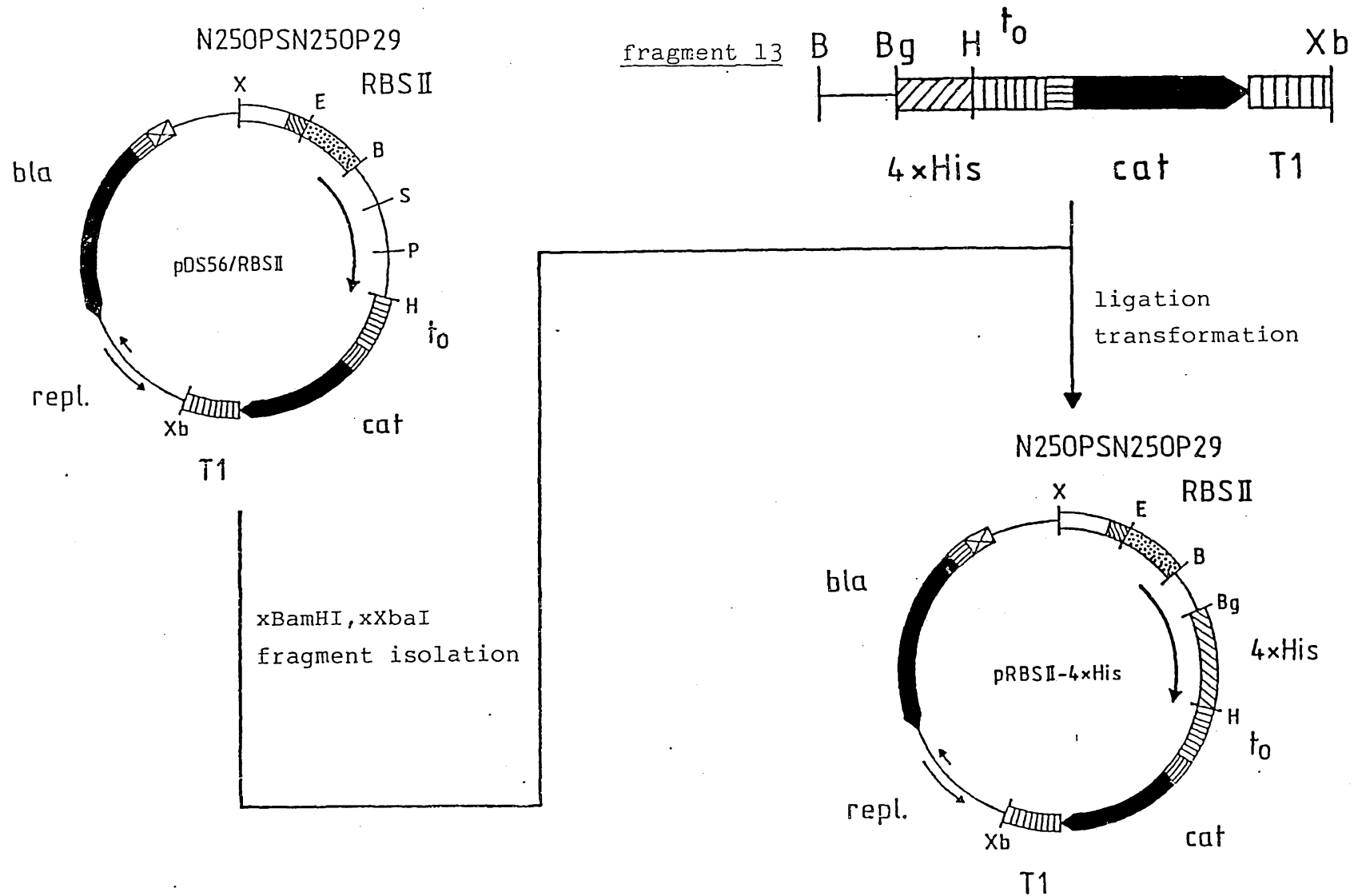


Fig. 34

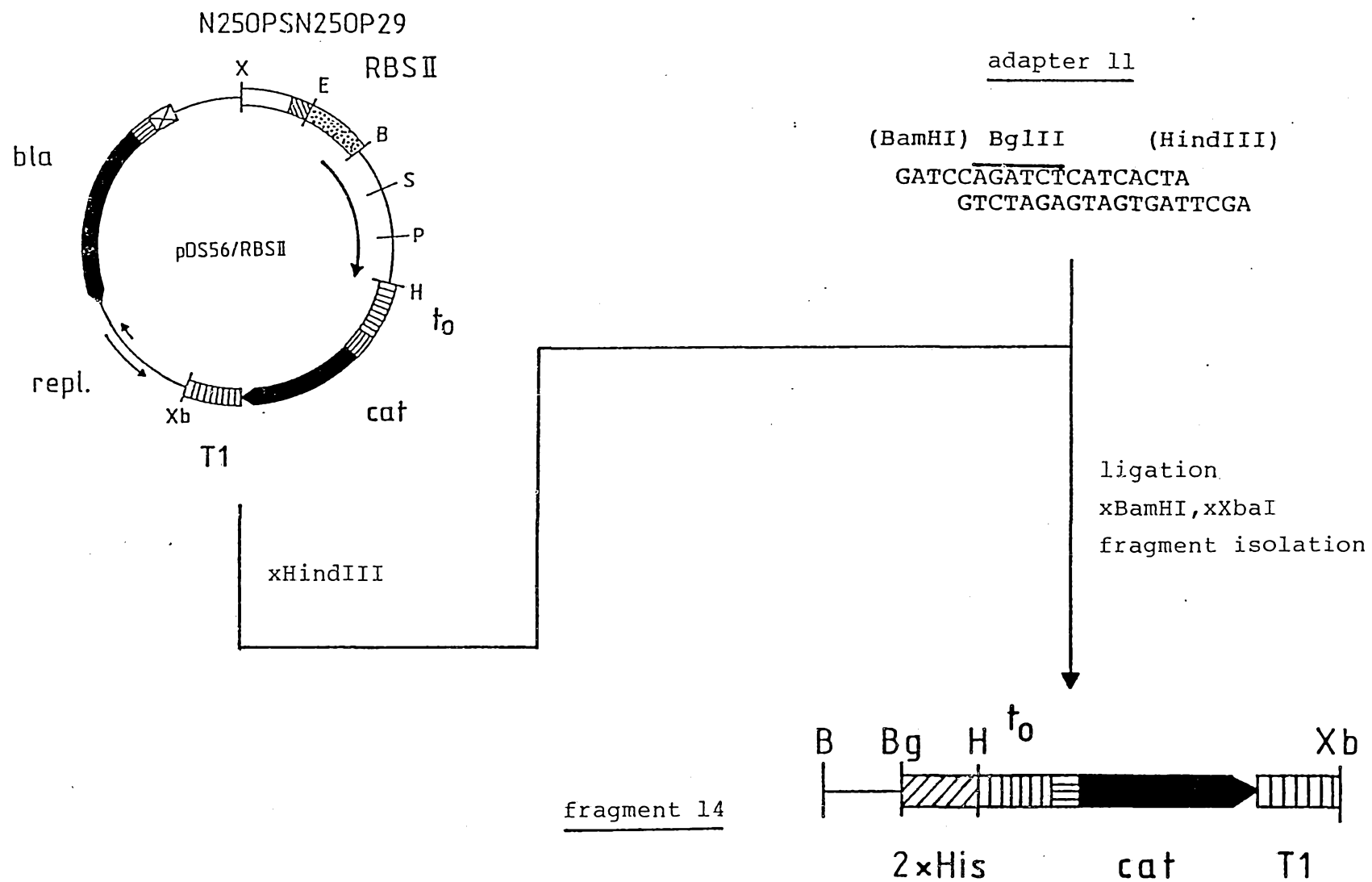


Fig. 35

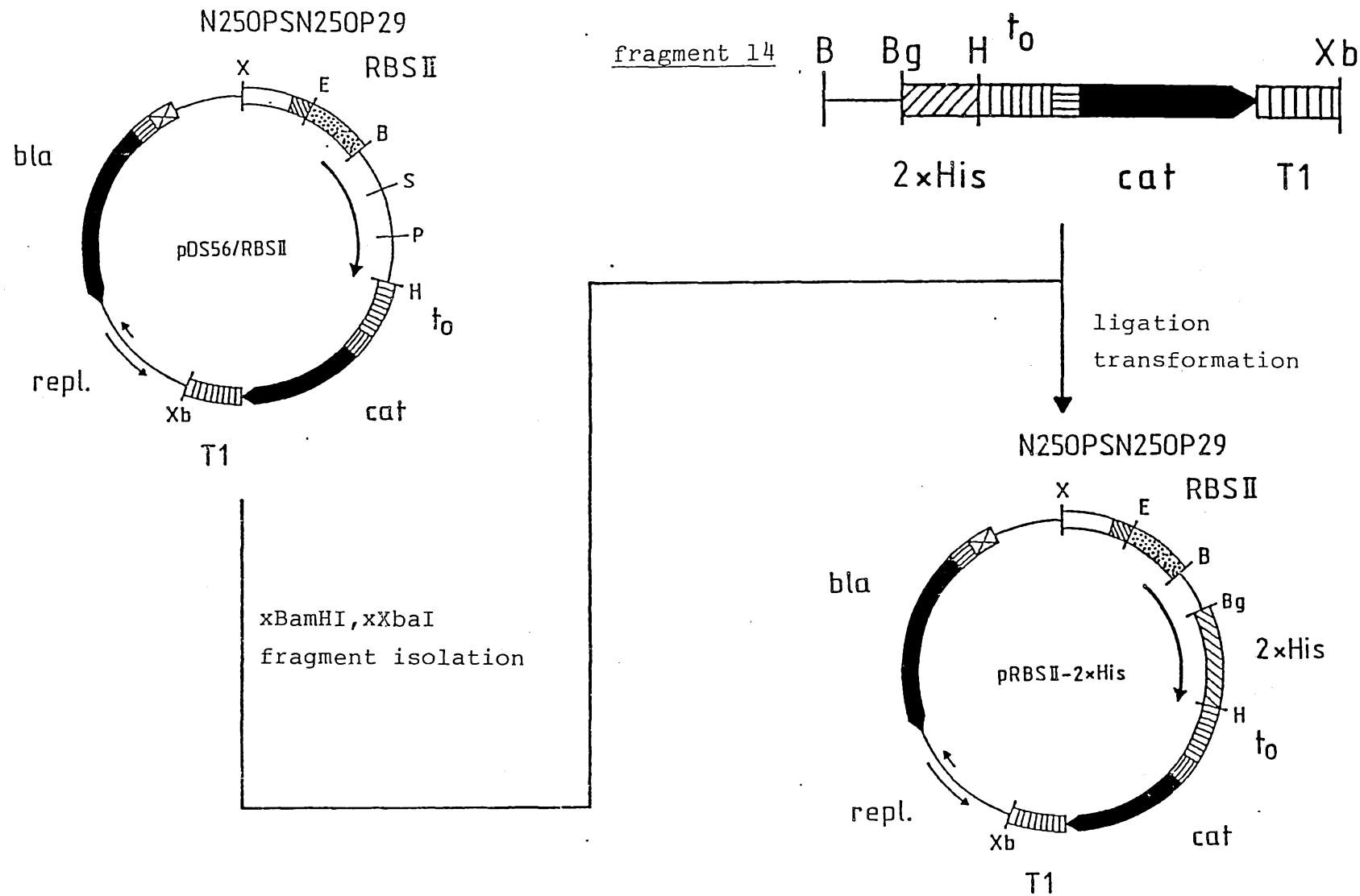


Fig. 36

Fig. 37

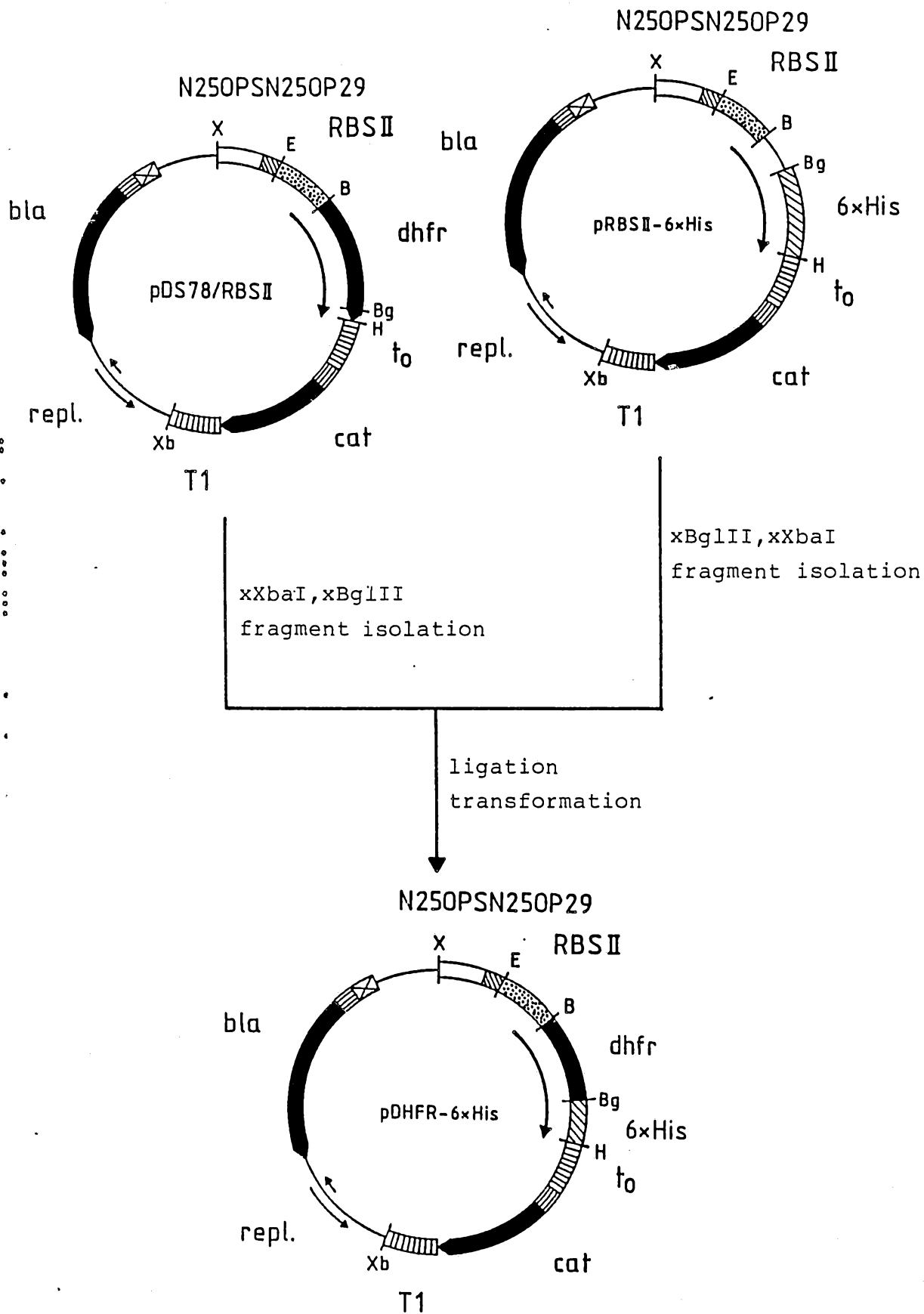


Fig. 38

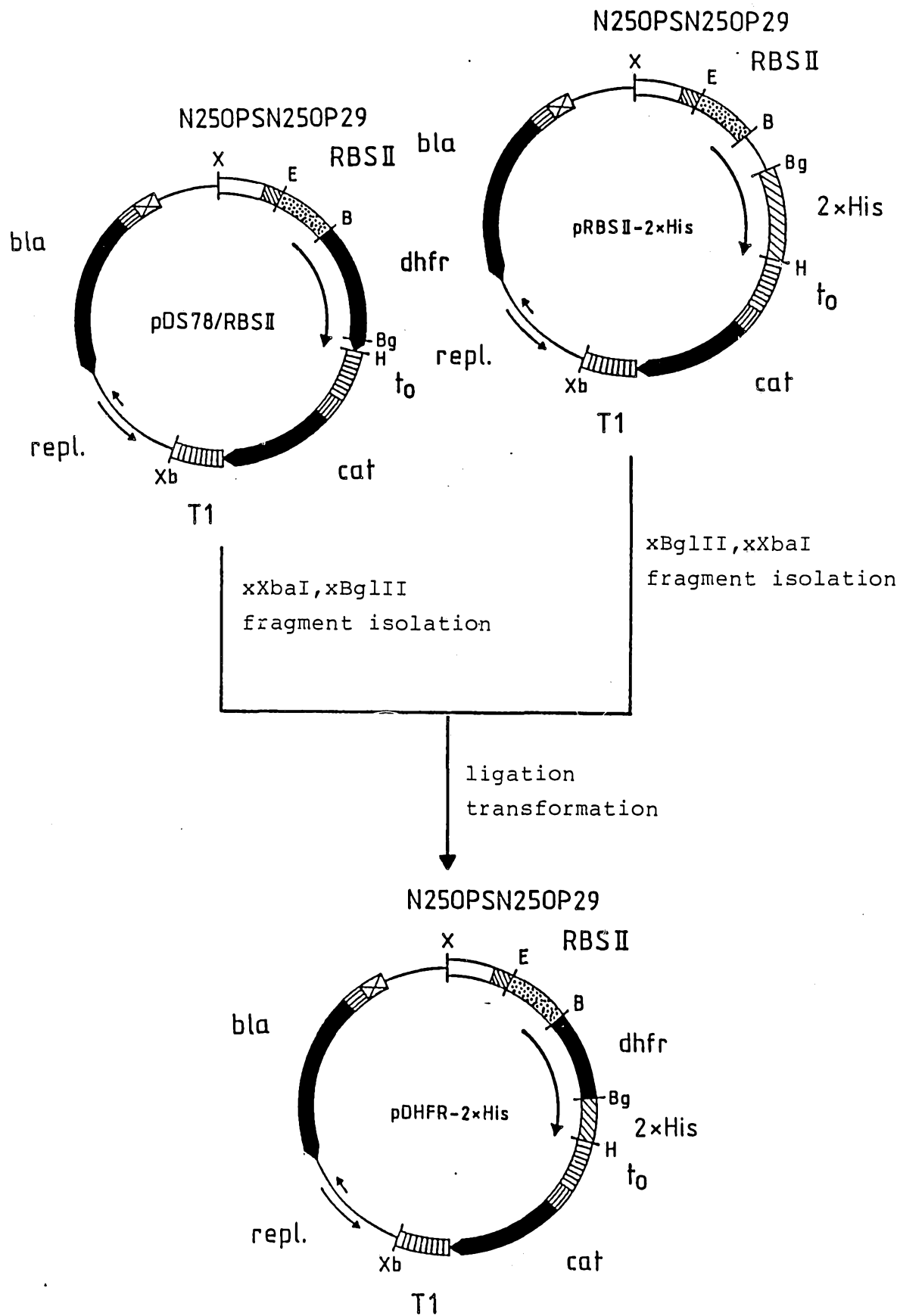


Fig. 39

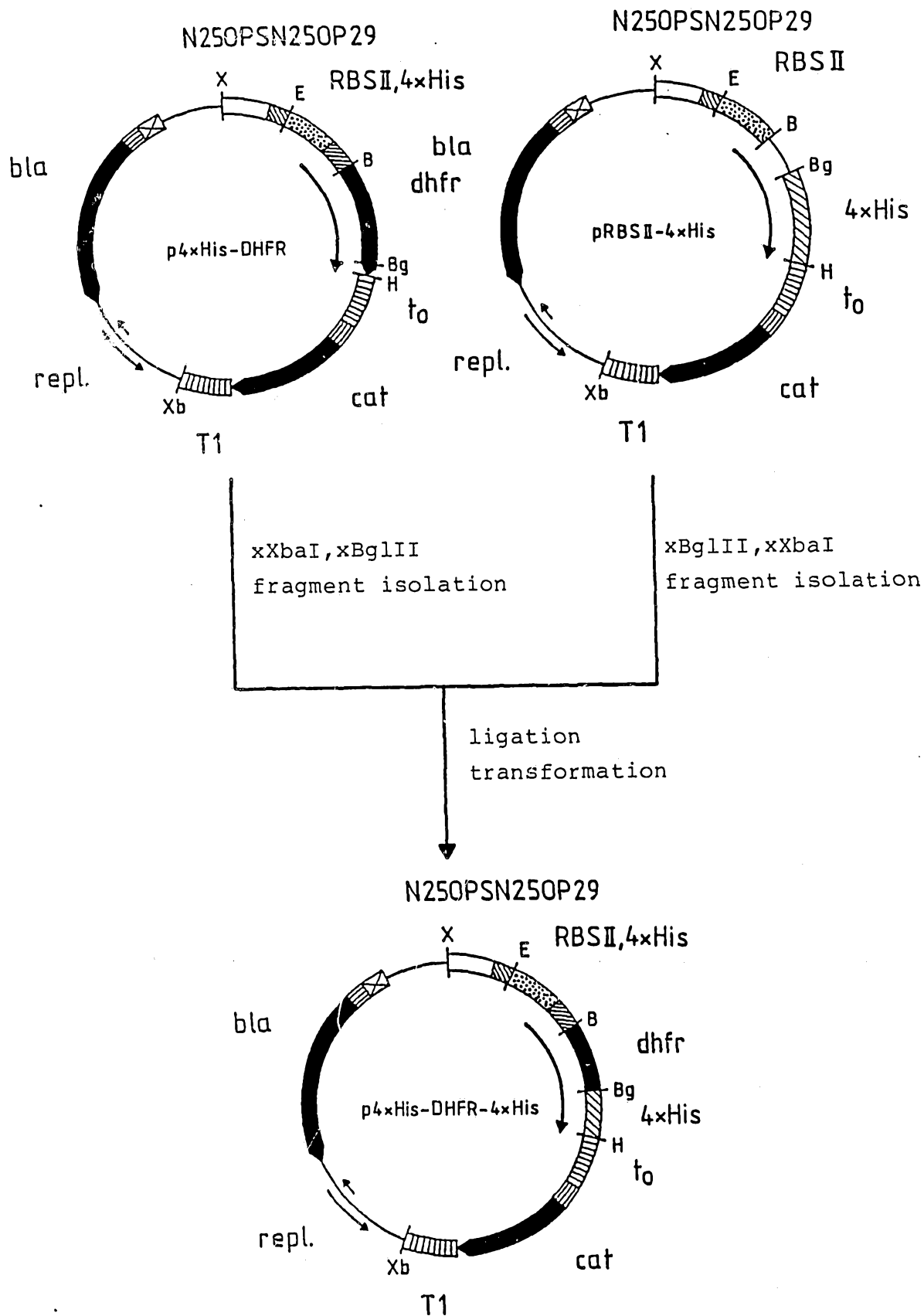


Fig. 40

			10				20				30				40			
ATG	CAG	GAC	CCA	TAT	GTA	AAA	GAA	GCA	GAA	AAC	CTT	AAG	AAA	TAT				
<u>MET</u>	Gln	Asp	Pro	Tyr	Val	Lys	Glu	Ala	Glu	Asn	Leu	Lys	Lys	Tyr				
		50				60				70				80				90
TTT	AAT	GCA	GGT	CAT	TCA	GAT	GTA	GCG	GAT	AAT	GGA	ACT	CTT	TTC				
Phe	Asn	Ala	Gly	His	Ser	Asp	Val	Ala	Asp	Asn	Gly	Thr	Leu	Phe				
			100				110				120				130			
TTA	GGC	ATT	TTG	AAG	AAT	TGG	AAA	GAG	GAG	AGT	GAC	AGA	AAA	ATA				
Leu	Gly	Ile	Leu	Lys	Asn	Trp	Lys	Glu	Glu	Ser	Asp	Arg	Lys	Ile				
		140				150				160				170				180
ATG	CAG	AGC	CAA	ATT	GTC	TCC	TTT	TAC	TTC	AAA	CTT	TTT	AAA	AAC				
<u>MET</u>	Gln	Ser	Gln	Ile	Val	Ser	Phe	Tyr	Phe	Lys	Leu	Phe	Lys	Asn				
			190				200				210				220			
TTT	AAA	GAT	GAC	CAG	AGC	ATC	CAA	AAG	AGT	GTG	GAG	ACC	ATC	AAG				
Phe	Lys	Asp	Asp	Gln	Ser	Ile	Gln	Lys	Ser	Val	Glu	Thr	Ile	Lys				
		230				240				250				260				270
GAA	GAC	ATG	AAT	GTC	AAG	TTT	TTC	AAT	AGC	AAC	AAA	AAG	AAA	CGA				
Glu	Asp	<u>MET</u>	Asn	Val	Lys	Phe	Phe	Asn	Ser	Asn	Lys	Lys	Lys	Arg				
			280				290				300				310			
GAT	GAC	TTC	GAA	AAG	CTG	ACT	AAT	TAT	TCG	GTA	ACT	GAC	TTG	AAT				
Asp	Asp	Phe	Glu	Lys	Leu	Thr	Asn	Tyr	Ser	Val	Thr	Asp	Leu	Asn				
		320				330				340				350				360
GTC	CAA	CGC	AAA	GCA	ATA	CAT	GAA	CTC	ATC	CAA	GTG	ATG	GCT	GAA				
Val	Gln	Arg	Lys	Ala	Ile	His	Glu	Leu	Ile	Gln	Val	<u>MET</u>	Ala	Glu				

Fig. 40 (continued)

			370			380			390			400		
CTG	TCG	CCA	GCA	GCT	AAA	ACA	GGG	AAG	CGA	AAA	AGG	AGT	CAG	ATG
Leu	Ser	Pro	Ala	Ala	Lys	Thr	Gly	Lys	Arg	Lys	Arg	Ser	Gln	<u>MET</u>

	410			420			430		
CTG	TTT	CGA	GGT	CGA	AGA	GCA	TCC	CAG	
Leu	Phe	Arg	Gly	Arg	Arg	Ala	Ser	Gln	

Fig. 41

			10			20			30			40		
ATG	CAT	CAC	GCC	GGC	ATC	GAA	GGG	CGC	CAA	GAT	CCA	TAT	GTA	AAA
<u>MET</u>	His	His	Ala	Gly	Ile	Glu	Gly	Arg	Gln	Asp	Pro	Tyr	Val	Lys
	50			60			70			80			90	
GAA	GCA	GAA	AAC	CTT	AAG	AAA	TAT	TTT	AAT	GCA	GGT	CAT	TCA	GAT
Glu	Ala	Glu	Asn	Leu	Lys	Lys	Tyr	Phe	Asn	Ala	Gly	His	Ser	Asp
		100			110			120			130			
GTA	GCG	GAT	AAT	GGA	ACT	CTT	TTC	TTA	GGC	ATT	TTG	AAG	AAT	TGG
Val	Ala	Asp	Asn	Gly	Thr	Leu	Phe	Leu	Gly	Ile	Leu	Lys	Asn	Trp
	140			150			160			170			180	
AAA	GAG	GAG	AGT	GAC	AGA	AAA	ATA	ATG	CAG	AGC	CAA	ATT	GTC	TCC
Lys	Glu	Glu	Ser	Asp	Arg	Lys	Ile	<u>MET</u>	Gln	Ser	Gln	Ile	Val	Ser
		190			200			210			220			
TTT	TAC	TTC	AAA	CTT	TTT	AAA	AAC	TTT	AAA	GAT	GAC	CAG	AGC	ATC
Phe	Tyr	Phe	Lys	Leu	Phe	Lys	Asn	Phe	Lys	Asp	Asp	Gln	Ser	Ile
	230			240			250			260			270	
CAA	AAG	AGT	GTG	GAG	ACC	ATC	AAG	GAA	GAC	ATG	AAT	GTC	AAG	TTT
Gln	Lys	Ser	Val	Glu	Thr	Ile	Lys	Glu	Asp	<u>MET</u>	Asn	Val	Lys	Phe
		280			290			300			310			
TTC	AAT	AGC	AAC	AAA	AAG	AAA	CGA	GAT	GAC	TTC	GAA	AAG	CTG	ACT
Phe	Asn	Ser	Asn	Lys	Lys	Lys	Arg	Asp	Asp	Phe	Glu	Lys	Leu	Thr
	320			330			340			350			360	
AAT	TAT	TCG	GTA	ACT	GAC	TTG	AAT	GTC	CAA	CGC	AAA	GCA	ATA	CAT
Asn	Tyr	Ser	Val	Thr	Asp	Leu	Asn	Val	Gln	Arg	Lys	Ala	Ile	His

Fig. 41 (continued)

			370			380			390			400		
GAA	CTC	ATC	CAA	GTG	ATG	GCT	GAA	CTG	TCG	CCA	GCA	GCT	AAA	ACA
Glu	Leu	Ile	Gln	Val	<u>MET</u>	Ala	Glu	Leu	Ser	Pro	Ala	Ala	Lys	Thr

410				420				430			440			450
GGG	AAG	CGA	AAA	AGG	AGT	CAG	ATG	CTG	TTT	CGA	GGT	CGA	AGA	GCA
Gly	Lys	Arg	Lys	Arg	Ser	Gln	<u>MET</u>	Leu	Phe	Arg	Gly	Arg	Arg	Ala

TCC CAG
Ser Gln

Fig. 42

			10				20			30			40		
ATG	CAT	CAC	GCC	GGC	GAT	GAC	GAT	GAC	AAA	CAA	GAT	CCA	TAT	GTA	
<u>MET</u>	His	His	Ala	Gly	Asp	Asp	Asp	Asp	Lys	Gln	Asp	Pro	Tyr	Val	
	50			60			70			80			90		
AAA	GAA	GCA	GAA	AAC	CTT	AAG	AAA	TAT	TTT	AAT	GCA	GGT	CAT	TCA	
Lys	Glu	Ala	Glu	Asn	Leu	Lys	Lys	Tyr	Phe	Asn	Ala	Gly	His	Ser	
	100			110			120			130					
GAT	GTA	GCG	GAT	AAT	GGA	ACT	CTT	TTC	TTA	GGC	ATT	TTG	AAG	AAT	
Asp	Val	Ala	Asp	Asn	Gly	Thr	Leu	Phe	Leu	Gly	Ile	Leu	Lys	Asn	
	140			150			160			170			180		
TGG	AAA	GAG	GAG	AGT	GAC	AGA	AAA	ATA	ATG	CAG	AGC	CAA	ATT	GTC	
Trp	Lys	Glu	Glu	Ser	Asp	Arg	Lys	Ile	<u>MET</u>	Gln	Ser	Gln	Ile	Val	
	190			200			210			220					
TCC	TTT	TAC	TTC	AAA	CTT	TTT	AAA	AAC	TTT	AAA	GAT	GAC	CAG	AGC	
Ser	Phe	Tyr	Phe	Lys	Leu	Phe	Lys	Asn	Phe	Lys	Asp	Asp	Gln	Ser	
	230			240			250			260			270		
ATC	CAA	AAG	AGT	GTG	GAG	ACC	ATC	AAG	GAA	GAC	ATG	AAT	GTC	AAG	
Ile	Gln	Lys	Ser	Val	Glu	Thr	Ile	Lys	Glu	Asp	<u>MET</u>	Asn	Val	Lys	
	280			290			300			310					
TTT	TTC	AAT	AGC	AAC	AAA	AAG	AAA	CGA	GAT	GAC	TTC	GAA	AAG	CTG	
Phe	Phe	Asn	Ser	Asn	Lys	Lys	Lys	Arg	Asp	Asp	Phe	Glu	Lys	Leu	

Fig. 42 (continued)

320		330		340		350		360
ACT AAT TAT TCG GTA ACT GAC TTG AAT GTC CAA CGC AAA GCA ATA								
Thr Asn Tyr Ser Val Thr Asp Leu Asn Val Gln Arg Lys Ala Ile								
	370		380		390		400	
CAT GAA CTC ATC CAA GTG ATG GCT GAA CTG TCG CCA GCA GCT AAA								
His Glu Leu Ile Gln Val <u>MET</u> Ala Glu Leu Ser Pro Ala Ala Lys								
	410		420		430			
ACA GGG AAG CGA AAA AGG AGT CAG ATG CTG								
Thr Gly Lys Arg Lys Arg Ser Gln <u>MET</u> Leu								

Fig. 43

			10				20			30			40		
ATG	CAT	CAC	GCC	GGC	ATC	GAA	GGG	CGC	CAA	AAC	CCA	TAT	GTA	AAA	
<u>MET</u>	His	His	Ala	Gly	Ile	Glu	Gly	Arg	Gln	Asn	Pro	Tyr	Val	Lys	
	50			60				70			80			90	
GAA	GCA	GAA	AAC	CTT	AAG	AAA	TAT	TTT	AAT	GCA	GGT	CAT	TCA	GAT	
Glu	Ala	Glu	Asn	Leu	Lys	Lys	Tyr	Phe	Asn	Ala	Gly	His	Ser	Asp	
		100			110			120			130				
GTA	GCG	GAT	AAT	GGA	ACT	CTT	TTC	TTA	GGC	ATT	TTG	AAG	AAT	TGG	
Val	Ala	Asp	Asn	Gly	Thr	Leu	Phe	Leu	Gly	Ile	Leu	Lys	Asn	Trp	
	140			150			160			170			180		
AAA	GAG	GAG	AGT	GAC	AGA	AAA	ATA	ATG	CAG	AGC	CAA	ATT	GTC	TCC	
Lys	Glu	Glu	Ser	Asp	Arg	Lys	Ile	<u>MET</u>	Gln	Ser	Gln	Ile	Val	Ser	
		190			200			210			220				
TTT	TAC	TTC	AAA	CTT	TTT	AAA	AAC	TTT	AAA	GAT	GAC	CAG	AGC	ATC	
Phe	Tyr	Phe	Lys	Leu	Phe	Lys	Asn	Phe	Lys	Asp	Asp	Gln	Ser	Ile	
	230			240			250			260			270		
CAA	AAG	AGT	GTG	GAG	ACC	ATC	AAG	GAA	GAC	ATG	AAT	GTC	AAG	TTT	
Gln	Lys	Ser	Val	Glu	Thr	Ile	Lys	Glu	Asp	<u>MET</u>	Asn	Val	Lys	Phe	
		280			290			300			310				
TTC	AAT	AGC	AAC	AAA	AAG	AAA	CGA	GAT	GAC	TTC	GAA	AAG	CTG	ACT	
Phe	Asn	Ser	Asn	Lys	Lys	Lys	Arg	Asp	Asp	Phe	Glu	Lys	Leu	Thr	

Fig. 43 (continued)

320		330		340		350		360
AAT TAT TCG GTA ~~AGT~~ CAC TTG AAT GTC CAA CGC AAA GCA ATA CAT								
Asn Tyr Ser Val Thr Asp Leu Asn Val Gln Arg Lys Ala Ile His								
	370		380		390		400	
GAA CTC ATC CAA GTG ATG GCT GAA CTG TCG CCA GCA GCT AAA ACA								
Glu Leu Ile Gln Val <u>MET</u> Ala Glu Leu Ser Pro Ala Ala Lys Thr								
	410		420		430			
GGG AAG CGA AAA AGG AGT CAG ATG CTG								
Gly Lys Arg Lys Arg Ser Gln <u>MET</u> Leu								