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

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2736498 A1 2010/03/18

(21) 2 736 498

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2009/09/11
(87) Date publication PCT/PCT Publication Date: 2010/03/18
(85) Entrée phase nationale/National Entry: 2011/03/08
(86) N° demande PCT/PCT Application No.: US 2009/056637
(87) N° publication PCT/PCT Publication No.: 2010/030861
(30) Priorité/Priority: 2008/09/11 (US61/096,185)

(51) Cl.Int./Int.Cl. *C12N 15/63* (2006.01), *C07K 14/47* (2006.01), *C12N 15/00* (2006.01), *C12N 15/12* (2006.01), *C12P 21/02* (2006.01)

(71) Demandeur/Applicant:
CALRETEX, LLC, US

(72) Inventeurs/Inventors:
MICHALAK, MAREK, CA;
DABROSKA, MONIKA, CA

(74) Agent: ROBIC

(54) Titre : COMPOSITIONS ET METHODES DE PURIFICATION DE LA CALRETICULINE
(54) Title: COMPOSITIONS AND METHODS FOR PURIFYING CALRETICULIN

(57) Abrégé/Abstract:
The present invention relates to compositions, methods for expressing, and related methods for purifying calretic-ulin that is free of an affinity label or tag (i.e., non- tagged calreticulin). The invention provides useful methods for commercial production of human calreticulin in a bacterial expression system such as Escherichia coli.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 March 2010 (18.03.2010)

PCT

(10) International Publication Number
WO 2010/030861 A2

(51) **International Patent Classification**: Not classified

(21) **International Application Number**:
PCT/US2009/056637

(22) **International Filing Date**:
11 September 2009 (11.09.2009)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/096,185 11 September 2008 (11.09.2008) US

(71) **Applicant** *(for all designated States except US)*: **CALRETEX, LLC** [US/US]; c/o Cahn Medical Technologies LLC, 14 West 68th St., Apt. 12, New York, NY 10023 (US).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **MICHALAK, Marek** [CA/CA]; 3748-117 Street, Edmonton, Alberta, T6J 1S7 (CA). **DABROSKA, Monika** [CA/CA]; 283 River Point, Edmonton, Alberta, T5A 4Z3 (CA).

(74) **Agents**: **LUDWIG, S., Peter** et al.; Darby & Darby P.C., P.O. Box 770, Church Street Station, New York, NY 10008-0770 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2010/030861 A2

(54) **Title**: COMPOSITIONS AND METHODS FOR PURIFYING CALRETICULIN

(57) **Abstract**: The present invention relates to compositions, methods for expressing, and related methods for purifying calreticulin that is free of an affinity label or tag (*i.e.*, non- tagged calreticulin). The invention provides useful methods for commercial production of human calreticulin in a bacterial expression system such as *Escherichia coli.*

COMPOSITIONS AND METHODS FOR PURIFYING CALRETICULIN

This application claims priority to U.S. provisional application No. 61/096,185, filed September 11, 2008, the contents of which are hereby incorporated by reference in their entirety.

FIELD OF THE INVENTION

[0001] The present invention relates to compositions and methods for expressing non-tagged human calreticulin and methods for purification of the non-tagged protein.

BACKGROUND OF THE INVENTION

[0002] Purified proteins are produced in systems for a wide range of applications in biology and biotechnology. These include research into cellular and molecular function, production of proteins as biopharmaceuticals or diagnostic reagents, and modification of the traits or phenotypes of livestock and crops.

[0003] Affinity protein tags are used as tools for purifying proteins from crude extracts. Protein tags are peptide sequences genetically grafted onto a recombinant protein. Often these tags are removable by chemical agents or by enzymatic means, such as proteolysis or intein splicing. Tags are attached to proteins for various purposes. Elutable affinity tags to purify proteins can be utilized in diverse cell systems including in *Escherichia coli*, yeast, Drosophila, and HeLa extracts.

[0004] Affinity tags are appended to proteins so that they can be purified from their crude biological source using an affinity technique. These include chitin binding protein (CBP), maltose binding protein (MBP), and glutathione-s-transferase (GST). The poly(His) tag is the most widely-used protein tag and it binds to metal matrices.

[0005] Solubilization tags are used, especially for recombinant proteins expressed in chaperone-deficient species such as *E. coli*, to assist in the proper folding in proteins and keep them from precipitating. These include thioredoxin (TRX) and poly(NANP). Some affinity tags have a dual role as a solubilization agent, such as MBP, and GST.

[0006] Chromatography tags are used to alter chromatographic pr
to afford different resolution across a particular separation technique. Often, these consist of polyanionic amino acids, such as FLAG tag.

[0007] Epitope tags are short peptide sequences which are chosen because high-affinity antibodies can be reliably produced in many different species. These are usually derived from viral genes, which explain their high immunoreactivity. Epitope tags include V5-tag, c-myc-tag, and HA-tag. These tags are particularly useful for western blotting and immunoprecipitation experiments, although they also find use in antibody purification.

[0008] Fluorescence tags are used to give visual readout on a protein. GFP and its variants are the most commonly used fluorescence tags. More advanced applications of GFP include using it as a folding reporter (fluorescent if folded, colorless if not).

[0009] Protein tags find many other usages, such as specific enzymatic modification (such as biotin ligase tags) and chemical modification (FLaSH) tag. Often tags are combined to produce multifunctional modifications of the protein. However, with the addition of each tag comes the risk that the native function of the protein may be abolished or compromised by interactions with the tag.

[0010] For pharmaceutical production, tagged proteins are undesirable and require an extra step of removal of the tags after the purification step. The expression and production of the protein product with the tag can cause conformation changes in the protein resulting in diminished activity, even after the removal of the tag. Additionally, it is difficult to completely remove the tagged protein, leading to the presence of residual impurities in the protein sample. Thus, there is a need for expression and purification systems for producing heterologous proteins without relying on affinity tags for purification. Additionally, there is a need for economical and efficient methods for producing heterologous proteins in bacterial systems.

SUMMARY OF THE INVENTION

[0011] The present invention relates to compositions and methods for expressing human calreticulin and methods for purification of the protein product that is free of an affinity tag or label.

[0012] In certain embodiments, the invention encompasses a method for producing calreticulin comprising: (a) culturing a microorganism having a calreticulin gene under suitable culture conditions; (b) inducing expression of the calreticulin gene to produce a calreticulin protein product that is free of an affinity label or tag; and (c) purifying the calreticulin product. In certain embodiments, the microorganism is bacteria. In yet additional embodiments, the bacteria is *Escherichia coli*.

[0013] In certain embodiments, the calreticulin gene is human calreticulin. In yet additional embodiments, the calreticulin gene comprises SEQ ID NO:3.

[0014] In certain embodiments, the purifying comprises at least one ion chromatography purification followed by at least one hydrophobic chromatography purification. In certain embodiments, the purifying comprises two sequential ion chromatography purifications followed by at least one hydrophobic chromatography purification. In certain embodiments, the hydrophobic chromatography step comprises utilizing 20-25% ammonium sulfate for elution.

[0015] In yet additional embodiments, the invention relates to a method for expressing calreticulin in a prokaryotic cell, which comprises:

a) culturing a prokaryotic cell comprising a vector under conditions conducive to gene expression, wherein the vector comprises a nucleic acid encoding calreticulin, an origin of replication, and at least one selectable marker, and wherein the nucleic acid is operatively associated with an expression control sequence;

(b) inducing expression of the calreticulin gene to produce a calreticulin protein product that is free of an affinity label or tag; and

(c) purifying said calreticulin protein product.

[0016] In certain embodiments, the prokaryotic cell is *E. coli*. In certain embodiments

the selectable marker is Kanamycin or Ampicillin. In yet additio
purifying comprises at least one ion chromatography purification followed by at least one hydrophobic chromatography purification. In certain embodiments, the purifying comprises two sequential ion chromatography purifications followed by at least one hydrophobic chromatography purification. In certain embodiments, the hydrophobic chromatography step comprises utilizing 20-25% ammonium sulfate for elution.

[0017] In yet further embodiments, the invention encompasses a vector comprising an origin of replication, at least one selectable marker, and a multiple cloning site comprising a nucleic acid encoding calreticulin, and wherein the nucleic acid is operatively associated with an expression control sequence. In certain embodiments, the calreticulin is human calreticulin. In additional embodiments, the calreticulin gene comprises SEQ ID NO:3. In certain embodiments, the selectable marker is Kanamycin or Ampicillin. In certain embodiments, the vector comprises the vector map of Figure 5. In yet additional embodiments, the origin of replication is an *E. coli* origin.

[0018] In yet further embodiments, the invention relates to a host cell transformed with any of the vectors described herein.

[0019] In yet additional embodiments, the invention encompasses isolated and purified calreticulin, prepared by any one of the methods described herein.

BRIEF DESCRIPTION OF DRAWINGS

[0020] Figures 1A-D are photos of Coomassie Blue gels with protein samples from Q Sepharose fractions.

[0021] Figures 2A-D are photos of SDS-PAGE gels with HTP purified human calreticulin fractions.

[0022] Figures 3A-D are photos of SDS-PAGE gels of Resource ISO purified human calreticulin fractions.

[0023] Figure 4 is a photo of an SDS-PAGE gel of purified human calreticulin (hCRT) samples ranging from 7.5μg to 30 μg of the protein.

[0024] Figure 5 is a vector map of the 5326 base pair pBAD-hCRT vector.

[0025] Figure 6 is the nucleotide sequence of the pBAD-hCRT including human calreticulin along with its amino acid sequence (SEQ ID NO:7) including the unique signal sequence from the pBAD construct. The boxed sequences correspond to the primers used for PCR or sequence analysis (SEQ ID NO:11 is GL96-seq.F; SEQ ID NO:12 is YQ15-seq.R; and SEQ ID NO:13 is Gl97-SEQ.R).

DESCRIPTION OF THE INVENTION

[0026] The present invention relates generally to purified proteins and methods for expressing a heterologous protein in a bacterial expression system without utilizing an affinity tag or label. In particular, the present invention provides a method for producing large quantities of calreticulin that is free of an affinity tag. The present invention also includes various aspects of biological materials and intermediates useful in the production of such non-tagged calreticulin by biological production.

[0027] The invention encompasses a method for producing calreticulin comprising: (a) culturing a microorganism having a calreticulin gene under suitable culture conditions; (b) inducing expression of the calreticulin gene to produce a calreticulin protein product that is free of an affinity label or tag; and (c) purifying the calreticulin product. In particular embodiments, the method utilizes a cDNA construct for calreticulin that terminates with a stop codon, such that the sequences encoding the expression tag(s) (i.e., c-myc and HIS-tag) from the expression vector are not translated and are not incorporated into the expressed protein.

[0028] Additionally, in a further embodiment, the present methods utilize a combination of at least two sequential ion chromatography purification steps followed by one hydrophobic chromatography purification step. In certain preferred embodiments, the hydrophobic chromatography step comprises utilizing 20-25% ammonium sulfate for elution. The ammonium sulfate elution step resulted in an excellent purification profile and was surprising in view of the failure of such a step to work in many other hydrophobic chromatography methods.

[0029] Even though high amounts of tagged calreticulin can be obtained using standard expression techniques, attempts to remove the tag generate impurities and may cause problems for its use in pharmaceutical/clinical applications.

[0030] The purification of calreticulin using the presently descril and three-step chromatography purification procedure provides a process for manufacturing significant quantities (suitable for large scale manufacture of a commercial pharmaceutical product) of purified human calreticulin lacking an affinity tag or label that is suitable for pharmaceutical use and analysis.

[0031] Calreticulin (CRT), also known as calregulin, is a well conserved chaperone protein that is found in most cells and has a variety of functions. It binds calcium with high capacity and interacts with newly synthesized proteins in the ER to facilitate their folding and to prevent aggregation. It is also expressed in cytotoxic granules where it is believed to regulate the activity of the cytolytic molecule, perforin. The cDNA sequence encoding the of human calreticulin precursor has been determined and has GenBank Accession No. NM_004343 (SEQ ID NO:3, 1899 bp). The amino acid sequence of the human calreticulin precursor protein has also been described at GenBank Accession No. NP_004334 (SEQ ID NO:4, 417 amino acids). Other names for calreticulin include: CRT, SSA, cC1qR, FLJ26680.

[0032] It is noted that the human calreticulin signal sequence MLLSVPLLLGLLGLAVA (SEQ ID NO:8) from the precursor protein is cleaved to produce the mature protein starting with the residues EPAVY... (SEQ ID NO:9). The unique signal sequence MKKLLFAIPPLVVPFYSHSTMELE (SEQ ID NO:10) preceding the mature calreticulin sequence is co-translationally cleaved to produce a secreted mature protein in the pBAD construct.

[0033] Calreticulin was first identified in skeletal muscle sarcoplasmic reticulum. Recently, it has been shown that calreticulin modulates the expression of hormonally regulated genes and has also been localized in other cellular organelles, supporting the idea that it functions outside the endoplasmic reticulum. Calreticulin is a multifunctional protein that acts as a major $Ca^{(2+)}$-binding (storage) protein in the lumen of the endoplasmic reticulum. It is also found in the nucleus, suggesting that it may have a role in transcription regulation. Calreticulin binds to the synthetic peptide KLGFFKR (SEQ ID NO:5), which is almost identical to an amino acid sequence in the DNA-binding domain of the superfamily of nuclear receptors. Calreticulin binds to antibodies in certain sera of systemic lupus and Sjogren patients which contain anti-Ro/SSA antibodies, it is highly conserved among species,

and it is located in the endoplasmic and sarcoplasmic reticulum where

The amino terminus of calreticulin interacts with the DNA-binding domain of the glucocorticoid receptor and prevents the receptor from binding to its specific glucocorticoid response element. Calreticulin can inhibit the binding of androgen receptor to its hormone-responsive DNA element and can inhibit androgen receptor and retinoic acid receptor transcriptional activities in vivo, as well as retinoic acid-induced neuronal differentiation. Thus, calreticulin can act as an important modulator of the regulation of gene transcription by nuclear hormone receptors. Increased autoantibody titer against human calreticulin is found in infants with complete congenital heart block of both the IgG and IgM classes.

[0034] Three putative structural domains were identified in calreticulin: a globular N-terminal domain, a praline-rich P domain, and an acidic C-terminal tail domain. High- and low-affinity binding of Ca^{2+} to calreticulin is restricted to the P and C domains, respectively. The globular N-terminal domain was shown to have affinity for the cytoplasmic domain of the alpha subunit of integrins and for a family of steroid receptors.

[0035] Suitable biological systems for producing calreticulin include prokaryotic systems.

[0036] The preferred prokaryote is *Escherichia coli* (*E. coli*). *E. coli* is well established as an industrial microorganism used in the production of metabolites (amino acids, vitamins) and several recombinant proteins. The entire *E. coli* genome has also been sequenced, and the genetic systems are highly developed. [Fujisaki, et. al, J. Biochem. 108, 995-1000 (1990); Lois et al., Proc. Natl. Acad. Sci. USA, 95, 2105-2110 (1998); Hemmi et al. , J. Biochem., 123, 1088-1096 (1998)].

[0037] Suitable organisms useful in producing calreticulin are available from numerous sources, such as the American Type Culture Collection (ATCC), Rockville, Md., Culture Collection of Algae (UTEX), Austin, Tex., the Northern Regional Research Laboratory (NRRL), Peoria, Ill. and the E. coli Genetic Stock Center (CGSC), New Haven, Conn. The present invention is not limited by the host cells employed.

[0038] Such expression organisms i.e., engineered host cells, can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants, or amplifying the calreticulin gene. Culture conditions, such as temperature,

pH and the like, are those previously used with the host cell selected fo
be apparent to those skilled in the art and in the references cited herein, including, e.g., Sambrook, Ausubel and Berger, as well as e.g., Freshney (1994) Culture of Animal Cells: A Manual of Basic Technique, 3.sup.rd ed. (Wiley-Liss, New York) and the references cited therein.

[0039] In certain aspects of the invention, the cells used in the cell culture are genetically modified to increase the yield of calreticulin. Cells may be genetically modified by genetic engineering techniques (i.e., recombinant technology), classical microbiological techniques, or a combination of such techniques and can also include naturally occurring genetic variants. Some of such techniques are generally disclosed, for example, in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Labs Press. The reference Sambrook et al., ibid., is incorporated by reference herein in its entirety. A genetically modified microorganism can include a microorganism in which nucleic acid molecules have been inserted, deleted or modified (i.e., mutated; e.g., by insertion, deletion, substitution, and/or inversion of nucleotides), in such a manner that such modifications provide the desired effect of increased yields of calreticulin within the microorganism or in the culture supernatant.

[0040] As used herein, genetic modifications which result in a decrease in gene expression, in the function of the gene, or in the function of the gene product (i.e., the protein encoded by the gene) can be referred to as inactivation (complete or partial), deletion, interruption, blockage or down-regulation of a gene. For example, a genetic modification in a gene which results in a decrease in the function of the protein encoded by such gene, can be the result of a complete deletion of the gene (i.e., the gene does not exist, and therefore the protein does not exist), a mutation in the gene which results in incomplete or no translation of the protein (e.g., the protein is not expressed), or a mutation in the gene which decreases or abolishes the natural function of the protein (e.g., a protein is expressed which has decreased or no enzymatic activity). Genetic modifications which result in an increase in gene expression or function can be referred to as amplification, overproduction, overexpression, activation, enhancement, addition, or up-regulation of a gene. Addition of cloned genes to increase gene expression can include maintaining the cloned gene(s) on replicating plasmids or integrating the cloned gene(s) into the genome of the production organism. Furthermore, increasing the expression of desired cloned genes can include operatively linkin the cloned

gene(s) to native or heterologous transcriptional control elements.

[0041] In certain embodiments, expression of calreticulin is induced in bacterial transformants with isopropylthiogalactoside (IPTG). In certain embodiments, the transformation mix is plated on agar plates made in L-Broth containing 100 μg/ml ampicillin.

[0042] Calreticulin has been shown to be useful in promoting the scarless healing of wounds in humans (See U.S. Patent No. 5,591,716) and is therefore a desirable pharmaceutical agent for this purpose. Calreticulin may also have other useful pharmaceutical applications. Thus, a procedure for producing large quantities of pure (tag-free) calreticulin (for use in a commercial/pharmaceutical product) would be of value. Calreticulin produced according to the methods described herein can be scaled up for production in fermenters by growing the expression strain containing for example, the pBAD/gIII A –hCRT vector under suitable conditions for inducing expression of calreticulin.

Molecular Biology

[0043] Conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art may be employed, in accordance with this invention. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual,* Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook *et al.*, 1989"); *DNA Cloning: A Practical Approach,* Volumes I and II (D.N. Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed. 1984); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds. (1985)); *Transcription And Translation* (B.D. Hames & S.J. Higgins, eds. (1984)); *Animal Cell Culture* (R.I. Freshney, ed. (1986)); *Immobilized Cells And Enzymes* (IRL Press, (1986)); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); F.M. Ausubel *et al.* (eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994).

[0044] The term "selection marker or selectable marker" is typically used to refer to a gene and/or protein whose presence can be detected directly or indirectly in a cell, for example, a gene and/or a protein that inactivates a selection agent and protects the host cell from the agent's lethal or growth-inhibitory effects (e.g., an antibiotic resistance gene and/or protein). Another possibility is that the selection marker induces fluorescence or a color

deposit (e.g., green fluorescent protein and derivatives, luciferase, or all

[0045] The term "selection agent" is typically defined as a chemical compound that is able to kill or retard the growth of host cells (e.g., an antibiotic).

[0046] The term "selection" is typically defined as the process of using a selection marker/selectable marker and a selection agent to identify host cells with specific genetic properties (e.g., that the host cell contains a transgene integrated into its genome).

[0047] "Amplification" of DNA as used herein denotes the use of polymerase chain reaction (PCR) or other methods to increase the concentration of a particular DNA sequence within a mixture of DNA sequences. For a description of PCR see Saiki *et al.*, Science 1988, 239:487.

[0048] A "nucleic acid molecule" refers to the phosphate ester polymeric form of ribonucleosides (adenosine, guanosine, uridine or cytidine; "RNA molecules") or deoxyribonucleosides (deoxyadenosine, deoxyguanosine, deoxythymidine, or deoxycytidine; "DNA molecules"), or any phosphoester analogs thereof, such as phosphorothioates and thioesters, in either single stranded form, or a double-stranded helix. Double stranded DNA-DNA, DNA-RNA and RNA-RNA helices are possible. The term nucleic acid molecule, and in particular DNA or RNA molecule, refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, *inter alia*, in linear (*e.g.*, restriction fragments) or circular DNA molecules, plasmids, and chromosomes. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (*i.e.*, the strand having a sequence homologous to the mRNA). A "recombinant DNA molecule" is a DNA molecule that has undergone a molecular biological manipulation.

[0049] A "polynucleotide" or "nucleotide sequence" is a series of nucleotide bases (also called "nucleotides" in a nucleic acid, such as DNA and RNA, and means any chain of two or more nucleotides. A nucleotide sequence typically carries genetic information, including the information used by cellular machinery to make proteins and enzymes. These terms include double or single stranded genomic and cDNA, RNA, any synthetic and genetically

manipulated polynucleotide, and both sense and anti-sense polynucl
sense stands are being represented herein). This includes single- and double-stranded molecules, *i.e.*, DNA-DNA, DNA-RNA and RNA-RNA hybrids, as well as "protein nucleic acids" (PNA) formed by conjugating bases to an amino acid backbone. This also includes nucleic acids containing modified bases, for example thio-uracil, thio-guanine and fluoro-uracil.

[0050] The nucleic acids herein may be flanked by natural regulatory (expression control) sequences, or may be associated with heterologous sequences, including promoters, internal ribosome entry sites (IRES) and other ribosome binding site sequences, enhancers, response elements, suppressors, signal sequences, polyadenylation sequences, introns, 5'- and 3'- non-coding regions, and the like. The nucleic acids may also be modified by many means known in the art. Non-limiting examples of such modifications include methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, and internucleotide modifications such as, for example, those with uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoroamidates, carbamates, etc.) and with charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, etc.). Polynucleotides may contain one or more additional covalently linked moieties, such as, for example, proteins (*e.g.*, nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), intercalators (*e.g.*, acridine, psoralen, etc.), chelators (*e.g.*, metals, radioactive metals, iron, oxidative metals, etc.), and alkylators. The polynucleotides may be derivatized by formation of a methyl or ethyl phosphotriester or an alkyl phosphoramidate linkage. Furthermore, the polynucleotides herein may also be modified with a label capable of providing a detectable signal, either directly or indirectly. Exemplary labels include radioisotopes, fluorescent molecules, biotin, and the like.

[0051] A “promoter” or "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the

binding of RNA polymerase. The promoter may be operatively
expression control sequences, including enhancer and repressor sequences.

[0052] A "coding sequence" or a sequence "encoding" an expression product, such as a RNA, polypeptide, protein, or enzyme, is a nucleotide sequence that, when expressed, results in the production of that RNA, polypeptide, protein, or enzyme, *i.e.,* the nucleotide sequence encodes an amino acid sequence for that polypeptide, protein or enzyme. A coding sequence for a protein may include a start codon (usually ATG) and a stop codon.

[0053] The term "gene", also called a "structural gene", means a DNA sequence that codes for or corresponds to a particular sequence of amino acids which comprise all or part of one or more proteins or enzymes, and may or may not include regulatory DNA sequences, such as promoter sequences, which determine for example the conditions under which the gene is expressed. Some genes, which are not structural genes, may be transcribed from DNA to RNA, but are not translated into an amino acid sequence. Other genes may function as regulators of structural genes or as regulators of DNA transcription.

[0054] A coding sequence is "under the control of" or "operatively associated with" expression control sequences in a cell when RNA polymerase transcribes the coding sequence into RNA, particularly mRNA, which is then trans-RNA spliced (if it contains introns) and translated into the protein encoded by the coding sequence.

[0055] The term "expression control sequence" refers to a promoter and any enhancer or suppression elements that combine to regulate the transcription of a coding sequence. In a preferred embodiment, the element is an origin of replication.

[0056] The terms "vector", "cloning vector", and "expression vector" mean the vehicle by which a DNA or RNA sequence (*e.g.* a foreign gene) can be introduced into a host cell, so as to transform the host and promote expression (*e.g.* transcription and translation) of the introduced sequence.

[0057] Vectors typically comprise the DNA of a transmissible agent, into which foreign DNA is inserted. A common way to insert one segment of DNA into another segment of DNA involves the use of specific enzymes called restriction enzymes that cleave DNA at specific sites (specific groups of nucleotides) called restriction sites. A "cassette" refers to a DNA coding sequence or segment of DNA that codes for an expression produc

inserted into a vector at defined restriction sites. The cassette restrictio
ensure insertion of the cassette in the proper reading frame. Generally, foreign DNA is inserted at one or more restriction sites of the vector DNA, and then is carried by the vector into a host cell along with the transmissible vector DNA. A segment or sequence of DNA having inserted or added DNA, such as an expression vector, can also be called a "DNA construct". A common type of vector is a "plasmid" which generally is a self-contained molecule of double-stranded DNA, usually of bacterial origin, that can readily accept additional (foreign) DNA and which can be readily introduced into a suitable host cell. A plasmid vector often contains coding DNA and promoter DNA and has one or more restriction sites suitable for inserting foreign DNA. Coding DNA is a DNA sequence that encodes a particular amino acid sequence for a particular protein or enzyme. Promoter DNA is a DNA sequence that initiates, regulates, or otherwise mediates or controls the expression of the coding DNA. Promoter DNA and coding DNA may be from the same gene or from different genes, and may be from the same or different organisms. Recombinant cloning vectors will often include one or more replication systems for cloning or expression, one or more markers for selection in the host, *e.g.* antibiotic resistance, and one or more expression cassettes. Vector constructs may be produced using conventional molecular biology and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, Fritsch & Maniatis, Molecular Cloning: A Laboratory Manual, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook et al., 1989"); DNA Cloning: A Practical Approach, Volumes I and II (D.N. Glover ed. 1985); F.M. Ausubel et al. (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (1994).

[0058] The terms "express" and "expression" mean allowing or causing the information in a gene or DNA sequence to become manifest, for example producing a protein by activating the cellular functions involved in transcription and translation of a corresponding gene or DNA sequence. A DNA sequence is expressed in or by a cell to form an "expression product" such as a protein. The expression product itself, *e.g.* the resulting protein, may also be said to be "expressed" by the cell. An expression product can be characterized as intracellular, extracellular or secreted. The term "intracellular" means something that is inside a cell. The term "extracellular" means something that is outside a cell. A substance is "secreted" by a cell if it appears in significant measure outside the cell, from somewhere on

or inside the cell.

[0059] The term "transfection" means the introduction of a foreign (*i.e.,* extrinsic or extracellular) gene or foreign nucleic acid into a cell. The term “transformation” means the introduction of a foreign gene, foreign nucleic acid, DNA or RNA sequence to a host cell, so that the host cell will express the introduced gene or sequence to produce a desired substance, typically a protein or enzyme coded by the introduced gene or sequence. The introduced gene or sequence may also be called a “cloned” or “foreign” gene or sequence, may include regulatory or control sequences, such as start, stop, promoter, signal, secretion, or other sequences used by a cell’s genetic machinery. The gene or sequence may include nonfunctional sequences or sequences with no known function. A host cell that receives and expresses introduced DNA or RNA has been “transformed” and is a "transformant" or a "clone." The DNA or RNA introduced to a host cell can come from any source, including cells of the same genus or species as the host cell, or cells of a different genus or species.

[0060] The term "host cell" means any cell of any organism that is selected, modified, transformed, grown, or used or manipulated in any way, for the production of a substance by the cell, for example the expression by the cell of a gene, a DNA or RNA sequence, a protein or an enzyme.

[0061] The term "expression system" means a host cell and compatible vector under suitable conditions, *e.g.* for the expression of a protein coded for by foreign DNA carried by the vector and introduced to the host cell. In a specific embodiment, the host cell of the present invention is a Gram-negative or Gram-positive bacteria. These bacteria include, but are not limited to, *Escherichia coli* (*E. coli).*

[0062] The term "heterologous" refers to a combination of elements not naturally occurring. For example, heterologous DNA refers to DNA not naturally located in the cell, or in a chromosomal site of the cell. Preferably, the heterologous DNA includes a gene foreign to the cell. A heterologous expression regulatory element is such an element operatively associated with a different gene than the one it is operatively associated with in nature. In the context of the present invention, a gene encoding a protein of interest is heterologous to the vector DNA in which it is inserted for cloning or expression, and it is heterologous to a host cell containing such a vector, in which it is expressed.

[0063] The terms "mutant" and "mutation" mean any detectal material, *e.g.* DNA, or any process, mechanism, or result of such a change. This includes gene mutations, in which the structure (*e.g.*, DNA sequence) of a gene is altered, any gene or DNA arising from any mutation process, and any expression product (*e.g.*, protein or enzyme) expressed by a modified gene or DNA sequence. The term “variant” may also be used to indicate a modified or altered gene, DNA sequence, enzyme, cell, etc., *i.e.*, any kind of mutant.

[0064] "Sequence-conservative variants" of a polynucleotide sequence are those in which a change of one or more nucleotides in a given codon position results in no alteration in the amino acid encoded at that position.

[0065] "Function-conservative variants" are those in which a given amino acid residue in a protein or enzyme has been changed without altering the overall conformation and function of the polypeptide, including, but not limited to, replacement of an amino acid with one having similar properties (such as, for example, polarity, hydrogen bonding potential, acidic, basic, hydrophobic, aromatic, and the like). Amino acids with similar properties are well known in the art. For example, arginine, histidine and lysine are hydrophilic-basic amino acids and may be interchangeable. Similarly, isoleucine, a hydrophobic amino acid, may be replaced with leucine, methionine or valine. Such changes are expected to have little or no effect on the apparent molecular weight or isoelectric point of the protein or polypeptide. Amino acids other than those indicated as conserved may differ in a protein or enzyme so that the percent protein or amino acid sequence similarity between any two proteins of similar function may vary and may be, for example, from 70% to 99% as determined according to an alignment scheme such as by the Cluster Method, wherein similarity is based on the MEGALIGN algorithm. A "function-conservative variant" also includes a polypeptide or enzyme which has at least 60 % amino acid identity as determined by BLAST or FASTA algorithms, preferably at least 75%, more preferably at least 85%, most preferably at least 90%, and which has the same or substantially similar properties or functions as the native or parent protein or enzyme to which it is compared.

[0066] As used herein, the terms "homologous" and “homology” refer to the relationship between proteins that possess a "common evolutionary origin," including proteins from superfamilies (*e.g.*, the immunoglobulin superfamily) and homologous proteins

from different species (*e.g.*, myosin light chain, etc.) (Reeck *et al.*, Cell
proteins (and their encoding genes) have sequence homology, as reflected by their sequence similarity, whether in terms of percent similarity or the presence of specific residues or motifs at conserved positions.

[0067] Accordingly, the term "sequence similarity" refers to the degree of identity or correspondence between nucleic acid or amino acid sequences of proteins that may or may not share a common evolutionary origin (*see* Reeck *et al.*, *supra*). However, in common usage and in the instant application, the term "homologous," when modified with an adverb such as "highly," may refer to sequence similarity and may or may not relate to a common evolutionary origin.

[0068] In a specific embodiment, two DNA sequences are "substantially homologous" or "substantially similar" when at least about 80%, and most preferably at least about 90 or 95% of the nucleotides match over the defined length of the DNA sequences, as determined by sequence comparison algorithms, such as BLAST, FASTA, DNA Strider, etc. An example of such a sequence is an allelic or species variant of the specific genes of the invention. Sequences that are substantially homologous can be identified by comparing the sequences using standard software available in sequence data banks, or in a Southern hybridization experiment under, for example, stringent conditions as defined for that particular system.

[0069] Similarly, in a particular embodiment, two amino acid sequences are "substantially homologous" or "substantially similar" when greater than 80% of the amino acids are identical, or greater than about 90% are similar. Preferably, the amino acids are functionally identical. Preferably, the similar or homologous sequences are identified by alignment using, for example, the GCG (Genetics Computer Group, Program Manual for the GCG Package, *Version 7*, Madison, Wisconsin) pileup program, or any of the programs described above (BLAST, FASTA, etc.).

[0070] A nucleic acid molecule is "hybridizable" to another nucleic acid molecule, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid molecule can anneal to the other nucleic acid molecule under the appropriate conditions of temperature and solution ionic strength (*see* Sambrook *et al.*, *supra*). The conditions of temperature and ionic strength determine the "stringency" of the hybridiz

preliminary screening for homologous nucleic acids, low stringency hy
corresponding to a T_m (melting temperature) of 55°C, can be used, *e.g.*, 5x SSC, 0.1% SDS, 5X Denhardt's, and no formamide; or 30% formamide, 5x SSC, 0.5% SDS, 5X Denhardt's). Moderate stringency hybridization conditions correspond to a higher T_m, *e.g.*, 40% formamide, with 5x or 6x SCC, 5x Denhardt's. High stringency hybridization conditions correspond to the highest T_m, *e.g.*, 50% formamide, 5x or 6x SCC, 5x Denhardt's. SCC is a 0.15M NaCl, 0.015M Na-citrate buffer. 5x Denhardt's is 0.1% ficoll, 0.1%g polyvinylpyrrolidone (MW40,000), 0.1%g BSA (w/v). Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of T_m for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher T_m) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating T_m have been derived (*see* Sambrook *et al.*, *supra*, 9.50-9.51). For hybridization with shorter nucleic acids, *i.e.*, oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (*see* Sambrook *et al.*, *supra*, 11.7-11.8). A minimum length for a hybridizable nucleic acid is at least about 10 nucleotides; preferably at least about 15 nucleotides; and more preferably the length is at least about 20 nucleotides.

[0071] In a specific embodiment, the term "standard hybridization conditions" refers to a T_m of 55°C, and utilizes conditions as set forth above. In a preferred embodiment, the T_m is 60°C; in a more preferred embodiment, the T_m is 65°C. In a specific embodiment, "high stringency" refers to hybridization and/or washing conditions at 68°C in 0.2x SSC, at 42°C in 50% formamide, 4x SSC, or under conditions that afford levels of hybridization equivalent to those observed under either of these two conditions.

[0072] As used herein, the term "oligonucleotide" refers to a nucleic acid, generally of at least 10, preferably at least 15, and more preferably at least 20 nucleotides, preferably no more than 100 nucleotides, that is hybridizable to a genomic DNA molecule, a cDNA molecule, or an mRNA molecule encoding a gene, mRNA, cDNA, or other nu

interest. Oligonucleotides can be labeled, *e.g.*, with ^{32}P-nucleotides or label, such as biotin, has been covalently conjugated. In one embodiment, a labeled oligonucleotide can be used as a probe to detect the presence of a nucleic acid. In another embodiment, oligonucleotides (one or both of which may be labeled) can be used as PCR primers, either for cloning full length or a fragment of the gene, or to detect the presence of nucleic acids encoding the protein. In a further embodiment, an oligonucleotide of the invention can form a triple helix with a DNA molecule. Generally, oligonucleotides are prepared synthetically, preferably on a nucleic acid synthesizer. Accordingly, oligonucleotides can be prepared with non-naturally occurring phosphoester analog bonds, such as thioester bonds, etc.

[0073] A "prokaryotic cell" or "prokaryote" refers to a cell without discrete nucleus and with single, circular DNA molecules within the cytoplasm. Prokaryotes include but are not limited to cells of bacteria and blue green algae.

[0074] Abbreviations are as follows: HPLC refers to high-performance liquid chromatography; GC and %GC refer to the number and percentage of guanine and cytosine bases; CFU refers to colony forming units.

EXAMPLES

[0075] The present invention will be better understood by reference to the following Examples, which are provided as exemplary of the invention, and not by way of limitation.

General Materials and Methods

[0076] DNA restriction and modification enzymes and T4 DNA ligase were obtained from New England Biolabs. Plasmid DNA was isolated using commercial kits and DNA fragments were purified using commercial kits (Qiagen, Valencia, CA). Competent *E. coli* cells (i.e., DH5α, BL21 cells) were obtained from Novagen (Merck/EMD, San Diego, California). TOP10F' Rosetta cells for expression studies for human calreticulin were obtained from Invitrogen (Carlsbad, California). Standard cloning and molecular biology kits and related techniques were used according to manufacturer's specifications and with buffers and reagents supplied by the manufacturer. Antibiotics were purchased from Sigma.

EXAMPLE 1

Cloning of the Human Calreticulin cDNA

[0077] PCR reactions were carried out under standard, manufacturers conditions unless otherwise noted. Likewise, except where noted, standard molecular techniques such as restriction endonuclease digestions and ligations may be carried out under any suitable standard conditions.

[0078] The PCR-driven amplification of human calreticulin cDNA (ATCC clone # 3767728, GenBank accession # AW239363) was carried out using the following sets of primers:

Forward: 5'-ATA CTC GAG GAG CCT GCC GTC TAC-3' (SEQ ID NO:1)

Reverse: 5'-AGG GAA TTC CTA CAG CTC GTC C-3' (SEQ ID NO:2)

[0079] The underlined EcoR1 and restriction sites were used for cloning. In certain experiments, the primers (SEQ ID NO:11) GL96-seq.F: ATGCCATAGCATTTTTATCCA and (SEQ ID NO:13) Gl97-SEQ.R: ACAGATTAAATCAGAACGCA were used.

[0080] Human calreticulin was amplified according to the following procedure:

1. Human calreticulin cDNA was amplified from plasmid pPIC9-human CRT (ATCC Accession No.: AW239363) using the following PCR reaction conditions:

10 x buffer	5 µl
10 mM dNTP:	1.5 µl
50 mM $MgSO_4$:	1 µl
10 µM primer1 (SEQ ID NO:1 or 11):	1.5 µl
10 µM primer2 (SEQ ID NO:2 or 13):	1.5 µl
Template DNA:	2 µl (100 ng)
Pfx (Invitrogen):	0.5 µl
H_2O:	to final volume 50 µl

[0081]

PCR cycles:	94°C	3 min	
	94°C	30 s	30 cycles
	58°C	30 s	
	68°C	1 min 30s	
	72°C	3 min, then keep in 4°C	

2. Run 1% agarose gel to confirm PCR product (expected size of 1899 nucleotides, about a

1900 base pair band on gel).

[0082] The calreticulin PCR product was confirmed by sequencing and then digested and ligated to a suitable vector according to the following procedures:

3. The hCRT PCR product was digested with *Xho* I and *EcoR* I under the following conditions:

10 x EcoR I buffer:	5 µl
Xho I:	1 µl
EcoR I:	1 µl
PCR product:	45 µl

Incubated at 37°C overnight.

[0083] The pBAD gIII A vector was prepared by digesting vector DNA with *Xho* I and *EcoR* I under the following conditions:

10 x EcoR I buffer:	5 µl
Xho I:	1 µl
EcoR I:	1 µl
pBAD vector DNA:	500 ng
H_2O:	to final volume 50 µl

Incubated at 37°C overnight.

4. The digested PCR product and vector were purified using a gel purification kit according to the manufacturer's instructions (Qiagen).
5. A 1% agarose gel was run to examine the concentration of PCR product and vector.
6. The purified PCR product and vector were ligated under the following conditions:

10 x ligation buffer:	2 µl
Purified PCR product:	100 ng
Purified pBAD vector:	50 ng
T4 ligase:	1 µl
H_2O:	to final volume 50 µl

[0084] The ligation reaction was incubated at room temperature for 4 hours.

7. Competent *E. coli* cells were transformed as follows:

Add 5 µl ligation product into 50 µl DH5α competent *E. coli* cells;

Incubate mixture on ice for 30 min;

Heat shock at 42°C for 90 seconds;

Incubate on ice for 2 min;

Add product to glass tube with 250 µl LB and grow for 1 hour in 37°C

shaker;

Plate total content to LB-Amp plates;

Grow plates overnight at 37°C.

8. Picked colonies and screened for positive clones using antibiotic resistance, restriction digestion, and nucleotide sequence analysis. The nucleotide sequence analysis, based on reference SEQ ID NO:3 was also carried out on the positive clones to verify the calreticulin insert (various primers were used for sequencing including: SEQ ID NO:11 (GL96-seq.F): ATGCCATAGCATTTTTATCCA; SEQ ID NO:12; YQ15-seq.R (1101): GAGATGGACGGAGAGTGGGA; and SEQ ID NO:13 (Gl97-SEQ.R): ACAGATTAAATCAGAACGCA. The DH5α construct was used as the stock strain. The pBAD/gIII-A hCRT vector was transformed into BL21 expression cells from Novagen (Merck/EMD, San Diego, California) for expression of human calreticulin, as described below.

EXAMPLE 2

Expressing and Purifying Human Calreticulin in *E. coli*

[0085] pBAD expression vectors encoding human calreticulin (pBADhCRT) were made as described in Example 1. While both amp and kan resistant vectors are available, the Kan resistant clone has slightly increased level of expression and is preferred by the FDA for pharmaceutical purposes. While numerous suitable expression strains can be used, the *E. coli* strain used for expression in the present examples, was Rosetta (obtained from Invitrogen (Carlsbad, California) TOP10F`. Prior to protein purification, the pH of each buffer was confirmed. For protein purification, all work was performed in the cold room unless specified. Protease inhibitor cocktail + PMSF (phenylmethanesulphonylfluoride) + Benzamidine (for details see buffers composition section below) were added to all buffers prior to use. The protease inhibitors are optional and may be omitted for the last chromatography step.

Buffers And Solutions For Protein Expression And Purification.

[0086] **Lysate Buffer (L-Buffer)**
150 mM NaCl
100 mM KH_2PO_4, pH-7.1
2 mM $MgCl_2$

Q SEPHAROSE CHROMATOGRAPHY COLUMN BUFFERS

[0087] **Column details:** HiLoad 26/10 Fast Flow column was used from GE Healthcare Bio-Sciences cat # 17-1062-01

Q Sepharose Binding buffer (BBQ)
150 mM NaCl
100 mM KH_2PO_4, pH-7.1
2 mM $MgCl_2$

Q Sepharose Elution buffer (EBQ)
1000 mM NaCl
100 mM KH_2PO_4, pH-7.1
2 mM $MgCl_2$

HYDROXYLAPATITE CHROMATOGRAPHY COLUMN BUFFERS

Column details: Hydroxylapatite Bio-gel HTP Gel, Bio-Rad cat# 130-0420

Binding buffer (BBH)
10 mM K_2HPO_4, pH-7.0

Wash buffer (WBH) (use 2 column volumes for washing)
80 mM K_2HPO_4, pH-7.0

Elution buffer (EBH) (for continuous salt gradient from 80 – 1000 mM)
80 mM K_2HPO_4, pH-7.0
1000 mM K_2HPO_4, pH-7.0

Resource Iso (Isopropyl) Chromatography Column Buffers

[0088] **Column details:** RESOURCE ISO 1 ml column from GE Healthcare Bio-Sciences
cat # 17-1185-01

Binding buffer (BBI) Buffers:

40% Ammonium Sulfate in 10 mM K_2HPO_4, pH- 7.2

Elution buffer (EBI)

10 mM K_2HPO_4, pH-7.2

Protein storage Buffer

10 mM Tris-HCl, pH-7.0
1 mM EDTA

[0089] The protein may also be stored in calcium containing buffers as long as calcium concentration does not exceed 10 mM.

Protease Inhibitor Solutions

Protease inhibitor cocktail (2000x stock):

[0090] The following recipe is for a 100 ml stock solution in 100% EtOH. Store at -20°C as 2 ml aliquots.

Aprotinin, Sigma Cat # A-1153	10 mg
Phosphoramidone, Sigma Cat # R-9382	5 mg
TLCK, Sigma Cat # T-7254	10 mg
TPCK, Sigma Cat # T-4376	20 mg
APMSF, Sigma Cat # A-6664	10 mg
E-64, Sigma Cat # E-3132	10 mg
Leupeptin, Sigma Cat # L-2884	5 mg
I Pepstatin, Sigma Cat # P-4265	2 mg

50 µl of the protease inhibitor cocktail/100 ml of buffer was used in the experiments described herein.

PMSF solution

[0091] Stock solution: 100 mM PMSF in 95% ETOH (Store chilled, i.e., in the refrigerator).

500 µl of PMSF stock solution/100 ml of buffer is used in the experiments described herein.

Benzamidine solution

[0092] Stock solution: 200 mM Benzamidine in 95% ETOH. (Store chilled, i.e., in the

refrigerator). 250 µl of benzamidine stock solution/100 ml of buffer was
experiments described herein.

E. coli Culture And Induction Of Protein Expression

[0093] **Chemicals/Solutions**
Kanamycin (Kan) final concentration 30 µg/ml. Add 5 µl/ml (filter sterilized stock 60 mg/ml in ddH_2O_2, store at -20°C).

Chloramphenicol (CM): final concentration 130 µg/ml. Add 30 µl/ml (filter sterilized stock 60 mg/ml, store at -20°C in 95%ETOH).

Arabinose: final concentration 0.02% (filter sterilized stock 20% in ddH_2O_2, store at -20°C). Use L-(+) Arabinose from Sigma, cat# A-3256 for the induction.

Calreticulin Expression in *E. coli*

[0094] **Day 1.** Streaked the LB/Kan/CM plate with *E. coli* harboring a pBAD-hCRT expression vector, incubated overnight (O/N – 18 -20 hrs) at 37°C.

Day 2. Inoculated 25 ml of LB/Kan/CM with single picked colony and incubated cells in a bacterial shaker overnight (O/N – 18 -20 hrs) at 37°C and at 200 rpm.

Day 3. Added 10 ml of *E.coli* O/N culture (point 2) at 1:50 dilution to 500 ml of LB/Kan/CM media in 2L flask and incubated in a bacterial shaker at 37°C, at 200 rpm for approximately 3-4 hrs until the turbidity of the culture reaches OD_{600}.of 0.5-0.6.

Induction: Added 0.5 ml of 20% (stock) L-Arabinose and continue dto incubate in a bacterial shaker at 30°C, 200 rpm for additional 4 hrs.

[0095] ***Centrifugation:*** Beckman centrifuge model #J2-21M using 250 ml centrifuge bottles; Rotor #JA-14, at 6,000 rpm (5,523 xg) for 10 min at 4°C.

Re-suspended a pellet in 30 ml ice cold L-Buffer containing protease inhibitors.

French pressed the resuspended protein pellets twice at 1,000 psi, (French pressure Cell Press, American Instrument Company P1603013600), keep sample on ice. Retained 20 µl aliquot for SDS-PAGE analysis.

Centrifugation: Beckman centrifuge model #J2-21, 2 x 40 ml tubes, rot
(13,776 xg) for 20 min at 4°C.

[0096] Saved the supernatant (30 ml) containing calreticulin from the French press and kept at 4°C.

[0097] Added protease inhibitors to the calreticulin samples. Filtered the supernatant through a0.22 μm filter using a syringe unit filter and proceeded to Q Sepharose chromatography step. Retained 20 μl aliquot for SDS-PAGE analysis.

[0098] Gently rinsed the pellet with a sterile ddH_2O to remove any trace of supernatant and re-suspend in 5 ml of L-Buffer. Retain 20 μl aliquot for SDS-PAGE analysis.

[0099] Standard Laemmli gel electrophoresis was used to test for expression of calreticulin (e.g., SDS-PAGE, 10 % acrylamide). Ten microliters of 3 x Sample Buffer was added to 20 μl sample mix and loaded on SDS-PAGE in the following amounts:

A-sample after French press 15 μl

B-sample of the supernatant 15 μl

C-sample of the pellet 5 μl

Purification Using Q Sepharose Column Chromatography

[00100] Before starting chromatography, the column was washed with 2-3 column volumes of binding buffer (BBQ). The following FPLC settings were used for Q Sepharose separation: Unicorn 5.11.

[00101] SDS-PAGE (10% acrylamide) was performed on every other fraction from the Q Sepharose column to identify the exact elution profile for calreticulin. Fractions containing human calreticulin were combined. Figures 1A-D are photos of SDS-PAGE gels showing the protein fractions from Q Sepharose chromatography. The Molecular Weight Standards (MWS or MW) used are Low range Bio-Rad cat# 161-0304 as follows:

→97,400 daltons Phosphorylase b

→66,200 daltons Serum albumin

→45,000 daltons Ovalbumin

→31,000 daltons Carbonic anhydrose

→21,500 daltons Trypsin inhibitor

→14,400 daltons Lysozyme

[00102] The samples in the gels shown in Figs. 1A-D, a representative Q Sepharose chromatography purification using the an expression vector according to the present invention (pBAD-hCRT), are as follows:

Gel A (lanes from left to right)

L,B,FT,W1,W2,W3.W4, B12,B10,B10,B8,B6,B4,B2,C1,C3

Gel B (lanes from left to right)

C5,L,C7,C9,C11,D12,D10,D8,D6,D4,D2,E1,E3,E5,E7

Gel C (lanes from left to right)

E9,E11,L,F12,F10,F8,F6,F4,F2,G1,G3,G5,G7,G9,G11

Gel D (lanes from left to right)

H12,H8,H4,L,H2,I1,I3,I5,I6,I8,I8,I10,J12,J10,J8,J6

[00103] The sample abbreviations stand for:
L=Low molecular weight standards
B=sample before the column
FT=flow through
W=washes (1, 2, 3, 4)
B12-J6 =fractions from Q Sepharose column

[00104] Each gel was loaded with 20 µl of each fraction/lane. Fractions C6 to E12 were pooled for further separation.

[00105] Pooled fractions C6 to E12 were dialyzed against 2 changes of 10 mM K_2HPO_4 pH-7.0 (for dialysis sample:buffer (v:v) ratio should be maintained at 1:1,000). A minimum of 6 hours was utilized for each buffer change. Dialysis can continue for longer, but does not have to exceed 8 hrs for each buffer change. Dialysis tubing used: Fisher Scientific Cat# 21-152-14.

[00106] Samples were removed from dialysis tubing and filtered through a 0.22 µm syringe unit filter. The filtered samples were then subjected to hydroxylapatide chromatography. At this point, the sample should be clear, if not, centrifuge: Beckman centrifuge model #J2-21M; 40 ml centrifuge bottles; Rotor #JA-17, 10,000 rpm (13,776 xg)

for 10 min at 4°C. Discard pellet, collect supernatant and proce chromatography.

Hydroxylapatite (Htp) Column Chromatography

Preparation Of Htp Beads And Column Chromatography

[00107] Using fresh beads each time for Htp chromatography is important for optimum results. Three columns of the same size were needed to process the Q Sepharose sample isolated as described above. The columns were prepared as follows:

1. 7 g of dry powder Bio-Gel HTP hydroxylapatite was weighed: column capacity 10 mg of protein/g. When hydrated, Bio-Gel HTP hydroxylapatite occupies approximately 2-3 ml per gram. 7 g of HPT in BBH buffer provided approximately 19-20 ml.

2. De-fining the beads. Added 1 part of HTP powder and 6 parts of binding buffer (BBH). Swirled gently, settled for 10 min and decanted the supernatant. Repeated this step 2 times.

3. Prepared slurry in binding buffer in a ratio of 75% settled gel to 25% buffer. Loaded the mixture onto a glass column: diameter: 20 mm; length: 13 cm.

4. Loaded 1/3 of Q Sepharose purified calreticulin onto each HTP column (divided equally among the 3 columns).

[00108] Before starting the program, wash the column with 2-3 column volumes of Binding buffer (BBH). The following FPLC settings were used for HTP column chromatography: Unicorn 5.11.

[00109] Figures 2A-D show SDS-PAGE gels of HTP purified human calreticulin.

The Molecular Weight Standards were used - Low range Bio-Rad cat# 161-0304:

→97,400 daltons Phosphorylase b
→66,200 daltons Serum albumin

→45,000 daltons Ovalbumin
→31,000 daltons Carbonic anhydrose
→21,500 daltons Trypsin inhibitor
→14,400 daltons Lysozyme

[00110] The Figures 2A-D samples from an exemplary HTP chromatography (following the Q-Sepharose purification step) purification correspond from left to right to the following:

Gel #A
L,B,FT,W,A9,A10,A11,A12,B12,B11,B10,B9,B8,B7,B6,

Gel#B
B5,L,B4,B2,C1,C3,C5,C7,C10,D11,D9,D6,D4,D1,E5,
Fractions from B11 to C10 were pooled.

Gel #C
A11,A12,L,Before(B),B12,B11,B10,B9,B8,B7,B6,B5,b4,B3,B2,

Gel#D
B1,LC2,C4,L,C6,C8,C10,D12,D8,D6,D4,D2,E1,E7,
Fractions from A12 to C11 were pooled.

[00111] The sample abbreviations stand for:

L=molecular weight standards
B=sample before the column (Gel #A and #C)
FT=flow through
W=wash

Each gel was loaded with 20 µl of each fraction/lane.

FPLC - RESOURCE ISO (ISOPROPYL) CHROMATOGRAPHY

[00112] The hydroxylapatite fractions containing calreticulin were combined and made saturated with solid ammonium sulfate to 40% (v/w). Two hundred forty-three grams of solid ammonium sulfate was then added to 1 L of solution. The solution was mixed gently and filter sterilized (Millipore 0.22 µm). The filtered sample was loaded into the Resource ISO column (hydrophobic chromatography step) with a flow rate of 1 ml/min. 1 ml resource ISO column, has a loading capacity of ~25 mg. Before starting the ISO chromatography step, the column was washed with 2-3 column volumes of Binding buffer (BBI). The ISO chromatography was performed as follows:

1. Column was stored in 20% ETOH.
2. Washed column first with ddH_2O, then equilibrate with a binding buffer
3. Washed with elution buffer.
4. Washed again with binding buffer.
5. Eluted the protein with a linear gradient of buffer BBI and buffer EBI, used 50 ml of each. Collected 0.5 ml fractions.
6. Analyzed every second fraction on SDS-PAGE (20 µl sample) (10% acrylamide).

Calreticulin eluted at approximately 20-25% ammonium sulfate.

The following FPLC setting was used for Resource ISO column chromatography: Unicorn 5.11.

[00113] Figures 3A-D show SDS-PAGE gels of Resource ISO purified calreticulin (i.e., calreticulin purified following the Q-Sepharose and HTP chromatography). The Molecular Weight Standards used are the Low range Bio-Rad cat# 161-0304:

→97,400 daltons Phosphorylase b
→66,200 daltons Serum albumin
→45,000 daltons Ovalbumin
→31,000 daltons Carbonic anhydrose
→21,500 daltons Trypsin inhibitor
→14,400 daltons Lysozyme

[00114] The samples from an exemplary Resource ISO chromatography correspond to the following lanes from left to right in Figs. 3A-D:

Gel A run#1

L,B,A11,A12,B12,B11,B10,B9,B8,B7,B6,B5,B4,B3,B2,B1

Fractions from A12 to B6 were pooled.

Gel B run#2

C12,L,D12 ,D11,D10,D9,D8,D7,D6,D5,D4,D3,D2,D1,E1

Fractions from C12 to D5 were pooled.

Gel C run#3

E12,F12,LF1,F10,F9,F8,F7F6,F5,F4,F3,F2,F1,G1

Fractions from F12 to F4 were pooled.

GelD run#4

H11,H10H9,L,H8,H7,H6H5,H4,H3,H2,H1,I1,I2,I3

Fractions from H11 to H4 were pooled.

The abbreviations are as follow:

L=molecular weight standards

B=sample before the column

FT=flow through

W=wash

The gels were loaded with 20 µl of each fraction/lane.

Filter Sterilization And Measurement Of Protein Concentration

[00115] Following purification using any of the chromatography methods described herein, the protein samples can be filter sterilized using any suitable method.

1. In a preferred embodiment, the protein samples were filter sterilized using Millex Sterile Syringe Driven Filter Unit-0.22 µm from Millipore, cat# SLGV 033RS. This unit uses a 10 ml syringe, through which the calreticulin sample is pushed.

2. Next, the filter unit was rinsed by filtering protein storage buffer to recover any protein that may remain on the filter.

3. Any suitable means for concentrating the protein may be used. In a preferred embodiment, an Amicon Ultra15 Centrifuge Filter Unit from Millipore - cat #UFC9 03024, regenerated cellulose 30MWCO unit was utilized. The amicon was centrifuged at a maximum speed of 4,800 rpm (3,838 xg) using, for example, a Beckman centrifuge GS-15R, swing bucket rotor # S4180.

4. Preferably, the protein is concentrated to a small volume (typically less than 1 ml). Next, the same filter unit was rinsed with exchange elution buffer (filter sterilized) using the same filter unit (repeating steps 1 and 2).

5. Next, the filter unit was filled with Protein storage Buffer (filter sterilized) and centrifuged at a maximum speed of about 4,800 rpm (3,838 xg), using for example, a Beckman centrifuge GS-15R, swing bucket rotor # S4180. This step was repeated 3 times.

6. Next the concentrated protein sample was removed from the

filter unit; collected in the tube (4°C), and combined with other sam
was retained for measuring protein concentration.

7. The sample was frozen in liquid nitrogen as 200 µg aliquots.

8. The frozen samples were stored at -80°C.

Calculation Of Protein Concentration

[00116] Using Absorbance @ A280 nm

$$\frac{A280}{a280 \times b}$$

Molecular mass of hCRT is 49142.6g/L
b=1cm
a280= 82975 for h CRT

Example:

5 µl of sample in 995 µl of Low-Buffer (1:200)

A280 is 0.1256 x 200=25.12

25.12/82975x1=0.0003027M

0.0003027 M x 49142.6(g/L)/1M=14.875g/L

[00117] Figure 4 shows a Coomassie Blue stained gel of purified human calreticulin in the amounts of 7.5 µg, 15µg and 30µg of human CRT, purified as described above. The gel shows final purified samples after the two ion chromatography purifications (i.e., Q-Sepharose followed by HRT chromatography) followed by one hydrophobic chromatography purification (ISO chromatography with 20-25% ammonium sulfate elution).

[00118] In a preferred embodiment, the two chromatography steps followed by the hydrophobic chromatography step, are performed immediately following one another. However, in certain circumstances it is useful to store a sample (at -20 or other suitable temperature, or even lyophilized (dry powder) after any one of the chromatography steps

prior to completing the three step purification of calreticulin. Thus, encompasses methods where the chromatography steps are performed after storage of an intermediate sample after any one or more of the purifications. Finally, it is contemplated that under certain conditions more than two ion chromatography steps may be desired. Thus, the present invention also encompasses methods where at least two ion chromatography steps are followed by at least one hydrophobic chromatography purification step, to yield a desired purified calreticulin protein product.

[00119] In certain embodiments, performing one ion chromatography step followed by one hydrophobic chromatography step will provide a desired purified calreticulin product. In yet additional embodiments, an ammonium sulfate precipitation (20-25%) of *E. coli* extracted protein followed directly by hydrophobic chromatography will provide a desired purified calreticulin product.

[00120] The Molecular Weight Standards used were the Low range Bio-Rad cat# 161-0304:

→97,400 daltons Phosphorylase b
→66,200 daltons Serum albumin
→45,000 daltons Ovalbumin
→31,000 daltons Carbonic anhydrose
→21,500 daltons Trypsin inhibitor
→14,400 daltons Lysozyme

[00121] Figure 5 shows a vector map of pBAD/gIII-human CRT. Figure 6 shows the nucleotide sequence of the pBAD-hCRT plasmid (SEQ ID NO:6) including human calreticulin along with its amino acid sequence (SEQ ID NO:7) including the unique signal sequence from the pBAD construct. In the vector, human calreticulin cDNA was inserted into the EcoR1-Xba1 sites in the multiple cloning site. In a unique feature, the calreticulin cDNA terminates with a stop codon, therefore, c-myc and HIS-tag are not translated but left in the vector. Thus, the calreticulin is secreted and expressed without an affinity tag or label.

[00122] Alternatively, removing the nucleotide sequence encoding the HIS and c-myc tag and utilizing the stop codon present in the vector, is another option for eliminating read-through of the vector for producing a non-tagged calreticulin product.

[00123] Other features of the vector are as follow (and according to Invitrogen, Carlsbad, CA):

*ara*BAD promoter region: bases 4-276

Gene III secretion signal: bases 319-373

*rrn*B transcriptional termination region: bases 595-752

Ampicillin resistance gene (ORF): bases 595-752

pBR322 origin: bases 2037-2710

AraC ORF: bases 3241-4139.

* * *

[00124] The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and the accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

[00125] It is further to be understood that values are approximate, and are provided for description.

[00126] Patents, patent applications, publications, procedures, and the like are cited throughout this application, the disclosures of which are incorporated herein by reference in their entireties.

WHAT IS CLAIMED IS:

1. A method for producing calreticulin comprising: (a) culturing a microorganism having a calreticulin gene under suitable culture conditions; (b) inducing expression of the calreticulin gene to produce a calreticulin protein product that is free of an affinity tag or label; and (c) purifying the calreticulin product.

2. The method of claim 1, wherein the microorganism is bacteria.

3. The method of claim 2, wherein the bacteria is *Escherichia coli*.

4. The method of claim 1, wherein the calreticulin gene is human calreticulin.

5. The method of claim 1, wherein the calreticulin gene comprises SEQ ID NO:3.

6. The method of claim 1, wherein the purifying step comprises at least one ion chromatography purification followed by at least one hydrophobic chromatography purification.

7. The method of claim 1, wherein the purifying step comprises two sequential ion chromatography purifications followed by at least one hydrophobic chromatography purification.

8. The method of claim 6 or 7, wherein the hydrophobic chromatography step comprises utilizing 20-25% ammonium sulfate for elution.

9. The method of claim 6 or 7, wherein the chromatography steps are performed immediately following one another.

10. The method of claim 6 or 7, wherein the chromatography steps are performed after

storage of an intermediate sample after any one or more of the purifications.

11. A method for expressing calreticulin in a prokaryotic cell, which comprises:
(a) culturing a prokaryotic cell comprising a vector under conditions conducive to gene expression, wherein the vector comprises a nucleic acid encoding calreticulin, an origin of replication, and at least one selectable marker, and wherein the nucleic acid is operatively associated with an expression control sequence;
(b) inducing expression of the calreticulin gene to produce a calreticulin protein product that is free of an affinity tag or label; and
(c) purifying said calreticulin protein product.

12. The method of claim 11, wherein the prokaryotic cell is *E. coli*.

13. The method of claim 11, wherein the nucleic acid encodes a termination signal.

14. The method of claim 11, wherein the selectable marker is Kanamycin or Ampicillin.

15. The method of claim 11, wherein the purifying comprises at least one ion chromatography purification followed by at least one hydrophobic chromatography purification.

16. The method of claim 11, wherein the purifying comprises two sequential ion chromatography purifications followed by one hydrophobic chromatography purification.

17. The method of claim 15 or 16, wherein the hydrophobic chromatography step comprises utilizing 20-25% ammonium sulfate for elution.

18. A vector comprising an origin of replication, at least one selectable marker, and a multiple cloning site comprising a nucleic acid encoding a signal sequence and calreticulin, and wherein said nucleic acid is operatively associated with an expression control sequence, and wherein said nucleic acid encodes a termination signal.

19. The vector of claim 18, wherein the calreticulin is human calreticulin.

20. The vector of claim 18, wherein the nucleic acid comprises SEQ ID NO:3.

21. The vector of claim 18, wherein the selectable marker is Kanamycin or Ampicillin.

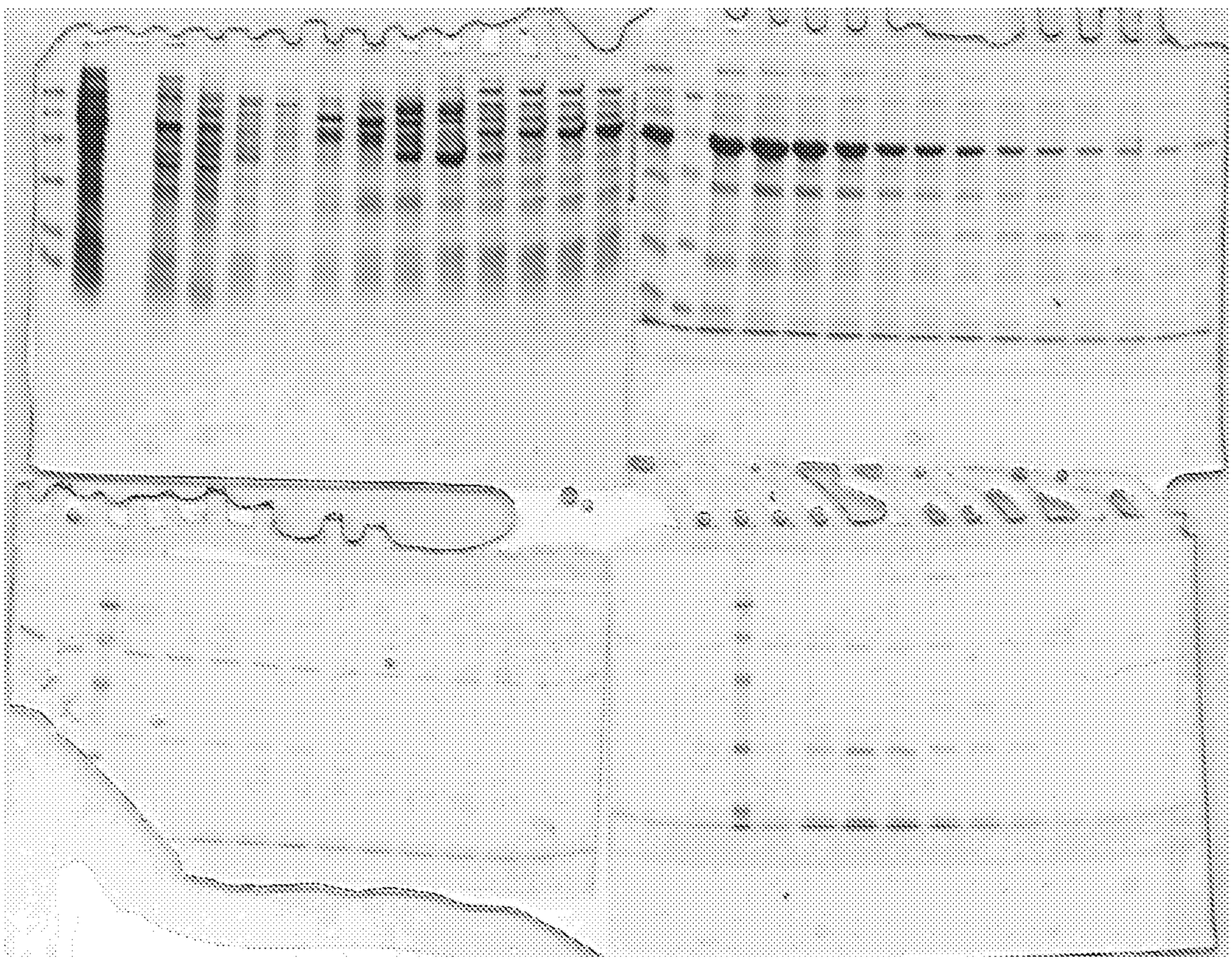
22. The vector of claim 18, wherein the vector comprises the vector map of Figure 5.

23. The vector of claim 18, wherein the origin of replication is an *E. coli* origin.

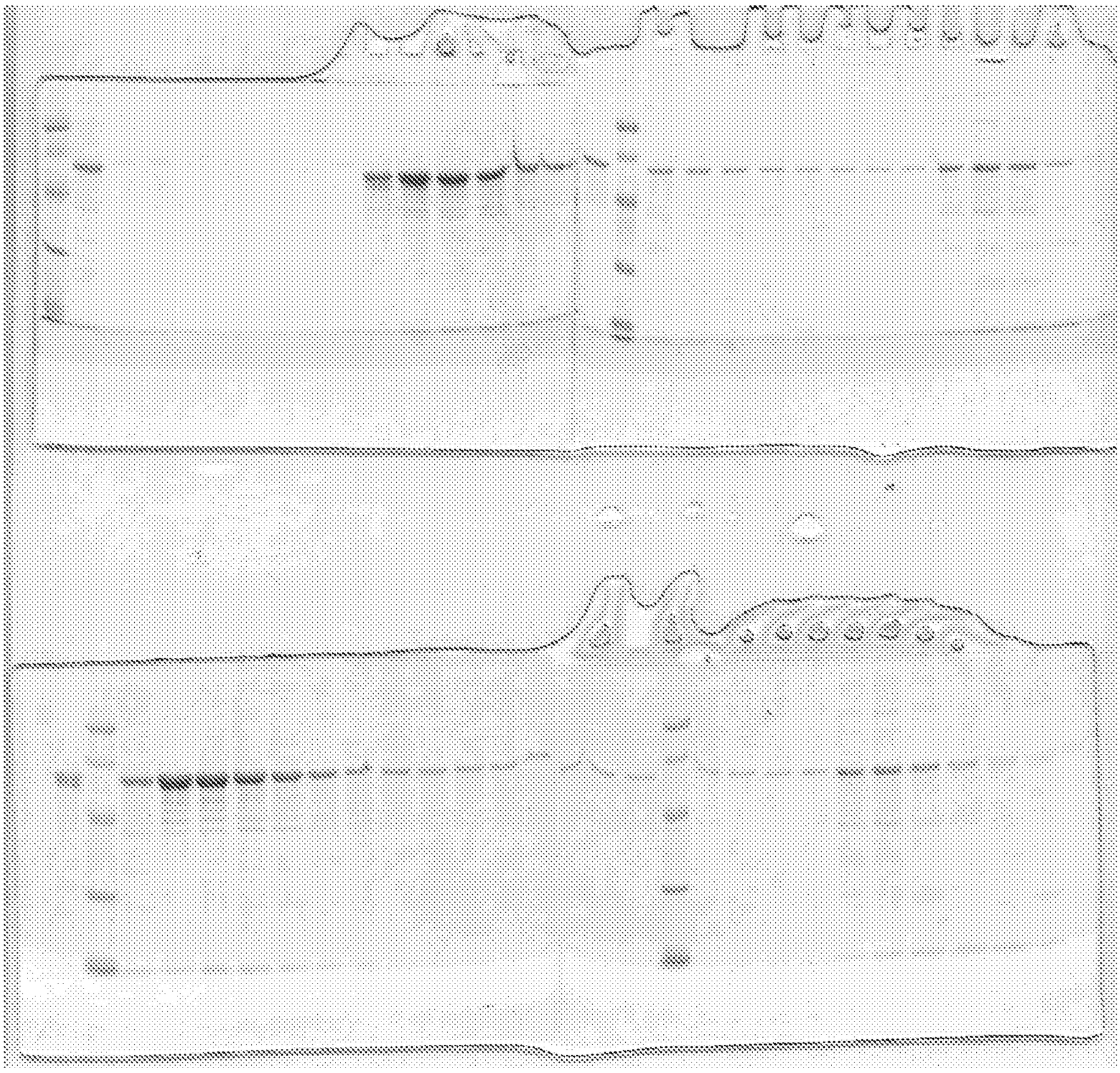
24. A host cell transformed with the vector of any one of claims 18-23.

25. Isolated calreticulin, prepared by any one of the methods of claims 1-17.

FIGURES 1A-D



FIGURES 2A-D



FIGURES 3A-D

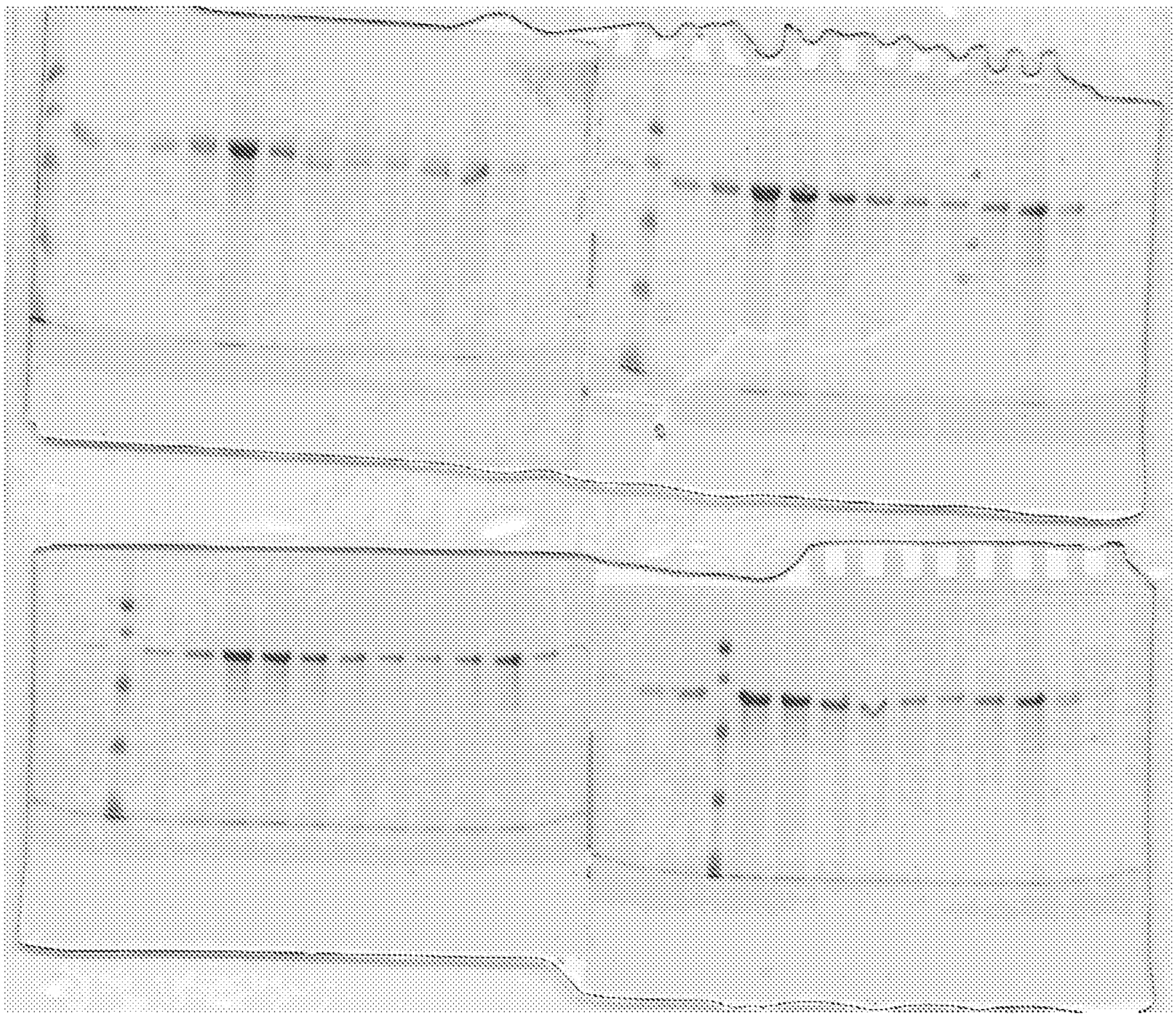


FIGURE 4

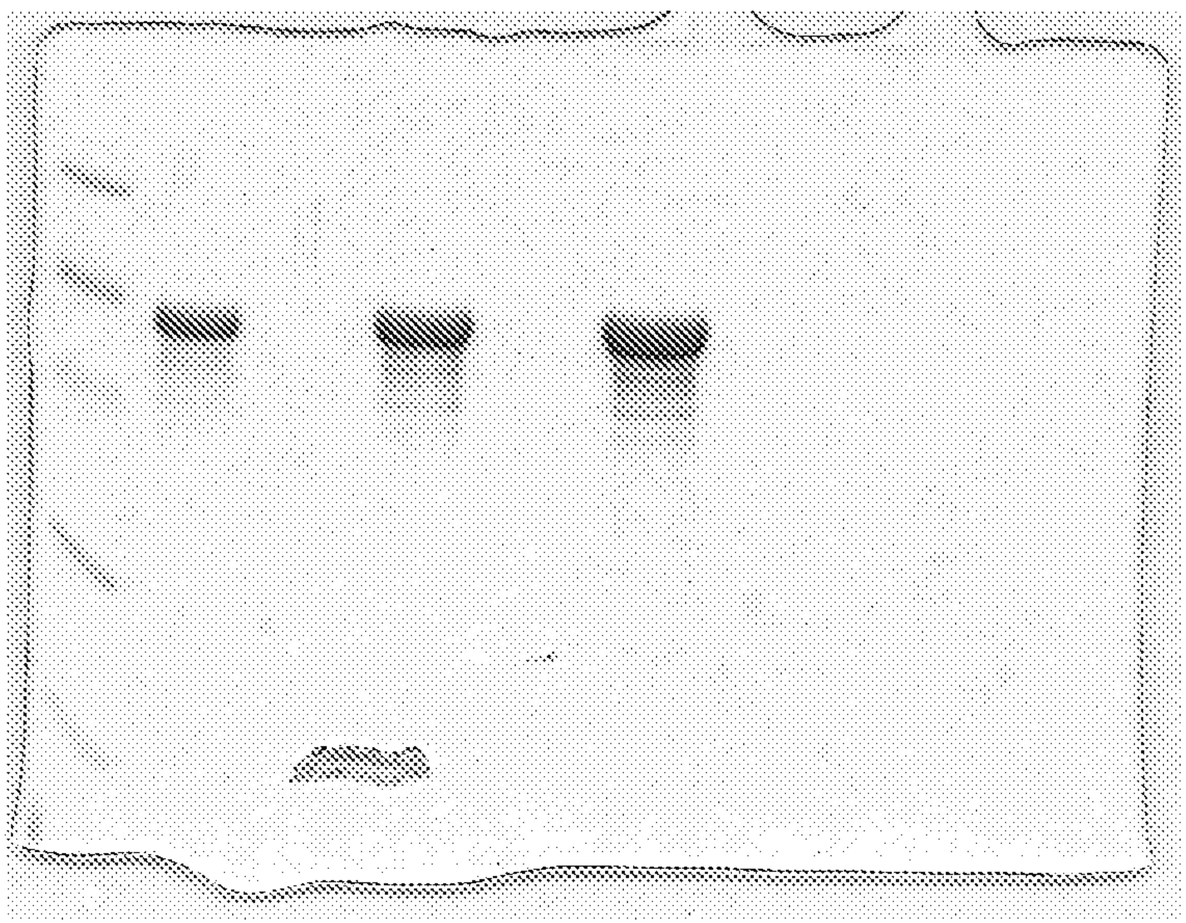


FIGURE 5

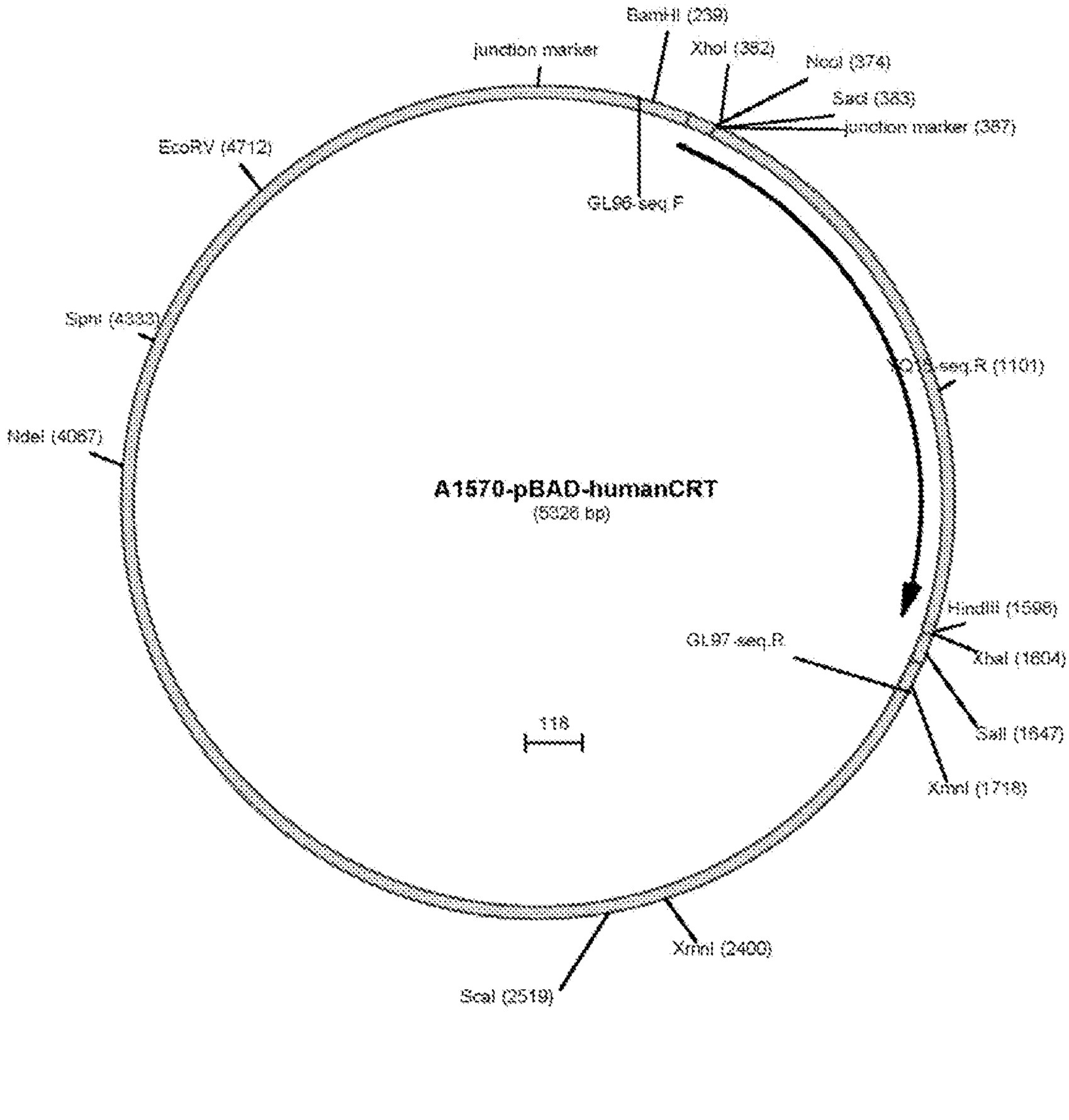

BamHI (239)
junction marker
XhoI (382)
NcoI (374)
SacI (383)
junction marker (387)
EcoRV (4712)
GL96-seq.F
SpnI (4338)
pQE-seq.R (1101)
NdeI (4067)
A1570-pBAD-humanCRT
(5826 bp)
HindIII (1588)
GL97-seq.R
XbaI (1604)
SalI (1647)
XmnI (1718)
118
XmnI (2400)
ScaI (2519)

FIGURE 6

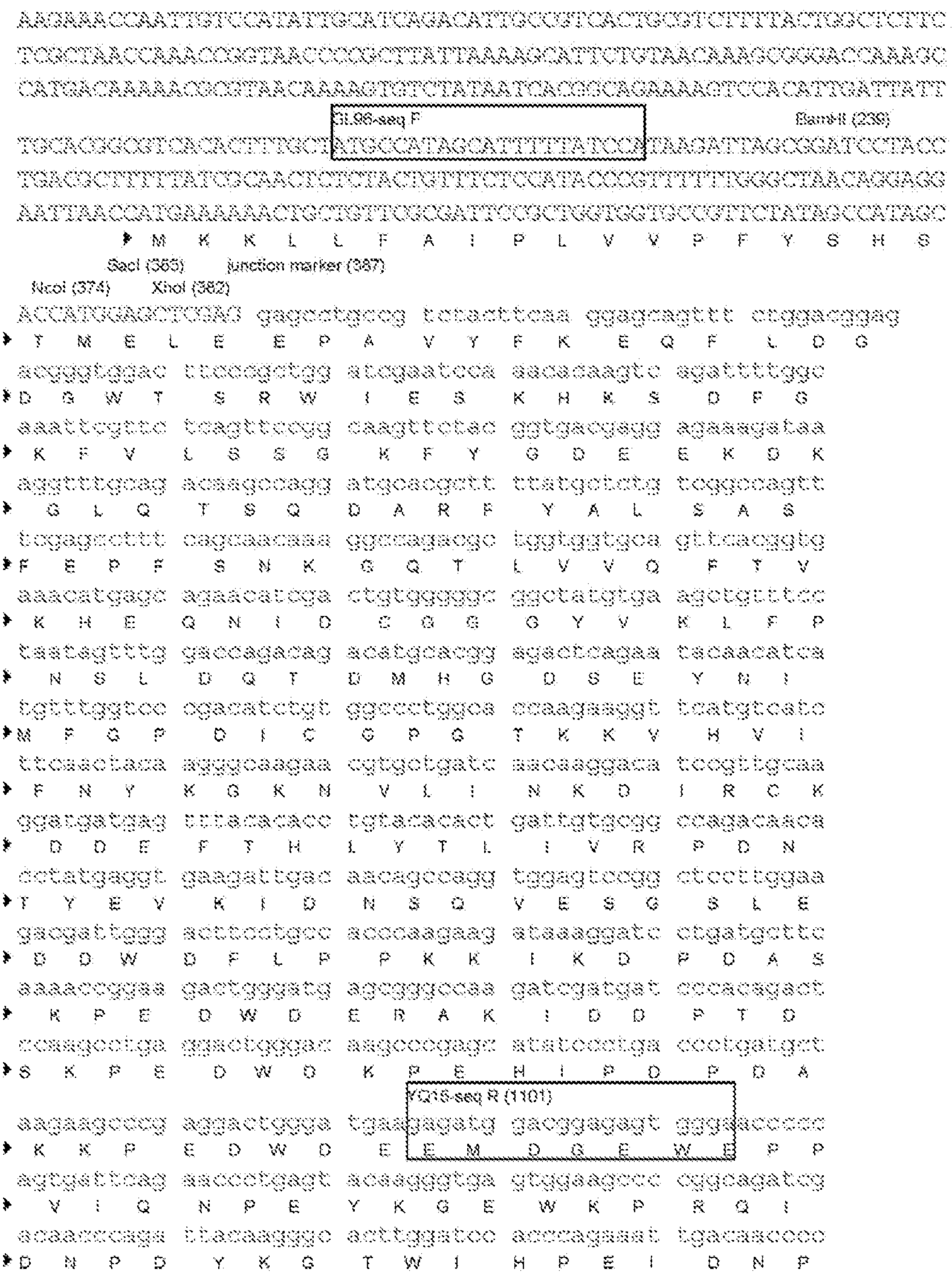

SL96-seq F
BamHI (239)
MKKLLFAIPLVVPFYSHS
SacI (365)
junction marker (387)
NcoI (374)
XhoI (382)
YQ15-seq R (1101)

FIGURE 6

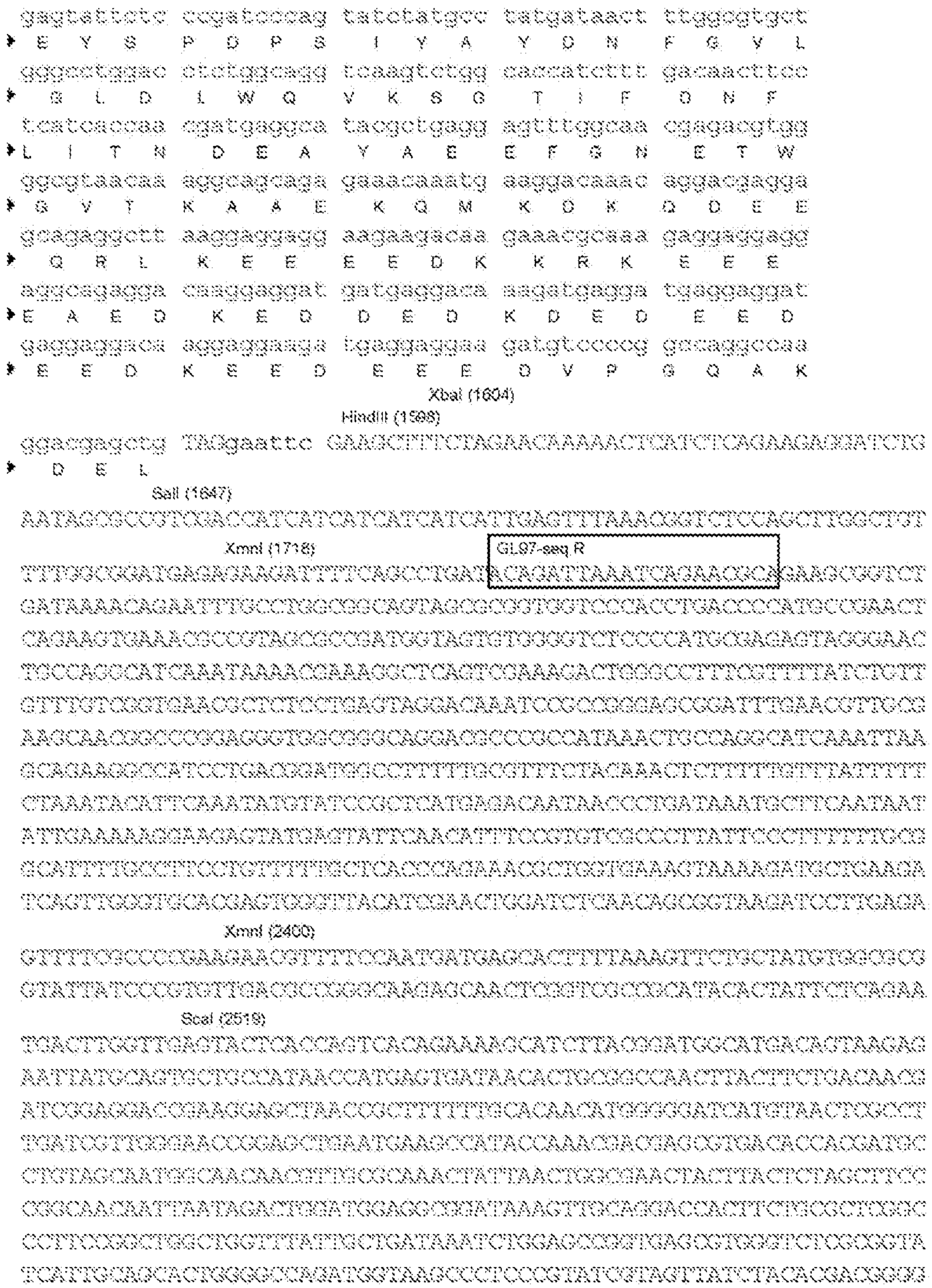

XbaI (1604)
HindIII (1598)
SalI (1647)
XmnI (1718)
GL97-seq R
XmnI (2400)
ScaI (2519)

FIGURE 6

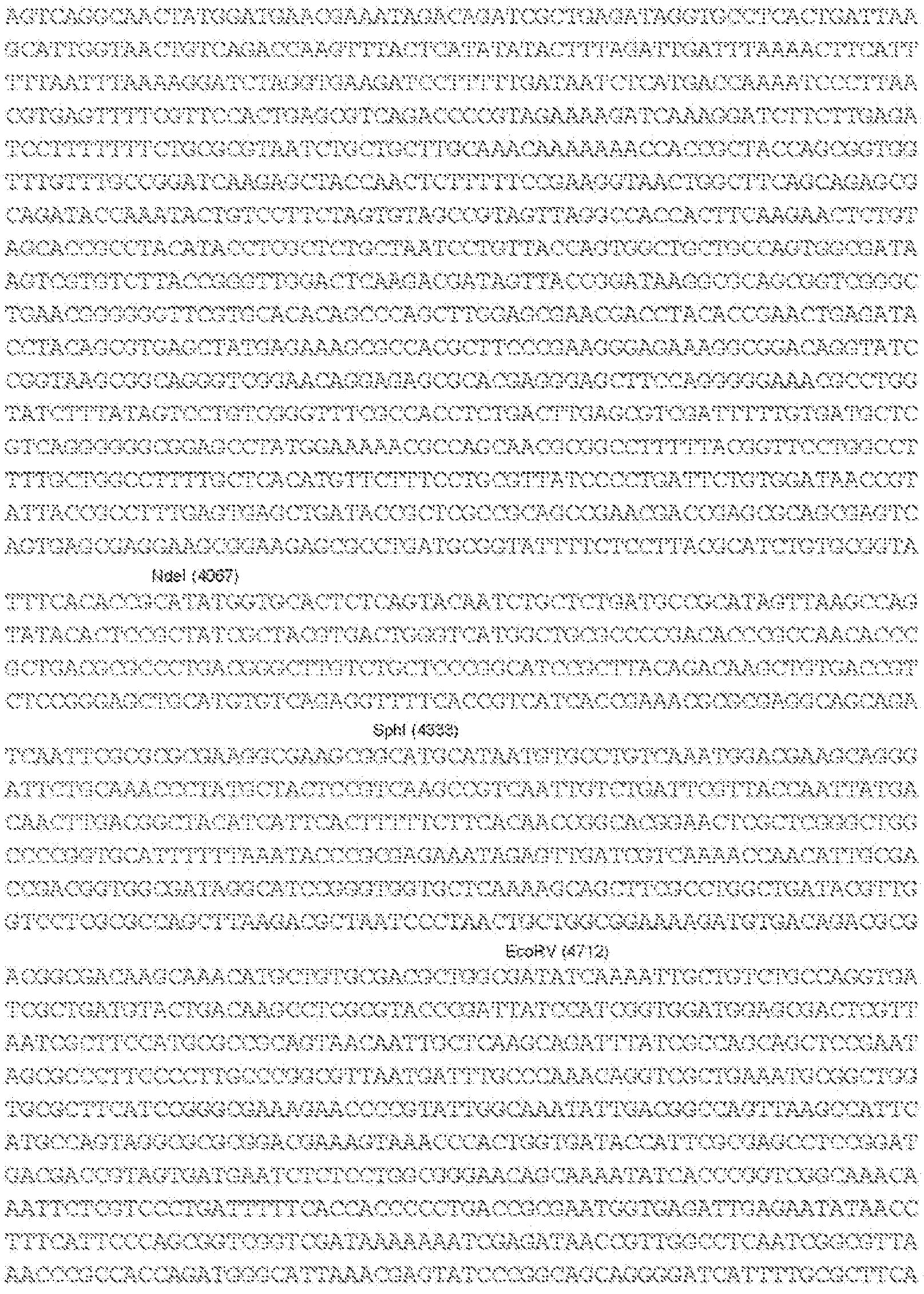


FIGURE 6