Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2248653 C 2003/08/12

(11)(21) 2 248 653

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 1997/04/25
(87) Date publication PCT/PCT Publication Date: 1997/11/06
(45) Date de délivrance/Issue Date: 2003/08/12
(85) Entrée phase nationale/National Entry: 1998/09/09
(86) N° demande PCT/PCT Application No.: US 1997/007039
(87) N° publication PCT/PCT Publication No.: 1997/041244
(30) Priorités/Priorities: 1996/05/01 (08/644,664) US;
1996/12/06 (08/761,277) US

(51) Cl.Int.^{6}/Int.Cl.^{6} C12N 15/69, A61K 39/395, A61K 39/39, C12N 15/85, C12N 15/64, C12N 5/10

(72) Inventeur/Inventor:
DENNEY, DAN W., JR., US

(73) Propriétaire/Owner:
GENITOPE CORPORATION, US

(74) Agent: SMART & BIGGAR

(54) Titre : VACCINS SERVANT A TRAITER LES LYMPHOMES ET LA LEUCEMIE
(54) Title: VACCINES FOR TREATMENT OF LYMPHOMA AND LEUKEMIA

(57) Abrégé/Abstract:
The present invention provides multivalent vaccines for the treatment of B-cell malignancies (e.g., lymphomas and leukemias). The present invention also provides methods for the production of custom vaccines, including multivalent vaccines for the treatment of immune cell tumors malignancies as well as methods of treating immune cell tumors using custom vaccines. The present invention further provides improved methods for the amplification and expression of recombinant genes in cells. The methods of the present invention permit the isolation of cell lines which have co-amplified input recombinant sequences which encode an amplifiable marker, one or more expression vectors encoding a protein of interest and optionally a selectable marker. The present invention also provides compositions comprising amplified T lymphoid cell lines.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/86, 15/63, 15/00, C12P 19/24, A61K 39/395, C07H 21/04

A1

(11) International Publication Number: WO 97/41244

(43) International Publication Date: 6 November 1997 (06.11.97)

(21) International Application Number: PCT/US97/07039

(22) International Filing Date: 25 April 1997 (25.04.97)

(30) Priority Data:

08/644,664	1 May 1996 (01.05.96)	US
08/761,277	6 December 1996 (06.12.96)	US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/761,277 (CIP)
Filed on 6 December 1996 (06.12.96)

(71) Applicant *(for all designated States except US)*: GENITOPE CORPORATION [US/US]; Building 1, Suite 185, 3000 Sand Hill Road, Menlo Park, CA 94025 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: DENNEY, Dan, W., Jr. [US/US]; 1003 Palomar Drive, Redwood City, CA 94062 (US).

(74) Agents: CARROLL, Peter, G. et al.; Medlen & Carroll, LLP, Suite 2200, 220 Montgomery Street, San Francisco, CA 94104 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: VACCINES FOR TREATMENT OF LYMPHOMA AND LEUKEMIA

(57) Abstract

The present invention provides multivalent vaccines for the treatment of B-cell malignancies (e.g., lymphomas and leukemias). The present invention also provides methods for the production of custom vaccines, including multivalent vaccines for the treatment of immune cell tumors malignancies as well as methods of treating immune cell tumors using custom vaccines. The present invention further provides improved methods for the amplification and expression of recombinant genes in cells. The methods of the present invention permit the isolation of cell lines which have co-amplified input recombinant sequences which encode an amplifiable marker, one or more expression vectors encoding a protein of interest and optionally a selectable marker. The present invention also provides compositions comprising amplified T lymphoid cell lines.

WO 97/41244 PCT/US97/07039

VACCINES FOR TREATMENT OF LYMPHOMA AND LEUKEMIA

FIELD OF THE INVENTION

The present invention generally relates to improved methods for the amplification and expression of recombinant genes in cells. The amplified cells provide large quantities of recombinant proteins suitable for immunotherapy for treatment of lymphomas and leukemias.

BACKGROUND OF THE INVENTION

As an increasing number of genes are isolated and developed for the expression of a wide array of useful polypeptide drugs, there is an increasing need to enhance the efficiencies and economies of the process. It is advantageous to obtain such polypeptides from mammalian cells since such polypeptides or proteins are generally correctly folded, appropriately modified and completely functional, often in marked contrast to those proteins as expressed in bacterial cells.

When large amounts of product are required, it is necessary to identify cell clones in which the vector sequences are maintained (*i.e.*, retained) during cell proliferation. Such stable vector maintenance can be achieved either as a consequence of integration of the vector into the DNA of the host cell or by use of a viral replicon such as bovine papillomavirus (BPV).

The use of viral vectors such as BPV-based vectors for the generation of stable cell lines expressing large amounts of a recombinant protein has been successful in some cases; however, the use of viral vectors is limited by the fact that the viral vectors are restricted in the cell types in which they can replicate. Furthermore expression levels and episomal maintenance of the viral vector can be influenced by the DNA sequences inserted into the vector.

Where the vector has been integrated into the genomic DNA of the host cell to improve stability, the copy number of the vector DNA, and concomitantly the amount of product which could be expressed, can be increased by selecting for cell lines in which the vector sequences have been amplified after integration into the DNA of the host cell.

A known method for carrying out such a selection procedure is to transform a host cell with a vector comprising a DNA sequence which encodes an enzyme which is inhibited by a known drug. The vector may also comprise a DNA sequence which encodes a desired

protein. Alternatively the host cell may be co-transformed with a second vector which comprises the DNA sequence which encodes the desired protein.

The transformed or co-transformed host cells are then cultured in increasing concentrations of the known drug hereby selecting drug-resistant cells. It has been found that one common mechanism leading to the appearance of mutant cells which can survive in the increased concentrations of the otherwise toxic drug is the over-production of the enzyme which is inhibited by the drug. This most commonly results from increased levels of its particular mRNA, which in turn is frequently caused by amplification of vector DNA and hence gene copies.

It has also been found that when drug resistance is caused by an increase in copy number of the vector DNA encoding the inhibitable enzyme, there is a concomitant increase in the copy number of the vector DNA encoding the desired protein in the DNA of the host cell. There is thus an increased level of production of the desired protein.

The most commonly used system for such co-amplification uses dihydrofolate reductase (DHFR) as the inhibitable enzyme. This enzyme can be inhibited by the drug methotrexate (MTX). To achieve co-amplification, a host cell which lacks an active gene which encodes DHFR is either transformed with a vector which comprises DNA sequences encoding DHFR and a desired protein or co-transformed with a vector comprising a DNA sequence encoding DHFR and a vector comprising a DNA sequence encoding the desired protein. The transformed or co-transformed host cells are cultured in media containing increasing levels of MTX, and those cell lines which survive are selected.

The co-amplification systems which are presently available suffer from a number of disadvantages. For instance, it is generally necessary to use a host cell which lacks an active gene encoding the enzyme which can be inhibited. This tends to limit the number of cell lines which can be used with any particular co-amplification system.

For instance, there are at present, only two cell lines known which lack the gene encoding DHFR and both of these cell lines are derivatives of the CHO-K1 cell line. These $DHFR^-$ CHO cell lines cannot be used to express certain protein products at high levels because CHO cells lack specialized postranslational modification pathways. For example, the production of functional human protein C requires that the cell possess the vitamin K-dependent γ-carboxylation pathway; CHO cells cannot properly modify the human protein C protein [Walls *et al.*, (1989) Gene 81:139].

Attempts to use DHFR genes as dominant selectable markers in other cell lines (*i.e.*, cell lines synthesizing wild type levels of DHFR) has not proved satisfactory. For instance, a MTX-resistant mutant DHFR or a DHFR gene under the control of a very strong promoter can act as a dominant selectable marker in certain cell types but such high concentrations of MTX are required that it has not been possible to achieve high copy numbers by selection for gene amplification using current methodologies.

Another approach to allow the use of DHFR as a dominant selectable marker in $DHFR^+$ cell lines is the use of both the DHFR gene and a gene encoding a selectable marker, such as the hygromycin phosphotransferase (*hyg*) gene, in addition to the gene of interest [Walls, *et al.* (1989), *supra*]. This approach is used to circumvent the problem of amplification of the endogenous *dhfr* gene during selection with MTX. The cells are transfected with DNA encoding the three genes and the cells are first selected for their ability to grow in hygromycin. The cells are then selected for the ability to grow in increasing concentrations of MTX. While this approach allows for the co-amplification of genes in $dhfr^+$ cell lines, present protocols show that the *dhfr* gene is amplified to a higher degree than the gene of interest with successive rounds of amplification (*i.e.*, stepwise increases in MTX concentration). For example, in several amplified clones the *dhfr* gene was present at approximately 100 copies while the gene of interest was present at only 20 copies.

Clearly, the art needs improved methods which would consistently provide for the coincidental amplification of the amplifiable marker and the gene of interest in a variety of cell lines. Furthermore, the art needs a means of amplifying DNA sequences of interest which is efficient, reproducible and which is not limited to the use of specialized enzyme deficient host cell lines or to a limited number of cell lines.

Improved methods which consistently provide for the coincidental amplification of the amplifiable marker and the gene of interest in a variety of cell lines and which are efficient and reproducible would allow the production of custom tumor-specific vaccines on a scale commensurate with patient demand. Current methods for producing custom tumor vaccines for the treatment of B-cell lymphoma are insufficient to meet current and anticipated future demand.

SUMMARY OF THE INVENTION

The present invention provides methods for the production of cell lines containing amplified copies of recombinant DNA sequences. Because the amplified cell lines contain

several different recombinant DNA sequences (*e.g.*, the amplification vector, one or more expression vectors and optionally a selection vector) which are coordinately amplified, the cell lines are the to have co-amplified the input or exogenous DNA sequences. The methods of the present invention permit the efficient isolation of the desired amplified cell lines with a considerable savings in time relative to existing amplification protocols. The gene amplification methods of the present invention permit the production of custom vaccines, including multivalent vaccines, which are useful for the treatment of immune cell tumors (*e.g.*, lymphomas and leukemias).

In one embodiment, the present invention provides a multivalent vaccine comprising at least two recombinant variable regions of immunoglobulin molecules derived from B-cell lymphoma cells, wherein the cells express at least two different immunoglobulin molecules, the immunoglobulin molecules differing by at least one idiotope. The invention is not limited by the context in which the recombinant variable regions are utilized; the variable regions may be present within an entire recombinant immunoglobulin (Ig) molecule, they may be present on Fab, Fab' or $F(ab')_2$ fragments (which may be generated by cleavage of the recombinant Ig molecule or they may be produced using molecular biological means) or they may be present on single chain antibody (Fv) molecules. In a preferred embodiment, the multivalent vaccine comprises at least two recombinant immunoglobulin molecules comprising the recombinant variable regions derived from the lymphoma cells.

In one embodiment, the immunoglobulin molecules comprising recombinant variable regions derived from a patient's lymphoma cells are covalently linked to an immune-enhancing cytokine. The linkage of the cytokine to the Ig molecule may be achieved by a variety of means known to the art including conventional coupling techniques (*e.g.*, coupling with dehydrating agents such as dicyclohexylcarbodiimide (DCCI), ECDI and the like), the use of linkers capable of coupling through sulfhydryl groups, amino groups or carboxyl groups (available from Pierce Chemical Co., Rockford, IL), by reductive amination. In addition, the covalent linkage may be achieved by molecular biological means (*e.g.*, the production of a fusion protein using an expression vector comprising a nucleotide sequence encoding the recombinant Ig operably linked to a nucleotide sequence encoding the desired cytokine).

The invention is not limited by the immune-enhancing cytokine employed. In a preferred embodiment, the cytokine is selected from the group consisting of granulocyte-macrophage colony stimulating factor, interleukin-2 and interleukin-4.

In one embodiment, the multivalent vaccines of the present invention comprise at least one pharmaceutically acceptable excipient. The invention is not limited by the nature of the excipient employed. The pharmaceutical compositions of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

In a preferred embodiment, the multivalent vaccine further comprises an adjuvant. When the vaccine is to be administered to a human subject, adjuvants approved for use in humans are employed (*e.g.*, SAF-1, alum, etc.). Additional adjuvants suitable for use in humans, include but are not limited to QS21 (Aquila Biopharmaceuticals, Inc., Worcester, MA), Optivax™ (Vaxcell Inc., Norcross, GA) and adjuvants manufactured by Ribi Immunochem Research, Inc. (Hamilton, MT). The recombinant Ig proteins (including fragments of Ig proteins) which comprise the multivalent vaccine may be conjugated to a carrier protein such as KLH.

The present invention further provides a composition comprising a T lymphoid cell line having an amplified gene, the amplified gene being amplified from a plurality of exogenous integrated nucleic acid, the exogenous nucleic acid comprising a recombinant oligonucleotide having a sequence encoding an inhibitable enzyme operably linked to a heterologous promoter. The compositions of the present invention are not limited by the nature of the inhibitable enzyme encoded by the recombinant oligonucleotide. In a preferred embodiment, the recombinant oligonucleotide having a sequence encoding an inhibitable enzyme encodes an active dihydrofolate reductase.

In yet another preferred embodiment, the T lymphoid cell line containing amplified copies of an exogenous (*i.e.*, input) recombinant oligonucleotide having a sequence encoding an inhibitable enzyme operably linked to a heterologous promoter is the BW5147.G.1.4 cell line. In another preferred embodiment, the T lymphoid cell line containing amplified copies of an exogenous (*i.e.*, input) recombinant oligonucleotide having a sequence encoding an

inhibitable enzyme operably linked to a heterologous promoter further comprises an integrated recombinant oligonucleotide comprising a gene encoding a protein of interest (*i.e.*, an expression vector). In yet another preferred embodiment, the T lymphoid cell line containing amplified copies of an exogenous (*i.e.*, input) recombinant oligonucleotide having a sequence encoding an inhibitable enzyme operably linked to a heterologous promoter further comprises an integrated recombinant oligonucleotide comprising a gene encoding a selectable marker (*i.e.*, a selection vector).

The present invention further provides a method for co-amplifying a first recombinant oligonucleotide having a sequence which encodes the amino acid sequence of a protein of interest and a second recombinant oligonucleotide having a sequence encoding an inhibitable enzyme operably linked to a heterologous promoter, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and iii) a eukaryotic parent cell line; b) introducing 400 to 500 micrograms of the expression vector [*i.e.*, a total of 400-500 μg of the expression vector(s)] and 20 to 30 micrograms of the amplification vector into the parent cell line to generate transformed cells; c) introducing the transformed cells into a first aqueous solution containing an inhibitor capable of inhibiting the first inhibitable enzyme, wherein the concentration of the inhibitor present in the first aqueous solution is sufficient to prevent growth of the parent cell line; and d) identifying a transformed cell capable of growth in the first aqueous solution, wherein the transformed cell capable of growth contains an amplified number of copies of the expression vector and an amplified number of copies of the amplification vector.

The method of the present invention is not limited by the nature of the means employed to introduce the vectors into the parent cell line. The art is well aware of numerous methods which allow the introduction of exogenous DNA sequences into mammalian cells, including but not limited to electroporation, microinjection, lipofection, protoplast fusion, liposome fusion and the like. In a preferred embodiment, the vectors are introduced into the parent cell line by electroporation.

The present invention is not limited by the nature of the cell line chosen as the parent cell line; a variety of mammalian cell lines may be employed including CHO cell lines and variants thereof, mouse L cells and BW5147 cells and variants thereof. The chosen cell line

grow in either an attachment-dependent or attachment-independent manner. In a preferred embodiment. the parent cell line is a T lymphoid cell line; a particularly preferred T lymphoid cell line is the BW5147.G.1.4 cell line.

In a preferred embodiment, the parent cell line contains an endogenous gene encoding a second inhibitable enzyme; the first and second inhibitable enzymes may be, but are not required to be, the same. In another embodiment, the second inhibitable enzyme is selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase. In another preferred embodiment, the concentration of inhibitor present in the first aqueous solution is four to six-fold the concentration required to prevent the growth of the parent cell line.

In a preferred embodiment, the method further comprises providing a selection vector encoding a selectable gene product which is introduced into the parent cell line together with the expression vector and the amplification vector. The method is not limited by the selectable marker chosen. In a preferred embodiment, the selection vector encodes an active enzyme selected from the group comprising hypoxanthine guanine phosphoribosyltransferase, hygromycin G phosphotransferase, xanthine-guanine phosphoribosyltransferase and aminoglycoside 3' phosphotransferase. When the selection vector encodes an active hypoxanthine guanine phosphoribosyltransferase, the transformed cells may be grown in an aqueous solution that comprises hypoxanthine and azaserine.

In another preferred embodiment, the method further comprises, following the introduction of the selection, expression and amplification vectors, the additional step of introducing the transformed cells into a second aqueous solution which requires the expression of the selectable gene product prior to introducing the transformed cell into the first aqueous solution containing an inhibitor capable of inhibiting the inhibitable enzyme. The selection, expression and/or amplification vectors may be linearized prior to introduction into the parent cell line.

The method is not limited by the nature of the amplification vector employed. In a preferred embodiment, the amplification vector encodes an active enzyme selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase. In another preferred embodiment the inhibitor is selected from the group consisting of methotrexate, 2'-deoxycoformycin, methionine sulphoximine, albizziin and β-aspartyl hydroxamate.

The present invention also provides a method, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and iii) a eukaryotic parent cell line; b) treating the expression vector and the amplification vector with a restriction enzyme to create a linearized expression vector and a linearized amplification vector (*i.e.*, each vector is digested with a restriction enzyme that cuts once within the vector and further does not cut within the transcription unit encoding the amplifiable marker or gene of interest; that is the site of cutting is within the plasmid backbone); c) introducing the linearized expression vector(s) and the linearized amplification vector into the parent cell line to generate transformed cells; d) introducing the transformed cells into a first aqueous solution, the first aqueous solution comprising an inhibitor capable of inhibiting the first inhibitable enzyme, wherein the concentration of the inhibitor present in the first aqueous solution is sufficient to prevent growth of the parent cell line; and e) identifying a transformed cell capable of growth in the first aqueous solution, wherein the transformed cell capable of growth contains an amplified number of copies of the expression vector and an amplified number of copies of the amplification vector.

The method of the present invention is not limited by the nature of the means employed to introduce the vectors into the parent cell line; in a preferred embodiment, the vectors are introduced into the parent cell line by electroporation.

The present invention is not limited by the nature of the cell line chosen as the parent cell line; a variety of mammalian cell lines may be employed including CHO cell lines and variants thereof, mouse L cells and BW5147 cells and variants thereof. The chosen cell line grow in either an attachment-dependent or attachment-independent manner. In a preferred embodiment. the parent cell line is a T lymphoid cell line; a particularly preferred T lymphoid cell line is the BW5147.G.1.4 cell line.

In a preferred embodiment, the parent cell line contains an endogenous gene encoding a second inhibitable enzyme; the first and second inhibitable enzymes may be, but are not required to be, the same. In another embodiment, the second inhibitable enzyme is selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase. In another preferred embodiment, the concentration of

inhibitor present in the first aqueous solution is four to six-fold the concentration required to prevent the growth of the parent cell line.

In a preferred embodiment, the method further comprises providing a selection vector encoding a selectable gene product which is introduced into the parent cell line together with the expression vector and the amplification vector. The method is not limited by the selectable marker chosen. In a preferred embodiment, the selection vector encodes an active enzyme selected from the group comprising hypoxanthine guanine phosphoribosyltransferase, hygromycin G phosphotransferase, xanthine-guanine phosphoribosyltransferase and aminoglycoside 3' phosphotransferase. When the selection vector encodes an active hypoxanthine guanine phosphoribosyltransferase, the transformed cells may be grown in an aqueous solution that comprises hypoxanthine and azaserine.

In a preferred embodiment, the method further comprises providing a selection vector encoding a selectable gene product which is introduced into the parent cell line together with the expression vector and the amplification vector. The method is not limited by the selection vector employed. In a preferred embodiment, the selection vector encodes an active enzyme selected from the group comprising hypoxanthine guanine phosphoribosyltransferase, hygromycin G phosphotransferase, xanthine-guanine phosphoribosyltransferase and aminoglycoside 3' phosphotransferase. The selection vector is preferably linearized prior to introduction into the parent cell line.

In another preferred embodiment, the method further comprises, following the introduction of the selection, expression and amplification vectors, the additional step of introducing the transformed cells into a second aqueous solution, the second aqueous solution requiring the expression of the selectable gene product for growth of the transformed cells, prior to introducing the transformed cells into the first aqueous solution.

The method is not limited by the nature of the amplification vector employed. In a preferred embodiment, the amplification vector encodes an active enzyme selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase. In another preferred embodiment the inhibitor is selected from the group consisting of methotrexate, 2'-deoxycoformycin, methionine sulphoximine, albizziin and β-aspartyl hydroxamate.

In a preferred embodiment, the method further comprises the steps of: f) introducing the transformed cell capable of growth in the first aqueous solution into a third aqueous solution, the third aqueous solution comprising the inhibitor capable of inhibiting the first

inhibitable enzyme and wherein the concentration of the inhibitor present in the third aqueous solution is sixteen-fold to thirty-six fold the concentration of the inhibitor required to prevent the growth of the parent cell line; and g) identifying at least one transformed cell capable of growth in the third aqueous solution.

In another preferred embodiment, 20 to 30 micrograms of the amplification vector and a total of 400 to 500 micrograms of the expression vector(s) are introduced into the parent cell line. In another preferred embodiment, 10 to 15 micrograms of the selection vector, 20 to 30 micrograms of the amplification vector and a total of 400 to 500 micrograms of the expression vector(s) are introduced into the parent cell line.

The present invention also provides a method, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and iii) a eukaryotic parent cell line; b) introducing the expression vector and the amplification vector into the parent cell line to generate transformed cells; c) introducing the transformed cells into a first aqueous solution, the first aqueous solution comprising an inhibitor capable of inhibiting the first inhibitable enzyme and wherein the concentration of the inhibitor present in the first aqueous solution is four-fold to six-fold the concentration of the inhibitor required to prevent the growth of the parent cell line; and d) identifying at least one transformed cell capable of growth in the first aqueous solution, wherein the transformed cell capable of growth contains an amplified number of copies of the expression vector and an amplified number of copies of the amplification vector. The method of the present invention is not limited by the nature of the means employed to introduce the vectors into the parent cell line; in a preferred embodiment, the vectors are introduced into the parent cell line by electroporation.

The present invention is not limited by the nature of the cell line chosen as the parent cell line; a variety of mammalian cell lines may be employed including CHO cell lines and variants thereof, mouse L cells and BW5147 cells and variants thereof. The chosen cell line grow in either an attachment-dependent or attachment-independent manner. In a preferred embodiment, the parent cell line is a T lymphoid cell line; a particularly preferred T lymphoid cell line is the BW5147.G.1.4 cell line.

In a preferred embodiment, the parent cell line contains an endogenous gene encoding a second inhibitable enzyme; the first and second inhibitable enzymes may be, but are not

required to be. the same. In another embodiment, the second inhibitable enzyme is selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase.

In a preferred embodiment, the method further comprises providing a selection vector encoding a selectable gene product which is introduced into the parent cell line together with the expression vector and the amplification vector. The method is not limited by the selectable marker chosen. In a preferred embodiment, the selection vector encodes an active enzyme selected from the group comprising hypoxanthine guanine phosphoribosyltransferase, hygromycin G phosphotransferase, xanthine-guanine phosphoribosyltransferase and aminoglycoside 3' phosphotransferase. When the selection vector encodes an active hypoxanthine guanine phosphoribosyltransferase, the transformed cells may be grown in an aqueous solution that comprises hypoxanthine and azaserine.

In a preferred embodiment, the method further comprises providing a selection vector encoding a selectable gene product which is introduced into the parent cell line together with the expression vector and the amplification vector. The method is not limited by the selection vector employed. In a preferred embodiment, the selection vector encodes an active enzyme selected from the group comprising hypoxanthine guanine phosphoribosyltransferase, hygromycin G phosphotransferase, xanthine-guanine phosphoribosyltransferase and aminoglycoside 3' phosphotransferase.

In a preferred embodiment, the method further comprises, following the introduction of the expression, amplification and selection vectors, the additional step of introducing the transformed cells into a second aqueous solution, the second aqueous solution requiring the expression of the selectable gene product for growth of the transformed cells, prior to introducing the transformed cells into the first aqueous solution.

The method is not limited by the nature of the amplification vector employed. In a preferred embodiment, the amplification vector encodes an active enzyme selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase. In another preferred embodiment the inhibitor is selected from the group consisting of methotrexate, 2'-deoxycoformycin, methionine sulphoximine, albizziin and β-aspartyl hydroxamate.

In a preferred embodiment, the method further comprises the steps of: e) introducing the transformed cell capable of growth in the first aqueous solution into a third aqueous solution, the third aqueous solution comprising the inhibitor capable of inhibiting the first

inhibitable enzyme and wherein the concentration of the inhibitor present in the third aqueous solution is sixteen-fold to thirty-six-fold the concentration of the inhibitor required to prevent the growth of the parent cell line; and f) identifying at least one transformed cell capable of growth in the third aqueous solution.

In another preferred embodiment, 20 to 30 micrograms of the amplification vector and a total of 400 to 500 micrograms of the expression vector(s) are introduced into the parent cell line. In a preferred embodiment, 10 to 15 micrograms of the selection vector, 20 to 30 micrograms of the amplification vector and a total of 400 to 500 micrograms of the expression vector(s) are introduced into the parent cell line. The selection, expression and/or amplification vectors are preferably linearized prior to introduction into the parent cell line.

The present invention further provide a method, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; iii) a selection vector comprising a third recombinant oligonucleotide having a sequence which encodes a selectable gene product; and iv) a eukaryotic parent cell line; b) introducing the expression vector, the amplification vector and the selection vector into the cell line to generate transformed cells; c) introducing the transformed cells into a first aqueous solution, the first aqueous solution requiring the expression of the selectable gene product for growth of the transformed cells; d) identifying at least one transformed cell capable of growth in the first aqueous solution; e) introducing the transformed cell capable of growth in the first aqueous solution into a second aqueous solution, the second aqueous solution comprising an inhibitor capable of inhibiting the first inhibitable enzyme, wherein the concentration of the inhibitor present in the second aqueous solution is sufficient to prevent growth of the parent cell line; and f) identifying at least one transformed cell capable of growth in the second aqueous solution, wherein the transformed cell capable of growth contains an amplified number of copies of the expression vector and an amplified number of copies of the amplification vector.

The method of the present invention is not limited by the nature of the means employed to introduce the vectors into the parent cell line; in a preferred embodiment, the vectors are introduced into the parent cell line by electroporation. The method of the present invention is not limited by the parent cell line employed. In a preferred embodiment, the parent cell line is a T lymphoid cell line, most preferably the BW5147.G.1.4 cell line. The

selection, expression and/or amplification vectors are preferably linearized prior to introduction into the parent cell line.

In a preferred embodiment, the parent cell line contains an endogenous gene encoding a second inhibitable enzyme; the first and second inhibitable enzymes may be, but are not required to be, the same. In another embodiment, the second inhibitable enzyme is selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase.

In a preferred embodiment, 10 to 15 micrograms of the selection vector, 20 to 30 micrograms of the amplification vector and a total of 400 to 500 micrograms of the expression vector(s) are introduced into the parent cell line.

In another preferred embodiment, the concentration of inhibitor present in the second aqueous solution is four-fold to six-fold the concentration required to prevent the growth of the parent cell line. In a preferred embodiment, the method further comprises the steps of: g) introducing the transformed cell capable of growth in the second aqueous solution into a third aqueous solution, the third aqueous solution comprising the inhibitor capable of inhibiting the first inhibitable enzyme and wherein the concentration of the inhibitor present in the third aqueous solution is sixteen-fold to thirty-six-fold the concentration of the inhibitor required to prevent the growth of the parent cell line; and h) identifying at least one transformed cell capable of growth in the third aqueous solution.

The method is not limited by the selection vector employed. In a preferred embodiment. the selection vector encodes an active enzyme selected from the group comprising hypoxanthine guanine phosphoribosyltransferase, hygromycin G phosphotransferase, xanthine-guanine phosphoribosyltransferase and aminoglycoside 3' phosphotransferase.

The method is not limited by the nature of the amplification vector employed. In a preferred embodiment, the amplification vector encodes an active enzyme selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase. In another preferred embodiment the inhibitor is selected from the group consisting of methotrexate, 2'-deoxycoformycin, methionine sulphoximine, albizziin and β-aspartyl hydroxamate.

The present invention also provides a method, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector

comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; iii) a selection vector comprising a third recombinant oligonucleotide having a sequence which encodes a selectable gene product; and iv) a eukaryotic parent cell line; b) introducing the expression vector, the amplification vector and the selection vector into the cell line to generate transformed cells; c) introducing the transformed cells into a first aqueous solution, the first aqueous solution requiring the expression of the selectable gene product for growth of the transformed cells; d) identifying at least one individual clone of transformed cells capable of growth in the first aqueous solution; e) introducing the individual clone capable of growth in the first aqueous solution into a second aqueous solution, the second aqueous solution comprising an inhibitor capable of inhibiting the first inhibitable enzyme, wherein the concentration of the inhibitor present in the first aqueous solution is sufficient to prevent growth of the parent cell line; and f) identifying at least one individual clone capable of growth in the second aqueous solution, wherein the clone capable of growth contains an amplified number of copies of the expression vector and an amplified number of copies of the amplification vector. The method of the present invention is not limited by the nature of the means employed to introduce the vectors into the parent cell line; in a preferred embodiment, the vectors are introduced into the parent cell line by electroporation. The selection, expression and/or amplification vectors are preferably linearized prior to introduction into the parent cell line.

The present invention is not limited by the nature of the cell line chosen as the parent cell line; a variety of mammalian cell lines may be employed including CHO cell lines and variants thereof, mouse L cells and BW5147 cells and variants thereof. The chosen cell line grow in either an attachment-dependent or attachment-independent manner. In a preferred embodiment, the parent cell line is a T lymphoid cell line; a particularly preferred T lymphoid cell line is the BW5147.G.1.4 cell line.

In a preferred embodiment, the parent cell line contains an endogenous gene encoding a second inhibitable enzyme; the first and second inhibitable enzymes may be, but are not required to be, the same. In another embodiment, the second inhibitable enzyme is selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase.

In a preferred embodiment, 10 to 15 micrograms of the selection vector, 20 to 30 micrograms of the amplification vector and a total of 400 to 500 micrograms of the expression vector(s) are introduced into the parent cell line.

In another preferred embodiment, the concentration of inhibitor present in the second aqueous solution is four-fold to six-fold the concentration required to prevent the growth of the parent cell line. In a preferred embodiment, the method further comprises the steps of: g) introducing the transformed cell capable of growth in the second aqueous solution into a third aqueous solution, the third aqueous solution comprising the inhibitor capable of inhibiting the first inhibitable enzyme and wherein the concentration of the inhibitor present in the third aqueous solution is sixteen-fold to thirty-six-fold the concentration of the inhibitor required to prevent the growth of the parent cell line; and h) identifying at least one transformed cell capable of growth in the third aqueous solution.

The method is not limited by the selection vector employed. In a preferred embodiment, the selection vector encodes an active enzyme selected from the group comprising hypoxanthine guanine phosphoribosyltransferase, hygromycin G phosphotransferase, xanthine-guanine phosphoribosyltransferase and aminoglycoside 3' phosphotransferase.

The method is not limited by the nature of the amplification vector employed. In a preferred embodiment, the amplification vector encodes an active enzyme selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase. In another preferred embodiment the inhibitor is selected from the group consisting of methotrexate, 2'-deoxycoformycin, methionine sulphoximine, albizziin and β-aspartyl hydroxamate.

The present invention provides a method, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and iii) a T lymphoid parent cell line; b) introducing the expression vector and the amplification vector into the T lymphoid parent cell line to generate transformed cells; c) introducing the transformed cells into a first aqueous solution, the first aqueous solution comprising an inhibitor capable of inhibiting the first inhibitable enzyme, wherein the concentration of the inhibitor present in the first aqueous solution is sufficient to prevent the growth of the parent cell line; and d) identifying a transformed cell capable of growth in the first aqueous solution, wherein the transformed cell capable of growth contains an amplified number of copies of the expression vector and an amplified number of copies of the amplification vector. The method of the present invention

is not limited by the nature of the means employed to introduce the vectors into the parent cell line: in a preferred embodiment, the vectors are introduced into the parent cell line by electroporation. The method of the present invention is not limited by the T lymphoid parent cell line employed. In a preferred embodiment, the T lymphoid cell line is the BW5147.G.1.4 cell line.

In a preferred embodiment, the parent cell line contains an endogenous gene encoding a second inhibitable enzyme; the first and second inhibitable enzymes may be, but are not required to be, the same. In another embodiment, the second inhibitable enzyme is selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase.

In another preferred embodiment, the concentration of inhibitor present in the first aqueous solution is four-fold to six-fold the concentration required to prevent the growth of the parent cell line. In a preferred embodiment, the method further comprises the steps of: e) introducing the transformed cell capable of growth in the first aqueous solution into a third aqueous solution, the third aqueous solution comprising the inhibitor capable of inhibiting the first inhibitable enzyme and wherein the concentration of the inhibitor present in the third aqueous solution is sixteen-fold to thirty-six-fold the concentration of the inhibitor required to prevent the growth of the parent cell line; and f) identifying at least one transformed cell capable of growth in the third aqueous solution.

In a preferred embodiment, the method further comprises providing a selection vector encoding a selectable gene product which is introduced into the parent cell line together with the expression vector and the amplification vector. The method is not limited by the selection vector employed. In a preferred embodiment, the selection vector encodes an active enzyme selected from the group comprising hypoxanthine guanine phosphoribosyltransferase, hygromycin G phosphotransferase, xanthine-guanine phosphoribosyltransferase and aminoglycoside 3' phosphotransferase.

In another preferred embodiment, the method further comprises, following the introduction of the expression, amplification and selection vectors, the additional step of introducing the transformed cells into a second aqueous solution, the second aqueous solution requiring the expression of the selectable gene product for growth of the transformed cells, prior to introducing the transformed cells into the first aqueous solution.

The method is not limited by the nature of the amplification vector employed. In a preferred embodiment, the amplification vector encodes an active enzyme selected from the

group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase. In another preferred embodiment the inhibitor is selected from the group consisting of methotrexate, 2'-deoxycoformycin, methionine sulphoximine, albizziin and β-aspartyl hydroxamate.

In one preferred embodiment, 20 to 30 micrograms of the amplification vector and a total of 400 to 500 micrograms of the expression vector(s) are introduced into the parent cell line. In another preferred embodiment, 10 to 15 micrograms of the selection vector, 20 to 30 micrograms of the amplification vector and a total of 400 to 500 micrograms of the expression vector(s) are introduced into the parent cell line.

The present invention further provides a method, comprising: a) providing: i) a vector comprising a first recombinant oligonucleotide (*i.e.*, polynucleotide) having a sequence encoding the amino acid sequence of a protein of interest and a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and ii) a T lymphoid parent cell line; b) introducing the vector into the T lymphoid parent cell line to generate transformed cells; c) introducing the transformed cells into a first aqueous solution, the first aqueous solution comprising an inhibitor capable of inhibiting the first inhibitable enzyme, wherein the concentration of the inhibitor present in the first aqueous solution is sufficient to prevent the growth of the parent cell line; and d) identifying a transformed cell capable of growth in the first aqueous solution, wherein the transformed cell capable of growth contains an amplified number of copies of the vector. It is not necessary that the cell containing an amplified number of copies of the vector contain amplified copies of the entire vector. Rather, the transformed cell need only have amplified those portions of the vector necessary for the expression of the amplificable marker (*i.e.*, sequences encoding the first inhibitable enzyme operably linked to a heterologous promoter) and the gene(s) of interest (*i.e.*, gene(s) encoding the protein(s) of interest). The vector may comprise more than one recombinant oligonucleotide encoding a protein of interest and the vector may further comprise a selectable marker.

The method of the present invention is not limited by the nature of the means employed to introduce the vector into the parent cell line; in a preferred embodiment, the vector is introduced into the parent cell line by electroporation. The method of the present invention is not limited by the T lymphoid parent cell line employed. In a preferred embodiment, the T lymphoid cell line is the BW5147.G.1.4 cell line.

In a preferred embodiment, the parent cell line contains an endogenous gene encoding a second inhibitable enzyme; the first and second inhibitable enzymes may be, but are not required to be, the same. In another embodiment, the second inhibitable enzyme is selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase.

In another preferred embodiment, the concentration of inhibitor present in the first aqueous solution is four-fold to six-fold the concentration required to prevent the growth of the parent cell line. In a preferred embodiment, the method further comprises the steps of: e) introducing the transformed cell capable of growth in the first aqueous solution into a second aqueous solution, the second aqueous solution comprising the inhibitor capable of inhibiting the first inhibitable enzyme and wherein the concentration of the inhibitor present in the second aqueous solution is sixteen-fold to thirty-six-fold the concentration of the inhibitor required to prevent the growth of the parent cell line; and f) identifying at least one transformed cell capable of growth in the second aqueous solution.

In a preferred embodiment, the method further comprises providing a selection vector encoding a selectable gene product which is introduced into the parent cell line together with the vector comprising the first and second recombinant oligonucleotides. The method is not limited by the selection vector employed. In a preferred embodiment, the selection vector encodes an active enzyme selected from the group comprising hypoxanthine guanine phosphoribosyltransferase, hygromycin G phosphotransferase, xanthine-guanine phosphoribosyltransferase and aminoglycoside 3' phosphotransferase.

In another preferred embodiment, the method further comprises, following the introduction of the vector comprising the first and second recombinant oligonucleotides and the selection vector, the additional step of introducing the transformed cells into a third aqueous solution, the third aqueous solution requiring the expression of the selectable gene product for growth of the transformed cells, prior to introducing the transformed cells into the first aqueous solution.

The method is not limited by the nature of the second recombinant oligonucleotide employed. In a preferred embodiment, the second recombinant oligonucleotide encodes an active enzyme selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase. In another preferred embodiment the inhibitor is selected from the group consisting of methotrexate, 2'-deoxycoformycin, methionine sulphoximine, albizziin and β-aspartyl hydroxamate.

In one preferred embodiment, 5 to 30 micrograms of the vector comprising the first and second recombinant oligonucleotides are introduced into the parent cell line. In another preferred embodiment, 10 to 15 micrograms of the selection vector, and 5 to 30 micrograms of the vector comprising the first and second recombinant oligonucleotides are introduced into the parent cell line. In another prefered embodiment, the vector comprising the first and second recombinant oligonucleotides is linearized prior to introduction into the parent cell line. When employed, the selection vector may also be linearized prior to introduction into the parent cell line.

In another embodiment, the selectable marker and the gene(s) of interest (*i.e.*, the first recombinant oligonucleotide) are contained on a single vector and 400-500 μg of this vector are employed in conjunction with 5-20 μg of an amplification vector.

The present invention also provides a method of producing a vaccine for treatment of B-cell lymphoma comprising: a) providing: i) malignant cells isolated from a patient having a B-cell lymphoma; ii) an amplification vector comprising a recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; iii) a eukaryotic parent cell line; b) isolating from the malignant cells nucleotide sequences encoding at least one V_H region and at least one V_L region, the V_H and V_L regions derived from immunoglobulin molecules expressed by the malignant cells; c) inserting the nucleotide sequences encoding the V_H and V_L regions into at least one expression vector; d) introducing the expression vector(s) and the amplification vector into the parent cell to generate one or more transformed cells; e) introducing the transformed cell(s) into a first aqueous solution containing an inhibitor capable of inhibiting the inhibitable enzyme wherein the concentration of the inhibitor present in the first aqueous solution is sufficient to prevent growth of the parent cell line; and f) identifying a transformed cell capable of growth in the first aqueous solution, wherein the transformed cell(s) capable of growth expresses the V_H and V_L regions. In a preferred embodiment, the transformed cell capable of growth in the first aqueous solution contains an amplified number of copies of the expression vector(s) and an amplified number of copies of the amplification vector.

In another preferred embodiment, the nucleotide sequences encoding the V_H and C_L regions comprise at least two V_H and at least two C_L regions (in this manner, a multivalent vaccine is produced). In another embodiment, the immunoglobulin comprising the V_H and V_L regions derived from immunoglobulin molecules expressed by the malignant cells that is

produced by the transformed (and amplified) cells is isolated (*e.g.*, purified) from the cell culture medium.

The method of the present invention is not limited by the nature of the means employed to introduce the vectors into the parent cell line. The art is well aware of numerous methods which allow the introduction of exogenous DNA sequences into mammalian cells, including but not limited to electroporation, microinjection, lipofection, protoplast fusion, liposome fusion and the like. In a preferred embodiment, the vectors are introduced into the parent cell line by electroporation.

The present invention is not limited by the nature of the cell line chosen as the parent cell line; a variety of mammalian cell lines may be employed including CHO cell lines and variants thereof, mouse L cells and BW5147 cells and variants thereof. The chosen cell line grow in either an attachment-dependent or attachment-independent manner. In a preferred embodiment. the parent cell line is a T lymphoid cell line; a particularly preferred T lymphoid cell line is the BW5147.G.1.4 cell line.

In another embodiment, the method of the present invention employs a parent cell line which contains an endogenous gene encoding a second inhibitable enzyme (*e.g.*, the genome of the parent cell line contains an endogenous gene comprising a coding region encoding a second inhibitable enzyme which is operably linked to the promoter naturally linked to this coding region (*i.e.*, the endogenous promoter for this gene). A contrast is made between the input or exogenous recombinant sequences encoding the first inhibitable enzyme and an endogenous gene encoding an inhibitable enzyme. The endogenous gene sequences will be expressed under the control of the endogenous promoter. Typically, the amplification vector will comprise a sequence encoding an inhibitable enzyme operably linked to a heterologous (*i.e.*, not the endogenous) promoter. The sequences encoding the first and the second inhibitable enzyme may encode the same or a different enzyme. Furthermore, when the same enzyme is encoded by the two sequences (*i.e.*, the recombinant and the endogenous sequences). these sequences may be derived from the same or a different source (*i.e.*, the recombinant sequence may encode an enzyme isolated from a mouse cell and may introduced into a mouse cell line which contains an endogenous gene encoding the same enzyme; alternatively. the recombinant sequence may encode an enzyme derived from a different species than that of the parent cell line (*e.g.*, the recombinant sequence may encode a rat DHFR and may be introduced into a parent mouse cell line which expresses the mouse

DHFR). The amplifiable gene (or marker) and the selectable marker may be present on the same vector: alternatively, they may be present on two separate vectors.

In one embodiment the second inhibitable enzyme expressed by the parent cell line is selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase, asparagine synthetase.

In another embodiment, the method of the present invention the concentration of inhibitor present in the first aqueous solution (*e.g.*, tissue culture medium) used to allow identification of the transformed cell(s) containing amplified copies of the amplification vector and amplified copies of the expression vector(s) is four-fold to six-fold the concentration required to prevent the growth of the parent cell line. It is well understood by those skilled in the art that only those sequences present on the amplification vector and expression vector(s) which are required for the expression of the inhibitable enzyme and the protein(s) of interest, respectively, need to be amplified. However, it is also well understood that any vector backbone sequences linked to the sequences required for expression of the inhibitable enzyme or protein(s) of interest may also be amplified (and typically are) during the co-amplification process.

In still another embodiment, the method of the present invention further comprises providing a selection vector encoding a selectable gene product (*i.e.*, a selectable marker) which is introduced into the parent cell line together with the expression vector and the amplification vector (alternatively, the selectable marker may be present on the same vector which contains the amplifiable marker). The invention is not limited by the nature of the selectable gene product employed. The selectable gene product employed may be a dominant selectable marker including but not limited to hygromycin G phosphotransferase (*e.g.*, the *hyg* gene product), xanthine-guanine phosphoribosyltransferase (*e.g.*, the *gpt* gene product) and aminoglycoside 3' phosphotransferase (*e.g.*, the *neo* gene product). Alternatively, the selectable marker employed may require the use of a parent cell line which lacks the enzymatic activity encoded by the selectable marker such as hypoxanthine guanine phosphoribosyltransferase, thymidine kinase or carbamoyl-phosphate synthetase-aspartate transcarbamoylase-dihydrooratase. In a particularly preferred embodiment, the selection vector encodes an active hypoxanthine guanine phosphoribosyltransferase. When the selection vector encodes an active hypoxanthine guanine phosphoribosyltransferase, the second aqueous solution which requires the expression of this selectable gene product comprises hypoxanthine

and azaserine. The selection, expression and/or amplification vectors are preferably linearized prior to introduction into the parent cell line.

In another embodiment, the method of the present invention further comprises following the introduction of the vectors (*i.e.*, the amplification, expression and selection vectors), the additional step of introducing the transformed cell into a second aqueous solution which requires the expression of the selectable gene product prior to introducing the transformed cell into a first aqueous solution containing an inhibitor capable of inhibiting the inhibitable enzyme.

The method of the present invention is not limited by the nature of the inhibitable enzyme encoded by the amplification vector; the art is well of aware of numerous amplifiable markers. In a preferred embodiment, the amplification vector encodes an active enzyme selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase, asparagine synthetase.

In another preferred embodiment, the inhibitor used to select for a transformed cell expressing the inhibitable enzyme encoded by the amplification vector is selected from the group consisting of methotrexate, 2'-deoxycoformycin, methionine sulphoximine, albizziin and β-aspartyl hydroxamate.

The present invention further provides a method of treating B-cell lymphoma, comprising: a) providing: i) a subject having a B-cell lymphoma; ii) a multivalent vaccine comprising at least two recombinant variable regions of immunoglobulin molecules derived from the subjects's B-cell lymphoma cells, wherein the cells express at least two different immunoglobulin molecules, the immunoglobulin molecules differing by at least one idiotope; b) administering the multivalent vaccine to the subject. In a preferred embodiment, the vaccine comprises at least two recombinant immunoglobulin molecules comprising the recombinant variable regions derived from the lymphoma cells. In a preferred embodiment, the method employs a multivalent vaccine which further comprises an adjuvant. When the vaccine is to be administered to a human subject, adjuvants approved for use in humans are employed (*e.g.*, alum). In a preferred embodiment, the adjuvant is Syntex™ adjuvant formulation 1 (SAF-1). Additional adjuvants suitable for use in humans, include but are not limited to QS21 (Aquila Biopharmaceuticals, Inc., Worcester, MA), Optivax™ (Vaxcell Inc., Norcross, GA) and adjuvants manufactured by Ribi Immunochem Research, Inc. (Hamilton, MT). The recombinant Ig proteins (including fragments of Ig proteins) which comprise the multivalent vaccine may be conjugated to a carrier protein such as KLH.

The present invention provides a method of treating B-cell lymphoma, comprising: a) providing: i) a subject having a B-cell lymphoma; ii) a multivalent vaccine comprising at least two recombinant variable regions of immunoglobulin molecules derived from the subjects's B-cell lymphoma cells, wherein the cells express at least two different immunoglobulin molecules, the immunoglobulin molecules differing by at least one idiotope; and iii) dendritic cells isolated from the subject; b) incubating the dendritic cells *in vitro* with the multivalent vaccine to produce autologous antigen-pulsed dendritic cells; c) administering intravenously the pulsed dendritic cells to the subject; and d) following the administration of the pulsed dendritic cells, administering the multivalent vaccine to the subject. In a preferred embodiment, the vaccine comprises at least two recombinant immunoglobulin molecules comprising the recombinant variable regions.

The present invention further provides a method of treating B-cell lymphoma, comprising: a) providing: i) a subject having a B-cell lymphoma; ii) a vaccine produced according to the methods of the present invention; and b) administering the vaccine to the subject.

Still further, the present invention provides a method of treating a subject having an immune cell tumor, comprising: a) providing: i) immune cell tumor cells isolated from a subject, the tumor cells expressing an idiotype protein on the cell membrane; ii) an amplification vector comprising a first recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; iii) a eukaryotic parent cell line; b) isolating nucleotide sequences encoding at least one idiotype protein expressed on the surface of the tumor cells; c) inserting the nucleotide sequences encoding the idiotype protein(s) into at least one vector to produce at least one expression vector capable of expressing the idiotype protein(s); d) introducing the expression vector(s) into the parent cell to generate one or more transformed cells; e) introducing the transformed cell into a first aqueous solution containing an inhibitor capable of inhibiting the inhibitable enzyme, wherein the concentration of the inhibitor present in the first aqueous solution is sufficient to prevent growth of the parent cell line; f) identifying a transformed cell capable of growth in the first aqueous solution, wherein the transformed cell capable of growth contains an amplified number of copies of the expression vector and an amplified number of copies of the amplification vector and wherein the transformed cell produces the idiotype protein(s) encoded by the expression vector(s); g) isolating the idiotype protein(s) produced by the transformed cell; and h) administering the isolated idiotype protein(s) to the subject.

The method of the present invention is not limited by the nature of the tumor cells. In one embodiment, the tumor cells are T lymphoid cells and the idiotype protein is a T cell receptor or fragment thereof. In another embodiment, the tumor cells are B lymphoid cells and the idiotype protein is an immunoglobulin or fragment thereof.

According to one aspect of the present invention, there is provided a multivalent vaccine comprising at least two recombinant variable regions of immunoglobulin molecules derived from B-cell lymphoma cells, wherein said at least two variable regions are from immunoglobulin molecules that differ by at least one idiotope.

According to another aspect of the present invention, there is provided a cell of a T lymphoid cell line comprising a first amplified recombinant oligonucleotide having a sequence encoding an inhibitable enzyme operably linked to a heterologous promoter, and a second amplified recombinant oligonucleotide having a sequence encoding an amino acid sequence of a protein of interest.

According to still another aspect of the present invention, there is provided a method of co-amplifying recombinant oligonucleotides, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and iii) a eukaryotic parent cell line; b) introducing 400 to 500 micrograms of said expression vector and 20 to 30 micrograms of said amplification vector into said parent cell line to generate transformed cells;

c) introducing said transformed cells into a first aqueous solution containing an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent growth of said parent cell line; and d) identifying a transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

According to yet another aspect of the present invention, there is provided a method of co-amplifying recombinant oligonucleotides, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and iii) a T lymphoid parent cell line; b) treating said expression vector and said amplification vector with a restriction enzyme to create a linearized expression vector and a linearized amplification vector; c) introducing said linearized expression vector and said linearized amplification vector into said parent cell line to generate transformed cells, wherein a ratio ranging from 10:1 to 15:1 of said expression vector to said amplification vector is employed; d) introducing said transformed cells into a first aqueous solution containing an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent growth of said parent cell line; and e) identifying a transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable

of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

According to a further aspect of the present invention, there is provided a method of co-amplifying recombinant oligonucleotides, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and iii) a T lymphoid parent cell line; b) introducing said expression vector and said amplification vector into said parent cell line to generate transformed cells; c) introducing said transformed cells into a first aqueous solution, said first aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme and wherein the concentration of said inhibitor present in said first aqueous solution is four-fold to six-fold the concentration of said inhibitor required to prevent the growth of said parent cell line; and d) identifying at least one transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

According to yet a further aspect of the present invention, there is provided a method of co-amplifying recombinant oligonucleotides, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector

comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; iii) a selection vector comprising a third recombinant oligonucleotide having a sequence which encodes a selectable gene product; and iv) a T lymphoid cell line containing an endogenous gene encoding a second inhibitable enzyme; b) introducing said expression vector, said amplification vector and said selection vector into said cell line to generate transformed cells, wherein a ratio ranging from 20:2:1 to 25:2:1 of said expression vector to said amplification vector to said selection vector is employed; c) introducing said transformed cells into a first aqueous solution, said first aqueous solution requiring the expression of said selectable gene product for growth of said transformed cells; d) identifying at least one transformed cell capable of growth in said first aqueous solution; e) introducing said transformed cell capable of growth into said first aqueous solution in a second aqueous solution, said second aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent growth of said parent cell line; and f) identifying at least one transformed cell capable of growth in said second aqueous solution, wherein said transformed cell capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

According to still a further aspect of the present invention, there is provided a method of co-amplifying recombinant oligonucleotides, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector

74667-99

comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; iii) a selection vector comprising a third recombinant oligonucleotide having a sequence which encodes a selectable gene product; and iv) a T lymphoid cell line containing an endogenous gene encoding a second inhibitable enzyme; b) introducing 400-500 micrograms of said expression vector, 20-30 micrograms of said amplification vector and 10-15 micrograms of said selection vector into said cell line to generate transformed cells; c) introducing said transformed cells into a first aqueous solution, said first aqueous solution requiring the expression of said selectable gene product for growth of said transformed cells;
d) identifying at least one individual clone of transformed cells capable of growth in said first aqueous solution;
e) introducing said individual clone capable of growth in said first aqueous solution into a second aqueous solution, said second aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent growth of said parent cell line; and f) identifying at least one individual clone capable of growth in said second aqueous solution, wherein said clone capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

According to another aspect of the present invention, there is provided a method of co-amplifying recombinant oligonucleotides, comprising: a) providing: i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest; ii) an amplification vector comprising a second recombinant oligonucleotide having a

sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and iii) a T lymphoid parent cell line; b) introducing said expression vector and said amplification vector into said T lymphoid parent cell line to generate transformed cells, wherein a ratio of at least 10:1 of said expression vector to said amplification vector is employed; c) introducing said transformed cells into a first aqueous solution, said first aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent the growth of said parent cell line; and d) identifying a transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

According to yet another aspect of the present invention, there is provided a method of co-amplifying recombinant oligonucleotides, comprising: a) providing: i) a vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest and a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and ii) a T lymphoid parent cell line; b) introducing said vector into said T lymphoid parent cell line to generate transformed cells; c) introducing said transformed cells into a first aqueous solution, said first aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent the growth of said parent cell line; and d) identifying a transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable

of growth contains an amplified number of copies of said vector.

According to another aspect of the present invention, there is provided a method of producing a vaccine for treatment of B-cell lymphoma, comprising: a) providing: i) malignant cells isolated from a patient having a B-cell lymphoma; ii) an amplification vector comprising a recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; iii) a eukaryotic parent cell line; b) isolating from said malignant cells nucleotide sequences encoding at least one V_H region and at least one V_L region, said V_H and V_L regions derived from immunoglobulin molecules expressed by said malignant cells; c) inserting said nucleotide sequences encoding said V_H and V_L regions into at least one expression vector; d) introducing said at least one expression vector and said amplification vector into said parent cell to generate transformed cells; e) introducing said transformed cells into a first aqueous solution, said first aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent growth of said parent cell line; and f) identifying a transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable of growth expresses said V_H and V_L regions.

According to still another aspect of the present invention, there is provided a use of a multivalent vaccine comprising at least two recombinant variable regions of immunoglobulin molecules derived from a subject's B-cell lymphoma cells, wherein said at least two variable regions are from immunoglobulin molecules that differ by at least one idiotope, for treating B-cell lymphoma.

74667-99

According to yet another aspect of the present invention, there is provided the use of a multivalent vaccine comprising at least two recombinant variable regions of immunoglobulin molecules derived from a subject's B-cell lymphoma cells, wherein said at least two variable regions are from immunoglobulin molecules that differ by at least one idiotope, in the manufacture of a medicament for the treatment of said subject's B-cell lymphoma.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the map of the expression vector pSSD5. Selected restriction enzyme sites are indicated.

Figure 2 shows the map of the expression vector pSSD7. Selected restriction enzyme sites are indicated.

Figure 3 shows the map of the expression vector pSRαSD5. Selected restriction enzyme sites are indicated.

Figure 4 shows the map of the expression vector pSRαSD7. Selected restriction enzyme sites are indicated.

Figure 5 shows the map of the expression vector pMSD5. Selected restriction enzyme sites are indicated.

Figure 6 shows the map of the expression vector pMSD7. Selected restriction enzyme sites are indicated.

Figure 7 shows the map of the expression vector pHEF1αASD5. Selected restriction enzyme sites are indicated.

Figure 8 shows the map of the expression vector pHEF1αASD7. Selected restriction enzyme sites are indicated.

Figure 9 shows the map of the expression vector pHEF1αBSD5. Selected restriction enzyme sites are indicated.

74667-99

Figure 10 shows the map of the expression vector pHEF1αBSD7. Selected restriction enzyme sites are indicated.

Figure 11 shows the map of the expression vector pMSD5-HPRT. Selected restriction enzyme sites are indicated.

Figure 12 shows the map of the expression vector pSSD7-DHFR. Selected restriction enzyme sites are indicated.

Figure 13 shows the map of the expression vector pJFE 14. Selected restriction enzyme sites are indicated.

Figure 14 shows the map of the expression vector pJFE 14ΔIL10. Selected restriction enzyme sites are indicated.

Figure 15 shows the map of the expression vector pSRαSD-DRα-DAF. Selected restriction enzyme sites are indicated.

Figure 16 shows the map of the expression vector pSRαSD-DRβ1-DAF. Selected restriction enzyme sites are indicated.

Figure 17 is a histogram showing the clone 5 cells selected for growth in hypoxanthine and azaserine stained with the L243 monoclonal antibody.

Figure 18 is a histogram showing the clone 5 cells selected for growth in 80 nM MTX stained with the L243 monoclonal antibody.

Figure 19 is a histogram showing the clone 5 cells selected for growth in 320 nM MTX stained with the L243 monoclonal antibody.

Figure 20 is a histogram showing the clone 5 cells selected for growth in 1 μM MTX stained with the L243 monoclonal antibody.

Figure 21 shows the map of the expression vector pSRαSD9. Selected restriction enzyme sites are indicated.

Figure 22 shows the map of the expression vector pSRαSD9CG3C. Selected restriction enzyme sites are indicated.

Figure 23 shows the map of the expression vector pSRαSD9CG4C. Selected restriction enzyme sites are indicated.

Figure 24 shows the map of the expression vector pSRαSDCKC. Selected restriction enzyme sites are indicated.

Figure 25 shows the map of the expression vector pSRαSDCL2C. Selected restriction enzyme sites are indicated.

Figure 26 shows the map of the selection and amplification vector pM-HPRT-SSD9-DHFR. Selected restriction enzyme sites are indicated.

DEFINITIONS

To facilitate understanding of the invention, a number of terms are defined below.

The term "recombinant DNA molecule" as used herein refers to a DNA molecule which is comprised of segments of DNA joined together by means of molecular biological techniques.

The terms "in operable combination" or "operably linked" as used herein refers to the linkage of nucleic acid sequences in such a manner that a nucleic acid molecule capable of directing the synthesis of a desired protein molecule is produced. When a promoter sequence is operably linked to sequences encoding a protein, the promoter directs the expression of mRNA which can be translated to produce a functional form of the encoded protein. The term also refers to the linkage of amino acid sequences in such a manner that a functional protein is produced.

DNA molecules are said to have "5' ends" and "3' ends" because mononucleotides are reacted to make oligonucleotides in a manner such that the 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor in one direction via a phosphodiester linkage. Therefore, an end of an oligonucleotides is referred to as the "5' end" if its 5' phosphate is not linked to the 3' oxygen of a mononucleotide pentose ring and as the "3' end" if its 3' oxygen is not linked to a 5' phosphate of a subsequent mononucleotide pentose ring. As used herein, a nucleic acid sequence, even if internal to a larger oligonucleotide, also may be said to have 5' and 3' ends. In either a linear or circular DNA molecule, discrete elements are referred to as being "upstream" or 5' of the "downstream" or 3' elements. This terminology reflects the fact that transcription proceeds in a 5' to 3' fashion along the DNA strand. The promoter and enhancer elements which direct transcription of a linked gene are generally located 5' or upstream of the coding region (enhancer elements can exert their effect even when located 3' of the promoter element and the coding region). Transcription termination and polyadenylation signals are located 3' or downstream of the coding region.

The term "an oligonucleotide having a nucleotide sequence encoding a gene" means a DNA sequence comprising the coding region of a gene or, in other words, the DNA sequence which encodes a gene product. The coding region may be present in either a cDNA or genomic DNA form. Suitable control elements such as enhancers/promoters, splice junctions, polyadenylation signals, etc. may be placed in close proximity to the coding region of the gene if needed to permit proper initiation of transcription and/or correct processing of the primary RNA transcript. Alternatively, the coding region utilized in the expression vectors of the present invention may contain endogenous enhancers/promoters, splice junctions, intervening sequences, polyadenylation signals, etc. or a combination of both endogenous and exogenous control elements.

The term "recombinant oligonucleotide" refers to an oligonucleotide created using molecular biological manipulations, including but not limited to, the ligation of two or more oligonucleotide sequences generated by restriction enzyme digestion of a polynucleotide sequence, the synthesis of oligonucleotides (*e.g.*, the synthesis of primers or oligonucleotides) and the like.

The term "recombinant oligonucleotide having a sequence encoding a protein operably linked to a heterologous promoter" or grammatical equivalents indicates that the coding region encoding the protein (*e.g.*, an enzyme) has been joined to a promoter which is not the promoter naturally associated with the coding region in the genome of an organism (i.e., it is linked to an exogenous promoter). The promoter which is naturally associated or linked to a coding region in the genome is referred to as the "endogenous promoter" for that coding region.

The term "transcription unit" as used herein refers to the segment of DNA between the sites of initiation and termination of transcription and the regulatory elements necessary for the efficient initiation and termination. For example, a segment of DNA comprising an enhancer/promoter, a coding region, and a termination and polyadenylation sequence comprises a transcription unit.

The term "regulatory element" as used herein refers to a genetic element which controls some aspect of the expression of nucleic acid sequences. For example, a promoter is a regulatory element which facilitates the initiation of transcription of an operably linked coding region. Other regulatory elements are splicing signals, polyadenylation signals, termination signals, etc. (defined *infra*).

The term "expression vector" as used herein refers to a recombinant DNA molecule containing a desired coding sequence and appropriate nucleic acid sequences necessary for the expression of the operably linked coding sequence in a particular host organism. Nucleic acid sequences necessary for expression in prokaryotes include a promoter, optionally an operator sequence, a ribosome binding site and possibly other sequences. Eukaryotic cells are known to utilize promoters, enhancers, and termination and polyadenylation signals.

Transcriptional control signals in eucaryotes comprise "promoter" and "enhancer" elements. Promoters and enhancers consist of short arrays of DNA sequences that interact specifically with cellular proteins involved in transcription [Maniatis, *et al.*, Science 236:1237 (1987)]. Promoter and enhancer elements have been isolated from a variety of eukaryotic sources including genes in yeast, insect and mammalian cells and viruses (analogous control

elements, *i.e.*, promoters, are also found in prokaryotes). The selection of a particular promoter and enhancer depends on what cell type is to be used to express the protein of interest. Some eukaryotic promoters and enhancers have a broad host range while others are functional in a limited subset of cell types [for review see Voss, *et al.*, Trends Biochem. Sci., 11:287 (1986) and Maniatis, *et a*l., *supra* (1987)]. For example, the SV40 early gene enhancer is very active in a wide variety of cell types from many mammalian species and has been widely used for the expression of proteins in mammalian cells [Dijkema, *et al.*, EMBO J. 4:761 (1985)]. Two other examples of promoter/enhancer elements active in a broad range of mammalian cell types are those from the human elongation factor 1α gene [Uetsuki *et al.*, J. Biol. Chem., 264:5791 (1989); Kim *et al.*, Gene 91:217 (1990); and Mizushima and Nagata, Nuc. Acids. Res., 18:5322 (1990)] and the long terminal repeats of the Rous sarcoma virus [Gorman *et al.*, Proc. Natl. Acad. Sci. USA 79:6777 (1982)] and the human cytomegalovirus [Boshart *et al.*, Cell 41:521 (1985)].

The term "promoter/enhancer" denotes a segment of DNA which contains sequences capable of providing both promoter and enhancer functions (for example, the long terminal repeats of retroviruses contain both promoter and enhancer functions). The enhancer/promoter may be "endogenous" or "exogenous" or "heterologous." An endogenous enhancer/promoter is one which is naturally linked with a given gene in the genome. An exogenous (heterologous) enhancer/promoter is one which is placed in juxtaposition to a gene by means of genetic manipulation (*i.e.*, molecular biological techniques).

The presence of "splicing signals" on an expression vector often results in higher levels of expression of the recombinant transcript. Splicing signals mediate the removal of introns from the primary RNA transcript and consist of a splice donor and acceptor site [Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York (1989) pp. 16.7-16.8]. A commonly used splice donor and acceptor site is the splice junction from the 16S RNA of SV40.

Efficient expression of recombinant DNA sequences in eukaryotic cells requires signals directing the efficient termination and polyadenylation of the resulting transcript. Transcription termination signals are generally found downstream of the polyadenylation signal and are a few hundred nucleotides in length. The term "poly A site" or "poly A sequence" as used herein denotes a DNA sequence which directs both the termination and polyadenylation of the nascent RNA transcript. Efficient polyadenylation of the recombinant transcript is desirable as transcripts lacking a poly A tail are unstable and are rapidly

degraded. The poly A signal utilized in an expression vector may be "heterologous" or "endogenous." An endogenous poly A signal is one that is found naturally at the 3' end of the coding region of a given gene in the genome. A heterologous poly A signal is one which is isolated from one gene and placed 3' of another gene. A commonly used heterologous poly A signal is the SV40 poly A signal. The SV40 poly A signal is contained on a 237 bp *Bam*HI/*Bcl*I restriction fragment and directs both termination and polyadenylation [Sambrook, *supra*, at 16.6-16.7]. This 237 bp fragment is contained within a 671 bp *Bam*HI/*Pst*I restriction fragment.

The term "stable transfection" or "stably transfected" refers to the introduction and integration of foreign DNA into the genome of the transfected cell. The term "stable transfectant" refers to a cell which has stably integrated foreign or exogenous DNA into the genomic DNA of the transfected cell.

The terms "selectable marker" or "selectable gene product" as used herein refer to the use of a gene which encodes an enzymatic activity that confers resistance to an antibiotic or drug upon the cell in which the selectable marker is expressed. Selectable markers may be "dominant": a dominant selectable marker encodes an enzymatic activity which can be detected in any mammalian cell line. Examples of dominant selectable markers include the bacterial aminoglycoside 3' phosphotransferase gene (also referred to as the *neo* gene) which confers resistance to the drug G418 in mammalian cells, the bacterial hygromycin G phosphotransferase (*hyg*) gene which confers resistance to the antibiotic hygromycin and the bacterial xanthine-guanine phosphoribosyl transferase gene (also referred to as the *gpt* gene) which confers the ability to grow in the presence of mycophenolic acid. Other selectable markers are not dominant in that their use must be in conjunction with a cell line that lacks the relevant enzyme activity. Examples of non-dominant selectable markers include the thymidine kinase (*tk*) gene which is used in conjunction with TK$^-$ cell lines, the carbamoyl-phosphate synthetase-aspartate transcarbamoylase-dihydroorotase (CAD) gene which is used in conjunction with CAD-deficient cells and the mammalian hypoxanthine-guanine phosphoribosyl transferase (*hprt*) gene which is used in conjunction with HPRT$^-$ cell lines. A review of the use of selectable markers in mammalian cell lines is provided in Sambrook *et*

al., *supra* at pp.16.9-16.15. It is noted that some selectable markers can be amplified and therefore can be used as amplifiable markers (*e.g.*, the CAD gene).

The term "amplification" or "gene amplification" as used herein refers to a process by which specific DNA sequences are disproportionately replicated such that the amplified gene becomes present in a higher copy number than was initially present in the genome. Gene amplification occurs naturally during development in particular genes such as the amplification of ribosomal genes in amphibian oocytes. Gene amplification may be induced by treating cultured cells with drugs. An example of drug-induced amplification is the methotrexate-induced amplification of the endogenous *dhfr* gene in mammalian cells [Schmike *et al.* (1978) Science 202:1051]. Selection of cells by growth in the presence of a drug (*e.g.*, an inhibitor of an inhibitable enzyme) may result in the amplification of either the endogenous gene encoding the gene product required for growth in the presence of the drug or by amplification of exogenous (*i.e.*, input) sequences encoding this gene product, or both.

The term "co-amplification" as used herein refers to the introduction into a single cell of an amplifiable marker in conjunction with other gene sequences (comprising one or more non-selectable genes such as those contained within an expression vector) and the application of appropriate selective pressure such that the cell amplifies both the amplifiable marker and the other, non-selectable gene sequences. The amplifiable marker may be physically linked to the other gene sequences or alternatively two separate pieces of DNA, one containing the amplifiable marker and the other containing the non-selectable marker, may be introduced into the same cell.

The term "amplifiable marker," "amplifiable gene" or "amplification vector" is used herein to refer to a gene or a vector encoding a gene which permits the amplification of that gene under appropriate growth conditions. Vectors encoding the dihydrofolate reductase (*dhfr*) gene can be introduced into appropriate cell lines (typically a *dhfr*$^-$ cell) and grown in the presence of increasing concentrations of the DHFR inhibitor methotrexate to select for cells which have amplified the *dhfr* gene. The adenosine deaminase (*ada*) gene has been used in analogous fashion to allow the amplification of *ada* gene sequences in cells selected for growth in the presence of ADA inhibitors such as 2'-deoxycoformycin. Examples of other genes which can be used as amplifiable markers in mammalian cells include the CAD gene (inhibitor: N-phosphonoacetyl-L-aspartic acid), the ornithine decarboxylase gene (inhibitor: difluoromethylornithine in medium lacking putrescine), and the asparagine synthetase gene

(inhibitors: albizziin or β-aspartyl hydroxamate in asparagine-free medium) [*see* Kaufman, Methods in Enzymol., 185:537 (1990) for a review].

The term "gene of interest" as used herein refers to the gene inserted into the polylinker of an expression vector whose expression in the cell is desired for the purpose of performing further studies on the transfected cell. The gene of interest may encode any protein whose expression is desired in the transfected cell at high levels. The gene of interest is not limited to the examples provided herein; the gene of interest may include cell surface proteins, secreted proteins, ion channels, cytoplasmic proteins, nuclear proteins (*e.g.*, regulatory proteins), mitochondrial proteins, etc.

The terms "nucleic acid molecule encoding," "DNA sequence encoding," and "DNA encoding" refer to the order or sequence of deoxyribonucleotides along a strand of deoxyribonucleic acid. The order of these deoxyribonucleotides determines the order of amino acids along the polypeptide (protein) chain. The DNA sequence thus codes for the amino acid sequence.

The vertebrate hematopoietic system comprises cells of the lymphoid and myeloid lineages. The myeloid lineage (or myeloid-erythroid lineage) gives rise to erythrocytes, basophils, neutrophils, macrophages, eosinophils and platelets. The lymphoid lineage gives rise to B lymphocytes, including plasma cells, and T lymphocytes.

The term "lymphoid" when used in reference to a cell line or a cell, means that the cell line or cell is derived from the lymphoid lineage and includes cells of both the B and the T lymphocyte lineages.

The terms "T lymphocyte" and "T cell"as used herein encompass any cell within the T lymphocyte lineage from T cell precursors (including Thy1 positive cells which have not rearranged the T cell receptor genes) to mature T cells (*i.e.*, single positive for either CD4 or CD8, surface TCR positive cells).

The terms "B lymphocyte" and "B cell" encompasses any cell within the B cell lineage from B cell precursors, such as pre-B cells ($B220^+$ cells which have begun to rearrange Ig heavy chain genes), to mature B cells and plasma cells. "Myeloma" cells or cell lines are malignant plasma cells or cell lines (and are thus in the B cell lineage, not the T cell lineage).

The terms "parent cell line" or "parental cell line" refers to a cell line prior to the addition of exogenous nucleic acid.

The term "transformed cells" refers to cells which contain exogenous DNA (*i.e.*, heterologous DNA introduced into the cells such as the introduction of an expression vector). Terms "transformed cell" and "transfected cell" are used herein interchangeably.

The term"aqueous solution" when used in reference to a solution used to grow a cell line refers to a solution containing compounds required to support the growth of the cells and may contain salts, buffering agents, serum or synthetic serum replacements. An aqueous solution capable of supporting the growth of a cell line is also referred to as "tissue culture medium" (*e.g.*, EMEM, DMEM, RMPI 1470, etc.).

An "aqueous solution which requires the expression of a selectable gene product" is a solution or tissue culture medium which forces a cell line to express a function or active form of the selectable gene product in order for the cells to survive in this medium (*e.g.*, the cell must express a functional HPRT when grown in medium containing hypoxanthine and azaserine). "Aqueous solutions which contain an inhibitor capable of inhibiting an inhibitable enzyme" expressed by a cell refers to medium containing an inhibitor (*e.g.*, methotrexate) which is capable of inhibiting an inhibitable enzyme (e.g., DHFR). The presence of the inhibitor in the medium requires the cell to express a functional or active form of the enzyme which is inhibited by the inhibitor in order to survive.

The "concentration of an inhibitor sufficient to prevent the growth of the parent cell line" refers to that concentration of inhibitor which must be present in the medium to achieve the killing of greater than 98% of the cells within 3 to 5 days after plating the parent cells in medium containing the inhibitor.

The term "amplified number of copies of a vector" refers to a cell line which has incorporated an exogenous or recombinant vector and has increased the number of copies of the vector present in the cell by virtue of the process of gene amplification.

The term "amplified gene" refers to a gene present in multiple copies in a cell line by virtue of gene amplification.

A cell which contains an "endogenous gene encoding an inhibitable enzyme" refers to cell which naturally (as opposed to by virtue of recombinant DNA manipulations) contains in its genomic DNA a gene encoding an inhibitable enzyme; the coding region of this gene will be operably linked to and under the control of its endogenous promoter.

The term "active enzyme" refers to an enzyme which is functional (*i.e.*, capable of carrying out the enzymatic function).

Immunoglobulin molecules consist of heavy (H) and light (L) chains, which comprise highly specific variable regions at their amino termini. The variable (V) regions of the H (V_H) and L (V_L) chains combine to form the unique antigen recognition or antigen combining site of the immunoglobulin (Ig) protein. The variable regions of an Ig molecule contain determinants (*i.e.*, molecular shapes) that can be recognized as antigens or idiotypes.

The term "idiotype" refers to the set of antigenic or epitopic determinants (*i.e.*, idiotopes) of an immunoglobulin V domain (*i.e.*, the antigen combining site formed by the association of the complementarity determining regions or V_H and V_L regions).

The term "idiotope" refers to a single idiotypic epitope located along a portion of the V region of an immunoglobulin molecule.

The term "anti-idiotypic antibody" or grammatical equivalents refers to an antibody directed against a set of idiotopes on the V region of an Ig protein.

A "multivalent vaccine" when used in reference to a vaccine comprising an idiotypic protein or fragment thereof (*e.g.*, immunoglobulin molecules or variable regions thereof, T cell receptor proteins or variable regions thereof) refers to a vaccine which contains at least two idiotypic proteins which differ by at least one idiotope. For example, a vaccine which contains two or more immunoglobulin molecules derived from a B-cell lymphoma where the immunoglobulin molecules differ from one another by at least one idiotope (*e.g.*, these immunoglobulins are somatic variants of one another) is a multivalent vaccine.

As used herein "recombinant variable regions of immunoglobulin molecules" refers to variable regions of Ig molecules which are produced by molecular biological means. As shown herein, the variable domain of the heavy and light chains may be molecularly cloned from lymphoma cells and expressed in a host cell (*e.g.*, by insertion into an expression vector followed by transfer of the expression vector into a host cell); variable domains expressed in this manner are recombinant variable regions of immunoglobulin molecules. The recombinant variable regions of immunoglobulin molecules may be expressed as an immunoglobulin molecule comprising the recombinant variable regions operably linked to the appropriate constant region (*i.e.*, C_H or C_L) (the constant region may comprise the constant region naturally associated with the recombinant variable region, as a Fab, $F(ab')_2$ or Fab' fragment comprising the variable domain of the heavy and light chains, the constant region of the light chain and a portion of the constant region of the heavy chain (the Fab, $F(ab')_2$ or Fab' fragments may be created by digestion of a recombinant immunoglobulin molecule or

alternatively, they may be produced by molecular biological means), or alternatively, as a single chain antibody or Fv protein.

"Single-chain antibodies" or "Fv" consist of an antibody light chain variable domain or region ("V_L") and heavy chain variable region ("V_H") connected by a short peptide linker. The peptide linker allows the structure to assume a conformation which is capable of binding to antigen [Bird *et al.*, (1988) *Science* 242:423 and Huston *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:5879].

A "recombinant variable region derived from a lymphoma cell" refers to a variable region which is molecularly cloned from RNA isolated from a lymphoma cell. The recombinant variable domain may be expressed as an entire immunoglobulin molecule or may be expressed as a fragment of an immunoglobulin molecule, including Fv molecules.

An "immune-enhancing cytokine" is a cytokine that is capable of enhancing the immune response when the cytokine is generated in situ or is administered to a mammalian host. Immune-enhancing cytokine include, but are not limited to, granulocyte-macrophage colony stimulating factor (G-CSF), interleukin-2 (IL-2), interleukin-3 (IL-3), interleukin-4 (IL-4) and interleukin-12 (IL-12).

An "adjuvant" is a compound which enhances or stimulates the immune response when administered with an antigen(s).

"Malignant cells isolated from a patient having a B-cell lymphoma" refers to the malignant or pathogenic B-cells found within the solid tumors characteristic of lymphoma (*e.g.*, lymph nodes and spleen containing the tumor cells).

DESCRIPTION OF THE INVENTION

The invention provides vectors and improved methods for the expression and co-amplification of genes encoding recombinant proteins in cultured cells. The description is divided into the following sections: I) Overview of Co-Amplification Methods; II) Expression Vectors; III) Amplification Vectors; IV) Selection Vectors; V) Cell Lines and Cell Culture; VI) Co-Transfection of Cell Lines; VII) Selection and Co-Amplification; VIII) Co-Amplification Without Prior Selection; IX) High-Level Expression of Interleukin 10 in Amplified Cell Lines; X) High-Level Expression of Human Class II MHC Antigens and T Cell Receptor Proteins in Amplified Cell Lines; and XI) Production of Custom Multivalent Vaccines For the Treatment of Lymphoma and Leukemia.

I. Overview Of Co-Amplification Methods

The present invention provides improved methods for the co-amplification of selectable and non-selectable genes in eukaryotic cell lines. The present invention allows, for the first time, the co-amplification of recombinant gene sequences in T lymphoid cell lines (*e.g.*, the BW5147.G.1.4 cell line).

The ability to amplify gene sequences in lymphoid cell lines (T or B lymphoid lines) is desirable for a number of reasons. These include the ability of these cells to secrete recombinant immunoglobulins and the ability to grow these suspension cell lines at high biomass in fermentators. To date amplification of input gene sequences has been reported only in B lymphoid cell lines (e.g., myeloma cell lines). Further, the ability to amplify genes in myeloma cell lines using the *dhfr* gene as the amplifiable marker have been problematic due to the endogenous DHFR activity in the myeloma cells. Successful amplification is reported to require the use of a MTX-resistant *dhfr* gene and the use of very high levels of MTX [Dorai and Moore (1987) J. Immunol. 139:4232]. In contrast, the present invention does not require the use of a MTX-resistant *dhfr* gene and permits the amplification of genes in T lymphoid cell lines.

A co-amplification scheme employing the glutamine synthetase (GS) gene has been described [U.S. Patent No. 5,122,464,

and Bebbington, *et al.* (1992) Bio/Technology 10:169]. This co-amplification scheme was developed in part to circumvent the need to use very high levels of MTX and a MTX-resistant dhfr gene to achieve co-amplification of genes in myeloma cells. The use of GS in co-amplification schemes has several drawbacks. First, the propensity of the endogenous GS locus in each cell line to be used must be examined to preclude the use of cell lines in which the endogenous GS locus will amplify at a frequency which makes the GS gene usable. Of four myeloma or hybridoma cell lines, examined, two of the four (50%) were found to be unsuitable host cells for the use of GS as a selectable marker (Bebbington, *et al., supra*). One of these two unsuitable cell lines, SP2/0, was found to amplify the endogenous GS locus.

A second drawback to the use of GS as a selectable and amplifiable marker is the amount of time required for the isolation of cell lines producing high levels of the non-selected gene product. A single round of amplification and recloning was reported to require 3 months using a myeloma cell line subjected to GS selection (Bebbington, *et al., supra*). Other selectable markers used in co-amplification protocols have been reported to require even

longer periods of time; selection of amplified myeloma cell lines using DHFR as the selectable marker takes up to 6 months [Dorai and Moore (1987) J. Immunol. 139:4232].

The present invention provides methods which allow the isolation of the desired amplified cell lines in a shorter period of time than permitted using existing co-amplification protocols. Multiple rounds of amplification can be achieved using the present invention in a period of about 3 months. The savings in time is realized, in part, by the use of cell lines which have rapid doubling times as the host cell line. In addition to shortening the period required for the generation of the desired amplified cell line, the present methods generate with high frequency amplified cell lines which have co-amplified the non-selectable gene(s) of interest as well as the amplifiable gene (*e.g.*, the *dhfr* gene).

In general the present invention involves the following steps:

1. Introduction of linearized plasmids comprising an expression vector(s) encoding a protein of interest, an amplification vector encoding an amplifiable marker (*e.g.*, the *dhfr* gene) and, optionally, a selection vector encoding a selectable marker (*e.g.*, HPRT) into a host cell line. Alternatively, a single vector comprising one or more genes of interest, an amplifiable marker and optionally a selectable marker may be employed. The host cell line will have a doubling time of 12 hours or less; a particularly preferred host cell line is the BW5147.G.1.4 cell line. The host cell prior to the introduction of the linearized vectors is referred to as the parental cell line. A preferred means of introducing the vector DNA into the host cell line is electroporation. The ratio of the amplification vector, non-selectable expression vector(s) and/or selection vector is important. A ratio of 1 (selectable vector): 2 (amplification vector): 20-25 [expression vector(s)] is employed. If a selectable marker is not employed a ratio of 1 (amplification vector): 10-15 [expression vector(s)] is used. The use of this ratio in conjunction with the electroporation of linearized vector DNA produces random concatemers of the transfected DNA vectors which contain a low percentage of the amplifiable gene. While not limiting the invention to any particular mechanism, it is believed that these random concatemers containing a low percentage of the amplifiable gene are less likely to generate an amplification unit composed primarily of the amplifiable marker. It is desirable to produce an amplification unit which contains primarily the expression vector(s) as this results in an amplified cell line which is expressing large quantities of the protein(s) of interest.

In contrast to existing transfection methods (including electroporation protocols), the methods of the present invention employ large quantities of DNA comprising the gene(s) of

interest (*i.e.*, the expression vector) [for a discussion of current electroporation methods see Ausubel *et al.*, *Current Protocols in Molecular Biology* (1995) John Wiley & Sons, Inc., at 9.3.1 to 9.3.6]. Using the methods of the present invention, a total of about 500 to 750 µg of DNA comprising the expression vector(s), the amplification vector and if employed, the selection vector in a total volume of 0.5 ml are introduced into approximately 2×10^7 cells in 0.5 ml of the electroporation buffer (final density of DNA is therefore 1 to 1.5 mg/ml). The use of large quantities of the expression vectors increases the frequency with which clones of cells expressing the gene products encoded by the exogenous DNA are isolated. Using the methods of the present invention about 20 to 25% of the selectants (or primary amplificants if no selection vector is employed) express the genes of interest at relatively high levels. In contrast, using conventional amounts of DNA (about 20 to 40 µg when introducing a single expression vector into the cells), only 1 to 5% of the selectants isolated express the gene of interest at relatively high levels.

2. When a selection vector is employed, the transfected cells are allowed to recover by growth in their normal growth medium for a short period (about 36 to 48 hours) and then they are placed in medium which requires the cells to express the selectable marker in order to survive (selective medium). The use of the selective medium facilitates the identification of cells which have taken up the transfected DNA. Colonies of cells which grow in the selective medium (selectants) are expanded and examined for the ability to express the protein of interest. Selectant clones which express the protein(s) of interest at high levels are then subjected to the amplification process.

3. Selectant clones expressing the protein(s) of interest at high levels are examined to determine their level of sensitivity to the inhibitor which inhibits the enzyme encoded by the amplifiable vector. The sensitivity of the parental cell line to the inhibitor is also determined. Selectants which survive growth in medium containing up to a 6-fold higher concentration (typically 4- to 6-fold higher) of the inhibitor than that required to kill the parental cell line are selected for further manipulation (the first round amplificants). [Any primary transfectant which has clearly taken up a transfected amplification vector (*e.g.*, one encoding DHFR) is suitable for continuation with the amplification protocols of the present invention. The presence of the transfected amplification vector is indicated by the ability of the primary transfectant to grow in medium containing the inhibitor at a level which is above the level required to kill the parental cell line.] The first round amplificants are examined for the expression of the protein(s) of interest. Cells which express low levels of the protein of

interest are discarded (as this indicates a lack of co-ordinate amplification of the amplifiable gene and the gene(s) of interest). Selectants which are capable of growing in medium containing greater than 6-fold the concentration of inhibitor which prevents the growth of the parental cell line are discarded. It has been found that selectants which are resistant to extremely high levels of the inhibitor typically do not yield amplified cell lines which express high quantities of the protein of interest. While not limiting the present invention to any particular mechanism, it is thought that resistance to extremely high levels of inhibitor at the first round of amplification is indicative of a cell line in which the amplifiable gene sequences readily separate away from the majority of the other input DNA sequences (*e.g.*, the expression vector) resulting the amplification of an amplified unit comprising primarily the amplifiable gene sequences.

4. The first round amplificants which are capable of growing in medium containing 4-fold to 6-fold higher concentrations of the inhibitor than that required to kill the parental cell line are grown in medium containing this level of inhibitor for 2 to 3 weeks. The cells are then grown in medium containing about 4- to 6-fold more of the inhibitor (*i.e.*, 16- to 36-fold the concentration which kills the parental cells) to generate the second round amplificants. The level of expression of the protein(s) of interest are examined in the second round amplificants; any clones which do not show an increase in expression of the protein(s) of interest which corresponds with the increased resistance to the inhibitor are discarded.

5. The amplified cell lines are subjected to subsequent rounds of amplification by increasing the level of inhibitor in the medium 4- to 6-fold for each additional round of amplification. At each round of amplification, the expression of the protein(s) of interest is examined. Typically any discordance between the level of resistance to the inhibitor and the level of expression of the protein(s) if interest is seen on the second round of amplification. Using the methods of the present invention more than 60% of the first round amplificants will co-amplify the gene(s) of interest and the amplifiable gene in the second round of amplification. All clones which co-amplified the gene(s) of interest and the amplifiable gene in the second round of amplification have been found to continue to coordinately amplify these gene sequences in all subsequent rounds of amplification until a maximum expression level was reached.

The following provides additional details regarding the various steps and components employed in the co-amplification protocols of the present invention.

II. Expression Vectors

The expression vectors of the invention comprise a number of genetic elements: A) a plasmid backbone; B) regulatory elements which permit the efficient expression of genes in eukaryotic cells -- these include enhancer/promoter elements, poly A signals and splice junctions; C) polylinkers which allow for the easy insertion of a gene (a selectable marker gene, an amplifiable marker gene or a gene of interest) into the expression vector; and D) constructs showing the possible combination of the genetic elements. These genetic elements may be present on the expression vector in a number of configurations and combinations.

A. Plasmid Backbone

The expression vector contains plasmid sequences which allow for the propagation and selection of the vector in procaryotic cells; these plasmid sequences are referred to as the plasmid backbone of the vector. While not intending to limit the invention to a particular plasmid, the following plasmids are preferred. The pUC series of plasmids and their derivatives which contain a bacterial origin of replication (the pMB1 replicon) and the β-lactamase or ampicillin resistance gene. The pUC plasmids, such as pUC18 (ATCC 37253) and pUC19 (ATCC 37254), are particularly preferred as they are expressed at high copy number (500-700) in bacterial hosts. pBR322 and its derivatives which contain the pMB1 replicon and genes which confer ampicillin and tetracycline resistance. pBR322 is expressed at 15-20 copies per bacterial cell. pUC and pBR322 plasmids are commercially available from a number of sources (for example, Gibco BRL, Gaithersburg, MD).

B. Regulatory Elements

i) Enhancer/Promoters

The transcription of each cDNA is directed by genetic elements which allow for high levels of transcription in the host cell. Each cDNA is under the transcriptional control of a promoter and/or enhancer. Promoters and enhancers are short arrays of DNA which direct the transcription of a linked gene. While not intending to limit the invention to the use of any particular promoters and/or enhancer elements, the following are preferred promoter/enhancer elements as they direct high levels of expression of operably linked genes in a wide variety of cell types. The SV40 and SRα enhancer/promoters are particularly preferred when the vector is to be transfected into a host cell which expresses the SV40 T antigen as these enhancer/promoter sequences contain the SV40 origin of replication.

a) The SV40 enhancer/promoter is very active in a wide variety of cell types from many mammalian species [Dijkema, R. *et al.*, EMBO J., 4:761 (1985)].

b) The SRα enhancer/promoter comprises the R-U5 sequences from the LTR of the human T-cell leukemia virus-1 (HTLV-1) and sequences from the SV40 enhancer/promoter [Takebe, Y. *et al.*, Mol. Cell. Biol., 8:466 (1988)]. The HTLV-1 sequences are placed immediately downstream of the SV40 early promoter. These HTLV-1 sequences are located downstream of the transcriptional start site and are present as 5' nontranslated regions on the RNA transcript. The addition of the HTLV-1 sequences increases expression from the SV40 enhancer/promoter.

c) The human cytomegalovirus (CMV) major immediate early gene (IE) enhancer/promoter is active in a broad range of cell types [Boshart *et al.*, Cell 41:521 (1985)]. The 293 cell line (ATCC CRL 1573) [J. Gen. Virol., 36:59 (1977), Virology 77:319 (1977) and Virology 86:10 (1978)], an adenovirus transformed human embryonic kidney cell line, is particularly advantageous as a host cell line for vectors containing the CMV enhancer/promoter as the adenovirus IE gene products increase the level of transcription from the CMV enhancer/promoter.

d) The enhancer/promoter from the LTR of the Moloney leukemia virus is a strong promoter and is active in a broad range of cell types [Laimins *et al.*, Proc. Natl. Acad. Sci. USA 79:6453 (1984)].

e) The enhancer/promoter from the human elongation factor 1α gene is abundantly transcribed in a very broad range of cell types [Uetsuki *et al.*, J. Biol. Chem., 264:5791 (1989) and Mizushima and Nagata, Nuc. Acids. Res. 18:5322 (1990)].

ii) Poly A Elements

The cDNA coding region is followed by a polyadenylation (poly A) element. The preferred poly A elements of the present invention are strong signals that result in efficient termination of transcription and polyadenylation of the RNA transcript. A preferred heterologous poly A element is the SV40 poly A signal (See SEQ ID NO:3). Another preferred heterologous poly A element is the poly A signal from the human elongation factor

1α (hEF1α) gene. (*See* SEQ ID NO:41). The invention is not limited by the poly A element utilized. The inserted cDNA may utilize its own endogenous poly A element provided that the endogenous element is capable of efficient termination and polyadenylation.

iii) Splice Junctions

The expression vectors also contain a splice junction sequence. Splicing signals mediate the removal of introns from the primary RNA transcript and consist of a splice donor and acceptor site. The presence of splicing signals on an expression vector often results in higher levels of expression of the recombinant transcript. A preferred splice junction is the splice junction from the 16S RNA of SV40. Another preferred splice junction is the splice junction from the hEF1α gene. The invention is not limited by the use of a particular splice junction. The splice donor and acceptor site from any intron-containing gene may be utilized.

C. Polylinkers

The expression vectors contain a polylinker which allows for the easy insertion of DNA segments into the vector. A polylinker is a short synthetic DNA fragment which contains the recognition site for numerous restriction endonucleases. Any desired set of restriction sites may be utilized in a polylinker. Two preferred polylinker sequences are the SD5 and SD7 polylinker sequences. The SD5 polylinker is formed by the SD5A (SEQ ID NO:1) and SD5B (SEQ ID NO:2) oligonucleotides and contains the recognition sites for *Xba*I, *Not*I, *Sfi*I, *Sac*II and *Eco*RI. The SD7 polylinker is formed by the SD7A (SEQ ID NO:4) and SD7B (SEQ ID NO:5) oligonucleotides and contains the following restriction sites: *Xba*I, *Eco*RI, *Mlu*I, *Stu*I, *Sac*II, *Sfi*I, *Not*I, *BssH*II and *Sph*I. The polylinker sequence is located downstream of the enhancer/promoter and splice junction sequences and upstream of the poly A sequence. Insertion of a cDNA or other coding region (*i.e.*, a gene of interest) into the polylinker allows for the transcription of the inserted coding region from the enhancer/promoter and the polyadenylation of the resulting RNA transcript.

D. Constructs

The above elements may be arranged in numerous combinations and configurations to create the expression vectors of the invention. The genetic elements are manipulated using standard techniques of molecular biology known to those skilled in the art [Sambrook, J. *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory

Press, New York (1989)]. Once a suitable recombinant DNA vector has been constructed, the vector is introduced into the desired host cell. DNA molecules are transfected into procaryotic hosts using standard protocols. Briefly the host cells are made competent by treatment with calcium chloride solutions (competent bacteria cells are commercially available and are easily made in the laboratory). This treatment permits the uptake of DNA by the bacterial cell. Another means of introducing DNA into bacterial cells is electroporation in which an electrical pulse is used to permit the uptake of DNA by bacterial cells.

Following the introduction of DNA into a host cell, selective pressure may be applied to isolate those cells which have taken up the DNA. Procaryotic vectors (plasmids) will contain an antibiotic-resistance gene, such as ampicillin, kanamycin or tetracycline resistance genes. The preferred pUC plasmids contain the ampicillin resistance gene. Growth in the presence of the appropriate antibiotic indicates the presence of the vector DNA.

For analysis to confirm correct sequences in the plasmids constructed, the ligation mixture may be used to transform suitable strains of *E. coli*. Examples of commonly used *E. coli* strains include the HB101 strain (Gibco BRL), TG1 and TG2 (derivatives of the JM101 strain), DH10B strain (Gibco BRL) or K12 strain 294 (ATCC No. 31446). Plasmids from the transformants are prepared, analyzed by digestion with restriction endonucleases and/or sequenced by the method of Messing *et al.*, Nuc. Acids Res., 9:309 (1981).

Plasmid DNA is purified from bacterial lysates by chromatography on Qiagen Plasmid Kit columns (Qiagen, Chatsworth, CA) according to the manufacturer's directions for large scale preparation.

Small scale preparation (*i.e.*, minipreps) of plasmid DNA is performed by alkaline lysis [Birnboim, H.C. and Doly, J. , Nuc. Acids. Res., 7:1513 (1979)]. Briefly, bacteria harboring a plasmid is grown in the presence of the appropriate antibiotic (for pUC-based plasmids ampicillin is used at 60 µg/ml) overnight at 37°C with shaking. 1.5 ml of the overnight culture is transferred to a 1.5 ml microcentrifuge tube. The bacteria are pelleted by centrifugation at 12,000*g* for 30 seconds in a microcentrifuge. The supernatant is removed by aspiration. The bacterial pellet is resuspended in 100 µl of ice-cold Solution I (50 mM glucose, 25 mM Tris-HCl, pH 8.0 and 10 mM EDTA pH 8.0). Two hundred µl of Solution II (0.2 N NaOH and 1% SDS) is added and the tube is inverted to mix the contents. 150 µl of ice-cold Solution III (3M sodium acetate adjusted to pH 4.8 with glacial acetic acid) is added and the tube is vortexed to mix the contents. The tube is then placed on ice for 3 to 5 minutes. The tube is then centrifuged at 12,000*g* for 5 minutes in a microcentrifuge and the

supernatant is transferred to a fresh tube. The plasmid DNA is precipitated using 2 volumes of ethanol at room temperature and incubating 2 minutes at room temperature (approximately 25°C). The DNA is pelleted by centrifugation at 12,000*g* for 5 minutes in a microcentrifuge. The supernatant is removed by aspiration and the DNA pellet is resuspended in a suitable buffer such as TE buffer (10 mM Tris-HCl, pH 7.6, 1 mM EDTA, pH 8.0).

Expression vector DNA purified by either chromatography on Qiagen columns or by the alkaline lysis miniprep method is suitable for use in transfection experiments.

III. Amplification Vectors

A vector encoding a structural gene which permits the selection of cells containing multiple or "amplified" copies of the vector encoding the structural gene is referred to as an amplification vector. The amplifiable gene is capable of responding either to an inhibitor or lack of an essential metabolite by amplification to increase the expression product (*i.e.*, the expression of the protein encoded by the amplifiable gene). The amplifiable gene may be characterized as being able to complement an auxotrophic host. For example, the gene encoding DHFR may be used as the amplifiable marker in conjunction with cells lacking the ability to express a functional DHFR enzyme. However, it is not necessary to use an auxotrophic host cell. In a preferred embodiment the host cell is not auxotrophic with respect to the amplifiable marker.

The invention is not limited by the use of a particular amplifiable gene. Various genes may be employed, such as the gene expressing DHFR, the CAD gene, genes expressing metallothioneins, the gene expressing asparagine synthetase, the gene expressing glutamine synthetase and genes expressing surface membrane proteins which offer drug resistance. By blocking a metabolic process in the cells with enzyme inhibitors, such as methotrexate, for DHFR or cytotoxic agents such as metals, with the metallothionein genes, or by maintaining a low or zero concentration of an essential metabolite, the cellular response will be amplification of the particular gene and flanking sequences [Kaufman and Sharp (1982) J. Mol. Biol. 159:601]. Because the process of gene amplification results in the amplification of the amplifiable marker and surrounding DNA sequences, it is possible to co-amplify gene sequences other than those encoding the amplifiable marker [Latt, *et al.* (1985) Mol. Cell. Biol. 5:1750]. The amplification of sequences encoding the gene of interest is accomplished by co-introducing sequences encoding the gene of interest and the amplifiable marker into the same host cell.

The gene encoding the protein of interest may be physically linked to the amplifiable marker by placing both coding regions with appropriate regulatory signals on a single vector (in addition. the gene encoding the protein of interest may be physically linked to both the amplifiable marker and the selectable marker by placing both the appropriate coding regions with appropriate regulatory signals on a single vector). However it is not necessary that both coding regions be physically located on the same vector. Because small vector molecules are easier to manipulate and give higher yields when grown in bacterial hosts, it is preferred that the gene of interest and the amplifiable marker gene be located on two separate plasmid vectors. Whether the amplifiable marker and the gene of interest are encoded on the same or separate vector plasmids, the vector molecules are linearized by digestion with a restriction enzyme prior to introduction of the vector DNAs into the host cell. The restriction enzyme utilized is selected for its ability to cut within the plasmid backbone of the vector but not cut within the regulatory signals or the coding region of the amplifiable marker or gene of interest.

The amplification vector is constructed by placing the desired structural gene encoding the amplifiable marker into an expression vector such that the regulatory elements present on the expression vector direct the expression of the product of the amplifiable gene. The invention is illustrated by the use of a structural gene encoding DHFR as the amplifiable marker. The DHFR coding sequences are placed in the polylinker region of the expression vector pSSD7 such that the DHFR coding region is under the transcriptional control of the SV40 enhancer/promoter. The invention is not limited by the selection of any particular vector for the construction of the amplification vector. Any suitable expression vector may be utilized. Particularly preferred expression vectors include pSSD5, pSSD7, pSRαSD5, pSRαSD7, pMSD5 and pMSD7. These expression vectors utilize regulatory signals which permit high level expression of inserted genes in a wide variety of cell types.

IV. Selection Vectors

An expression vector encoding a selectable marker gene is referred to as a selection vector. The selectable marker may be a dominant selectable marker. Examples of dominant selectable markers include the *neo* gene, the *hyg* gene and the *gpt* gene. The selectable marker may require the use of a host cell which lacks the ability to express the product encoded by the selectable marker. Examples of such non-dominant markers include the *tk* gene, the CAD gene and the *hprt* gene.

The invention is not limited to the use of a particular selectable marker or to the use of any selectable marker (besides the amplifiable marker) at all. In a preferred embodiment, the host cell used is a HPRT-deficient cell line and the amplifiable marker used is DHFR.

When an HPRT-deficient cell line is utilized and this cell line produces a functional DHFR enzyme, a selectable marker encoding the HPRT enzyme may be utilized. The host cell is co-transfected with plasmids containing a selectable marker (HPRT), an amplifiable marker (DHFR) and one or more proteins of interest. The transfected cells are then first selected for the ability to grow in HxAz medium (hypoxanthine and azaserine) which requires the expression of HPRT by the cell. Cells which have the ability to grow in HxAz medium have incorporated at least the selection vector encoding HPRT. Because the vector DNAs are linearized and then introduced into the host cell by electroporation (discussed below), cells which have taken up the HPRT vector are also likely to have taken up the vectors encoding DHFR and the protein of interest. This is because the linearized vectors form long concatemers or tandem arrays which integrate with a very high frequency into the host chromosomal DNA as a single unit [Toneguzzo, *et al.* (1988) Nucl. Acid Res. 16:5515].

The ability to select for transfected cells expressing HPRT facilitates the use of DHFR as the amplifiable marker in a cell line which is not DHFR-deficient. The use of the selectable marker allows the circumvention of the problem of amplification of the host cell's endogenous DHFR gene [Walls, J.D. *et al.*, (1989), *supra*]. However, as discussed below, the present invention can be practiced without using a selectable marker in addition to the amplification vector when cell lines which are not DHFR-deficient are employed.

The invention may be practiced such that no selectable marker is used. When the amplifiable marker is a dominant amplifiable marker such as the glutamine synthetase gene or where the host cell line lacks the ability to express the amplifiable marker (such as a $DHFR^-$ cell line) no selectable marker need be employed.

When employed, the selectable marker may be 1) provided on a separate vector (*i.e.*, a selection vector) or 2) may be present on a vector containing the gene(s) of interest and the amplifiable marker 3) may be provided on a vector containing the amplifiable marker (amplification/selection vector) or 4) may be provided on a vector containing the gene(s) of interest.

V. Cell Lines And Cell Culture

A variety of mammalian cell lines may be employed for the expression of recombinant proteins according to the methods of the present invention. Exemplary cell lines include CHO cell lines [*e.g.*, CHO-K1 cells (ATCC CCl 61; ATCC CRL 9618) and derivations thereof such as DHFR- CHO-KI cell lines (*e.g.*, CHO/dhFr-; ATCC CRL 9096), mouse L cells and BW5147 cells and variants thereof such as BW5147.3 (ATCC TIB 47) and BW5147.G.1.4 cells (ATCC TIB 48). The cell line employed may grow attached to a tissue culture vessel (*i.e*, attachment-dependent) or may grow in suspension (*i.e.*, attachment-independent).

BW5147.G.1.4 cells are particularly preferred for the practice of the present invention. BW5147.G.1.4 cells have a very rapid doubling time [*i.e.*, a doubling time of about 12 hours when grown in RPMI 1640 medium containing 10% Fetal Clone I™(Hyclone)]. The doubling time or generation time refers to the amount of time required for a cell line to increase the number of cells present in the culture by a factor of two. In contrast, the CHO-K1 cell line (from which the presently available dhfr- CHO-KI cell lines were derived) have a doubling time of about 21 hours when the cells were grown in either DMEM containing 10% Fetal Clone II™(Hyclone) or Ham's F-12 medium containing 10% Fetal Clone II.™

A rapid doubling time is advantageous as the more rapidly a cell line doubles, the more rapidly amplified variants of the cell line will appear and produce colonies when grown in medium which requires the expression of the amplifiable marker. Small differences (*i.e.*, 1-2 hours) in the doubling times between cell lines can translate into large difference in the amount of time required to select for a cell line having useful levels of amplification which result in a high level of expression of the non-selectable gene product. The speed with which a high expressing cell line can be isolated may be critical in certain situations. For example, the production of proteins to be used in clinical applications (*e.g.*, the production of tumor-related proteins to be used to immunize a cancer patient) requires that the protein of interest be expressed in useable quantities as quickly as possible so that maximum benefit to the patient is realized.

In addition, BW5147.G.1.4 cells permit the amplification of the non-selectable gene (which encodes the protein of interest) at a very high frequency. Using the methods of the present invention, about 80% of BW5147.G.1.4 cells which survive growth in the selective medium (*e.g.*, HxAz medium) will amplify the input DNA which contains the amplifiable marker and the DNA encoding the protein of interest (as measure by the ability of the cells to survive in medium containing MTX and the production of increased amounts of the protein of

interest). That is 80% of the cells which survive growth in the selective medium will survive growth in medium which requires the expression of the amplifiable marker. When cells are subjected to growth in medium containing a compound(s) which requires expression of the amplifiable marker (*e.g.*, growth in the presence of MTX requires the expression of DHFR), the cells which survive are said to have been subjected to a round of amplification. Following the initial or first round of amplification, the cells are placed in medium containing an increased concentration of the compounds which require expression of the amplifiable marker and the cells which survive growth in this increased concentration are said to have survived a second round of amplification. Another round of selection in medium containing yet a further increase in the concentration of the compounds which require expression of the amplifiable marker is referred to as the third round of amplification.

Of those transfected BW5147.G.1.4 clones which amplify in the first round of amplification (as measured by both the ability to grow in increased concentrations of MTX and an increased production of the protein of interest), about 2/3 also coordinately amplify the amplifiable gene as well as the gene encoding the protein of interest in the second round of amplification. All clones which coordinately amplified the amplifiable marker and the gene encoding the protein of interest in the second round of amplification have been found to coordinately amplify both genes in all subsequent rounds of amplification.

An additional advantage of using BW5147.G.1.4 cells is the fact that these cells are very hardy. A cell line is said to be hardy when it is found to be able to grow well under a variety of culture conditions and when it can withstand a certain amount of mal-treatment (*i.e.*, the ability to be revived after being allowed to remain in medium which has exhausted the buffering capacity or which has exhausted certain nutrients). Hardiness denotes that the cell line is easy to work with and it grows robustly. Those skilled in the art of tissue culture know readily that certain cell lines are more hardy than others; BW5147.G.1.4 cells are particularly hardy cells.

BW5147.G.1.4 cells may be maintained by growth in DMEM containing 10% FBS or RPMI 1640 medium containing 10% Fetal Clone I.™ CHO-K1 cells (ATCC CCl 61, ATCC CRL 9618) may be maintained in DMEM containing 10% Fetal Clone II (Hyclone), Ham's F12 medium containing 10% Fetal Clone II™or Ham's F12 medium containing 10% FBS and CHO/dhFr- cells (CRL 9096) may be maintained in Iscove's modified Dulbecco's medium containing 0.1 mM hypoxanthine, 0.01mM thymidine and 10% FBS. These cell lines are grown in a humidified atmosphere containing 5% CO_2 at a temperature of 37°C.

The invention is not limited by the choice of a particular host cell line. Any cell line can be employed in the methods of the present invention. Cell lines which have a rapid rate of growth or a low doubling time (*i.e.*, a doubling time of 15 hours or less) and which is capable of amplifying the amplifiable marker at a reasonable rate without amplification of the endogenous locus at a similar or higher rate are preferred. Cell lines which have the ability to amplify the amplifiable marker at a rate which is greater than the rate at which the endogenous locus is amplified are identified by finding that the ability of the cell to grow in increasing concentrations of the inhibitor (*i.e.*, the compound which requires the cell to express the amplifiable marker in order to survive) correlates with an increase in the copy number of the amplifiable marker (this may be measured directly by demonstrating an increase in the copy number of the amplifiable marker by Southern blotting or indirectly by demonstrating an increase in the amount of mRNA produced from the amplifiable marker by Northern blotting).

VI. Co-Transfection Of Cell Lines

Prior to introduction of vector DNA into a given cell line, the vector DNA is linearized using a restriction enzyme which cuts once within the vector sequences and which does not cut within the control or coding regions necessary for the expression of the encoded protein. Linearization of the DNA is advantageous as it promotes the integration of the vector DNA into the chromosomal DNA of the host cell line (free ends of DNA are recombinogenic). Furthermore, vector DNA must break in order to integrate into the genomic DNA of the host cell; linearization allows control over where this break occurs thereby preventing the loss of functional vector sequences by directing this break to a non-essential region of the vector. Additionally, linear DNA molecules tend to integrate into the genomic DNA of the host cell as a random head to tail concatemer (it is noted that circular DNA also tends to integrate as a head to tail concatemer; however, as discussed above, the circular DNA must break prior to integration). This obviates the need to construct a single large vector containing the selectable gene, amplifiable gene and the gene(s) of interest. Several smaller vectors may be co-transfected instead thereby essentially eliminating the likelihood that the vector will suffer a break in an essential region.

To generate a stable cell line expressing large quantities of a desired protein(s), the following vectors are introduced as linear DNA: 1) a selectable vector such as pMSD5-HPRT; 2) an amplifiable vector such as pSSD7-DHFR and 3) one or more vectors encoding a gene

of interest. This also results in a much higher ratio of copies of the expressed gene(s) of interest to amplifiable marker genes in the concatemer. The ratio of the selectable vector, amplifiable vector and the vector(s) encoding a protein(s) of interest is 1:2:20-50. Multiple vectors encoding separate proteins of interest are utilized when it is desirable to express multiple proteins in a single cell. This will be the case where the protein of interest is a multi-chain protein. For example, immunoglobulins are formed by the association of two heavy chains and two light chains; the heavy and light chains are encoded by separate genes. Expression of a functional immunoglobulin requires that the transfected cell express both the heavy and light chain genes. Up to six non-selectable/amplifiable plasmids (*i.e.*, encoding a gene of interest) may be used to transfect a given cell line.

Large quantities of the expression vector(s) are introduced into the cells along with the amplification and selection vectors. Typically 10 to 15 μg of the selectable vector (*e.g.*, pMSD5-HPRT), 20 to 30 μg of the amplification vector (*e.g.*, pSSD7-DHFR) and a total of 400 to 500 μg total of the expression vectors. For example, when two expression vectors are to be used, 200 to 250 μg of each of the two expression vectors (*i.e.*, plasmid encoding a gene of interest) are used in addition to the selection and amplification vectors. The maximum amount of DNA which can be electroporated under the conditions used herein is about 500 to 750 μg DNA (*i.e.*, the total amount or the sum of all vector DNAs). If 6 separate expression vectors are to be introduced into a cell in addition to the selection and amplification vectors, the following amounts of DNA are employed: 7.5 μg of the selection vector, 15 μg of the amplification vector and ~121 μg of each of the six expression vectors [the total amount of DNA is therefore ~750 μg per electroporation using 2 x 10^7 cells/ml in 0.5 ml of 1X HBS(EP)].

When a vector comprising the gene(s) of interest and an amplifiable marker is employed, 5-30 μg of of this type of vector are employed. When a vector comprising both an amplifiable and a selectable marker (*i.e.*, an amplification/selection vector) is employed in addition to an expression vector(s), the following amounts of DNA are used: 5-30 μg of the amplification/selection vector and a total of 400-500 μg of the expression vector(s). When a vector comprising one or more genes of interest, an amplifiable marker and a selectable marker are employed, 5-30 μg of of this type of vector are employed.

The vectors to be co-transfected into the cells are linearized using appropriate restriction enzymes (*i.e.*, enzymes which cut only within the plasmid backbone) in the same reaction tube. Following digestion with the appropriate restriction enzymes, the DNA is

precipitated using ethanol and resuspended in 0.5 ml of 1X HBS (EP) (20 mM HEPES, pH 7.0; 0.75 mM Na_2HPO_4/NaH_2PO_4, pH 7.0; 137 mM NaCl; 5 mM KCl and 1 gm/liter dextrose).

The linearized vector DNAs are preferentially introduced into the host cell by electroporation. Alternatively, the linearized vector DNAs may be introduced into the host cell by microinjection using techniques known to the art. The use of electroporation is preferred over other methods of introducing DNA into cells for a number of reasons: 1) efficiency of transfection. A number of attractive cell lines (*e.g.*, virtually any lymphoid cell line) are refractory to transformation via any other method (such as DEAE-dextran mediated transfection or calcium phosphate-DNA co-precipitation). Electroporation of these lines allows the ready isolation of as many independent transformants as might be reasonably required. 2) Electroporation preserves the integrity of the transfected DNA. DNA introduced by other methods (DEAE-dextran or $CaPO_4$) has been shown to acquire observable mutations at observable frequencies, posing a concern for therapeutically used proteins derived from these sorts of transfections [*See* for example, M.P. Calos *et al.* (1983) Proc. Natl. Acad. Sci. USA 80:3015; Kopchick and Stacey (1984) Mol. Cell. Biol. 4:240; Wake *et al.* (1984) Mol. Cell. Biol. 4:387; and Lebkowski *et al.* (1984) Mol. Cell. Biol. 4:1951]. Lebkowski *et al.*, *supra* reported a mutation frequency in DNA chemically introduced that was four orders of magnitude above the endogenous mutational frequency. In contrast, DNA introduced into cells via electroporation was found to have a mutation frequency equal to the background mutational frequency of the cell [Drinkwater and Klinedinst (1986) Proc. Natl. Acad. Sci. USA 83:3402]. 3) Cotransformation of several unlinked DNA molecules is readily achieved using electroporation. As demonstrated herein, a minimum of four unlinked DNAs can be cotransfected into cells by electroporation and a high frequency of the cells expressing the selectable marker will also express all of the other genes. 4) Electroporation is simple to perform. While microinjection of DNA avoids the increased mutation frequency observed using chemical introduction of DNA, microinjection of somatic cells is technically challenging and requires the use of expensive equipment. In contrast electroporation can be performed using fairly inexpensive equipment which may be prepared in house or purchased commercially.

Lymphoid cell lines have been very difficult to transfect with $CaPO_4$-mediated co-precipitation, although it has been achieved [Rice and Baltimore (1982) Proc. Natl. Acad. Sci. USA 79:7862 and Oi *et al.* (1983) Proc. Natl. Acad. Sci. USA 80:825]. In contrast,

transfection of numerous lymphoid cell lines has been achieved by electroporation with acceptably high transformation frequencies [Potter *et al.* (1984) Proc. Natl. Acad. Sci. USA 81: 7161; Boggs *et al.* (1986) Exp. Hematol. 14:988; Toneguzzo *et al.* (1986) Mol. Cell. Biol. 6:703 and Toneguzzo and Keating (1986) Proc. Natl. Acad. Sci. USA 83:3496]. Oi *et al.*, *supra* report a transformation frequency for BW5147 cells using $CaPO_4$-mediated co-precipitation and a gpt-expressing plasmid of 1 per 10^7 cells. Toneguzzo *et al.*, *supra* report a transformation frequency for BW5147 cells using electroporation and a gpt-expressing plasmid of 3.6 per 10^4 cells (a frequency greater than 3000-fold higher than that achieved using $CaPO_4$-mediated co-precipitation).

The host cells, typically BW5147.G.1.4 cells, are washed twice in ice-cold 1X HBS(EP) and resuspended at 2 x 10^7 cells/ml in 0.5 ml of 1X HBS(EP). The cells are then placed in a 1 ml cuvette (#67.746, Sarstedt, Inc., Princeton, NJ) which contains the linearized DNAs. The cuvette is placed on ice. The electroporation is performed at 225 volts using an ISCO Model 493 power supply (ISCO). The electroporation apparatus is constructed exactly as described in Chu, G. *et al.*, Nucl. Acids Res. 15:1311 (1987). The electroporation device is set on constant voltage (225V) at the 2X setting (*i.e.*, both capacitors are used). Alternatively, a commercially available electroporation device may be employed [*e.g.*, Gene Pulser™ (BioRad, Hercules, CA) with the Capacitance Extender set at 960 μFD]. Following electroporation, the cells are allowed to recover by incubation on ice for 5 to 15 minutes, typically 10 minutes.

VII. Selection And Co-Amplification

The electroporated cells are then transferred to a T75 flask (Falcon) containing 30 mls of RPMI 1640 medium (Irvine Scientific) supplemented with 10% fetal calf serum (FCS; HyClone) and 50 μg/ml gentamicin (Sigma). The cells are then incubated at 37° C in a humidified atmosphere containing 5% CO_2 for 36 to 48 hours. The cells are then transferred to 48 well plates (Costar) at 1 x 10^4 to 1 x 10^5 cells per well in selective medium. The use of selective medium facilitates the identification of cells which have taken up the transfected DNA. Cells which grow either in an attachment-dependent manner or an attachment-independent manner are plated in multiwell plates during growth in selective medium.

A variety of selectable markers may be used including both dominant selectable markers and markers which require the use of a cell line lacking a given enzyme. For example, cell lines lacking the enzyme HPRT can be used in conjunction with a vector

expressing the *hprt* gene. The transfected cells are then grown in the presence of hypoxanthine and azaserine (HxAz medium). Examples of dominant selectable markers which do not require the use of enzyme-deficient cell lines include the *neo* gene, the *hyg* gene and the *gpt* gene.

When pMSD5-HPRT is used as the selectable marker, the selective medium comprises RPMI 1640 medium containing 10% FCS, 100 μM hypoxanthine (Hx) (Sigma) and 2 μg/ml azaserine (Az) (Sigma). After approximately 11 days, positive wells (*i.e.*, wells containing cells capable of growth in the selective medium) are visible and the colonies are removed to 24 well plates. The positive colonies are picked from the 48 well plates from about day 11 to about 3 weeks following the addition of selective medium.

Positive colonies removed from the 48 well plates are placed into 24 well plates (Costar) in RPMI 1640 medium containing 10% dialyzed FCS (HyClone) and 100 μM Hx. The use of dialyzed serum at this point increases the speed and frequency of co-amplification of the input DNA in the transfectants. Hypoxanthine is retained in the culture medium for a few passages until the azaserine is diluted to non-toxic concentrations.

The transfected cells which survived growth in selective medium are then checked to see if they are expressing the genes of interest. This may be done by any suitable assay including cell surface staining, a bioassay for activity, ELISA or immunoprecipitation followed by polyacrylamide gel electrophoresis. For example if the gene(s) of interest encode a cell surface molecule, the transfected cells are analyzed by staining with an antibody specific for the vector-encoded cell surface molecule. The presence of the antibody on the surface of the transfected cell is detected by fluorescence microscopy (the specific antibody is either directly conjugated to a fluorochrome or a fluorescienated secondary antibody is utilized). The best expressing clones are then checked to determine their level of sensitivity to MTX. Typically 6 to 18, more preferably 12, clones are checked.

The parental (*i.e.*, non-transfected) BW5147.G.1.4 cells barely grow in the presence of 10 nM MTX. By visual inspection 3 to 5 days after plating, greater than about 98 percent of the parental BW5147.G.1.4 cells are killed when 1×10^4 cells are placed in 2 ml of medium containing 20 nM MTX in the well of a 24 well plate (this level of MTX is referred to as the growth cut off for the parental BW5147.G.1.4 cell line). At 30 nM MTX, colonies of BW5147.G.1.4 cells are seen at a frequency of less than 10^{-7}.

The transfected and selected cells ("selectants") are plated in a range of MTX concentrations ranging from 10 to 100 nM; the cells are plated at a density of 1 to 5×10^4

cells per well in a 24 well plate (Costar); the selectants are plated at the same density of cells as was used to determine the level of MTX at which > about 98% of the parental cells were killed. This is done because MTX irreversibly binds to DHFR so that the number of cells present in a given volume effects the concentration of MTX required to kill the cells; that is if a higher density of cell is used, a higher concentration of MTX will be required to kill about 98% of the cells [For example when the parental cells are plated at a density of 1 x 10^4 cells/2 ml medium in the well of a 24 well plate 20 nM MTX is sufficient to kill >98% cells in a 3 to 5 day assay. If the density is increased two-fold (1 x 10^4 cells in ml medium), 25 nM MTX is required for >98% killing. If 5 x 10^4 cells are placed in 2 ml of medium in the well of a 24 well plate, 30 nM MTX is required to achieve >98% killing.]

Clones of selectants typically show growth cut offs of 30 to 60 nM MTX (that is greater than about 98% of the selectants are killed when placed in medium containing 30 to 60 nM MTX when the plates are visually inspected 3 to 5 days after plating in medium containing this level of MTX). Cells from each selectant of interest which shows MTX resistance above the parental BW5147.G.1.4 cells (*e.g.*, above 20 to 30 nM MTX) are plated at 10^4 cells per well of a 48 well plate (Costar) in 0.5 ml of RPMI 1640 containing 10% dialyzed FCS and MTX (hereinafter medium-MTX). Several concentrations of MTX are used: 20 nM, 40 nM and 60 nM above each clones' upper level of MTX resistance (*i.e.*, if the upper level of MTX resistance is 30 nM then the following concentrations may be used: 50 nM, 70 nM and 90 nM); these levels of MTX are typically 4-fold to 6-fold the level of MTX required to kill greater than about 98% of the parental cells. Any selectants which are capable of growth in medium containing greater than 90 nM MTX are discarded; it has been observed that selectants which are capable of growing in such high levels of MTX tend to preferentially amplify the amplification vector at the expense of the expression vector(s).

After 7 to 10 days, the wells are fed with 0.5 ml medium-MTX. Initial amplificants are picked between 2 to 6 weeks (typically 3 to 5 weeks) after plating in medium-MTX. The clones are then analyzed again for expression of the gene(s) of interest using the appropriate assay (*i.e.*, staining with antibodies for cell surface expression, ELISA, bioassays for activity, immunoprecipitation and PAGE, etc.).

Typically a $HPRT^+$ clone is plated at a concentration of 50 to 80 nM MTX (this represent the first round of amplification). The clone is grown for 2 to 3 weeks and then the level of MTX is increased to 200 nM to 480 nM (a 4 fold increase; this represents the second round of amplification). After another 2 to 4 weeks, the level of MTX is increased to 1 to 2

μM MTX (another 4 to 6 fold increase; this represents the third round of amplification). Any clones which show an increased resistance to MTX without a corresponding increase in expression of the gene(s) of interest is discarded. Typically any discordance is seen on the second round of amplification; such clones prove to be unstable and are undesirable.

The methods of the present invention allow, for the first time, the co-amplification of transfected DNA sequences in BW5147 cells. In addition, the methods of the present invention provide improved methods for the co-amplification of DNA sequences in cell lines. Of the selectants that are expressing the gene(s) of interest, most (*i.e.*, greater than 80%), if not all, will co-amplify the amplifiable marker (*e.g*, the *dhfr* gene which confers resistance to MTX) and the gene(s) of interest in the first round of amplification. More than 60% of the first round amplificants will co-amplify the gene(s) of interest in the second round in addition to *dhfr* gene sequences. To date, using the methods of the present invention, no clones have been obtained that amplify the gene(s) of interest in the second round of amplification that then fail to continue to coordinately amplify in all subsequent rounds until a maximum expression level is reached. Thus, the methods of the present invention result in a much higher frequency of coordinate co-amplification of gene sequences than has been reported for other methods of co-amplification such as that reported by Walls *et al.* (1989) Gene 81:139 or by Kaufman *et al.* (1985) Mol. Cell. Biol. 5:1750 when single clones were examined. In addition to providing a means for achieving a very high frequency of coordinate co-amplification of gene sequences, the methods of the present invention also provide methods which produce the desired amplificants with a considerable time savings relative to existing methods. The method of the present invention avoids the time-consuming step of searching through pools of primary transformants which have been subjected to a round of amplification to find those few clones within the pool which are expressing the protein of interest at high levels.

The following modifications to the above-described amplification protocol have been found to decrease the amount of time required for the first round of amplification by 2 to 3 weeks. First, the original transfectants are selected by growth in RPMI 1640 medium containing 100 μM Hx, 2 μg/ml Az and 10% dialyzed FCS. Second, the original transfectants are fed at about 10 days following electroporation with 0.5 ml per well (in a 48 well plate) of RPMI 1640 medium containing 10% dialyzed FCS, 100 μM Hx and 10 nm MTX; this yields a final concentration in each well of the 48 well plate of 5 nM MTX. The

net effect of the growth of the transfected cells in medium containing dialyzed FCS and 5 nM MTX is to give the cells which have undergone amplification events a selective advantage.

VIII. Co-Amplification Without Prior Selection

The amplified cell lines of the present invention may be generated using only an amplification vector in addition to the expression vector(s) (*i.e.*, the use of a selection vector is not required). Cell lines which do not lack a functional gene product corresponding to the enzyme encoded by the amplification vector and which can be successfully employed without the use of a selectable marker in addition to the amplifiable marker are those cell lines in which the background level of amplification of the endogenous gene (*e.g.*, the endogenous *dhfr* gene when DHFR is used as the amplifiable marker) is low enough that amplification of the input amplifiable gene (*i.e.*, the amplification vector) occurs preferentially.

When it desired that no selection step be employed, the above protocols are modified as follows. The amplification vector and expression vector(s) are linearized and electroporated into the parental cell line using a ratio of 1:10-15 (amplification vector:expression vector). Again large amounts of DNA are introduced, preferably by electroporation, into the cells. Typically, 20 μg of the amplification vector is used and 200 to 250 μg each of two expression vectors (or 400 to 500 μg of a single expression vector). If a vector containing both the gene(s) of interest and the amplifiable marker is employed, 5-30 μg, more preferably 5-10 μg, of this type of vector is employed.

Following electroporation, the transfected cells are allowed to recover for 36 to 48 hours as described above. The transformed cells are then transferred to 48 well plates at a density of no more than 1 x 10^6 cells per well in medium containing 4-fold to 6-fold the concentration of inhibitor required to prevent the growth of the parental cells. Using the BW5147.G.1.4 cell line, the expected frequency of generating a primary transformant which contains enough copies of the input amplifiable gene to allow the isolation of a first round amplificant capable of growth in medium containing 4- to 6-fold the level of inhibitor required to prevent growth of the parental BW5147.G.1.4 cells is approximately 1 in 10^8 to 1 in 10^{10} cells. Accordingly, at least 5 x 10^8 to 1 x 10^{11} cells must be plated in medium containing elevated levels of the inhibitor to permit the isolation of several first round amplificants. Cells capable of growing in 4- to 6-fold the level of inhibitor required to prevent growth of the parental cells are examined for the ability to express the protein(s) of interest; those clones expressing high levels of the protein of interest are subjected to

subsequent rounds of amplification as described above. Any clones which do not display a coordinate increase in the level of expression of the protein(s) of interest and the level of resistance to the inhibitor at any amplification step are discarded.

The ability to generate amplified cell lines without the need to employ a selection vector reduces the amount of time required to produced the desired amplified cell line. However, the use of a selection vector and the initial selection step is advantageous particularly when working with cell lines which have a high background frequency of amplification of the endogenous locus corresponding to the amplifiable gene present on the amplification vector. Even when working with a cell line which does not a have a high background level of amplification of the endogenous gene, the use of a selection vector and an initial selection step is advantageous because it allows one to work with only the primary selectants expressing the highest levels of the gene(s) of interest. This reduces the amount of time and effort required to generate amplified cell lines expressing very high levels of the protein(s) of interest.

IX. High-Level Expression Of Interleukin 10 In Amplified Cell Lines

Using the methods of the present invention, cell lines were isolated which express large quantities of interleukin 10 (IL-10). IL-10 is a cytokine produced by TH_2 cells (type 2 helper T cells), macrophages/monocytes, and some B cells. IL-10 acts to inhibit the synthesis of cytokines by activated TH_1 cells, activated macrophages and natural killer cells [Mosmann (1993) Ann. Rev. Immunol. 11: 165 and Mosmann (1994) Advances in Immunol. 56: 1]. Studies have shown that IL-10 expression is positively correlated with graft outcome in transplantation [Bromberg (1995 Curr. Op. Immunol. 7:639]. Accordingly, there is interest in using IL-10 therapeutically. Therapeutic use of IL-10, of course, requires the ability to produce large quantities of IL-10.

Presently, there are two commercial sources of murine IL-10. Genzyme Diagnostics (Cambridge, MA) sells 5 mg of IL-10 produced in *E. coli* produced for $295.00 (cat#2488-01, ~2500 units). Biosource International (Camarillo, CA) sells 5 mg of IL-10 produced in *E. coli* for $245.00 (cat# PMC-0104, ~2500 units). The methods of the present invention were used to isolate cell line which produces about 75,000 units per milliliter of culture supernatant. Using the lower commercial price for IL-10, these cells produce about $7,350,000.00 worth of IL-10 per liter in a static culture. These amplified cell lines yield about 150 mg of IL-10 protein per liter in static culture; thus, the unpurified culture

supernatant from these amplified cell lines provides a much more pure source of IL-10 than do presently available commercial preparations.

X. High-Level Expression Of Human Class II MHC Antigens And T Cell Receptor Proteins In Amplified Cell Lines

The human class II MHC antigens, HLA-DR, and their corresponding mouse analogs, the Ia antigens, are cell surface membrane glycoproteins which mediate the recognition of non-self molecules (*i.e.*, antigens) by T lymphocytes. Class II molecules display fragments of foreign antigens on the surface of antigen presenting cells which include macrophages, dendritic cells, B lymphocytes and activated T lymphocytes. When MHC-restricted, antigen-specific T lymphocytes interact with antigen presenting cells bearing class II molecules bound to antigen, an immune response is generated.

Class II antigens comprise two chains, an α chain and a β chain. Both chains must be expressed in the same cell in order for the class II molecule to be transported to the surface of the cell. The β chain is highly polymorphic and this polymorphism results in heritable differences in immune responsiveness. In certain class II MHC molecules (*e.g.*, mouse IA, human DQ and DP), the α chain is also highly polymorphic. Given the central role that class II molecules play in the immune response, including rejection of transplanted tissue and heritable susceptibility to autoimmune disease, studies of the interaction of class II molecules with foreign antigen and with T lymphocytes have been undertaken. These studies of the physical-chemical interaction of class II molecules with antigen require the availability of large quantities of purified soluble class II molecules. In addition, the use of class II molecules complexed with specific peptides has been suggested for the treatment of autoimmune disease [Sharma, *et al.* (1991) Proc. Natl. Acad. Sci. USA 88:11465].

In order to provide such reagents, chimeric human DR molecules were expressed at high levels on the surface of amplified cell lines using the selection amplification method of the invention. The DR molecules are cleaved from the cell surface to produce soluble DR molecules by treatment with an enzyme capable of cleaving either a phosphatidylinositol linkage or a thrombin site which is present on the chimeric DR molecule.

A similar approach allows the production of large quantities of soluble T cell receptor (TCR) molecules or immunoglobulin (Ig) molecules. Like, class II molecules, TCR and Ig molecules comprise heterodimers (*i.e.*, two different chains associate to form the TCR or Ig molecule displayed on the cell surface; it is noted that both cell surface and soluble forms of

Ig molecules exist in nature and for patient immunization one would produce soluble Ig). The methods of the present invention permit the production of large quantities of soluble forms of class II MHC molecules and TCR to be produced in a rapid manner. This allowing for the production of customized tumor cell vaccines comprising soluble TCR for the treatment of lymphoma and leukemia patients as well as the production of soluble class II MHC molecules for the treatment of autoimmune disease.

XI. Production of Custom Multivalent Vaccines For The Treatment of Lymphoma and Leukemia

The existing approach toward vaccination (*i.e.*, active immunotherapy) of B-cell lymphoma and leukemia involves the production of a custom vaccine comprising autologous immunoglobulin idiotype which corresponds to the most abundant antibody molecule expressed on the surface of the B-cell tumor. An analogous approach for the treatment of T-cell lymphomas and leukemias would involve the production of a custom vaccine comprising autologous T cell receptor (TCR) idiotype which corresponds to the most abundant TCR molecule expressed on the surface of the T-cell tumor.

It is known in B-cells that the variable regions of rearranged immunoglobulin (Ig) genes accumulate point mutations following antigenic stimulation (Ig). This process, known as somatic mutation or somatic hypermutation, is responsible for affinity maturation of humoral immune responses [Tonegawa (1983) Nature 302:575, Teillaud *et al.* (1983) Science 222:721, Griffiths *et al.* (1984) Nature 312:272 and Clarke *et al.* (1985) J. Exp. Med. 161:687]. During affinity maturation, antibodies of higher affinity emerge as the immune response proceeds (*i.e.*, progression from primary to secondary to tertiary responses). A comparison of the antibody produced during the immune response reveals that mutations accumulate from the late stage of primary responses onward; these mutations cluster in the second complementarity determining region (CDR2) region of the Ig molecule (*i.e.*, within the hypervariable regions within the antigen-binding site). Somatic mutation does not occur in T cells.

Somatic variants are known to exist within the population of cells comprising certain B-cell tumors (*e.g.*, low grade or follicular B-cell lymphomas); thus, while these tumors are clonal at the level of Ig gene rearrangements (including nucleotide sequence at the V-D-J junctions) they are in fact quasi-clonal with respect to the nucleotide or amino acid sequence of their heavy chain V regions [Cleary ML *et al.* (1986), Cell 44:97 and Levy S *et al.* (1988)

J. Exp. Med. 168:475]. It is thought that following the transformation event(s) which gives rise to the B-cell tumor, somatic mutation persists. Analysis of B-cell lymphomas reveals that about 1 to 5% of the cells within the tumor contain somatic mutations.

The fact that somatic variants exist within a B-cell tumor has implications for immunotherapy of these tumors. For example, treatment of B-cell lymphoma with anti-idiotype antibodies was shown to produce an initial partial response in patients; however, idiotype variant tumor cells (idiotype negative) emerged at the original tumor site [Cleary ML *et al.* (1986), *supra*; Bahler DW and Levy R (1992) Proc. Natl. Acad. Sci. USA 89:6770; Zelenetz AD *et al.* (1992) J. Exp. Med. 176:1137; and Zhu D *et al.* (1994) Brit. J. Haematol. 86:505]. It is thought that these idiotype variant tumor cells were already present before treatment with the monoclonal anti-idiotype antibody and that they were allowed to proliferate after the selective removal of the idiotype positive tumor cells. These clinical trials showed the drawback of targeting a single epitope on the tumor cell.

In order to improve the immunotherapy of B-cell tumors, active immunization with autologous tumor derived Ig or Ig subfragments has been employed. It is hoped that the use of the Ig or Ig subfragments as an immunogen would produce a T cell response and antibodies directed against a number of different epitopes found within the tumor-specific Ig. In this type of treatment the Ig (or idiotype fragment of the Ig) of the patient's tumor cell is expressed While this approach has the advantage that multiple epitopes are targeted, it still suffers from the fact that a single Ig (or subfragment) is used as the immunogen and therefore the possibility exists that tumor cells expressing somatic variants of the predominant Ig will escape and proliferate. To produce the tumor Ig-idiotype protein used for immunization with existing methodologies, lymphoma cells removed by surgical biopsy are fused with the heterohybridoma cell line K6H6/B5 which has lost the ability to secrete endogenous Ig. Hybrid cells which secrete Ig corresponding to the immunophenotype of the tumor sample are expanded and the secreted Ig is purified for use as a vaccine [Kwak *et al.* (1992), *supra*]. This technique is referred to as "rescue fusion." The Ig produced by rescue fusion represents a single Ig derived from the patient's tumor; this Ig is presumably the predominant Ig expressed by the tumor. Thus, vaccines produced by rescue fusion are monovalent and do not represent the full complexity of Ig expressed by tumors which contain somatic variants.

Clinical trials using tumor Ig-idiotype protein produced by rescue fusion to vaccinate B-cell lymphoma patients are in progress. These trials are showing impressive clinical improvements for these tumors which remain essentially incurable with conventional therapy

(*i.e.*, chemotherapy). This custom vaccine therapy is used following a course of conventional chemotherapy (employed to reduce the tumor burden). The clinical improvements are seen when comparing patients treated solely with conventional chemotherapy with patients who received custom vaccines following chemotherapy. Among the patients who have been treated with custom vaccines and followed for a lengthy period of time (about 8 years), one has recently relapsed. Although not confirmed at this time, it is possible that this relapse is due to the outgrowth of tumor cells bearing somatic variants of the tumor Ig-idiotype protein used in the vaccine.

In addition to the failure to provide a multivalent vaccine representative of all Ig variants present in the patients tumor, the rescue fusion technique has other draw backs. This technique requires a large number of tumor cells which are obtained by surgical biopsy of enlarged lymph nodes in the patient. Some B-cell lymphoma patients do not present with enlarged lymph nodes suitable for surgical biopsy and therefore cannot be treated using vaccines produced by the rescue fusion technique. Furthermore, the production of each custom heterohybridoma cell line secreting the patients Ig takes between 2 to 8 months (average of 4 months) and is labor intensive; in laboratories having extensive experience with the rescue fusion technique, the rate of vaccine production is about 25 custom vaccines per technician per year. This rate of producing custom vaccines is not sufficient to meet the existing and growing patient demand.

Ideally, the method for producing custom tumor specific vaccines could be performed on a small number of cells (*i.e.*, from a fine needle biopsy), would produce a multivalent vaccine representative of the full complexity of the patient's tumor, would be fast (average of 2-3 months) and would be less labor intensive than existing methods such that a single technician could produce at least a hundred custom vaccines per year.

The methods described herein (Examples 9 and 10) provide a means to produce custom vaccines, including multivalent vaccines, from small numbers of cells quickly and efficiently. The ability to use a small sample size permits the production of custom vaccines for patients lacking enlarged lymph nodes suitable for surgical biopsy. In addition to expanding the pool of patients who can be treated with custom vaccines, the ability to use fine needle biopsy material obviates the need for surgery for all lymphoma patients (at least with respect to the collection of cells for the production of custom vaccines).

EXPERIMENTAL

The following examples serve to illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof.

In the experimental disclosure which follows, the following abbreviations apply: M (molar); mM (millimolar); μM (micromolar); nM (nanomolar); mol (moles); mmol (millimoles); μmol (micromoles); nmol (nanomoles); gm (grams); mg (milligrams); μg (micrograms);pg (picograms); L (liters); ml (milliliters); μl (microliters); cm (centimeters); mm (millimeters); μm (micrometers); nm (nanometers); °C (degrees Centigrade); AMP (adenosine 5'-monophosphate); cDNA (copy or complementary DNA); DNA (deoxyribonucleic acid); ssDNA (single stranded DNA); dsDNA (double stranded DNA); dNTP (deoxyribonucleotide triphosphate); RNA (ribonucleic acid); PBS (phosphate buffered saline); *g* (gravity); OD (optical density); HEPES (N-[2-Hydroxyethyl]piperazine-N-[2-ethanesulfonic acid]); HBS (HEPES buffered saline); SDS (sodium dodecylsulfate); Tris-HCl (tris[Hydroxymethyl]aminomethane-hydrochloride); Klenow (DNA polymerase I large (Klenow) fragment); rpm (revolutions per minute); EGTA (ethylene glycol-bis(ß-aminoethyl ether) N, N, N', N'-tetraacetic acid); EDTA (ethylenediaminetetracetic acid); bla (ß-lactamase or ampicillin-resistance gene); ORI (plasmid origin of replication); lacI (lac repressor); Amicon (Amicon Corp., Beverly, MA); ATCC (American Type Culture Collection, Rockville, MD); Becton Dickinson (Becton Dickinson Immunocytometry Division, San Jose CA); Costar (Costar, Cambridge, MA); Falcon (division of Becton Dickinson Labware, Lincoln Park, NJ); FMC (FMC Bioproducts, Rockland, ME); Gibco/BRL (Gibco/BRL. Grand Island , NY); HyClone (HyClone. Logan. UT); Sigma (Sigma Chemical Co., St. Louis, MO); NEB (New England Biolabs, Inc., Beverly, MA); Operon (Operon Technologies. Alameda, CA); Perkin-Elmer (Perkin-Elmer, Norwalk, CT); Pharmacia (Pharmacia Biotech, Pisacataway, NJ); Promega (Promega Corp., Madison , WI); Sarstedt (Sarstedt, Newton, NC); Stratagene (Stratagene, LaJolla, CA); U.S. Biochemicals (United States Biochemical, Cleveland, OH); and Vector (Vector Laboratories, Burlingame. CA).

EXAMPLE 1

Construction Of Expression Vectors

In order to construct the expression vectors of the invention a number of intermediate vectors were first constructed.

74667-99

Construction Of pSSD5 And pSSD7

pSSD5 and pSSD7 contain the following elements from SV40: the enhancer/promoter region, the 16S splice donor and acceptor and the poly A site. Vectors containing the SV40 enhancer/promoter sequences will replicate extrachromosomally in cell lines which express the SV40 large T antigen as the SV40 enhancer/promoter sequences contain the SV40 origin of replication.

A polylinker containing the recognition sequences for several restriction enzymes is located between the splice acceptor and poly A sequences. The polylinker allows for the easy insertion of a gene of interest. The gene of interest will be under the transcriptional control of the SV40 enhancer/promotor. pSSD5 and pSSD7 differ only in the sequences of the polylinker (sequences listed below). The polylinker of pSSD5 contains the following restriction sites: *Xba*I, *Not*I, *Sfi*I, *Sac*II and *Eco*RI. The polylinker of pSSD7 contains the following restriction sites: *Xba*I, *Eco*RI, *Mlu*I, *Stu*I, *Sac*II, *Sfi*I, *Not*I, *Bss*HII and *Sph*I.

pSSD5 was constructed by digestion of the plasmid pL1 [Okayama and Berg, Mol. Cell. Biol., 3:280 (1983)] with *Pst*I and *Hin*dIII. All restriction enzymes were obtained from New England Biolabs and were used according to the manufacturer's directions. The plasmid pcDV1 [Okayama and Berg, *supra*] was digested with *Hin*dIII and *Bam*HI. Both digests were electrophoresed on a 0.8% low melting temperature agarose gel (SeaPlaque,™ FMC). A 535 bp DNA fragment from the pL1 digest containing the SV40 enhancer/promoter and 16S splice junctions was cut out of the gel. A 2.57 kb DNA fragment from the pcDV1 digest containing the SV40 polyadenylation signals and the pBR322 backbone was cut out of the gel. The two gel slices were combined in a microcentrifuge tube and the agarose was removed by digestion with β-Agarase I (NEB) followed by isopropanol precipitation according to the manufacturer's directions. The DNA pellets were dried and resuspended in 20 μl of TE.

Two synthetic oligonucleotides (Operon), SD5A [5'-TCTAGAGCGGCCGCG GAGGCCGAATTCG-3' (SEQ ID NO:1)] and SD5B [5'-GATCCGAATTCGGCCT CCGCGGCCGCTCTAGATGCA-3' (SEQ ID NO:2)] were added in equal molar amounts to the resuspended DNA fragments. Ligation buffer (10X concentrate, NEB) was added to a 1X concentration, 80 units of T4 DNA ligase was added and the ligation was placed at 14°C overnight. Competent *E. coli* cells were transformed with the ligation mixture and a plasmid was isolated that consisted of the DNA fragments from pL1 and pcDV1 with a novel polylinker connecting the fragments. The resulting plasmid was named pSSD.

The ~ 670 bp *Bam*HI/*Pst*I fragment containing the SV40 poly A sequences (SV40 map units 2533 to 3204; SEQ ID NO:3) was removed from SV40 DNA and cloned into pUC19 digested with *Bam*HI and *Pst*I. The resulting plasmid was then digested with *Bcl*I (corresponds to SV40 map unit 2770). The ends were treated with the Klenow enzyme (NEB) and dNTPs to create blunt ends. Unphosphorylated *Pvu*II linkers (NEB) were ligated to the blunted ends and the plasmid was circularized to create pUCSSD. The SV40 poly A sequences can be removed from pUCSSD as a *Bam*HI/*Pvu*II fragment.

pSSD5 was constructed by ligating together the following three fragments: 1) the 1873 bp *Ssp*I/*Pvu*II fragment from pUC19; this provides the plasmid backbone; 2) the 562 bp fragment containing the SV40 enhancer/promoter and 16S splice junction and the polylinker from pSSD: this fragment was obtained by digestion of pSSD with *Ssp*I and partial digestion with *Bam*HI followed by isolation on low melting agarose and recovery as described above; and 3) the 245 bp *Bam*HI/*Pvu*II fragment from pUCSSD (this fragment contains the SV40 poly A sequences). The three fragments were mixed together and ligated using T4 DNA ligase (NEB) to create pSSD5. The map of pSSD5 is shown in Figure 1.

To create pSSD7, pSSD5 was digested with *Xba*I and *Eco*RI. The synthetic oligonucleotide pair SD7A and SD7B (Operon) was ligated into the cut pSSD5 thereby generating the SD7 polylinker. The sequence of SD7A is 5'-CTAGAATTC ACGCGTAGGCCTCCGCGGCCGCGCGCATGC-3' (SEQ ID NO:4). The sequence of SD7B is 5'-AATTGCATGCGCGCGGCCGCGGAGGCCTACGCGTGA ATT-3' (SEQ ID NO:5). The map of pSSD7 is shown in Figure 2.

Construction Of pSRαSD5 And pSRαSD7

pSRαSD5 and pSRαSD7 contain the SRα enhancer/promoter followed by the 16S splice junction of SV40 and either the polylinker formed by the oligonucleotide pair SD5A/SD5B or SD7A/SD7B. The polylinker is followed by the SV40 poly A sequences. A gene of interest can be inserted into the polylinker and the expression of the inserted gene will be under the control of the human SRα enhancer/promoter. The SRα enhancer/promoters a hybrid enhancer/promoter comprising human T cell leukemia virus 1 5' untranslated sequences and the SV40 enhancer [Takebe *et al.*, Mol. Cell. Biol., 8:466 (1988)]. The SRα enhancer/promoter is reported to increase expression from the SV40 enhancer/promoter by ten-fold in host cells. This enhancer/promoter is active in a broad range of cell types. Vectors containing the SRα enhancer/promoter will replicate in cells expressing SV40 large T

antigen as the SV40 origin of replication is present within the SRα enhancer/promoter sequences.

The SRα enhancer/promoter was removed from pcDL-SRα296 by digestion with *Hind*III and *Xho*I. The ~640 bp *Hind*III/*Xho*I fragment (SEQ ID NO:6) was recovered from a low melting agarose gel as described above. This ~640 bp fragment was inserted into either pSSD5 or pSSD7 digested with *Hind*III and *Xho*I (removes the SV40 enhancer/promoter from pSSD5 or pSSD7). The map of pSRαSD5 is shown in Figure 3. The map of pSRαSD7 is shown in Figure 4.

Construction Of pMSD5 And pMSD7

pMSD5 and pMSD7 contain the long terminal repeat (LTR) from the Moloney murine leukemia virus followed by the 16S splice junction of SV40 and either the polylinker formed by the oligonucleotide pair SD5A/SD5B or SD7A/SD7B. The polylinker is followed by the SV40 poly A sequences. A gene can be inserted into the polylinker and the expression of the inserted gene will be under the control of the Moloney LTR.

The Moloney LTR was removed from a plasmid containing Moloney murine leukemia viral DNA [Shinnick *et al.*, Nature 293:543 (1981)] by digestion of the plasmid with *Cla*I (corresponds to Moloney map unit 7674). The ends were made blunt by incubation with Klenow and dNTPs. Unphosphorylated *Hind*III linkers (NEB) were ligated onto the blunt ends. This treatment destroyed the *Cla*I site and replaced it with a *Hind*III site. The plasmid was then digested with *Sma*I (corresponds to Moloney map unit 8292) and unphosphorylated *Xho*I linkers were ligated onto the ends. The resulting plasmid now contains a *Xho*I site replacing the *Sma*I site at Moloney map unit 8292 and a *Hind*III site replacing the *Cla*I site at Moloney map unit 7674. The plasmid was then digested with *Xho*I and *Hind*III. The resulting *Xho*I/*Hind*III fragment containing the Moloney LTR (SEQ ID NO:7) was inserted into pSSD5 digested with *Xho*I and *Hind*III (this removes the SV40 enhancer/promoter and 16S splice junction from pSSD5) to yield pMSD5. The map of pMSD5 is shown in Figure 5.

To create pMSD7, the Moloney LTR on the *Xho*I/*Hind*III fragment was inserted into pSSD7 digested with *Xho*I and *Hind*III. The map of pMSD7 is shown in Figure 6.

74667-99

Construction Of Vectors Containing The Human Elongation Factor 1α Enhancer/Promoter

The human elongation factor 1α enhancer/promoter is abundantly transcribed in a very broad range of cell types. Vectors containing two versions of this active enhancer/promoter were constructed: 1) a long version containing ~1.45 kb of sequences located upstream of the initiation codon and continuing through the first intron to provide a splice junction and 2) a short version containing 475 bp of sequences upstream of the initiation codon. The short version of the promoter is termed the "A" version and the long version is termed the "B" version.

A. Construction Of pHEF1αASD5 And pHEF1αASD7

pHEF1αASD5 and pHEF1αASD7 contain the short version of the human elongation factor 1α enhancer/promoter [Uetsuki *et al.*, J. Biol. Chem., 264:5791 (1989) and Mizushima and Nagata. Nuc. Acids. Res., 18:5322 (1990)]. The human elongation factor 1α enhancer/promoter is abundantly transcribed in a very broad range of cell types including L929, HeLa. CHU-2 and COS cells.

The human elongation factor 1α enhancer/promoter (nucleotides 125 to 600 of the human elongation factor 1α gene; SEQ ID NO:8) was isolated from human genomic DNA as follows. Genomic DNA was isolated from the MOU cell line (GM 08605, NIGMS Human Genetic Mutant Cell Repository, Camden, NJ) using standard techniques [Sambrook *et al.*, *supra* at pp. 9.16-9.23]. The MOU cell line is an Epstein-Barr virus transformed human B cell line.

Two synthetic oligonucleotide primers (Operon) were used to prime the polymerase chain reaction (PCR) for the isolation of an ~ 475 bp fragment containing the human elongation factor 1α enhancer/promoter (SEQ ID NO:8). U.S. Patent Nos. 4,683,195, 4,683,202 and 4,965,188 cover PCR methodology.

The 5' primer, designated HEF1αL5, contains the following sequence: 5'-AAGCTTTGGAGCTAAGCCAGCAAT-3' (SEQ ID NO:9). The 3' primer, designated HEF1αL3A. contains the following sequence: 5'-CTCGAGGCGGCAA ACCCGTTGCG-3' (SEQ ID NO:10). PCR conditions were as reported in Saiki *et al.*, Science 239:487 (1988). Briefly, 10 μg MOU genomic DNA and 1 μM final concentration of each primer were used in a 400 μl PCR reaction. Reaction conditions were 94°C for 1 minute, 60°C for 1 minute, 72°C for 1.5 minutes, 30 cycles. *Taq* DNA polymerase was obtained from Perkin-Elmer.

The primer pair generates a 475 bp fragment having a *Hind*III site at the 5' end and a *Xho*I site at the 3' end. The PCR reaction products were electrophoresed on a low melting agarose gel and the 475 bp fragment was recovered as described above. The recovered fragment was digested with *Hind*III and *Xho*I and inserted into either pSSD5 or pSSD7 digested with *Hind*III and *Xho*I to yield pHEF1αASD5 and pHEF1αASD7, respectively. The maps of pHEF1αASD5 and pHEF1αASD7 are shown in Figure 7 and 8, respectively.

B. Construction Of pHEF1αBSD5 And pHEF1αBSD7

pHEF1αBSD5 and pHEF1αBSD7 were constructed as described above for pHEF1αASD5 and pHEF1αASD7 with the exception that the HEF1αL3B primer was used instead of the HEF1αL3A primer with the HEF1αL5 primer to generate a ~1.45 kb fragment containing the human elongation factor 1α enhancer/promoter and a splice donor and acceptor from the human elongation factor 1α gene. The ~1.45 kb fragment corresponds to map units 125 to 1567 in the human elongation factor 1α gene (SEQ ID NO:11). The sequence of HEF1αL3B is 5'-TCTAGAGTTTTCACG ACACCTGA-3' (SEQ ID NO:12). The HEF1αL3B primer generates a *Xba*I site at the 3' end of the ~1.45 kb fragment. This fragment was digested with *Hind*III and *Xba*I and inserted into either pSSD5 or pSSD7 digested with *Hind*III and *Xba*I to generate pHEF1αBSD5 or pHEF1αBSD7, respectively. Digestion of pSSD5 and pSSD7 with *Hind*III and *Xba*I removes the SV40 enhancer/promoter and the SV40 16S splice junction. These SV40 sequences are replaced with the human elongation factor 1α enhancer/promoter and a splice donor and acceptor from the human elongation factor 1α gene. The maps of pHEF1αBSD5 and pHEF1αBSD7 are shown in Figures 9 and 10, respectively.

EXAMPLE 2

Construction Of The Selection Vector pMSD5-HPRT

pMSD5-HPRT contains a full length cDNA clone encoding the mouse HPRT enzyme under the transcriptional control of the Moloney LTR. The Moloney LTR contains a strong enhancer/promoter which is active in a broad range of cell types [Laimins *et al.*, Proc. Natl. Acad. Sci. USA 79:6453 (1984)]. The pMSD5-HPRT expression vector is used as the selective plasmid (or selective or selectable marker) when HPRT$^-$ cell lines, such as BW5147.G.1.A, are used as the recipient cell line for the generation of stable transformants.

HPRT⁻ cell lines cannot grow in medium containing hypoxanthine, aminopterin or azaserine and thymidine (HAT medium). The addition of a functional HPRT gene by gene transfer allows the cells which have integrated the vector DNA encoding the HPRT gene to grow in HAT medium.

a. Isolation Of A Full Length Mouse HPRT cDNA

A cDNA library was prepared from poly A^+ mRNA isolated from C6VL cells [Allison *et al.*, J. Immunol., 129:2293 (1982)] using standard techniques [Sambrook *et al.*, *supra* at 7.26-7.29]. cDNA was generated from the mRNA and inserted into the expression vector λgt10 using standard techniques [Huynh, *et al.*, in *DNA Cloning: A Practical Approach* (D.M. Glover, ed.), Vol. 1, IRL Press Oxford (1985), pp. 49-78]. The full-length mouse HPRT cDNA was isolated using a full-length human HPRT cDNA clone containing an approximately 1.4 kb *Pst*I-*Bam*HI restriction fragment as a probe [pcD-HPRT; Jolly *et al.* (1983) Proc. Natl. Acad. Sci. USA 80:477]. The full length cDNA clone was digested with *Not*I and *Eco*RI to generate a 1.3 kb fragment containing the coding region of HPRT (the coding region of the mouse HPRT is listed in SEQ ID NO:13; the amino acid sequence encoded within SEQ ID NO:13 is listed in SEQ ID NO:14).

pMSD5 (described in Example 1) was digested with *Not*I and *Eco*RI and the 1.3 kb *Not*I/*Eco*RI fragment containing the mouse HPRT cDNA was inserted to generate pMSD5-HPRT. The map of pMSD5-HPRT is shown in Figure 11.

EXAMPLE 3

Construction Of The Amplification Vector pSSD7-DHFR

pSSD7-DHFR contains a full length copy of the mouse DHFR cDNA under the transcriptional control of the SV40 enhancer/promoter. This promoter/enhancer is active in a wide variety of cell types from many mammalian species [Dijkema *et al.*, EMBO J., 4:761 (1985)]. pSSD7-DHFR is referred to as the amplifiable marker as the use of this vector allows the selection of cell lines which have amplified the vector sequences by selecting for cell which can grow in increasing concentrations of MTX.

The mouse DHFR cDNA was isolated from double stranded cDNA generated from liver RNA using the PCR as follows. Poly A^+ RNA was isolated from the liver of (Balb/c x C57Bl/6) F1 mice using standard techniques. First strand cDNA was synthesized from the

poly A^+ RNA in a final reaction volume of 100 µl. The following reagents were added in order: 35.6 µl H_2O, 5 µl poly A^+ RNA (1 µg) and 1.4 µl SBNSSdT primer (1 µg). The sequence of the SBNSSdT primer is 5'-GCATGCGCGCGGCCGCGGAGGC-TTTTTTTTTTTTTTTTTT-3' (SEQ ID NO:15). The water, primer and RNA were heated at 60°C for 2 minutes then placed on ice. Forty µl of all four dNTPs at 5 mM each, 10 µl 10X reverse transcriptase salts (1.0 M Tris-HCl, pH 8.3, 0.5 M KCl, 0.1 M $MgCl_2$, 0.1 M DTT), 2 µl RNasin™ (Promega) and 5 µl AMV reverse transcriptase (Molecular Genetic Resources, Tampa, FL). The reaction was run at 41°C for 3 hours. The reaction was stopped by incubation at 65°C for 10 minutes.

The reaction components were transferred to a Centricon™ 100 tube (Amicon) and 2.1 ml of 5 mM Tris-HCl, pH 8.3 was added. The tube was centrifuged at 300 rpm (~700*g*) for 4 minutes at 10°C. 2.2 ml of Tris-HCl, pH 8.3 was added and the tube was centrifuged again as above. This washing step was repeated and then the tube was inverted and centrifuged at 2500 rpm for 5 minutes at 10°C to recover the first strand cDNA (volume ~50 µl). Second strand cDNA was synthesized as follows. 96 µl H_2O and 20 µl 10X rTth RTase buffer (900 mM KCl, 100 mM Tris-HCl, pH 8.3) was added to the first strand cDNA. In a separate tube the following components were mixed: 20 µl 10 mM $MnCl_2$, 4 µl of each of the four dNTPs at 10 mM and 10 µl rTth reverse transcriptase (Perkin-Elmer). Both mixtures were heated to 60°C and the second mixture was added to the cDNA mixture. The reaction was carried out at 60°C for 10 minutes. The reaction was stopped by addition of 25 µl chelating buffer [50% glycerol (v/v), 1 mM KCl, 100 mM Tris-HCl, pH 8.3, 7.5 mM EGTA, 0.5% Tween™ 20] and the mixture was placed on ice.

The reaction mixture was then transferred to a Centricon 100 tube and 2.1 ml of 5 mM Tris-HCl, pH 7.5 was added. The tube was centrifuged at 5500 rpm for 30 minutes at 10°C. 2.2 ml of Tris-HCl, pH 7.5 was added and the tube was centrifuged again as above. This washing step was repeated and then the tube was inverted and centrifuged at 2500 rpm for 5 minutes at 10°C to recover the double stranded cDNA (volume ~50 µl). The cDNA was precipitated with ethanol, resuspended in sterile H_2O and quantitated by absorption at 260 and 280 nm.

Two hundred pg of double stranded cDNA was used in a 400 µl PCR reaction. The primer set used to prime the PCR was: muDHFR.A: 5'-CGGCAAC GCGTGCCATCATGGTTCGAC-3' (SEQ ID NO:16) and muDHFR.B: 5'-CGGCA GCGGCCGCATAGATCTAAAGCCAGC-3' (SEQ ID NO:17). The PCR reaction conditions

were as reported in Saiki *et al.*, Science 239:487 (1988). Briefly, the reaction was run at 94°C for 1 minute, 55°C for 1 minute, 72°C for 1.5 minutes and 30 cycles were performed. *Taq* DNA polymerase was obtained from Perkin-Elmer and the reaction buffer used was that recommended by the manufacturer. The primer pair generates a 671 bp fragment having a *Mlu*I site at the 5' end and a *Not*I site at the 3' end (SEQ ID NO:18; the amino acid sequence encoded by SEQ ID NO:18 is listed in SEQ ID NO:19). The PCR reaction products were digested with *Mlu*I and *Not*I and electrophoresed on a low melting temperature agarose gel (SeaPlaque,™ FMC). The 671 bp fragment was cut out of the gel and the agarose was removed by digestion with β-Agarase I (NEB) followed by isopropanol precipitation according to the manufacturer's directions.

The 671 bp fragment was inserted into pSSD7 which was digested with *Mlu*I and *Not*I to generate pSSD7-DHFR. The map of pSSD7-DHFR is shown in Figure 12.

EXAMPLE 4

Construction Of The Expression Vector pJFE 14ΔIL10

pJFE 14ΔIL10 contains a full length cDNA clone encoding the mouse interleukin 10 (IL-10) protein under the transcriptional control of the SRα enhancer/promoter. As discussed above, the SRα enhancer/promoter is active in a broad range of cell types. pJFE 14ΔIL10 is used to direct the expression of the IL-10 gene in transfected cells (*i.e.*, pJFE 14ΔIL10 expresses IL-10 as the gene of interest).

a. Construction Of pJFE 14ΔIL10

The plasmid pJFE14 [Elliott *et al.* (1990) Proc. Natl. Acad. Sci USA 87:6363] was constructed by combining DNA fragments from the plasmids pSSD, pcDL-SRα296 [Takebe *et al.* (1988) Mol. Cell. Biol. 8:466] and pCDM8 [Seed (1987) Nature 329:840]. pSSD was cut with *Hin*dIII and *Xho*I and a 2.77 kb fragment was isolated from an agarose gel. pcD-SRα296 was cut with *Hin*dIII and *Xho*I and an ~ 640 bp fragment was isolated from an agarose gel. The two gel-purified DNA fragments were ligated together to generate the plasmid pSRαSD. pSRαSD was cut with *Xba*I and *Not*I and a 3.4 kb fragment was isolated from an agarose gel. pCMD8 was cut with *Xba*I and *Not*I and a 440 bp fragment was isolated. The 3.4 kb and 440 bp *Xba*I/*Not*I fragments were ligated together to generate pJEL14. A schematic of pJFE14 is shown in Figure 13.

The ΔIL10 cDNA was generated from a full-length mouse cDNA clone, F115 [Moore *et al.* (1990) Science 248:1230] using the PCR. The pcDSRα-F115 clone was linearized with *Bam*HI, which cuts out the cDNA insert. A PCR reaction was run using AmpliTaq™ DNA Polymerase (Perkin Elmer) and buffer supplied by the manufacturer according to their suggested conditions. The primers used in the PCR were IL10Δ-5' [5'-ATATATCTAGACCACCATGCCTGGCTCAGCACTG-3' (SEQ ID NO:20)] and IL10Δ-3' [5'-ATTATTGCGGCCGCTTAGCTTTTCATTTTGAT CAT-3' (SEQ ID NO:21)]. The PCR reaction was run at 94°C, 1 min, 72°C, 1 min, 46°C, 1 min for 30 cycles. The PCR generated DNA has deleted essentially all of the non- coding sequences and placed an optimal Kozak sequence just 5' to the initiator ATG of the IL-10 gene sequences. The PCR generated DNA was extracted with phenol:$CHCl_3$ (1:1) and then with $CHCl_3$. The DNA was ethanol precipitated. pelleted in a microcentrifuge and resuspended in TE. The DNA was cut with *Xba*I and *Not*I. pJFE14 was cut with *Xba*I and *Not*I. Both digestion mixtures were run on a low melt agarose gel. The 550 bp ΔIL10 band and the 3.4 kb pJFE14 band were cut out of the gel and combined in a tube. The DNAs were co-extracted from the agarose, ligated together and transformed into the bacteria DH5α. Colonies were picked and the clone pJFE14-ΔIL10 was identified. A schematic map of pJFE14-ΔIL10 is shown in Figure 14.

EXAMPLE 5

Construction Of pSRαSD5-DRα-DAF

pSRαSD5-DRα-DAF contains a cDNA clone encoding a chimeric mouse DRα gene. In this chimeric protein, the extracellular domain of the DRα protein is joined to sequences derived from the decay accelerating factor (DAF) gene. The DAF sequences provide a glycophosphatidylinositol linkage which allows the chimeric protein to be cleaved from the surface of the cell (cell surface expression requires the expression of the DRβ chain in the same cell) by treatment of the cell with phospholipase C.

a. Construction Of The Phagemid Vector pDAF20

To generate pSRαSD5-DRα-DAF and pSRαSD5-DRβ1-DAF (Example 6), a vector containing sequences encoding a portion of decay accelerating factor (DAF) which anchors DAF to the cell surface via a glycophosphatidylinositol linkage was constructed. pDAF20 was constructed as follows.

74667-99

Two micrograms of pBluescript™ KS(-) (Stratagene) was cut with *Eco*RV (NEB). TE buffer was added to such that the final volume was 200 µl. Spermine was added to a final concentration of 1.4 mM and the DNA was allowed to precipitate for 20 minutes on ice. The precipitated DNA was then pelleted by centrifugation for 10 min. in a microcentrifuge and the spermine was washed from the pellet exactly as described [Hoopes and McClure (1988) Nucleic Acids Res. 9:5493]. Briefly, the pellet was dispersed in extraction buffer [75% EtOH, 1X Buffer 2 (0.3M sodium acetate, 0.01M magnesium acetate)] by vortexing; the dispersed pellet was then left on ice for 1 hour. The pellet was collected by centrifugation for 10 min. in a microcentrifuge. The pellet was dried at room temperature and resuspended in 14 µl H_2O. On ice, 250 ng each of DAFa (SEQ ID NO:22) and DAFb (SEQ ID NO:23) unphosphorylated oligonucleotides were added to the resuspended DNA. The DNA-oligonucleotide mixture was then brought to a final concentration of 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM DTT and 1 mM rATP in a final reaction volume of 20 µl. Eighty units of T4 DNA ligase (NEB) was added and the ligation mixture was placed at 14°C overnight. The ligation mixture was then heated to 65°C for 10 min. NaCl was added to a final concentration of 50 mM and the DNA was digested with *Eco*RV (NEB). An aliquot of the DNA was then used to transform competent HB101.

Clones were picked and miniprep DNA was examined by restriction enzyme digestion. A clone, called DAF20, was isolated that has the DAF sequence cloned in the *Eco*RV site of pBluescript™ KS(-) with the *Xba*I at one end of the DAF sequence adjacent to the *Eco*RI site in the polylinker and away from the *Hin*dIII site in the polylinker. The sequence of the pDAF20 polylinker region containing the DAF insert is listed in SEQ ID NO:24.

The resulting plasmid pDAF20 contains DNA encoding the final 37 amino acids of the form of DAF that is anchored to the cell surface by a glycophosphatidylinositol (PI) linkage [Caras *et al.* (1987) Nature 325:545]. Chimeric proteins containing these 37 amino acids at their C-terminus, can be expressed on the cell surface of mammalian (and insect) cells with this PI anchor. This anchor can be readily cleaved and the protein solubilized from the cell surface using phosphatidylinositol-specific phospholipase C [Caras *et al.* (1987) Science 238:1280].

Phosphatidylinositol-specific phospholipase C was purified from *Bacillus thuringiensis* (ATCC 10792) exactly as described [Kupke *et al.* (1989) Eur. J. Biochem. 185:151]; phosphatidylinositol-specific phospholipase C is available commercially (*e.g.*, Sigma).

The use of soluble class II molecules complexed with specific peptides has been suggested for the treatment of autoimmune disease [Sharma, *et al.* (1991) Proc. Natl. Acad. Sci. USA 88:11465]. Such therapy requires that ample quantities of soluble class II molecules be available. The present invention allows large quantities of soluble class II molecules to be produced from cells expressing class II molecules on the cell surface wherein these molecules are anchored to the cell via the PI anchor provided by sequences derived from DAF. Alternatively, soluble forms of cell surface proteins can be produced according to the methods of the present invention using DNA sequences encoding chimeric class II molecules containing a thrombin cleavage site between the extracellular domain and the transmembrane domain of each chain comprising the class II heterodimer.

b. Isolation Of A Full-Length HLA DRα cDNA

A cDNA library was prepared from poly A^+ mRNA isolated from IBw4 cells (GM03104B, NIGMS Human Genetic Mutant Cell Repository at the Coriell Institute for Medical Research, Camden, NJ) using standard techniques [Sambrook *et al., supra* at 7.26-7.29]. cDNA was generated from the mRNA and inserted into the cloning vector λgt10 using standard techniques [Huynh *et al.*, in *DNA Cloning: A Practical Approach* (D.M. Glover, ed.), vol. 1, IRL Press Oxford (1985), pp. 49-78]. A full-length DRα cDNA was isolated from the library using a partial DRα cDNA as a probe; the partial DRα cDNA was contained within pDRα1 [Stetler *et al.* (1982) Proc. Natl. Acad. Sci. USA 79:5966]. The resulting full-length DRα cDNA was contained on a 1.2 kb *Not*I/*Eco*RI fragment.

c. Construction Of SRαSD5-DRα-DAF

An in-frame connection between the extracellular coding sequence of DRα and the DAF sequence was performed using site-directed *in vitro* deletional mutagenesis [Kunkel *et al.* (1987) Methods in Enzymology 154:367]. The mutational, bridging oligonucleotide encodes the desired connection.

The full length DRα cDNA was subcloned as a *Not*I-*Eco*RI fragment into pDAF20 (section a above). The pDAF20-DRα was isolated and transformed into the bacteria BW313 [Kunkel *et al.* (1987), *supra*]. A colony was then grown overnight in LB containing 100 μg/ml ampicillin. The overnight culture was diluted 1:10 in a final volume of 6 ml and grown at 37°C. After 1 hour, 400 μl of a stock of helper phage R408 [Russel *et al.* (1986) Gene 45:333] having a titer of approximately 1 x 10^{11} pfu/ml was added to the culture and the

culture was grown at 37°C for approximately 8 hours. One point four (1.4) ml aliquots of the culture were then placed into 4 microcentrifuge tubes and spun in a microcentrifuge 5 min at 4°C. One point one (1.1) ml of each supernatant was transferred to fresh microcentrifuge tubes containing 150 µl of 20% PEG(6000), 2.5 M NaCl. The contents of the tubes were mixed and allowed to stand at room temp. for at least 20 min. Precipitated, ssDNA containing phage particles were pelleted in a microcentrifuge for 5 min at 4°C. Care was taken to remove all the PEG-containing supernatant from the pellets. The four pellets were resuspended in a total of 200 µl of 300 mM NaOAc, pH 7 and extracted with an equal volume of phenol:$CHCl_3$ (1:1) twice, and then once with $CHCl_3$. Two volumes of ethanol was added to the supernatant and chilled to -20°C. The ssDNA was pelleted in a microcentrifuge 20 min at 4°C. The pellet was dried and resuspended in 10 µl TE buffer.

The bridging oligonucleotide was phosphorylated in a volume of 20 µl containing 50 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 10 mM DTT, 1 mM rATP and 65 ng of the RADAF2 oligonucleotide (SEQ ID NO:25) with 8 units of T4 DNA polynucleotide kinase (Pharmacia) at 37°C for 1 hour. To anneal the bridging oligonucleotide to the ssDNA template. 1.1 µl of the phosphorylated RADAF2 oligonucleotide (SEQ ID NO:25) and 5 µl of the ssDNA prep were mixed in a final volume of 15 µl of 40 mM Tris-HCl (pH 7.5), 20 mM $MgCl_2$, 50 mM NaCl, heated to 70°C and allowed to cool to room temp. on the bench top. In the reaction tube, the concentrations of the buffers were adjusted to give, in a final volume of 95 µl, 16.8 mM Tris-HCl, pH 7.5, 11.6 mM $MgCl_2$, 7.9 mM NaCl, 10.5 mM DTT and 1.1 mM rATP. Four units of T4 DNA ligase (NEB) and 3.8 units of Sequenase™ (US Biochemicals) were added to the reaction, which was incubated at room temp. for 5 min and 37°C for 1 hour. The reaction was adjusted to 58 mM NaCl and heated at 65°C for 10 min. The tube was cooled to 37°C and the DNA cut with *Eco*RI and *Xba*I. An aliquot of DNA was transformed into *E. coli* strain TG2 and plated on ampicillin-containing plates. A clone that showed the proper deletion of DNA between the desired connection of the DRα and DAF sequences was isolated. This clone was sequenced to confirmed the presence of the desired sequences using standard techniques. The coding region for the DRα-DAF protein is listed in SEQ ID NO:26; the amino acid sequence encoded by SEQ ID NO:26 is listed in SEQ ID NO:27.

The plasmid containing the correct DRα-DAF construct was cut with *Hin*dIII. The ends generated by *Hin*dIII digestion were made blunt with Klenow enzyme and unphosphorylated *Eco*RI linkers were ligated onto the blunt ends using standard techniques. The DNA was transformed into competent *E. coli* and clones which contained the DRα-DAF

sequences as a *Not*I-*Eco*RI fragment were isolated. The DRα-DAF DNA was then subcloned into the pSRαSD5 plasmid as a *Not*I-*Eco*RI fragment to generate pSRαSD5-DRα-DAF. The map of pSRαSD5-DRα-DAF is shown in Figure 15.

EXAMPLE 6

Construction Of pSRαSD5-DRβ1-DAF

pSRαSD5-DRβ1-DAF contains a cDNA clone encoding a chimeric mouse DRβ1-DAF gene. In this chimeric protein, the extracellular domain of the DRβ1 protein is joined to sequences derived from the DAF gene. The DAF sequences provide a glycophosphatidylinositol linkage which allows the chimeric protein to be cleaved from the surface of the cell (cell surface expression requires the expression of the DRα chain in the same cell) by treatment of the cell with phospholipase C.

a. Isolation Of A Full-Length DRβ1 cDNA

A cDNA library was prepared from poly A^+ mRNA isolated from IBw4 cells (GM03104B, NIGMS Human Genetic Mutant Cell Repository at the Coriell Institute for Medical Research, Camden, NJ) using standard techniques [Sambrook *et al., supra* at pp. 7.26-7.29]. cDNA was generated from mRNA and inserted into the cloning vector λgt10 using standard techniques [Huynh *et al.*, in *DNA Cloning: A Practical Approach* (D.M. Glover, ed.), vol. 1, IRL Press Oxford (1985), pp. 49-78]. A full-length DRβ1 cDNA clone was isolated from the library using a full length DRβ cDNA probe which was contained within the plasmid p2918.4 [Bell *et al.* (1985) Proc. Natl. Acad. Sci. USA 82:3405]. The resulting full-length DRβ1 clone was contained on a 1.2 kb *Not*I/*Eco*RI fragment.

b. Construction Of pSRαSD5-DRβ1-DAF

An in-frame connection between the extracellular coding sequence of DRβ and the DAF sequence was performed using site-directed *in vitro* deletional mutagenesis [Kunkel *et al.* (1987), *supra*] as described in Example 5c.

The full length DRβ1 cDNA (section a above) was subcloned into pDAF20 (Ex. 5a) as a *Not*I-*Eco*RI fragment to generate pDAF20-DRβ1. pDAF20-DRβ1 DNA was isolated and transformed into the *E. coli* strain BW313. A colony was then grown overnight in LB containing 100 μg/ml ampicillin. The overnight culture was diluted and incubated with helper

phage as described in Example 5c to generate single-stranded pDAF20-DRβ1 DNA. The ssDNA was precipitated and resuspended in TE buffer as described in Example 5c.

The bridging oligonucleotide, RQBDAF2 (SEQ ID NO:28), was phosphorylated as described in Example 5c. To anneal the bridging oligonucleotide to the ssDNA template, 1.1 µl of phosphorylated RADAF2 and 5 µl of the ssDNA prep were mixed, heated and cooled as described in Example 5c. The reaction mixture was adjusted to give, in a final volume of 95 µl, a concentration of 16.8 mM Tris-HCl (pH 7.5), 11.6 mM $MgCl_2$, 7.9 mM NaCl, 10.5 mM DTT and 1.1 mM rATP. Four units of T4 DNA ligase (NEB) and 3.8 units of Sequenase™ (US Biochemicals) were added to the reaction, which was incubated at room temp. for 5 min and 37°C for 1 hour. The reaction was adjusted to 58 mM NaCl and heated at 65°C for 10 min. The tube was cooled to 37°C and the DNA digested with *Eco*RI and *Xba*I. An aliquot of the digested DNA was used to transform *E. coli* strain TG2. The transformed cells were plated on plates containing ampicillin. A clone that showed the proper deletion of DNA between the desired connection of the DRβ1 and DAF sequences was isolated. The presence of the desired sequences was confirmed by DNA sequencing using standard techniques. The coding region for the DRβ1-DAF protein is listed in SEQ ID NO:29; the amino acid sequence encoded by SEQ ID NO:29 is listed in SEQ ID NO:30.

The plasmid containing the correct DRβ1-DAF construct was cut with *Hin*dIII. The DNA was blunted with Klenow enzyme and *Eco*RI linkers were added to the blunted ends using standard techniques. The DNA was transformed into bacteria that contained the DRβ1-DAF as a *Not*I-*Eco*RI fragment were isolated. The DRβ1-DAF DNA was subcloned into pSRαSD5 as a *Not*I-*Eco*RI fragment to generate pSRαSD5-DRβ1-DAF. The map of pSRαSD5-DRβ1-DAF is shown in Figure 16.

EXAMPLE 7

High-Level Expression Of Recombinant IL-10 In Lymphoid Cells

High levels of IL-10 were expressed in BW5147.G.1.4 cells (a T lymphoid cell line) by co-amplification of the following three plasmids: 1) the expression vector pJFE 14ΔIL10 which encodes mouse IL10; 2) the selection vector pMSD5-HPRT which encodes the HPRT enzyme and 3) the amplification vector pSSD7-DHFR which encodes the mouse DHFR enzyme. The plasmids were introduced into BW5147.G.1.4 cells by electroporation. The plasmid DNA was isolated from bacterial cells using CsCl density gradient centrifugation.

The plasmids were prepared for electroporation as follows. First, the plasmids were linearized in the same reaction tube. 200 µg of pJFE 14ΔIL10 was digested with *Sal*I. Ten µg of pMSD5-HPRT was digested with *Sal*I. Twenty µg of pSSD7-DHFR was digested with *Sal*I. *Sal*I was obtained from New England BioLabs and restriction digests were performed according to the manufacturer's instructions. The linearized plasmids were then precipitated with ethanol and resuspended in 0.5 ml of 1X HBS(EP) buffer [20 mM HEPES (pH 7.0); 0.75 mM Na_2HPO_4/NaH_2PO_4 (pH 7.0); 137 mM NaCl; 5 mM KCl and 1 gm/l dextrose].

BW5147.G.1.4 cells were grown in RPMI 1640 medium (Gibco/BRL) containing 10% FCS (HyClone) and 50 µg/ml gentamycin (Sigma). Prior to electroporation, the cells were washed twice in ice cold 1xHBS(EP) buffer and resuspended at 2 x 10^7 cells/ml in 0.5 ml of 1X HBS(EP). The cells were then placed in a 1 ml cuvette (Sarstedt) which contained the linearized DNAs in 0.5 ml of 1X HBS(EP). The cuvette was placed on ice. The electroporation was performed at 225 volts using an ISCO Model 493 power supply. The electroporation apparatus was constructed exactly as described [Chu, G. *et al.*, (1987) Nucl. Acids Res. 15:1311]. The electroporation device was set on constant voltage (225V) at the 2X setting (*i.e.*, both capacitors were used). Following electroporation, the cells were allowed to recover by incubation on ice for 5 to 15 minutes.

The electroporated cells were then transferred to a T75 flask (Falcon) containing 30 ml of RPMI 1640 medium containing 10% FCS and 50 µg/ml gentamycin. The cells were placed in a humidified atmosphere containing 5% CO_2 at 37°C for 36 hours. The cells were then plated in 24 well plates (Falcon, Lincoln Park, NJ) at a density of 1 x 10^4 cells/well in selective medium [RPMI 1640 containing 10% FCS, 100 µM hypoxanthine (Sigma) and 2 µg/ml azaserine (Sigma)]. Each well contained 0.5 ml of selective medium. One week after plating the cells in the 24 well plates, 0.5 ml of fresh selective medium was added.

$HPRT^+$ colonies (*i.e.*, wells containing growing cells or positive wells) were visible after approximately 10 days. At day 13 (with the day of electroporation being day zero) 100 µl of culture supernatant was removed and assayed for the presence of mouse IL10 using an ELISA assay performed as described [Mosmann et al. (1990) J. Immunol. 145:2938]. The monoclonal antibody (mcab) SXC1 (PharMingen, San Diego, CA) was used as the capture antibody and biotinylated mcab SXC2 [the mcab JESS-2A5 (PharMingen) may be used in place of SXC2] was used as the detection antibody. Briefly, 20 µl of mcab SXC1 at a concentration of 2 µg/ml in PBS was allowed to bind to the wells of flexible vinyl 96 well plates (Falcon) by incubating for 30 min to 3 hours at 37°C. Excess protein binding sites

were then blocked by adding 200 µl/well PBS,10% FCS. After 30 minutes of blocking at 37°C, the plates were washed with PBS, 0.1% Tween™ 20 (ICN Biochemicals, Aurora, OH). Samples to be tested were added at 50 µl/well and incubated 1 hour at 37°C. Plates were washed with PBS, 0.1% Tween 20 and 20 µl/well of PBS,0.1% Tween 20, 1 µg/ml biotinylated mcab SXC2 was added. The plates were incubated 30 min. at 37°C. The supernatants were removed and the plates were washed with PBS, 0.1% Tween 20. A 1/5000 dilution of streptavidin-horseradish peroxidase conjugate (Jackson Immunoresearch Laboratories, West Grove, PA) in PBS, 0.1% Tween™ 20, 0.1% BSA was added at 50 µl/well and incubated 30 min. at 37°C. The plates were then exhaustively washed with PBS, 0.1% Tween™ 20 and 100 µl/well of 44 mM NaH_2PO_4, 28 mM Citric Acid, 0.003% H_2O_2, 1 mg/ml 2,2' azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (Sigma) was added. The optical densities (ODs) of the wells were measured after 1 hour using a VMAX microplate reader (Molecular Devices, Menlo Park, CA) with a test wavelength of 405 nm and a reference of 490 nm.

The cells from wells containing 1 to 3 apparent clones and which contained greater than or equal to 100 units IL10/ml were cloned by limiting dilution using standard techniques [Cloning by Limiting Dilution, in Current Protocols in Immunology (J.E. Coligan *et al.*, eds.) John Wiley & Sons, New York, section 2.5.10]. For the limit dilution cloning, the cells were plated at 2 cells or 4 cells per well in a 96 well plate (Falcon) in selective medium; one 96 well plate was set up for each cell density (2 or 4 cells/well). In total, 16 independent colonies were cloned by limit dilution.

Eight days after limit dilution cloning was initiated, isolated colonies were picked from each of the limit dilution plates; these colonies were transferred to a 96 well plate; each well contained 5 ml RPMI 1640 containing 10% dialyzed FCS (HyClone) and 100 µM hypoxanthine. The use of dialyzed serum at this point increases the speed and frequency of amplification of the transfectants; hypoxanthine is added to the medium at this point as it is required for the growth of the cells for a few passages until the azaserine level is diluted to a negligible concentration.

Two days later, 100 µl of culture supernatant was tested for the presence of IL-10 using an ELISA as described above. The two best-producing clones from each of the original wells (*e.g.*, the 24 well plate) were chosen for further manipulation. In total 19 clones (termed selectants as these clones have survived growth in selective medium but have not yet been subjected to amplification by growth in the presence of methotrexate) were chosen.

Five days after the transfer of the isolated colonies (cloned by limit dilution) to 96 well plates. the colonies were transferred to 24 well plates and allowed to expand. The expanded colonies were then transferred to 5 ml flasks (Falcon) containing 5 ml of RPMI 1640 medium containing 10% dialyzed FCS. The clones produced between 100 and 200 units/ml of IL-10.

The selected clones were then subjected to amplification by growing the cells in the presence of methotrexate. The 19 clones were each tested for their sensitivity to methotrexate (MTX). Five x 10^4 cells from each clone was placed into a well in a series of 24 well plates. The clones were grown in the presence of RPMI 1640 medium containing 10% dialyzed FCS and either 3, 10, 30, 60 or 90 nM MTX. Six clones were able to grow in the presence of greater than or equal to 30 nM MTX; these clones were retained.

The six clones resistant to ≥ 30 nM MTX were plated in T25 flasks (Falcon) containing 5 ml of RPMI 1640 medium containing 10% dialyzed FCS and either 90, 150 or 210 nM MTX. Three flasks were set up for each clone. The clones were allowed to grow for 15 days at these three concentrations of MTX and then supernatants were taken from each flask and assayed for IL-10 production using an ELISA as above. All clones from flasks containing 90 or 150 nM MTX produced between 800 and 1200 units/ml of IL10. The best producing clone from each of the six original MTX^r clones was selected (one from a 90 nM MTX flask and the rest from 150 nM MTX flasks). These clones were then expanded to 5 mls in medium containing the appropriate concentration of MTX (over a 6 day period). The clones were then transferred into medium containing either 450, 750 or 1050 nM MTX. Sixteen days later supernatants from clones growing in the presence of 1050 nm MTX were assayed for IL-10 production. The clones were found to produce between 12,000 and 76,000 units/ml of IL-10 (one clone produced 12,000 u/ml, one clone produced 15,000 u/ml and eight clones produced between 50,000 and 76,000 u/ml).

The two clones producing the highest levels of IL-10 were chosen; these clones were designated as 9-2 and 11-2. Clones 9-2 and 11-2 were then grown in the presence of 5 μM MTX for 3 weeks, expanded and then frozen. Cultures were frozen as follows. Thirty milliliters of media containing cells at a density of 6 to 10 x 10^5 cells per ml were pelleted in a 50 ml conical tube (Falcon) at 500 x *g* for 5 minutes. The supernatant was poured off and the cells were resuspended in 7.5 ml of Freezing Media (40% FCS, 53% RPMI 1640, 7% DMSO) and placed in 5 freezing vials (Nunc, Naperville, IL). The cells were placed in a -70°C freezer for 24 to 96 hours and then transferred to liquid nitrogen for long term storage.

Aliquots of each clone were thawed after approximately 2 months, re-tested for IL-10 production and grown continuously in the presence of 5 μM MTX. These two clones (9-2 and 11-2) continue to produce between 64,000 to 86,000 units/ml of IL10.

The levels of expression of IL10 were roughly equivalent when the cells were grown at 1 or 5 μM MTX (compare 76,000 at 1 μM to 64-86,000 at 5 μM). The use of concentrations of MTX greater than 5 μM appeared to make the cells grow more slowly so that the total yield of protein was no greater than that obtained by growing the cells in the presence of 1 to 5 μM MTX.

It should be noted that selective pressure to maintain the expression of the HPRT protein (*i.e.*, growth in the presence of medium containing hypoxanthine and azaserine) was not used after the cells were transferred into medium containing MTX with no loss of IL-10 expression. Furthermore, because the level of IL-10 continued to rise with increasing concentrations of MTX, the endogenous DHFR gene is not likely to be amplified in the MTX^r cells. In other words, the increase in MTX-resistance is due to the amplification of the exogenous DHFR gene present on the amplification vector pSSD7-DHFR.

EXAMPLE 8

High-Level Expression Of DR Class II MHC In Lymphoid Cells

High levels of DR class II MHC molecules were expressed on the surface of BW5147.G.1.4 cells by co-amplification of the following four plasmids: 1) the expression vector pSRαSD5-DRα-DAF which encodes the alpha chain of the human DR molecule linked to a DAF tail; 2) the expression vector pSRαSD5-DRβ1-DAF which encodes the beta chain of the human DR molecule linked to a DAF tail; 3) the selection vector pMSD5-HPRT which encodes the HPRT enzyme and 3) the amplification vector pSSD7-DHFR which encodes the mouse DHFR enzyme. The plasmids were introduced into BW5147.G.1.4 cells by electroporation. The plasmid DNAs were isolated from bacterial cells using the standard technique of CsCl density gradient centrifugation.

The isolated plasmid DNAs were prepared for electroporation as follows. First the plasmids were linearized in the same reaction tube. All four plasmids were linearized with *Sal*I. The following amounts of plasmid were used: 200 μg of pSRαSD5-DRα-DAF; 200 μg of pSRαSD5-DRβ1-DAF; 10 μg of pMSD5-HPRT and 25 μg of pSSD7-DHFR. The

linearized plasmids were then precipitated with ethanol and resuspended in 0.5 ml of 1xHBS(EP) buffer.

BW5147.G.1.4 cells were grown in RPMI-1640 medium containing 10% FCS and 50 µg/ml gentamicin. Prior to electroporation the cells were washed twice in ice cold 1X HBS(EP) buffer and resuspended at a density of 2 x 10^7 cells/ml in 0.5 ml of 1X HBS(EP). The cells were then placed in a 1 ml cuvette (Sarstedt) which contained the linearized DNAs in 0.5 ml of 1X HBS(EP). The cuvette was placed on ice. The electroporation was performed as described above.

After electroporation the cells were allowed to recover by incubation on ice and then they were placed in a T75 flask (Falcon) containing 30 ml of RPMI-1640 medium containing 10% FCS and 50 µg/ml gentamicin. The cells were placed in a humidified atmosphere containing 5% CO_2 at 37°C and grown in bulk culture for 36 hours. The cells were then plated into four 48 well plates (Costar) at a density of 10^4 cells/well in 0.5 ml selective medium [RPMI 1640 containing 10% FCS, 100 µM hypoxanthine (Sigma) and 2 µg/ml azaserine (Sigma)]. The use of a cell density of 1 x 10^4 ensures that any colonies which arise are derived from a single cell; that is this density provides for limit dilution cloning. Any remaining cells were plated at a density of 1 x 10^5 cells/well in 0.5 ml of selective medium. One week after plating in the 48 well plates an additional 0.5 ml of selective medium was added..

Wells containing clones capable of growth in the selective medium (selectants) were visible after 8 days. Positive colonies (*i.e.*, positive for growth in selective medium) were picked into 12 well plates (Costar) containing 4 ml of RPMI 1640 containing 10% dialyzed FCS (HyClone) and 100 µM hypoxanthine 10-12 days after the application of selective medium. The use of dialyzed serum at this point increases the speed and frequency of amplification of the selectants; hypoxanthine is added to the medium at this point as it is required for the growth of the cells for a few passages until the azaserine level is diluted to a negligible concentration. The cells were allowed to grow for 3-4 days in the 12 well plates.

Colonies which grew in the presence of hypoxanthine and azaserine (selectants) were checked for the ability to express the DR molecule on the surface of the cell by staining cells with the monoclonal antibody L243. L243 binds specifically to the human HLA-DR antigens [Lampson and Levy, J. Immunol., 125:293 (1980)].

The antibody was prepared as follows. Hybridoma L243 was grown and the culture supernatant collected using standard techniques [Harlow and Lane, eds., *Antibodies: A*

Laboratory Manual, Cold Spring Harbor Press, New York (1988), pp. 272, 276]. The monoclonal antibodies were purified from the hybridoma supernatants. L243 was purified on a Protein A-Sepharose™ column (Pharmacia) using the protocol supplied by the manufacturer. The purified monoclonal antibody was then biotinylated using standard techniques [*Antibodies: A Laboratory Manual*, *supra* at p. 341]. Biotin was obtained from Vector. Biotinylated L243 was used at a dilution of 1:200.

The cells were stained as follows. The contents of the wells on the 12 well plates were gently mixed by pipeting the medium. One to 2 ml of the cell suspension was removed; this sample size contains 1-3 x 10^6 cells. The cells were pelleted by centrifugation at 1000 rpm for 4 minutes at 4°C. One hundred µl of L243 diluted into staining media (10 mM HEPES, pH 7.0, 5% calf serum, 4 mM sodium azide in Hanks balanced salt solution) was added. The cells were incubated for 20 minutes on ice. The cells were then washed by adding 1 ml of staining media and then the cells were underlaid with 1 ml of calf serum. The cells were pelleted through the serum by centrifugation at 1000 rpm for 4 minutes at 4°C. The supernatant was removed by aspiration. The cells were then suspended in 100 µl of fluorescein isothiocyanate (FITC) conjugated avidin (Vector, used at 1:50 dilution). The cells were incubated for 20 minutes on ice. The cells were then washed as described above.

The supernatant was removed and the cells were suspended in 200 µl of staining media containing 2 µg/ml propidium iodide. Propidium iodide is excluded from living cells but taken up by dead or dying cells. The addition of propidium iodide allows the exclusion of dead cells (propidium iodide-bright cells) from the analysis. The cells were filtered through nylon screen (Nitex™ nylon monofilament, 48 micron mesh, Fairmont Fabrics, Hercules, CA) prior to analysis on a FACScan™ (Becton-Dickinson). An aliquot of parental BW5147.G.1.4 cells (*i.e.*, not transfected) was stained as above to provide a negative control.

Figure 17 shows the results of staining a representative selectant clone, clone 5, with L243. Figure 17 is a histogram showing the log of fluorescein (x axis) plotted against the relative number of cells in the sample. Cells which express the DR molecule on the surface of the BW5147.G.1.4 cell appear as fluorescein bright cells due to staining of the cell surface with biotylinated-L243 followed by FITC-avidin. As shown in Figure 17, all of the cells in clone 5 express the transfected DR molecule. The fact that surface expression of the DR molecule is seen shows that both the α and the β chain DR constructs are expressed inside clone 5.

Eight selectant clones having the highest levels of expression of DR were chosen for further manipulation. These eight selectant clones were then tested for their sensitivity to MTX. Each clone was plated at a density of 2 x 10^4 cells/well in a 24 well plate. Each well contained 1 ml of medium containing RPMI-1640, 10% dialyzed FCS and MTX. The clones were grown in the presence of either 3, 10, 30, 60 or 90 nM MTX. Non-transfected BW5147.G.1.4 cells were also grown in the above range of MTX as a control. Clones which grew in MTX levels at least 2-3 fold higher than that tolerated by the parental BW5147.G.1.4 (typically less than or equal to 10 nM MTX) were selected for further analysis. Four of the selectant clones grew in greater than or equal to 30 nM MTX and were retained; these clones are the primary transfectants chosen for amplification. All 4 clones which grew in > 30 nM MTX were analyzed for the ability to express DR molecules on the surface by an ELISA. The cell surface ELISA was performed as follows.

Between 5 and 20 x 10^4 cells/well were put into a U-bottom 96 well plate. The cells were pelleted in a centrifuge using a plate carrier at 1000 rpm for 3 min at 4°C. The supernatant was flicked from the wells, the cells dispersed from their pellets by tapping and the plate was placed on ice. Fifty microliters of a 1/200 dilution of biotinylated mcab L243 (Becton-Dickinson) in staining media [Hank's Basic Salt Solution (Irvine Scientific), 10 mM HEPES, pH 7, 5% calf serum] was added to each well. The cells were incubated with the biotinylated mcab for 20 min on ice. Ice cold staining media was added to a final volume of 200 µl/well. The cells were pelleted and the supernatant flicked out and the pellets dispersed as described above. The cells were washed twice more with 200 µl/well of ice cold staining media. Fifty microliters of a 1/1000 dilution of Horseradish peroxidase conjugated Avidin (Vector Laboratories, Burlingame, CA) was added per well and incubated on ice for 20 min. Ice cold staining media was added to a final volume of 200 µl/well. The cells were pelleted and the supernatant flicked out and the pellets dispersed as described above. The cells were washed three more with 200 µl/well of ice cold staining media. After the final wash, the plate was again tapped to disperse the cell pellets and each well received 200 µl of freshly made OPD Substrate Solution [16 mM Citric Acid, 34 mM Sodium Citrate, 0.01% H_2O_2, 1 mg/ml O-phenylene diamine dihydrochloride (Sigma)]. The plate was allowed to sit at room temp for 10 to 20 min. The cells were then pelleted at 1000 rpm for 3 min at 4°C. One hundred microliters of supernatant from each well was transferred to a fresh, flat bottom 96 well plate (Costar) and the plate was read on a VMAX™ microplate reader (Molecular Devices, Menlo Park, CA) at a wavelength of 450 nm.

All four clones expressed the DR molecule as judged by ELISA analysis. Each of these four clones was grown in the highest MTX level at which obvious growth still occurred as determined by the test for MTX sensitivity above; the levels ranged from 30 to 80 nM MTX. The clones were then again checked for the ability to express DR on the cell surface by staining with L243 and FACS analysis as above. One out four first round amplificants, clone 5, showed both an increased resistance to MTX and the best corresponding increase in DR expression (all four clones showed increased DR expression). The histogram of cells from clone 5 grown in 80 nM MTX is shown in Figure 18. In Figure 18 the log of fluorescein (x axis) is plotted against the relative number of cells in the sample. Growth in 80 nM MTX represents the first round of amplification for clone 5.

The three clones which grew in higher levels of MTX but which did not show a high coincidental increase in the expression of DR were discarded. Clone 5 was retained and subjected to further rounds of amplification by grow in increasing concentrations of MTX. Figures 19 and 20 show histograms of cells from clone 5 grown in 320 nM and 1 μM MTX, respectively. The cells were stained with L243 and analyzed on a FACScan as described above. As is shown in Figures 19 and 20, clone 5 continued to show a coincidental increase in DR expression and increased MTX-resistance. Integration of the area under the peaks of fluorescence from each of Figures 17-20 showed that clone 5 achieved a 30-fold increase in DR expression between the initial selectant stage and the third round of amplification (1 μM MTX^r).

Continued analysis of clone 5 demonstrated that it is extremely stable. Clone 5 grown in 1 μM MTX (referred to as the 1 μM MTX amplificant of clone 5) can be grown for 2 to 3 weeks in medium lacking MTX without any apparent drop in expression of DR (as judged by cell surface ELISA assays).

EXAMPLE 9

Production Of Large Quantities Of Soluble T Cell Receptor And Class II MHC Molecules

Tumors of B and T cells (*i.e.*, lymphomas and leukemias) are often clonal in nature and therefore the Ig or TCR carried on the surface of the tumor cell can serve as a tumor-specific antigen. Soluble forms of the tumor-specific Ig have been used to immunize patients in order to invoke an immune response against the tumor cell [Kwak *et al.* (1992) N. Engl. J.

Med. 327:1209 and Hsu *et al.* (1996) Nature Med. 2:52]. The therapeutic use of soluble forms of a patient's tumor-specific antigen requires that large quantities of the soluble antigen be produced in a short period of time so that immunization of the patient can be carried out quickly (*i.e.*, before the patient's disease progress to a point that therapy is pointless). Large quantities of soluble class II MHC molecules are required to allow treatment of autoimmune disease using soluble class II molecules complexed with specific peptides [Sharma, *et al. supra*].

The methods of the present invention allow the production of large quantities of soluble forms of class II MHC molecules and TCR to be produced in a rapid manner. These methods allow for the production of customized tumor cell vaccines comprising soluble TCR for the treatment of lymphoma and leukemia patients as well as the production of soluble class II MHC molecules for the treatment of autoimmune disease. DNA sequences encoding the chains comprising the extracellular domains of the TCR or class II MHC molecules expressed by the patient's tumor cells are cloned using the PCR. These sequences are joined to sequences encoding a thrombin cleavage site followed by the transmembrane and cytoplasmic domains of either the α or β chain of a mammalian class II MHC heterodimer. The sequences encoding each chain of the chimeric TCR or class II MHC molecules (*i.e.*, the genes of interest) are inserted into any of the SD7 vectors described herein (*e.g.*, pSRαSD7; Ex. 1) and the resulting vectors are co-transfected into BW5147.G.1.4 cells along with an amplification vector (*e.g.*, pSSD7-DHFR; Ex. 3) and, if so desired, a selection vector (*e.g.*, pMSD5-HPRT; Ex. 2). The transfected cells will express the chimeric TCR or class II MHC molecules on the cell surface. The transfected cells are subjected to selection and/or amplification in order to produce amplified cell lines which express large quantities of the chimeric TCR or class II MHC molecules on the cell surface. These chimeric proteins can be cleaved from the cell surface to produce soluble TCR or class II MHC molecules by digestion with thrombin.

The following discussion illustrates the production of soluble TCR or class II MHC proteins using amplified cell lines. An analogous approach can be used to produce soluble forms of any multi-chain cell surface protein.

a. Construction Of Vectors Encoding Chimeric TCR Chains

Sequences encoding chimeric α chain of a TCR are constructed which comprise (from the amino- to carboxyl-termini) the extracellular domains of the α chain of a TCR followed

by 21 amino acids derived from the thrombin receptor which comprise a thrombin cleavage site followed by 41 amino acids comprising the transmembrane and cytoplasmic domains of the class II MHC molecule DRα. An analogous construct is used to construct a chimeric β chain of a TCR comprising (from the amino- to carboxyl-termini) the extracellular domains of the β chain of a TCR followed by 21 amino acids derived from the thrombin receptor which comprise a thrombin cleavage site followed by 42 amino acids comprising the transmembrane and cytoplasmic domains of the class II MHC molecule DRβ1. Any mammalian class II MHC αβ pair can be used to provide sequences encoding the transmembrane and cytoplasmic domains of the MHC molecule which permit the association of the chimeric TCR chains. While, the number of amino acid residues comprising the transmembrane and cytoplasmic domains of the α and β chains of the class II MHC molecules differs by one, both MHC junctions are at the third amino acid residue from the beginning of the transmembrane domain. This arrangement preserves the glutamate residue from the α chain and the lysine from the β chain which have been shown to have a positive effect upon heterodimer formation of class II MHC molecules [Cosson and Bonifacino (1992) Science 258:659].

A vector containing sequences encoding the thrombin and class II MHC sequences is constructed by synthesizing the DNA sequences listed in SEQ ID NO:31 and SEQ ID NO:33. The amino acid sequence encoded by SEQ ID NO:31 is listed in SEQ ID NO:32 and amino acid sequence encoded by SEQ ID NO:33 is listed in SEQ ID NO:34.

SEQ ID NO:31 encodes the thrombin site-DRα chimeric sequence and SEQ ID NO:33 encodes the thrombin site-DRβ1 chimeric sequence. Inspection of these sequences shows that the sequences at the 5' end which encodes the thrombin site contains the recognition site for the following restriction enzymes: *Bam*HI, *Pvu*I and *Fsp*I. A *Not*I site is located at the 3' end of the thrombin site-DRβ_1 chimeric sequences. The synthetic DNA is inserted into any suitable vector (*e.g.*, pUC 18 or pUC 19) as a *Bam*HI-*Not*I fragment. The thrombin site encoded by these sequences is very efficiently cleaved by thrombin due to the presence of the hirudin-like domain following the thrombin cleavage site [Vu *et al.* (1991) Cell 64:1057 and Vu *et al.* (1991) Nature 353:674].

DNA sequences encoding TCR chains are isolated from double-stranded cDNA generated from a cell line or a patient's tumor (double-stranded cDNA may be generated using the protocol set forth in Example 3; oligo d(T) may be used to prime first strand cDNA synthesis in place of the SBNSSdT primer). The double stranded cDNA is then used in PCRs which contain primer pairs designed to amplify either the α chain or the β chain of the

human TCR. The PCR is conducted using 1 unit/100 µl reaction Pfu polymerase (Stratagene) in the reaction buffer provided by Stratagene, 5 ng/100 µl of a cloned template or 25 ng/100 µl of ds-cDNA derived from polyA+ RNA isolated from a cell line or tumor, 0.1 mM of each of the four dNTPs and 0.5 µM of each primer. The PCR is cycled at 94°C for 15 sec followed by 60°C for 30 sec followed by 75°C for 2 min for 21 cycles.

The 5' primer used to amplify TCR sequences contains the following restriction sites at the 5' end of the primer: *Xba*I, *Eco*RI and *Mlu*I followed 18-21 nucleotides comprising a consensus sequence derived from the V regions of human TCRs. Therefore the 5' primer will comprise sets of degenerate primers having the following sequence: 5'-TCTAGAATTCACGCGT$(N)_{18-21}$-3' (SEQ ID NO:81), where N is any nucleotide and the 18-21 nucleotide stretch represents a consensus V region sequence. The following 3' primer is used in conjunction with the above-described consensus 5' primer to amplify the extracellular domains of human TCR α chains: 5'-CGATCGTGGATCCAAGTTTAGGTTCGTATCTGTTTCAAA-3' (SEQ ID NO:35). The 3' connection for the TCR α chain is made after the asparagine which appears at position 110 of the constant (C) region of the α chain. The following 3' primer is used in conjunction with the above-described consensus 5' primer to amplify the extracellular domains of human TCR β chains: 5'-CGATCGAGGATCC AAGATGGTGGCAGACAGGACC-3' (SEQ ID NO:36). The 3' connection for the TCR α chain is made after the isoleucine which appears at position 147 of the C region of the β chain. These 3' primers are designed such that in both cases (*i.e.*, for both the α and the β chain of the TCR) the connection between the extracellular domains of the TCR with the thrombin site is made at the fourth amino acid residue from the apparent beginning of the respective transmembrane regions of the TCR chains. Both 3' primers contain recognition sites for *Pvu*I and *Bam*HI at their 5' ends. The restriction sites located at the 5' ends of the primers allows the resulting PCR products comprising a TCR chain to be removed as a *Xba*I or *Eco*RI or *Mlu*I (5' end)-*Bam*HI or *Pvu*I (3' end) fragment and joined with the appropriate thrombin-transmembrane DNA sequence [as a *Bam*HI or *Pvu*I (5' end)-*Not*I (3' end) fragment] and inserted into any of the SD7 vectors (*e.g.*, pSRαSD7). The resulting expression vectors (one for each of the α chains and the β chains of the chimeric TCR) are co-transfected using electroporation into BW5147.G.1.4 cells along with the amplification vector pSSD7-DHFR (Ex. 3) and the selection vector pMSD5-HPRT (Ex. 2). The amount of each plasmid DNA to be used (the plasmids are linearized before electroporation), the conditions for electroporation, selection and amplification are

described above. The resulting amplified cell lines will express the chimeric TCR heterodimer on the surface of the cell. The TCR is solubilized by digestion of the cells with thrombin. The thrombin solubilized extracellular domains will have 3 (TCR β) or 4 (TCR α) novel amino acids at the C-termini.

b. Construction Of Vectors Encoding Chimeric Class II MHC Chains

Sequences encoding a chimeric α chain of a class II MHC protein are constructed which comprise (from the amino- to carboxyl-termini) the extracellular domains of the α chain of DRα followed by 21 amino acids derived from the thrombin receptor which comprise a thrombin cleavage site followed by 41 amino acids comprising the transmembrane and cytoplasmic domains of the class II MHC molecule DRα. An analogous construct is used to construct a chimeric β chain of a class II MHC protein comprising (from the amino- to carboxyl-termini) the extracellular domains of the β chain of $DR\beta_1$ followed by 21 amino acids derived from the thrombin receptor which comprise a thrombin cleavage site followed by 42 amino acids comprising the transmembrane and cytoplasmic domains of the class II MHC molecule $DR\beta_1$.

Sequences encoding the extracellular domains of the α and β chains of a class II MHC heterodimer are isolated using the PCR as described above with the exception that the following primer pairs are used in the PCR. Sequences encoding the extracellular domain of DRα are amplified using 5'-ACGCGTCCACCATGGCC ATAAGTGGAGTCCCT-3' (SEQ ID NO:37) (this primer contains a *Mlu*I site at the 5' end) and 5'-GGATCCAACTCTGTAGTCTCTGGGAGAG-3' (SEQ ID NO:38) (this primer contains a *Bam*HI site at the 5' end). The use of these primers allows the connection of the extracellular domain of DRα with the thrombin site-transmembrane sequences (described above) after amino acid 191, a glutamate residue in the mature (*i.e.*, after the removal of the signal sequence) DRα protein.

Sequences encoding the extracellular domain of $DR\beta_1$ are amplified using: 5'-ACGCGTCCACCATGGTGTGTCTGAAGCTCCTG-3' (SEQ ID NO:39) (this primer contains a *Mlu*I site at the 5' end) and 5'-GGATCCAACTTGCTCTGTGCA GATTCAGA-3' (SEQ ID NO:40) (this primer contains a *Bam*HI site at the 5' end). The use of these primers allows the connection of the extracellular domain of DRβ with the thrombin site-

transmembrane sequences (described above) after amino acid 198, a lysine residue, in the mature DRβ protein.

The restriction sites located at the 5' ends of the primers allows the resulting PCR products comprising the class II MHC chains to be removed as a *Mlu*I (5' end)-*Bam*HI (3' end) fragment and joined with the appropriate thrombin-transmembrane DNA sequence [as a *Bam*HI (5' end)-*Not*I (3' end) fragment] and inserted into any of the SD7 vectors (*e.g.*, pSRαSD7). The resulting expression vectors (one for each of the α chains and the β chains of the chimeric class II MHC protein) are co-transfected using electroporation into BW5147.G.1.4 cells along with the amplification vector pSSD7-DHFR (Ex. 3) and the selection vector pMSD5-HPRT (Ex. 2). The amount of each plasmid DNA to be used (the plasmids are linearized before electroporation), the conditions for electroporation, selection and amplification are described above. The resulting amplified cell lines will express the chimeric class II heterodimer on the surface of the cell. The class II MHC heterodimer is solubilized by digestion of the cells with thrombin.

EXAMPLE 10

Production of Custom Multivalent Vaccines For The Treatment of Lymphoma and Leukemia

The existing approach toward vaccination (*i.e.*, active immunotherapy) of B-cell lymphoma and leukemia involves the production of a custom vaccine comprising autologous immunoglobulin idiotype which corresponds to the most abundant antibody molecule expressed on the surface of the B-cell tumor. An analogous approach for the treatment of T-cell lymphomas and leukemias would involve the production of a custom vaccine comprising autologous T cell receptor (TCR) idiotype which corresponds to the most abundant TCR molecule expressed on the surface of the B-cell tumor.

Existing methods for the production of custom vaccines for the treatment of B-cell lymphoma employ the "rescue fusion" technique. The rescue fusion technique involves the removal of lymphoma cells by surgical biopsy. The tumor cells are then fused with the heterohybridoma cell line K6H6/B5 which has lost the ability to secrete endogenous Ig. Hybrid cells which secrete Ig corresponding to the immunophenotype of the tumor sample are expanded and the secreted Ig is purified for use as a vaccine [Kwak *et al.* (1992), *supra*]. The Ig produced by rescue fusion represents a single Ig derived from the patient's tumor; this

Ig is presumably the predominant Ig expressed by the tumor. Thus, vaccines produced by rescue fusion are monovalent and do not represent the full complexity of Ig expressed by tumors which contain somatic variants.

In order to produce multivalent custom vaccines from small numbers of cells quickly and efficiently, the gene amplification techniques described in the preceding examples are employed. In this example, methods for the production of tumor-specific Ig derived from a B-cell lymphoma patient are provided. However, the general approach outlined herein is applicable for the production of tumor-specific proteins generally (*i.e.*, production of soluble TCR for treatment of T cell tumors, production of Ig for treatment of B cell leukemias, etc.).

In this novel approach, the variable regions corresponding to the patient's Ig (V_H and V_L) are molecularly cloned and joined to an appropriate constant region gene contained within an expression vector. Expression plasmids containing the patient's V_H region(s) joined to either a Cγ3 or Cγ4 sequence and expression plasmids containing the patient's V_L region(s) joined to either a Cκ or Cλ2 sequence are cotransfected (via electroporation) along with the selectable and amplifiable marker pM-HPRT-SSD9-DHFR into the desired cell line (*e.g.*, BW5147.G.1.4). The transfected cells are then subjected to selection and amplification as described in the preceding examples. The method outlined below permits the production of a multivalent vaccine which reflects the degree of somatic variation found within the patient's tumor. These novel multivalent vaccine preparations provide superior vaccines for the treatment of B-cell lymphoma and should reduce the rate of relapse observed when the current generation of monovalent vaccines are employed.

a) Construction of Expression and Selection/Amplification Plasmids

For the following constructions, unless otherwise stated, all enzymes are obtained from New England Biolabs (NEB) and used in conjunction with the buffers and reaction conditions recommended by the manufacturer.

i) Construction of pSRαSD9

Two micrograms of pSRαSD7 (Ex. X) is cut with *Sal*I and *Hin*dIII (NEB enzymes, buffers & conditions). The plasmid is spermine precipitated (Ex. 5) and resuspended in 34 µl H_2O and 4 µl of 10x T4 DNA ligase buffer. Equal molar amounts (6.3 ng each) of the unphosphorylated oligonucleotides SXAPH5 (SEQ ID NO:42) and SXAPH3 (SEQ ID NO:43) are added. The reaction is chilled on ice, 400 units of T4 DNA ligase is added and the tube is placed at 14°C overnight. The ligation is transformed into bacteria and clones screened for

the presence of the added *Asc*I & *Pac*I restriction sites. The resulting plasmid is called pSRαSD9. Figure 21 provides a schematic map of pSRαSD9.

ii) Construction of pSRαSD9CG3C, pSRαSD9CG4C, pSRαSD9CKC and pSRαSD9CL2C

The plasmids pSRαSD9CG3C, pSRαSD9CG4C, pSRαSD9CKC and pSRαSD9CL2C contain sequences encoding the Cγ3, Cγ4, Cκ or Cλ2 constant regions, respectively. The constant regions contained within these expression vectors are encoded by synthetic DNA sequences which encode the same amino acid sequences as that found in the native proteins; however, the DNA sequences have been modified to utilize codons which are found most frequently in highly expressed mammalian proteins [Haas *et al.* (1996) Curr. Biol. 6:315 and Zolotukhin *et al.* (1996) J. Virol. 70:4646]. The DNA sequence encoding the Cγ3 region is listed in SEQ ID NO:44; the amino acid sequence encoded by SEQ ID NO:44 is listed in SEQ ID NO:45. The DNA sequence encoding the Cγ4 region is listed in SEQ ID NO:46; the amino acid sequence encoded by SEQ ID NO:46 is listed in SEQ ID NO:47. The DNA sequence encoding the Cκ region is listed in SEQ ID NO:48; the amino acid sequence encoded by SEQ ID NO:48 is listed in SEQ ID NO:49. The DNA sequence encoding the Cλ2 region is listed in SEQ ID NO:50; the amino acid sequence encoded by SEQ ID NO:50 is listed in SEQ ID NO:51.

Double stranded DNA corresponding to SEQ ID NOS:44, 46, 48 and 50 are synthesized (Operon Technologies). Each synthetic DNA sequence is cut with *Not*I and *Bgl*II, run through a 0.8% SeaPlaque™ Agarose gel (FMC) and recovered using β-agarase as described below. Each C region sequence is ligated to the two DNA restriction fragments generated from pSRαSD9 as follows. A 2 µg aliquot of pSRαSD9 is cut with *Hin*dIII and *Bam*HI and a 2314bp band is isolated. A second 2 µg aliquot of pSRαSD9 is cut with *Hin*dIII and *Not*I and an 854 bp band is isolated. These fragments are isolated by running each digest on a 0.8% SeaPlacque Agarose (FMC), the appropriate bands are cut out and combined in a microfuge tube. The agarose is remove by β-Agarase (NEB) digestion and the DNA is recovered by isopropanol precipitation exactly as indicated by NEB.

The ligation of SEQ ID NO:44 (digested with *Not*I and *Bgl*II) with the above fragments of pSRαSD9 generates pSRαSD9CG3C (map shown in Figure 22). The ligation of SEQ ID NO:45 (digested with *Not*I and *Bgl*II) with the above fragments of pSRαSD9 generates pSRαSD9CG4C (map shown in Figure 23). The ligation of SEQ ID NO:46 (digested with *Not*I and *Bgl*II) with the above fragments of pSRαSD9 generates

pSRαSD9CKC (map shown in Figure 24). The ligation of SEQ ID NO:47 (digested with *Not*I and *Bgl*II) with the above fragments of pSRαSD9 generates pSRαSD9CL2C (map shown in Figure 25).

iii) Construction of pM-HPRT-SSD9-DHFR

pM-HPRT-SSD9-DHFR contains the *hprt* gene under the control of the Moloney enhancer/promoter and the *dhfr* gene under the control of the SV40 enhancer/promoter. pM-HPRT-SSD9-DHFR is constructed by first subcloning the HPRT cDNA (Ex. 2) into pMSD8 (described below) to create pMSD8-HPRT. The small DNA fragment located between the *Sal*I and *Hin*dIII sites on pMSD8-HPRT is then replaced with a sequence containing *Asc*I and *Pac*I sites as follows. pMSD8-HPRT is digested with *Sal*I and *Hin*dIII and the SXAPH5 and SAXPH3 oligonucleotides (SEQ ID NOS:42 and 43) are ligated to the ends of the digested pMSD8-HPRT (as described in section i above) to create pMSD9-HPRT. The ~2450 bp *Sal*I-*Cla*I fragment containing the *Asc*I and *Pac*I sites, the Moloney enhancer/promoter, the HPRT cDNA and the EF1α poly A region is inserted between the *Sal*I and *Cla*I sites of pSSD7-DHFR (Ex.) to generate pM-HPRT-SSD9-DHFR. Figure 26 provides a map of pM-HPRT-SSD9-DHFR.

pMSD8 is similar to pMSD5 but contains the poly A site from the human elongation factor 1α gene. pMSD8 was constructed as follows: A 292 bp fragment containing the poly A site from the human elongation factor 1α gene (SEQ ID NO:78) was isolated from MOU cell (GM 08605, NIGMS Human Genetic Mutant Cell Repository, Camden, NJ) genomic DNA using PCR. MOU genomic DNA was isolated using conventional techniques. The PCR was conducted using 10 µg MOU genomic DNA and 1 µM final concentration of each primer in a 400 µl reaction. Reaction conditions were 94°C for 1 minute, 60°C for 1 minute, 72°C for 1.5 minutes, 30 cycles. *Taq* DNA polymerase was obtained from Perkin-Elmer. The following oligonucleotides were used to prime the PCR: 5EF1αPolyA: 5' GAATTCTTTTTTGCGTGTGGCAG 3' (SEQ ID NO:79) and 3EF1αPolyA: 5' ATCGATATTCCTTCCCCTTCC 3' (SEQ ID NO:80). The 3EF1αPolyA oligonucleotide generates a *Cla*I site at the 3' end of the poly A site and the 5EF1αPolyA oligonucleotide generates an *Eco*RI site at the 5' end of the poly A site. Digestion of the PCR product with *Eco*RI and *Cla*I yields a 292 bp *Eco*RI/*Cla*I fragment.

pSSD5 (Ex. 1) was digested with *Pvu*II and a *Cla*I linkers (NEB, unphosphorylated) were ligated to the *Pvu*II ends to convert the *Pvu*II site located at the 3' end of the SV40 poly A site to a *Cla*I site. The resulting construct was then digested with *Sal*I and *Cla*I and the ~2.1

kb fragment containing the plasmid backbone (*e.g.*, the Amp^R gene and plasmid ORI) was isolated and ligated to an ~870 bp *Sal*I/*Eco*RI fragment containing the Moloney enhancer/promoter, splice donor/acceptor and polylinker isolated from pMSD5 (Ex. 1) together with the 292 bp *Eco*RI/*Cla*I fragment containing the poly A site of the human elongation factor 1α gene to generate pMSD8.

b) Collection of Tumor Cells

Cells are collected by either surgical biopsy of enlarged lymph nodes or by fine needle biopsy of effected lymph nodes. The biopsy sample is rapidly frozen on dry ice.

c) Isolation of RNA From Tumor Cells

RNA is isolated from the biopsy sample by using a variety of standard techniques or commercially available kits. For example, kits which allow the isolation of RNA from tissue samples are available from Qiagen, Inc. (Chatsworth, CA) and Stratagene (LaJolla, CA), respectively. Total RNA may be isolated from tissues and tumors by a number of methods known to those skilled in the art and commercial kits are available to facilitate the isolation. For example, the RNeasy® kit (Qiagen Inc., Chatsworth, CA) provides protocol, reagents and plasticware to permit the isolation of total RNA from tissues, cultured cells or bacteria, with no modification to the manufacturer's instructions, in approximately 20 minutes. Should it be desirable to further enrich for messenger RNAs, the polyadenylated RNAs in the mixture may be specifically isolated by binding to an oligo-deoxythymidine matrix, through the use of a kit such as the Oligotex® kit (Qiagen). Comparable isolation kits for both of these steps are available through a number of commercial suppliers.

In addition, RNA may be extracted from samples, including biopsy specimens, conveniently by lysing the homogenized tissue in a buffer containing 0.22 M NaCl, 0.75 mM $MgCl_2$, 0.1 M Tris-HCl, pH 8.0, 12.5 mM EDTA, 0.25% NP40, 1% SDS, 0.5 mM DTT, 500 u/ml placental RNAse inhibitor and 200 µg/ml Proteinase K. Following incubation at 37°C for 30 min, the RNA is extracted with phenol:chloroform (1:1) and the RNA is recovered by ethanol precipitation.

A particularly preferred method for the isolation of total cellular RNA from patient tumor samples is the RNAzol™ method (Teltest, Inc., Friendswood, TX) which is performed according to the manufacturer's instructions.

d) Cloning of Ig Genes from Tumor Cells

Because the first and third complementarity determining regions (CDRs) of rearranged immunoglobulin genes are flanked by conserved sequences, it is possible to design PCR primers capable of amplifying cDNA for the variable regions from mRNA derived from Ig-expressing tumor cells without any specific knowledge of the nucleotide sequence of that specific antibody. Primers suitable for isolating the variable regions from a patient's tumor are provided below.

Using total cellular RNA isolated from the tumor, double stranded (ds) cDNA is generated as described in Example 3 with the exception that 20 μg of total cellular RNA is used instead of poly A^+ RNA. Five percent of the ds cDNA preparation is used for each PCR reaction. [Alternatively, ds cDNA may be produced using the technique of RT-PCR (reverse transcription-PCR); kits which permit the user to start with tissue and produce a PCR product are available from Perkin Elmer (Norwalk, CT) and Stratagene (LaJolla, CA). The RT-PCR technique generates a single-stranded cDNA corresponding to a chosen segment of the coding region of a gene by using reverse transcription of RNA; the single-stranded cDNA is then used as template in the PCR].

PCR reactions are carried out in a final volume of 50 μl and contain 1x *Pfu* Buffer (Stratagene). all 4 dNTPs at 100μM each, primers at 0.5μM each, *Pfu* polymerase (Stratagene) and 5% of the ds cDNA preparation. The reactions are thermocycled as follows: 94°C, 15 sec; 60°C, 30 sec; 75°C, 1.5 min for 15-30 cycles. Aliquots (5 μl) are removed after 15, 20, 25 and 30 cycles to examine the appearance of the primary PCR product. Preparative reactions of 200 μl using the correct V region primers will be then run for cloning purposes.

Prior to conducting a PCR reaction to obtain Ig sequences from a patient's tumor, the tumor is immunophenotyped using commercially available antibodies to determine the heavy chain and light chain isotypes; this allows the number of PCRs to be minimized. For example, if the Ig expressed by the patient's tumor utilizes a μ heavy chain and a κ light chain, then PCR reactions described below which contain Cγ and Cλ primers need not be run. However, the use of PCR primers corresponding to heavy and light chain isotypes which are not utilized (according to the immunophenotyping results) by the patient's tumor serves as a convenient means to confirm the immunophenotyping results.

PCR primers utilized to clone variable regions of the patient's tumor-specific Ig are summarized below in Tables 1 through 3:

TABLE 1

Heavy Chain Primers:

VH1L	5'-TCT AGA ATT CAC GCG TCC ACC ATG GAC TGG ACC TGG AG-3'	SEQ ID NO:52
VH2L	5'-TCT AGA ATT CAC GCG TCC ACC ATG GAC ACA CTT TGC TAC AC -3'	SEQ ID NO:53
VH3L	5'-TCT AGA ATT CAC GCG TCC ACC ATG GAG TTT GGG CTG AGC TGG-3'	SEQ ID NO:54
VH4L	5'-TCT AGA ATT CAC GCG TCC ACC ATG AAA CAC CTG TGG TTC TTC CT-3'	SEQ ID NO:55
VH5L	5'-TCT AGA ATT CAC GCG TCC ACC ATG GGG TCA ACC GCC ATC CT-3'	SEQ ID NO:56
VH6L	5'-TCT AGA ATT CAC GCG TCC ACC ATG TCT GTC TCC TTC CTC ATC TT-3'	SEQ ID NO:57
C_γ	5'-GCC TGA GTT CCA CGA CAC CGT CAC-3'	SEQ ID NO:58
C_μ	5'-GGG GAA AAG GGT TGG GGC GGA TGC-3'	SEQ ID NO:59
JH1245	5'-GAG GGG CCC TTG GTC GAC GCT GAG GAG ACG GTG ACC AGG -3	SEQ ID NO:60
JH3	5'-GAG GGG CCC TTG GTC GAC GCT GAA GAG ACG GTG ACC ATT G -3'	SEQ ID NO:61
JH6	5'-GAG GGG CCC TTG GTC GAC GCT GAG GAG ACG GTG ACC GTG-3'	SEQ ID NO:62

TABLE 2

Kappa Chain Primers:

VκI	5'-TCT AGA ATT CAC GCG TCC ACC ATGGAC ATG AGG GTC CCC GCT CAG-3'	SEQ ID NO:63
VκII	5'-TCT AGA ATT CAC GCG TCC ACC ATG AGG CTC CCT GCT CAG C-3'	SEQ ID NO:64
VκIII	5'-TCT AGA ATT CAC GCG TCC ACC ATG GAA GCC CCA GCG CAG CTT-3'	SEQ ID NO:65
VκIV	5'-TCT AGA ATT CAC GCG TCC ACC ATG GTG TTG CAG ACC CAG GT-3'	SEQ ID NO:66
VκV	5'-TCT AGA ATT CAC GCG TCC ACC ATG GGG TCC CAG GTT CAC CT-3'	SEQ ID NO:67
VκVIa	5'-TCT AGA ATT CAC GCG TCC ACC ATG TTG CCA TCA CAA CTC ATT G-3'	SEQ ID NO:68
VκVIb	5'-TCT AGA ATT CAC GCG TCC ACC ATG GTG TCC CCGTTG CAA TT-3'	SEQ ID NO:69
Cκ	5'-GGT TCC GGA CTT AAG CTG CTC ATC AGA TGG CGG G-3'	SEQ ID NO:70

TABLE 3

Lambda Chain Primers:

VL1	5'-TCT AGA ATT CAC GCG TCC ACC ATG GCC TGCTCT CCT CTC CTC CT-3'	SEQ ID NO:71
VL2	5'-TCT AGA ATT CAC GCG TCC ACC ATG GCC TGG GCT CTG CTG CTC CT-3'	SEQ ID NO:72
VL3	5'-TCT AGA ATT CAC GCG TCC ACC ATG GCC TGG ATC CTT CTC CTC CTC-3'	SEQ ID NO:73
VL4	5'-TCT AGA ATT CAC GCG TCC ACC ATG GCC TGG ACC CCT CTC TGG CTC-3'	SEQ ID NO:74
VL6	5'-TCT AGA ATT CAC GCG TCC ACC ATG GCC TGG GCC CCA CTA CT-3'	SEQ ID NO:75
VL8	5'-TCT AGA ATT CAC GCG TCC ACC ATG GCC TGG ATG ATG CTT CTC CT-3'	SEQ ID NO:76
Cλ	5'-GGC GCC GCC TTG GGC TGA CCT AGG ACG GT-3'	SEQ ID NO:77

The VH1-6L primers contain recognition sites for *Xba*I, *Eco*RI and *Mlu*I at their 5' ends. The three JH primers contain recognition sites for *Apa*I and *Sal*I at their 5' ends. The seven Vκ primers contain recognition sites for *Xba*I, *Eco*RI and *Mlu*I at their 5' ends. The Cκ primer contains recognition sites for *Bsp*EI and *Afl*III at the 5' end. The six VL primers

contain recognition sites for *Xba*I, *Eco*RI and *Mlu*I at their 5' ends. The Cλ primer contains recognition sites for *Kas*I and *Avr*II at the 5' end.

For each tumor sample, five V_H PCR reactions are run. Each V_H reaction will contain the Cμ and Cγ primers. The Cμ primer (SEQ ID NO:59) should result in ~590 bp product for the heavy chain V (V_H) region expressed in an IgM positive tumor. The Cγ primer (SEQ ID NO:58) should result in ~480 bp product for the heavy chain V region expressed in an IgG positive tumor. The VH1, VH2, VH3, and VH4 primers (SEQ ID NOS: 52-55 , respectively) are used in separate PCR reactions and the VH5 and VH6 primers (SEQ ID NOS:56 and 57, respectively) are used together in the same reaction. The V_H primer(s), which when used in connection with a C_H region primer, gives a PCR product of the expected size is then be used in three separate PCR reactions containing either the JH1245, JH3 or JH6 primers (SEQ ID NOS:60-62. respectively) to generate a PCR product corresponding to the variable (V), diversity (D) and joining (J) regions present in the Ig(s) expressed by the patient's tumor. The VDJ reaction product is then subcloned into the pSRαSD9CG3C vector or pSRαSD9CG4C vector using the 5' *Xba*I, *Eco*RI or *Mlu*I sites and the 3' *Sal*I or *Apa*I sites to provide an expression vector encoding the patient's heavy chain variable domain linked to either a γ3 or γ4 constant domain. As is understood by those in the art, the PCR product is subcloned into the expression vector using restriction enzymes which lack sites internal to the PCR product (*i.e.*, within the Ig sequences). The PCR products are digested with restriction enzymes that have sites located within the PCR primers to confirm that the PCR product lacks an internal site for a given restriction enzyme prior to subcloning the PCR product into the desired expression vector. It is anticipated that the 5' *Mlu*I site can be employed for each PCR product given that *Mlu*I sites are very infrequently found in the genome; however the 5' primers also contain *Xba*I and *Eco*RI sites in the event a particular PCR product contains an internal *Mlu*I site. The following restriction enzymes (which have recognition sites in the above-described 3' PCR primers) are examined first for their inability to cut internally to the PCR products: *Sal*I for heavy chain PCR products; *Afl*II for kappa light chain PCR products; *Avr*II for lambda light chain PCR products. As discussed above, each 3' PCR primer provides alternative restriction enzyme sites.

With regard to choosing an expression vector, the pSRαSD9CG3C vector is initially chosen as Cγ3 is the least frequently used isotype in humans (Cγ4 is the next least frequently utilized isotype, with Cγ1 and Cγ2 being the most frequently used isotypes) and therefore ELISAs performed following immunization with a vaccine comprising Cγ3 are easier to

conduct and interpret as the patient's anti-idiotype response will mainly consist of the γ1 and γ2 isotypes. However, Cγ4 may be chosen over Cγ3 if a given Cγ3 construct produces an Ig protein which tends to fall out of solution upon purification.

For each tumor sample, five Vκ PCR reactions are run. Each Vκ PCR reaction will contain the Cκ primer (SEQ ID NO:70). The VκI, VκII, and VκIII primers (SEQ ID NOS:63-65, respectively) will be run in separate reactions. The VκIV and VκV primers (SEQ ID NOS:66 and 67, respectively) are combined in one PCR reaction and the VκVIa and VκVIb primers (SEQ ID NOS:68 and 69, respectively) in another. The PCR reaction which yields a PCR product of the expected size (~ 480 bp) is used as the source of DNA encoding the variable domain derived from the light chain of the patient's Ig. The positive reaction product is subcloned into the pSRαSD9CKC vector using the 5' *Xba*I, *Eco*RI or *Mlu*I sites and the 3' *Afl*II or *Bsp*EI sites.

For each tumor sample, six Vλ PCR reactions are run. Each Vκ PCR reaction will contain the Cλ primer (SEQ ID NO:77). The VL1, VL2, VL3, VL4, VL6 and VL8 primers (SEQ ID NOS:71-76, respectively) are used in separate reactions. The PCR reaction which yields a PCR product of the expected size (~ 420 bp) is used as the source of DNA encoding the variable domain derived from the light chain of the patient's Ig. The positive reaction product will be subcloned into the pSRαSD9CL2C vector using the 5' *Xba*I, *Eco*RI or *Mlu*I sites and the 3' *Avr*II or *Kas*I sites. It is understood by those skilled in the art that the tumor cells will express either a κ or a λ light chain. Therefore, it is expected that a PCR product will be recovered from either the Vκ or Vλ PCRs but not from both.

e) Expression and Amplification of Tumor-Specific Ig in Mammalian Cells

Once expression vectors containing sequences derived from the variable regions of the heavy and light chains found in the patient's tumor are constructed, these plasmids are used to transform *E. coli* using conventional techniques. Between 18 and 24 colonies from each subcloning are screened for heavy and light chain inserts as appropriate by restriction enzyme analysis of miniprep DNA (from 1-1.5 ml cultures). Equal aliquots of the positive subclones are used to inoculate larger cultures (~250 mls) from which the DNA for electroporation is prepared. This allows for the isolation of the somatic variants in the tumor population and result in transfectants (*e.g.*, BW5147.G.1.4 transfectants) expressing these somatic variants.

To further define the presence of somatic variants, 20 μl PCR reactions are run using ~100 pg of each miniprep DNA and the appropriate V region and C region primers. Digestion

of the resulting PCR products with several four base recognition restriction enzymes allows the differentiation of somatic variants. In addition, DNA sequencing can be performed on individual subclones to demonstrate the presence of somatic variants within the pool of subclones containing the cloned heavy and light chain variable regions.

Plasmids encoding the chimeric heavy and light chains derived from the patient's Ig are electroporated along with pM-HPRT-SSD9-DHFR into BW5147.G.1.4 cells as follows. The Ig expression plasmids (which comprise a mixture of vectors containing the somatic variants found within the tumor Ig) are linearized by digestion with *Asc*I or *Pac*I. pM-HPRT-SSD9-DHFR is linearized with *Asc*I or *Pac*I. pM-HPRT-SSD9-DHFR and the Ig expression plasmids are used at a ratio of 1:20-50. Approximately 15 μg of pM-HPRT-SSD9-DHFR (10-20 μg) is used while a total of ~500 μg of the expression vectors are used. The linearized plasmids are digested, precipitated and resuspended in 0.5 ml electroporation buffer [*i.e.*, 1X HBS(EP)] as described in Example 7. The linearized plasmids in 0.5 ml electroporation buffer are mixed with 2 x 10^7 cells (*e.g.*, BW5147.G.1.4) in 0.5 ml electroporation buffer and electroporated as described in Example 7. The cells are then grown in selective medium followed by growth in medium containing MTX as described in Examples 7 and 8. Clones which grow in the selective medium are checked for the ability to express the cloned Ig proteins using standard methods (*e.g.*, by ELISA). Primary selectants expressing high levels of the cloned Ig proteins are then grown in medium containing MTX as described in Examples 7 and 8 to amplify the transfected genes. The presence of the selectable and amplifiable markers on a single piece of DNA (*i.e.*, pM-HPRT-SSD9-DHFR), obviates concerns that primary transfectants (*i.e.*, cells capable of growing in medium containing Hx and Az) which express the genes of interest (*i.e.*, the Ig proteins) at high levels have failed to integrate a DHFR gene.

f) Purification of Tumor-Specific Ig From Amplified Cell Lines

The tumor-specific Ig expressed by the amplified cell lines (using either the pSRαSD9CG3C or pSRαSD9CG4C vectors) is purified by chromatography of culture supernatants on Protein G Sepharose™ (Pharmacia); Protein G binds to both IgG_3 and IgG_4. The chromatography is conducted according to the manufacturer's instructions. When the tumor-specific Ig is produced using the pSRαSD9CG4C vector, Protein A Sepharose (Pharmacia) may also be employed for purification.

g) Administration of Tumor-Specific Ig (Multivalent Vaccine)

The purified tumor immunoglobulin-idiotype protein may be conjugated to a protein carrier such as keyhole limpet hemocyanin (KLH) (Calbiochem, San Diego, CA) prior to administration to the patient. If the immunoglobulin-idiotype protein is to be conjugated with KLH, the KLH is depleted of endotoxin using methods known to the art [Kwak *et al.* (1992), *supra*]. For example, the KLH is applied to a QAE Zeta Prep™ 15 disk (LKB, Broma, Sweeden) to produce a preparation of KLH containing less than 1000 endotoxin units per milliliter. Equal volumes of filter sterilized purified KLH and purified immunoglobulin-idiotype protein (each at 1 mg/ml) are mixed together. Sterile glutaraldehyde is added at a final concentration of 0.1%. The Ig-KLH conjugate is then dialyzed extensively against physiologic saline to remove excess glutaraldehyde.

Purified immunoglobulin-idiotype protein (conjugated or unconjugated) is mixed with an immunologic adjuvant such as SAF-1™ (Syntex adjuvant formulation 1; Roche) or other adjuvant presently or subsequently approved for administration to humans [*e.g.*, QS-21 (PerImmune, Inc., Rockville, MD)]. The purified immunoglobulin-idiotype protein is emulsified in the desired adjuvant and injected subcutaneously at 0, 2, 6, 10 and 14 weeks. Booster injects may be given at 24 and 28 weeks. Each injection contains 0.5 mg of purified, tumor-specific idiotype immunoglobulin (which may be conjugated 1:1 with KLH).

An alternative to the use of KLH as a foreign carrier protein to boost the immune response to the immunoglobulin idiotype protein is the use of a fusion protein comprising idiotype protein and a cytokine (*e.g.*, GM-CSF, IL-2 or IL-4) [PCT International Application PCT/US93/09895, Publication No. WO 94/08601 and Tao and Levy (1993) Nature 362:755 and Chen *et al.* (1994) J. Immunol. 153:4775]. In these fusion proteins, sequences encoding the desired cytokine are added to the 3' end of sequences encoding the immunoglobulin idiotype protein. The present invention contemplates the use of idiotype-cytokine fusion proteins for the treatment of B-cell lymphoma. The sequences encoding the heavy chain of the patient's immunoglobulin protein are cloned as described above and inserted into an expression vector containing sequences encoding the desired cytokine such that a fusion protein comprising, from amino- to carboxy-terminus, the heavy chain of the patient's tumor-specific immunoglobulin and the desired cytokine.

An alternative to the use of foreign carrier proteins, cytokines, or immunologic adjuvants is the use of autologous dendritic cells pulsed with the purified immunoglobulin-idiotype protein [see for example, Hsu *et al.* (1996), *supra* and PCT International Application

PCT/US91/01683. Publication No. WO 91/13632]. Methods for the isolation of human dendritic cells from peripheral blood are known to the art [Mehta *et al.* (1994) J. Immunol. 153:996 and Takamizawa *et al.* (1995) J. Clin. Invest. 95:296]. Briefly, the patient is leukapheresed using a cell separator (COBE). Peripheral blood mononuclear cells (PBMCs) are collected by separation through Ficoll-Hypaque™ (Pharmacia). Monocytes are then removed by centrifugation through discontinuous Percoll™ (Pharmacia) gradients. The monocyte-depleted PBMCs are then placed in medium (RPMI 1640 containing 10% autologous patient serum) containing idiotype protein (2 μg/ml). Following incubation for 24 hours at 37°C in a humidified atmosphere containing 10% CO_2, the dendritic cells are separated from lymphocytes by sequential centrifugation through 15% and 14% (wt/vol) metrizamide gradients. The preparation is then incubated for 14-18 hours in medium containing 50 μg/ml idiotype protein. The cells are then washed to remove free antigen (*i.e.*, idiotype protein) and placed in sterile saline containing 5% autologous serum and administered intravenously.

Each patient is followed to determine the production of idiotype-specific antibody; the *in vitro* proliferative responses (to KLH, if used, and to immunoglobulin idiotype using 0 to 100 μg of soluble protein per milliliter in 5 day *in vitro* cultures) of PBMCs isolated from the treated patients may also be determined. These assays are conducted immediately before each immunization and 1 to 2 months following the last immunization. Patients are monitored for disease activity by physical examination, routine laboratory studies and routine radiographic studies. Regression of lymph nodes or cutaneous lymphomatous masses may be confirmed by computed tomography (CT). In addition, residual disease may be measured using a tumor-specific CDR3 analysis as described by Hsu *et al.* (1996), *supra*.

h) Treatment of T-cell Tumors

Vaccines comprising soluble T cell receptor (TCR) proteins derived from a patient's T cell tumor (*i.e.*, a T cell leukemia or lymphoma) are produced using the methods described in Example 9 with the exception that pM-HPRT-SSD9-DHFR is used in place of separate selection and amplification vectors as described above. The thrombin solubilized TCR proteins are purified by chromatography on a resin comprising a monoclonal antibody (mcab) directed against a monomorphic determinant on human αβ TCRs [*e.g.*, mcab T10B9.1A-31 (Pharmingen, San Diego, CA); mcab BMA031 (Immunotech, Westbrook, ME); mcabs

BW242/412. 8A3 or 3A8 (Serotec, Washington, DC). Antibodies directed against monomorphic (*i.e.*, invariant) determinants on TCRs recognize all αβ TCRs.

The purified tumor-specific idiotype TCR protein is administered as described above for the purified tumor-specific idiotype Ig protein (*i.e.*, mixing with an immunologic adjuvant, conjugation to a protein carrier, the use of TCR-cytokine fusion proteins, the use of dendritic cells pulsed with the purified TCR protein such as SAF-1, etc.). Patients are followed to determine the production of idiotype-specific antibody as described above. Patients are monitored for disease activity by physical examination, routine laboratory studies and routine radiographic studies.

From the above, it is clear that the present invention provides improved methods for the amplification and expression of recombinant genes in cells. The resulting amplified cell lines provide large quantities of recombinant proteins in a short period of time. The ability to produce large quantities of recombinant proteins in a short period of time is particularly advantageous when proteins unique to a patient's tumors are to be used for therapeutic purposes, such as for vaccination.

Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: GENITOPE CORPORATION

(ii) TITLE OF INVENTION: VACCINES FOR TREATMENT OF LYMPHOMA AND LEUKEMIA

(iii) NUMBER OF SEQUENCES: 81

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SMART & BIGGAR

(B) STREET: P.O. BOX 2999, STATION D

(C) CITY: OTTAWA

(D) STATE: ONT

(E) COUNTRY: CANADA

(F) ZIP: K1P 5Y6

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: ASCII (text)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: CA 2,248,653

(B) FILING DATE: 25-APR-1997

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/644,664

(B) FILING DATE: 01-MAY-1996

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/761,277

(B) FILING DATE: 06-DEC-1996

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: SMART & BIGGAR

(B) REGISTRATION NUMBER:

(C) REFERENCE/DOCKET NUMBER: 74667-99

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (613)-232-2486

(B) TELEFAX: (613)-232-8440

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCTAGAGCGG CCGCGGAGGC CGAATTCG 28

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GATCCGAATT CGGCCTCCGC GGCCGCTCTA GATGCA 36

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 677 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGATCCAGAC ATGATAAGAT ACATTGATGA GTTTGGACAA ACCACAACTA GAATGCAGTG 60

AAAAAAATGC TTTATTTGTG AAATTTGTGA TGCTATTGCT TTATTTGTAA CCATTATAAG 120

CTGCAATAAA CAAGTTAACA ACAACAATTG CATTCATTTT ATGTTTCAGG TTCAGGGGGA 180

GGTGTGGGAG GTTTTTTAAA GCAAGTAAAA CCTCTACAAA TGTGGTATGG CTGATTATGA 240

TCATGAACAG ACTGTGAGGA CTGAGGGGCC TGAAATGAGC CTTGGGACTG TGAATCAATG 300

CCTGTTTCAT GCCCTGAGTC TTCCATGTTC TTCTCCCCAC CATCTTCATT TTTATCAGCA 360

TTTTCCTGGC TGTCTTCATC ATCATCATCA CTGTTTCTTA GCCAATCTAA AACTCCAATT 420

CCCATAGCCA CATTAAACTT CATTTTTTGA TACACTGACA AACTAAACTC TTTGTCCAAT 480

CTCTCTTTCC ACTCCACAAT TCTGCTCTGA ATACTTTGAG CAAACTCAGC CACAGGTCTG 540

TACCAAATTA ACATAAGAAG CAAAGCAATG CCACTTTGAA TTATTCTCTT TTCTAACAAA 600

AACTCACTGC GTTCCAGGCA ATGCTTTAAA TAATCTTTGG GCCTAAAATC TATTTGTTTT 660

ACAAATCTGG CCTGCAG 677

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CTAGAATTCA CGCGTAGGCC TCCGCGGCCG CGCGCATGC 39

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AATTGCATGC GCGCGGCCGC GGAGGCCTAC GCGTGAATT 39

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 633 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CAAGCTTGCT GTGGAATGTG TGTCAGTTAG GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG 60

GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CAGGTGTGGA AAGTCCCCAG 120

GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA ACCATAGTCC 180

CGCCCCTAAC TCCGCCCATC CCGCCCCTAA CTCCGCCCAG TTCCGCCCAT TCTCCGCCCC 240

ATGGCTGACT AATTTTTTTT ATTTATGCAG AGGCCGAGGC CGCCTCGGCC TCTGAGCTAT 300

TCCAGAAGTA GTGAGGAGGC TTTTTTGGAG GCCTAGGCTT TTGCAAAAAG CTCCTCGAGC 360

TCGCATCTCT CCTTCACGCG CCCGCCGCCC TACCTGAGGC CGCCATCCAC GCCGGTTGAG 420

TCGCGTTCTG CCGCCTCCCG CCTGTGGTGC CTCCTGAACT GCGTCCGCCG TCTAGGTAAG 480

TTTAGAGCTC AGGTCGAGAC CGGGCCTTTG TCCGGCGCTC CCTTGGAGCC TACCTAGACT 540

CAGCCGGCTC TCCACGCTTT GCCTGACCCT GCTTGCTCAA CTCTACGTCT TTGTTTCGTT 600

TTCTGTTCTG CGCCGTTACA GATCGCCTCG AGG 633

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 635 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CAAGCTTGCG ATTAGTCCAA TTTGTTAAAG ACAGGATATC AGTGGTCCAG GCTCTAGTTT 60

TGACTCAACA ATATCACCAG CTGAAGCCTA TAGAGTACGA GCCATAGATA AAATAAAAGA 120

TTTTATTTAG TCTCCAGAAA AAGGGGGGAA TGAAAGACCC CACCTGTAGG TTTGGCAAGC 180

TAGCTTAAGT AACGCCATTT TGCAAGGCAT GGAAAAATAC ATAACTGAGA ATAGAGAAGT 240

TCAGATCAAG GTCAGGAACA GATGGAACAG CTGAATATGG GCCAAACAGG ATATCTGTGG 300

TAAGCAGTTC CTGCCCCGGC TCAGGGCCAA GAACAGATGG AACAGCTGAA TATGGGCCAA 360

ACAGGATATC TGTGGTAAGC AGTTCCTGCC CCGGCTCAGG GCCAAGAACA GATGGTCCCC 420

AGATGCGGTC CAGCCCTCAG CAGTTTCTAG AGAACCATCA GATGTTTCCA GGGTGCCCCA 480

AGGACCTGAA ATGACCCTGT GCCTTATTTG AACTAACCAA TCAGTTCGCT TCTCGCTTCT 540

GTTCGCGCGC TTCTGCTCCC CGAGCTCAAT AAAAGAGCCC ACAACCCCTC ACTCGGGGCG 600

CCAGTCCTCC GATTGACTGA GTCGCCCCCT CGAGG 635

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 483 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AAGCTTTGGA GCTAAGCCAG CAATGGTAGA GGGAAGATTC TGCACGTCCC TTCCAGGCGG 60

CCTCCCCGTC ACCACCCCCC CCAACCCGCC CCGACCGGAG CTGAGAGTAA TTCATACAAA 120

AGGACTCGCC CCTGCCTTGG GGAATCCCAG GGACCGTCGT TAAACTCCCA CTAACGTAGA 180

ACCCAGAGAT CGCTGCGTTC CCGCCCCCTC ACCCGCCCGC TCTCGTCATC ACTGAGGTGG 240

AGAAGAGCAT GCGTGAGGCT CCGGTGCCCG TCAGTGGGCA GAGCGCACAT CGCCCACAGT 300

CCCCGAGAAG TTGGGGGGAG GGGTCGGCAA TTGAACCGGT GCCTAGAGAA GGTGGCGCGG 360

GGTAAACTGG GAAAGTGATG TCGTGTACTG GCTCCGCCTT TTTCCCGAGG GTGGGGGAGA 420

ACCGTATATA AGTGCAGTAG TCGCCGTGAA CGTTCTTTTT CGCAACGGGT TTGCCGCCTC 480

GAG 483

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AAGCTTTGGA GCTAAGCCAG CAAT 24

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTCGAGGCGG CAAACCCGTT GCG 23

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1451 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AAGCTTTGGA GCTAAGCCAG CAATGGTAGA GGGAAGATTC TGCACGTCCC TTCCAGGCGG 60

CCTCCCCGTC ACCACCCCCC CCAACCCGCC CCGACCGGAG CTGAGAGTAA TTCATACAAA 120

AGGACTCGCC CCTGCCTTGG GGAATCCCAG GGACCGTCGT TAAACTCCCA CTAACGTAGA 180

ACCCAGAGAT CGCTGCGTTC CCGCCCCCTC ACCCGCCCGC TCTCGTCATC ACTGAGGTGG 240

AGAAGAGCCA TGCGTGAGGC TCCGGTGCCC GTCAGTGGGC AGAGCGCACA TCGCCCACAG 300

TCCCCGAGAA GTTGGGGGGA GGGGTCGGCA ATTGAACCGG TGCCTAGAGA AGGTGGCGCG 360

GGGTAAACTG GGAAAGTGAT GTCGTGTACT GGCTCCGCCT TTTTCCCGAG GGTGGGGGAG 420

AACCCGTATA TAAGTGCAGT AGTCGCCGTG AACGTTCTTT TTCGCAACGG GTTTGCCGCC 480

AGAACACAGG TAAGTGCCGT GTGTGGTTCC CGCGGGCCTG GCCTCTTTAC GGGTTATGGC 540

CCTTGCGTGC CTTGAATTAC TTCCACGCCC CTGGCTGCAG TACGTGATTC TTGATCCCGA 600

GCTTCGGGTT GGAAGTGGGT GGGAGAGTTC GAGGCCTTGC GCTTAAGGAG CCCCTTCGCC 660

TCGTGCTTGA GTTGAGGCCT GGCCTGGGCG CTGGGGCCCC CGCGTGCGAA TCTGGTGGCA 720

CCTTCGCGCC TGTCTCGCTG CTTTCGATAA GTCTCTAGCC ATTTAAAATT TTTGATGACC 780

TGCTGCGACG CTTTTTTTCT GGCAAGATAG TCTTGTAAAT GCGGGCCAAG ATCTGCACAC 840

TGGTATTTCG GTTTTTGGGG CCGCGGGCGG CGACGGGGCC CGTGCGTCCC AGCGCACATG 900

TTCGGCGAGG CGGGGCCTGC GAGCGCGGCC ACCGAGAATC GGACGGGGGT AGTCTCAAGC 960

TGGCCGGCCT GCTCTGGTGC CTGGCCTCGC GCCGCCGTGT ATCGCCCCGC CCTGGGCGGC 1020

AAGGCTGGCC CGGTCGGCAC CAGTTGCGTG AGCGGAAAGA TGGCCGCTTC CCGGCCCTGC 1080

TGCAGGGAGC TCAAAATGGA GGACGCGGCG CTCGGGAGAG CGGGCGGGTG AGTCACCCAC 1140

ACAAAGGAAA AGGGCCTTTC CGTCCTCAGC CGTCGCTTCA TGTGACTCCA CGGAGTACCG 1200

GGCGCCGTCC AGGCACCTCG ATTAGTTCTC GAGCTTTTGG AGTACGTCGT CTTTAGGTTG 1260

GGGGGAGGGG TTTTATGCGA TGGAGTTTCC CCACACTGAG TGGGTGGAGA CTGAAGTTAG 1320

GCCAGCTTGG CACTTGATGT AATTCTCCTT GGAATTTGCC CTTTTTGAGT TTGGATCTTG 1380

GTTCATTCTC AAGCCTCAGA CAGTGGTTCA AAGTTTTTTT CTTCCATTTC AGGTGTCGTG 1440

AAAACTCTAG A 1451

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TCTAGAGTTT TCACGACACC TGA 23

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1289 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 88..741

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
TTACCTCACT GCTTTCCGGA GCGGTAGCAC CTCCTCCGCC GGCTTCCTCC TCAGACCGCT      60

TTTTGCCGCG AGCCGACCGG TCCCGTC ATG CCG ACC CGC AGT CCC AGC GTC          111
                                 Met Pro Thr Arg Ser Pro Ser Val
                                   1               5

GTG ATT AGC GAT GAT GAA CCA GGT TAT GAC CTA GAT TTG TTT TGT ATA        159
Val Ile Ser Asp Asp Glu Pro Gly Tyr Asp Leu Asp Leu Phe Cys Ile
     10                  15                  20

CCT AAT CAT TAT GCC GAG GAT TTG GAA AAA GTG TTT ATT CCT CAT GGA        207
Pro Asn His Tyr Ala Glu Asp Leu Glu Lys Val Phe Ile Pro His Gly
 25                  30                  35                  40

CTG ATT ATG GAC AGG ACT GAA AGA CTT GCT CGA GAT GTC ATG AAG GAG        255
Leu Ile Met Asp Arg Thr Glu Arg Leu Ala Arg Asp Val Met Lys Glu
                 45                  50                  55

ATG GGA GGC CAT CAC ATT GTG GCC CTC TGT GTG CTC AAG GGG GGC TAT        303
Met Gly Gly His His Ile Val Ala Leu Cys Val Leu Lys Gly Gly Tyr
             60                  65                  70

AAG TTC TTT GCT GAC CTG CTG GAT TAC ATT AAA GCA CTG AAT AGA AAT        351
Lys Phe Phe Ala Asp Leu Leu Asp Tyr Ile Lys Ala Leu Asn Arg Asn
         75                  80                  85

AGT GAT AGA TCC ATT CCT ATG ACT GTA GAT TTT ATC AGA CTG AAG AGC        399
Ser Asp Arg Ser Ile Pro Met Thr Val Asp Phe Ile Arg Leu Lys Ser
     90                  95                  100

TAC TGT AAT GAT CAG TCA ACG GGG GAC ATA AAA GTT ATT GGT GGA GAT        447
Tyr Cys Asn Asp Gln Ser Thr Gly Asp Ile Lys Val Ile Gly Gly Asp
105                 110                 115                 120

GAT CTC TCA ACT TTA ACT GGA AAG AAT GTC TTG ATT GTT GAA GAT ATA        495
Asp Leu Ser Thr Leu Thr Gly Lys Asn Val Leu Ile Val Glu Asp Ile
                125                 130                 135
```

```
ATT GAC ACT GGT AAA ACA ATG CAA ACT TTG CTT TCC CTG GTT AAG CAG        543
Ile Asp Thr Gly Lys Thr Met Gln Thr Leu Leu Ser Leu Val Lys Gln
            140                 145                 150

TAC AGC CCC AAA ATG GTT AAG GTT GCA AGC TTG CTG GTG AAA AGG ACC        591
Tyr Ser Pro Lys Met Val Lys Val Ala Ser Leu Leu Val Lys Arg Thr
        155                 160                 165

TCT CGA AGT GTT GGA TAC AGG CCA GAC TTT GTT GGA TTT GAA ATT CCA        639
Ser Arg Ser Val Gly Tyr Arg Pro Asp Phe Val Gly Phe Glu Ile Pro
    170                 175                 180

GAC AAG TTT GTT GTT GGA TAT GCC CTT GAC TAT AAT GAG TAC TTC AGG        687
Asp Lys Phe Val Val Gly Tyr Ala Leu Asp Tyr Asn Glu Tyr Phe Arg
185                 190                 195                 200

AAT TTG AAT CAC GTT TGT GTC ATT AGT GAA ACT GGA AAA GCC AAA TAC        735
Asn Leu Asn His Val Cys Val Ile Ser Glu Thr Gly Lys Ala Lys Tyr
                205                 210                 215

AAA GCC TAAGATGAGC GCAAGTTGAA TCTGCAAATA CGAGGAGTCC TGTTGATGTT        791
Lys Ala

GCCAGTAAAA TTAGCAGGTG TTCTAGTCCT GTGGCCATCT GCCTAGTAAA GCTTTTTGCA     851

TGAACCTTCT ATGAATGTTA CTGTTTTATT TTTAGAAATG TCAGTTGCTG CGTCCCCAGA     911

CTTTTGATTT GCACTATGAG CCTATAGGCC AGCCTACCCT CTGGTAGATT GTCGCTTATC     971

TTGTAAGAAA AACAAATCTC TTAAATTACC ACTTTTAAAT AATAATACTG AGATTGTATC    1031

TGTAAGAAGG ATTTAAAGAG AAGCTATATT AGTTTTTTAA TTGGTATTTT AATTTTTATA    1091

TATTCAGGAG AGAAAGATGT GATTGATATT GTTAATTTAG ACGAGTCTGA AGCTCTCGAT    1151

TTCCTATCAG TAACAGCATC TAAGAGGTTT TGCTCAGTGG AATAAACATG TTTCAGCAGT    1211

GTTGGCTGTA TTTTCCCACT TTCAGTAAAT CGTTGTCAAC AGTTCCTTTT AAATGCAAAT    1271

AAATAAATTC TAAAAATT                                                  1289
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 218 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
Met Pro Thr Arg Ser Pro Ser Val Val Ile Ser Asp Asp Glu Pro Gly
 1               5                  10                  15

Tyr Asp Leu Asp Leu Phe Cys Ile Pro Asn His Tyr Ala Glu Asp Leu
            20                  25                  30

Glu Lys Val Phe Ile Pro His Gly Leu Ile Met Asp Arg Thr Glu Arg
        35                  40                  45

Leu Ala Arg Asp Val Met Lys Glu Met Gly Gly His His Ile Val Ala
    50                  55                  60

Leu Cys Val Leu Lys Gly Gly Tyr Lys Phe Phe Ala Asp Leu Leu Asp
65                  70                  75                  80
```

Tyr Ile Lys Ala Leu Asn Arg Asn Ser Asp Arg Ser Ile Pro Met Thr
85 90 95

Val Asp Phe Ile Arg Leu Lys Ser Tyr Cys Asn Asp Gln Ser Thr Gly
100 105 110

Asp Ile Lys Val Ile Gly Gly Asp Asp Leu Ser Thr Leu Thr Gly Lys
115 120 125

Asn Val Leu Ile Val Glu Asp Ile Ile Asp Thr Gly Lys Thr Met Gln
130 135 140

Thr Leu Leu Ser Leu Val Lys Gln Tyr Ser Pro Lys Met Val Lys Val
145 150 155 160

Ala Ser Leu Leu Val Lys Arg Thr Ser Arg Ser Val Gly Tyr Arg Pro
165 170 175

Asp Phe Val Gly Phe Glu Ile Pro Asp Lys Phe Val Val Gly Tyr Ala
180 185 190

Leu Asp Tyr Asn Glu Tyr Phe Arg Asn Leu Asn His Val Cys Val Ile
195 200 205

Ser Glu Thr Gly Lys Ala Lys Tyr Lys Ala
210 215

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCATGCGCGC GGCCGCGGAG GCTTTTTTTT TTTTTTTTTT 40

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGGCAACGCG TGCCATCATG GTTCGAC 27

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGCAGCGGC CGCATAGATC TAAAGCCAGC 30

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 671 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..573

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
ACGCGTGCCA TC ATG GTT CGA CCA TTG AAC TGC ATC GTC GCC GTG TCC          48
              Met Val Arg Pro Leu Asn Cys Ile Val Ala Val Ser
                1               5                  10

CAA AAT ATG GGG ATT GGC AAG AAC GGA GAC CTA CCC TGG CCT CCG CTC          96
Gln Asn Met Gly Ile Gly Lys Asn Gly Asp Leu Pro Trp Pro Pro Leu
        15                  20                  25

AGG AAC GAG TTC AAG TAC TTC CAA AGA ATG ACC ACA ACC TCT TCA GTG         144
Arg Asn Glu Phe Lys Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser Val
     30                  35                  40

GAA GGT AAA CAG AAT CTG GTG ATT ATG GGT AGG AAA ACC TGG TTC TCC         192
Glu Gly Lys Gln Asn Leu Val Ile Met Gly Arg Lys Thr Trp Phe Ser
 45                  50                  55                  60

ATT CCT GAG AAG AAT CGA CCT TTA AAG GAC AGA ATT AAT ATA GTT CTC         240
Ile Pro Glu Lys Asn Arg Pro Leu Lys Asp Arg Ile Asn Ile Val Leu
                 65                  70                  75

AGT AGA GAA CTC AAA GAA CCA CCA CGA GGA GCT CAT TTT CTT GCC AAA         288
Ser Arg Glu Leu Lys Glu Pro Pro Arg Gly Ala His Phe Leu Ala Lys
             80                  85                  90

AGT TTG GAT GAT GCC TTA AGA CTT ATT GAA CAA CCG GAA TTG GCA AGT         336
Ser Leu Asp Asp Ala Leu Arg Leu Ile Glu Gln Pro Glu Leu Ala Ser
         95                  100                 105

AAA GTA GAC ATG GTT TGG ATA GTC GGA GGC AGT TCT GTT TAC CAG GAA         384
Lys Val Asp Met Val Trp Ile Val Gly Gly Ser Ser Val Tyr Gln Glu
     110                 115                 120

GCC ATG AAT CAA CCA GGC CAC CTT AGA CTC TTT GTG ACA AGG ATC ATG         432
Ala Met Asn Gln Pro Gly His Leu Arg Leu Phe Val Thr Arg Ile Met
 125                 130                 135                 140

CAG GAA TTT GAA AGT GAC ACG TTT TTC CCA GAA ATT GAT TTG GGG AAA         480
Gln Glu Phe Glu Ser Asp Thr Phe Phe Pro Glu Ile Asp Leu Gly Lys
                 145                 150                 155

TAT AAA CTT CTC CCA GAA TAC CCA GGC GTC CTC TCT GAG GTC CAG GAG         528
Tyr Lys Leu Leu Pro Glu Tyr Pro Gly Val Leu Ser Glu Val Gln Glu
             160                 165                 170

GAA AAA GGC ATC AAG TAT AAG TTT GAA GTC TAC GAG AAG AAA GAC             573
Glu Lys Gly Ile Lys Tyr Lys Phe Glu Val Tyr Glu Lys Lys Asp
         175                 180                 185
```

TAACAGGAAG ATGCTTTCAA GTTCTCTGCT CCCCTCCTAA AGCTATGCAT TTTTATAAGA 633

CCATGGGACT TTTGCTGGCT TTAGATCTAT GCGGCCGC 671

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 187 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Val Arg Pro Leu Asn Cys Ile Val Ala Val Ser Gln Asn Met Gly
1 5 10 15

Ile Gly Lys Asn Gly Asp Leu Pro Trp Pro Pro Leu Arg Asn Glu Phe
20 25 30

Lys Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser Val Glu Gly Lys Gln
35 40 45

Asn Leu Val Ile Met Gly Arg Lys Thr Trp Phe Ser Ile Pro Glu Lys
50 55 60

Asn Arg Pro Leu Lys Asp Arg Ile Asn Ile Val Leu Ser Arg Glu Leu
65 70 75 80

Lys Glu Pro Pro Arg Gly Ala His Phe Leu Ala Lys Ser Leu Asp Asp
85 90 95

Ala Leu Arg Leu Ile Glu Gln Pro Glu Leu Ala Ser Lys Val Asp Met
100 105 110

Val Trp Ile Val Gly Gly Ser Ser Val Tyr Gln Glu Ala Met Asn Gln
115 120 125

Pro Gly His Leu Arg Leu Phe Val Thr Arg Ile Met Gln Glu Phe Glu
130 135 140

Ser Asp Thr Phe Phe Pro Glu Ile Asp Leu Gly Lys Tyr Lys Leu Leu
145 150 155 160

Pro Glu Tyr Pro Gly Val Leu Ser Glu Val Gln Glu Glu Lys Gly Ile
165 170 175

Lys Tyr Lys Phe Glu Val Tyr Glu Lys Lys Asp
180 185

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATATATCTAG ACCACCATGC CTGGCTCAGC ACTG 34

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATTATTGCGG CCGCTTAGCT TTTCATTTTG ATCAT 35

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 134 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGTCTAGAGC CAAATAAAGG AAGTGGAACC ACTTCAGGTA CTACCCGTCT TCTATCTGGG 60

CACACGTGTT TCACGTTGAC AGGTTTGCTT GGGACGCTAG TAACCATGGG CTTGCTGACT 120

TAGGCATCGA ATTC 134

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 134 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GAATTCGATG CCTAAGTCAG CAAGCCCATG GTTACTAGCG TCCCAAGCAA ACCTGTCAAC 60

GTGAAACACG TGTGCCCAGA TAGAAGACGG GTAGTACCTG AAGTGGTTCC ACTTCCTTTA 120

TTTGGCTCTA GACC 134

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 300 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TAATACGACT CACTATAGGG CGAATTGGAG CTCCACCGCG GTGGCGGCCG CTCTAGAACT 60

AGTGGATCCC CCGGGCTGCA GGAATTCGAT GGTCTAGAGC CAAATAAAGG AAGTGGAACC 120

ACTTCAGGTA CTACCCGTCT TCTATCTGGG CACACGTGTT TCACGTTGAC AGGTTTGCTT 180

GGGACGCTAG TAACCATGGG CTTGCTGACT TAGGCATCGA ATTCATCAAG CTTATCGATA 240

CCGTCGACCT CGAGGGGGGG CCCGGTACCC AGCTTTTGTT CCCTTTAGTG AGGGTTAATT 300

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CCACTTCCTT TATTTGGGAG AGGGCTTG 28

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 747 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..744

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
ATG GCC ATA AGT GGA GTC CCT GTG CTA GGA TTT TTC ATC ATA GCT GTG      48
Met Ala Ile Ser Gly Val Pro Val Leu Gly Phe Phe Ile Ile Ala Val
 1               5                  10                  15

CTG ATG AGC GCT CAG GAA TCA TGG GCT ATC AAA GAA GAA CAT GTG ATC      96
Leu Met Ser Ala Gln Glu Ser Trp Ala Ile Lys Glu Glu His Val Ile
            20                  25                  30

ATC CAG GCC GAG TTC TAT CTG AAT CCT GAC CAA TCA GGC GAG TTT ATG     144
Ile Gln Ala Glu Phe Tyr Leu Asn Pro Asp Gln Ser Gly Glu Phe Met
        35                  40                  45

TTT GAC TTT GAT GGT GAT GAG ATT TTC CAT GTG GAT ATG GCA AAG AAG     192
Phe Asp Phe Asp Gly Asp Glu Ile Phe His Val Asp Met Ala Lys Lys
    50                  55                  60

GAG ACG GTC TGG CGG CTT GAA GAA TTT GGA CGA TTT GCC AGC TTT GAG     240
Glu Thr Val Trp Arg Leu Glu Glu Phe Gly Arg Phe Ala Ser Phe Glu
65                  70                  75                  80

GCT CAA GGT GCA TTG GCC AAC ATA GCT GTG GAC AAA GCC AAC TTG GAA     288
Ala Gln Gly Ala Leu Ala Asn Ile Ala Val Asp Lys Ala Asn Leu Glu
                85                  90                  95

ATC ATG ACA AAG CGC TCC AAC TAT ACT CCG ATC ACC AAT GTA CCT CCA     336
Ile Met Thr Lys Arg Ser Asn Tyr Thr Pro Ile Thr Asn Val Pro Pro
            100                 105                 110

GAG GTA ACT GTG CTC ACG AAC AGC CCT GTG GAA CTG AGA GAG CCC AAC     384
Glu Val Thr Val Leu Thr Asn Ser Pro Val Glu Leu Arg Glu Pro Asn
        115                 120                 125
```

```
GTC CTC ATC TGT TTC ATA GAC AAG TTC ACC CCA CCA GTG GTC AAT GTC        432
Val Leu Ile Cys Phe Ile Asp Lys Phe Thr Pro Pro Val Val Asn Val
    130                 135                 140

ACG TGG CTT CGA AAT GGA AAA CCT GTC ACC ACA GGA GTG TCA GAG ACA        480
Thr Trp Leu Arg Asn Gly Lys Pro Val Thr Thr Gly Val Ser Glu Thr
145                 150                 155                 160

GTC TTC CTG CCC AGG GAA GAC CAC CTT TTC CGC AAG TTC CAC TAT CTC        528
Val Phe Leu Pro Arg Glu Asp His Leu Phe Arg Lys Phe His Tyr Leu
                165                 170                 175

CCC TTC CTG CCC TCA ACT GAG GAC GTT TAC GAC TGC AGG GTG GAG CAC        576
Pro Phe Leu Pro Ser Thr Glu Asp Val Tyr Asp Cys Arg Val Glu His
            180                 185                 190

TGG GGC TTG GAT GAG CCT CTT CTC AAG CAC TGG GAG TTT GAT GCT CCA        624
Trp Gly Leu Asp Glu Pro Leu Leu Lys His Trp Glu Phe Asp Ala Pro
        195                 200                 205

AGC CCT CTC CCA AAT AAA GGA AGT GGA ACC ACT TCA GGT ACT ACC CGT        672
Ser Pro Leu Pro Asn Lys Gly Ser Gly Thr Thr Ser Gly Thr Thr Arg
    210                 215                 220

CTT CTA TCT GGG CAC ACG TGT TTC ACG TTG ACA GGT TTG CTT GGG ACG        720
Leu Leu Ser Gly His Thr Cys Phe Thr Leu Thr Gly Leu Leu Gly Thr
225                 230                 235                 240

CTA GTA ACC ATG GGC TTG CTG ACT TAG                                   747
Leu Val Thr Met Gly Leu Leu Thr
                245
```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 248 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
Met Ala Ile Ser Gly Val Pro Val Leu Gly Phe Phe Ile Ile Ala Val
 1               5                  10                  15

Leu Met Ser Ala Gln Glu Ser Trp Ala Ile Lys Glu Glu His Val Ile
            20                  25                  30

Ile Gln Ala Glu Phe Tyr Leu Asn Pro Asp Gln Ser Gly Glu Phe Met
        35                  40                  45

Phe Asp Phe Asp Gly Asp Glu Ile Phe His Val Asp Met Ala Lys Lys
    50                  55                  60

Glu Thr Val Trp Arg Leu Glu Glu Phe Gly Arg Phe Ala Ser Phe Glu
65                  70                  75                  80

Ala Gln Gly Ala Leu Ala Asn Ile Ala Val Asp Lys Ala Asn Leu Glu
                85                  90                  95

Ile Met Thr Lys Arg Ser Asn Tyr Thr Pro Ile Thr Asn Val Pro Pro
            100                 105                 110

Glu Val Thr Val Leu Thr Asn Ser Pro Val Glu Leu Arg Glu Pro Asn
        115                 120                 125
```

```
Val Leu Ile Cys Phe Ile Asp Lys Phe Thr Pro Pro Val Val Asn Val
    130                 135                 140

Thr Trp Leu Arg Asn Gly Lys Pro Val Thr Thr Gly Val Ser Glu Thr
145                 150                 155                 160

Val Phe Leu Pro Arg Glu Asp His Leu Phe Arg Lys Phe His Tyr Leu
                165                 170                 175

Pro Phe Leu Pro Ser Thr Glu Asp Val Tyr Asp Cys Arg Val Glu His
            180                 185                 190

Trp Gly Leu Asp Glu Pro Leu Leu Lys His Trp Glu Phe Asp Ala Pro
        195                 200                 205

Ser Pro Leu Pro Asn Lys Gly Ser Gly Thr Thr Ser Gly Thr Thr Arg
    210                 215                 220

Leu Leu Ser Gly His Thr Cys Phe Thr Leu Thr Gly Leu Leu Gly Thr
225                 230                 235                 240

Leu Val Thr Met Gly Leu Leu Thr
                245
```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```
CCACTTCCTT TATTTGGTGC AGATTCAG                                        28
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 786 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..783

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
ATG GTG TGT CTG AAG CTC CCT GGA GGC TCC TGC ATG ACA GCG CTG ACA        48
Met Val Cys Leu Lys Leu Pro Gly Gly Ser Cys Met Thr Ala Leu Thr
 1               5                  10                  15

GTG ACA CTG ATG GTG CTG AGC TCC CGA CTG GCT TTG GCT GGG GAC ACC        96
Val Thr Leu Met Val Leu Ser Ser Arg Leu Ala Leu Ala Gly Asp Thr
            20                  25                  30

CGA CCA CGT TTC TTG TGG CAG CTT AAG TTT GAA TGT CAT TTC TTC AAT       144
Arg Pro Arg Phe Leu Trp Gln Leu Lys Phe Glu Cys His Phe Phe Asn
        35                  40                  45
```

```
GGG ACG GAG CGG GTG CGG TTG CTG GAA AGA TGC ATC TAT AAC CAA GAG        192
Gly Thr Glu Arg Val Arg Leu Leu Glu Arg Cys Ile Tyr Asn Gln Glu
     50                  55                  60

GAG TCC GTG CGC TTC GAC AGC GAC GTG GGG GAG TAC CGG GCG GTT GAG        240
Glu Ser Val Arg Phe Asp Ser Asp Val Gly Glu Tyr Arg Ala Val Glu
 65                  70                  75                  80

GAG CTG GGG CGG CCT GAT GCC GAG TAC TGG AAC AGC CAG AAG GAC CTC        288
Glu Leu Gly Arg Pro Asp Ala Glu Tyr Trp Asn Ser Gln Lys Asp Leu
                 85                  90                  95

CTG GAG CAG AAG CGG GGC CAG GTG GAC AAT TAC TGC AGA CAC AAC TAC        336
Leu Glu Gln Lys Arg Gly Gln Val Asp Asn Tyr Cys Arg His Asn Tyr
            100                 105                 110

GGG GTT GGT GAG AGC TTC ACA GTG CAG CGG CGA GTT GAG CCT AAG GTG        384
Gly Val Gly Glu Ser Phe Thr Val Gln Arg Arg Val Glu Pro Lys Val
        115                 120                 125

ACT GTG TAT CCT TCA AAG ACC CAG CCC CTG CAG CAC CAC AAC CTC CTG        432
Thr Val Tyr Pro Ser Lys Thr Gln Pro Leu Gln His His Asn Leu Leu
    130                 135                 140

GTC TGC TCT GTG AGT GGT TTC TAT CCA GGC AGC ATT GAA GTC AGG TGG        480
Val Cys Ser Val Ser Gly Phe Tyr Pro Gly Ser Ile Glu Val Arg Trp
145                 150                 155                 160

TTC CGG AAC GGC CAG GAA GAG AAG GCT GGG GTG GTG TCC ACG GGC CTG        528
Phe Arg Asn Gly Gln Glu Glu Lys Ala Gly Val Val Ser Thr Gly Leu
                165                 170                 175

ATC CAG AAT GGA GAT TGG ACC TTC CAG ACC CTG GTG ATG CTG GAA ATA        576
Ile Gln Asn Gly Asp Trp Thr Phe Gln Thr Leu Val Met Leu Glu Ile
            180                 185                 190

GTT CCT CGG AGT GGA GAG GTT TAC ACC TGC CAA GTG GAG CAC CCA AGT        624
Val Pro Arg Ser Gly Glu Val Tyr Thr Cys Gln Val Glu His Pro Ser
        195                 200                 205

GTG ACG AGC CCT CTC ACA GTG GAA TGG AGA GCA CGG TCT GAA TCT GCA        672
Val Thr Ser Pro Leu Thr Val Glu Trp Arg Ala Arg Ser Glu Ser Ala
    210                 215                 220

CCA AAT AAA GGA AGT GGA ACC ACT TCA GGT ACT ACC CGT CTT CTA TCT        720
Pro Asn Lys Gly Ser Gly Thr Thr Ser Gly Thr Thr Arg Leu Leu Ser
225                 230                 235                 240

GGG CAC ACG TGT TTC ACG TTG ACA GGT TTG CTT GGG ACG CTA GTA ACC        768
Gly His Thr Cys Phe Thr Leu Thr Gly Leu Leu Gly Thr Leu Val Thr
                245                 250                 255

ATG GGC TTG CTG ACT TAG                                               786
Met Gly Leu Leu Thr
            260
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 261 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Val Cys Leu Lys Leu Pro Gly Gly Ser Cys Met Thr Ala Leu Thr
1 5 10 15

Val Thr Leu Met Val Leu Ser Ser Arg Leu Ala Leu Ala Gly Asp Thr
20 25 30

Arg Pro Arg Phe Leu Trp Gln Leu Lys Phe Glu Cys His Phe Phe Asn
35 40 45

Gly Thr Glu Arg Val Arg Leu Leu Glu Arg Cys Ile Tyr Asn Gln Glu
50 55 60

Glu Ser Val Arg Phe Asp Ser Asp Val Gly Glu Tyr Arg Ala Val Glu
65 70 75 80

Glu Leu Gly Arg Pro Asp Ala Glu Tyr Trp Asn Ser Gln Lys Asp Leu
85 90 95

Leu Glu Gln Lys Arg Gly Gln Val Asp Asn Tyr Cys Arg His Asn Tyr
100 105 110

Gly Val Gly Glu Ser Phe Thr Val Gln Arg Arg Val Glu Pro Lys Val
115 120 125

Thr Val Tyr Pro Ser Lys Thr Gln Pro Leu Gln His His Asn Leu Leu
130 135 140

Val Cys Ser Val Ser Gly Phe Tyr Pro Gly Ser Ile Glu Val Arg Trp
145 150 155 160

Phe Arg Asn Gly Gln Glu Glu Lys Ala Gly Val Val Ser Thr Gly Leu
165 170 175

Ile Gln Asn Gly Asp Trp Thr Phe Gln Thr Leu Val Met Leu Glu Ile
180 185 190

Val Pro Arg Ser Gly Glu Val Tyr Thr Cys Gln Val Glu His Pro Ser
195 200 205

Val Thr Ser Pro Leu Thr Val Glu Trp Arg Ala Arg Ser Glu Ser Ala
210 215 220

Pro Asn Lys Gly Ser Gly Thr Thr Ser Gly Thr Thr Arg Leu Leu Ser
225 230 235 240

Gly His Thr Cys Phe Thr Leu Thr Gly Leu Leu Gly Thr Leu Val Thr
245 250 255

Met Gly Leu Leu Thr
260

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 189 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..186

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
TTG GAT CCA CGA TCG TTT CTA TTG CGC AAT CCA AAT GAT AAG TAC GAA          48
Leu Asp Pro Arg Ser Phe Leu Leu Arg Asn Pro Asn Asp Lys Tyr Glu
  1               5                  10                  15

CCA TTT TGG GAA GAT ACT ACA GAG AAC GTG GTG TGT GCC CTG GGC CTG          96
Pro Phe Trp Glu Asp Thr Thr Glu Asn Val Val Cys Ala Leu Gly Leu
            20                  25                  30

ACT GTG GGT CTG GTG GGC ATC ATT ATT GGG ACC ATC TTC ATC ATC AAG         144
Thr Val Gly Leu Val Gly Ile Ile Ile Gly Thr Ile Phe Ile Ile Lys
        35                  40                  45

GGA GTG CGC AAA AGC AAT GCA GCA GAA CGC AGG GGG CCT CTG                  186
Gly Val Arg Lys Ser Asn Ala Ala Glu Arg Arg Gly Pro Leu
    50                  55                  60

TAA                                                                     189
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 62 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
Leu Asp Pro Arg Ser Phe Leu Leu Arg Asn Pro Asn Asp Lys Tyr Glu
  1               5                  10                  15

Pro Phe Trp Glu Asp Thr Thr Glu Asn Val Val Cys Ala Leu Gly Leu
            20                  25                  30

Thr Val Gly Leu Val Gly Ile Ile Ile Gly Thr Ile Phe Ile Ile Lys
        35                  40                  45

Gly Val Arg Lys Ser Asn Ala Ala Glu Arg Arg Gly Pro Leu
    50                  55                  60
```

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 192 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..189

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
TTG GAT CCA CGA TCG TTT CTA TTG CGC AAT CCA AAT GAT AAG TAC GAA          48
Leu Asp Pro Arg Ser Phe Leu Leu Arg Asn Pro Asn Asp Lys Tyr Glu
  1               5                  10                  15

CCA TTT TGG GAA GAT CAG AGC AAG ATG CTG AGT GGA GTC GGG GGC TTC          96
Pro Phe Trp Glu Asp Gln Ser Lys Met Leu Ser Gly Val Gly Gly Phe
            20                  25                  30
```

```
GTG CTG GGC CTG CTC TTC CTT GGG GCC GGG CTG TTC ATC TAC TTC AGG        144
Val Leu Gly Leu Leu Phe Leu Gly Ala Gly Leu Phe Ile Tyr Phe Arg
         35                  40                  45

AAT CAG AAA GGA CAC TCT GGA CTT CAG CCA ACA GGA TTC CTG AGC        189
Asn Gln Lys Gly His Ser Gly Leu Gln Pro Thr Gly Phe Leu Ser
     50                  55                  60

TGA                                                                192
```

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```
Leu Asp Pro Arg Ser Phe Leu Leu Arg Asn Pro Asn Asp Lys Tyr Glu
 1               5                  10                  15

Pro Phe Trp Glu Asp Gln Ser Lys Met Leu Ser Gly Val Gly Gly Phe
            20                  25                  30

Val Leu Gly Leu Leu Phe Leu Gly Ala Gly Leu Phe Ile Tyr Phe Arg
        35                  40                  45

Asn Gln Lys Gly His Ser Gly Leu Gln Pro Thr Gly Phe Leu Ser
    50                  55                  60
```

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```
CGATCGTGGA TCCAAGTTTA GGTTCGTATC TGTTTCAAA                          39
```

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```
CGATCGAGGA TCCAAGATGG TGGCAGACAG GACC                               34
```

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

ACGCGTCCAC CATGGCCATA AGTGGAGTCC CT 32

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GGATCCAACT CTGTAGTCTC TGGGAGAG 28

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

ACGCGTCCAC CATGGTGTGT CTGAAGCTCC TG 32

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GGATCCAACT TGCTCTGTGC AGATTCAGA 29

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 292 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GAATTCTTTT TTGCGTGTGG CAGTTTTAAG TTATTAGTTT TTAAAATCAG TACTTTTTAA 60

TGGAAACAAC TTGACCAAAA ATTTGTCACA GAATTTTGAG ACCCATTAAA AAAGTTAAAT 120

GAGAAACCTG TGTGTTCCTT TGGTCAACAC CGAGACATTT AGGTGAAAGA CATCTAATTC 180

TGGTTTTACG AATCTGGAAA CTTCTTGAAA ATGTAATTCT TGAGTTAACA CTTCTGGGTG 240

GAGAATAGGG TTGTTTTCCC CCCACATAAT TGGAAGGGGA AGGAATATCG AT 292

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TCGATGGCGC GCCTTAATTA 20

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AGCTTAATTA AGGCGCGCCA 20

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1147 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 7..1137

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GCGGCC GCG TCG ACC AAG GGC CCC AGC GTG TTC CCC CTG GCC CCC TGC 48
Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys
1 5 10

TCC CGC AGC ACC AGC GGC GGC ACC GCC GCC CTG GGC TGC CTG GTG AAG 96
Ser Arg Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys
15 20 25 30

GAC TAC TTC CCC GAG CCC GTG ACC GTG AGC TGG AAC AGC GGC GCC CTG 144
Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu
35 40 45

```
ACC AGC GGC GTC CAC ACC TTC CCC GCC GTG CTG CAG TCC AGC GGC CTG        192
Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu
                50                  55                  60

TAC TCC CTG AGC AGC GTG GTG ACC GTG CCC AGC AGC AGC CTG GGC ACC        240
Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr
            65                  70                  75

CAG ACC TAC ACC TGC AAC GTG AAC CAC AAG CCC AGC AAC ACC AAG GTG        288
Gln Thr Tyr Thr Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val
        80                  85                  90

GAC AAG CGC GTG GAG CTG AAG ACC CCC CTG GGC GAC ACC ACC CAC ACC        336
Asp Lys Arg Val Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr His Thr
 95                 100                 105                 110

TGC CCC CGC TGC CCC GAG CCC AAG AGC TGC GAC ACC CCT CCC CCC TGC        384
Cys Pro Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys
                115                 120                 125

CCC CGC TGC CCC GAG CCC AAG AGC TGC GAC ACC CCT CCC CCC TGC CCC        432
Pro Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro
            130                 135                 140

CGC TGC CCC GAG CCC AAG AGC TGC GAC ACC CCT CCC CCC TGC CCC CGC        480
Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
        145                 150                 155

TGC CCC GCC CCC GAG CTG CTG GGC GGC CCC AGC GTG TTC CTG TTC CCC        528
Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro
    160                 165                 170

CCC AAG CCC AAG GAC ACC CTG ATG ATC TCC CGC ACC CCC GAG GTG ACC        576
Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr
175                 180                 185                 190

TGC GTG GTG GTG GAC GTG AGC CAC GAG GAC CCC GAG GTG CAG TTC AAG        624
Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Gln Phe Lys
                195                 200                 205

TGG TAC GTG GAC GGC GTG GAG GTG CAT AAC GCC AAG ACC AAG CCC CGC        672
Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg
            210                 215                 220

GAG GAG CAG TAC AAC AGC ACC TTC CGC GTG GTG AGC GTG CTG ACC GTG        720
Glu Glu Gln Tyr Asn Ser Thr Phe Arg Val Val Ser Val Leu Thr Val
        225                 230                 235

CTG CAC CAG GAC TGG CTG AAC GGC AAG GAG TAC AAG TGC AAG GTG AGC        768
Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser
    240                 245                 250

AAC AAG GCC CTG CCC GCC CCC ATC GAG AAG ACC ATC TCC AAG ACC AAG        816
Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys
255                 260                 265                 270

GGC CAG CCC CGC GAG CCC CAG GTG TAC ACC CTG CCC CCC AGC CGC GAG        864
Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu
                275                 280                 285

GAG ATG ACC AAG AAC CAG GTG AGC CTG ACC TGC CTG GTG AAG GGC TTC        912
Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe
            290                 295                 300

TAC CCC AGC GAC ATC GCC GTG GAG TGG GAG AGC AGC GGC CAG CCC GAG        960
Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln Pro Glu
        305                 310                 315
```

```
AAC AAC TAC AAC ACC ACC CCC CCC ATG CTG GAC AGC GAC GGC AGC TTC      1008
Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe
    320                 325                 330

TTC CTG TAC AGC AAG CTG ACC GTG GAC AAG AGC CGC TGG CAG CAG GGC      1056
Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly
335                 340                 345                 350

AAC ATC TTC TCC TGC AGC GTG ATG CAT GAG GCC CTG CAC AAC CGC TTC      1104
Asn Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn Arg Phe
                355                 360                 365

ACC CAG AAG AGC CTG AGC CTG AGC CCC GGC AAG TGATAGATCT               1147
Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
            370                 375
```

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 377 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```
Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
 1               5                  10                  15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
            20                  25                  30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
        35                  40                  45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
    50                  55                  60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
65                  70                  75                  80

Tyr Thr Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
                85                  90                  95

Arg Val Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr His Thr Cys Pro
            100                 105                 110

Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
        115                 120                 125

Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg Cys
    130                 135                 140

Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg Cys Pro
145                 150                 155                 160

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
                165                 170                 175

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
            180                 185                 190

Val Val Asp Val Ser His Glu Asp Pro Glu Val Gln Phe Lys Trp Tyr
        195                 200                 205
```

```
Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
    210                 215                 220

Gln Tyr Asn Ser Thr Phe Arg Val Val Ser Val Leu Thr Val Leu His
225                 230                 235                 240

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
                245                 250                 255

Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys Gly Gln
            260                 265                 270

Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met
        275                 280                 285

Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
    290                 295                 300

Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln Pro Glu Asn Asn
305                 310                 315                 320

Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe Phe Leu
                325                 330                 335

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Ile
            340                 345                 350

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn Arg Phe Thr Gln
        355                 360                 365

Lys Ser Leu Ser Leu Ser Pro Gly Lys
    370                 375
```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 999 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 9..989

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
GCGGCCGC GCG TCG ACC AAG GGC CCC AGC GTG TTC CCC CTG GCC CCC TGC        50
         Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys
           1               5                  10

AGC CGC AGC ACC AGC GAG AGC ACC GCC GCC CTG GGC TGC CTG GTG AAG        98
Ser Arg Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys
 15                  20                  25                  30

GAC TAC TTC CCC GAG CCC GTG ACC GTG AGC TGG AAC AGC GGC GCC CTG       146
Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu
                35                  40                  45

ACC AGC GGC GTG CAC ACC TTC CCC GCC GTG CTG CAG AGC AGC GGC CTG       194
Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu
            50                  55                  60
```

```
TAC TCC CTG AGC AGC GTG GTG ACC GTG CCC AGC AGC AGC CTG GGC ACC        242
Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr
          65                  70                  75

AAG ACC TAC ACC TGC AAC GTG GAC CAC AAG CCC AGC AAC ACC AAG GTG        290
Lys Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val
      80                  85                  90

GAC AAG CGC GTG GAG AGC AAG TAC GGC CCC CCC TGC CCC AGC TGC CCC        338
Asp Lys Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser Cys Pro
 95                 100                 105                 110

GCC CCC GAG TTC CTG GGC GGC CCC AGC GTG TTC CTG TTC CCC CCC AAG        386
Ala Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
                115                 120                 125

CCC AAG GAC ACC CTG ATG ATC AGC CGC ACC CCC GAG GTG ACC TGC GTG        434
Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
            130                 135                 140

GTG GTG GAC GTG AGC CAG GAG GAC CCC GAG GTG CAG TTC AAC TGG TAC        482
Val Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr
        145                 150                 155

GTG GAC GGC GTG GAG GTG CAT AAC GCC AAG ACC AAG CCC CGC GAG GAG        530
Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
    160                 165                 170

CAG TTC AAC AGC ACC TAC CGC GTG GTG AGC GTG CTG ACC GTG CTG CAC        578
Gln Phe Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
175                 180                 185                 190

CAG GAC TGG CTG AAC GGC AAG GAG TAC AAG TGC AAG GTG TCC AAC AAG        626
Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
                195                 200                 205

GGC CTG CCC AGC AGC ATC GAG AAG ACC ATC AGC AAG GCC AAG GGC CAG        674
Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
            210                 215                 220

CCC CGC GAG CCC CAG GTG TAC ACC CTG CCC CCC AGC CAG GAG GAG ATG        722
Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met
        225                 230                 235

ACC AAG AAC CAG GTG AGC CTG ACC TGC CTG GTG AAG GGC TTC TAC CCC        770
Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
    240                 245                 250

AGC GAC ATC GCC GTG GAG TGG GAG AGC AAC GGC CAG CCC GAG AAC AAC        818
Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
255                 260                 265                 270

TAC AAG ACC ACC CCC CCC GTG CTG GAC AGC GAC GGC AGC TTC TTC CTG        866
Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
                275                 280                 285

TAC AGC CGC CTG ACC GTG GAC AAG AGC CGC TGG CAG GAG GGC AAC GTG        914
Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val
            290                 295                 300

TTC TCC TGC TCC GTG ATG CAT GAG GCC CTG CAC AAC CAC TAC ACC CAG        962
Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
        305                 310                 315

AAG AGC CTG AGC CTG AGC CTG GGC AAG TGATAGATCT                        999
Lys Ser Leu Ser Leu Ser Leu Gly Lys
    320                 325
```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 327 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
1 5 10 15

Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Lys Thr
65 70 75 80

Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp Lys
85 90 95

Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser Cys Pro Ala Pro
100 105 110

Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
115 120 125

Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val
130 135 140

Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp
145 150 155 160

Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe
165 170 175

Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp
180 185 190

Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu
195 200 205

Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
210 215 220

Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys
225 230 235 240

Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
245 250 255

Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
260 265 270

Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
275 280 285

Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser
290 295 300

```
Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
305                 310                 315                 320

Leu Ser Leu Ser Leu Gly Lys
                325
```

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 337 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 9..326

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

```
GCGGCCGC ACT GTG GCT GCA CCA TCT GTC TTC ATC TTC CCG CCA TCT GAT        50
         Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp
           1               5                  10

GAG CAG CTT AAG TCC GGA ACC GCC AGC GTG GTG TGC CTG CTG AAC AAC        98
Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn
 15                  20                  25                  30

TTC TAC CCC CGC GAG GCC AAG GTG CAG TGG AAG GTG GAC AAC GCC CTC       146
Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu
                 35                  40                  45

CAG AGC GGC AAC TCC CAG GAG AGC GTG ACC GAG CAG GAC AGC AAG GAC       194
Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp
            50                  55                  60

AGC ACC TAC AGC CTG AGC AGC ACC CTG ACC CTG AGC AAG GCC GAC TAC       242
Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr
        65                  70                  75

GAG AAG CAC AAG GTG TAC GCC TGC GAG GTG ACC CAT CAG GGC CTG AGC       290
Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser
     80                  85                  90

AGC CCC GTG ACC AAG AGC TTC AAC CGG GGC GAG TGC TAGTGAGATC            336
Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
 95                  100                 105

T                                                                     337
```

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 106 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```
Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
  1               5                  10                  15
```

```
Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
                20                  25                  30

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
            35                  40                  45

Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
        50                  55                  60

Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
65                  70                  75                  80

His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro
                85                  90                  95

Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
            100                 105
```

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 346 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 9..335

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

```
GCGGCCGC ACC GTC CTA GGT CAG CCC AAG GCG GCG CCC AGC GTG ACC CTG      50
         Thr Val Leu Gly Gln Pro Lys Ala Ala Pro Ser Val Thr Leu
           1               5                  10

TTC CCC CCC AGC AGC GAG GAG CTG CAG GCC AAC AAG GCC ACC CTG GTG      98
Phe Pro Pro Ser Ser Glu Glu Leu Gln Ala Asn Lys Ala Thr Leu Val
 15                  20                  25                  30

TGC CTG ATC AGC GAC TTC TAC CCC GGG GCC GTG ACC GTG GCC TGG AAG     146
Cys Leu Ile Ser Asp Phe Tyr Pro Gly Ala Val Thr Val Ala Trp Lys
                35                  40                  45

GCC GAC AGC AGC CCC GTG AAG GCC GGC GTG GAG ACC ACC ACC CCC AGC     194
Ala Asp Ser Ser Pro Val Lys Ala Gly Val Glu Thr Thr Thr Pro Ser
            50                  55                  60

AAG CAG AGC AAC AAC AAG TAC GCC GCC AGC AGC TAC CTG AGC CTG ACC     242
Lys Gln Ser Asn Asn Lys Tyr Ala Ala Ser Ser Tyr Leu Ser Leu Thr
        65                  70                  75

CCC GAG CAG TGG AAG AGC CAC CGC AGC TAC AGC TGC CAG GTC ACC CAC     290
Pro Glu Gln Trp Lys Ser His Arg Ser Tyr Ser Cys Gln Val Thr His
    80                  85                  90

GAG GGC AGC ACC GTG GAG AAG ACC GTG GCC CCC ACC GAG TGC AGC         335
Glu Gly Ser Thr Val Glu Lys Thr Val Ala Pro Thr Glu Cys Ser
 95                  100                 105

TAGTGAGATC T                                                        346
```

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 109 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

```
Thr Val Leu Gly Gln Pro Lys Ala Ala Pro Ser Val Thr Leu Phe Pro
 1               5                  10                  15

Pro Ser Ser Glu Glu Leu Gln Ala Asn Lys Ala Thr Leu Val Cys Leu
            20                  25                  30

Ile Ser Asp Phe Tyr Pro Gly Ala Val Thr Val Ala Trp Lys Ala Asp
        35                  40                  45

Ser Ser Pro Val Lys Ala Gly Val Glu Thr Thr Thr Pro Ser Lys Gln
    50                  55                  60

Ser Asn Asn Lys Tyr Ala Ala Ser Ser Tyr Leu Ser Leu Thr Pro Glu
65                  70                  75                  80

Gln Trp Lys Ser His Arg Ser Tyr Ser Cys Gln Val Thr His Glu Gly
                85                  90                  95

Ser Thr Val Glu Lys Thr Val Ala Pro Thr Glu Cys Ser
            100                 105
```

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

TCTAGAATTC ACGCGTCCAC CATGGACTGG ACCTGGAG 38

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TCTAGAATTC ACGCGTCCAC CATGGACACA CTTTGCTACA C 41

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

TCTAGAATTC ACGCGTCCAC CATGGAGTTT GGGCTGAGCT GG 42

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

TCTAGAATTC ACGCGTCCAC CATGAAACAC CTGTGGTTCT TCCT 44

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

TCTAGAATTC ACGCGTCCAC CATGGGGTCA ACCGCCATCC T 41

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

TCTAGAATTC ACGCGTCCAC CATGTCTGTC TCCTTCCTCA TCTT 44

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

GCCTGAGTTC CACGACACCG TCAC 24

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GGGGAAAAGG GTTGGGGCGG ATGC 24

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

GAGGGGCCCT TGGTCGACGC TGAGGAGACG GTGACCAGG 39

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GAGGGGCCCT TGGTCGACGC TGAAGAGACG GTGACCATTG 40

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

GAGGGGCCCT TGGTCGACGC TGAGGAGACG GTGACCGTG 39

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

TCTAGAATTC ACGCGTCCAC CATGGACATG AGGGTCCCCG CTCAG 45

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

TCTAGAATTC ACGCGTCCAC CATGAGGCTC CCTGCTCAGC 40

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

TCTAGAATTC ACGCGTCCAC CATGGAAGCC CCAGCGCAGC TT 42

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

TCTAGAATTC ACGCGTCCAC CATGGTGTTG CAGACCCAGG T 41

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

TCTAGAATTC ACGCGTCCAC CATGGGGTCC CAGGTTCACC T 41

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

TCTAGAATTC ACGCGTCCAC CATGTTGCCA TCACAACTCA TTG 43

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TCTAGAATTC ACGCGTCCAC CATGGTGTCC CCGTTGCAAT T 41

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

GGTTCCGGAC TTAAGCTGCT CATCAGATGG CGGG 34

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

TCTAGAATTC ACGCGTCCAC CATGGCCTGC TCTCCTCTCC TCCT 44

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

TCTAGAATTC ACGCGTCCAC CATGGCCTGG GCTCTGCTGC TCCT 44

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

TCTAGAATTC ACGCGTCCAC CATGGCCTGG ATCCTTCTCC TCCTC 45

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

TCTAGAATTC ACGCGTCCAC CATGGCCTGG ACCCCTCTCT GGCTC 45

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

TCTAGAATTC ACGCGTCCAC CATGGCCTGG GCCCCACTAC T 41

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

TCTAGAATTC ACGCGTCCAC CATGGCCTGG ATGATGCTTC TCCT 44

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

GGCGCCGCCT TGGGCTGACC TAGGACGGT 29

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 292 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

GAATTCTTTT TTGCGTGTGG CAGTTTTAAG TTATTAGTTT TTAAAATCAG TACTTTTTAA 60

TGGAAACAAC TTGACCAAAA ATTTGTCACA GAATTTTGAG ACCCATTAAA AAAGTTAAAT 120

GAGAAACCTG TGTGTTCCTT TGGTCAACAC CGAGACATTT AGGTGAAAGA CATCTAATTC 180

TGGTTTTACG AATCTGGAAA CTTCTTGAAA ATGTAATTCT TGAGTTAACA CTTCTGGGTG 240

GAGAATAGGG TTGTTTTCCC CCCACATAAT TGGAAGGGGA AGGAATATCG AT 292

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

GAATTCTTTT TTGCGTGTGG CAG 23

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

ATCGATATTC CTTCCCCTTC C 21

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(ix) FEATURE:
(A) NAME/KEY: misc_difference
(B) LOCATION: replace(17, "")
(D) OTHER INFORMATION: /note= "The residue at this position can be repeated 18-21 times."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

TCTAGAATTC ACGCGTN 17

74667-99

CLAIMS:

1. A multivalent vaccine comprising at least two recombinant variable regions of immunoglobulin molecules derived from B-cell lymphoma cells, wherein said at least two variable regions are from immunoglobulin molecules that differ by at least one idiotope.

2. The vaccine of claim 1, wherein said vaccine comprises at least two recombinant immunoglobulin molecules comprising said recombinant variable regions derived from said lymphoma cells.

3. The vaccine of claim 2, wherein said recombinant immunoglobulin molecules are covalently linked to an immune-enhancing cytokine.

4. The vaccine of claim 3, wherein said cytokine is selected from the group consisting of granulocyte-macrophage colony stimulating factor, interleukin-2 and interleukin-4.

5. The multivalent vaccine of any one of claims 1 to 4, further comprising at least one pharmaceutically acceptable excipient.

6. The multivalent vaccine of any one of claims 1 to 5, further comprising an adjuvant.

7. A cell of a T lymphoid cell line comprising a first amplified recombinant oligonucleotide having a sequence encoding an inhibitable enzyme operably linked to a heterologous promoter, and a second amplified recombinant oligonucleotide having a sequence encoding an amino acid sequence of a protein of interest.

74667-99

8. The cell of claim 7, wherein said recombinant oligonucleotide having a sequence encoding an inhibitable enzyme encodes an active dihydrofolate reductase.

9. The cell of claim 7 or claim 8, wherein said T lymphoid cell line is the BW5147.G.1.4 cell line.

10. The cell of any one of claims 7 to 9, further comprising an integrated recombinant oligonucleotide comprising a gene encoding a selectable marker.

11. A method of co-amplifying recombinant oligonucleotides, comprising:

a) providing:

i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest;

ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and

iii) a eukaryotic parent cell line;

b) introducing 400 to 500 micrograms of said expression vector and 20 to 30 micrograms of said amplification vector into said parent cell line to generate transformed cells;

c) introducing said transformed cells into a first aqueous solution containing an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent growth of said parent cell line; and

d) identifying a transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

12. A method of co-amplifying recombinant oligonucleotides, comprising:

a) providing:

i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest;

ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and

iii) a T lymphoid parent cell line;

b) treating said expression vector and said amplification vector with a restriction enzyme to create a linearized expression vector and a linearized amplification vector;

c) introducing said linearized expression vector and said linearized amplification vector into said parent cell line to generate transformed cells, wherein a ratio ranging from 10:1 to 15:1 of said expression vector to said amplification vector is employed;

d) introducing said transformed cells into a first aqueous solution containing an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of

said inhibitor present in said first aqueous solution is sufficient to prevent growth of said parent cell line; and

e) identifying a transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

13. A method of co-amplifying recombinant oligonucleotides, comprising:

a) providing:

i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest;

ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and

iii) a T lymphoid parent cell line;

b) introducing said expression vector and said amplification vector into said parent cell line to generate transformed cells;

c) introducing said transformed cells into a first aqueous solution, said first aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme and wherein the concentration of said inhibitor present in said first aqueous solution is four-fold to six-fold the concentration of said inhibitor required to prevent the growth of said parent cell line; and

d) identifying at least one transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

14. A method of co-amplifying recombinant oligonucleotides, comprising:

a) providing:

i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest;

ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter;

iii) a selection vector comprising a third recombinant oligonucleotide having a sequence which encodes a selectable gene product; and

iv) a T lymphoid parent cell line containing an endogenous gene encoding a second inhibitable enzyme;

b) introducing said expression vector, said amplification vector and said selection vector into said cell line to generate transformed cells, wherein a ratio ranging from 20:2:1 to 25:2:1 of said expression vector to said amplification vector to said selection vector is employed;

c) introducing said transformed cells into a first aqueous solution, said first aqueous solution requiring the expression of said selectable gene product for growth of said transformed cells;

d) identifying at least one transformed cell capable of growth in said first aqueous solution;

e) introducing said transformed cell capable of growth in said first aqueous solution into a second aqueous solution, said second aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said second aqueous solution is sufficient to prevent growth of said parent cell line; and

f) identifying at least one transformed cell capable of growth in said second aqueous solution, wherein said transformed cell capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

15. A method of co-amplifying recombinant oligonucleotides, comprising:

a) providing:

i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest;

ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter;

iii) a selection vector comprising a third recombinant oligonucleotide having a sequence which encodes a selectable gene product; and

iv) a T lymphoid parent cell line containing an endogenous gene encoding a second inhibitable enzyme;

b) introducing 400-500 micrograms of said expression vector, 20-30 micrograms of said amplification vector and 10-15 micrograms of said selection vector into said cell line to generate transformed cells;

c) introducing said transformed cells into a first aqueous solution, said first aqueous solution requiring the expression of said selectable gene product for growth of said transformed cells;

d) identifying at least one individual clone of transformed cells capable of growth in said first aqueous solution;

e) introducing said individual clone capable of growth in said first aqueous solution into a second aqueous solution, said second aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said second aqueous solution is sufficient to prevent growth of said parent cell line; and

f) identifying at least one individual clone capable of growth in said second aqueous solution, wherein said clone capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

16. A method of co-amplifying recombinant oligonucleotides, comprising:

a) providing:

i) at least one expression vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest;

74667-99

ii) an amplification vector comprising a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and

iii) a T lymphoid parent cell line;

b) introducing said expression vector and said amplification vector into said T lymphoid parent cell line to generate transformed cells, wherein a ratio of at least 10:1 of said expression vector to said amplification vector is employed;

c) introducing said transformed cells into a first aqueous solution, said first aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent the growth of said parent cell line; and

d) identifying a transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

17. A method of co-amplifying recombinant oligonucleotides, comprising:

a) providing:

i) a vector comprising a first recombinant oligonucleotide having a sequence encoding the amino acid sequence of a protein of interest and a second recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter; and

ii) a T lymphoid parent cell line;

b) introducing said vector into said T lymphoid parent cell line to generate transformed cells;

c) introducing said transformed cells into a first aqueous solution, said first aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent the growth of said parent cell line; and

d) identifying a transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable of growth contains an amplified number of copies of said vector.

18. A method of producing a vaccine for treatment of B-cell lymphoma, comprising:

a) providing:

i) malignant cells isolated from a patient having a B-cell lymphoma;

ii) an amplification vector comprising a recombinant oligonucleotide having a sequence encoding a first inhibitable enzyme operably linked to a heterologous promoter;

iii) a eukaryotic parent cell line;

b) isolating from said malignant cells nucleotide sequences encoding at least one V_H region and at least one V_L region, said V_H and V_L regions derived from immunoglobulin molecules expressed by said malignant cells;

c) inserting said nucleotide sequences encoding said V_H and V_L regions into at least one expression vector;

d) introducing said at least one expression vector and said amplification vector into said parent cell to generate transformed cells;

e) introducing said transformed cells into a first aqueous solution, said first aqueous solution comprising an inhibitor capable of inhibiting said first inhibitable enzyme, wherein the concentration of said inhibitor present in said first aqueous solution is sufficient to prevent growth of said parent cell line; and

f) identifying a transformed cell capable of growth in said first aqueous solution, wherein said transformed cell capable of growth expresses said V_H and V_L regions.

19. The method of claim 18, wherein transformed cell capable of growth contains an amplified number of copies of said expression vector and an amplified number of copies of said amplification vector.

20. The method of claim 18 or claim 19, wherein nucleotide sequences encoding said V_H and V_L regions comprise at least two V_H and at least two V_L regions.

21. The method of any one of claim 11 and 18 to 20, wherein said parent cell line is a T lymphoid cell line.

22. The method of any one of claims 11, 12, 13, 16, 17 and 18 to 21, wherein said parent cell line contains an endogenous gene encoding a second inhibitable enzyme.

23. The method of any one of claims 14, 15 and 22, wherein said second inhibitable enzyme is selected from the

group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase.

24. The method of any one of claims 11, 12, 16, 17 and 18 to 21, wherein said concentration of inhibitor present in said first aqueous solution is four-fold to six-fold the concentration required to prevent the growth of said parent cell line.

25. The method of claim 14 or 15, wherein said concentration of inhibitor present in said second aqueous solution is four-fold to six-fold the concentration required to prevent the growth of said parent cell line.

26. The method of claim 14 or claim 15, wherein said first and said second inhibitable enzyme are the same.

27. The method of any one of claims 11, 12, 13, 16 and 18 to 21, further comprising providing a selection vector encoding a selectable gene product which is introduced into said parent cell line together with said expression vector and said amplification vector.

28. The method of claim 17, further comprising providing a selection vector encoding a selectable gene product which is introduced into said parent cell line together with said vector comprising said first and second recombinant oligonucleotides.

29. The method of any one of claims 14, 15, 27 and 28, wherein said selection vector encodes an active enzyme selected from the group consisting of hypoxanthine guanine phosphoribosyltransferase, hygromycin G phosphotransferase, xanthine-guanine phosphoribosyltransferase and aminoglycoside 3' phosphotransferase.

74667-99

30. The method of claim 29, wherein said selection vector encodes an active hypoxanthine guanine phosphoribosyltransferase.

31. The method of claim 14 or 15, wherein said first aqueous solution requiring the expression of said selectable gene product comprises hypoxanthine and azaserine.

32. The method of claim 27 or 28, further comprising, following the introduction of said selection, expression and amplification vectors, the additional step of introducing said transformed cells into a second aqueous solution, said second aqueous solution requiring the expression of said selectable gene product for growth of said transformed cells, prior to introducing said transformed cells into said first aqueous solution.

33. The method of any one of claims 11 to 16 and 18, wherein said amplification vector encodes an active enzyme selected from the group consisting of dihydrofolate reductase, glutamine synthetase, adenosine deaminase and asparagine synthetase.

34. The method of claim 33, wherein said inhibitor is selected from the group consisting of methotrexate, 2'-deoxycoformycin, methionine sulphoximine, albizziin and β-aspartyl hydroxamate.

35. The method of claim 14 or 25, further comprising the steps of:

i) introducing said transformed cell capable of growth in said first aqueous solution into a third aqueous solution, said third aqueous solution comprising said inhibitor capable of inhibiting said first inhibitable enzyme and wherein the concentration of said inhibitor present in said third

aqueous solution is sixteen-fold to thirty-six-fold the concentration of said inhibitor required to prevent the growth of said parent cell line; and

ii) identifying at least one transformed cell capable of growth in said third aqueous solution.

36. The method of claim 26, further comprising the steps of:

i) introducing said transformed cell capable of growth in said second aqueous solution into a third aqueous solution, said third aqueous solution comprising said inhibitor capable of inhibiting said first inhibitable enzyme and wherein the concentration of said inhibitor present in said third aqueous solution is sixteen-fold to thirty-six-fold the concentration of said inhibitor required to prevent the growth of said parent cell line; and

ii) identifying at least one transformed cell capable of growth in said third aqueous solution.

37. The method of any one of claims 12, 13, 16 and 19, wherein 20 to 30 micrograms of said amplification vector and a total of 400 to 500 micrograms of said expression vector are introduced into said parent cell line.

38. The method of claim 14 or 27, wherein 10 to 15 micrograms of said selection vector, 20 to 30 micrograms of said amplification vector and a total of 400 to 500 micrograms of said expression vector are introduced into said parent cell line.

39. The method of claim 17, wherein 5 to 30 micrograms of said vector are introduced into said parent cell line.

40. The method of claims 16 or 21, wherein said T lymphoid cell line is the BW5147.G.1.4 cell line.

41. The method of any one of claims 11, 13, 16 and 18, wherein said expression and amplification vectors are linearized prior to introduction into said parent cell line.

42. The method of claim 14 or 15, wherein said expression, amplification and selection vectors are linearized prior to introduction into said parent cell line.

43. The method of claim 17, wherein said vector is linearized prior to introduction into said parent cell line.

44. The method of claim 22, wherein said first and said second inhibitable enzyme are the same.

45. A use of a multivalent vaccine comprising at least two recombinant variable regions of immunoglobulin molecules derived from a subject's B-cell lymphoma cells, wherein said at least two variable regions are from immunoglobulin molecules that differ by at least one idiotope, for treating B-cell lymphoma.

46. The use of claim 45, wherein said vaccine comprises at least two recombinant immunoglobulin molecules comprising said recombinant variable regions derived from said lymphoma cells.

47. The use of claim 45 or claim 46, wherein said vaccine further comprises an adjuvant.

48. The use of claim 47, wherein said adjuvant is selected from the group consisting of Syntex™ adjuvant formulation 1, QS21, and Optivax™.

74667-99

49. A use of a multivalent vaccine comprising at least two recombinant variable regions of immunoglobulin molecules derived from a subject's B-cell lymphoma cells, wherein said at least two variable regions are from immunoglobulin molecules that differ by at least one idiotope, in the manufacture of a medicament for the treatment of said subject's B-cell lymphoma.

SMART & BIGGAR
OTTAWA, CANADA

PATENT AGENTS

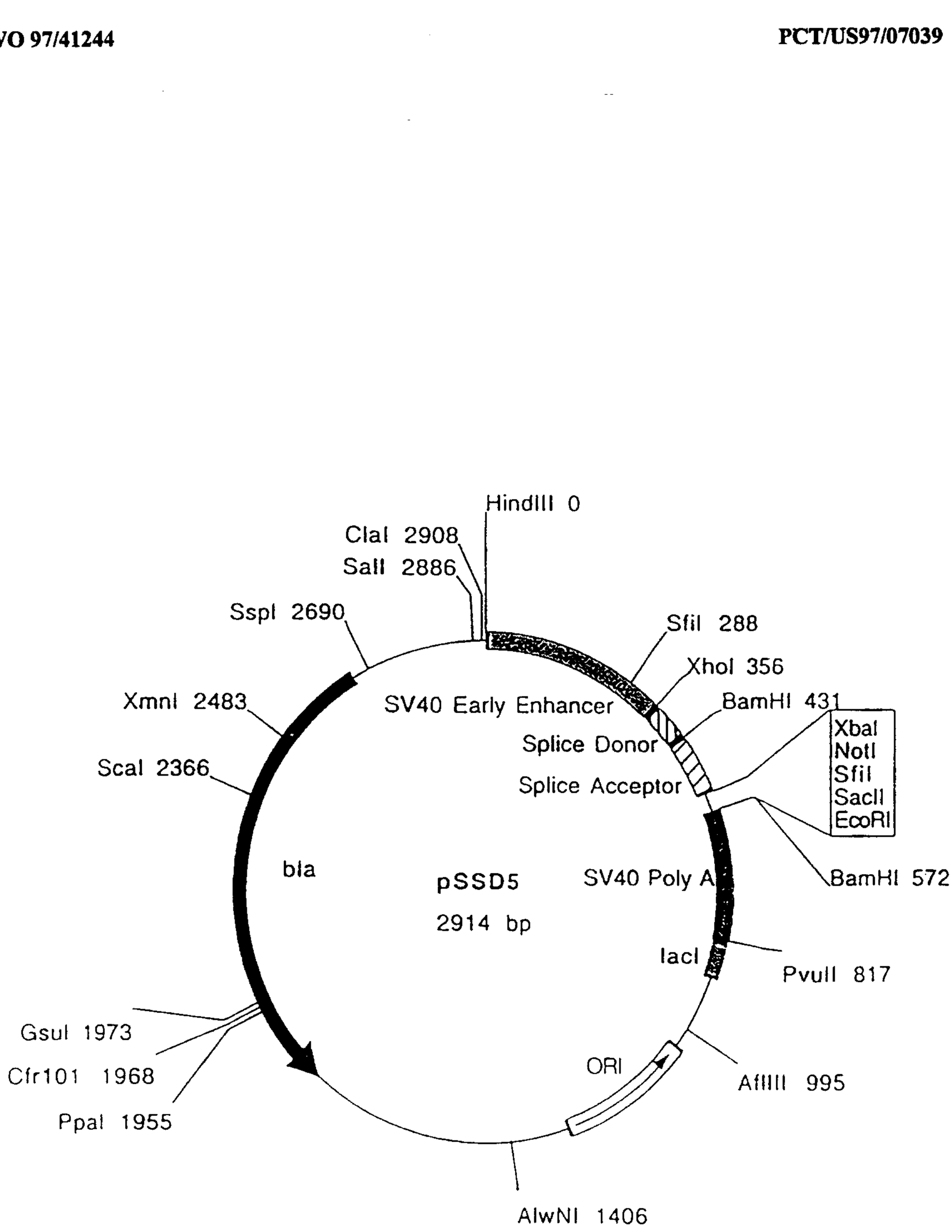
HindIII 0
ClaI 2908
SalI 2886
SspI 2690
SfiI 288
XhoI 356
BamHI 431
XmnI 2483
SV40 Early Enhancer
Splice Donor
Splice Acceptor
XbaI
NotI
SfiI
SacII
EcoRI
ScaI 2366
bla
pSSD5
2914 bp
SV40 Poly A
BamHI 572
lacI
PvuII 817
Gsul 1973
Cfr101 1968
Ppal 1955
ORI
AflIII 995
AlwNI 1406

FIGURE 1

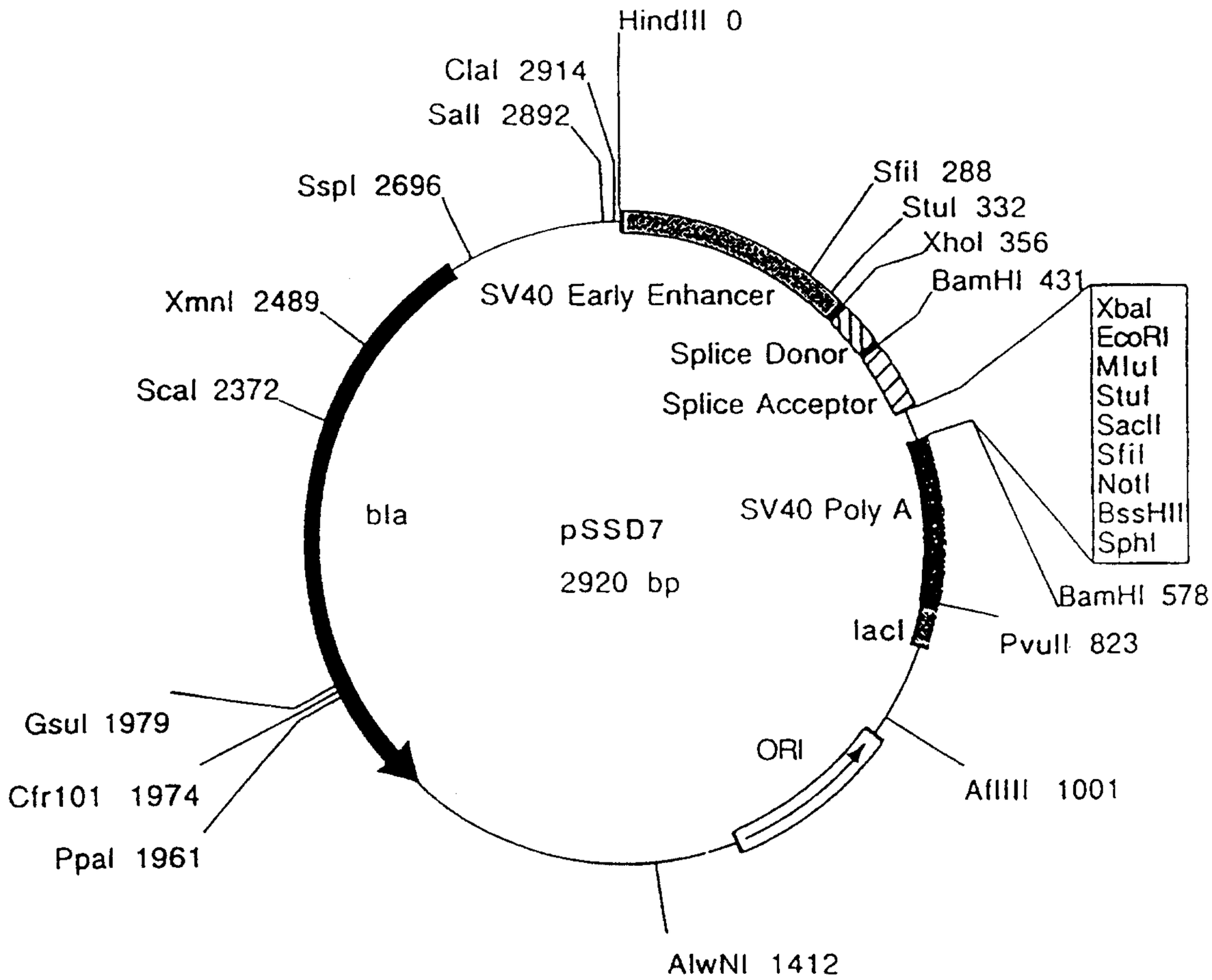
HindIII 0
ClaI 2914
SalI 2892
SspI 2696
SfiI 288
StuI 332
XhoI 356
BamHI 431
SV40 Early Enhancer
XmnI 2489
ScaI 2372
Splice Donor
Splice Acceptor
XbaI
EcoRI
MluI
StuI
SacII
SfiI
NotI
BssHII
SphI
bla
pSSD7
2920 bp
SV40 Poly A
BamHI 578
lacI
PvuII 823
GsuI 1979
Cfr10I 1974
PpaI 1961
ORI
AflIII 1001
AlwNI 1412

FIGURE 2

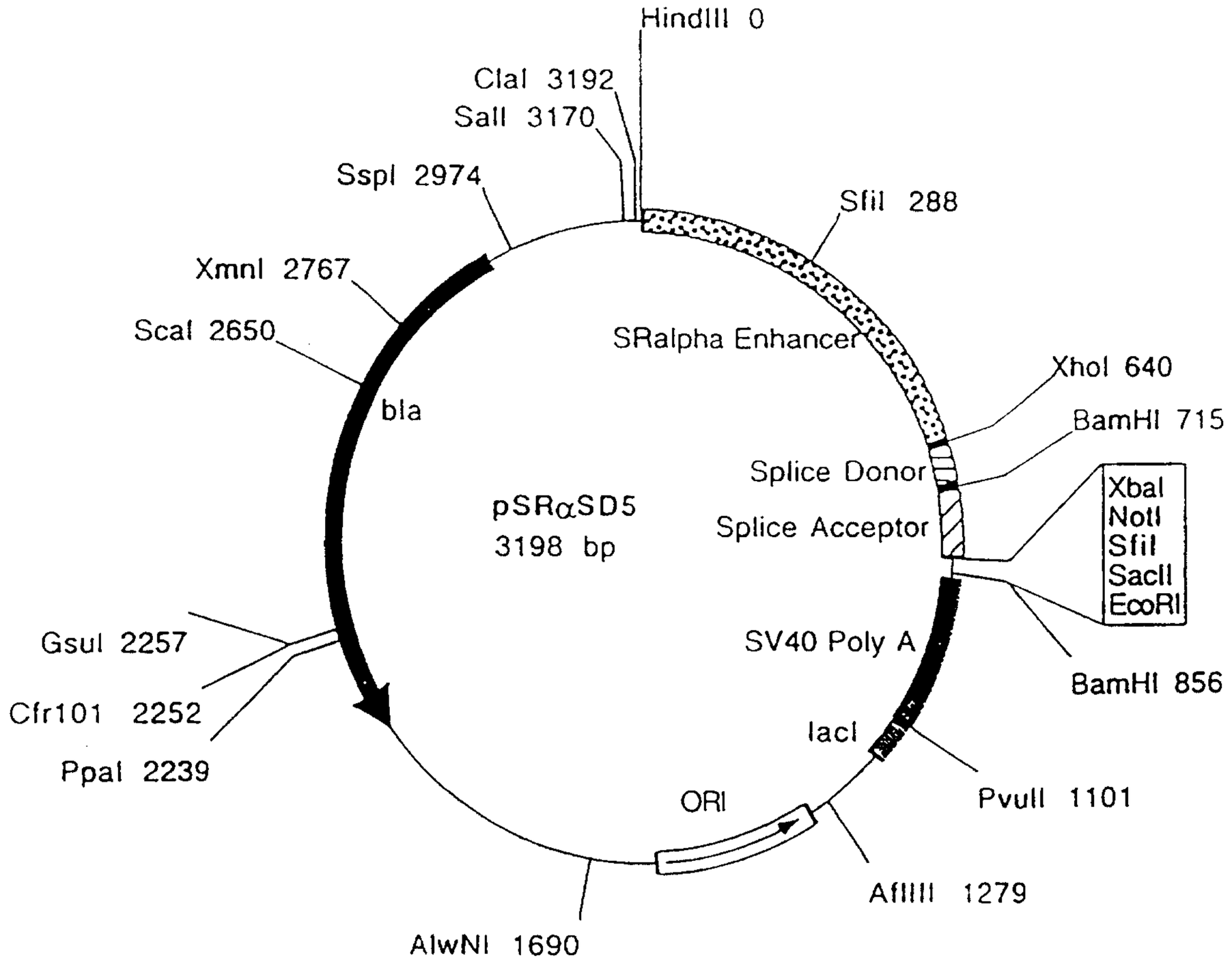
HindIII 0
ClaI 3192
SalI 3170
SspI 2974
SfiI 288
XmnI 2767
ScaI 2650
SRalpha Enhancer
XhoI 640
BamHI 715
bla
Splice Donor
pSRαSD5
3198 bp
Splice Acceptor
XbaI
NotI
SfiI
SacII
EcoRI
SV40 Poly A
BamHI 856
GsuI 2257
Cfr101 2252
PpaI 2239
lacI
PvuII 1101
ORI
AflIII 1279
AlwNI 1690

FIGURE 3

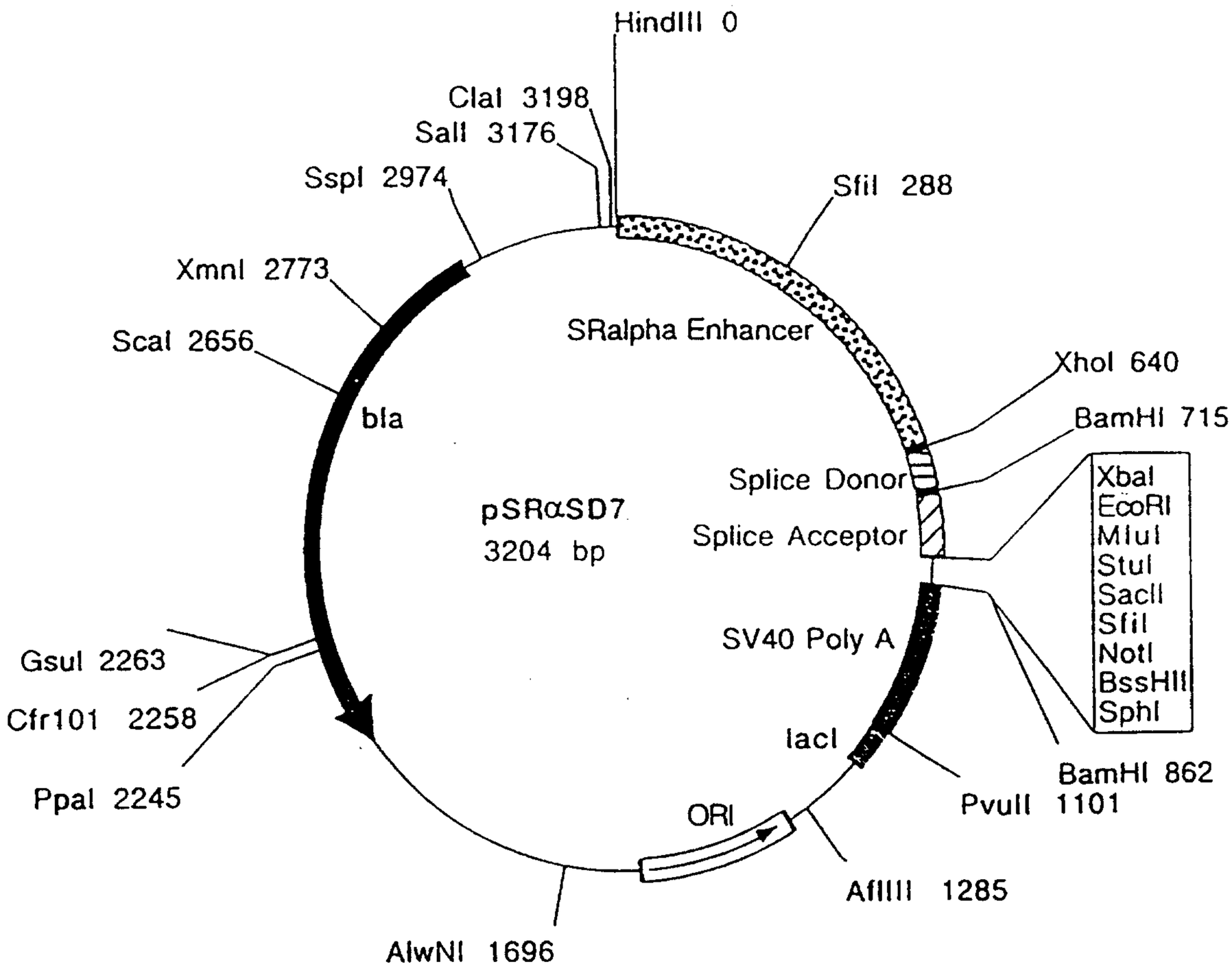
HindIII 0
ClaI 3198
SalI 3176
SspI 2974
SfiI 288
XmnI 2773
ScaI 2656
SRalpha Enhancer
XhoI 640
BamHI 715
bla
Splice Donor
Splice Acceptor
pSRαSD7
3204 bp
XbaI
EcoRI
MluI
StuI
SacII
SfiI
NotI
BssHII
SphI
SV40 Poly A
GsuI 2263
Cfr101 2258
lacI
BamHI 862
Ppal 2245
PvuII 1101
ORI
AflIII 1285
AlwNI 1696

FIGURE 4

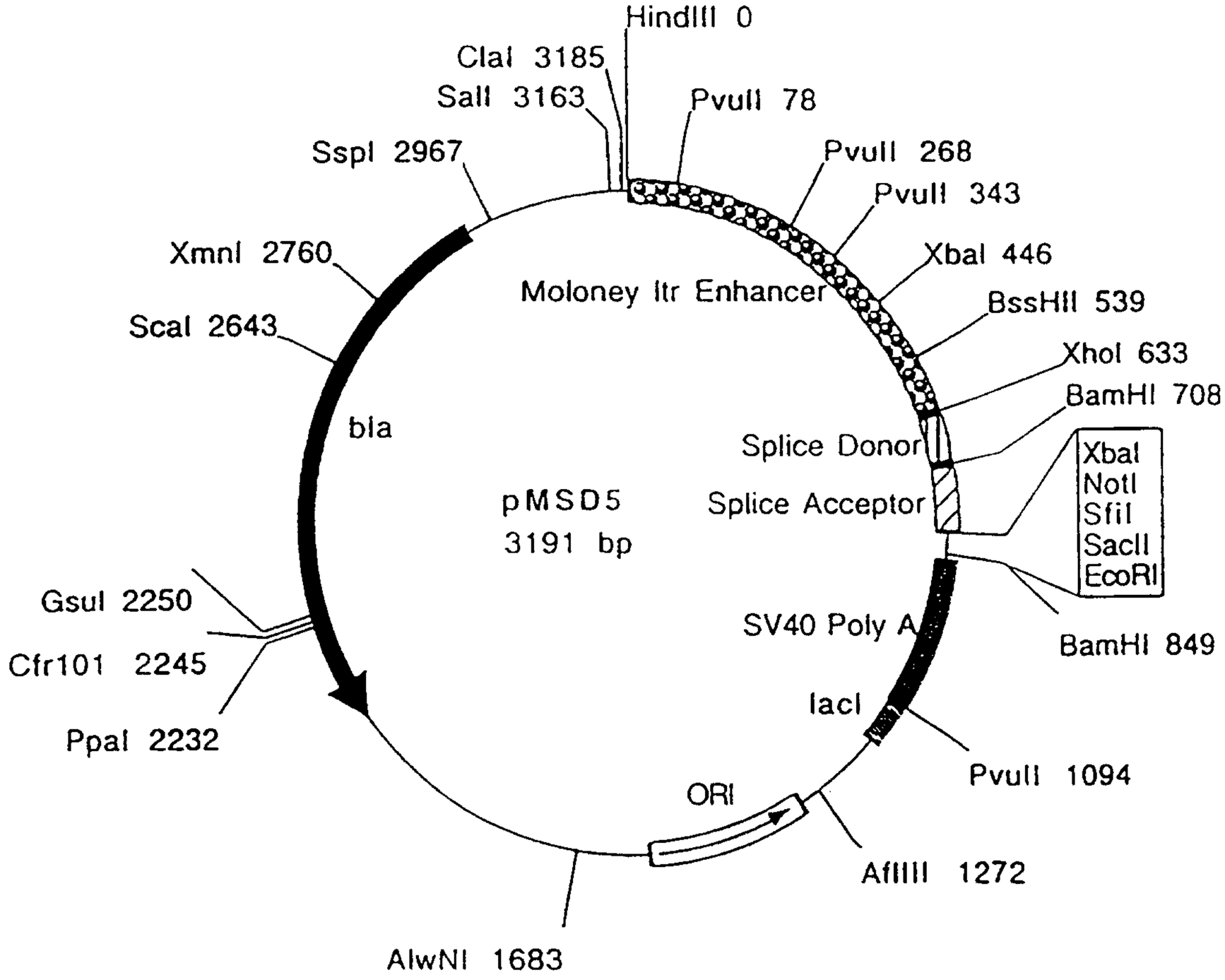
HindIII 0
ClaI 3185
SalI 3163
PvuII 78
SspI 2967
PvuII 268
PvuII 343
XmnI 2760
XbaI 446
Moloney ltr Enhancer
BssHII 539
ScaI 2643
XhoI 633
BamHI 708
bla
Splice Donor
XbaI
NotI
SfiI
SacII
EcoRI
pMSD5
3191 bp
Splice Acceptor
GsuI 2250
SV40 Poly A
BamHI 849
Cfr101 2245
lacI
Ppal 2232
PvuII 1094
ORI
AflIII 1272
AlwNI 1683

FIGURE 5

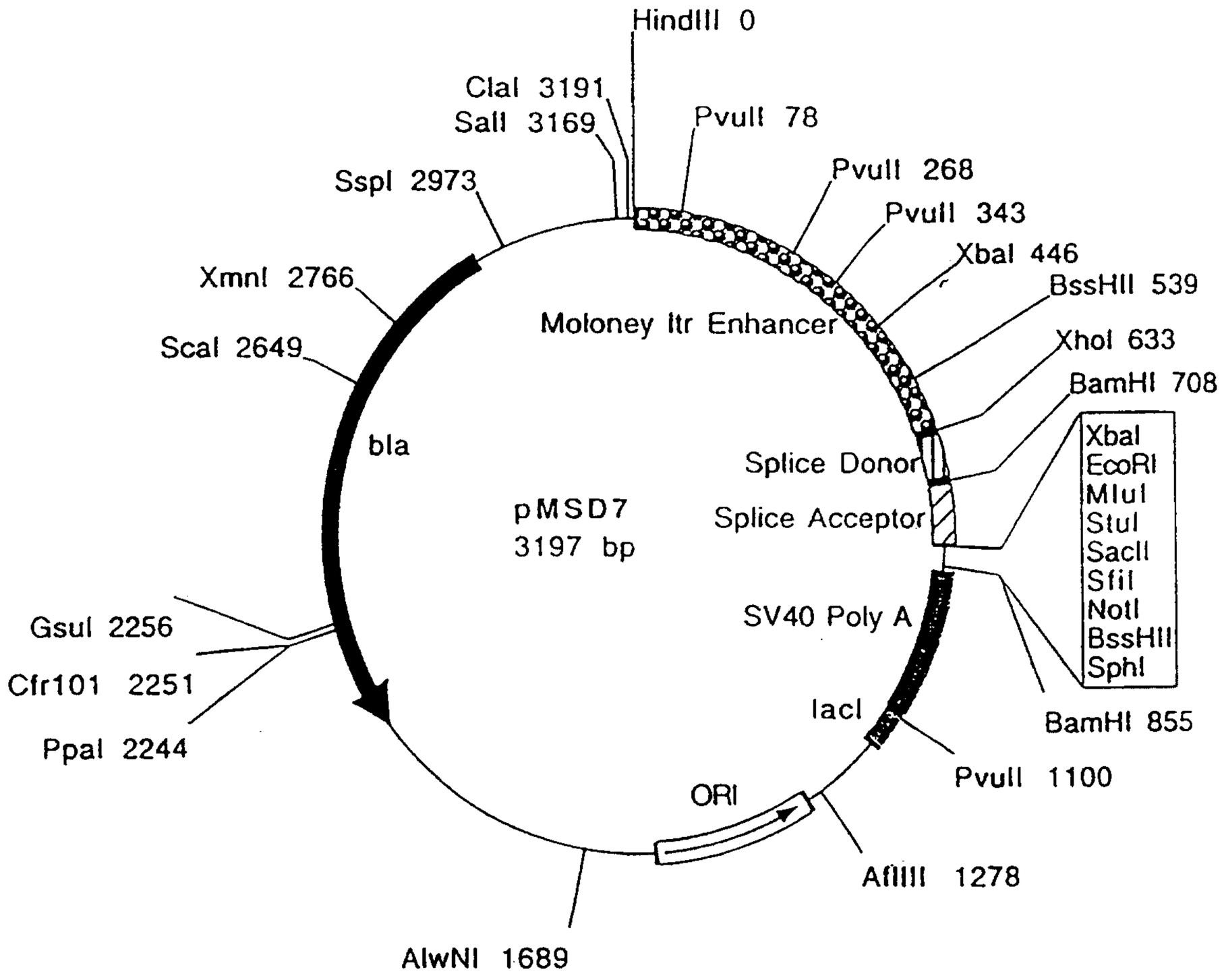
HindIII 0
ClaI 3191
SalI 3169
PvuII 78
PvuII 268
PvuII 343
XbaI 446
BssHII 539
XhoI 633
BamHI 708
SspI 2973
XmnI 2766
ScaI 2649
Moloney ltr Enhancer
bla
pMSD7
3197 bp
Splice Donor
Splice Acceptor
SV40 Poly A
XbaI
EcoRI
MluI
StuI
SacII
SfiI
NotI
BssHII
SphI
BamHI 855
GsuI 2256
Cfr101 2251
PpaI 2244
lacI
PvuII 1100
ORI
AflIII 1278
AlwNI 1689

FIGURE 6

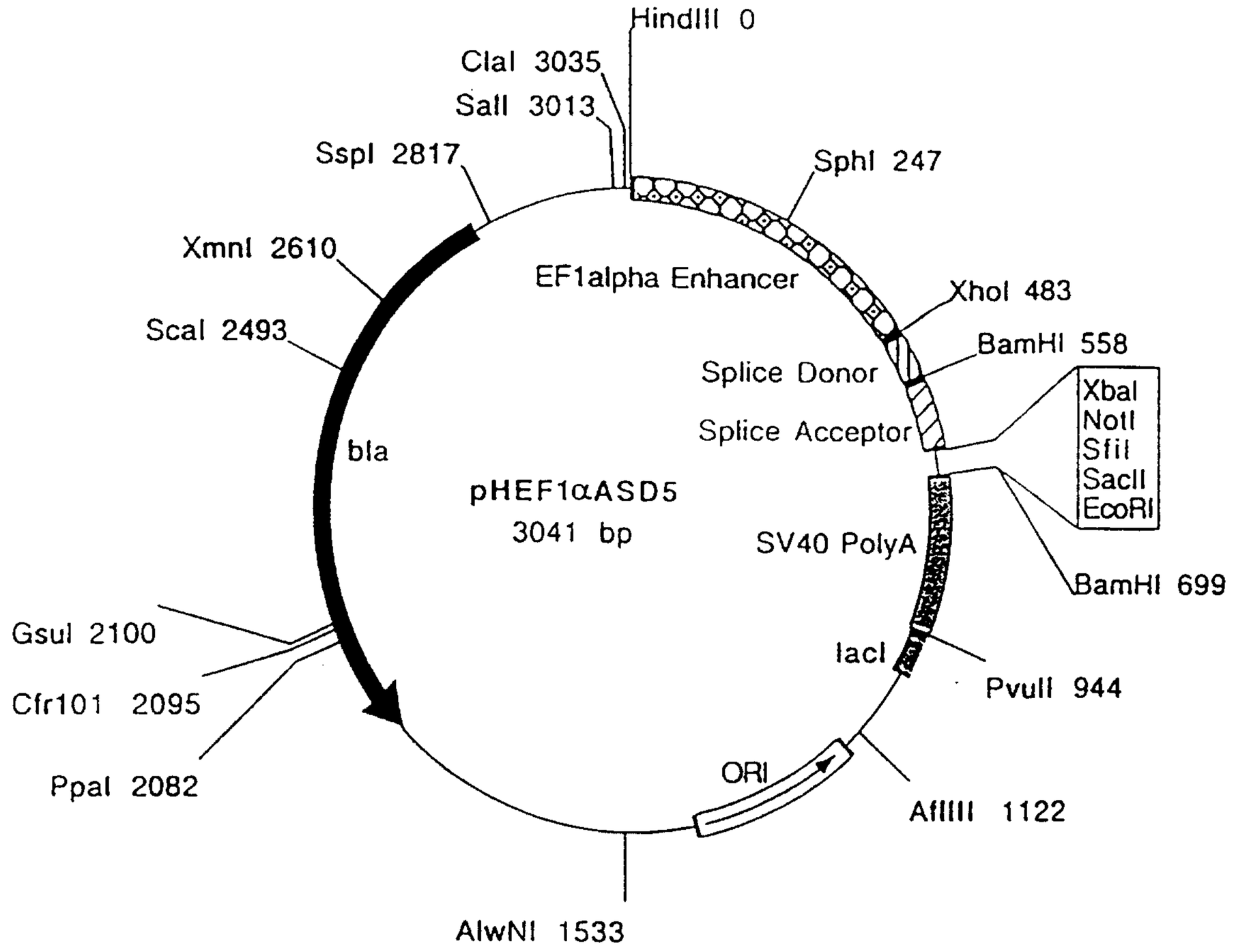
HindIII 0
ClaI 3035
SalI 3013
SspI 2817
SphI 247
XmnI 2610
EF1alpha Enhancer
XhoI 483
ScaI 2493
BamHI 558
Splice Donor
Splice Acceptor
XbaI
NotI
SfiI
SacII
EcoRI
bla
pHEF1αASD5
3041 bp
SV40 PolyA
BamHI 699
GsuI 2100
lacI
PvuII 944
Cfr10I 2095
Ppal 2082
ORI
AflIII 1122
AlwNI 1533

FIGURE 7

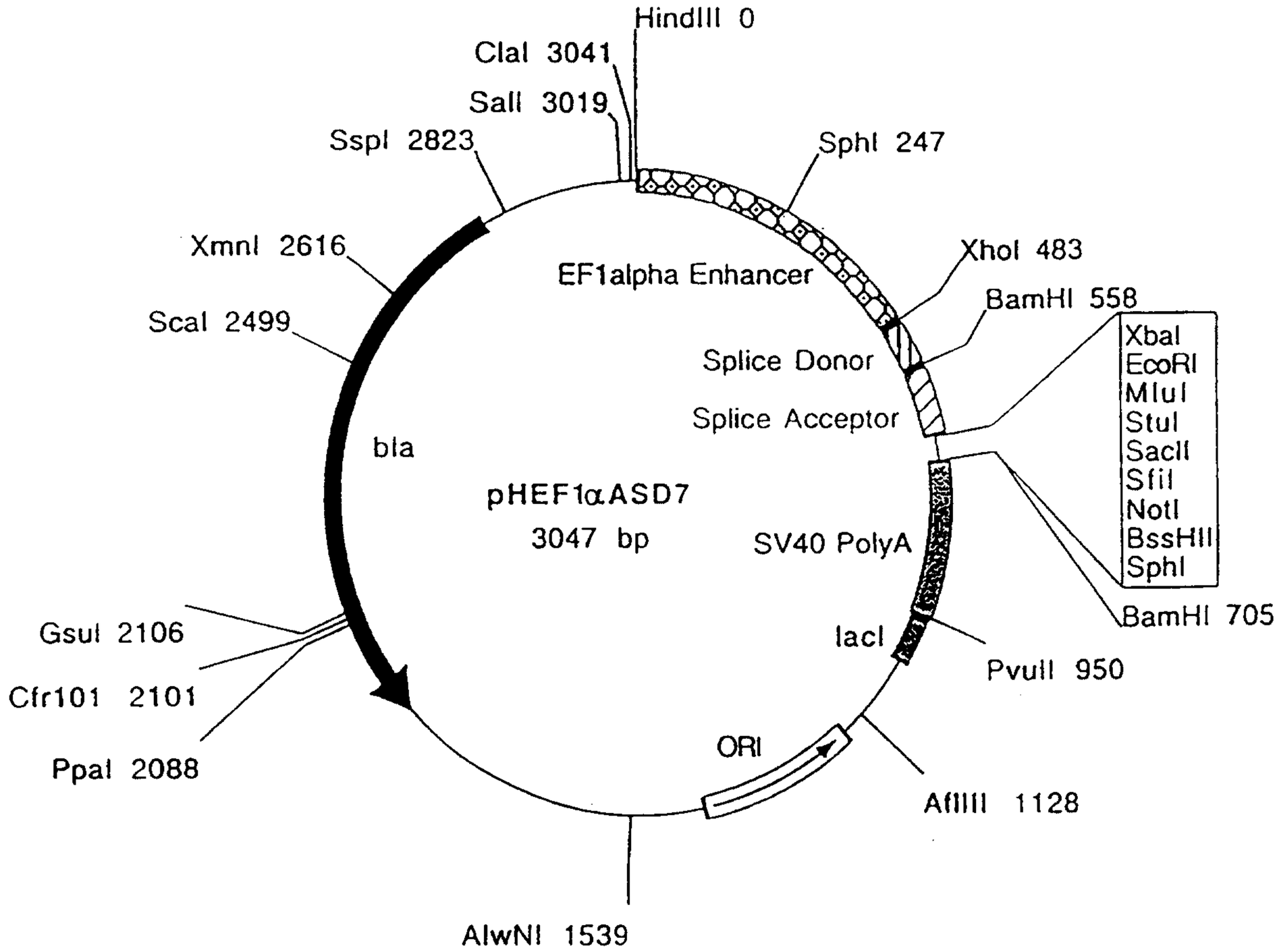
HindIII 0
ClaI 3041
SalI 3019
SspI 2823
SphI 247
XmnI 2616
XhoI 483
EF1alpha Enhancer
BamHI 558
ScaI 2499
Splice Donor
Splice Acceptor
XbaI
EcoRI
MluI
StuI
SacII
SfiI
NotI
BssHII
SphI
bla
pHEF1αASD7
3047 bp
SV40 PolyA
BamHI 705
GsuI 2106
lacI
PvuII 950
Cfr10I 2101
ORI
Ppal 2088
AflIII 1128
AlwNI 1539

FIGURE 8

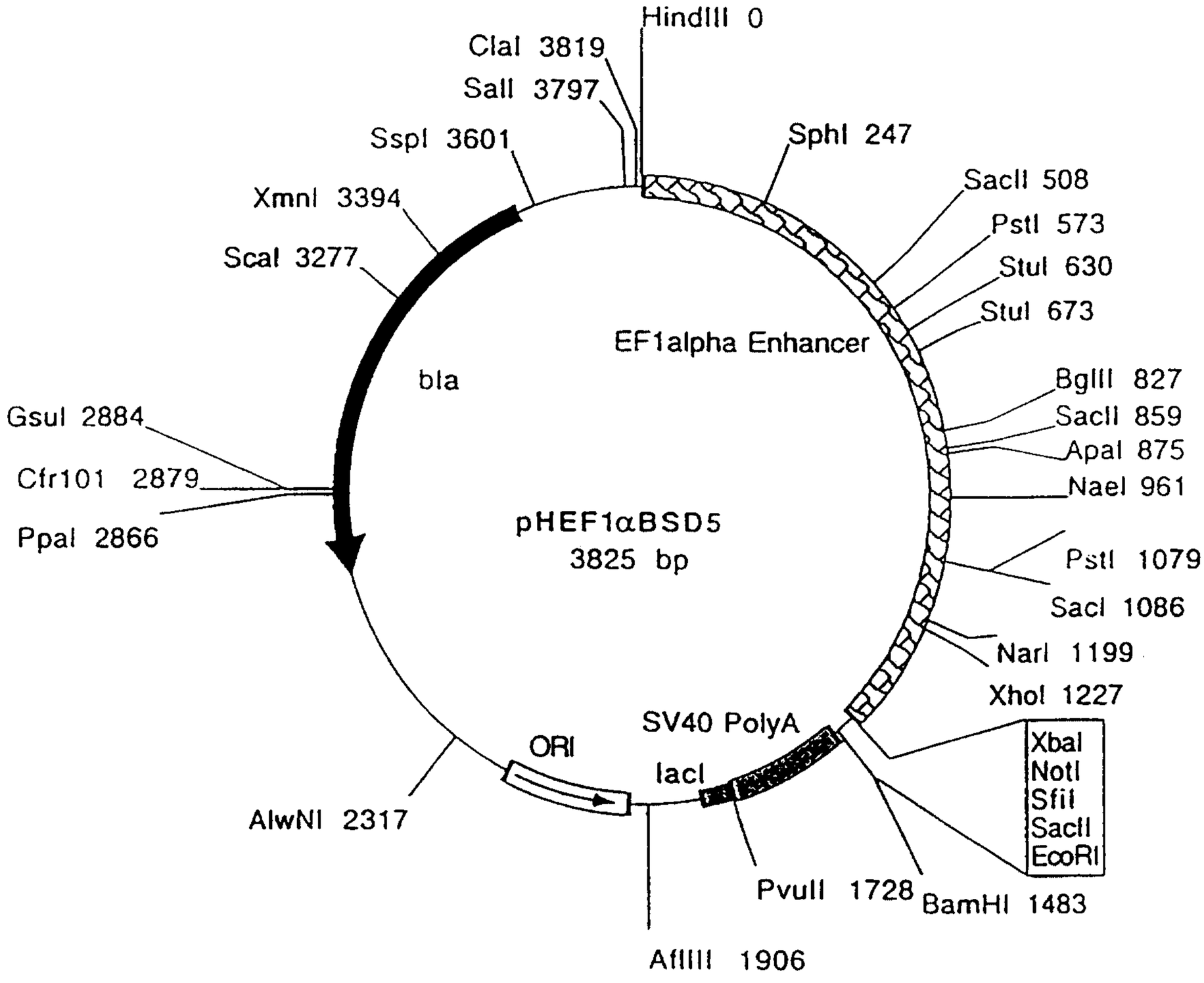
HindIII 0
ClaI 3819
SalI 3797
SspI 3601
SphI 247
XmnI 3394
SacII 508
ScaI 3277
PstI 573
StuI 630
StuI 673
EF1alpha Enhancer
bla
BglII 827
SacII 859
ApaI 875
GsuI 2884
NaeI 961
Cfr101 2879
Ppal 2866
pHEF1αBSD5
3825 bp
PstI 1079
SacI 1086
NarI 1199
XhoI 1227
SV40 PolyA
ORI
lacI
XbaI
NotI
SfiI
SacII
EcoRI
AlwNI 2317
PvuII 1728
BamHI 1483
AflIII 1906

FIGURE 9

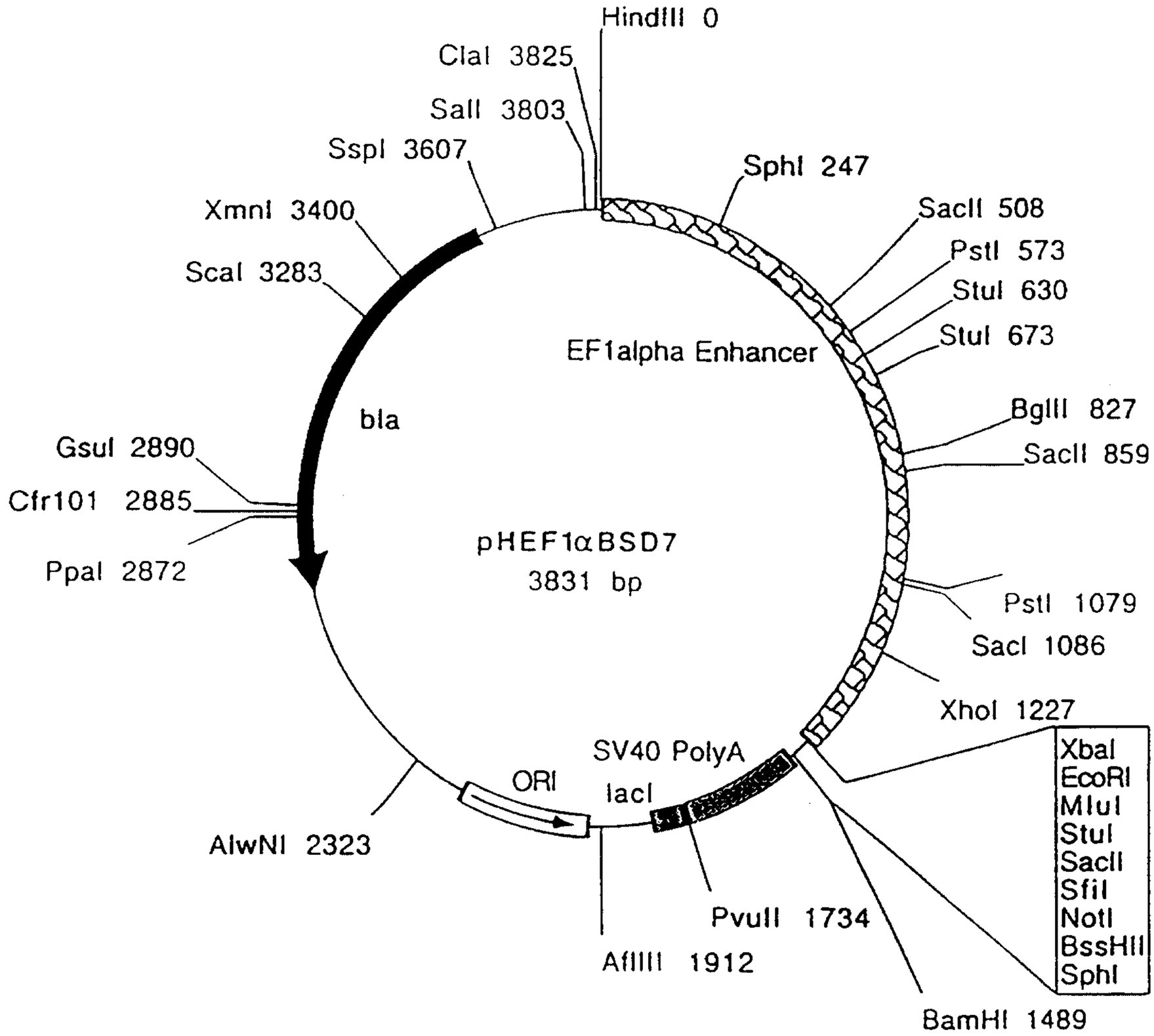
HindIII 0
ClaI 3825
SalI 3803
SspI 3607
XmnI 3400
ScaI 3283
SphI 247
SacII 508
PstI 573
StuI 630
StuI 673
EF1alpha Enhancer
BglII 827
SacII 859
bla
GsuI 2890
Cfr101 2885
Ppal 2872
pHEF1αBSD7
3831 bp
PstI 1079
SacI 1086
XhoI 1227
XbaI
EcoRI
MluI
StuI
SacII
SfiI
NotI
BssHII
SphI
SV40 PolyA
ORI
lacI
AlwNI 2323
PvuII 1734
AflIII 1912
BamHI 1489

FIGURE 10

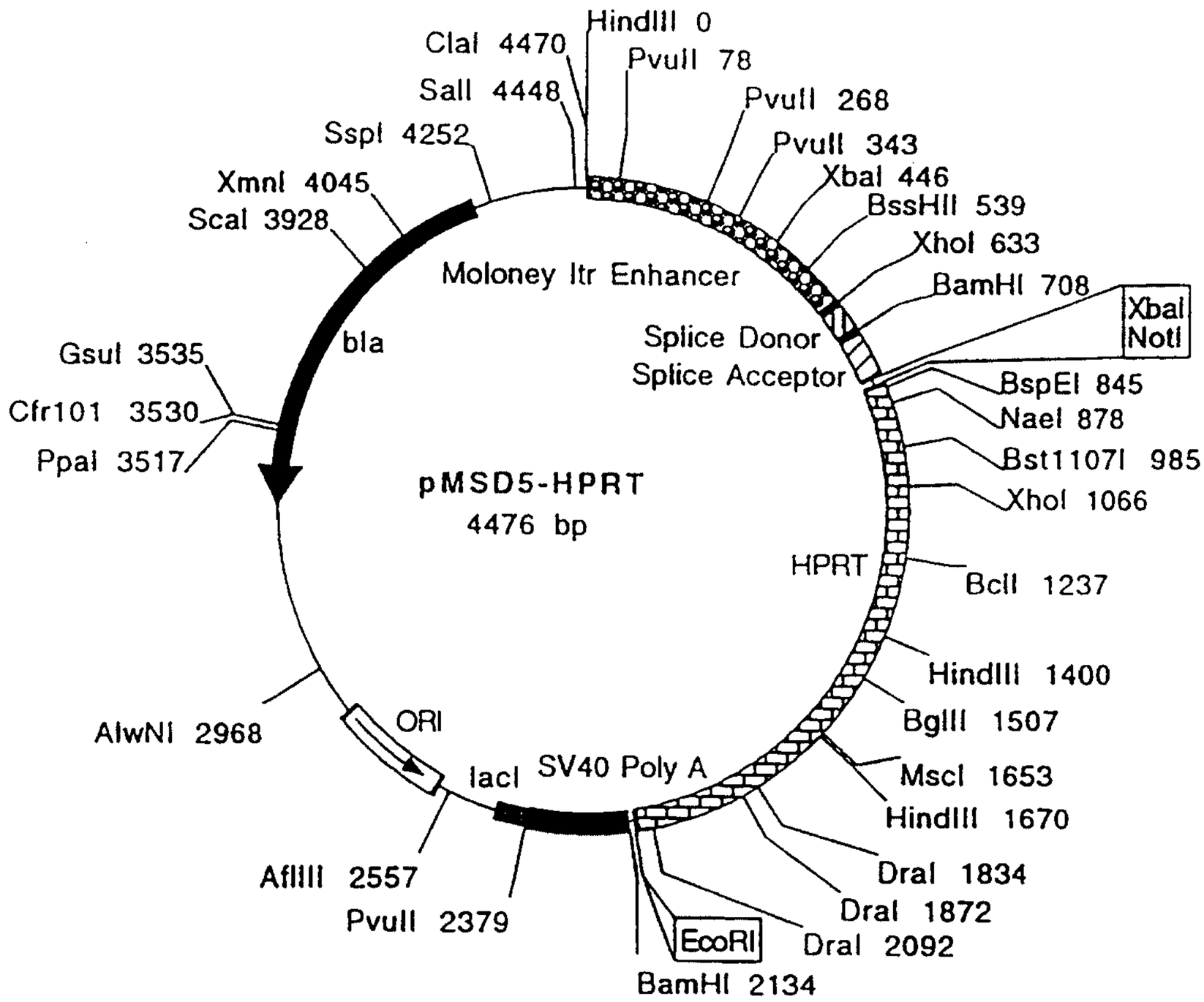
HindIII 0
ClaI 4470
PvuII 78
SalI 4448
PvuII 268
SspI 4252
PvuII 343
XmnI 4045
XbaI 446
ScaI 3928
BssHII 539
XhoI 633
Moloney ltr Enhancer
BamHI 708
XbaI
NotI
Splice Donor
Splice Acceptor
bla
GsuI 3535
BspEI 845
Cfr101 3530
NaeI 878
Ppal 3517
Bst1107I 985
pMSD5-HPRT
4476 bp
XhoI 1066
HPRT
BclI 1237
HindIII 1400
AlwNI 2968
ORI
BglII 1507
lacI
SV40 Poly A
MscI 1653
HindIII 1670
AflIII 2557
DraI 1834
PvuII 2379
DraI 1872
EcoRI
DraI 2092
BamHI 2134

FIGURE 11

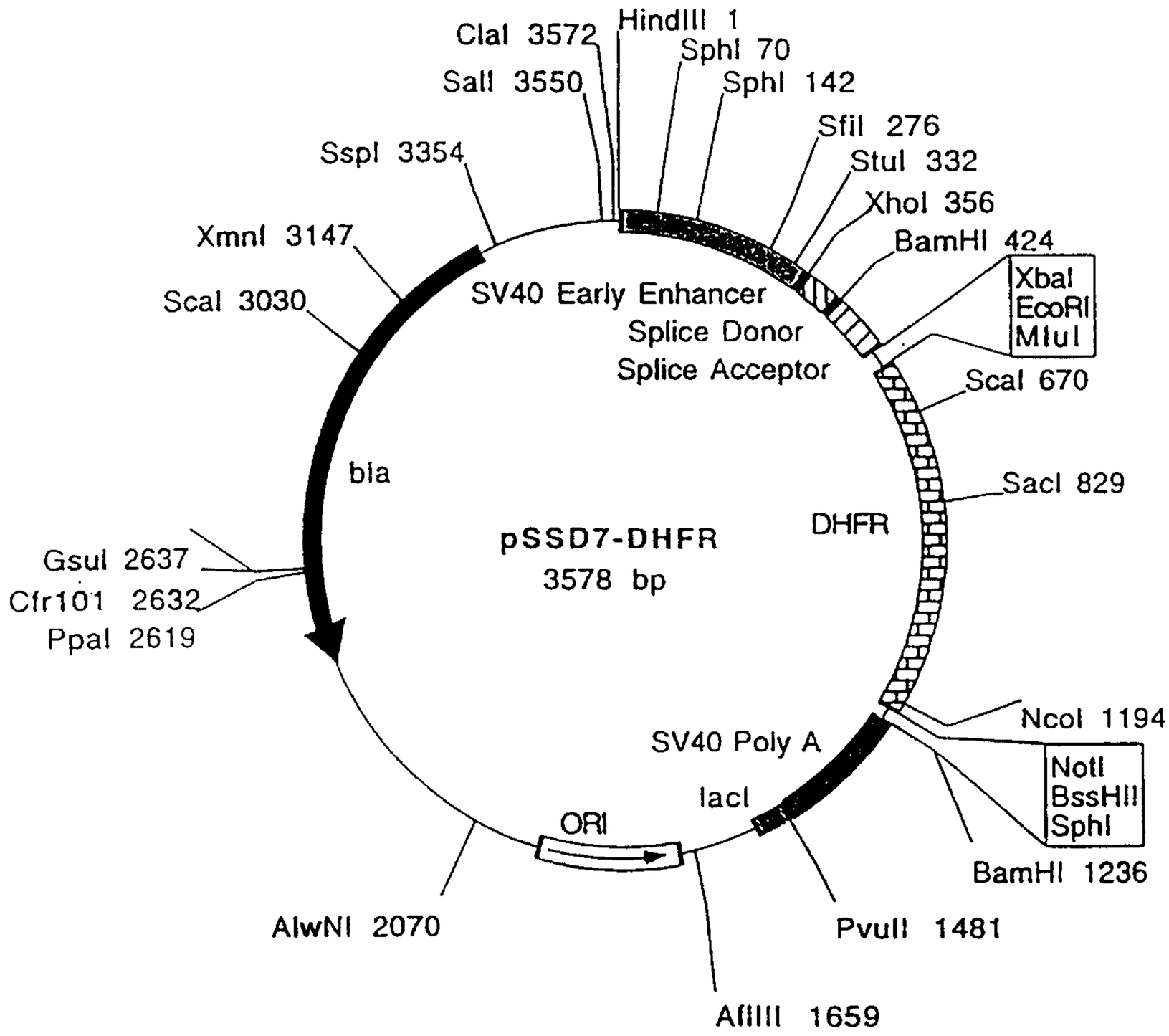
HindIII 1
ClaI 3572
SphI 70
SalI 3550
SphI 142
SfiI 276
SspI 3354
StuI 332
XhoI 356
BamHI 424
XmnI 3147
XbaI
EcoRI
MluI
ScaI 3030
SV40 Early Enhancer
Splice Donor
Splice Acceptor
ScaI 670
bla
SacI 829
DHFR
pSSD7-DHFR
3578 bp
GsuI 2637
Cfr10I 2632
PpaI 2619
NcoI 1194
SV40 Poly A
NotI
BssHII
SphI
lacI
ORI
BamHI 1236
AlwNI 2070
PvuII 1481
AflIII 1659

FIGURE 12

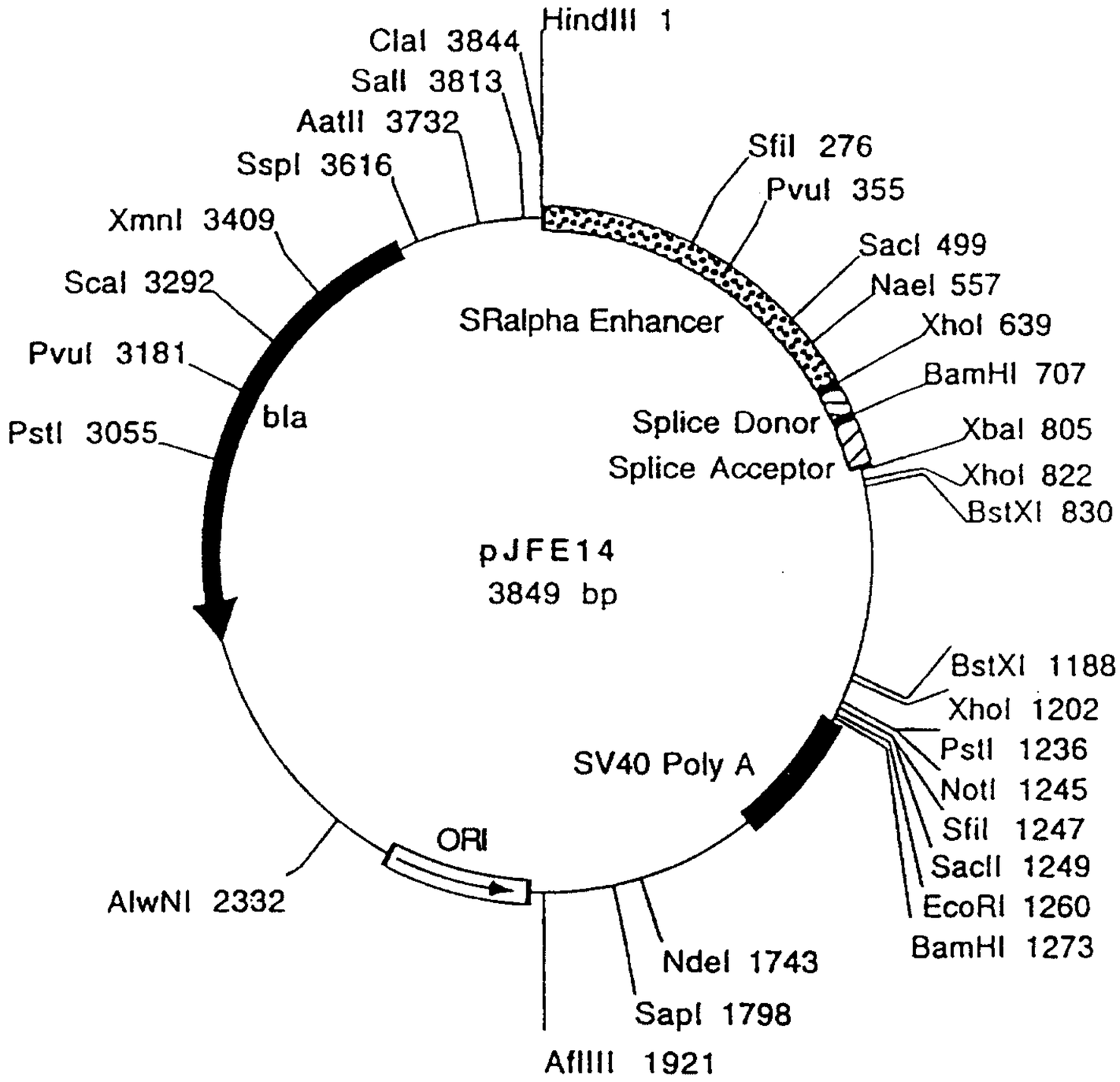
HindIII 1
ClaI 3844
SalI 3813
AatII 3732
SspI 3616
XmnI 3409
ScaI 3292
PvuI 3181
PstI 3055
bla
SfiI 276
PvuI 355
SacI 499
NaeI 557
XhoI 639
BamHI 707
XbaI 805
XhoI 822
BstXI 830
SRalpha Enhancer
Splice Donor
Splice Acceptor
pJFE14
3849 bp
BstXI 1188
XhoI 1202
PstI 1236
NotI 1245
SfiI 1247
SacII 1249
EcoRI 1260
BamHI 1273
SV40 Poly A
ORI
AlwNI 2332
NdeI 1743
SapI 1798
AflIII 1921

FIGURE 13

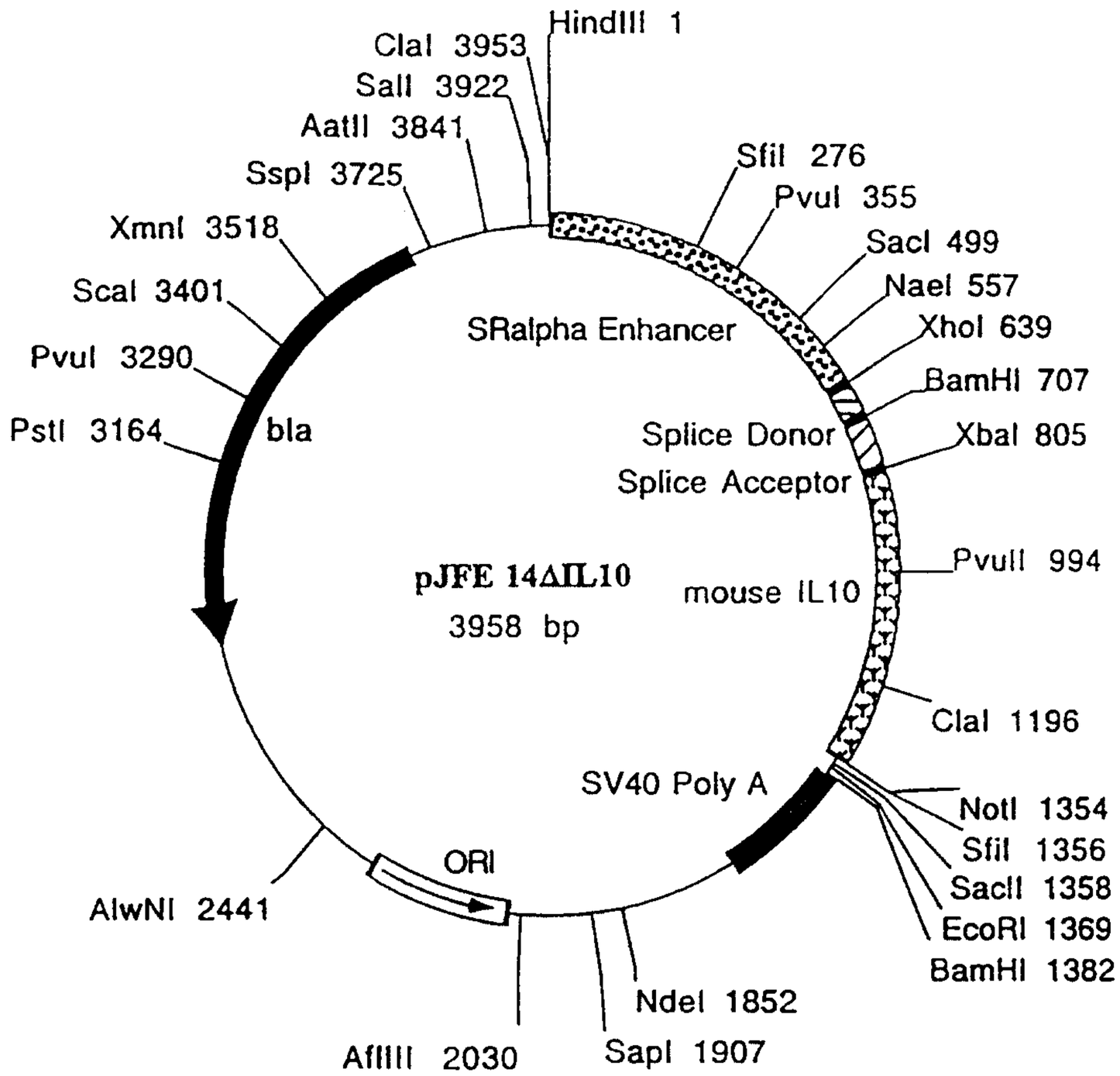
HindIII 1
ClaI 3953
SalI 3922
AatII 3841
SspI 3725
XmnI 3518
ScaI 3401
PvuI 3290
PstI 3164
bla
AlwNI 2441
AflIII 2030
ORI
NdeI 1852
SapI 1907
SV40 Poly A
SfiI 276
PvuI 355
SacI 499
NaeI 557
XhoI 639
BamHI 707
XbaI 805
SRalpha Enhancer
Splice Donor
Splice Acceptor
pJFE 14ΔIL10
3958 bp
mouse IL10
PvuII 994
ClaI 1196
NotI 1354
SfiI 1356
SacII 1358
EcoRI 1369
BamHI 1382

FIGURE 14

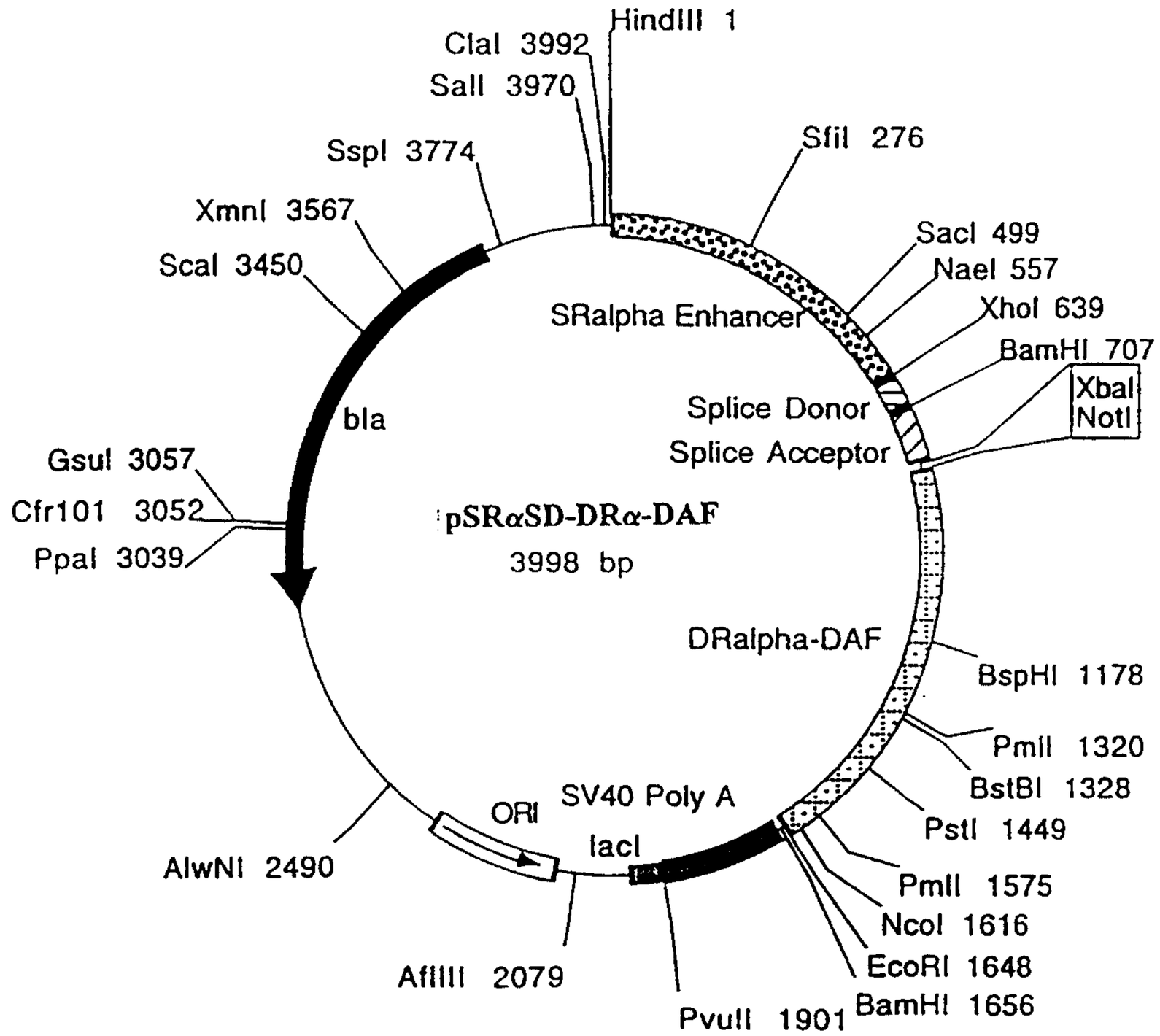
HindIII 1
ClaI 3992
SalI 3970
SspI 3774
XmnI 3567
ScaI 3450
SfiI 276
SacI 499
NaeI 557
XhoI 639
BamHI 707
XbaI
NotI
SRalpha Enhancer
Splice Donor
Splice Acceptor
bla
GsuI 3057
Cfr10I 3052
PpaI 3039
pSRαSD-DRα-DAF
3998 bp
DRalpha-DAF
BspHI 1178
PmlI 1320
BstBI 1328
PstI 1449
PmlI 1575
NcoI 1616
EcoRI 1648
BamHI 1656
SV40 Poly A
ORI
lacI
AlwNI 2490
AflIII 2079
PvuII 1901

FIGURE 15

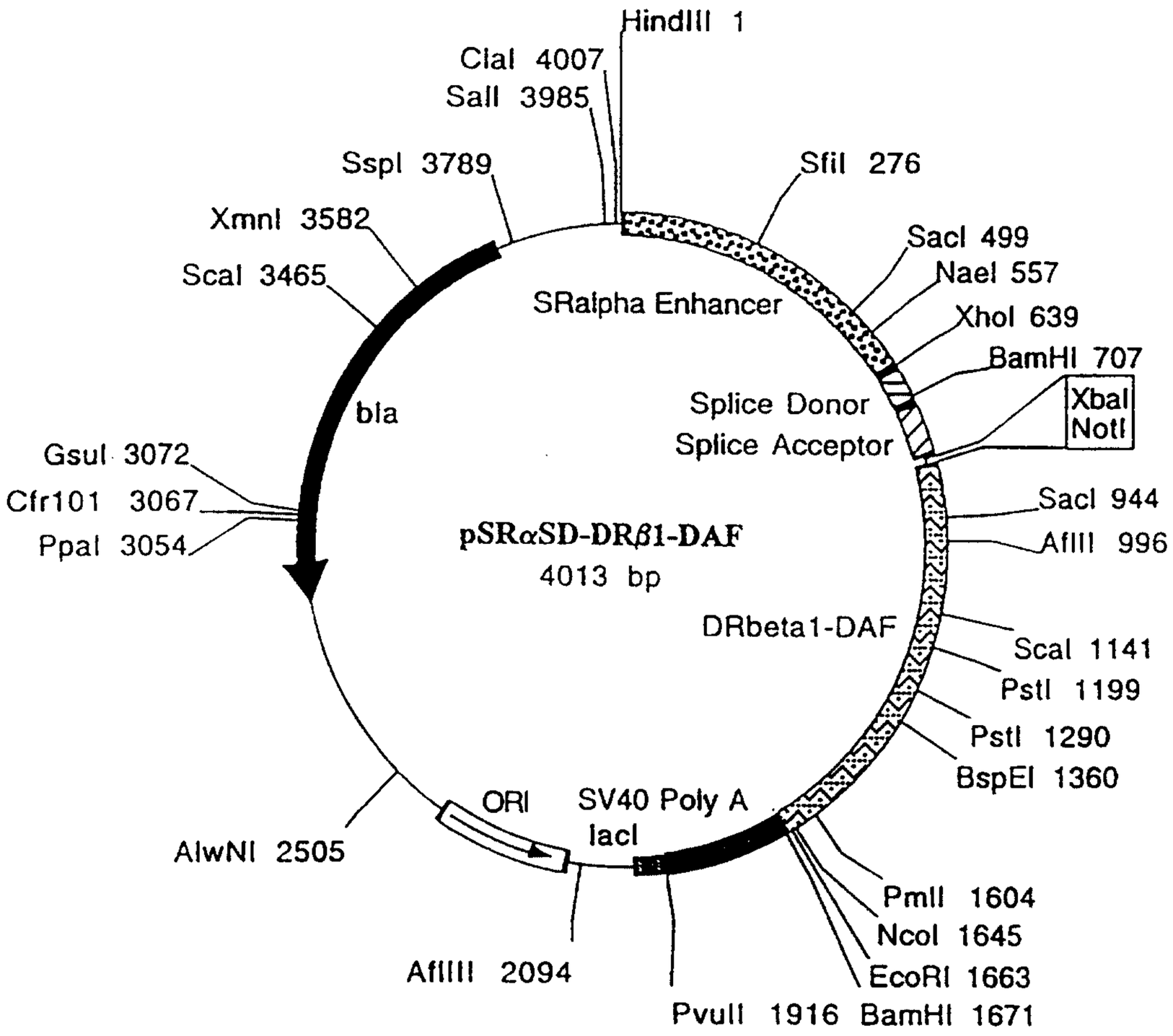
HindIII 1
ClaI 4007
SalI 3985
SspI 3789
SfiI 276
XmnI 3582
SacI 499
ScaI 3465
NaeI 557
SRalpha Enhancer
XhoI 639
BamHI 707
bla
Splice Donor
XbaI
NotI
Splice Acceptor
GsuI 3072
Cfr101 3067
SacI 944
Ppal 3054
pSRαSD-DRβ1-DAF
AflII 996
4013 bp
DRbeta1-DAF
ScaI 1141
PstI 1199
PstI 1290
BspEI 1360
ORI
SV40 Poly A
lacI
AlwNI 2505
PmlI 1604
NcoI 1645
AflIII 2094
EcoRI 1663
PvuII 1916
BamHI 1671

FIGURE 16

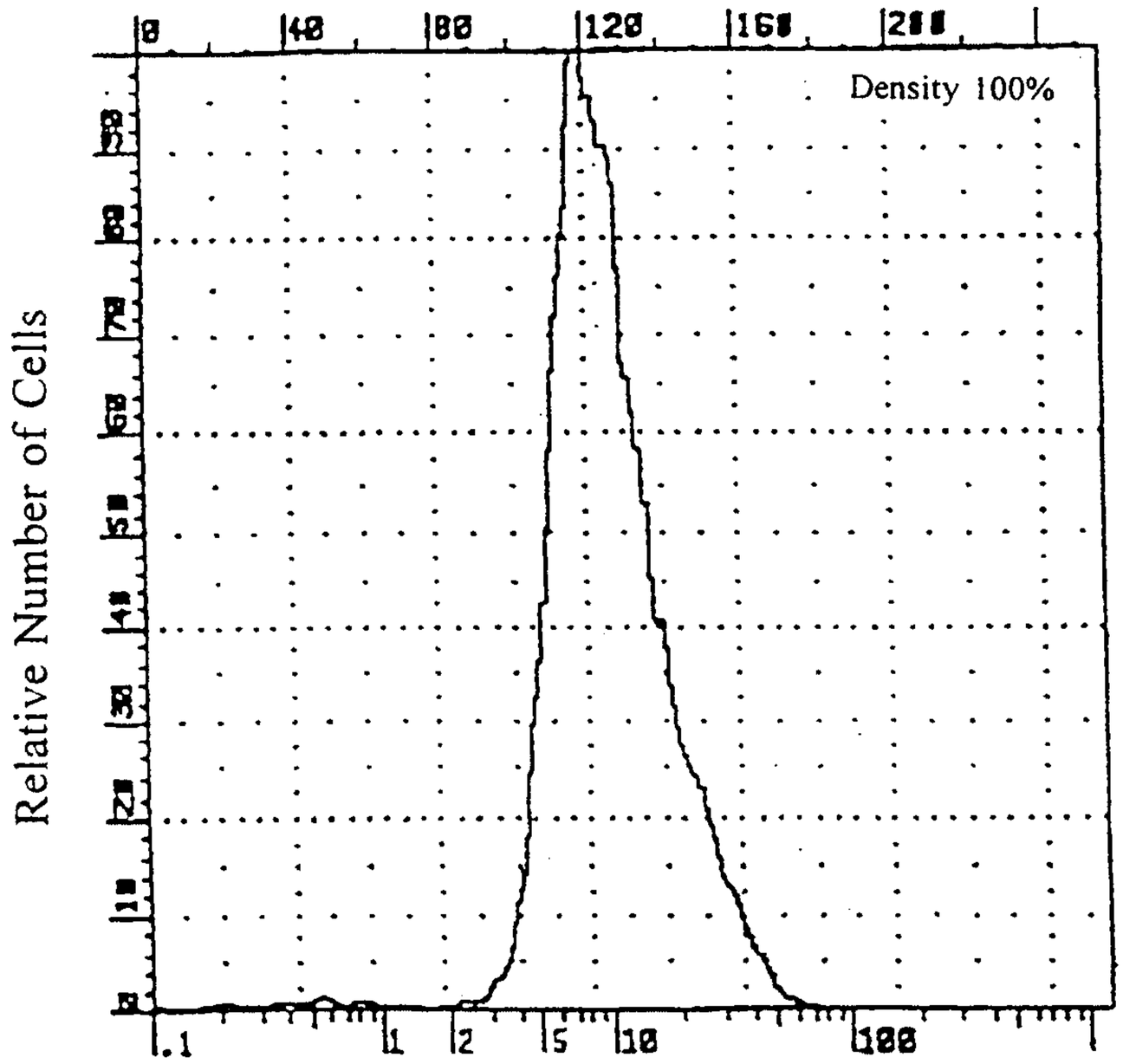

FIGURE 17

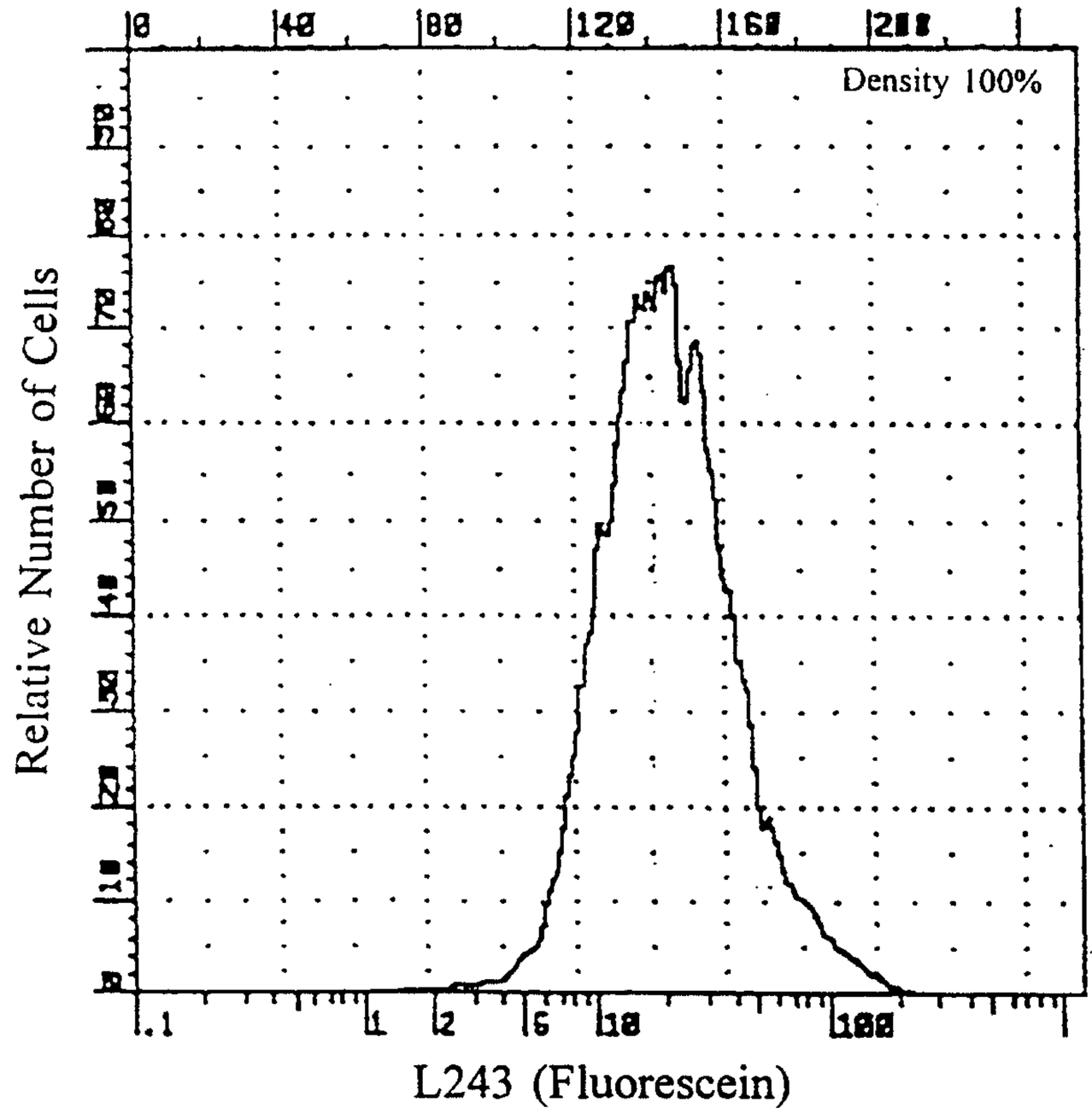

FIGURE 18

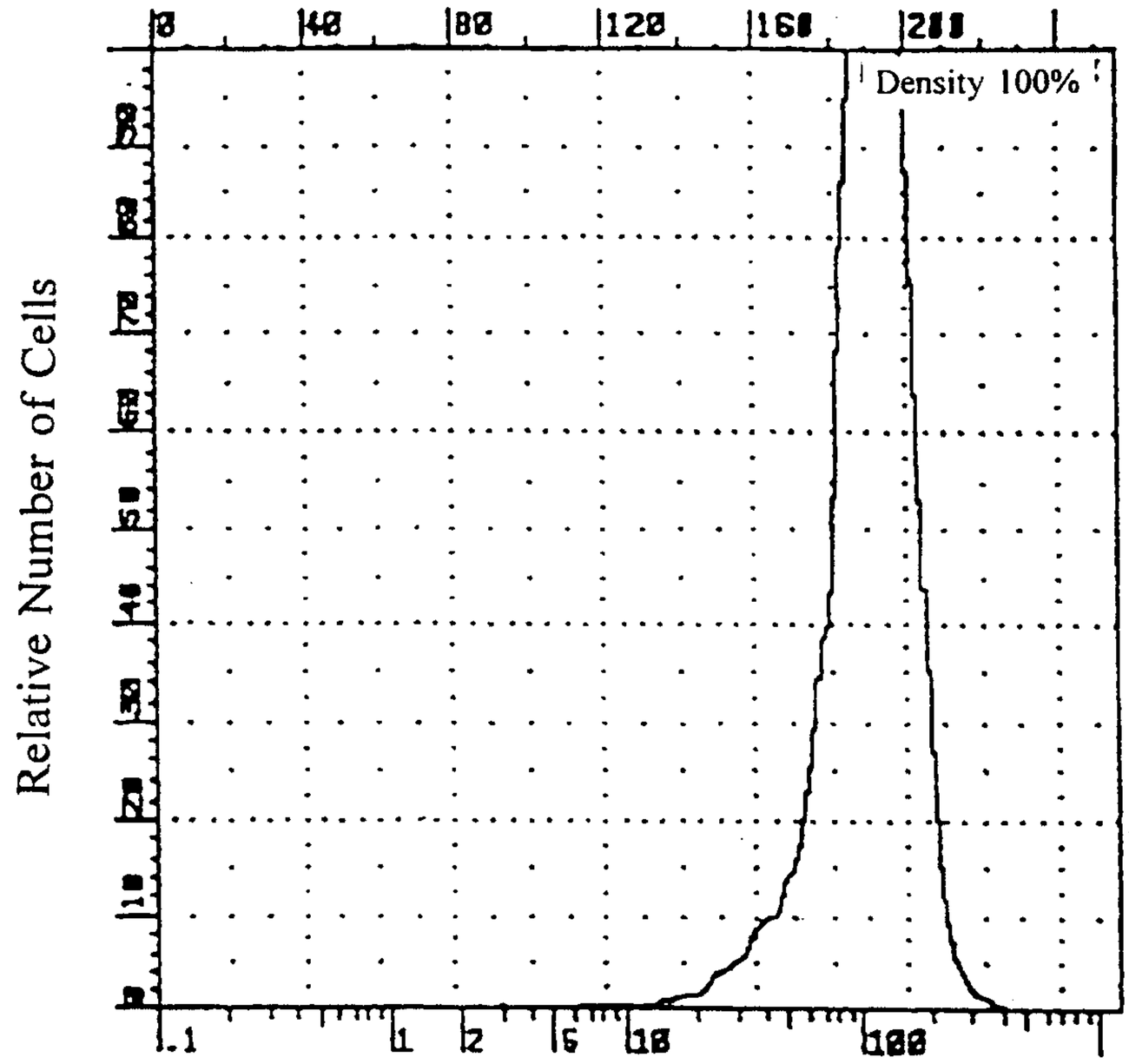

FIGURE 19

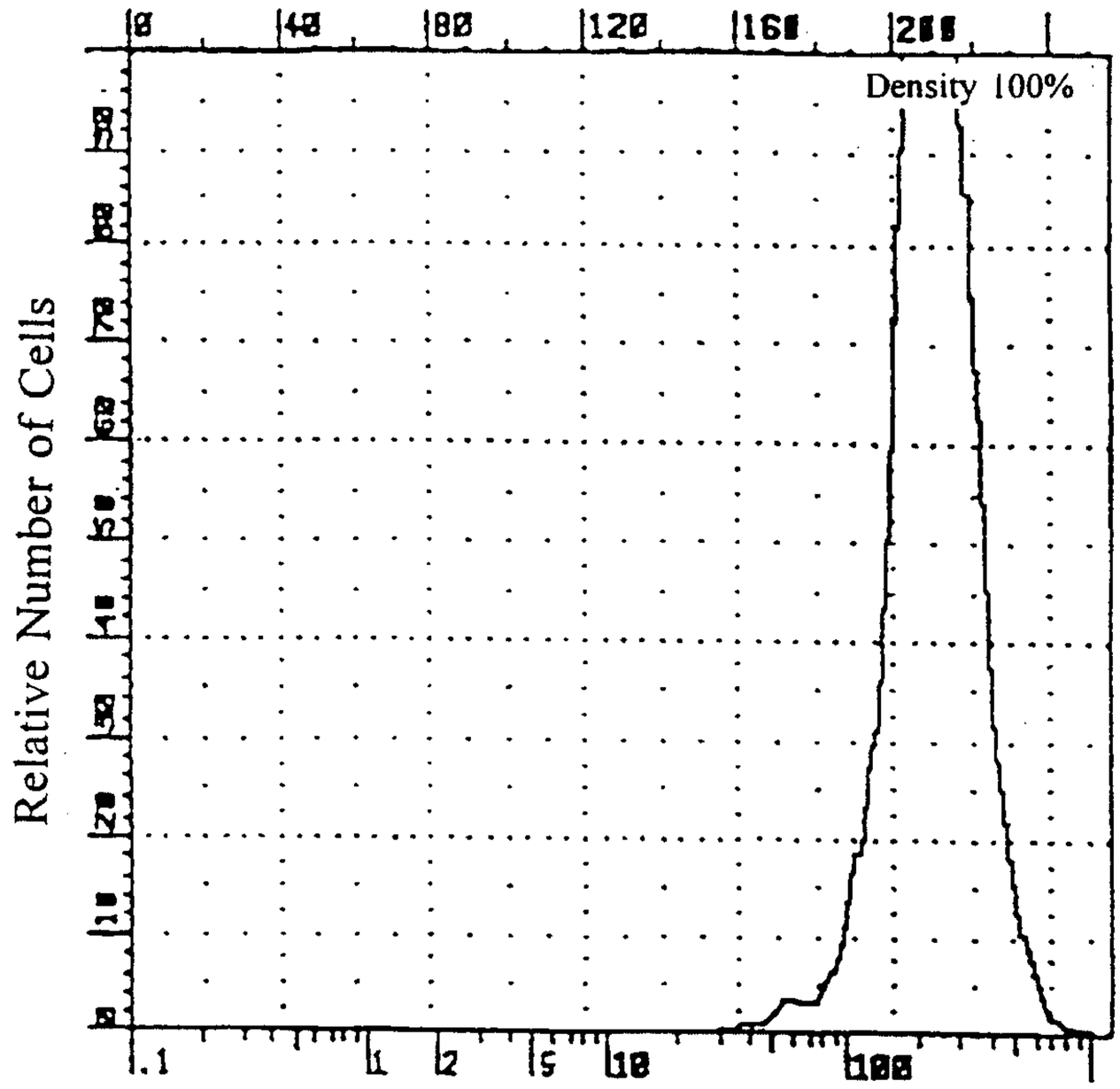

FIGURE 20

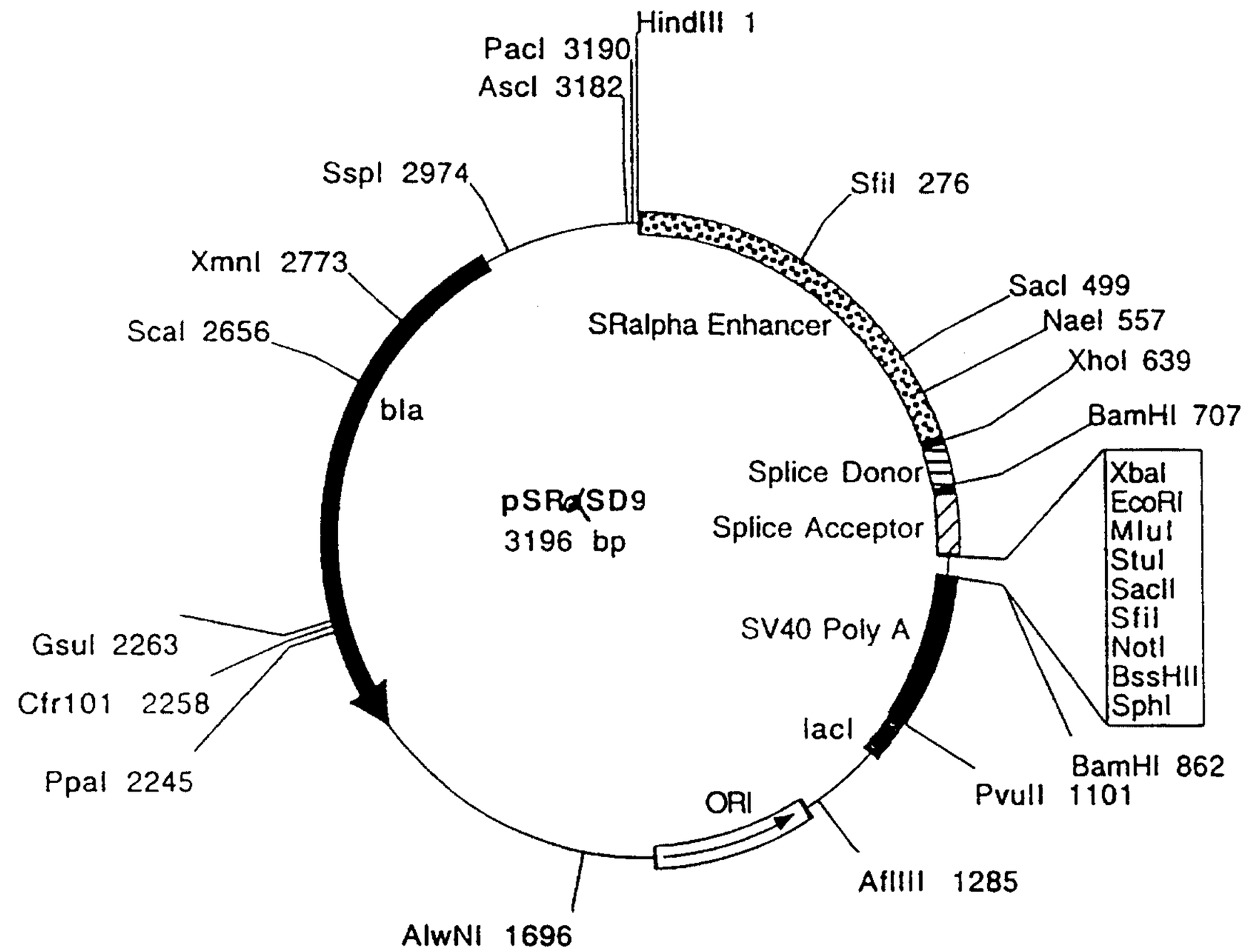
HindIII 1
PacI 3190
AscI 3182
SspI 2974
SfiI 276
XmnI 2773
SacI 499
ScaI 2656
NaeI 557
SRalpha Enhancer
XhoI 639
bla
BamHI 707
Splice Donor
pSRαSD9
3196 bp
Splice Acceptor
XbaI
EcoRI
MluI
StuI
SacII
SfiI
NotI
BssHII
SphI
GsuI 2263
SV40 Poly A
Cfr101 2258
lacI
BamHI 862
Ppal 2245
PvuII 1101
ORI
AflIII 1285
AlwNI 1696

FIGURE 21

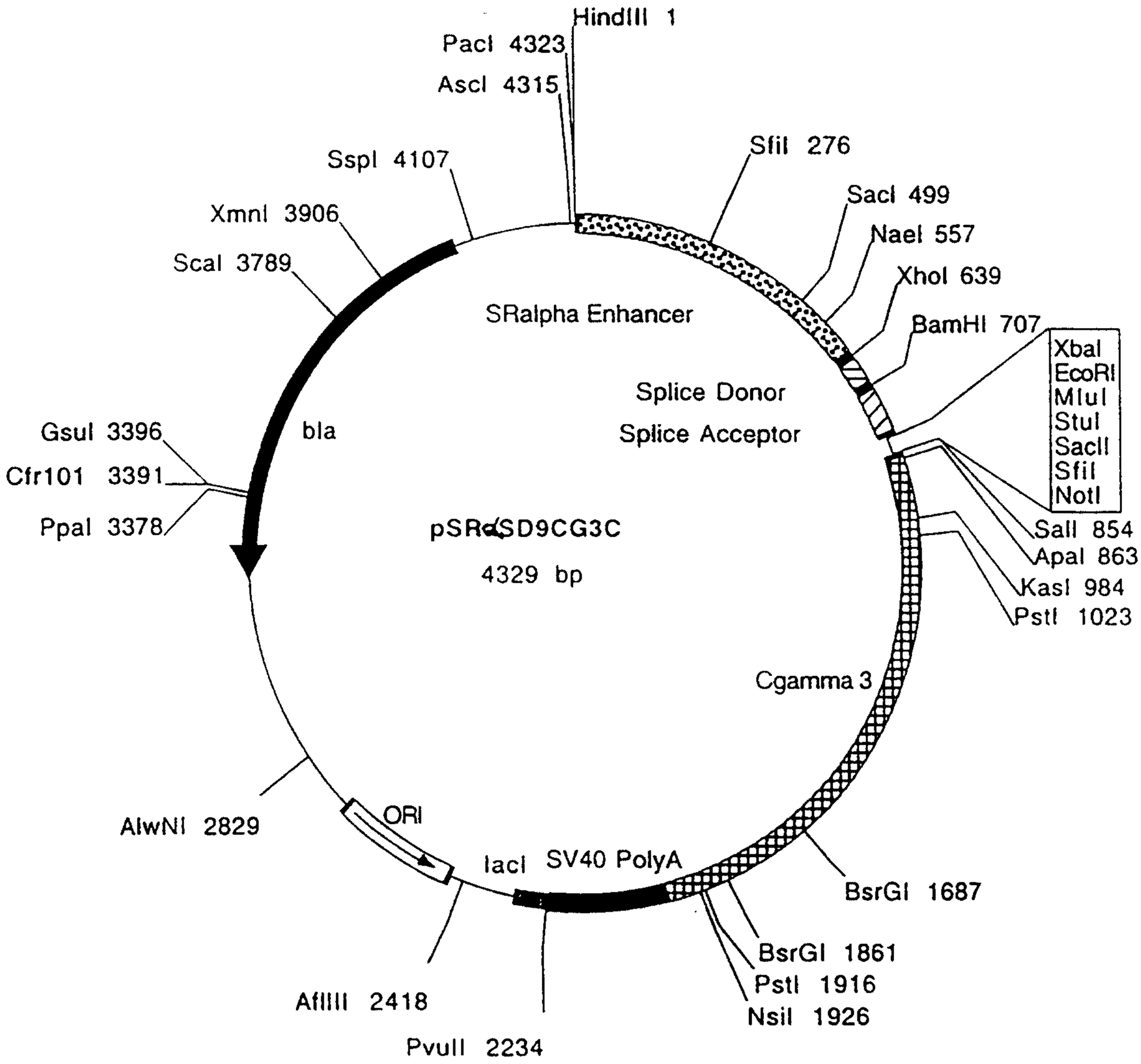
HindIII 1
PacI 4323
AscI 4315
SspI 4107
XmnI 3906
ScaI 3789
SfiI 276
SacI 499
NaeI 557
XhoI 639
BamHI 707
SRalpha Enhancer
Splice Donor
Splice Acceptor
XbaI
EcoRI
MluI
StuI
SacII
SfiI
NotI
SalI 854
ApaI 863
KasI 984
PstI 1023
Cgamma 3
BsrGI 1687
BsrGI 1861
PstI 1916
NsiI 1926
GsuI 3396
Cfr10I 3391
PpaI 3378
bla
pSRαSD9CG3C
4329 bp
AlwNI 2829
ORI
lacI
SV40 PolyA
AflIII 2418
PvuII 2234

FIGURE 22

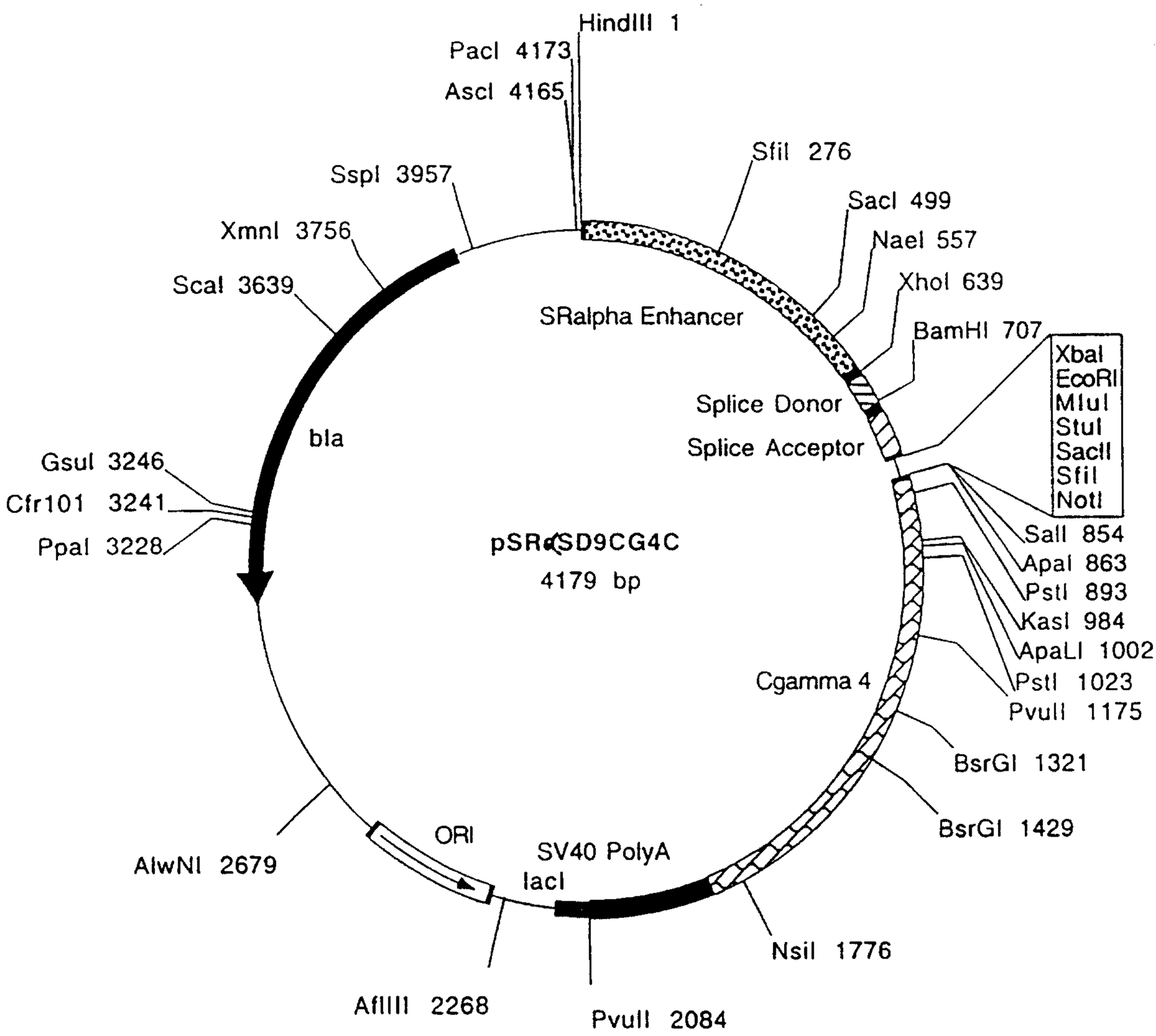
HindIII 1
PacI 4173
AscI 4165
SfiI 276
SspI 3957
SacI 499
XmnI 3756
NaeI 557
ScaI 3639
XhoI 639
SRalpha Enhancer
BamHI 707
XbaI
EcoRI
MluI
StuI
SacII
SfiI
NotI
Splice Donor
Splice Acceptor
bla
GsuI 3246
Cfr10I 3241
PpaI 3228
pSRαSD9CG4C
4179 bp
SalI 854
ApaI 863
PstI 893
KasI 984
ApaLI 1002
PstI 1023
PvuII 1175
Cgamma 4
BsrGI 1321
BsrGI 1429
AlwNI 2679
ORI
SV40 PolyA
lacI
NsiI 1776
AflIII 2268
PvuII 2084

FIGURE 23

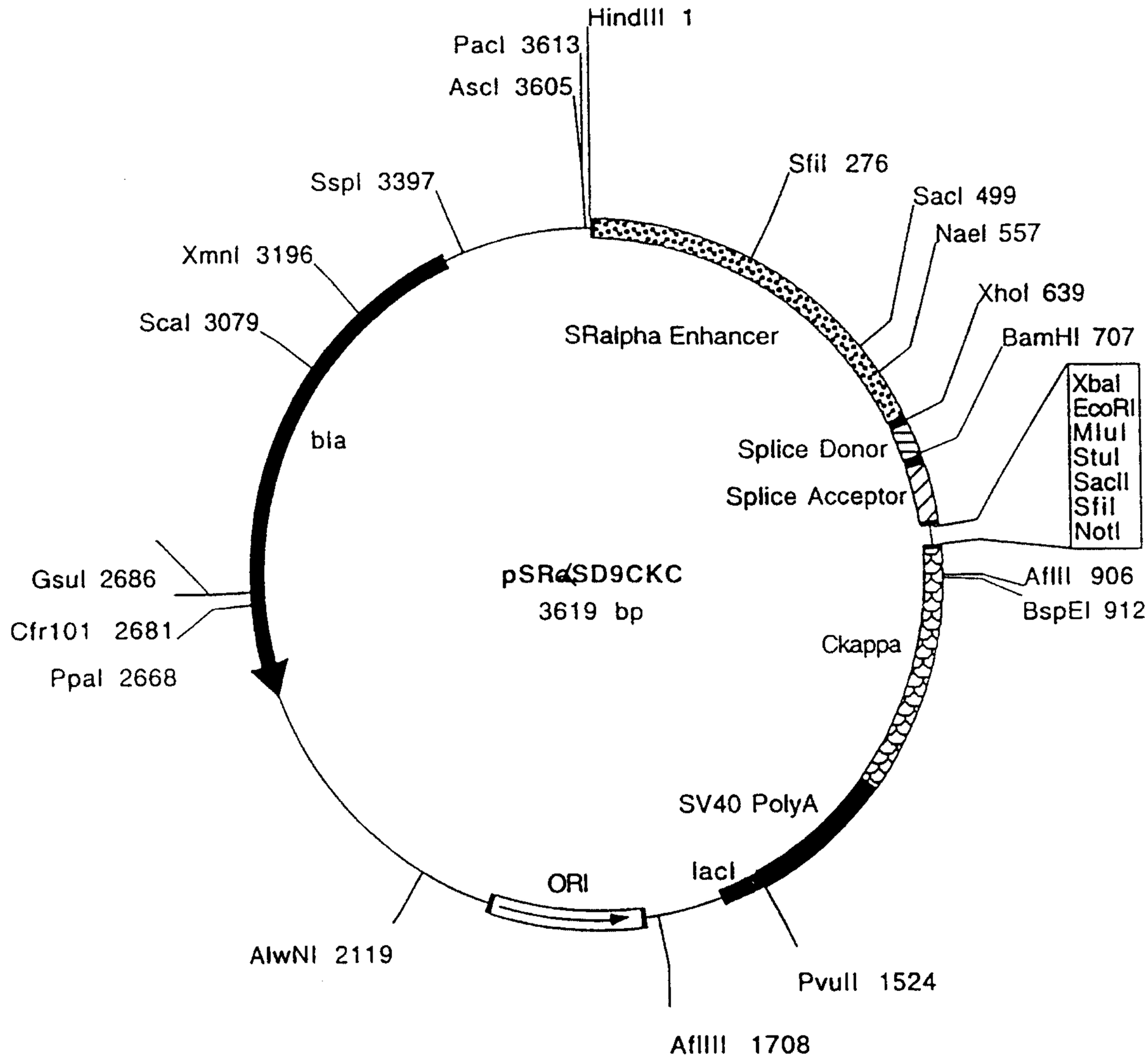
HindIII 1
PacI 3613
AscI 3605
SfiI 276
SspI 3397
SacI 499
NaeI 557
XmnI 3196
XhoI 639
ScaI 3079
SRalpha Enhancer
BamHI 707
XbaI
EcoRI
MluI
StuI
SacII
SfiI
NotI
bla
Splice Donor
Splice Acceptor
GsuI 2686
pSRαSD9CKC
3619 bp
AflII 906
BspEI 912
Cfr10I 2681
Ckappa
PpaI 2668
SV40 PolyA
ORI
lacI
AlwNI 2119
PvuII 1524
AflIII 1708

FIGURE 24

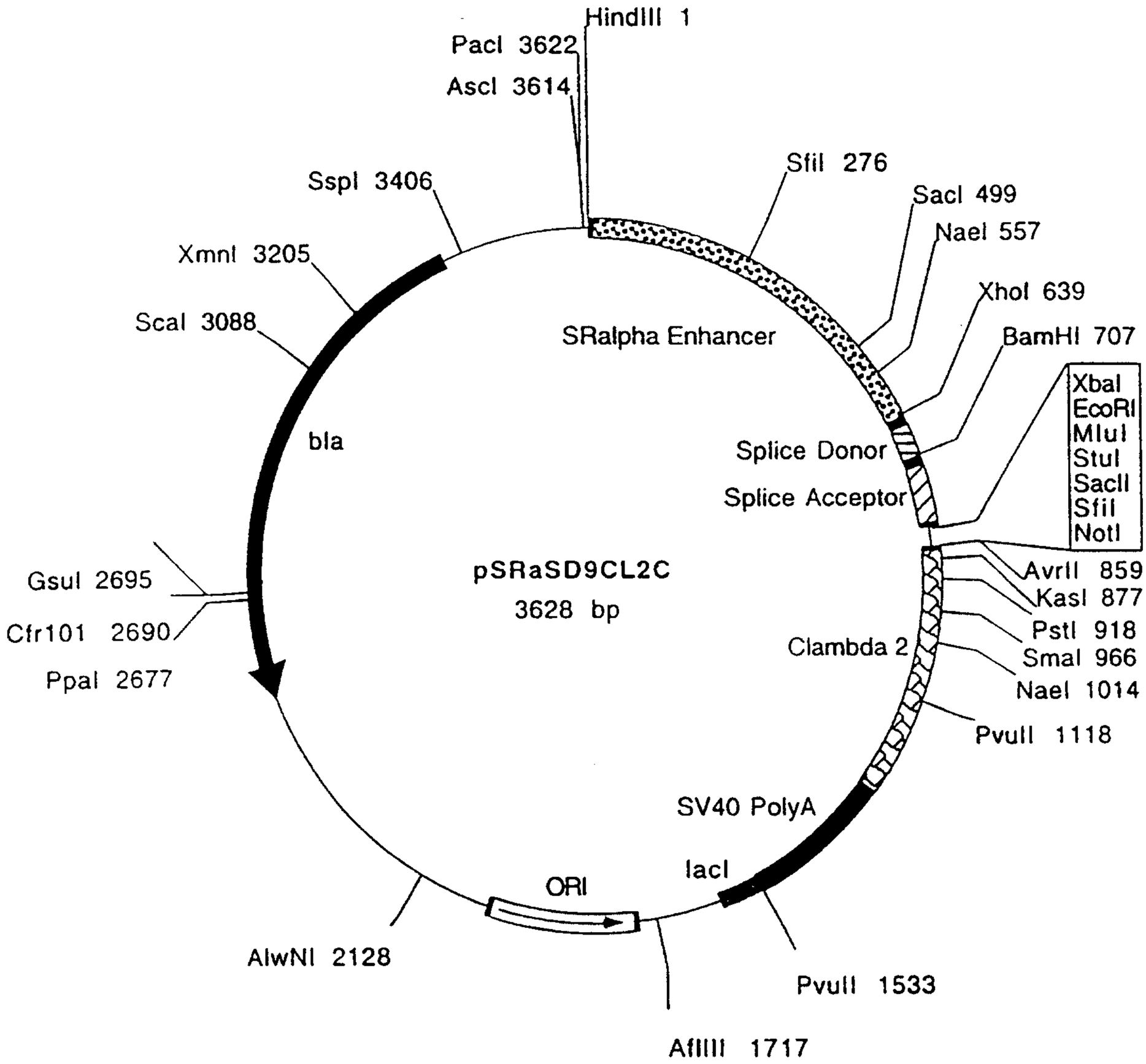
HindIII 1
PacI 3622
AscI 3614
SspI 3406
XmnI 3205
ScaI 3088
SfiI 276
SacI 499
NaeI 557
XhoI 639
BamHI 707
SRalpha Enhancer
XbaI
EcoRI
MluI
StuI
SacII
SfiI
NotI
Splice Donor
Splice Acceptor
bla
pSRaSD9CL2C
3628 bp
AvrII 859
KasI 877
PstI 918
SmaI 966
NaeI 1014
PvuII 1118
Clambda 2
GsuI 2695
Cfr101 2690
PpaI 2677
SV40 PolyA
lacI
ORI
AlwNI 2128
PvuII 1533
AflIII 1717

FIGURE 25

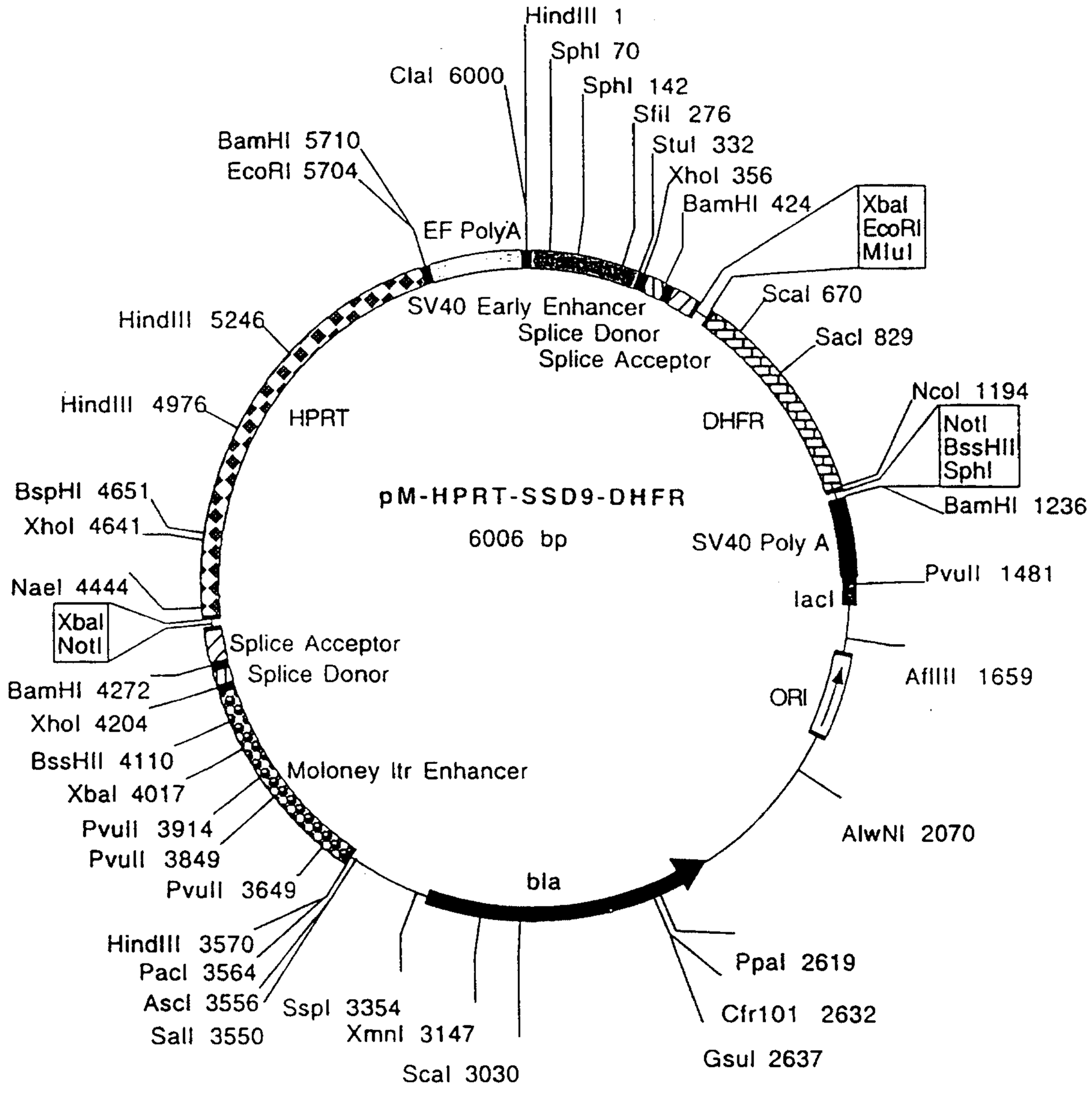
HindIII 1
SphI 70
SphI 142
SfiI 276
StuI 332
XhoI 356
BamHI 424
XbaI
EcoRI
MluI
ScaI 670
SacI 829
NcoI 1194
NotI
BssHII
SphI
BamHI 1236
PvuII 1481
AflIII 1659
AlwNI 2070
Ppal 2619
Cfr101 2632
GsuI 2637
ClaI 6000
BamHI 5710
EcoRI 5704
EF PolyA
SV40 Early Enhancer
Splice Donor
Splice Acceptor
DHFR
SV40 Poly A
lacI
ORI
bla
HindIII 5246
HindIII 4976
HPRT
BspHI 4651
XhoI 4641
NaeI 4444
XbaI
NotI
Splice Acceptor
Splice Donor
BamHI 4272
XhoI 4204
BssHII 4110
XbaI 4017
PvuII 3914
PvuII 3849
PvuII 3649
HindIII 3570
PacI 3564
AscI 3556
SalI 3550
SspI 3354
XmnI 3147
ScaI 3030
Moloney ltr Enhancer
pM-HPRT-SSD9-DHFR
6006 bp

FIGURE 26