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

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
30 August 2007 (30.08.2007)

(10) International Publication Number
WO 2007/096192 A2

(51) **International Patent Classification:**
C12N 5/10 (2006.01) *A61K 39/00* (2006.01)
C12N 15/79 (2006.01)

(21) **International Application Number:**
PCT/EP2007/001606

(22) International Filing Date:
23 February 2007 (23.02.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/776,391 23 February 2006 (23.02.2006) US

(71) **Applicant** *(for all designated States except US):* **ERA BIOTECH, S.L.** [ES/ES]; Pare Cientific de Barcelona, Josep Samitier 1-5, E-08028 Barcelona (ES).

(72) Inventors; and

(75) Inventors/Applicants *(for US only):* **HEIFETZ, Peter Bernard** [US/US]; 10805 Birch Bluff Avenue, San Diego, CA 92131 (US). **LLOMPART ROYO, Blanca** [ES/ES]; Ronda Sant Antoni, 3. Pral 2°, E-0801 1 Barcelona (ES). **MARZÁBAL LUNA, Pablo** [ES/ES]; Felip Ii, 226. 7°-2ª, E-08027 Barcelona (ES). **BASTIDA VIRGILI, Miriam** [ES/ES]; Sant Antoni Mª Claret, 2. 2° - 2ª, E-08750 Molins De Rei - Barcelona (ES). **LUDEVID MUGICA, Maria Dolores** [ES/ES]; Esglesia 8 Bis, E-08960 Sant lust Desvern - Barcelona (ES). **TORRENT QUETGLAS, Margarita** [ES/ES]; Betania 3, 1° 3ª, E-08023 Barcelona (ES). **O CONNOR, Kevin James** [IE/ES]; Ciudad De Consuegra, 4. 3°-1ª, E-08820 El Prat De Llobregat - Barcelona (ES). **PALLISSE BERGWERF, Roser** [ES/ES]; loan Borras, 54, E-08197 Valldoreix (ES). **LLOP, Mª Inmaculada** [ES/ES]; Frederica Montseny N° 26, Bajos, E-08980 St. Feliu De Llobregat (ES).

(74) Agent: ARIAS SANZ, Juan; ABGPatentes, S.L., Orense, 68 7th Floor, E-28020 Madrid (ES).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, **BR,** BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, 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, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, 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, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, 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*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PRODUCTION OF BIOLOGICALLY ACTIVE PROTEINS

(57) Abstract: A fusion protein that is expressed in a recombinant protein body-like assembly (RPBLA) in host eukaryotic cells and organisms is disclosed. More particularly, a biologically active polypeptide fused to a protein sequence that mediates the induction of RPBLA formation is expressed and accumulated in host cells after transformation with an appropriate vector. The eukaryotic host cell does not produce protein bodies in the absence of the fusion protein. Methods for preparing and using the RPBLAs and the fusion protein are also disclosed, as are nucleic acid molecules that encode the fusion proteins.

WO 2007/096192 A2

PRODUCTION OF BIOLOGICALLY ACTIVE PROTEINS

CROSS-REFERENCE TO RELATED APPLICATION

This application claims benefit of provisional application Serial No. 60/776,391 that was filed on February 23, 2006.

TECHNICAL FIELD

The present invention contemplates the production of biologically active recombinant peptides and proteins, collectively referred to as polypeptides, in eukaryotic cells and organisms as host systems. More particularly, a biologically active polypeptide is fused to a protein body-inducing sequence (PBIS) that mediates the induction of recombinant protein body-like assemblies (RPBLA) to form a fusion protein that is stably expressed and accumulated in the host system as an RPBLA after transformation of the host cells with an appropriate vector.

BACKGROUND ART

The production of recombinant proteins for therapeutic, nutraceutical or industrial uses has enjoyed great success over the past decade. Introduction of heterologous genes having a desired nucleotide sequence leads to expression of a polypeptide or protein having the corresponding desired amino acid residue sequence or primary structure. In many instances, however, the protein or polypeptide expressed has had the amino acid residue sequence of the naturally-produced material, but has lacked the biological activity of that material.

Biological activity, given the proper primary structure of the expressed product, can be a function of the product having the proper folding and internal

hydrogen, Van der Waals, ionic and disulfide bonding, and also having proper post-translational modification, as for instance glycosylation . For example, disulfide bond formation occurs spontaneously in the lumen of the endoplasmic reticulum (ER), but not in the cytosol of prokaryotes, which makes bacterial cells such as *E. coli* cells poor hosts for the synthesis of correctly-folded mammalian proteins that are normally stabilized by disulfide bonds. Disulfide bond formation can occur in the periplasmic space of *E. coli* were PDI-like proteins are functional (Fernandez, et al., 2001. *MoI. Microbiol.* Λpr 40(2):332-346), however the oxi-redox system is not very efficient .

A particular case in point relates to erythropoietin (EPO), a protein that stimulates red blood cell production. Recombinant EPO is disclosed in US Patent No. 4,703,008 to Lin. The patent discloses activities for EPO protein expressed from *E. coli, s. cerevisiae ,* and mammalian Chinese hamster ovary (CHO) and African green monkey kidney (COS-I) cells. Although anti-EPO antisera immunoreacted with EPO expressed by each cell type, only the proteins expressed from mammalian cells exhibited substantial *in vivo* biological activity as EPO, and similar concentrations by antibody assay, *in vitro* and *in vivo* assays. The mammalian-expressed protein is that used to treat humans .

It is believed that those differences in biological activity were a function of glycosylation in that *E. coli,* a prokaryote, cannot glycosylate its expressed proteins. Yeast cells are eukaryotes, but their glycosylation pattern for secreted proteins is different from a mammal's. On the other hand, the CHO and COS-I cells used to provide protein of substantial biological

activity were mammalian and the protein expressed therefrom was useful. Published studies of glycosylated and unglycosylated EPO indicate that glycosylation plays a critical role in stabilizing erythropoietin to denaturing conditions. Narhi et al., (1991) *J. Biol. Cham.* **266(34)** :23022-23026. In addition, it has been reported that *in vivo* life time and activity of EPO can be related to the glycosylation of the molecule.

Eukaryotic cells are therefore greatly preferred for recombinant production of therapeutic, industrial and other useful proteins of eukaryotic origin. Different eukaryotic cells and organisms have been shown to be able to produce active protein-based therapeutics. Unfortunately, the high costs frequently derived from low recombinant protein production levels and/or from protein isolation and purification procedures, can invalidate their industrial application. Active research is done to improve both production levels and purification procedures by different approaches.

One way of improving the efficiency of recombinant protein isolation is by means of intracellular concentration. One of these approaches is the random aggregation of recombinant proteins into non-secreted inclusion bodies which can be separated from lysed cells by density-based purification techniques. Inclusion bodies are amorphous protein deposits found in bacteria. Structural characterization studies showed that the insoluble nature of the inclusion bodies may be due to the hydrophobic mtermolecular interactions of non-native folded proteins (Seshadri et al., 1999, *Methods Enzymol.* 309:559-576). The general strategy used to recover active proteins from inclusion bodies requires the solubilization of the protein to disrupt the random aggregates followed by one or more

chemical refolding steps. This is an important problem to be solved because the renaturing efficiency of denaturated proteins can be limited, mostly if the protein contains disulf ide-bonds (Clare, Ed., Apr. 2001 *Curr. 0pm. Biotechnol.* **12** (2) :202-207).

More particularly, strong denaturants such as high concentration of chaotropic agents (i.e. urea and guanidinium hydrochloride) are used to solubilize unfolded proteins that accumulate in aggregates. The denaturants are thereafter dialyzed away in an attempt to refold the protein in a natural conformation. Biological activity of such refolded proteins is usually much less than that of the native-formed protein.

Protein bodies (PBs) are naturally-occurring structures in certain plant seeds that have evolved to concentrate storage proteins intracellularly in eukaryotic cells while retaining correct folding and biological activity. Protein bodies (PBs) share some of the characteristics of the inclusion bodies from bacteria. They are dense, and contain a high quantity of aggregated proteins that are tightly packed by hydrophobic interactions [Momany et al., 2006 *J Agric. Food Chem .* **Jan 25;54(2)** :543-547 and Garrat, et al,. 1993 *Proteins* **Jan;15(l)** :88-99] . Moreover, the presence of a large quantity of disulf ide-bonds in some of the PBIS, as for instance RX3, [Ludevid, et al., 1984 *Plant MoI. Biol.* 3:227-234 and Kawagoe et al., 2005 *Plant Cell* Apr **17(4)** :1141-1153] , which are probably involved in PB formation and stabilization, could represent an additional difficulty to produce a biologically active, native-folded protein, particularly a protein that contains cysteine residues .

The observation of biological activity without the need for refolding and renaturation of a wide variety of proteins produced in synthetic PBs in non-yeast eukaryotic hosts was therefore unexpected.

A new technology based on the fusion of a plant seed storage protein domain with the protein of interest (WO 2004/003207) has been developed to increase the stability and accumulation of recombinant proteins in higher plants. These storage proteins are specific to plant seeds wherein they stably accumulate in protein bodies (Galili et al., 1993, *Trends Cell Biol* 3:437-442) .

The storage proteins are inserted into the lumen of the endoplasmic reticulum (ER) via a signal peptide and are assembled either in the endoplasmic reticulum developing specific organelles called ER-derived protein bodies (ER-PBs) or in protein storage vacuoles (PSV) (Okita et al., 1996 *Annu . Rev. Plant Physiol MoI. Biol. 41:321*-350; Herman et al., 1999 *Plant Cell* 11:601-613; Sanderfoot et al., 1999 *Plant Cell* 11:629-642) . Full-length recombinant storage proteins have also been described to assemble in PB-like organelles in non-plant host systems as Xenopus oocytes.

Expression of cereal prolamms (the most abundant cereal storage proteins) has been described in *Xenopus* oocytes after injection of the corresponding mRNAs . This system has been used as a model to study the targeting properties of these storage proteins (Simon et al., 1990, *Plant Cell* 2:941-950; Altschuler et al., 1993, *Plant Cell* 5:443-450; Torrent et al., 1994, *Planta* 192:512-518) and to test the possibility of modifying the 19 kDa α-zem, a maize prolamin, by introducing the essential amino acids lysine and tryptophan into its sequence, without altering its stability (Wallace et al, 1988, *Science* 240:662-664).

Zems, the complex group of maize prolamins, have also been produced in yeast with various objectives. Coraggio et al., 1988, *Eur J Cell Biol* 47:165-172, expressed native and modified α-zeins in yeast to study targeting determinants of this protein. Kim et al., 2002, *Plant Cell* 14: 655-672, studied the possible σ-, β-, γ- and δ-zem interactions that lead to protein body formation. To address this question, they transformed yeast cells with cDNAs encoding these proteins. In addition, those authors constructed zein-GFP fusion proteins to determine the subcellular localization of zem proteins in the yeast cells but did not observe formation of dense, concentrated structures characteristic of *bona fide* PBs. It is worth to noting that Kim et al., 2002, *Plant Cell* 14: 655-672, concluded that yeast is not a good model to study zem interactions because zems, by themselves, were poorly accumulated in transformed yeast. The yeast cells were also used as a model to study the mechanisms that control the transport and protein body deposition of the wheat storage proteins called gliadins (Rosenberg et al., 1993, *Plant Physiol* 102:61-69).

Biological activity is particularly relevant for vaccines, which must induce a correct immune response in an immunized human or other animal. Several new vaccines are composed of synthetic, recombinant, or highly purified subunit immunogens (antigens) that are thought to be safer than whole-inactivated or live-attenuated vaccines. However, the absence of adjυvanting immunomodulatory components associated with attenuated or killed vaccines often results in weaker immunogenicity for such vaccines.

Immunologic adjutants are agents that enhance specific immune responses to vaccines. An immunologic adjuvant can be defined as any substance or formulation

that, when incorporated into a vaccine, acts generally to accelerate, prolong, or enhance the quality of specific immune responses to vaccine antigens. The word *adjuvant* is derived from the Latin verb *adjuvare*, which means to help or aid. Adjuvant mechanisms of action include the following: (1) increasing the biological or immunologic half-life of vaccine immunogens; (2) improving antigen delivery to antigen-presenting cells (APCs), as well as antigen processing and presentation by the APCs; and (3) inducing the production of immunomodulatory cytokines.

Phagocytosis involves the entry of large particles, such us apoptotic cells or whole microbes. The capacity of the cells to engulf large particles likely appeared as a nutritional function in unicellular organisms; however complex organisms have taken advantage of the phagocytic machinery to fulfil additional functions. For instance, the phagocytosis of antigens undertaken by the macrophages, the B-cells or the dendritic cells represents a key process in innate and adaptive immunity. Indeed, phagocytosis and the subsequent killing of microbes in phagosomes form the basis of an organism's innate defense against intracellular pathogens. Furthermore, the degradation of pathogens in the phagosome lumen and the production of antigenic peptides, which are presented by phagocytic cells to activate specific lymphocytes, also link phagocytosis to adaptive immunity (Jutras et al., 2005 *Annual Review in Cell Development Biology.* 21:511-27).

The proteins present on engulfed particles encounter an array of degrading proteases in phagosomes. Yet, this destructive environment generates peptides that are capable of binding to MHC class II molecules. Newly formed antigen-MHC class II complexes are delivered to the cell surface for presentation to CD4+ T cells (Boes et al, .

2002. *Nature* 418:983-988) . The activation of these cells induces the Th2 subset of cytokines such as IL-d and IL-S that help B cells to proliferate and differentiate, and is associated with humoral-type immune response.

A large body of evidence indicates that, in addition to the clear involvement of the MHC class II pathway in the immune response against phagocytosed pathogens, antigens from pathogens, including mycobacteria, *Salmonella, Brucella*, and *Leishmama* , can elicit an antigen cross-presentation. That is to say, the presentation of engulfed antigen by phagocytosis by the MHC class I-dependent response promotes the proliferation of CD8+ cytotoxic T cells (Ackerman et al., 2004 *Nature Immunology* 5(7): 678-684 Kaufmann et al., 2005 *Current Opinions in Immunology* 17 (1) :79-87).

Dendritic cells play a central antigen presentation role to induce the immune system (Blander et al., *Nature Immunology* 2006 10:1029-1035) . Although rare, dendritic cells are the most highly specialised APC, with ability both to instigate and regulate immune reactivity (Lau et al. 2003 *Gut* 52:307-314) . Although dendritic cells are important in presenting antigens, particularly to initiate primary immune responses, macrophages are the APC type most prominent in inflammatory sites and specialized for clearing necrotic and apoptotic material. Macrophages can act not only as APC, but can also perform either pro- or anti-inf lammatory roles, dependent on the means by which they are activated.

Considering that APC plays a central role in the induction and regulation of the adaptive immunity (humoral and cellular) , the recognition and phagocytosis of the antigen by those cells can be considered a key step in the immunization process. A wide variety of techniques based

on the uptake of fluorescent particles have been developed to study phagocytosis by the macrophages (Vergne et al, . 1998 *Analytical Biochemistry* **255** .127-132) .

An important aspect in veterinary vaccines is the genetic diversity of the species being considered and the requirement for generic systems that work across different species. To a large degree, this diversity limits the use of molecular targeting techniques to cell surface markers and immune modulators such as cytokines, because for many species including wildlife, only minimal knowledge of these molecules is available. Thus, adjuvants that rely on universal activation signals of the innate immune response (i.e. that are identical in different species) are to be preferred. Taking these requirements into consideration, particulate vaccine delivery systems are well suited for veterinary and wildlife vaccine strategies (Scheerlinck et al., 2004 *Methods* 40:118-124) .

As is discussed in greater detail hereinafter, the present invention discloses that the expression of a fusion protein comprised of (i) a protein sequence that mediates induction of recombinant protein body-like assemblies (RPBLAs) linked to (ii) a biologically active polypeptide (protein of interest or target) induces the accumulation of those RPBLAs in cells of eukaryotic organisms such as plants, fungi, algae and animals, producing a biologically active target (protein) .

BRIEF SUMMARY OF THE INVENTION

The present invention provides a system and method for producing a fusion protein containing a protein body-inducing sequence (PBIS) and a biologically active peptide or protein (often collectively referred to herein as a polypeptide or target) of interest in eukaryotic

cells. The fusion proteins containing the polypeptide of interest stably accumulate as recombinant protein body-like assemblies (RPBLAs) in the eukaryotic cells, which can be plant, animal, fungal or algal cells.

Cells of higher plants are preferred eukaryotic host cells in some embodiments, whereas cells of lower plants such as algae are preferred in other embodiments, cells of animals such as mammals and insects are preferred eukaryotic host cells in further embodiments and fungi are preferred eukaryotic host cells in still other embodiments. The fusion protein can be expressed constitutively or preferentially in particular cells in multi-cellular eukaryotes. The PBISs are able to mediate the induction of RPBLA formation and fusion protein entry and/or accumulation in these organelles, with appropriate folding and/or post-translational modifications such as basal glycosylation and disulfide bond formation to provide biological activity to the expressed peptide or protein of interest (targets) .

Thus, a eukaryotic host cell that contains a biologically active recombinant fusion protein within recombinant protein body-like assemblies (RPBLAs) is contemplated as one aspect of the present invention. The fusion protein contains two sequences linked together in which one sequence is a protein body-inducing sequence (PBIS) and the other is the sequence of at least 20 amino acid residues of a biologically active polypeptide. The biologically active polypeptide, as found in nature, can be heterologous to the recited eukaryotic host cells and is thus expressed in a second cell type that is different from the first-mentioned eukaryotic host cell, or it is produced synthetically. In addition, the eukaryotic host cell does not produce PBs in the absence of the fusion protein.

Thus, it is the expression of the fusion protein and the PBIS that causes the host cell to form protein body-like assemblies or RPBLAs.

In a particular embodiment, the nucleic acid sequence used for transformation comprises (i) a nucleic acid sequence coding for a PBIS, and (ii) a nucleic acid sequence comprising the nucleotide sequence coding for a product of interest. In one embodiment, the 3' end of nucleic acid sequence (i) is linked to the 5' end of said nucleic acid sequence (ii). In another embodiment, the 5' end of nucleic acid sequence (i) is linked to the 3' end of nucleic acid sequence (ii). Thus, the PBIS sequence can be at the N-termmus or the C-termmus of the fusion protein. It is to be understood that all of the DNA linkages discussed herein for the expression of a fusion protein are such that the two components of the fusion protein are expressed in frame.

The biologically active polypeptide of the fusion protein exhibits at least 25 percent, preferably at least 50 percent, more preferably 75 percent, and most preferably at least 90 percent of the biological activity of the same polypeptide isolated from the above second cell type in an assay of the activity of that polypeptide.

In another particular embodiment, the nucleic acid sequence used for transformation comprises, in addition to the before-mentioned nucleic acid sequences (i) and (ii), a nucleic acid sequence comprising the nucleotide sequence coding for a linker or spacer amino acid sequence. The spacer amino acid sequence can be an amino acid sequence cleavable, or not cleavable, by enzymatic or autoproteolytic or chemical means. In a particular embodiment, the nucleic acid sequence (in) is placed between the nucleic acid sequences (i) and (ii), e.g., the

3' end of nucleic acid sequence (in) is linked to the 5' end of said nucleic acid sequence (ii). In another embodiment, the 5' end of said nucleic acid sequence (in) is linked to the 3' end of nucleic acid sequence (ii).

Also, in a particular embodiment, the nucleic acid sequence used for transformation purposes encodes a sequence in accord with patent application WO 2004003207, wherein the nucleic acid sequence coding for the ammo acid sequence that is specifically cleavable by enzymatic or chemical means is present or absent. In a further embodiment, the fusion proteins can be a direct fusion between the PBIS and the peptide or protein of interest.

In a further embodiment, the method of the invention further comprises the isolation and purification of the biologically active fusion protein.

In another embodiment, the method of the invention further comprises the isolation and purification of the fusion protein, and obtaining a biologically active fusion protein. Thus, where the fusion protein is tightly assembled and enclosed within a membrane, it can be difficult to illustrate that the polypeptide is biologically active. As a consequence, the biological activity can be assayed after removal of the membrane, and if it is required, the solubilization of the fusion protein. A method of preparing a biologically active polypeptide is therefore contemplated.

In this method, recombinant protein body-like assemblies (RPBLAs) are provided that comprise a membrane-enclosed fusion protein. The RPBLAs are usually present in a generally spherical form having a diameter of about 0.5 to about 3 microns (μ), but in some instances are amorphous m shape and can vary widely in dimensions, but are still derived from the ER. The fusion protein contains two

sequences linl-ed together in which one sequence is a protein body-inducing sequence (PBIS) and the other is a biologically active polypeptide. The RPBLAs are contacted with an aqueous buffer containing a membrane-disassembling amount of a detergent (surfactant). That contact is maintained for a time period sufficient to disassemble the membrane and at a temperature that does not denature the biologically active polypeptide to separate the membrane and fusion protein. The separated fusion protein is thereafter collected in a usual manner, or can be acted upon further without collection.

In some embodiments, the separated fusion protein exhibits the biological activity of the biologically active polypeptide. In other embodiments, biological activity of the polypeptide is exhibited after the fusion protein is dissolved or dispersed in an appropriate buffer. In yet other embodiments, the fusion protein has to be cleaved into its constituent parts before biological activity of the polypeptide is exhibited. Thus, the biologically active polypeptide can be linked to the PBIS by a spacer amino acid sequence that is cleavable by enzymatic or chemical means. Then, upon cleavage, the biologically active polypeptide exhibits biological activity when cleaved from the PBIS of the fusion protein. In some embodiments, the fusion protein retains its activity even when still incorporated into the intact RPBLA.

In another embodiment, the biologically active polypeptide contains at least two N-linked glycosylation sequences .

In yet another preferred embodiment, the polypeptide of interest is fused to a natural or modified storage protein, as for instance, natural or modified prolamine or prolamin domains.

In another embodiment, the RPBLAs containing the biologically active polypeptide are used as a delivery system for the biologically active polypeptide. The benefits of this invention could be applied in drug delivery, vaccines and nutrition.

In yet another embodiment, the RPBLAs containing a polypeptide antigen can be used as a delivery system to provide ad]uvanticity (increase the immune response) . The administration of these RPBLAs can represent an improvement in the immunization parameters such as the speed, quantity, quality and duration of the immunization. The beneficial effect of administrating antigens in RPBLAs can be achieved because (i) the antigen is encapsulated and remains longer in the blood or in the gastrointestinal tract (slow release effect) and/or (ii) the antigen is better exposed to the immune system (RPBLAs as an antigen presentation vehicle) and/or (in) the presence of adjuvant molecules in the RPBLAs preparations, and/or (iv) the RPBLAs are carriers able to cross membranes that themselves provide ad]uvanticity, and/or others.

Thus, another aspect of the invention is a vaccine or inoculum (immunogenic composition) that comprises an immunogenic effective amount of RPBLAs that contain biologically active is recombinant fusion protein dissolved or dispersed in a pharmaceutically acceptable diluent. The recombinant fusion protein contains two sequences linked together in which one sequence is a PBIS and the other is a biologically active polypeptide to which an immunological response is to be induced by said vaccine or inoculum. The pharmaceutically acceptable diluent composition typically also contains water. In another embodiment an RPBLA not incorporating an antigen but

possessing active adjuvant properties is co-delivered with an isolated antigen to induce an immunological response.

In another embodiment, the PBIS can be used as a carrier to cross membranes. In a specific embodiment the PBIS is ZERA (RX3) or a fragment of it.

The present invention has several benefits and advantages .

One benefit is that use of the invention enables relatively simple and rapid expression of a desired recombinant biologically active protein in an eukaryotic cell of choice.

An advantage of the invention is that it provides a source of readily obtainable and purifiable recombinant biologically active protein due to the unique properties of the expression in RPBLAs.

Another benefit of the invention is that the fusion protein-containing RPBLAs can be used for delivery of vaccines, including oral delivery vaccine.

Another advantage of the present invention is that the fusion protein-containing RPBLAs can be used as is in an immunogen in an injectable vaccine.

Another advantage of the present invention is that RPBLAs can be used as insulators, membrane bond structures that isolate the expressed polypeptide from the rest of the cell components. These insulators protect the cell from the polypeptide activity, and the polypeptide from the cell, increasing the accumulation rate. Thus, difficult biologically-active polypeptides that are toxic and/or labile can be successfully expressed.

Still further benefits and advantages will be apparent to the skilled worker from the discussion that follows .

BRIEF DESCRIPTION OF THE DRAWINGS

In the drawings forming a portion of this disclosure,

Fig. 1, Panel A is the schematic representation of the constructs used for the CHO cells transfection studies. The construct pECFP-Nl corresponds to the control expressing the ECFP in the cytosol. The pRX3-ECFP and pRX3-Gx5-ECFP are the constructs expressing the fusion protein RX3-ECFP, in the absence or presence of a spacer formed by five glycine ammo acids (Gx5), respectively. The p22aZ-ECFP is the constructs coding for the maize alpha zein (22KDa) fused to ECFP. On the bottom, the pcDNAS.l(-) (Invitrogen) based vectors are represented along with several constructs discussed hereinafter. Panel B shows the schematic representation of binary vectors for plant transformation (upper) and the baculovirus vectors for insect infection (bottom) . "RX3" = N-terminal proline- π ch gamma-zein sequence; "(GIy) x5" = spacer formed by five glycines; "ECFP" = enhanced cyan fluorescent protein gene; "P_{CMV}" = human cytomegalovirus promoter; "P_{PH}" = Polyhed π n promoter; "P_{SV40}" = SV40 early promoter; "CaMV35S x2" = Double cauliflower mosaic virus promoter; "P_{Cbhi}" = *major* cellulase promoter; "t35S" = Cauliflower mosaic virus terminator; "TEV" = Translational enhancer of the tobacco etch virus ; "SV40 ter" = SV40 terminator; "HSV ter" = herpes simplex virus thymidine kinase polyadenylation signal; "cbhl ter" = major cellulase polyadenylation signal; "Kana/Neo" = kanamycin/neomycin resistance gene; "Amp R" = Ampicilin resistance gene; "Gentamicine" = Gentamicin resistance gene "SP_{cbhl}" = major cellulase signal peptide; "On fl" = fl single strand DNA origin; "On pUC" = plasmid replication origin; "BGH ter" = Bovine growth

hormone terminator; "P BLA" = beta lactamase gene promoter; "GFP" = Green fluorescent protein; "DsRED" = Dicosoma red fluorescent protein; "hGH" = human growth hormone; "EGF" = human epidermal growth factor; "EK" = bovine enterokmase; "GUS" = Glucuronidase; "RTB" = Lectin subunit of ricin *{Ricinus comunis)* ; "Casp2" = Human Caspase 2; "Casp3" = Human Caspase 3; "Int" = Ssp DNAb intern from New England Biolabs; "mint" = mutated version of Ssp DNAb intern (Aspl54 → Ala substitution) .

Fig. 2 shows immunoblots from subcellular fractionation studies of CHO cells transfected with pRX3-ECFP, pRX3-G-ECFP and pECFP-Nl as a control (Panel A)/ p3. l-RX3-hGH, p3.1-RX3, p3.1-RX3-EK, p3.1-RX3-C3, p3.1-RX3-C2, p3.1-RX3-GUS and p3.l-RX3-I-hGH plasmids (Panel B). In panel B the immunoblot from subcellular fractionation studies of tobacco plants agroinf iltrated with pB-RX3-RTB are also shown. Panel C corresponds to subcellular fractionation studies of insect larvae infected with pF-RX3-DsRED and pF-DsRED as a control. Transfected cell homogenates were loaded on step sucrose gradients, and after centrif ugation, the accumulation of the corresponding fusion proteins in the supernatant, interphase and pellet fractions was analyzed by immunoblot. The molecular weights and the antibody used in the immunoblot are indicated on the right. H, homogenate loaded in the density gradient; S, supernatant/ F_X, upper interphase of the X% w/w sucrose cushion/ P, pellet under 56% sucrose cushion.

Fig. 3 is a confocal microscopy photograph in six panels showing the localization of the fusion proteins RX3-ECFP (panel A), RX3-Gx5-ECFP (panel B), 22aZ-ECFP (panel D), RX3-GFP (panel E) and RX3-DsRED (panel F) in RPBLAs within transfected CHO cells. Some of the RPBLA structures containing the active (fluorescent) fusion proteins are

indicated by arrows. The localization of the ECFP in the cycosol and the nucleus (panel C) in CHO cells transfected by pECFP-Nl are shown as a control. "N" = nucleus.

Fig. 4 is a confocal microscopy photograph in four panels showing the localization of fluorescent RX3 fusion proteins in different hosts. In panel A is shown the confocal optical sections of epidermal leaf tissue from tobacco plants co-agroinf iltrated with pB-RX3-GFP and a binary vector coding for HcPRO, a suppressor of gene silencing. It can be observed a lot fluorescent RPBLAs containing the active RX3-GFP fusion protein. On the right, in Panel B, the merging of the RX3-GFP fluorescence and the contrast phase shows the high percentage of transiently transfected cells. The projection of optical sections of SF9 insect cells infected with pF-RX3-DsRED is shown in Panel C. One micrometer optical sections of fat tissue from insect larvae infected with pF-RX3-DsRED are shown in Panel D. Some of the RPBLA structures containing the active (fluorescent) fusion proteins are indicated by arrows .

Fig. 5 is a photograph in six panels (A-F) showing the localization of RX3 fusion proteins inside RPBLAs in CHO cells, four days after their transf ection . Optical microscopy was used to show CHO cells expressing RX3-hGH (Panels A and B) immunolocalized by using anti-RX3 and anti-hGH serum, respectively. Panel C shows RX3 protein immunolocalization with RX3 antiserum. Anti-hGH serum was used in Panel D to lmmunolocalize the RX3-I-hGH fusion protein. The incubation of CHO cells expressing RX3-GUS fusion protein with RX3 antiserum is shown in Panel E. Smaller RPBLAs were observed in CHO cells expressing RX3-EK, incubated with anti-RX3 serum (Panel F). The endoplasmic reticulum (ER) and the RPBLAs are indicated.

Fig. 6 shows western blots that illustrate the induction of Ssp DNAb intern self-cleavage after RX3-I-hGH fusion protein solubilization from a RPBLAs preparation by low speed centrifugation. Panels A and B illustrate the self-cleavage of the RX3-I-hGH (wild type Ssp DNAb intein) fusion protein, after solubilization. The RX3-Im-hGH (mutated Ssp DNAb intern) fusion protein was included as a negative control. Equivalent volumes of the samples were loaded per lane, and the western blot was performed with anti-RX3 serum (Panel A) or anti-hGH serum (Panel B). The full length fusion proteins are indicated with white arrowheads and the products of the Ssp DNAb intern self-cleavage (RX3-I in Panel A, and hGH in Panel B) are indicated with black arrowheads. Panel C illustrates the comparison of RX3-I-hGH fusion protein self-cleavage efficiency after 0.1% SDS (Sl) and biphasic (S2) solubilization. Equivalent volumes of the samples were loaded per lane, except TO that was overloaded 4-folds. The incubation with anti-hGH serum shows the full length fusion protein RX3-I-hGH (white arrowhead) and the liberated hGH (black arrowhead). "S" = Soluble fraction; "U" = insoluble fraction; "TO" = Sample before induction of intern self-cleavage.

Fig. 7 contain micrographs that show the uptake and processing of RX3-DsRED RPBLAs from insect larvae by macrophages. In panel A, confocal microscopy analysis of macrophages 1 hour after incubation with insect RX3-DsRED RPBLAs is shown. On the left, 2 macrophages can be observed by phase contrast microscopy. On the right, is shown the merged image of DsRED fluorescence (black arrowheads) and the self-fluorescence of the macrophages (white arrowheads) from 1 micrometer optical section of the same cells. The observation of the nucleus (N) in this

optical section indicates that the RPBLAs have been taken up and are now intracellular. Panel B shows the time course study (1 and 10 hours) of DsRED fluorescence emitted by the macrophages, after incubation for 1 hour with RPBLAs containing RX3-DsRED. On the left, the phase contrast microscopy shows the presence of macrophages. On the right, the DsRED fluorescence of 1 micrometer optical sections shows the presence of undigested RPBLAs at 1 hour (white arrowhead) and a more homogeneous DsRED fluorescence pattern at 10 hours indicative of digested and dispersed RPBLAs. The inset image corresponds to a higher magnification of the undigested RPBLAs observed at 1 hour.

Fig. 8 contains micrographs that show the uptake of RX3-DsRED RPBLAs from insect larvae by dendritic cells. The photographs correspond to dendritic cells incubated with RPBLAs (Panel A) and membrane-less RPBLAs (Panel B) over time (2, 5 and 10 hours). In the upper of each panel the phase contrast shows the presence of dendritic cells. At the bottom, the DsRED fluorescence from the same dendritic cells shows the presence of RPBLAs absorbed to the plasma membrane (2 hours) or phagocytosed inside the cell (5 and 10 hours). "N" = nucleus.

DETAILED DESCRIPTION OF THE INVENTION

A contemplated recombinant biologically active polypeptide is a portion of a fusion protein that forms recombinant protein body-like assemblies (RPBLAs), frequently membrane-enclosed, in the host cells in which they are expressed. RPBLA formation is induced by a protein body-inducing sequence (PBIS) comprised of a signal peptide and storage protein domain that forms high density deposits inside the cells. These dense deposits can accumulate in the cytosol, an endomenbrane system

organelle, mitochondria, plastic! or can be secreted. With the exception of certain cereal plant seeds, the eukaryotic host cell does not itself produce protein bodies (PBs) in the absence of the fusion protein. Thus, it is the expression of the fusion protein and its PBIS portion that causes the host cell to form protein body-like assemblies or RPBLAs.

A contemplated fusion protein comprises two polypeptide sequences linked together directly or indirectly by a peptide bond, in which one sequence is that of a protein body-inducing sequence (PBIS) linked to the second sequence that is a biologically active polypeptide product (e.g., peptide or protein) of interest (target) . The biologically active polypeptide, as found in nature, is heterologous to the recited eukaryotic host cells and is thus expressed in a second cell type that is different from the first-mentioned eukaryotic host cell, or it is produced synthetically. That is, the biologically active polypeptide is heterologous to the recited eukaryotic host cells. PBIS are protein or polypeptide amino acid sequences that mediate the induction of RPBLA formation and the protein entry and/or accumulation in organelles such as the ER. The fusion protein when free and separated from the PBIS exhibits a biological activity similar to that of the polypeptide.

The biologically active polypeptide of the fusion protein exhibits at least 25 percent, preferably at least 50 percent, more preferably at least 75 percent and most preferably at least 90 percent of the biological activity of the same polypeptide isolated from the above second cell type, or synthesized *m vitro.* A material is considered "biologically active" or "bioactive" if it has interaction

with or effect on any metabolite, protein, receptor, organelle, cell or tissue in an organism.

These biological activities can be readily determined and quantified using standard techniques for determining the activity of that polypeptide. For example, assay results for biological activity between the polypeptide isolated from the second cell type, or synthesized *in vitro,* and the expressed polypeptide can be compared. When comparing the activity of a fusion protein, the proportion of that material provided by the PBIS and any linker sequence are taken into account in the assay comparison. Biological activity can be exhibited by the expressed RPBLAs, the fusion protein as a protein free of a surrounding membrane or as a target polypeptide that is free of its PBIS.

In a particular embodiment, the nucleic acid sequence used for transformation comprises (i) a nucleic acid sequence coding for a PBIS, and (ii) a nucleic acid sequence comprising the nucleotide sequence coding for a product of interest. In one embodiment, the 3' end of nucleic acid sequence (i) is linked to the 5' end of said nucleic acid sequence (ii). In another embodiment, the 5' end of nucleic acid sequence (i) is linked to the 3' end of nucleic acid sequence (ii). Thus, the PBIS sequence can be at the N-termmus or the C-terminus of the fusion protein. It is to be understood that all of the DNA linkages discussed herein for the expression of a fusion protein are such that the two components of the fusion protein are expressed in frame.

Most protein bodies have round-shaped (generally spherical) structures, with diameters of about 0.5 to about 3.0 μ. When expressed in animal cells, the RPBLAs are generally spherical in shape, have diameters of about 0.5

to about 3 microns (μ) and have a surrounding membrane. RPBLAs expressed in plants are also usually generally spherical, have diameters of about 0.5 to about 2 μ, and are surrounded by a membrane. RPBLAs expressed in either plants, animals or fungi are derived from the ER if targeted there by an ER-specific secretion signal and accumulate externally to the ER envelope of the host cell following assembly. It is noted that EGF-containmg RPBLAs expressed in the ER of plant cells were not generally spherical, and were amorphous in shape and of non-uniform size .

The recombinant protein body-like assemblies have a predetermined density that can differ among different fusion proteins, but is known for a particular fusion protein being prepared. That predetermined density of the RPBLAs is typically greater than that of substantially all of the endogenous host cell proteins present in the homogenate, and is typically about 1.1 to about 1.35 g/ml. The high density of novel RPBLAs is due to the general ability of the recombinant fusion proteins to assemble as multimers and accumulate into ordered aggregates associated with membranes. The contemplated RPBLAs are expressed in eukaryotes and can be characterized by their densities as noted above, and their size and shape.

The polypeptide portion of the fusion protein is believed to obtain its biological activity from folding within the ER and in some instances from glycosylation in the ER. Interestingly, most plants, animals such as mammals and single celled eukaryotes such as fungi, N-glycosylate proteins in the same pattern based upon the tripeptide glycosylation sequence Asn-X-Ser or Asn-X-Thr, where "X" is any amino acid residue but proline. Thus, a Glc3Mang (GIcNAc) 2 N-linked polypeptide is formed initially,

and is trimmed back after formation to a Man7_g (GIcNAc) 2 N-linked polypeptide that can be excreted to the Golgi or retained within the ER. This basal glycosyla tion is remarkably similar across eukaryotic genera. Further post-translational modification such as host-specif icterrainal glycosylation can occur in the Golgi for proteins not maintained in RPBLAs as are the fusion proteins contemplated here

In this method, recombinant protein body-like assemblies (RPBLAs) are provided that frequently comprise a membrane-enclosed fusion protein ordered assembly, are preferably present in a generally spherical form having a diameter of about 0.5 to about 3 microns. The fusion protein contains two sequences linked together in which one sequence is a protein body-inducing sequence (PBIS) and the other is a biologically active polypeptide. The RPBLAs are contacted with an aqueous buffer containing a membrane-disassembling amount of a detergent (surfactant) . That contact is maintained for a time period sufficient to disassemble the membrane and at a temperature that does not denature the biologically active polypeptide (e.g., above freezing to about 40^0C) to separate the membrane and fusion protein. The separated fusion protein is thereafter collected in a usual manner, or can be acted upon further without collection. Illustrative useful surfactants include Triton-X 100, CHAPS and the like as are will known in biochemistry for solubilizing lipids.

The separated fusion protein is typically in an insoluble form due to the interactions among the PBIS portions of the fusion protein mediated at least in part by the presence of cysteine residues. However, the polypeptide of interest is complexed with eukaryotic chaperones and foldases derived from the ER and hence is

held in a correctly folded conformation despite being tethered to the assembled (and hence insoluble) PBIS domain. The PBIS-PBIS interactions can be disrupted and the fusion protein solubilized by contacting the fusion protein with an aqueous buffer that contains a reducing agent such as dithiothreitol or 2-mercaptoethanol or β-mercaptoethanol (β-ME). Conditions are chosen so as to not disrupt and unfold the attached biologically active protein of interest. The separated, solubilized fusion protein that contains the biologically active polypeptide is then collected or otherwise used. In addition, the two portions of the fusion can be cleaved from each other upon solubilization. It is to be understood that that cleavage need not be at the exact borders between the two portions.

In some embodiments, the separated fusion protein exhibits the biological activity of the biologically active polypeptide. In other embodiments, the fusion protein is dissolved or dispersed in a suitable buffer to exhibit the biological activity of the polypeptide. For example, as discussed in detail hereinafter, human growth hormone (hGH) expressed in RPBLAs in mammalian cells and solubilized as a fusion protein exhibited significant activity and also as a cleaved polypeptide exhibited activities substantially similar to that of the native polypeptide.

In yet other embodiments, for the fusion protein has to be cleaved into its constituent parts before biological activity of the polypeptide is exhibited. Thus, the biologically active polypeptide can be linked to the PBIS by a by a spacer amino acid sequence that is cleavable by enzymatic or chemical means. Then, upon cleavage from the BPIS of the fusion protein and assay, the target (biologically active) polypeptide exhibits biological activity. Studies discussed hereinafter illustrate

biological activity of the T-20 polypeptide cleaved from its fusion partner and produced in plants.

Protein Body-inducing Sequences

A contemplated protein body-inducing sequences (PBIS) and the host cell are preferably of different biological phyla. Thus, the PBIS is preferably from a higher plant, a spermatophyte, whereas the host cell is a eukaryote that is other than a spermatophyte and can be an animal cell, as for instance mammalian or insect cells, a fungus, or an algal cell, all of which are of different phyla from spermatophytes . A PBIS and the host cell can also be from the same phylum so that both can be from a higher plant, for example. Illustrative, non-limiting examples of PBIS include storage proteins or modified storage proteins, as for instance, prolamms or modified prolamine, prolamm domains or modified prolamin domains. Prolamms are reviewed in Shewry et al., 2002 *J. Exp. Bot.* 53**(370)** :947-958 . Preferred PBIS are those of prolamin compounds such as gamma-zem, alpha-zem, delta-zem, beta-zein, rice prolamin and the gamma-gliadm that are discussed hereinafter.

A PBIS also includes a sequence that directs a protein towards the endoplasmic reticulum (ER) of a plant cell. That sequence often referred to as a leader sequence or signal peptide can be from the same plant as the remainder of the PBIS or from a different plant or an animal or fungus. Illustrative signal peptides are the 19 residue gamma-zem signal peptide sequence shown in WO 2004003207 (US 20040005660), the 19 residue signal peptide sequence of alpha-gliadm or 21 residue gamma-gliadm signal peptide sequence (see, Altschuler et al., 1993 *Plant Cell* 5:443-450; Sugiyama et al., 1986 *Plant Sci.* 44:205-

209; and Rafalski et al., 1984 *EMBO* J 3(6):1409-1415 and the citations therein.) The pathogenesis-related protein of PRlO class includes a 25 residue signal peptide sequence that is also useful herein. Similarly functioning signal peptides from other plants and animals are also reported in the literature.

The characteristics of the signal peptides responsible for directing the protein to the ER have been extensively studied (von Heijne et al., 2001 *Biochim. Biophys. Acta* Dec 12 1541(1-2):114-119). The signal peptides do not share homology at a primary structure, but have a common tripartite structure: a central hydrophobic h-region and hydrophilic N- and C-terminal flanking regions. These similarities, and the fact that proteins are translocated through the ER membrane using apparently common pathways, permits interchange of the signal peptides between different proteins or even from different organisms belonging to different phyla (See, Examples 1 and 2 hereinafter, and Martoglio et al., 1998 *Trends Cell Biol.* Oct; 8(10):410-415). Thus, a PBIS can include a signal peptide of a protein from a phylum different from higher plants.

Gamma-Zein, a maize storage protein whose DNA and amino acid residue sequences are shown hereinafter, is one of the four maize prolamms and represents 10-15 percent of the total protein in the maize endosperm. As other cereal prolamms, alpha- and gamma-zems are biosynthesized in membrane-bound polysomes at the cytoplasmic side of the rough ER, assembled within the lumen and then sequestered into ER-derived protein bodies (Herman et al., 1999 *Plant Cell* 11:601-613; Ludevid et al., 1984 *Plant MoI. Biol.* 3:277-234; Torrent et al., 1986 *Plant MoI. Biol.* 7:93-403).

Gamma-Zein is composed of four characteristic domains i) a peptide signal of 19 amino acids, ii) the repeat domain containing eight units of the hexapeptide PPPVHL (SEQ ID NO:1) [(53 amino acid residues *{aa)*], in) the ProX domain where proline residues alternate with other amino acids (29 *aa)* and iv) the hydrophobic cysteine rich C-termmal domain (111 aa) .

The ability of gamma-zein to assemble in ER-derived RPBLAs is not restricted to seeds. In fact, when gamma-zem-gene was constitutively expressed in transgenic Arabidopsis plants, the storage protein accumulated within ER-derived PBLS in leaf mesophyl cells (GeIi et al., 1994 *Plant Cell* 6:1911-1922) . Looking for a signal responsible for the gamma-zein deposition into the ER-derived protein bodies (prolamms do not have KDEL signal), it has been demonstrated that the proline-πch N-terminal domain including the tandem repeat domain was necessary for ER retention. In this work, it was also suggested that the C-termmal domain could be involved in protein body formation, however, recent data (WO2004003207A1) demonstrate that the proline-rich N-terminal domain is necessary and sufficient to retain in the ER and to induce the protein body formation. However, the mechanisms by which these domains promote the protein body assembly are still unknown, but evidence from *m vitro* studies suggests that the N-terminal portion of gamma-zem is able to self-assemble into ordered structures.

It is preferred that a gamma-zein-based PBIS include at least one repeat and the amino-terminal nine residues of the ProX domain, and more preferably the entire Pro-X domain. The C-termmal portion of gamma-zem is not needed, but can be present. Those sequences are shown in

US 20040005660 and designated as RX3 and *PA,* respectively, and are noted hereinafter.

Inasmuch as protein bodies are appropriately so-named only in seeds, similar structures produced in other plant organs and in non-higher plants are referred to generally as synthetic PBs or recombinant protein body-like assemblies (RPBLAs) .

Zems are of four distinct types: alpha, beta, delta, and gamma. They accumulate in a sequential manner in the ER-derived protein bodies during endosperm development. Beta-zein and delta-zein do no accumulate in large amount in maize PBs, but they were stable in the vegetative tissues and were deposited in ER-derived protein body-like structures when expressed in tobacco plants (Bagga et al., 1997 *Plant Cell* Sep 9(9) :1683-1696) . This result indicates that beta-zem, as well as delta-zein, can induce ER retention and protein body formation.

The wheat prolamin storage proteins, gliadms, are a group of K/HDEL-less proteins whose transport via the ER appears to be complex. These proteins sequester in to the ER where they are either retained and packaged into dense protein bodies, or are transported from the ER via the Golgi into vacuoles. (Altschuler et al., 1993 *Plant Cell* 5:443-450.)

The gliadms appear to be natural chimeras, containing two separately folded autonomous regions. The N-terminus is composed of about 7 to about 16 tandem repeats rich in glutamme and proline. The sequence of the tandem repeats varies among the different gliadms, but are based on one or the other consensus sequences PQQPFPQ (SEQ ID NO:47), PQQQPPFS (SEQ ID NO:48) and PQQPQ (SEQ ID NO: 49). The C-termmal region of the protein contains six to eight cysteines that form intramolecular disulfide

bonds. The work of the Altschuler et al. group indicates that the N-terminal region and consensus sequences are responsible for PB formation m the ER from gamma-gliadm . (Altschuler et al., 1993 *Plant Cell* 5:443-450.)

Illustrative other useful prolamin-type sequences are shown in the Table below along with their GenBank identifiers .

PROTEIN NAME	GENBANK ID
α-Zeın (22kD)	M86591
Albumın (32 kD)	X70153
γ-Zeın (27kD)	X53514
γ-Zeın (50kD)	AF371263
δ-Zeın (18kD)	AF371265
7S Globulın or Vıcılın type	NM 113163
11S Globulın or Legumın type	DQ256294
Prolamın 13 kD	AB016504
Prolamın 16 kD	AY427574
Prolamın 10 kD	AF294580
γ-Glıadın	M36999
γ-Glıadın precursor	AAA34272

Further useful sequences are obtained by carrying out a BLAST search in the all non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF (excluding environmental samples) data base as described in Altschul et al., 1997 *Nucleic Acids Res.* 25:3389-3402 using a query such as those shown below:

RX3 query (SEQ ID NO: 2)
Alpha-zein (SEQ ID NO: 3)
Rice prolamin query (SEQ ID NO: 4)

An illustrative modified prolamin includes (a) a signal peptide sequence, (b) a sequence of one or more copies of the repeat domain hexapeptide PPPVHL (SEQ ID NO:

(b) of the protein gamma-zem, the entire domain containing eight hexapeptide units; and (c) a sequence of all or part of the ProX domain of gamma-zem. Illustrative specific modified prolamms include the polypeptides identified below as R3, RX3 and P4 whose DNA and ammo acid residue sequences are also shown below.

Particularly preferred prolamms include gamma-zem and its component portions as disclosed in published application WO2004 003207, the rice rP13 protein and the 22 kDa maize alpha-zem and its N-terminal fragment. The DNA and ammo acid residue sequences of the gamma-zem, rice and alpha-zem proteins are shown below.

Gamma-zem of 27 kD
DNA Sequence (SEQ ID NO: 5)
Protein Sequence (SEQ ID NO: 6)

RX3
DNA Sequence (SEQ ID NO: 7)
Protein Sequence (SEQ ID NO: 8)

R3
DNA Sequence (SEQ ID NO: 9)
Protein Sequence (SEQ ID NO: 10)

P4
DNA Sequence (SEQ ID NO: 11)
Protein Sequence (SEQ ID NO: 12)

XlO
DNA Sequence (SEQ ID NO: 13)
Protein Sequence (SEQ ID NO: 14)

rP13 - rice prolamin of 13 kD homologous to the clone - AB016504 Sha et al., 1996 *Biosci. Biotacnnol . Biochem.* 60(2) :335-337; Wen et al., 1993 *Plant Physiol.* **101(3)** :1115-1116; Kawagoe et al., 2005 Plant Cell 17 (4) :114 1-1153 ; Mullins et al., 2004 J . *Agric. Food Chem .* 52 (8) :2242-224 6 ; Mitsukawa et al., 1999 *Biosci. Biotechnol . Biochem.* **63(11)** :1851-1858

Protein Sequence (SEQ ID NO: 15)

DNA Sequence (SEQ ID NO: 16)

22aZt N-terminal fragment of the maize alpha-zein of 22 kD - V01475 Kim et al., 2002 *Plant Cell* 14 (3) :655-672 ; Woo et al., 2001 *Plant Cell* **13** (10) :2297-2317 ; Matsushima et al., 1997 *Biochim. Biophys . Acta* **1339 (1)** :14-22 ; Thompson et al., 1992 *Plant MoI. Biol.* 18 (4) :827-833 .

Protein Sequence (full length) (SEQ ID NO: 17)

DNA Sequence (full length) (SEQ ID NO: 18)

Gamma-Gliadin precursor- AAA34272- Scheets et al., 1988 Plant Sci. 57:141-150.

Protein Sequence (SEQ ID NO: 19)

DNA Sequence (M36999) (SEQ ID NO:20)

Beta zein - AF371264 - Woo et al., (2001) Plant Cell 13 (10), 2297-2317.

DNA (SEQ ID NO: 21)

Protein (SEQ ID NO: 22)

Delta zein 10 kD- AF371266- Woo et al., (2001) Plant Cell 13 (10), 2297-2317. and Kirihara et al., (1988) Gene. Nov 30;71 (2) :359-70.

DNA (SEQ ID NO: 23)

Protein (SEQ ID NO:24)

Signal Peptides

Gamma-Zem (SEQ ID NO:25)

Alpha-Gliadin (SEQ ID NO:26)

Gamma-Gliadin (SEQ ID NO:27)

PRlO (SEQ ID NO:28)

Proteins of Interest

Examples of polypeptides or proteins of interest (targets) include any protein having therapeutic, nutraceutical , agricultural, biocontrol, or industrial uses. Illustrative activities of such proteins include (a) light capture and emission as are provided by green fluorescent protein (GFP) , enhanced cyan fluorescent protein (ECFP), red fluorescent protein (DsRED) and the like; (b) enzymatic activity as can be associated with primary and secondary intracellular signaling and metabolic pathways, is exemplified by enterokinase, beta-glucuromdase (GUS), phytase, carbonic anhydrase, and industrial enzymes (hydrolases, glycosidases, cellulases, oxido-reductases , and the like); (c) protein-protein, protein-receptor, and protem-ligand interaction such as, for example antibodies (mAbs such as IgG, IgM, IgA, etc.) and fragments thereof, hormones [calcitonin, human growth hormone (hGH) , epidermal growth factor (EGF) and the like] , protease inhibitors, antibiotics, antimicrobials, HIV entry inhibitors [Ryser et al., 2005 *Drug Discov Today.* **Aug. 15/10(16)** :1085-1094] , collagen, human lactoferrin, and cytokines; (d) protein and peptides antigens for vaccines (human immunodeficiency virus, HIV; hepatitis B

pre-surface, surface and core antigens, Foot and Mouth Disease Virus (FMDV) structural polyprotein gene P1 [Dus Santos ec al., 2005 *Vaccine.* **Mar** 7;23(15):1838-1843] T cell stimulating peptides of US Patent No. 4,882,145, gastroenteritis corona virus, human papilloma virus, and the like); (e) protein-non protein interactions such as, phytohaemaggl υtinin (PHA), the Ricin Toxin subunit B (RTB) and other lectins.

Assays for the bioactivity of such expressed polypeptides are well known in the art and are available in one or more publications. For example, the ECFP (enhanced cyan fluorescent protein) activity can be measured by quantifying the fluorescence emitted at a 470-530 nm wavelength when the protein has been exited at 458 nm. See, Richards et al., 2003 *Plant Cell Rep.* 22:117-121. The enzymatic activity of enterokinase (EK), for example, can be measured with two different approaches. The activity can be determined by analyzing the cleavage of a fusion protein containing the enterokinase specific cleavage site by western blot, as discussed in the Invitrogen Life Technologies catalog (E180-01 and E180-2), and also by quantifying the EK activity using fluorogenic peptide substrate for EK (Sigma G-5261, CAS® RN 70023-02-8); enzyme activity is measured by an increase of fluorescence (excitation at 337 nm, emission at 420 nm) caused by the release of β-naphthylamine from the peptide over time. See, LaVallie et al., 1993 J. *Biol. Chem.* **268**(31):23311-23317. The activity of the enzyme beta-glucuronidase (GUS) can be measured by the conversion of the substrate MUG (4-methyl umbelliferyl glucuronide) to the product MU. This product can be quantified by measuring the fluorescence with excitation at 365 nm, emission at 455 nm on a spectrof luorimeter . See, Pai-Hsiang et al., 2001 J. *Plant*

Physiol. **158(2)** :247-254; and Jefferson et al., 1987 *EMBO J* 6:3901-3907. Phytase assays are carried out by the quantification of inorganic *ortho* phosphates liberated from the AAM reagent consisting of acetone, 5.0 N sulfuric acid, and 10 mM ammonium molybdate. See, Ullah et al., 1999 *Biochem. Biophys . Res. Commun .* **264** (1) :201-206 . Similar assays are available for other biological proteins. The RTB activity assays can be performed by measuring the binding of RTB to asialofetuin, lactose and galactose, as described in Reed et al., 2005 *Plant Cell Rep.* **Apr,-24(1)** :15-24 .

The EGF is a growth factor involved in fibroblasts proliferation. The EGF activity can be assayed by the quantification of the induction of DNA synthesis measured by incorporation of the pyrimidme analog 5-bromo-2'-deoxyuridine (BrdU), instead of thymidine, into the DNA of proliferating cells using the cell proliferation ELISA kit [Oliver, et al., 2004 *Am. J. Physiol. Cell Physiol.* 286:1118-1129; Catalog no. 1647229, Roche Diagnostics, Mannheim, Germany]

It is noted that light capture and emission constitutes a separate and special type of "biological activity" in that such activity does not provide therapeutic, nutraceutical, agricultural, biocontrol, or industrial use as do the other types of activity noted above. The polypeptides of this class of targets are included herein as biologically active because they share some of the required secondary, tertiary and quaternary structural features that are possessed by the target molecules that provide therapeutic, nutraceutical, biocontrol, or industrial uses. These proteins are useful, however, as reporter molecules in many types of assays or screens used in the analysis or discovery of biologically

important molecules, and their luminescent activity requires the presence of correct secondary and tertiary protein structure. It is possibly more accurate to refer to the group of targets as those polypeptides that are biologically active and/or luminescently active.

Illustrative DNA and amino acid residue sequences for illustrative proteins of interest are provided below.

ECFP
DNA(SEQ ID NO:29)
protein (SEQ ID NO:30)

GUS1381
DNA (SEQ ID NO: 31)
protein (SEQ ID NO:32)

GUS1391Z
DNA (SEQ ID NO: 33)
protein (SEQ ID NO: 34)

Salmon calcitonin BAC57417
Protein sequence (SEQ ID NO: 35)
DNA sequence (SEQ ID NO: 36)

hEGF - Construction based in the AAF85790 without the signal peptide
Protein sequence (SEQ ID NO: 37)
DNA sequence (SEQ ID NO: 38)

hGH - Construction based in the P01241 without the signal peptide
Protein sequence (SEQ ID NO: 39)

DNA sequence (SEQ ID NO: 40)

In another embodiment, the recombinant fusion protein further comprises in addition to the sequences of the PBIS and product of interest, a spacer ammo acid sequence. The spacer ammo acid sequence can be an ammo acid sequence cleavable by enzymatic or chemical means or not cleavable. By "not cleavable" it is meant that cleavage of the spacer does not occur without destruction of some or all of the biologically active polypeptide.

In a particular embodiment, the spacer ammo acid sequence is placed between the PBIS and biologically active polypeptide. An illustrative amino acid sequence is cleavable by a protease such as an enterokmase, Arg--C endoprotease, Glu--C endoprotease, Lys--C endoprotease, Factor Xa, SUMO proteases [Tauseef et al., 2005 *Protein Expr. Puπ f.* 2005 Sep 43(1) :l-9] and the like. Alternatively, the spacer ammo acid sequence corresponds to an auto-cleavable sequence such as the FMDV viral auto-processing 2A sequence, interns such as the Ssp DNAb intern and the like as are commercially available from New England Biolabs and others. The use of an intern linker sequence is preferred as such sequences can be selectively induced to cause protein splicing and thereby eliminate themselves from an expressed, recovered, protein. Interns are particularly interesting since they do not require large protein enzymes to reach their target site in order to cleave the PBIS from the protein of interest. This property may be particularly useful for direct isolation of proteins of interest from intact RPBLAs. Alternatively, an ammo acid sequence is encoded that is specifically cleavable by a chemical reagent, such as, for example, cyanogen bromide that cleaves at methionine residues.

In a further embodiment, the nucleic acid sequence used for transformation purposes is as disclosed according to co-assigned patent application WO 2004003207, with or without the nucleic acid sequence coding for the cleavable amino acid sequence.

Methods of Preparation

In a preferred embodiment, the fusion proteins are prepared according to a method that comprises transforming an eukaryotic host cell system such as an animal, animal cell culture, plant or plant cell culture, fungus culture, insect cell culture or algae culture with a nucleic acid (DNA or RNA) sequence comprising (i) a first nucleic acid coding for a PBIS that is operatively linked in frame to (ii) a second nucleic acid sequence comprising the nucleotide sequence coding for a polypeptide product of interest that is biologically active; that is, the nucleic acid sequence that encodes the PBIS is chemically bonded (peptide bonded) to the sequence that encodes the polypeptide of interest such that both polypeptides are expressed from their proper reading frames and the protein of interest is biologically active. It is also contemplated that appropriate regulatory sequences be present on either side of the nucleic acid sequences that encode the PBIS and protein of interest as is discussed hereinafter. Such control sequences are well known and are present in commercially available vectors. The use of indirect means of introducing DNA, such as via viral transduction or infection, is also contemplated, and shall be used interchangeably with direct DNA delivery methods such as transfection .

The transformed host cell or entity is maintained for a time period and under culture conditions suitable for

expression of the fusion protein and assembly of the expressed fusion protein into recombinant protein body-like assemblies (RPBLAs). Upon expression, the resulting fusion protein accumulates in the transformed host-system as high density recombinant protein body-like assemblies. The fusion protein can then be recovered from the host cells or the host cells containing the fusion protein can be used as desired, as for an animal food containing an added nutrient or supplement. The fusion protein can be isolated as part of the RPBLAs or free from the RPBLAs.

Culture conditions suitable for expression of the fusion protein are typically different for each type of host entity or host cell. However, those conditions are known by skilled workers and are readily determined. Similarly, the duration of maintenance can differ with the host cells and with the amount of fusion protein desired to be prepared. Again, those conditions are well known and can readily be determined in specific situations. Additionally, specific culture conditions can be obtained from the citations herein.

In one embodiment, the 3' end of the first nucleic acid sequence (i) is linked (bonded) to the 5' end of the second nucleic acid sequence (ii). In other embodiment, the 5' end of the first nucleic acid sequence (i) is linked (bonded) to the 3' end of the second nucleic acid sequence (ii). In another embodiment, the PBIS comprises a storage protein or a modified storage protein, a fragment or a modified fragment thereof.

In another particular embodiment, a fusion protein is prepared according to a method that comprises transforming the host cell system such as an animal, animal cell culture, plant, plant cell culture, fungus or algae with a nucleic acid sequence comprising, in addition to the

nucleic acid sequences (i) and (11) previously mentioned, an in frame nucleic acid sequence (111) that codes for a spacer amino acid sequence. The spacer amino acid sequence can be an amino acid sequence cleavable by enzymatic or chemical means or not cleavable, as noted before. In one particular embodiment, the nucleic acid sequence (in) is placed between said nucleic acid sequences (i) and (ii), e.g., the 3' end of the third nucleic acid sequence (in) is linked to the 5' end of the second nucleic acid sequence (ii). In another embodiment, the 5' end of the third nucleic acid sequence (in) is linked to the 3' end of the second nucleic acid sequence (ii).

A nucleic acid sequence (segment) that encodes a previously described fusion protein molecule or a complement of that coding sequence is also contemplated herein. Such a nucleic acid segment is present in isolated and purified form in some preferred embodiments.

In living organisms, the ammo acid residue sequence of a protein or polypeptide is directly related via the genetic code to the deoxyribonucleic acid (DNA) sequence of the gene that codes for the protein. Thus, through the well-known degeneracy of the genetic code additional DNAs and corresponding RNA sequences (nucleic acids) can be prepared as desired that encode the same fusion protein ammo acid residue sequences, but are sufficiently different from a before-discussed gene sequence that the two sequences do not hybridize at high stringency, but do hybridize at moderate stringency.

High stringency conditions can be defined as comprising hybridization at a temperature of about 50°-55°C in 6XSSC and a final wash at a temperature of 68°C in 1-3XSSC. Moderate stringency conditions comprise hybridization at a temperature of about 50°C to about 65°C

in 0.2 to 0.3 M NaCl, followed by washing at about 50°C to about 55°C in 0.2X SSC, 0.1% SDS (sodium dodecyl sulfate) .

A nucleic sequence (DNA sequence or an RNA sequence) that (1) itself encodes, or its complement encodes, a fusion protein containing a protein body-inducing sequence (PBIS) and a polypeptide of interest is also contemplated herein. As is well-known, a nucleic acid sequence such as a contemplated nucleic acid sequence is expressed when operatively linked to an appropriate promoter in an appropriate expression system as discussed elsewhere herein. This nucleic acid sequence can be delivered directly or indirectly (via an appropriate vector organism such as a virus or bacterium) to the host eukaryotic cell, and can be integrated stably into the host nuclear or organellar genome, or transiently expressed without genome integration.

Different hosts often have preferences for a particular codon to be used for encoding a particular amino acid residue. Such codon preferences are well known and a DNA sequence encoding a desired fusion protein sequence can be altered, using *in vitro* mutagenesis for example, so that host-preferred codons are utilized for a particular host in which the fusion protein is to be expressed.

A recombinant nucleic acid molecule such as a DNA molecule, comprising a vector containing one or more regulatory sequences (control elements) such as a promoter suitable for driving the expression of the gene in a compatible eukaryotic host cell organism operatively linked to an exogenous nucleic acid segment (e.g., a DNA segment or sequence) that defines a gene that encodes a contemplated fusion protein, as discussed above, is also contemplated in this invention. More particularly, also

contemplated is a recombinant DNA molecule that comprises a vector comprising a promoter for driving the expression of the fusion protein in host organism cells operatively linked to a DNA segment that defines a gene encodes a protein body-inducing sequence (PBIS) linked to a polypeptide of interest. That recombinant DNA molecule, upon suitable transfection and expression in a host eukaryotic cell, provides a contemplated fusion protein as RPBLAs.

As is well known in the art, so long as the required nucleic acid, illustratively DNA sequence, is present, (including start and stop signals), additional base pairs can usually be present at either end of the DNA segment and that segment can still be utilized to express the protein. This, of course, presumes the absence in the segment of an operatively linked DNA sequence that represses expression, expresses a further product that consumes the fusion protein desired to be expressed, expresses a product that consumes a wanted reaction product produced by that desired fusion protein, or otherwise interferes with expression of the gene of the DNA segment.

Thus, so long as the DNA segment is free of such interfering DNA sequences, a DNA segment of the invention can be about 500 to about 15,000 base pairs in length. The maximum size of a recombinant DNA molecule, particularly an expression vector, is governed mostly by convenience and the vector size that can be accommodated by a host cell, once all of the minimal DNA sequences required for replication and expression, when desired, are present. Minimal vector sizes are well known. Such long DNA segments are not preferred, but can be used.

A DNA segment that encodes a before-described fusion protein can be synthesized by chemical techniques,

for example, the phosphotriester method of Matteucci et al., 1981 J. *Am. Cham. Soc.,* 103:3185. Of course, by chemically synthesizing the coding sequence, any desired modifications can be made simply by substituting the appropriate bases for those encoding the native ammo acid residue sequence. However, DNA segments including sequences specifically discussed herein are preferred.

DNA segments containing a gene encoding the fusion protein are preferably obtained from recombinant DNA molecules (plasmid vectors) containing that gene. A vector that directs the expression of a fusion protein gene in a host cell is referred to herein as an "expression vector" .

An expression vector contains expression control elements including the promoter. The fusion protein-coding gene is operatively linked to the expression vector to permit the promoter sequence to direct RNA polymerase binding and expression of the fusion protein-encoding gene. Useful in expressing the polypeptide coding gene are promoters that are inducible, viral, synthetic, constitutive as described by Paszkowski et al., 1989 *EMBO J.,* 3:2719 and Odell et al., 1985 *Nature,* 313:810, as well as temporally regulated, spatially regulated, and spatiotemporally regulated as given in Chua et al., 1989 *Science,* 244:174-181.

Expression vectors compatible with eukaryotic cells, such as those compatible with cells of mammals, algae or insects and the like, are contemplated herein. Such expression vectors can also be used to form the recombinant DNA molecules of the present invention. Eukaryotic cell expression vectors are well known in the art and are available from several commercial sources. Normally, such vectors contain one or more convenient restriction sites for insertion of the desired DNA segment

and promoter sequences Optionally, such vectors contain a selectable marker specific for use in eukaryotic cells.

Production of a fusion protein by recombinant DNA expression in mammalian cells is illustrated hereinafter using a recombinant DNA vector that expresses the fusion protein gene in Chinese hamster ovary (CHO) host cells, Cosl monkey host and human 293T host cells. This is accomplished using procedures that are well known in the art and are described in more detail in Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratories (1989).

An insect cell system can also be used to express a contemplated fusion protein. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) or baculovirus is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in Trichoplusia larvae. The sequences encoding a fusion protein can be cloned into a non-essential region of the virus, such as the polyhedrm gene, and placed under control of the polyhedrin promoter. Successful insertion of a fusion protein sequence renders the polyhedrin gene inactive and produces recombinant virus lacking coat protein. The recombinant viruses can then be used to infect, for example, *s. Frugiperda* cells or *Trichoplusia* larvae in which the fusion protein can be expressed. E. Engelhard et al. (1994) *Proc. Natl. Acad. Sci., USA,* 91:3224-3227; and V. Luckow, "Insect Cell Expression Technology", pages 183-218, in Protein Engineering: Principles and Practice, J.L. Cleland et al. eds., Wiley-Liss, Inc, 1996). Heterologous genes placed under the control of the polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcNPV) are often

expressed at high levels during the late stages of infection .

Recombinant baculoviruses containing the fusion protein gene are constructed using the baculovirus shuttle vector system (Luckow et al., 1993 J. *Virol.*, 67:4566-4579], sold commercially as the Bac-To-Bac lvt baculovirus expression system (Life Technologies). Stocks of recombinant viruses are prepared and expression of the recombinant protein is monitored by standard protocols (O'Reilly et al., Baculovirus Expression Vectors: A Laboratory Manual , W.H. Freeman and Company, New York, 1992; and King et al., The Baculovirus Expression System: A Laboratory Guide, Chapman & Hall, London, 1992) . Use of baculovirus or other delivery vectors in mammalian cells, such as the 'BacMam' system described by T. Kost and coworkers (see, for example Merrihew et al., 2004 *Methods MoI Biol.* 246:355-365) , or other such systems as are known to those skilled in the art are also contemplated in the instant invention.

The choice of which expression vector and ultimately to which promoter a fusion protein-encoding gene is operatively linked depends directly on the functional properties desired, e.g., the location and timing of protein expression, and the host cell to be transformed. These are well known limitations inherent in the art of constructing recombinant DNA molecules. However, a vector useful in practicing the present invention can direct the replication, and preferably also the expression (for an expression vector) of the fusion protein gene included in the DNA segment to which it is operatively linked.

Typical vectors useful for expression of genes in cells from higher plants and mammals are well known in the art and include plant vectors derived from the tumor-

inducing (Ti) plasmid of *Agrobacterium tumefaciens* described by Rogers et al. (1987) *Meth. in Enzymol.,* 153:253-277 and mammalian expression vectors pKSV-10, above, and pCI-neo (Promega Corp., #E1841, Madison, WI). However, several other expression vector systems are known to function in plants including pCaMVCN transfer control vector described by Fromm et al. (1985) *Proc. Natl. Acad. Sci. USA,* 82:58-24. Plasmid pCaMVCN (available from Pharmacia, Piscataway, NJ) includes the cauliflower mosaic virus CaMV 35S promoter.

The above plant expression systems typically provide systemic or constitutive expression of an inserted transgene. Systemic expression can be useful where most or all of a plant is used as the source of RPBLAs and their fusion proteins. However, it can be more efficacious to express RPBLAs and their fusion protein contents in a plant storage organ such as a root, seed or fruit from which the particles can be more readily isolated or ingested.

One manner of achieving storage organ expression is to use a promoter that expresses its controlled gene in one or more preselected or predetermined non-photosynthetic plant organs. Expression in one or more preselected storage organs with little or no expression in other organs such as roots, seed or fruit versus leaves or stems is referred to herein as enhanced or preferential expression. An exemplary promoter that directs expression in one or more preselected organs as compared to another organ at a ratio of at least 5:1 is defined herein as an organ-enhanced promoter. Expression in substantially only one storage organ and substantially no expression in other storage organs is referred to as organ-specific expression; i.e., a ratio of expression products in a storage organ relative to another of about 100:1 or greater indicates

organ specificity. Storage organ-specific promoters are thus members of the class of storage organ-enhanced promoters .

Exemplary plant storage organs include the roots of carrots, taro or manioc, potato tubers, and the meat of fruit such as red guava, passion fruit, mango, papaya, tomato, avocado, cherry, tangerine, mandarin, palm, melons such cantaloupe and watermelons and other fleshy fruits such as squash, cucumbers, mangos, apricots, peaches, as well as the seeds of maize (corn), soybeans, rice, oil seed rape and the like.

The CaMV 35S promoter is normally deemed to be a constitutive promoter. However, research has shown that a 21-bp region of the CaMV 35S promoter, when operatively linked into another, heterologous usual green tissue promoter, the rbcS-3A promoter, can cause the resulting chimeric promoter to become a root-enhanced promoter. That 21-bp sequence is disclosed in U.S. Patent No. 5,023,179. The chimeric rbcS-3A promoter containing the 21-bp insert of U.S. Patent No. 5,023,179 is a useful root-enhanced promoter herein.

A similar root-enhanced promoter, that includes the above 21-bp segment is the -90 to +8 region of the CAMV 35S promoter itself. U.S. Patent No. 5,110,732 discloses that that truncated CaMV 35S promoter provides enhanced expression in roots and the radical of seed, a tissue destined to become a root. That promoter is also useful herein .

Another useful root-enhanced promoter is the –1616 to -1 promoter of the oil seed rape *(Brassica napus L.)* gene disclosed in PCT/GB92/004 16 (WO 91/13922 published Sep. 19, 1991). *E. coli* DH5. alpha. harboring plasmid pRlambdaS4 and bacteriophage lambda .beta .1 that contain

this promoter were deposited at the National Collection of Industrial and Marine Bacteria, Aberdeen, GB on Mar. 8, 1990 and have accession numbers NCIMB40265 and NCIMB40266. A useful portion of this promoter can be obtained as a 1.0 kb fragment by cleavage of the plasmid with Haelll.

A preferred root-enhanced promoter is the mannopme synthase (mas) promoter present in plasmid pKan2 described by DiRita and Gelvin (1987) MoJ. *Gen. Genet,* 207:233-241. This promoter is removable from its plasmid pKan2 as a Xbal-Xball fragment.

The preferred mannopme synthase root-enhanced promoter is comprised of the core mannopme synthase (mas) promoter region up to position -138 and the mannopme synthase activator from -318 to -213, and is collectively referred to as AmasPmas. This promoter has been found to increase production in tobacco roots about 10- to about 100-fold compared to leaf expression levels.

Another root specific promoter is the about 500 bp 5' flanking sequence accompanying the hydroxyprolme-rich glycopeprotem gene, HRGPnt3, expressed during lateral root initiation and reported by Keller et al. (1989) *Genes Dev.,* 3:1639-1646. Another preferred root-specific promoter is present in the about -636 to -1 5' flanking region of the tobacco root-specific gene ToRBF reported by Yamamoto et al. (1991) *Plant Cell,* 3:371-381. The cis-acting elements regulating expression are more specifically located by those authors in the region from about -636 to about -299 5' from the transcription initiation site. Yamamoto et al. reported steady state mRNA production from the ToRBF gene in roots, but not in leaves, shoot meristems or stems.

Still another useful storage organ-specific promoter are the 5' and 3' flanking regions of the fruit-

ripening gene E8 of the tomato, *Lycopersico ʔ esculentum.* These regions and their cDNA sequences are illustrated and discussed in Deikman et al. (1988) EMBO J., 7(11):3315-3320 and (1992) *Plant Physiol.,* 100:2013-2017.

Three regions are located in the 2181 bp of the 5' flanking sequence of the gene and a 522 bp sequence 3' to the poly (A) addition site appeared to control expression of the E8 gene. One region from -2181 to -1088 is required for activation of E8 gene transcription in unripe fruit by ethylene and also contributes to transcription during ripening. Two further regions, -1088 to -863 and -409 to -263, are unable to confer ethylene responsiveness in unripe fruit but are sufficient for E8 gene expression during ripening.

The maize sucrose synthase-1 (Sh) promoter that in corn expresses its controlled enzyme at high levels in endosperm, at much reduced levels in roots and not m green tissues or pollen has been reported to express a chimeric reporter gene, β-glucuromdase (GUS), specifically in tobacco phloem cells that are abundant in stems and roots. Yang et al. (1990) *Proc . Natl. Acad. Sci ., U.S.A.,* 87:4144-4148. This promoter is thus useful for plant organs such as fleshy fruits like melons, e.g. cantaloupe, or seeds that contain endosperm and for roots that have high levels of phloem cells.

Another exemplary tissue-specific promoter is the lectin promoter, which is specific for seed tissue. The lectin protein in soybean seeds is encoded by a single gene (LeI) that is only expressed during seed maturation and accounts for about 2 to about 5 percent of total seed mRNA. The lectin gene and seed-specific promoter have been fully characterized and used to direct seed specific expression in transgenic tobacco plants. See, e.g., Vodkin et al.

(1983) Cell, 34:1023 and Lindstrom et al. (1990) *Developmental Geietics,* 11:160.

A particularly preferred tuber-specif ic expression promoter is the 5' flanking region of the potato patatin gene. Use of this promoter is described in Twell et al. (1987) *Plant MoI. Biol.,* 9:365-375. This promoter is present in an about 406 bp fragment of bacteriophage LPOTI. The LPOTI promoter has regions of over 90 percent homology with four other patatin promoters and about 95 percent homology over all 400 bases with patatin promoter PGT5. Each of these promoters is useful herein. See, also, Wenzler et al. (1989) *Plant MoI. Biol.,* 12:41-50.

Still further higher plant organ-enhanced and organ-specific promoter are disclosed in Benfey et al. (1988) *Science,* 244:174-181.

Each of the promoter sequences utilized is substantially unaffected by the amount of RPBLAs in the cell. As used herein, the term "substantially unaffected" means that the promoter is not responsive to direct feedback control (inhibition) by the RPBLAs accumulated in transformed cells or transgenic plant.

Transfection of plant cells using *Agrobacterium tumefaciens* is typically best carried out on dicotyledonous plants. Monocots are usually most readily transformed by so-called direct gene transfer of protoplasts. Direct gene transfer is usually carried out by electroportation, by polyethyleneglycol-mediated transfer or bombardment of cells by microprojectiles carrying the needed DNA. These methods of transfection are well-known in the art and need not be further discussed herein. Methods of regenerating whole plants from transfected cells and protoplasts are also well-known, as are techniques for obtaining a desired

protein from plant tissues. See, also, U.S. Patents No. 5,618,988 and 5,679,880 and the citations therein.

A transgenic plant formed using *Agrobacterium* transformation, electroportation or other methods typically contains a single gene on one chromosome. Such transgenic plants can be referred to as being heterozygous for the added gene. However, inasmuch as use of the word "heterozygous" usually implies the presence of a complementary gene at the same locus of the second chromosome of a pair of chromosomes, and there is no such gene in a plant containing one added gene as here, it is believed that a more accurate name for such a plant is an independent segregant, because the added, exogenous chimer molecule-encoding gene segregates independently during mitosis and meiosis. A transgenic plant containing an organ-enhanced promoter driving a single structural gene that encodes a contemplated HBc chimeric molecule; i.e., an independent segregant, is a preferred transgenic plant.

More preferred is a transgenic plant that is homozygous for the added structural gene; i.e., a transgenic plant that contains two added genes, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually mating (selfing) an independent segregant transgenic plant that contains a single added gene, germinating some of the seed produced and analyzing the resulting plants produced for enhanced chimer particle accumulation relative to a control (native, non-transgenic) or an independent segregant transgenic plant. A homozygous transgenic plant exhibits enhanced chimer particle accumulation as compared to both a native, non-transgenic plant and an independent segregant transgenic plant.

It is to be understood that two different transgenic plants can also be mated to produce offspring that contain two independently segregating added, exogenous (heterologous) genes Selfing of appropriate progeny can produce plants that are homozygous for both added, exogenous genes that encode a chimeric HBc molecule. Back-crossing to a parental plant and out-crossing with a non-transgenic plant are also contemplated.

A transgenic plant of this invention thus has a heterologous structural gene that encodes a contemplated chimeric HBc molecule. A preferred transgenic plant is an independent segregant for the added heterologous chimeric HBc structural gene and can transmit that gene to its progeny. A more preferred transgenic plant is homozygous for the heterologous gene, and transmits that gene to all of its offspring on sexual mating.

The expressed RPBLAs and their fusion proteins can be obtained from the expressing host cells by usual means utilized in biochemical or biological recovery. Because the RPBLAs are dense relative to the other proteins present in the host cells, the RPBLAs are particularly amenable to being collected by centrifugation of a cellular homogenate .

Thus, regions of different density are formed in the homogenate to provide a region that contains a relatively enhanced concentration of the RPBLAs and a region that contains a relatively depleted concentration of the RPBLAs. The RPBLAs-depleted region is separated from the region of relatively enhanced concentration of RPBLAs, thereby purifying said fusion protein. The region of relatively enhanced concentration of RPBLAs can thereafter be collected or can be treated with one or more reagents or subjected to one or more procedures prior to isolation of

the RPBLAs or the fusion protein therein. In some embodiments, the collected RPBLAs are used as is, without the need to isolate the fusion protein, as where the RPBLAs are used as an oral vaccine. The fusion protein containing the biologically active polypeptide can be obtained from the collected RPBLAs by dissolution of the surrounding membrane in an aqueous buffer containing a detergent and a reducing agent as discussed previously. Illustrative reducing agents include 2-mercaptoethanol , thioglycolic acid and thioglycolate salts, dithiothreitol (DTT), sulfite or bisulfite ions, followed by usual protein isolation methods. Sodium dodecyl sulfate (SDS) is the preferred detergent, although other ionic (deoxycholate, [1]N-Lauroylsarcosine, and the like), non-ionic (TweenO 20, Nonidet® P-40, octyl glucoside and the like) and zwitterionic (CHAPS, Zwittergent™ 3-X serie and the like) surfactants can be used. A minimal amount of surfactant that dissolves or disperses the fusion protein is utilized.

VACCINES AND INOCULA

In yet another embodiment of the invention, RPBLAs are used as the immunogen of an inoculum or vaccine in a human patient or suitable animal host such as a chimpanzee, mouse, rat, horse, sheep, bovine, dog, cat or the like. An inoculum can induce a B cell or T cell response (stimulation) such as production of antibodies that immunoreact with the immunogenic epitope or antigenic determinant, or T cell activation to such an epitope, whereas a vaccine provides protection against the entity from which the immunogen has been derived via one or both of a B cell or T cell response.

The RPBLAs of a contemplated vaccine or inoculum appear to act upon antigen presenting cells (APCs) such as

dendritic cells and monocytes/ macrophages that engulf the RPBLAs and process their contents. In acting upon those cell types, the RPBLAs improve the antigen delivery to antigen-presenting cells. Those RPBLAs also improve the antigen processing and presentation to antigen-presenting cells .

Thus, the invention also contemplates a vaccine or inoculum that comprises an immunogenic effective amount of recombinant protein body-like assemblies (RPBLAs) that are dissolved or dispersed in a pharmaceutically acceptable diluent. The RPBLAs contain a recombinant fusion protein recombinant fusion protein that itself contains two sequences linked together in which one sequence is a protein body-inducing sequence (PBIS) and the other is a biologically active polypeptide to which an immunological response is to be induced by said vaccine or inoculum.

T cell activation can be measured by a variety of techniques. In usual practice, a host animal is inoculated with a contemplated RPBLA vaccine or inoculum, and peripheral mononuclear blood cells (PMBC) are thereafter collected. Those PMBC are then cultured *in vitro* in the presence of the biologically active polypeptide (T cell immunogen) for a period of about three to five days. The cultured PMBC are then assayed for proliferation or secretion of a cytokine such as IL-2, GM-CSF of IFN-γ. Assays for T cell activation are well known in the art. See, for example, U. S. Patent No. 5,478,726 and the art cited therein.

Using antibody formation as exemplary, a contemplated inoculum or vaccine comprises an immunogenically effective amount of RPBLAs that are dissolved or dispersed in a pharmaceutically acceptable diluent composition that typically also contains water.

When administered to a host animal in which an immunological response to the biologically active polypeptide is to be induced by the vaccine or inoculum such as a host animal in need of immunization or in which antibodies are desired to be induced such as a mammal (e.g., a mouse, dog, goat, sheep, horse, bovine, monkey, ape, or human) or bird (e.g., a chicken, turkey, duck or goose), an inoculum induces antibodies that immunoreact with one or more antigenic determinants of the target biologically active polypeptide.

The amount of RPBLA immunogen utilized in each immunization is referred to as an immunogenically effective amount and can vary widely, depending *inter alia,* upon the RPBLA immunogen, patient immunized, and the presence of an adjuvant in the vaccine, as discussed below. Immunogenically effective amounts for a (i) vaccine and an (ii) inoculum provide the (i) protection or (ii) antibody or T cell activity, respectively, discussed hereinbefore.

Vaccines or inocula typically contain a RPBLA immunogen concentration of about 1 microgram to about 1 milligram per inoculation (unit dose), and preferably about 10 micrograms to about 50 micrograms per unit dose. The term "unit dose" as it pertains to a vaccine or inoculum of the present invention refers to physically discrete units suitable as unitary dosages for animals, each unit containing a predetermined quantity of active material calculated to individually or collectively produce the desired immunogenic effect in association with the required diluent; i.e., carrier, or vehicle.

Vaccines or inocula are typically prepared from a recovered RPBLA immunogen by dispersing the immunogen, in particulate form, in a physiologically tolerable (acceptable) diluent vehicle such as water, saline,

phosphate-buffered saline (PBS), acetate-buffered saline (ABS), Ringer's solution, or the like to form an aqueous composition. The diluent vehicle can also include oleaginous materials such as peanut oil, squalane, or squalene as is discussed hereinafter.

The preparation of inocula and vaccines that contain proteinaceous materials as active ingredients is also well understood in the art. Typically, such inocula or vaccines are prepared as parenterals, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified, which is particularly preferred.

The immunogenically active RPBLAs are often mixed with excipients that are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, an inoculum or vaccine can contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents that enhance the immunogenic effectiveness of the composition.

The word "antigen" has been used historically to designate an entity that is bound by an antibody or receptor, and also to designate the entity that induces the production of the antibody. More current usage limits the meaning of antigen to that entity bound by an antibody or receptor, whereas the word "immunogen" is used for the entity that induces antibody production or binds to the receptor. Where an entity discussed herein is both immunogenic and antigenic, reference to it as either an immunogen or antigen is typically made according to its intended utility.

"Antigenic determinant" refers to the actual structural portion of the antigen that is immunologically bound by an antibody combining site or T-cell receptor. The term is also used interchangeably with "epitope".

As used herein, the term "fusion protein" designates a polypeptide that contains at least two amino acid residue sequences not normally found linked together in nature that are operatively linked together end-to-end (head-to-tail) by a peptide bond between their respective carboxy- and ammo-terminal amino acid residues. The fusion proteins of the present invention are chimers of a protein body-inducing sequence (PBIS) linked to a second sequence that is a biologically active polypeptide product (e.g., peptide or protein) of interest (target) .

Without further elaboration, it is believed that one skilled in the art can, using the preceding description and the detailed examples below, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limiting of the remainder of the disclosure in any way whatsoever.

Example 1: Accumulation of RX3-ECFP derived fusion proteins in dense fractions of transfected mammal cells

The polynucleotide sequence coding for the N-terminal gamma-zem coding sequence RX3 (WO2004003207) was fused directly, or through a linker consisting of five glycines, to the 5' end of the sequence encoding ECFP, a cyan fluorescent variant of GFP. Those constructs (Fig. IA) that code for the fusion proteins RX3-ECFP or RX3-Gx5-ECFP were introduced in CHO mammal cultured cells by the

Lipofectamine-based transfection method (Invitrogen). CHO cells transfected with plasmid pECFP-Nl (Clontech) containing the gene sequence of the cytosolic ECFP, were used as controls.

Transfected mammalian cell extracts were loaded on density step gradients and centrifuged. The accumulation of recombinant proteins in the different fractions was analyzed by immunoblot (Fig. 2A). The results shown in that figure indicate that RX3-ECFP and RX3-Gx5-ECFP appeared in fractions F42, F56 and P corresponding to dense RPBLAs, (Fig. 2A, lanes 3-5). This result demonstrates that the fusion proteins are able to assemble and induce the RPBLA formation. Some fusion protein was also detected in the supernatant fraction (Fig. 2A, lane 1), probably representing fusion proteins from the RPBLAs solubilized partially during the extraction process, or fusion proteins just synthesized that had not assembled.

Contrapπly, when mammalian cell extract transfected with the control plasmid pECFP-Nl was loaded on the same density step gradients, the ECFP protein was observed exclusively in the supernatant. No traces of ECFP were detected in the dense fractions indicating that the ECFP by itself is not able to aggregate and form PB like structures.

Example 2: Accumulation of active ECFP fused to PBIS domains in RPBLAs of transfected mammal cells

To determine if the fusion proteins RX3-ECFP, RX3-Gx5-ECFP and 22aZ-ECFP are active inside the RPBLAs, confocal microscopic analyses were performed in CHO cells transfected with the constructs that code for them (Fig

IA). Cyan fluorescent images were collected at 458 nm e>citation with the argon ion laser by using an emission window set at 470-530 nm. As shown in Fig. 3, the corresponding fusion proteins, RX3-ECFP (Fig. 3A) and RX3-Gx5-ECFP (Fig. 3B) and 22aZ-ECFP (Fig. 3D), were detected in the endoplasmic reticulum, indicating that the gamma-zem and the alpha-zein signal peptide is functional in mammal cells where it mediates the translocation of the fusion protein into the ER.

It is important to note that the fusion proteins surprisingly appear preferentially accumulated large and dense spherical structures that strongly resembled both natural PBs of cereal seed and RPBLAs in the heterologous systems visualised by immunodetection. The intense fluorescence observed in these structures indicates that the fusion protein remains properly folded, and therefore active, in spite of being highly packaged inside the RPBLAs. It is also important to note that RX3 domains, as well as other protein body inducing sequences (PBIS) responsible for the formation of PBs and PB-like structures contain multiple cysteines residues. Although it might be predicted that such cysteines could form disulfide bonds with target protein cysteines and hence interfere with the proper folding of the target proteins this was not observed to be the case. Both active target protein (ECFP fluorescence) and functional PBIS (formation of RPBLAs) were observed.

As a control, the construct pECFP-Nl was used to transfect CHO cells. The expression of a cytosolic ECFP showed a homogeneous fluorescence pattern all along the cell, including the nucleus (Fig. 3C).

Example 3: Subcellular localization of other fluorescent proteins fused to RX3 in CHO cells.

The sub-cellular localization of RX3-DsRED and RX3-GFP fusion proteins in transiently transfected CHO cells was analyzed by confocal microscopy to analyze whether other fluorescent proteins than ECFP fused to RX3 are properly folded inside the RPBLAs and bioactive. It is important to note that DsRED shares no homology to ECFP, which implies a completely different folding mechanism. Micrographs from the transfected cells were obtained by using a confocal laser scanning microscope (Leica TCS SP, Heidelberg, Germany) fitted with spectrophotometers for emission band wavelength selection. Green fluorescent images were collected at 488 nm excitation with the Argon ion laser by using an emission window set at 495-535 nm. Red fluorescent images were collected after 543 nm excitation with a HeNe laser and emission window 550-600. Optical sections were 0.5 µm thick.

The expression of RX3-GFP (Fig. 3E) and RX3-DsRED (Fig. 3F) fusion proteins in CHO cells produced a large amount of highly fluorescent round-shaped RPBLAs. These results confirm that both fusion proteins are properly folded and active inside the RPBLAs.

Example 4: Subcellular localization of fluorescent RX3 fusion proteins in plants and insects

In order to analyze whether host cells other than CHO cells can produce RPBLAs containing active fluorescent proteins fused to RX3 domains, tobacco plants were transiently transformed with RX3-GFP by syringe

agromf iltration . The analysis by coπfocal microscopy of the epidermal cells (Figs. 4A and 4B) showed the presence of a large amount of fluorescent RPBLAs. Similar results were obtained when transformed tobacco mesophyll cells were analyzed.

Similar results were obtained when *Spodoptera* SF9 insect cells or insect larvae *{Tπ choplusia ni)* were infected with baculovirus coding for the fusion protein RX3-DsRED. As shown in Fig. 4C the projection of optical sections of infected insect cells accumulated a large amount of fluorescent RPBLAs about 0.5 micrometers in diamerer containing the active RX3-DsRED fusion protein. Confocal analysis of infected larvae also showed an impressive amount of fluorescent RPBLAs in whatever tissue analyzed. In Fig. 4D, fat cells from infected larvae show RPBLAs containing active RX3-DsRED. Interestingly, DsRED fluorescence was not observed in insect haemolymph, suggesting that the expressed protein remained sequestered entirely within RPBLAs.

Example 5: Activity of RX3-hGH assembled in RPBLAs in CHO cells .

Studies were undertaken to determine the activity of human growth hormone (hGH) produced in RPBLAs. The hGH was chosen because this molecule contains 2 disulphide bonds that are important for the proper folding of the protein. The RX3 domain also contains cysteine residues involved in disulphide bonds that are essential for the assembly and stabilization of the RPBLAs, which could interfere in the proper folding of the hGH.

The p3.1-RX3-hGH construct was introduced into CHO cells by transient transfection with the lipofectamine protocol (Invitrogen) . Four days after transfection the

cells were fi,ed, permebealized and incubated with anti-RX3 or an anti-hGH antiserum (Figs. 5A and 5B, respectively) and the secondary antibody conjugated to Ale/a Fluor 488 (Invitrogen) . The presence of large RPBLAs (1-3 micrometers) containing the RX3-hGH fusion protein was observed by optical microscopy analysis independently of the primary antibody used.

In order to corroborate that the RPBLAs were dense organelles as was described previously, CHO cells expressing RX3-hGH were homogenized, and the homogenates loaded on a density step gradient and centrifuged as described elsewhere. The accumulation of RX3-hGH in the different fractions was analyzed by immunoblot . As can be seen, part of the fusion protein was present in the supernatant, representing non-assembled RX3-hGH, but most of the fusion protein was detected in fraction F56 corresponding to dense RPBLAs (Fig. 2B, lanes 2 and 5, respectively) .

This F56 fraction was diluted 3-fold in buffer PBP4 (100 mM Tris pH7.5, 50 mM KCl, 5 mM MgCl2, 5 mM EDTA) and centrifuged at 80000xg in a swinging-bucket to recover the RPBLAs in the pellet. The presence of hGH was quantified using an ELISA assay (Active© Human Growth Hormone ELISA - DSL-10-1 900 ; Diagnostic Systems Laboratories, Inc) , which was able to detect the hGH even in the presence of the intact RPBLA membrane.

This same sample was applied to a bioactivity assay (Active® Bioactive Human Growth Hormone ELISA - DSL-10-11100; Diagnostic Systems Laboratories, Inc) . This bioactivity assay is based on the capacity of properly folded hGH to interact to a hGH binding protein provided by the kit, this interaction being dependent on a functional conformation of the hGH. The sample gave a positive result

at 24 ng/ml of bioactive protein. The hGH proteins evidently were correctly folded and presented on the outer surface of the dense RPBLAs. Removal of the membrane surrounding the RPBLAs by washing the preparation with 50 mM Tris pH 8 and 1% Triton X-IOO and by sonicating at 50% amplitude and 50% cycle for 1 minute, repeated times 5 (Ikasomc U200S - IKA Labortechnik) resulted in greater specific activity (45 ng/ml) due to the exposure of additional hGH molecules on the surface of the aggregates.

In determining the activity of hGH fused to RX3, the fusion protein was solubilized from RPBLAs isolated by density gradient (F56, diluted 3-fold in buffer PBP4 and centrifuged at 80000xg in a swmging-bucket for 2 hours) . The fusion protein was solubilized in buffer S (Tris 50 mM, pH8 and 2% of β-ME) and sonicated (Clyde 5, Amplitude 50%, 1 minute, repeated five times; Ikasomc U200S - IKA Labortechnik) . After incubation at 37 ^{0}C for 2 hours, the sample was centrifuged at 5000xg for 10 minutes, and the supernatant containing the soluble RX3-hGH fusion protein was assayed to quantify and assess the bioactive component of the fraction. The amount of fusion protein in the supernatant was determined to be 250 ng/mL by ELISA (Active® Human Growth Hormone ELISA - DSL-10-1900; Diagnostic Systems Laboratories, Inc) . The protein assayed in the bioactivity ELISA assay (Active© Bioactive Human Growth Hormone ELISA - DSL-10-11100; Diagnostic Systems Laboratories, Inc) gave a result of 70 ng/ml, indicating that about 30 % of the RX3-hGH fusion protein is active. The loss of hGH activity could be a consequence of the high concentration of reducing agent used in the solubilization, or due to some impairing effect of the RX3 domain over the hGH or the hGH binding protein.

Finally, the RX3-hGH fusion protein was cleaved by a site specific protease to liberate the hGH from the fusion protein. The solubilized RX3-hGH fusion protein was diluted 2-fold and the digestion was performed with EKmay as described by the manufacturer (Invitrogen). After that, free hGH was isolated from the uncleaved fusion protein (insoluble) by centrifugation at lβOOOxg at 4°C for 1 hour. The soluble hGH was recovered from the supernatant and applied to the quantification and bioactivity assays from Diagnostic Systems Laboratories. Surprisingly, the results from both these kits gave the same value of 90 ng/ml for the quantification and bioactivity ELISA assays (Active® Human Growth Hormone ELISA - DSL-10-1900; Diagnostic Systems Laboratories, Inc) and Active® Bioactive Human Growth Hormone ELISA - DSL-10-11100; Diagnostic Systems Laboratories, Inc) indicating that all the protein present as detected by the quantification kit is also determined to be bioactive.

Summary table for the quantification and bioactivity of the hGH protein in all the formulations is presented below:

	Quantification	Bioactivity
Formulation	Amount ng/ml	Amount ng/ml
Intact RPBLAs	14	25
Membrane removed RPBLAs	35	45
Soluble RX3-hGH	250	70
Cleaved hGH	90	90

It is important to note that CHO cells stably transfected with the vector p3.1-RX3 were used as a

negative control. As shown in Fig. 2B, the expression of RX3 in CHO cells also accumulates in dense structures which can be isolated by density step gradient in F56 (Fig. 2B, lane 5). Moreover, optical analysis of CHO cells transfected with p3.1-RX3, showed that the RX3 protein accumulate in RPBLAs (Fig. 5C) These control RX3 RPBLA preparations and isolated RX3 protein showed no hGH activity in the ELISA bioactivity assay.

Example 6: Activity of DNAb intern after RX3-Int-hGH solubilization from RPBLAs from CHO cells

The polynucleotide sequence coding for the Ssp DNAb intern (New England Biolabs) was fused in frame to the 3' end of the RX3 sequence (WO2004003207), and to the 5' end of the hGH cDNA. The resulting construct was cloned into vector pcDNA3.1(-) [Fig. IA] to form vector p3.1-RX3-I-hGH. As a negative control, an inactive version of the same intern was produced by PCR where the ammo acid residue Aspl54 was mutated to Ala [Fig. IA] to form vector p3.l-RX3-Im-hGH . The Aspl54 amino acid residue has been reported to be essential for the Ssp DNAb self-cleavage activity (Mathys et al, *GENE* (1999) 231:1-13) .

Immunochemical analysis of CHO cells transfected with p3.l-RX3-I-hGH using anti-hGH antiserum revealed that the fusion protein RX3-Int-hGH accumulated in big round-shaped RPBLA, similar to the ones observed in CHO cells expressing RX3-hGH (compare Figs. 5B and 5D). This result indicates that the fusion protein containing the DNAb intern self-assembles and accumulates in the high density structures .

CHO cells transfected with p3.l-RX3-I-hGH were homogenized, the homogenates were loaded in density step gradients, and the fractions corresponding to the different densities were analyzed by immunoblot . Most of the RX3-I-hGH was detected in the fraction F56 corresponding to dense RPBLAs (Fig. 2B) . As for other RX3 fusion proteins, the presence of RX3-I-hGH fusion protein in the supernatant probably represents the un-assembled fusion protein contained in the ER and solubilized during the homogenization process.

Once it was demonstrated that the RX3-I-hGH accumulated in RPBLAs, these ER-derived organelles were isolated by low speed centrif ugation as described elsewhere herein. The centrif ugation of homogenates of CHO cells transfected with p3.l-RX3-I-hGH at 1500xg for 10 minutes permitted the separation of the non-assembled RX3-Int-hGH fusion proteins in the supernatant from the assembled in RPBLAs in the pellet. Equivalent studies were performed with CHO cells expressing the inactive RX3-mInt-hGH fusion protein.

The pellets containing the assembled RX3-Int-hGH and RX3-mInt-hGH fusion proteins were solubilized in Sl buffer (20 mM Tris pH7, 200 mM NaCl, 1 mM EDTA, 0.1% SDS and 0.1 mM TCEP) at 37°C for 2 hours, and the intern enzymatic activity was induced by incubation at 25°C for 48 hours after dialysis against the cleavage induction buffer: 20 mM Tris pH 7, 200 mM NaCl, 1 mM EDTA. After induction of intern self-cleavage, the composition was centrifuged at 16000xg for 10 minutes and the supernatant and the pellet analyzed by immunoblot using anti-RX3 and anti-hGH antiserum.

Both fusion proteins were solubilized, but only the fusion protein containing the active intern (RX3-Int-

hGH) was able to self-cleave (Figs. 6A and 6B, black arrowheads). The absence of self-cleavage of the mutated RX3-mInt-hGH fusion protein demonstrates that the self-cleavage observed with the RX3-Int-hGH is due to the specific activity of the intern, and not due to some endogenous protease activity co-purified during the RPBLAs isolation process.

To optimize the efficiency of intern self-cleavage, alternative solubilization protocols were assayed. The intern self-cleavage of the RX3-Int-hGH can be compared, after solubilization with the S1 buffer and the biphasic extraction protocol (S2) described elsewhere (Fig. 6C). From the ratio between the remaining of the full-length fusion protein and the appearance of the band corresponding to the liberated hGH, even though the biphasic extraction protocol was the more efficient permitting more than 50% of cleavage, it can be concluded that in both cases a large proportion of DNAb intern was active and able to self-cleave.

Example 7: Activity of RX3-EGF assembled in RPBLAs in tobacco plants

RPBLAs from transgenic tobacco plants expressing the RX3-EGF fusion protein were isolated by low speed centrifugation essentially as described in US11/289,264. The fusion protein was solubilized by sonication (Cycle 5, Amplitude 50%, 1 minute, repeated five times; Ikasonic U200S - IKA Labortechnik) in 50 mM Tris pH 8 and 2% of β-ME and incubation at 37°C for 2 hours. Afterwards, the solubilized material was centrifuged at 16000xg at 4°C for 30 minutes to discard the unsolubilized fusion protein in the pellet. The supernatant was dialyzed against 50 mM

Tris pH 8 to remove the β-ME, centrifuged once again at 16000xg at 4°C for 30 minutes, and the supernatant quantified by the hEGF kit from Biosource International Inc. (KHG0062) .

The bioactivity of EGF was analyzed by determining the proliferation rate (radioactive thymidine incorporation to DNA) of MDA-MB231 cells (breast cancer cells that overexpress EGF receptor) incubated with 1.2 ng/mL of RX3-EGF fusion protein. As a positive control, MDA-MB231 cells were incubated with 10 ng/mL of commercial EGF (Promega) or fetal calf serum (FCS) . The results, summarized in the following table are represented as percentage (%) of proliferation with regard to the basal proliferation rate of MB231 cells (100%), determined as the proliferation rate of these cells cultivated in the absence of EGF (deprived) .

Proliferation of MDA-MB231 cells

		% proliferation with respect to Deprived cells	
Sample	Concentration	Mean	STD
Deprived	-	100	-
FCS	-	145	1,27
EGF (Promega)	10 ng/mL	158	11,7
RX3-EGF	1,2 ng/mL	146	4

As expected, the supplementation of MB231 cell culture with commercial EGF (Promega) or the FCS produced a significant increase of the proliferation rate (158% and

145%, respectively) . Unexpectedly, the addition of 1.2 ng/mL of RX3-EGF also produced an increase of 146% of the proliferation rate. It is important to note that almost the same proliferation rate was observed with 10-fold more concentration of commercial EGF than with RX3-EGF. This surprising result could be explained by previous results showing that saturation of the proliferation rate of MB231 cell was observed at 5 ng/mL of the commercial EGF. Another possible explanation could be a more active conformation of EGF when fused to RX3. In any case, this result shows that RX3-EGF is at least as active as the commercial EGF (Promega).

Example 8: Activity of RX3-GUS assembled in RPBLAs in CHO cells .

The β-glucuronidase enzyme (GUS) is a broadly used reporter protein (Gilisen et al., *Transgenic Res.* (1998) 7(3):157-163) . The expression of an active RX3-GUS fusion protein in RPBLAs was a challenge, mainly by the presence of 9 cysteine ammo acid residues, and also because it is a large protein (about 70 kDa) .

The polynucleotide sequence coding for RX3 (WO2004003207) was fused in frame to the 5' end of the sequence encoding GUS (Fig. IA. RX3-GUS), and the resulting construct used to transfect CHO cells as described m Example 7 .

Immunochemical analysis of CHO cells transfected with p3.1-RX3-GUS incubated with anti-RX3 antiserum revealed the presence of large RPBLAs (Figure 5E). To verify the density of those RPBLAs, CHO cells transfected with the same plasmid were homogenized and afterwards loaded onto step density gradients. The analysis of the

different fractions by immunoblot showed that the fusion proteins localized in the higher dense fractions (Fig. 2B. F56), indicating that the RX3-GUS fusion proteins are able to assemble and accumulate in dense RPBLAs. It is important to note that no fusion protein was detected in the supernatant, meaning that almost all RX3-GUS is assembled in dense structures (RPBLAs).

Once it was demonstrated that the RX3-GUS accumulated in RPBLAs, the fusion protein was recovered from the F56 fraction (as described in Example 5 for RX3-hGH) and solubilized in 50 mM Tris, pH 8, β-ME 2% and SDS 0.1% at 37 ^{0}C for 2 hours. Afterwards, the solubilized material was centrifuged at 16000xg at room temperature for 10 minutes, and the supernatant containing the soluble disassembled RX3-GUS fusion protein was dialyzed at 4°C against a 50 mM Tris pH 8 solution over night (about 18 hours).

GUS activity test is based in the catalysis of metilumbelif eril-β-glucuronide acid (MUG) to the 4-metilumbelif erone (4-MU) fluorescent product, by the GUS enzyme (Jefferson et al. 1987 *EMBO J.* 6(13):3901-3907). Fifty µL of the solubilized RX3-GUS fusion protein (around 0.25 ng of RX3-GUS/ µL) were incubated in the presence of MUG at room temperature, and the appearance of 4-MU was carried out in a fluorimeter (excitation wavelength 355 nm; emission wavelength 420 nm). To rule out the possibility of measuring endogenous GUS-like activity present in the RPBLAs preparation from CHO cells, RPBLAs from CHO cells transfected with p3.1-RX3 were isolated, and once the RX3 protein was solubilized, this sample was included in the activity test as a control. The table below summarizes the results obtained:

Absorbance at 420 nrn

	RX3-GUS		RX3	
Time (minutes)	Mean	STD	Mean	STD
0	337	24	227	6 4
30	534	4 2	236	15
60	690	12 7	265	9 2
90	909	30 4	299	21 2
120	1049	38 9	309	10 6
160	1141	21 9	311	82

From the results shown in this table, it is clear that the RX3-GUS fusion protein remains active once solubilized from the RPBLAs. The specific activity of the RX3-GUS calculated from these experiments was 0.2 pmol of 4-MU/ min-1 * 12.5 ng-1 of RX3-GUS. No significant endogenous GUS-like activity was observed when the RX3 preparation was analyzed.

Example 9: Activity of RX3-EK assembled in RPBLAs in CHO cells

Bos taurus enterokmase (enteropeptidase) is a membrane-bound serine protease of the duodenal mucosa, involved in the processing of the trypsinogen to trypsin (DDDKi) with a chymotrypsin-like serine protease domain. The enteropeptidase is a disulfide linked two-chain peptide formed by the heavy chain (EK_HC - 120 kD) and the catalytic light chain (EK_{LC} - 47 kD). The catalytic subunit (here referred as EK) is almost as active and specific by itself as the whole holoenzyme (LaVallie et al. 1993 *J. Biol. Chem.* 268 (31) :23311-23317). It is important to point out that bovine EK has 4 disulphide bonds. Moreover, the N-terminal end of the protein is folded inside the protein, and it is essential for the proper folding of a functional EK. These two EK requirements make EK protein a

challenging protein to be expressed as an active protein in RPBLAs.

The polynucleotide sequence coding for RX3 (WO2004003207) was fused through a linker comprising the FXa cleavage site (IEGR) to the 5' end of the EK sequence, and cloned in pcDNA3.1(-) (Fig. IA, p3.1-RX3-EK) .

This construct was used in CHO cells transfection by the lipofectamme method (Invitrogen) . Immunochemistry analysis of those transfected cells with anti-RX3 antiserum revealed the presence of a large quantity of small RPBLAs. These organelles were to be seen all along the cytoplasm of the transfected cells, but the size usually did not exceed 0.5 micrometers (Fig. 5F) .

To verify the density of those small RPBLAs, CHO cells transfected with the same plasmid were homogenized and loaded in step density gradients. The RX3-EK fusion protein was localized in F56 fraction (Fig. 2B). The high density of the RX3-EK fusion protein assemblies suggests that this fusion protein accumulates in dense RPBLAs. It is important to note that no fusion protein was detected in the supernatant, meaning that almost all RX3-EK is assembled in dense structures (RPBLAs). Interestingly, the molecular weight of the RX3-EK fusion protein was estimated at 58 KDa, about 15 KDa higher than the theoretical molecular weight. This result suggests that the EK in the RPBLAs is highly glycosylated, as has been described for the natural protein (LaVaIlie et al., 1993 *J. Biol. Chem.* **268(31)** :23311-23317) .

The fusion protein was recovered from the F56 fraction (as described in Example 5 for RX3-hGH) and solubilized in 50 mM Tris, pH 8, β-ME 2% and SDS 0.1% at 37 ^{0}C for 2 hours. To increase the solubilization, the sample was sonicated at 50% amplitude and 50% cycle for 1 minute,

repeated 5 times (Ikasonic U200S - IKA Labortechnik), before SDS was added. Afterwards, the sample was centrifuged at 5000xg at room temperature for 10 minutes, and the supernatant containing the soluble disassembled RX3-EK fusion protein was dialyzed at 4°C against a 50 mM Tris pH 8 solution over night (about 18 hours). Then, the fusion protein was digested by FXa as described by the manufacturer (Quiagen), and the EK activity was measured by fluorimetric assay (Grant, et al., 1979 *Biochim. Biophys. Acta* 567:207-215). The liberated EK from the RX3-EK had enteropeptidase activity.

Example 10: Activity of RX3-Casp2 and RX3-Casp3 assembled in RPBLAs in CHO cells

Studies were undertaken to determine the activity of caspases produced in RPBLAs.

Caspases are a family of cysteine proteases that cleave with high specificity after an aspartic acid of a consensus sequence. They are the main executioners of the highly regulated process of apoptosis.

Caspases exist as inactive procaspases with a prodomam of variable length followed by a large subunit (p20) and a small subunit (plO). They are activated through proteolysis and mature active caspase consists of the heterotetramer $p20_2$-$pl0_2$ (Lavπ k et al., 2005 *J. CIm. Invest.* 115:2665-2671). Caspases are divided into initiator caspases and executioner caspases that differ in their mechanism of action. Caspase2 (initiator caspase) and caspase3 (executioner caspase) have been chosen as an example of proteins which are active in the RPBLAs (Baliga et al., 2004 *Cell Death and Differentiation* 11:1234-1241; Feeney et al., 2006 *Protein Expression and Purification*

47(1):311-318). Those proteins are especially challenging because they are synthesized as zymogens that, to become active, need to be self-cleaved and to form the heterotetramer.

The p3.1-RX3-C2 and p3.1-RX3-C3 constructs (Fig. 1) were introduced into CHO cells by transient transfection with the lipofectamine protocol (Invitrogen). Four days after transfection, to determine if caspases are accumulated in dense RPBLAs organelles, CHO cells expressing RX3-Casp2 or RX3-Casp2 were homogenized, loaded on a density step gradient and centrifuged as described elsewhere.

The accumulation of both RX3-caspases fusion proteins in the different fractions was analyzed by immunoblot (Fig. 2B). As it can be seen, most of the RX3-Casp2 or RX3-Casp2 fusion proteins sediment to fraction F56 and F42 corresponding to dense RPBLAs. This result indicates that these two fusion proteins are able to tightly assemble in dense structures.

In the immunoblot presented in Fig. 2B, only the full length fusion protein was shown, but bands of different molecular weight are present in this fraction. These bands being reactive to either anti-RX3 antibody or anti-CASP (SA-320 and SA-325, Biomol International) antibody correspond to the different Caspase subunits, indicating that autocatalytic activation has taken place inside RPBLAs. These observations indicate that Caspase2 and Caspase 3 are active *m vivo.*

The F56 and F42 fractions were diluted 4-fold in buffer PBP4 and centrifuged at 80000xg in a swinging-bucket to recover the RPBLAs in the pellet. The ER membrane surrounding this organelle was removed by washing the RPBLAs preparation with 50 mM Tris pH 8 and 1% Triton X-

100. Upon removal of the ER membrane, activity of caspase is assayed using the BIOMOL QuantiZymeTM Assay System, CASPASE-3 Cellular Activity Assay Kit PLUS-AK703 (caspase 3) and BIOMOL QuantiZymeTM Assay System, CASPASE-2 Cellular Activity Assay Kit PLUS-AK702 (caspase 2). This kit measures caspase activity colorimetrically with a specific substrate. The RX3-Casp2 and the RX3-Casp3 RPBLAs show Caspase activity.

In determining the activity of Caspases fused to RX3, the fusion protein is solubilized from RPBLAs isolated by density gradient (F56 and F42, diluted 4-fold in buffer PBP4 and centrifuged at 80000xg in a swinging-bucket). The fusion protein is solubilized in buffer CA (50 mM Hepes, pH 7.4, 100 mM NaCl, 1 mM EDTA, 100 mM DTT, 1 % CHAPS, 10 % glycerol) after sonication (50% amplitude and 50% cycle for 30 seconds, 5 times) . Solubilization is performed by a 2-hour incubation at 37°C and insoluble material is discarded by centrif ugation at lβOOOxg for 10 minutes. The supernatant containing the soluble RX3-casp fusion protein is dialyzed against caspase kit assay buffer (50 mM Hepes, pH 7.4, 100 mM NaCl, 1 mM EDTA, 10 mM DTT, 0.1 % CHAPS, 10 % glycerol) . Activity of the dialyzed sample containing RX3-Casp2 and RX3-Casp3 are assessed with the BIOMOL QuantiZymeTM Assay System, CASPASE-3 Cellular Activity Assay Kit PLUS-AK703 (caspase 3) and BIOMOL QuantiZymeTM Assay System, CASPASE-2 Cellular Activity Assay Kit PLUS-AK702 (caspase 2). Caspase 2 and Caspase 3 are active.

Example 11: Activity of RX3-RTB assembled m RPBLAs in agroinf iltrated tobacco plants

The polynucleotide sequence coding for RTB (Reed et al., 2005 *Plant Cell Report* 24:15-24) was fused in frame

to the 3' end of RX3 domain and cloned in a binary vector (pB-RX3-RTB) .

This construct was used in tobacco plants transformed by syringe agroinf lltration, as described elsewhere. The agroinf iltrated tobacco leaves were homogenized and loaded in step density gradients. The RX3-RTB fusion protein was localized in fractions F42 and F56 (Fig. 2B), suggesting that the fusion protein self-assembles and accumulates in dense RPBLAs. As described for RX3-EK, the RX3-RTB fusion protein isolated from the RPBLAs has a lower electrophoretic mobility compared to the theoretical molecular weight. This results supports that RTB can be glycosylated in RPBLAs.

The fusion protein was recovered from those dense fractions (as described in Example 5 for RX3-hGH) and solubilized in 50 mM Tris, pH 8, β-ME 0.8% at 37 ^{0}C for 2 hours. To increase the solubilization, the sample was sonicated at 50% amplitude and 50% cycle for 1 minute, repeated 5 times (Ikasonic U200S - IKA Labortechmk) . Afterwards, the sample was centrifuged at 5000xg at room temperature for 10 minutes, and the supernatant containing the soluble disassembled RX3-RTB was analyzed by ELISA for binding to the glycoprotein fetuin treated with sialydase to expose galactose-terminated glycans. The RX3-RTB binds to it.

Example 12: Plasmid construction for plant transformation

The coding sequences of human epidermal growth factor (hEGF) were obtained synthetically and were modified in order to optimize its codon usage for expression in plants .

hEGF protein (SEQ ID NO:41)

hEGF DNA (SEQ ID NO: 42)

The synthetic gene encoding the 53 amino acids of active hEGF was obtained by primer overlap extension PCR method, using 4 oligonucleotides of around 60 bases, with 20 overlapping bases. The synthetic hEGF cDNA included a 5' linker sequence corresponding to the Factor Xa specific cleavage site. The oligonucleotides were purified by polyacrylamide denaturing gel.

Synthetic hEGF cDNA was purified from agarose gel (Amersham) and cloned into pGEM vector (Promega) . The RX3 cDNA fragment (coding for an N-terminal domain of gamma-zein) containing cohesive ends of BspHI and Ncol, was inserted into the vector pCKGFPS65C (Reichel et al., 1996 *Proc. Natl. Acad. Sci. USA* 93:5888-5893) previously digested with Ncol (as described in patent application WO2004003207) . The sequence coding for EGF was fused in frame to the RX3 sequence. The constructs RX3-EGF was prepared by substitution of the GFP coding sequence for the EGF synthetic gene.

The resulting construct named pCRX3EGF contained a nucleic acid sequence that directs transcription of a protein as the enhanced 35S promoter, a translation enhancer as the tobacco etch virus (TEV) , the EGF coding

sequence and the 3' polyadenylation sequences from the cauliflower mosaic virus (CaMV). Effective plant transformation vector pl9RX3EGF was ultimately obtained by inserting the Hindi II/HindIII expression cassettes into the binary vector pBinl9 (Bevan, 1984 *Nucleic Acids Research* 12:8711-8721).

The cDNA encoding the alpha-zein of 22 kD (22aZ) and the rice prolamm of 13 kD (rP13) were amplified by RT-PCR from a cDNA library from maize W64A and Senia rice cultivar, respectively. The oligonucleotides used in the PCR reaction were:

22aZ-5' (SEQ ID NO: 43)

22aZ-3' (SEQ ID NO:44)

Ricel3Prol-5' (SEQ ID NO:45)

Ricel3Prol-3' (SEQ ID NO:46)

The corresponding PCR fragments were cloned in the pCRII vector (Invitrogen), sequenced and cloned in pUC18 vectors containing the enhanced CaMV 35S promoter, the TEV sequence and 3' ocs terminator. The pCRII-rP13 was digested by Sail and Ncol, and cloned in the pUC18RX3Ct, pUC18RX3hGH and pUC18RX3EGF plasmids digested by the same enzymes to obtain plasmid pUC18rP13EGF. The pCRII-22aZ was digested by Sall/Ncol and cloned in the pUC18RX3EGF plasmid digested by the same enzymes to obtain plasmid pUC1822aZtEGF. Finally, the pUC18-derived vector was cloned in pCambia 5300 by Hindlll/EcoRI .

The construct pBIN m-gfp4-ER, contain an optimized GFP for expression in plants (Haseloff et al., 1997 *Proc. Natl. Acad. Sci. USA* 94:2122-2127) . This construct was used as template for PCR amplification of the GFP. The oligonucleotides were designed to eliminate the signal peptide and HDEL motif present in the original sequence as well as to introduce the restriction sites for further cloning.

Primers:

GFP 5 (SEQ ID NO: 50)

GFP 3 (SEQ ID NO: 51)

The PCR product was cloned in a PCR cloning vector (PCROH Vector, Invitrogen)) and the sequence verified. The GFP fragment containing cohesive ends Rcal/BamHI was cloned into pUC18RX3hGH (US2006123509 (Al)), giving the cassette RX3-GFP in a pUC18 vector. This cassette was liberated by Hmdlll/BamHI digestion and subsequently inserted in a pCAMBIA 2300 vector (pB-RX3-GFP)

The RTB clone (GenBank accession no. X03179) was amplified by PCR (RTB5 and RTB3) and digested by Rcal/Smal. The digested PCR fragment was cloned in pUC18RX3hGH (US2006123509 (Al)) digested by Ncol/Smal to obtain pUC18RX3RTB. Then, this vector was digested by Hindlll/EcoRI and the liberated fragment cloned in a pCAMBIA 2300 vector digested by the same restriction enzymes (pB-RX3-RTB)

Primers :

RTB5 (SEQ ID NO:52)

RTB3 (SEQ ID NO:53)

Plant material

Tobacco *{Nicotiana tabacum var.* Wisconsin) plants were grown in an *in vitro* growth chamber at 24-26°C with a 16 hour photope π od. Adult plants were grown in greenhouse between at 18-28°C, humidity was maintained between 55 and 65% with average photoperiod of 16 hours.

Plantlets for agroinf iltration (Vaquero et al., 1999 *Proc. Natl. Acad. Sci., USA* 96(20):11128-11133; Kapila et al., 1997 *Plant Sci.* 122:101-108) method were grown from seeds for 4-6 weeks in the *m vitro* conditions described above.

Tobacco stable transformation

The binary vectors were transferred into LBA4404 strain of *A. tumefaciens.* Tobacco *(Nicotiana tobaccum, W38)* leaf discs were transformed as described by Draper and Hamil 1988, In: Plant Genetic Transformation and Gene Expression. A Laboratory Manual (Eds. Draper, J., Scott, R., Armitage, P. and Walden, R.), Blackwell Scientific Publications. Regenerated plants were selected on medium containing 200 mg/L kanamycm and transferred to a greenhouse. Transgenic tobacco plants having the highest transgene product levels were cultivated in order to obtain Tl and T2 generations.

Recombinant protein levels were detected by immunoblot. Total protein extracts from tobacco leaves were quantified by Bradford assay, separated onto 15% SDS-PAGE and transferred to nitrocellulose membranes using a Mini Trans-Blot Electrophoretic Transfer Cell (Bio Rad). Membranes were incubated with gamma-zein antiserum (dilution 1/7000) (Ludevid et al. 1985, *Plant Science*

41:41-48) and were then incubated with horseradish peroxidase-conjugated antibodies (dilution 1/10000, Amersham Pharmacia) . Immunoreactive bands were detected by enhanced chemiluminescence (ECL western blotting system, Amersham Pharmacia).

Tobacco Agroinfiltration

Vacuum agroinfiltration

Plantlets for agroinfiltration method were grown from seeds for 4-6 weeks in an *m vitro* growth chamber at 24-26°C with a 16 hour photoperiod.

A. tumefaciens strain LB4404 containing a desired construct was grown on LB medium (Triptone 10 g/1, yeast extract 5 g/1, NaCl 10 g/1) supplemented with kanamycm (50 mg/1) and rifampicme (100 mg/1) at 28°C with shaking (250 rpm) overnight (about 18 hours) . Agrobacteria were then inoculated in 30 ml of LB also supplemented with kanamycm (50 mg/1) and rifampicin (100 mg/1). After overnight culture at 28°C (about 18 hours), agrobacterial cells were collected by centrifugation for 10 minutes at 3000xg and resuspended in 10 ml of liquid MS medium with MES (Sigma Chemical) 4.9 g/1 and sucrose 30 g/1 at pH 5.8. Bacterial culture was adjusted to a final OD_{600} of 0.1 for agroinfiltration. Then, cell culture was supplemented with acetosyringone to a final concentration of 0.2 mM and incubated for 90 minutes at 28°C.

For agroinfiltration, the plantlets were totally covered with the suspension and vacuum was applied (100 KPa) for 5-6 seconds. The suspension was removed and plantlets maintained in a growth chamber at 24-26°C under a photoperiod of 16 hours for four days. Plant material was recovered and total protein extraction analyzed by immunoblot using anti-gamma-zem antibody.

Agrom filtration by syringe

Agrobacterium tumefaciens strain EHA 105 was grown at 28°C in L-broth supplemented with 50 µg ml/^{1} kanamycin and 50µg ml/^{1} rifampycin to stationary phase. Bacteria were sedimented by centrif ugation at 5000 g for 15 minutes at room temperature and resuspended in 1OmM MES buffer pH 5.6, 10 mM MgCl; and 200 µM acetosyringone to a final OD_{600} of 0.2. Cells were left m this medium for 3 h at room temperature. Individual *Agrobacterium* cultures carrying the RX3 constructs and the HC-Pro silencing supressor constructs (Goytia *et al.,*2006) were mixed together and infiltrated into the abaxial face of leaves of 2-4-week-old *Nicotiana benthamiana* plants (Voinnet *et al,* 2003) .

Example 13: Isolation (Purification) of RPBLAs by density gradient from transgenic plant vegetative tissues

The gene coding for RX3-EGF gamma-zein derived fusion proteins was introduced in tobacco plants via *Agrobacterium tumefaciens.* Transformed plants were analyzed by immunoblot to determine those plants with higher recombinant protein expression. The predominant lower bands of immunoblots correspond to the monomer form of fusion proteins and the higher bands to the dimers. The fusion proteins usually accumulate as multimers and the amount of monomers and oligomers detected in the immunoblots depends on the disulfide bond reduction level.

Tobacco leaf extracts were loaded on density step gradients and the accumulation of recombinant proteins m the different fractions was analyzed by immunoblot. The results indicate that RX3-EGF appeared in fractions

corresponding to dense RPBLAs. Most of these organelles exhibited densities higher than 1.2632 g/crn' and a significant portion of them show a density higher than 1.3163 g/cm^3.

These novel RPBLAs formed in tobacco leaves exhibit densities in the range of the natural maize protein bodies (Ludevid et al., 1984 *Plant MoI. Biol.* 3:227-234; Lending et al., 1989 *Plant Cell* 1:1011-1023), or are even more dense.

It was estimated that more than 90 percent of the recombinant protein was recovered in the dense RPBLAs fractions and pellet. Thus, isolation of RPBLAs by density appears to be a useful system to purify (concentrate) the fusion proteins.

To evaluate the purification of the recombinant protein RX3-EGF by RPBLAs isolation, the different density fractions were analyzed by silver stain. More than 90 percent of tobacco endogenous proteins were located in the soluble and the interphase fractions of the gradient, the fractions in which, RX3-EGF protein was absent or barely detected. Thus, soluble proteins and the bulk of proteins present in less dense organelles could be discarded by selecting one or two fractions of the gradient.

In respect to the degree of fusion proteins purification in the RPBLAs fractions, it was estimated that RX3-EGF protein represents approximately 80 percent of the proteins detected in the PBLS-containmg fractions. This result indicates that, using a RPBLAs isolation procedure, one can achieve an important enrichment of fusion proteins in only one step of purification.

Example 14: Recombinant proteins recovery in RPBLAs isolated from dry plant tissues

An important point in molecular farming is the presence of an easy means to store plant biomass. In this context, drying can provide a convenient method to lessen storage volume and preserve the product. Nevertheless, drying frequently promotes the degradation of the proteins of interest. The use of desiccated plants to isolate RPBLAs containing recombinant proteins would be of great interest for industrial purposes.

Transformed tobacco leaves accumulating RX3-EGF fusion protein as described above were dried as also discussed above. After 5 months of dry storage, the stability of recombinant proteins was analyzed. Protein extracts from equivalent amounts of wet (fresh) and dry leaf tissue were analyzed by immunoblot . The RX3-EGF protein was stable in desiccated transformed plants, the amount recovered in wet and dry plants being similar.

The distribution in step density gradients of RX3EGF fusion protein from homogenates of dried leaves was analyzed by immunoblot. The fusion protein was mainly recovered in dense structures exhibiting densities higher than 1.1868 g/cm^3 and 1.2632 g/cm^3.

Thus, recombinant proteins can be purified from dried tissues via isolation of RPBLAs thereby illustrating that transgenic plant collection and recombinant protein extraction and purification can be independent in time. In keeping with these results, gamma-zem fusion proteins were also accumulated in RPBLAs in rice seeds.

Example 15: Recombinant protein recovery by isolation of RPBLAs from transiently transformed tobacco plantlets

The transient expression systems can be a convenient tool to test the accumulation behavior of

recombinant proteins in a short period of time. Thus, the recombinant protein RX3-EGF was also expressed and accumulated in transiently transformed tobacco plantlets via agroinfiltration. The protein extracts from transformed plantlets analyzed by immunoblot show the characteristic complex electrophoretic pattern observed from stably-transformed plants, indicating that the fusion proteins assemble correctly using this method of transformation.

Example 16: Recovery of recombinant proteins by low and medium speed centrifugation

To simplify the procedure used to purify recombinant proteins via dense recombinant protein body-like assemblies, two additional alternative methods were performed: i) clarified homogenates were centrifuged through only one dense sucrose cushion and ii) clarified homogenates were simply centrifuged at low speed centrifugation (i.e. 1000-2500xg for 10 minutes).

In agreement with the previously described results, the RX3-EGF protein was recovered in high yields (more than 90%) in the pellets obtained after centrifugation through 1.1868 g/cm3 sucrose cushions. In addition, the purification of RX3-EGF protein was very high in that contaminant tobacco endogenous proteins were barely detected in the corresponding pellet.

The principal advantage of this method as compared to step density gradients lies in its easy scalability for industrial production of recombinant proteins. It should be noted that the cushion density as well other properties such as its viscosity and osmolarity can be adjusted in each case in order to optimize recovery and purification of the recombinant proteins.

In addition, low speed centrifugation (LSC) was also assayed to concentrate and purify fusion protein-containing protein body-like structures. The results indicated that, after lOOOxg for 10 minutes, practically all the RX3-EGF fusion protein was recovered in the pellet. But the staining of the proteins contained in this pellet revealed that the fusion protein was not highly purified as compared with that obtained after centrifugation through 1.1868 g/cm3 sucrose cushion.

Thereafter, the first pellet obtained by low speed centrifugation was washed by using a buffer containing 5% Triton® X-100. After washing, the sample was centrifuged at 12,000xg for 5 minutes and, interestingly, the bulk of contaminating proteins present in the Pl pellet were eliminated after washing and centrifugation and the new pellet contained a highly enriched RX3-EGF protein. It is noted that the amount as well the pattern of proteins noted in this study is similar to those obtained after washing the pellet obtained after centrifugation through the sucrose cushion in the Triton X-100-containmg buffer. The low speed centrifugation alternative is based on the high density of the structures containing fusion proteins and centrifugation conditions can be optimized for every target before to scale up.

Transgenic tobacco plants expressing fusion proteins that include EGF linked to rice prolamin or alpha-zem rather than RX3, rP13-EGF and the 22aZ-EGF, were produced by *Agrobacterium tumefasciens* transformation. The best expressers where determined by immunoblot using an antibody against the EGF, and those cell lines were used in a comparative analysis with tobacco plantlets agroinfiltrated with the same constructs. In all cases,

the RPBLAs where recovered in unique interface, suggesting that the RPBLAs are very dense and homogeneous.

Taking all these results together, it is clear that prolamins are able to induce high density RPBLAs, even when they are fused to other proteins. That is an unexpected result, mainly when almost no homology is observed between them. Moreover, there are some data suggesting that the prolamins interact to stabilize the protein bodies, and that some of them are not stable when expressed in vegetative tissue alone, as for instance alpha-zem (Coleman et al., 1996 *Plant Cell* 8:2335-2345)

Example 17: Extraction of recombinant proteins from isolated RPBLAs

It has been demonstrated that the isolation of dense recombinant PB-like assemblies is an advantageous method to recover recombinant proteins with high yield and high purification level from transgenic organisms. Here it is shown that these recombinant proteins can be extracted from the storage organelles.

After an overnight (about 18 hours) incubation of RPBLAs fractions at 37°C in a buffer containing a detergent and reducing agents (SB buffer that contained sodium borate 12.5 mM pH 8, 0.1% SDS and 2% β-mercaptoethanol ; treatment), RX3-EGF protein was solubilized. The extracted fusion protein was recovered in its soluble form. Afterwards, as a function of their application, the extracted proteins can be submitted to further purification or used as partially purified extracts.

Example 18: Plasmid construction for animal cell transformation

The RX3 sequence was amplified by PCR to obtain the cDNA fragments corresponding to RX3 and RX3-(GIy)x5. These fragments were digested by Sall/BamHI cloned in plasmid pECFP-Nl (Clontech) opened by the same enzymes to obtain pRX3-ECFP and pRX3-G-ECFP plasmids, respectively.

Primers :

SPfor (SEQ ID NO: 54)

RX3ECFP3' (SEQ ID NO:55)

RX3G5ECFP3' (SEQ ID NO: 56)

The p22aZ-ECFP vector corresponds to the following Hmdlll/Xbal DNA fragment in pEGFP-Nl plasmid (Clontech) (SEQ ID NO:57)

The GFP was obtained by PCR amplification of the plasmid pEGFP-Nl (Clontech) with specific oligonucleotides containing enzyme restriction sites for further cloning:

ECFP Ncol 5' (SEQ ID NO:58)

ECFPNl BamNotSac 3'(SEQ ID NO: 59)

The PCR product (GFP) was cloned in a PCR cloning vector (PCR®II Vector, Invitrogen) and the sequence verified. The GFP fragment was excised by NcoI/BamHI digestion and cloned into pUC18RX3hGH (US2006123509 (Al)), giving the cassette RX3-GFP in a pUCl δ vector. This cassette was liberated by Sall/BamHI digestion and

subsequently cloned into a pCDNA3.1(-) (Invitrogen) previously digested by XhoI/BamHI (p3.1-RX3-GFP)

A construct containing the coding sequence of an improved monomeric DS Red protein (mCherry; Shaner et al., 2004 *Nat. Biotechnol.* 22:1567-1572) was a template in a PCR reaction (mCherry Real 5'/ECFPNl BamNotSac 3).

mCherry Real 5' (SEQ ID NO: 60)

The PCR product (DsRed) was cloned in a PCR cloning vector (PCR®II Vector, Invitrogen)) and the sequence verified. The DsRed fragment was excised by Rcal/BamHI digestion and cloned into pUC18RX3hGH (US2006123509 (Al)), giving the cassette RX3-DsRed in a pUC18 vector. This cassette was liberated by Sall/BamHI digestion and subsequently cloned into a pCDNA3.1(-) (Invitrogen) previously digested by XhoI/BamHI (p3.1-RX3-DsRED)

To obtain a RX3 cDNA with a STOP codon at the 3' end, the RX3 fragment was amplified by PCR (SPFOR/RX3STOP) and digested by Sall/BamHI. The fragment was cloned in pcDNA3.1(-) digested by the same restriction enzymes to obtain p3.1-RX3.

RX3STOP3' (SEQ ID N0:61)

The cDNA encoding the hGH were fused to the RX3 N-termmal gamma-zem coding sequence (patent WO2004003207) and was introduced into the vector pcDNA3.1(-) (Invitrogen) as described elsewhere. In the resulting construct named p3.lRX3hGH, the fusion protein sequences were under the CMV promoter and the terminator pA BGH.

The Ssp DNAb intern from pTWINl plasmid (New England Biolabs) and the hGH cDNA were amplified by PCR. Both PCR fragments were fused in frame, also by PCR, digested by NcoI/BamHI and cloned in pUC18RX3hGH (US2006121573 (Al)) vector also digested by NcoI/BamHI. The RX3-Int-hGH insert was obtained by Sall/BamHI digestion of this intermediate vector and cloned in pcDNA3.1(-) (Invitrogen) digested by XhoI/BamHI. The resulting contruct was named p3.l-RX3-I-hGH . The PCR product was digested by BsRGI/BamHI and cloned in p3.l-RX3-I-hGH plasmid digested with the same restriction enzymes.

Primers :

5'DNAb (SEQ ID NO: 62)

3'DNAb (SEQ ID NO: 63)

DNAb-hGH: (SEQ ID NO: 64)

3'hGH (SEQ ID NO: 65)

As negative control of cleavage induction, an uncleavable Ssp DnaB was engineered. The mutated (Aspl54 → Alal54) Ssp DnaB intern fused in frame to the hGH was obtained by PCR from p3.l-RX3-I-hGH .

Primers :

IM-for (SEQ ID NO: 66)

IM-rev (SEQ ID NO: 67)

Full length cDNAs of human caspase-2 (IRAUp969A0210D6) and caspase-3 (IRATp970B0521D6) were acquired from RZPD GmbH (Berlin), from an original reference based at the Nacional Lawrence Livermore Library.

By PCR, the caspase-3 and the caspase-2 specific cleavage (DEVD and DEHD, respecively) site were added at 5' termini of the corresponding caspase sequence. It is important to note that amplified fragment corresponding to caspase-2 did not contain the pro-domain.

Casp3 forward (SEQ ID NO: 68)

Casp3 reverse (SEQ ID NO:69)

Casp2 for (SEQ ID NO:70)

Casp2 reverse (SEQ ID N0:71)

The amplified sequences were cloned into pUC18RX3hGH (US2006123509 (Al)) by digesting with Ncol and Kpnl. The resulting construct was then digested by Sall/Kpnl and cloned to a pCDNA3.1 (Invitrogen) vector digested by Xhol/Kpnl. The corresponding vectors were named (p3.1-RX3-C2 and p3.1-RX3-C3) .

The pUC18RX3hGH (US2006123509 (Al)) vector was digested by Hindlll/EcoRI, and the liberated insert cloned in pCambia2300 also digested by these enzymes. The corresponding vector was digested by Hmdlll/Ncol and the insert cloned in pCambial381 opened by Hindlll/Ncol (p4-17) . The DNA comprising the RX3-(gly)x5-GUS fragment was obtained by digesting p4-17 by BstEII, then filling in the overhang with klenow and finally digesting by Sail. This

fragment was cloned in pcDNA3.1(-) digested by XhoI/EcoRV to obtain the p3.1-RX3-GUS clone.

The p3.1-RX3-EK corresponds to the following Nhel/Hmdlll DNA fragment in pcDNA3.1(-) (Invitrogen) (SEQ ID NO: 72)

Example 19: Plasmid construction for insect infection

The RX3-DsRED fragment from p3.l-RX3-DsRED was digested by Xbal/Hindlll and cloned in pFastBacl (Invitrogen) digested also by these two enzymes in order to obtain pF-RX3-DsRED vector.

The DsRED cDNA was amplified by PCR from pF-RX3-DsRED by using the following primers:

bGH rev (SEQ ID NO:73)

bGH rev2 (SEQ ID NO:74)

To obtain the pF-DsRED vector, the PCR-amplif ied DNA fragment was digested by Xbal/Hmdlll and cloned in pFastBacl (Invitrogen) also digested by Xbal/Hindlll .

Example 20: Insect cell and larvae infection

Baculovirus and Larvae

The baculoviral expression vector system (pFastBac, Invitrogen) , was used as the basis vector for this work. The recombinant virus was produced and amplified as described by the manufacturer. Cabbage looper, *Trichoplusia ni,* eggs were obtained from Entopath,

Inc. (Easton, PA). The eggs were hatched according to the directions provided by the manufacturer, and fourth instar larvae were used for infection.

Larvae Infection

Various amounts of baculovirus stock solution, consisting of occluded recombinant virus were spread on the larval diet, which was ordered premade m Styrofoam cups from Entopath, Inc. (Easton, PA). The cups were covered and allowed to stand for an hour so that the virus was completely absorbed by the media. The fourth instar larvae were then placed into the cups (approximately 10 to 15 larvae per cup), and the cups were inverted. The larvae fed from the top (bottom of cup) so that fecal matter dropped on to the lid where it was discarded daily. The quantity of food was sufficient for at least 5 days of growth. Three to five larvae were collected daily for RX3-DsRED and DsRED analysis.

SF9 Infection

Spodoptera

Sf9 cells were obtained from Invitrogen (San Diego, CA, U.S.A.) and cultured as previously described (O'Reilly et al., 1992) using Grace's insect medium supplemented with lactalbumin hydrolysate, yeastolate, L-glutamme, 10 % heat-mactivated fetal bovine serum and 1% penicillin/streptomycin solution (Gibco). Cells were grown in either spinner flasks (Bellco Glass, Vmeland, NJ, U.S.A.) or 100 mm plastic tissue culture dishes (Falcon). Recombinant viruses were produced using the BaculoGold Transfection Kit (PharMmgen, San Diego, CA, U.S.A.). Single plaques were isolated and amplified two to four times to obtain a high-titre viral stock which was stored

at 4°C until use. For routine infection, Sf9 cells in Grace's medium were allowed to attach to the bottom of a 100 mm plastic culture dish (107 cells/dish) . After incubation for 15 mm to 1 h, a portion of viral stock was added and the cultures were maintained at 27 ^{0}C in a humidified air atmosphere. Commonly cells were used at 30-36 hours after infection.

Example 21: RPBLAs preparation from Mammal cells and insect larvae

Homogenization

Mammal cells

Transfected cells were recovered from culture plates by scraping and were suspended in the homogenization B medium (10 mM Tris-HCl pH 8.0, 0.9% NaCl, 5 mM EDTA with protease inhibitors). The cell suspension was taken into a 5 ml syringe fitted with a 23 gauge needle and it was taken up and expelled approximately 30 times. Cell rupture was monitored by a phase contrast microscope.

Insect larvae

Frozen Trichoplusia ni larvae expressing RX3-DsRED and DsRED proteins were homogenized in PBP5 buffer (20 mM Hepes pH 7,5, 5 mM EDTA) by polytron for 2 minutes at 13500 rpm and by Potter for 5 minutes in ice at 2000 rpm. This homogenate was centrifuged at 20Og 10 minutes to remove cuticle and tissue debris and the supernatant was loaded on a density step gradient.

RPBLAs isolation by density

RPBLAs from mammal cells and frozen insect larvae were isolated essentially as described for planes (density step gradient or low speed centrifugation).

Example 22: Solubilization by Triton X-114 based biphasic separation

Cell homogenates were diluted with PBS and centrifuged at 16,000xg for 15 minutes. The supernatant was removed and the pellet dried. It was added 2 ml of ice cold Solubilisation Buffer (5OmM Tris pH7, 5% Triton X-114, 20 mM TCEP, 20 mM NDSB195 and 10OmM $MgCl_2$) to the pellet, and afterwards 1 ml of PBS containing IM Urea, 10% Glycerol and 100 mM $MgCl_2$.
This composition was incubated on ice for 15 minutes with occasional vortexing. The suspension was then sonicated for 20 seconds X 4 at 50% potential, keeping it on ice between bursts for 1 minute to maintain the cold temperature. The suspension was then incubated at 37°C for 15 minutes to form the 2 phases. Three millilitres of 10% PEG were added to the lower hydrophobic layer (Triton X-114 rich) and the composition was incubated on ice for 20 minutes. Then, the solution was incubated at 37°C for 15 minutes to form the 2 phases again. The upper phase (4 ml) was recovered and stored for analysis.

Example 23: Immunolocalization

Immunocytochemistry using a fluorescent microscope (Vertical Eclipse Microscope Nikon E600A). Between 2 to 4 days after transfection, cells were fixed for 30 minutes in 1% paraformaldehyde solution and after

washing with phosphate saline buffer, incubated for 45 minutes with the antibody against: (i) hGH (dilution 1/150), (ii) EK (dilution 1/500), (in) RX3 (dilution 1/700) . In order to detect the antigen-antibody reaction, an incubation for 45 minutes with anti-rabbit conjugated to Alexa Fluor 488 (Invitrogen)

Confocal analysis were performed in a Confocal laser scanning microscope (Leica TCS SP, Heidelberg, Germany) fitted with spectrophotometers for emission band wavelength selection. Green fluorescent images were collected at 488 nm excitation with the Argon ion laser by using an emission window set at 495-535 nm. Red fluorescent images were collected after 543 nm excitation with a HeNe laser and emission window 550-600. Optical sections were 0.5 to 1 µm thick.

Example 24: Activity assays

EGF activity assay

MDA-MB231 cells (breast cancer cells that overexpress EGF receptor) are seeded in 96-well plates at 5,500 cells/well. Cells were allowed to adhere for 8 hours in growth medium with 10% FCS (Fetal calf serum) and then starved overnight in medium supplemented with 0.1% of FCS. Afterwards, the media is removed and the EGF (positive control) from Promega or the corresponding sample (solubilyzed RX3-EGF) is added at different concentrations. Then, the radioactive timidme is added to a final concentration of 0,5 µCi. Proliferation is studied at 48 hours after stimulation at 37 °C. Then, the cells are washed twice with cold PBS, and the cells are kept on ice to stop the cell metabolism. A 10% trichloroacetic acid (TCA) solution

is added, and the cells are incubated for 20 minutes at 4°C. Once the TCA solution is removed, the plates *are* washed twice with Ethanol at 70%, and the cells are incubated for 20 minutes at 37 ^{0}C in 0.5 rnL the lysis solution (2% CO3Na2, 0.IN NaOH and 10% SDS). Plates are mixed by vortex agitation and the sample is not measured before 12 hours to avoid undesired chemo-lurru niscent phenomena.

EK activity assay

The enzymatic activity was measured by fluorometric assay (Grantet al. (1979) *Biochim. Biophys. Acta* **567**:207-215). The reaction was initiated by adding the enzyme to 0.3 to 1.0 mM of the fluorogenic substrate GIy-(Asp)4-Lys-βnaphtylamide (Sigma) in 25 mM Tris-HCl (pH 8.4),
10 mM CaC12, 10% DMSO (Dimethyl sulfoxide) at 37°C. Free β-naphtylamme concentration was determined from the increment of fluorescence (λex = 337 nm and λem = 420 nm) continuously monitored for 1 mm. The activity was calculated as change in fluorescence over time.

GUS activity assay

GUS activity assay is based in the catalysis of metilumbelif eril-β-glucuronide acid (MUG) to the 4-metilumbelif erone (4-MU) fluorescent product, by the GUS enzyme (Jefferson RA, et al. (1987) EMBO J. 6(13): 3901-3907). 50 µL of solubilyzed RX3-GUS (or solubilyzed RX3 as a control) was added to 200 µL of Reaction buffer (50 mM of phosphate buffer pH7, 10 mM EDTA, 0,1% SDS and 0,1% Triton XlOO) plus 66 µL of Methanol. The substrate (MUG) was added to a final concentration of 10 mM. The standard was prepared by adding 0, 50, 100, 200, 300 or 500 pmols of 4-

MU (the product of the reaction) to 200 µL of Reaction buffer of the reaction (4-MU) .

The samples and the standard were mixed and they were measured in a fluorimeter (Victor, Perkin-Elmer) at λeκ = 355 nm and λem = 460 nm. The samples were measured each 30 minutes for 3 hours. The specific activity was calculated by the formula: GUS activity (pmols 4-MU/min-l*mg-l) = (λem(Tl) - (λem(TO))/(k * (Tl-TO)). "K" = ratio (Units of fluorescence) /(pmol 4-MU) .

RTB activity assay

fAsialofetuin-binding ELISA)

The functionality of RX3-RTB in the protein extracts from RPBLAs was determined via binding to asialof etum, the glycoprotein fetuin treated with sialydase to expose galactose-termmated glycans. Two hundred microliters of asialofetuin (Sigma) at a concentration of 300 mg/mL in modified PBS (mPBS) buffer (100 mM Na-phosphate, 150 mM NaCl, pH 7.0) was bound to the wells of an Immulon 4HBX (Fisher, Pittsburg, Pa.) microtiter plate for 1 hour at RT. The coating solution was discarded and the wells blocked with 200 ml 3% BSA, 0.1% Tween 20 in mPBS for 1 hour at RT. After the blocking solution was discarded, 100 ml of RTB standards and protein extracts (see below) were applied and incubated for 1 h at RT. The wells were then washed three times with 200 ml mPBS, 0.1% Tween 20. Rabbit anti-R. communis lectin (RCA60) polyclonal Ab (Sigma) at 1:4000 in blocking buffer (as above) was applied and incubated for 1 hour at RT. The wells were then washed as before. AP conjugated goat-anti-rabbit IgG (Bio-Rad) was applied at a dilution of 1:3000 in blocking buffer and incubated for 1 h at RT. The wells were washed three times as described above and 100 ml pNPP

(pnitrophenyl phosphate disodium salt) substrate (Pierce, Rockford, 111.) was applied. The reaction was stopped after 15 minutes by the addition of 50 µl of 2 N NaOH. Absorbance (A405) was read in a Bio-Tek EL808 Ultra Microplate Reader. Protein extracts were prepared at a ratio of 1 g FW leaf to 3 ml of Tris-acorbate buffer (above), and the samples compared against a standard curve consisting of serially diluted castor bean-derived RTB (Vector Labs, Burlingame, Calif.) in Tris-acorbate buffer, with the concentrations ranging from 5 ng to 500 ng per well.

Example 25: Enhanced uptake of RX3-DsRED assembled in RPBLAs from insect larvae by macrophages

The cDNΔ coding for RX3-DsRED and DsRED were cloned in the baculovirus FastBac vector (Invitrogen) to obtain pFB-RX3-DsRED and pFB-DsRED. These constructs were used to infect *Trichoplusia ni* larvae. Frozen larvae expressing RX3-DsRED and DsRED proteins were homogenized and loaded on a density step gradient. After centrifugation at 80000xg in a swinging-bucket for 2 hours, the analysis of the RX3-DsRED fusion protein and the control corresponding to DsRED expressed in the cytosol was performed by immunoblot (Fig. 2C). As expected, when expressed in the larval cells cytosol, the DsRED protein did not assemble in highly dense structures and was localized in the supernatant and the F35 fraction (Fig. 2C, lane 2 and 3). On the other hand, RX3-DsRED fusion protein was able to assemble and accumulate in dense structures that can be isolated from F56 (Fig. 2C, lane 5). As shown by confocal microscopy analysis in Example 4 (Fig. 4), the RX3-DsRED accumulated in round-shaped RPBLAs.

The RPBLAS of RX3-DsRED from F56 were diluted 3-fold in PBP5 (10 mM HEPES pH 7.4, 2 mM EDTA) and collected in the pellet by centrif ugation at 80000xg at 4°C in a swinging-bucket for 2 hours. The pellet was resuspended in PBS buffer and the number of RPBLAs was quantified by FACS. From 1 larva infected with the pFB-RX3-DsRED vector, approximately $IxIO^9$ RPBLAs particles were obtained at a concentration of 500,000 RPBLAs per microliter (µl).

It has been reported that antigen presentation by the antigen presentation cells (APC) such as the macrophages and dendritic cells is a key process necessary to induce the immune response (Greenberg et al, *Current Op. Immunology* (2002), 14:136-145). In this process, the APC phagocytoses the antigen, which is subsequently cleaved in small peptides in the phagolysosome. These peptides interact with the MHCII and are sorted to the plasma membrane to be presented to the cell- and antibody-mediated immunity responses (Villandagos et al., *Immunological Reviews* (2005) 207:101-205) .

To determine the antigenicity of RX3 fusion proteins present inside the RPBLAs, a macrophage cell culture was incubated with these organelles at different RPBLA/cell ratios (100:1 and 1000:1) . The macrophage cell cultures were grown on starved conditions or m the presence of (M-CSF) . These cell cultures were incubated with RPBLAs for 1 hour, and 1, 2, 5 and 10 hours after RPBLA removal, the macrophages were extensively washed with PBS and fixed with 2% paraformaldehyde. Afterwards, these fixed macrophages were analyzed by FACS to quantify the amount of fluorescent RPBLAs up taken by the macrophages as well as the percentage of macrophages that had phagocytosed the fluorescent RX3-DsRED RPBLAs

Percentage of Fluorescent Macrophages

Time (hours)	Starved (RPBLA/cell ratio 100 1)		M-CSF (RPBLA/cell ratios 100 1)		M-CSF (RPBLA/cell ratio 1000 1)	
	Mean	STD	Mean	STD	Mean	STD
zero	1.19	1 21	0 82	0 35	0 82	0 35
1	65 42	2 29	65 19	3 2	85 78	1 65
2	79 64	1 66	75 08	3 94	91 55	1 5
5	91 85	2 17	87 68	1 58	91 53	1 09
10	88 91	0 7	90 54	1 59	94 4	0 08

From these results, it is clear that the macrophages phagocytosed the RX3-DsRED RPBLAs with an unexpected avidity. Even at the lower RPBLAs/cells ratio (1:100) and in the presence of M-CSF, at 1 hour after RPBLAs addition, 65% of macrophages are fluorescent. Even, the presence of a mitogenic cytokine, such as M-CSF, which has a negative effect on macrophage phagocytosis can not impair significantly the RPBLAs uptake. At 5 hours, almost all (more than 80%) of the macrophages were fluorescent, meaning that the majority of the cells had up taken some RPBLAs from the medium.

When the amount of fluorescence associated with the macrophages was analyzed over time of incubation, the result was even more surprising. In any of the conditions analyzed (ratio RPBLAs/cells or presence of absence of M-CSF) no saturation effect on the capacity of the macrophages to uptake the RPBLAs was observed. If the results of the Tables above and below are compared at 5 and 10 hours of incubation, it is seen that almost all the macrophages are fluorescent, but there is a continuous increase in the total fluorescence associated to the macrophages. This result indicates that, the macrophages are phagocytosing a large quantity of fluorescent RPBLAs particles .

Time Dependent Macrophage Fluorescence

Time (hours)	Starved (RPBLA/cell ratio 100 1)		M-CSF (RPBLA/cell ratios 100 1)		M-CSF (RPBLA/cell ratio 1000 1)	
	Mean	STD	Mean	STD	Mean	STD
0	0 975	0 31	0 725	0 1	0 725	0 1
1	8 9	0 42	10 3	1 13	24	1 7
2	16 35	0 07	16.25	0 5	41 5	0 3
5	64 65	2.05	42 35	4 45	93 3	2 2
10	120 7	1 84	79 9	5 66	125 65	13 08

To demonstrate that RPBLAs containing the RX3-DsRED fusion protein were inside the macrophages and not simply adsorbed to plasma membrane, confocal microscopy analysis were performed. Fig. 7A (left panel) shows some of those macrophage cells incubated with RX3-DsRED particles (at 100:1) for 1 hour. On the left panel of the same figure, a section of 1 micrometer of the same cells shows the typical green auto-fluorescence of macrophages observed with a green filter (Fig. 7A, white arrowhead). The presence of the nucleus and the red-fluorescent RPBLAs particles (Fig. 7A, black arrowhead) in the same optical section indicated that the RPBLAs had been taken up inside the cells by phagocytosis.

Another important factor to be analyzed is the degradation of the immunogen once it has been phagocytosed by the macrophage. Antigen degradation is needed to produce the antigenic peptides that are presented on the MHCII receptor. The analysis of the DsRED fluorescent pattern of the macrophages over the time showed that the RPBLAs particles were actively digested.

Another set of micrographs shows that after 1 hour of incubation, the RPBLA particles were not fully degraded and could still be observed inside the cells (Fig. 7B, upper panels). After 10 hours, the red fluorescence pattern was more homogenous all along the cells, indicating

that the macrophages had begun to degrade the RPBLA particles (Fig. 7B, bottom panels) .

Example 26: Enhanced uptake of RX3-DsRED in RPBLAs from insect larvae by dendritic cells

Dendritic cells plays a central antigen presentation role to induce the immune system (Blander et al., *Nature Immunology* (2006) 10:1029-1035) . Although rare, dendritic cells are the most highly specialized APC, with ability both to instigate and regulate immune reactivity (Lau AH et al *Gut* 2003 52:307-314) . To asses the capacity of those cells to phagocyte RX3-DsRED fusion proteins assembled in RPBLAs from insect larvae, a dendritic cell culture was incubated with these organelles at a 100 RPBLAs/cell ratio. Two kinds of RPBLAs were prepared: (i) RPBLAs isolated as described before and (ii) the same RPBLAs through fully washed in 50 mM Tris pH 8, 1% Triton X-100, in order to remove the ER membrane. The dendritic cell cultures were grown on starved conditions in the presence of RPBLAs, and samples were analyzed at 0, 1, 2, 5 and 10 hours.

Percentage of Fluorescent Dendritic cells

Time (hours)	% of fluorescent dendritic cells			
	RPBLAs		Membrane -Ie ss RPBLAs	
	Mean	STD	Mean	STD
0	1.43	-	1.41	-
1	26.76	-	36.46	0.28
2	33.79	0.6	50.785	0.21
5	45.845	0.07	67.275	3.4
10	61.885	5.73	74 .97	4.17

Time Dependent Dendritic cells Fluorescence

Time (hours)	Fluorescence associated ot dendritic cells			
	RPBLAs		Membrane-less RPBLAs	
	Mean	STD	Mean	STD
0	0.5	-	1.1	-
1	3.1	-	5.1	0.28
2	3.55	0.6	5.05	0.21
5	25. 15	0.07	54	3.4
10	37.05	5.73	74.05	4.17

As can be concluded from Tables above, the dendritic cells show a surprising avidity for RPBLAs. As expected, they have a slower phagocytosis rate compared to the macrophages (compare the previous tables), as is described elsewhere. The percentage of fluorescent dendritic cells increases all along the time course analyzed, and no saturation effect was observed even at 10 hours after RPBLAs incubation. Similar conclusions can be drawn when the amount of fluorescence associated to the macrophages over time was analyzed.

The dendritic cells' capacity to take up the RPBLAs did not exhibit a saturation effect. This lack of effect can be explained by the fact that more and more dendritic cells are induced to phagocytosis (and becoming fluorescent) over time. Nevertheless, it is also possible that the phagocytosis capacity of those cells is not saturated, as have been observed with macrophages.

Unexpectedly, the FACS analysis of dendritic cells incubated with membrane-less RPBLAs showed a significantly higher percentage of fluorescent dendritic cells than the same cells incubated with membrane-containing RPBLAs. Moreover, the fluorescence of these dendritic cells was also higher as well. Similar results

were obtained using macrophages with membrane-less RPBLAs. This was somewhat surprising as it was expected that the presence of insect-derived membrane proteins in the membrane-containing RPBLAs would be recognized as foreign proteins by the murine dendritic cells, and hence enhance phagocytosis. It is thus apparent that insect-derived RPBLAs in the presence or absence of the surrounding membrane are very efficient antigen presentation vehicles.

To demonstrate that RPBLAs and membrane-less RPBLAs containing the RX3-DsRED fusion protein were taken up by the dendritic cells, optical microscopy analysis was done. Fig. 8A (upper) shows dendritic cells incubated for 2, 5 and 10 hours with RX3-DsRED RPBLAs (100:1 ratio). On the bottom of Fig. 8B, the red fluorescence of the DsRED protein illustrates the uptake of the RPBLAs by those cells. At 2 hours of incubation, some phagocytosis can be observed, but most of the RPBLAs are only adsorbed to the plasmatic membrane. At 5 hours, and even more at 10 hours, many phagocytosed red fluorescent RPBLAs were observed. Similar results were obtained when dendritic cells were incubated with membrane-less RPBLAs (Fig. 8B).

It is important to note that even at 10 hours of incubation with RPBLAs or membrane-less RPBLAs, most of the phagocytosed particles remain visible as particles, meaning that little proteolysis had take place. This observation agrees with previous observation showing that the kinetics of protease acquisition, and hence, of proteolysis is slower in dendritic cells than in macrophages (Lennon-Dum'enil et al. (2002) *J. Exp. Med.* 196:529-540). These conditions may limit the proteolysis of proteins in dendritic cells and favor the generation of peptide antigens of appropriate length for loading onto MHC class II molecules.

Example 27: Phagocytosis of macrophages and dendritic cells

Macrophages

Macrophages were obtained from marrow of mice Balb/C. Mice were sacrificed by a cervical dislocation and femur and tibia were extracted. The bones were cut and the marrow was extracted with DMEM medium using a syringe. The marrow was cultivated on a 150mm Petri plate with complete DMEM medium (supplemented with 20%FCS and 30% L-cell) . A macrophage culture of 99% purity was obtained after 7 days of incubation at 37°C.

The differentiated macrophages were cultivated in complete medium to give 350.000 cells per well. When the cells were adhered, the medium was removed and cells were incubated with new medium that contained RX3-DsRED RPBLAs from larvae. The experiment was done with 100 or 1000 particles: 1 cell. The number of particles (RPBLAs) was counted by Coulter Epics XL FACS using the Argon laser at 488nm for excitation and FL2 at 575nm +/-30 for emission. Flow-count from Beckman Coulter ref. 7547053 (lot 754896F) was used to check the flowing.

After different times (0, 1, 2, 5 and 10 hours) the medium was removed and two washings with PBS were performed. Cells were permitted to recuperate and then fixed by PBS with 2% of paraformaldehyde. The treated macrophages were stored at 4°C and the fluorescence was analyzed by FACS (with the same program used for the counting) .

To verify that RX3-DsRED particles were been phagocitated inside the cells an experiment of

immunocitochemis try was done. The differentiated macrophages (50.000 cells/ well) were incubated with 100:1 particles of RX3-DsRED for 1 hour. After incubation, cells were washed twice with PBS and fixed with PBS with 2% formaldehyde for 15 minutes. Treated cells were analyzed by confocal microscopy.

Dendritic cells

The marrow from Balb/C mice was cultivated with complete medium (DMEM, 10% FCS, 5ng/ml GM-CSF) for one day. In order to remove granulocytes, plates were agitated and medium was changed twice. Medium was then changed twice without agitation and incubated for 2 days to obtain immature denditric cells. Denditric cells were incubated with 100:1 particles of RX3-DsRED for 1, 5 and 10 hours. After treatments, cells were fixed with 2% paraformaldehyde, stored at 4°C and the fluorescence was analyzed by FACS.

Each of the patent applications, patents and articles cited herein is incorporated by reference. The use of the article "a" or "an" is intended to include one or more.

The foregoing description and the examples are intended as illustrative and are not to be taken as limiting. Still other variations within the spirit and scope of this invention are possible and will readily present themselves to those skilled in the art.

CLAIMS

1. A eukaryotic host cell that contains a recombinant fusion protein within recombinant protein body-like assemblies (RPBLAs),

said fusion protein containing two sequences linked together in which one sequence is a protein body-inducing sequence (PBIS) and the other is a biologically active polypeptide,

said biologically active polypeptide being naturally found in a second cell type that is different from said eukaryotic host cell,

said eukaryotic host cell not producing protein bodies in the absence of said fusion protein.

2. The host cell according to claim 1 wherein said fusion protein further includes a linker sequence between the protein body-inducing sequence and the sequence of the biologically active polypeptide.

3. The host cell according to any previous claim wherein the protein body-inducing sequence comprises a prolamin sequence.

4. The host cell according to claim 3 wherein the prolamin sequence is gamma-zein, alpha-zein, gamma-gliadm or rice prolamin.

5. The host cell according to claim 4 wherein the prolamin sequence is the gamma-zem RX3 sequence.

6. The host cell according to any previous claim wherein the protein body-inducing sequence further includes

a sequence that directs a protein towards the endoplasmic reticulum (ER) of a plant cell.

7. The host cell according to claim 6 wherein the protein body-inducing sequence that directs a protein towards the ER of a plant host cell is a signal peptide from the same plant as the remainder of the PBIS or from a different plant.

8. The host cell according to claim 7 wherein said signal peptide is selected from the group consisting of the 19 residue gamma-zein signal peptide sequence, the 19 residue signal peptide sequence of alpha-gliadm, the 21 residue gamma-gliadm signal peptide sequence and the pathogenesis-related protein of PRlO class 25 residue signal peptide sequence.

9. The host cell according to any previous claim wherein the biologically active polypeptide of said fusion protein exhibits at least 25% of the biological activity of the same polypeptide isolated from said second cell type.

10. The host cell according to any previous claim wherein the biologically active polypeptide of said fusion protein contains at least two N-linked glycosylation sequences .

11. The host cell according to any previous claim wherein the protein body-inducing sequence and the biologically active polypeptide of said fusion protein are linked together by a spacer amino acid sequence that is cleavable by enzymatic or chemical means or is not cleavable .

12. The host cell according to any previous claim wherein said host cells are algal cells.

13. The host cell according to any previous claim wherein said host cells are animal cells.

14. The host cell according to any previous claim wherein said host cells are higher plant cells.

15. The host cell according to any previous claim wherein said host cells are fungal cells.

16. Recombinant protein body-like assemblies (RPBLAs) that comprise a membrane-enclosed fusion protein, said fusion protein containing two sequences linked together in which one sequence is a protein body-inducing sequence (PBIS) and the other is a biologically active polypeptide .

17. The RPBLAs according to claim 16 that have a density of about 1.1 to about 1.35 g/ml.

18. The RPBLAs according to claims 16 or 17 wherein said PBIS includes a prolamm sequence.

19. The RPBLAs according to claim 18 wherein said prolamm sequence is selected from the group consisting of gamma-zein, alpha-zein, rice prolamm and gamma-gliadin .

20. The RPBLAs according to claim 19 wherein the prolamm sequence is the gamma-zem RX3 sequence.

21. The RPBLAs according to claims 16 to 20 wherein said PBIS also includes a sequence that directs a protein towards the endoplasmic reticulum (ER) of a plant cell.

22. The RPBLAs according to claim 21 wherein said sequence that directs a protein towards the ER of a plant cell is a signal peptide from the same plant as the remainder of the PBIS or from a different plant.

23. The RPBLAs according to claim 22 wherein said signal peptide is selected from the group consisting of 19 residue gamma-zein signal peptide sequence, the 19 residue signal peptide sequence of alpha-gliadm, the 21 residue gamma-gliadm signal peptide sequence and the pathogenesis-related protein of PRlO class 25 residue signal peptide sequence.

24. A method of preparing a biologically active polypeptide that comprises the steps of:

a) providing recombinant protein body-like assemblies (RPBLAs) that comprise a membrane-enclosed fusion protein, said fusion protein containing two sequences linked together in which one sequence is a protein body-inducing sequence (PBIS) and the other is a biologically active polypeptide;

b) contacting the RPBLAs with an aqueous buffer containing a membrane-disassembling amount of a surfactant;

c) maintaining said contact for a time period sufficient to disassemble the membrane and at a temperature that does not denature the biologically active polypeptide to separate the membrane from the fusion protein;

d) collecting the separated fusion protein that contains the biologically active polypeptide.

25. The method according to claim 24 wherein the separated fusion protein exhibits the biological activity of said biologically active polypeptide.

26. The method according to claim 24 or 25 wherein said biologically active polypeptide is linked to said PBIS by a by a spacer amino acid sequence that is cleavable by enzymatic or chemical means.

27. The method according to claim 26 wherein said biologically active polypeptide exhibits biological activity when cleaved from the BPIS of the fusion protein.

28. The method according to claim 24 to 27 wherein said RPBLAs are present in a generally spherical form having a diameter of about 0.5 to about 3 microns.

29. A vaccine or inoculum comprising an immunogenic effective amount of recombinant protein body-like assemblies (RPBLAs) that contain a recombinant fusion protein dissolved or dispersed in a pharmaceutically acceptable diluent, said recombinant fusion protein containing two sequences linked together in which one sequence is a protein body-inducing sequence (PBIS) and the other is a biologically active immunogenic polypeptide to which an immunological response is to be induced by said vaccine or inoculum.

30. The vaccine or inoculum according to claim 29 wherein said fusion protein further includes a linker

sequence between the protein body-inducing sequence and the sequence of the biologically active immunogenic polypeptide .

31. The vaccine or inoculum according to claims 29 or 30 wherein the PBIS comprises a prolamin sequence.

32. The vaccine or inoculum according to claim 31 wherein the prolamin sequence is gamma-zein, alpha-zein, gamma-gliadm or rice prolamin.

33. The vaccine or inoculum according to claim 31 wherein the prolamin sequence is the gamma-zem RX3 sequence .

34. The vaccine or inoculum according to claim 29 to 33 wherein said RPBLAs improve the antigen delivery to antigen-presenting cells.

35. The vaccine or inoculum according to claim 29 to 34 wherein said RPBLAs improve the antigen processing and presentation to antigen-presenting cells.

Figure 1

Fig. 1A

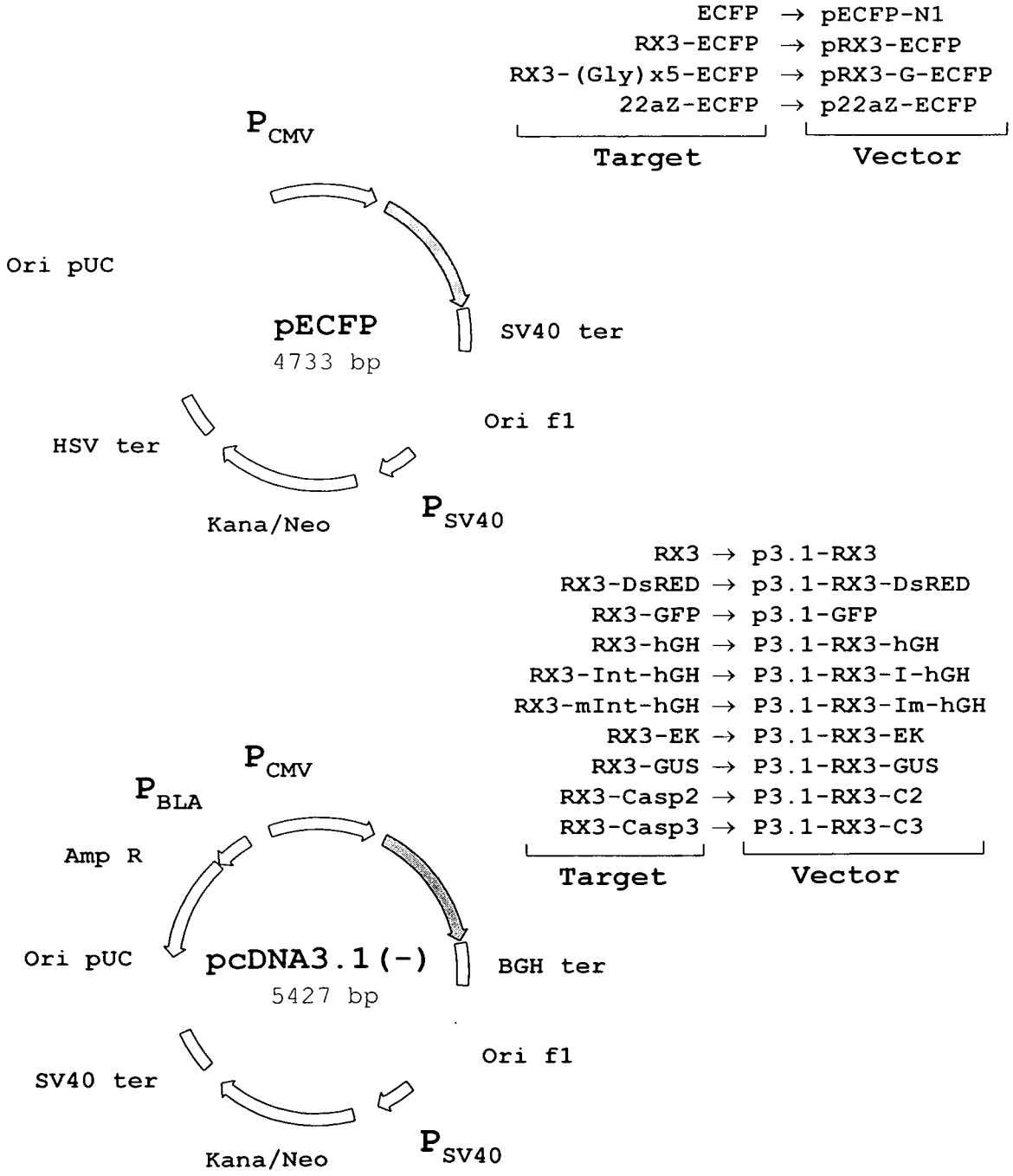

Figure 1

Fig 1B

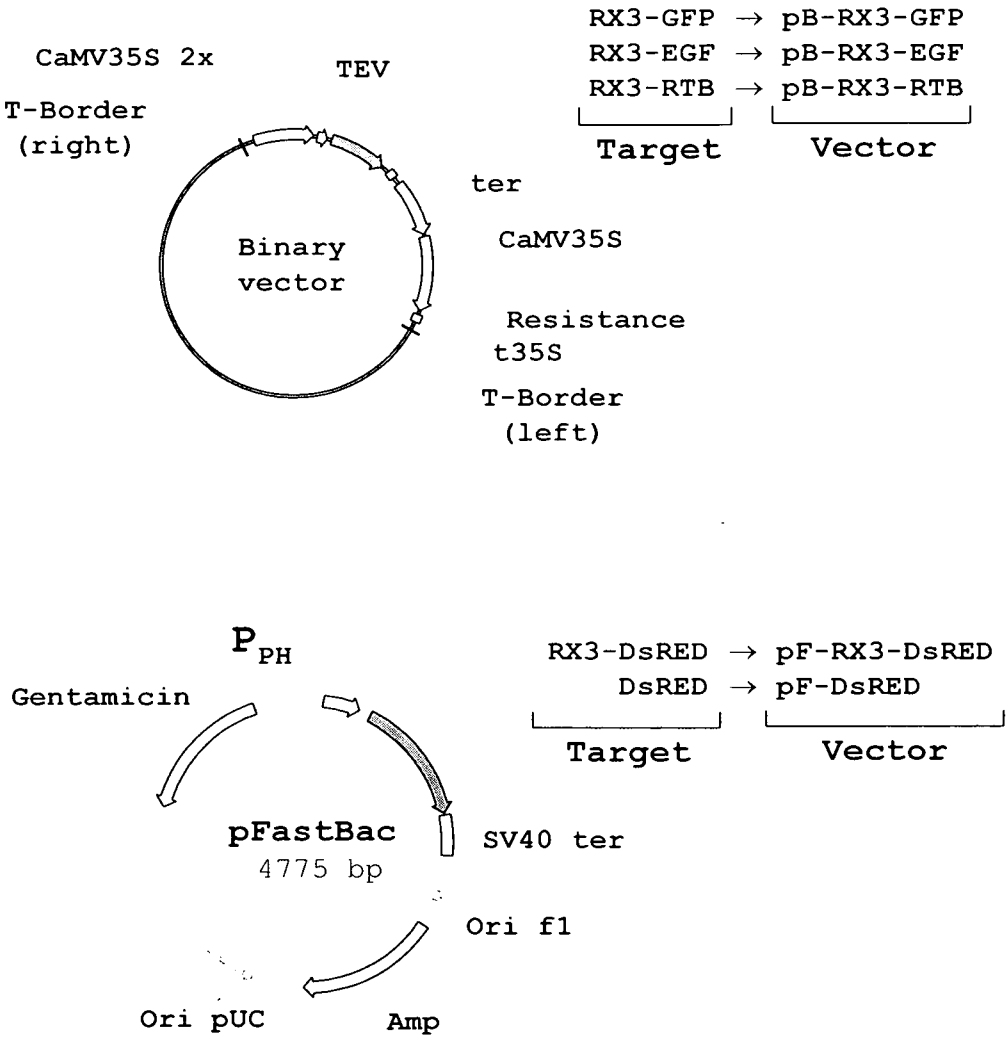

Figure 2

Fig 2A

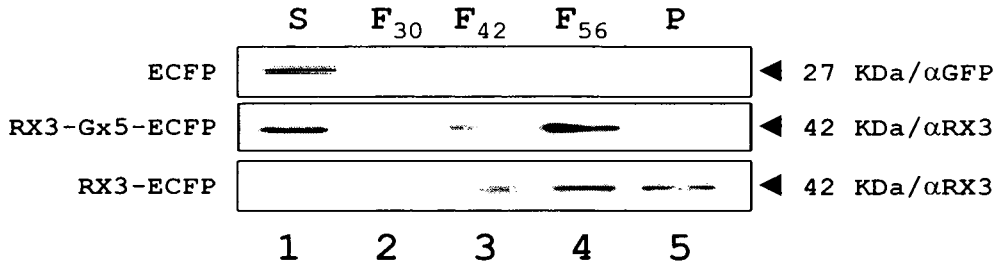

Fig 2B

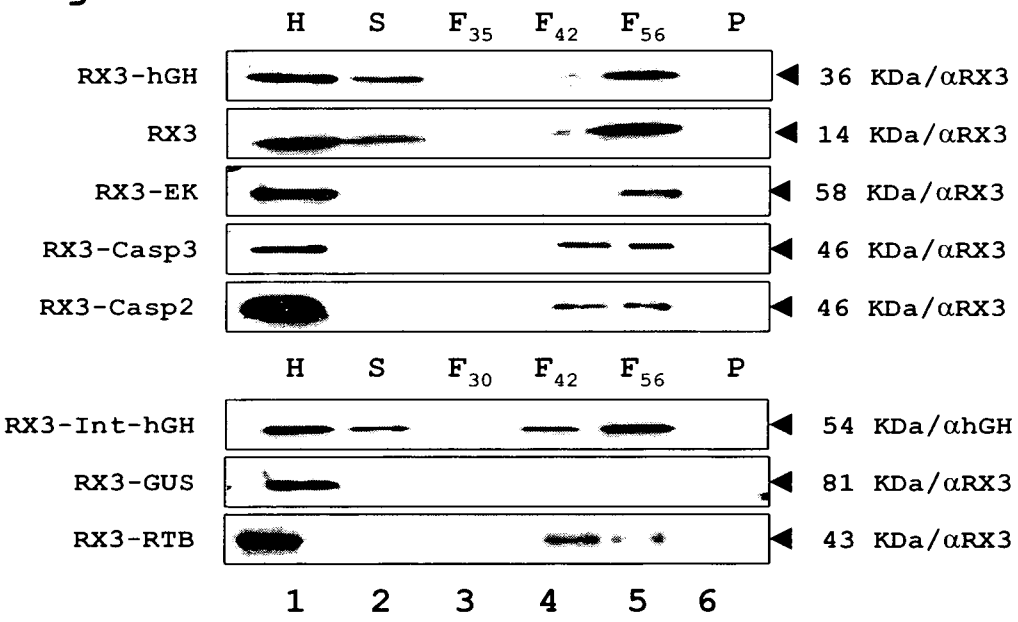

Fig 2C

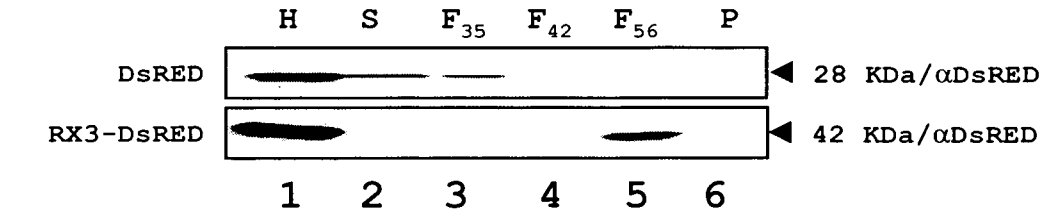

Figure 3

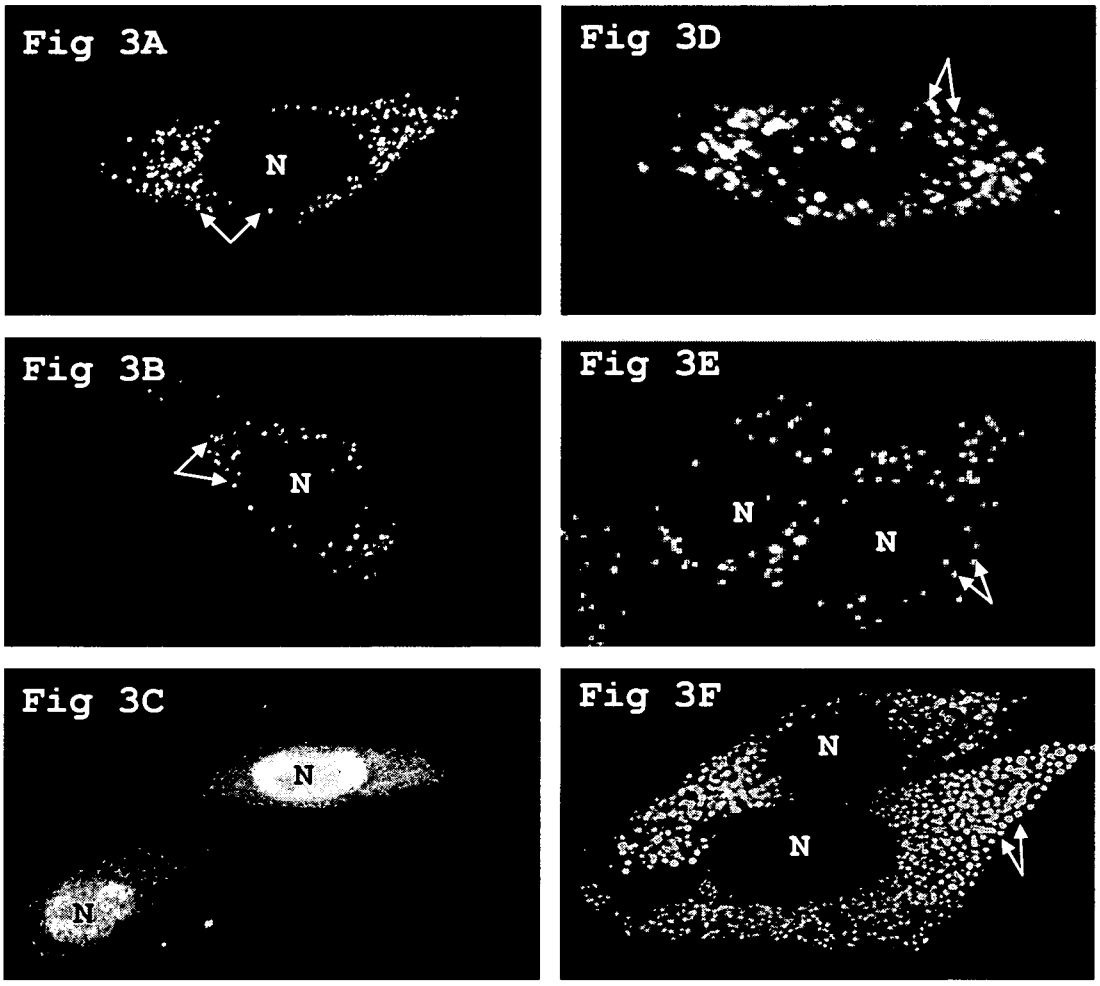

Figure 4

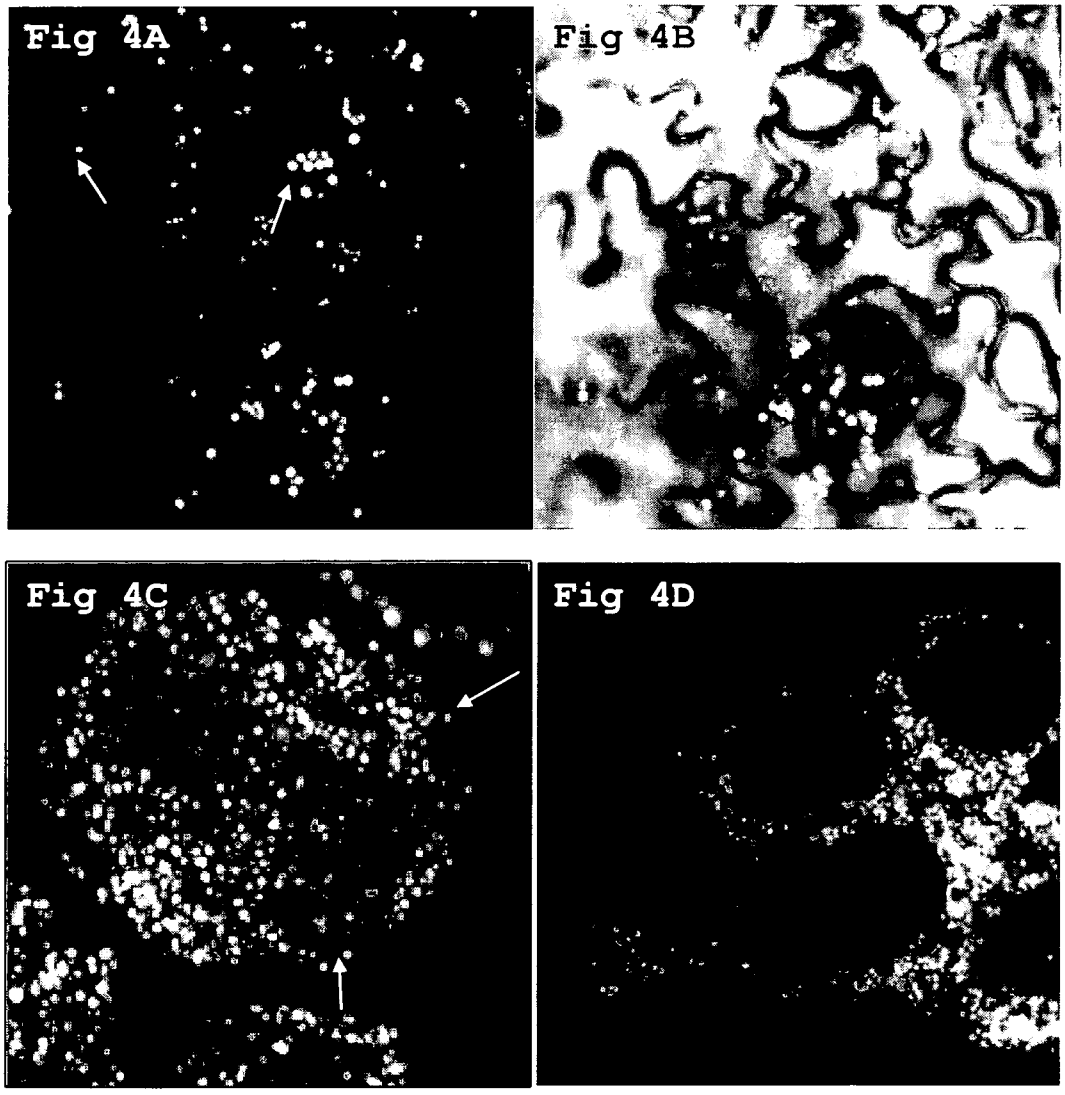

Figure 5

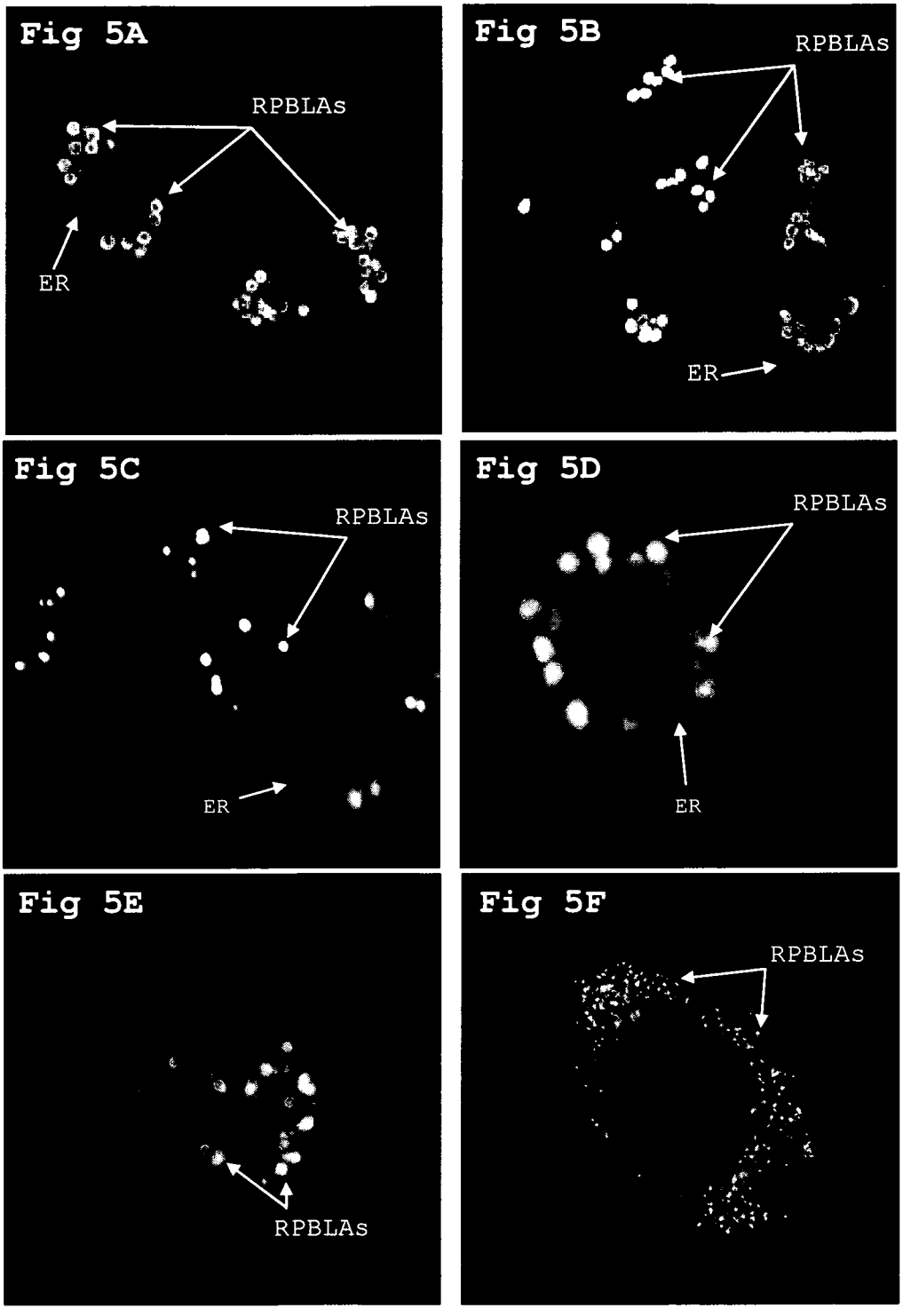
Fig 5A
RPBLAs
ER
Fig 5B
RPBLAs
ER
Fig 5C
RPBLAs
ER
Fig 5D
RPBLAs
ER
Fig 5E
RPBLAs
Fig 5F
RPBLAs

Figure 6

Fig 6A

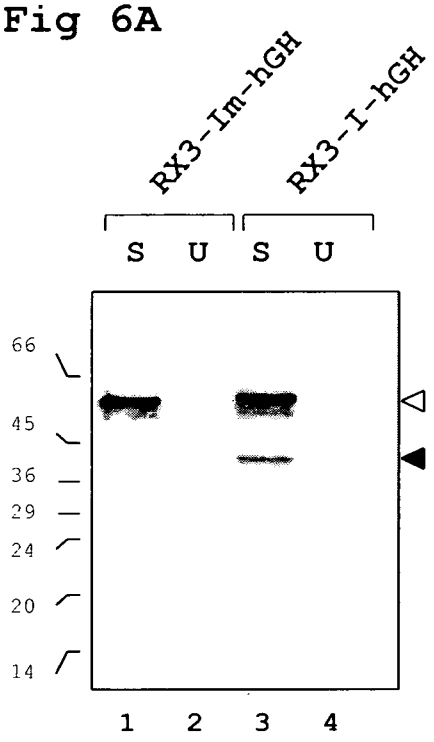

Fig 6B

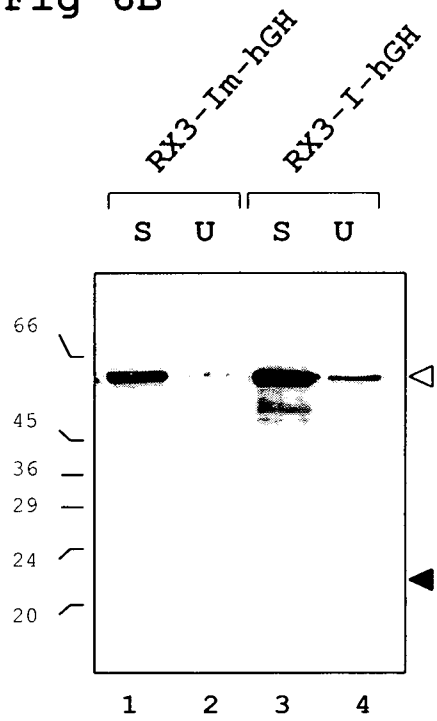

Fig 6C

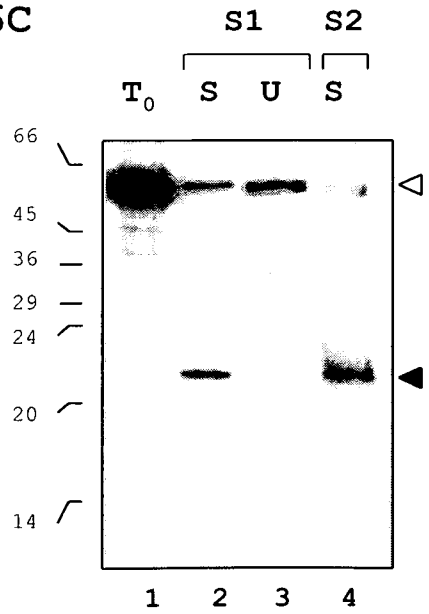

Figure 7

Fig 7A

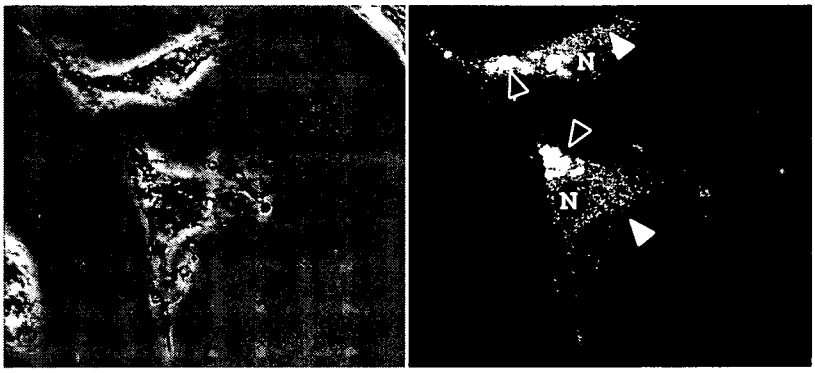

Fig 7B

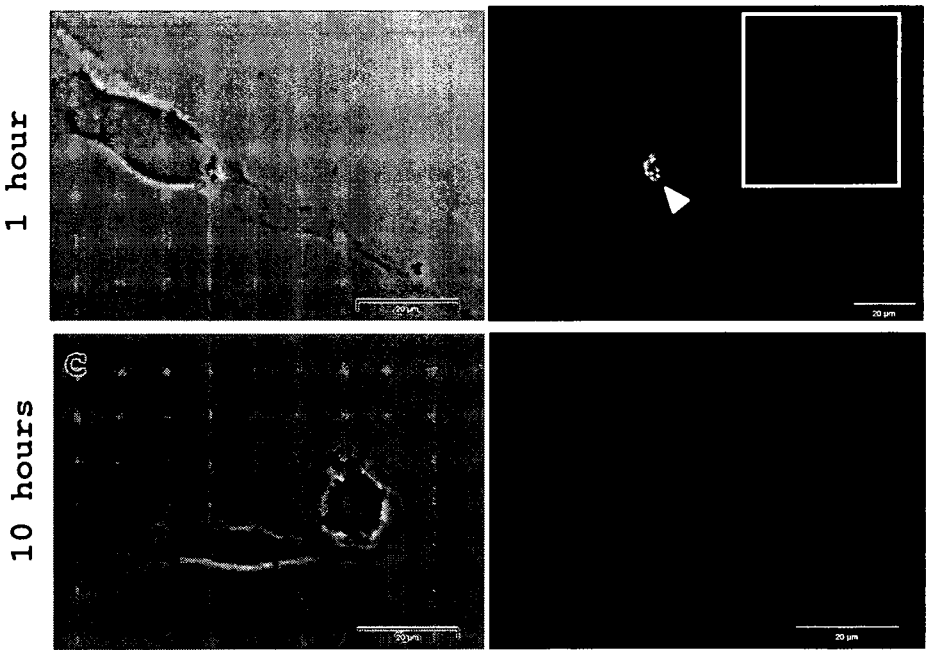

Figure 8

Fig 8A

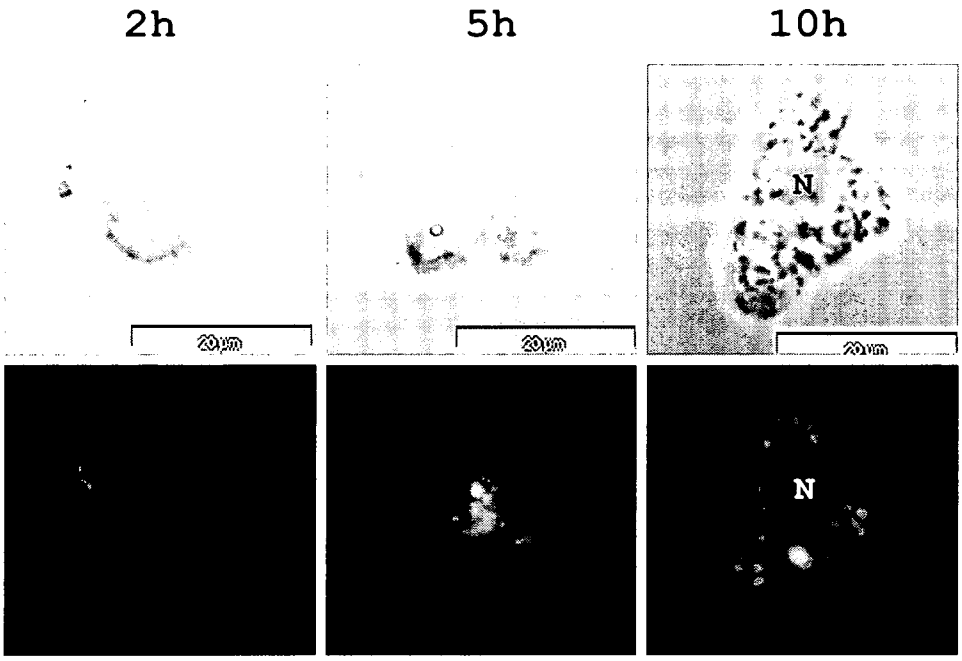

Fig 8B