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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 October 2016 (13.10.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/162385 A1

(51) International Patent Classification:
C07K 14/245 (2006.01) *G01N 33/68* (2006.01)
C07K 14/47 (2006.01)

(21) International Application Number:
PCT/EP2016/057543

(22) International Filing Date:
6 April 2016 (06.04.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
P 201530456 7 April 2015 (07.04.2015) ES

(71) Applicant: **CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS (CSIC)** [ES/ES]; C/ Serrano, 117, 28006 Madrid (ES).

(72) Inventors: **GIRALDO SUÁREZ, Rafael**; Centro De Investigaciones Biologicas (CIB), C/ Ramiro de Maeztu, 9, 28040 Madrid (ES). **MOLINA GARCÍA, Laura**; Centro De Investigaciones Biologicas (CIB), C/ Ramiro de Maeztu, 9, 28040 Madrid (ES).

(74) Agent: **PONS ARIÑO, Ángel**; Glorieta de Rubén Darío, 4, 28010 Madrid (ES).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

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

WO 2016/162385 A1

(54) Title: BACTERIAL SYSTEM FOR THE IDENTIFICATION OF AMYLOIDOGENIC PEPTIDES AND THE SCREENING OF INHIBITORS OF AMYLOIDOSIS

(57) Abstract: The invention relates to a bacterial system useful for the in vivo identification of amyloidogenic peptides and inhibitors of amyloid protein aggregation, based on the expression of a fusion protein that comprises the amyloidogenic peptide under study bound to a bacterial translation termination factor. Therefore, the invention provides such a fusion protein, a bacterial cell that expresses it, a method for identifying amyloidogenic peptides and a method for identifying inhibitors of amyloid protein aggregation in which the bacterial cell is cultured and the phenotypic changes produced as a consequence of the alteration in protein translation caused by the aggregation of the translation termination factor in the fusion protein are visualised.

BACTERIAL SYSTEM FOR THE IDENTIFICATION OF AMYLOIDOGENIC PEPTIDES AND THE SCREENING OF INHIBITORS OF AMYLOIDOSIS

DESCRIPTION

The present invention is framed within methods and systems for the *in vivo* identification of peptide sequences with amyloidogenic potential and molecules that inhibit amyloid protein aggregation, which are useful for the treatment and/or prevention of diseases characterised by the formation of seed aggregates. Therefore, the invention pertains to the fields of protein engineering and diagnostic and therapeutic methods for amyloid proteinopathies.

PRIOR ART

Amongst the structures that may be adopted by a protein, amyloid structures are the only ones whose thermodynamic stability may compete with that of the native state. In amyloids, one or more segments of a polypeptide chain are assembled, as β–strands, to form β–sheets which, after becoming laterally associated, produce fibrils of indefinite length and with variable morphologies. The amyloid state may be adopted by practically any protein, regardless of its sequence and type of three-dimensional folding, and may be modulated by both mutations and allosteric effectors (ligands). Amyloid aggregation follows nucleated polymerisation kinetics, wherein a discrete number of soluble protein molecules aggregate, thereby altering their conformation and acting as a “seeds” for the subsequent growth of the aggregate by recruiting additional molecules of the same protein and converting them into the amyloid state.

Amyloid aggregates, particularly oligomeric assembly intermediaries, are cytotoxic. Thus, amyloids are causal agents of a large number of diseases, both systemic (type II diabetes, haemodialysis-associated amyloidosis) and neurodegenerative (Alzheimer’s, Parkinson’s, spongiform encephalopathies or prionic diseases, Huntington’s, amyotrophic lateral sclerosis), all known generically as amyloid proteinopathies or amyloidosis. Consequently, there is a great demand for model systems that allow for the monitoring of, and intervention in, amyloidogenesis, i.e. the generation of amyloid protein assemblies and aggregates, by means of *in vitro* approaches and, especially, *in vivo* approaches.

In recent years, numerous computer algorithms have been disclosed that reliably predict the amyloidogenic potential of a given protein sequence (Ahmed and Kajava, 2013, FEBS Lett 587, 1089-1095), which may be verified *in vitro* by means of a number of biochemical and biophysical approaches (electron microscopy, X-ray diffraction, circular dichroism, infrared spectroscopy, interaction of fluorophores with an affinity for amyloids, etc.). In general, both predictions and the *in vitro* characterisation of the aggregation are robust and have contributed substantial molecular-level knowledge about the structure, the stability and the dynamics of amyloids.

However, at present the capacity for therapeutic intervention in patients suffering from amyloidosis is severely limited, and *post mortem* anatomic pathology analyses are even used to obtain a definitive diagnosis.

In regards to heterologous model systems, there has been significant progress in the experimental development of amyloidosis in “humanised” animals, generally by means of transgenesis of the genes involved in human pathologies; the mechanisms and pathways involved in amyloid pathology are studied on these animals and potential inhibitory molecules of amyloidogenesis are assayed. Worth mentioning amongst these model systems are the *Caenorhabditis elegans* nematode, the *Drosophila melanogaster* fly and, especially, the *Saccharomyces cerevisiae* yeast (Narayan et al., 2014, Nat Chem Biol 10, 911–920). Although the existence of prions has been described in this yeast, they do not trigger an amyloid proteinopathy, but are epigenetic determinants (non-Mendelian inheritance) of phenotypically selectable traits that confer greater adaptive flexibility to the cells that propagate them.

Another screening method for identifying amyloidogenic proteins and modulators of amyloid aggregation is the system based on the so-called *curli system* of the *E. coli* bacteria, through which amyloid aggregates migrate to the exterior by forming amyloid fibrils bound to the cell surface. Such fibrils are made up of two proteins, CsgA and CsgB, and may be detected by staining with marker probes, for example, Congo red, which is capable of detecting the presence of these fibrils. Thus, this cellular export system has been proposed for the production of extracellular amyloid fibrils composed of heterologous amyloidogenic proteins (from yeasts and humans). Specifically, this system is based on the expression in *E. coli* of a fusion

protein that comprises the signal sequence of CsgA ($CsgA_{ss}$) and the protein with amyloidogenic potential under study, in order to subsequently visualise its presence on the surface of the bacterium in the form of amyloid fibrils. Therefore, this system has been disclosed as a method for distinguishing amyloidogenic proteins and identifying modulators of amyloid aggregation (Viknesh Sivanathan and Ann Hochschild, 2012, Genes & Development, 26:2659–2667).

Moreover, a bacterial system has been disclosed for the screening of peptide aggregation inhibitors, which uses a fusion protein composed of a polypeptide that forms aggregates, for example Aβ42, fused to a reporter protein, such as GFP or similar. In the absence of inhibition by the molecule under study, aggregation of the polypeptide causes incorrect folding of the fusion protein and fluorescence is not emitted. On the contrary, compounds capable of causing inhibition allow for the correct folding of GFP in its native structure, which results in the emission of a detectable fluorescent signal (US20070077552).

During the past decade, the mechanism that activates plasmid replication by RepA family proteins in Gram-negative bacteria has been disclosed. Recently, it has been found that DNA binding promotes the *in vitro* assembly of a winged-helix domain (WH1) in the form of amyloid fibrils in RepA (Giraldo, 2007, Proc Natl Acad Sci USA 104, 17388-17393); the same occurs in proteins that cause spongiform encephalopaties and Parkinson's disease. Amyloidogenesis is inhibited by a molecule that interferes with the binding of RepA-WH1 to DNA (Gasset-Rosa et al., 2008, Nucleic Acids Res 36, 2249–2256). In recent years, it has been discovered that, when RepA-WH1 is expressed in *E. coli*, outside its natural functional context and fused to a fluorescent marker protein, it causes a synthetic amyloid proteinopathy (Fernández-Tresguerres et al., 2010, Mol Microbiol 77, 1456–1469). Although it is not an infectious agent (i.e. it is not susceptible to horizontal transmission), a reason for which it is considered to be a "prionoid", RepA-WH1 is capable of templating its amyloid conformation on molecules of the same protein, both *in vitro* (Fernández-Tresguerres et al., 2010, Mol Microbiol 77, 1456–1469) and *in vivo* (Molina-García and Giraldo, 2014, J. Bacteriol 196(14):2536-2542). Moreover, its vertical propagation from the mother cell to daughter cells has been characterised by means of microfluidics (Gasset-Rosa et al., 2014, Mol Microbiol 91, 1070-1087). Thus, the synthetic bacterial prionoid RepA-WH1 is a minimal, bio-safe model for unravelling the common

mechanisms and pathways of amyloid proteinopathies (Giraldo et al., 2011, Prion 5, 60–64).

On the other hand, in bacteria such as *E. coli*, translation termination is dependent upon the existence of RF1 termination factor, which recognises the UAG and UAA codons, and RF2 termination factor, which recognises UAA and UGA, being both factors released from the ribosomal A site through the action of a GTPase (RF3). Detailed studies on the structure of bacterial ribosomes in their different functional states have made it possible to characterise the molecular details of translation termination complexes, in particular, the mechanism of action of RFs. Both RF1 and RF2 are proteins that act as structural analogues of tRNAs, by occupying the ribosomal A site over the termination codons, thereby exposing amino acid residues that recognise the stop nucleotide triplet and projecting their respective N-terminal domains towards the exterior of the ribosome. Under conditions wherein the levels of RF1/RF2 are drastically reduced, when a stop codon is reached, there is a pause in the translation and a charged anti-terminator tRNA (aminoacylated) inefficiently replaces the termination factor, thereby allowing for stop codon read-through in a discrete fraction of the pause events. The strict quality control exerted in *E. coli*, both on the stability of incompletely translated messengers and on partially synthesised proteins following a stop or pause in the ribosomes (Keiler and Feaga, 2014, J Bacteriol 196, 2123-2130), may be a reason why translation termination in bacteria has not as yet been exploited for biotechnological purposes.

In sum, there is a need for systems that allow for the quick, reliable and simple identification, preferably *in vivo*, of amyloidogenic peptides of any origin, as well as modulator compounds of amyloid protein aggregation. Such systems would make it possible, not only to identify and classify peptides on the basis of their amyloidogenic capacity, but also to identify candidate molecules for the treatment and/or prevention of diseases wherein amyloid aggregates are involved.

DESCRIPTION OF THE INVENTION

The present invention provides a system that allows for the efficient *in vivo* identification in bacteria, preferably in the *Escherichia coli* bacterium, of the amyloid aggregation capacity of any heterologous peptide sequence. Moreover, this bacterial

system makes it possible to assay the action of various molecules as potential inhibitors of such protein aggregation process, i.e. the system is also useful for identifying molecules with anti-amyloidogenic activity that may be used as therapeutic agents against amyloid proteinopathies.

This system of the invention involves the generation of a completely synthetic reporter system in bacteria, based on the insertion, at one of the ends (N-terminal) of the bacterial translation termination factor RF1, encoded by the *prfA* gene, which recognises the UAG stop codon, of at least one copy of the amyloidogenic sequence under study (Figure 1), and the subsequent expression of the resulting fusion polypeptide in the bacterium. In the event that such amyloidogenic sequence forms aggregates, the amyloid aggregation of a substantial fraction of RF1 factor molecules takes place, which entails the inactivation thereof; consequently, the ribosomes do not stop the translation at a UAG premature codon introduced into a reporter gene. This leads to the synthesis of a functional reporter protein, whose activity or expression are detectable.

Therefore, this bacterial system allows for the *in vivo* identification of amyloid sequences originated from any proteome, including human proteome, of biomedical or biotechnological interest, and inhibitory molecules of protein amyloidosis, preferably in humans.

The examples of the present invention show that this bacterial system is useful for the *in vivo* monitoring of the inhibitory action exerted by different anti-amyloid compounds (with recognised anti-amyloid activity *in vitro*) on protein aggregates. Thus, this system allows not only for the *in vivo* monitoring of the amyloid potential of any peptide sequence, but also for the identification and classification of inhibitory compounds of amyloidosis, whilst also allowing for a quick, simple and economical *in vivo* evaluation of the dose-dependent degree of toxicity of the different compounds assayed.

Another advantage of this system is that it is sufficiently sensitive to classify and grade the molecules under study according to their higher or lower inhibitory activity on the amyloid aggregation process, since the activity of the protein encoded by the reporter gene is not only detectable, but also quantifiable.

Another advantage lies in the fact that, due to their extraordinary proliferation capacity, bacteria are ideal organisms to perform quick *in vivo* diagnoses of the amyloidogenic potential of protein sequences, and to conduct a preliminary selection of molecules with anti-amyloidogenic potential in libraries of natural or synthetic substances.

The system described herein is the first of its nature generated in bacteria, which places their high proliferation capacity, versatile manipulation and low cost at the service of the screening of molecules of interest for the treatment and/or prevention of amyloid proteinopathies. Therefore, the system of the present invention is the first synthetic bacterial system based on the modulation of aggregation-dependent translation termination in proteins, for the *in vivo* identification of peptide sequences with amyloid potential and of inhibitors of amyloidosis.

Consequently, one aspect of the invention relates to a fusion protein that comprises:

a. at least one copy of an amino acid sequence of a peptide with amyloidogenic potential, and
b. the amino acid sequence of the bacterial RF1 translation termination factor,

where the C-terminal end of the amino acid sequence of (a) is bound to the N-terminal end of the amino acid sequence of (b). Hereinafter, this protein will be referred to as "fusion protein of the invention".

The present invention is based on the principle of protein aggregation of a bacterial translation termination factor, an aggregation caused by an amyloidogenic peptide whereto it is bound, and its involvement in the read-through of a premature stop codon that is exclusively and specifically recognised by that factor, where the stop codon is included in the sequence that encodes a reporter gene. Therefore, this principle, on which the embodiment of the present invention is based in relation to the RF1 translation termination factor, may be extrapolated to the bacterial RF2 translation termination factor, which recognises the UGA stop codon. Thus, in the fusion protein of the invention, the amino acid sequence of (b) may be replaced by the amino acid

sequence of the bacterial RF2 translation termination factor. Therefore, the scope of the present invention also includes a fusion protein as the one described above wherein the amino acid sequence of (b) is that of the bacterial RF2 translation termination factor and the premature stop codon used in the reporter gene is UGA.

The "RF1 translation termination factor" or "Release factor RF1" is the prokaryotic factor that recognises the UAA and UAG stop codons. Thus, RF1 is a protein that allows for translation termination by recognising one of these two stop codons in an mRNA sequence on the ribosome. Preferably, RF1 originates from the *E. coli* bacterium, although the scope of the present invention also includes any functional and structural homologue of RF1 in any bacteria. This translation termination factor is encoded by the *prfA* gene. The amino acid sequence of this translation termination factor is that described in UniProtKB under accession number C4ZTQ2.

The "RF2 translation termination factor" is the prokaryotic factor that recognises the UAA and UGA stop codons in the mRNA on the ribosome. Thus, RF2 is a protein that allows for translation termination by recognising one of these two stop codons in an mRNA sequence. Preferably, RF2 originates from the *E. coli* bacterium, although the scope of the present invention also includes any functional and structural homologue of RF2 in any bacteria. This translation termination factor is encoded by the *prfB* gene. The amino acid sequence of this translation termination factor is that described in UniProtKB under access number C5A0G3.

Both RF1 and RF2 are proteins that occupy the ribosomal A site over the termination codons, exposing amino acid residues that recognise the stop nucleotide triplet and projecting their respective N-terminal domains towards the exterior of the ribosome. For this reason, in the present invention, the N-terminal domain is preferred for the binding to the amino acid sequence of a peptide with amyloidogenic potential. Thus, the N-terminal end of the amino acid sequence of RF1/2 of (b) is bound to the C-terminal end of the amino acid sequence of the peptide (or peptide repeats), with amyloidogenic potential, of (a).

"Fusion proteins" or "chimeric proteins" are proteins designed from the binding of two or more genes that originally encode independent proteins or peptides. The translation of the fusion gene results in a simple polypeptide with properties derived

from each of the original proteins. Fusion proteins are artificially created by means of recombinant DNA technology, which typically involves deleting the stop codon of a cDNA sequence that encodes the first protein and inserting the cDNA sequence of a second protein into a reading frame through ligation or extension by PCR. This DNA-generated sequence will be expressed in a cell as a single protein. This fusion protein may include the complete sequences of both original proteins or only a portion of each. For example, without being limited thereto, the fusion protein of the invention may comprise, instead of the complete amino acid sequence of the RF1 or RF2 factor, one or more of the amino acid sequences of the RF1 or RF2 factors necessary for translation termination.

In another preferred embodiment, the fusion protein of the invention is further bound to an advantageous sequence for it to be purified once it has been generated through recombinant DNA or DNA synthesis technologies, for example, without being limited thereto, a histidine tail, more preferably comprising between six and ten histidines, a GST protein or a FLAG peptide. These sequences allow for the fusion protein to be isolated or purified by means of, for example, affinity column chromatography with Ni or Co resins.

The fusion protein of the present invention, and the variants or derivatives thereof, may be synthesised, for example, without being limited thereto, *in vitro*. For example, by means of solid-phase peptide synthesis or recombinant DNA approaches. The fusion protein of the invention may be produced through recombinant DNA or DNA synthesis technologies and may contain, for example, without being limited thereto, a signal sequence or another polypeptide with a protease cleavage site, for example, without being limited thereto, at the N-terminal end.

The fusion protein of the invention may present variants. These variants refer to limited variations in the amino acid sequence that make it possible to maintain the functionality of the fusion protein. This means that the reference sequence and the sequence of the variant are similar overall, and identical in many regions thereof. These variations are generated by means of substitutions, deletions and/or additions. Such substitutions are performed with conserved amino acids. Conserved amino acids are amino acids with similar lateral chains and properties in regards to, for example, hydrophobicity or aromaticity. These substitutions include, without being limited thereto,

substitutions between glutamic acid (Glu) and aspartic acid (Asp), between lysine (Lys) and arginine (Arg), between asparagine (Asn) and glutamine (Gln), between serine (Ser) and threonine (Thr), and/or between the amino acids that make up the alanine (Ala), leucine (Leu), valine (Val) and isoleucine (Ile) group. The variations may be artificially generated, for example, by means of mutagenesis or direct synthesis. These variations do not cause essential modifications in the essential characteristics or properties of the fusion protein. For this reason, the scope of the present invention includes, in addition to the fusion proteins whose amino acid sequence is identical or homologous to the sequences described in the present invention, others that differ from those described herein by a discrete number of residues (e.g.: between 1 and 15) due to the aforementioned mutations.

On the other hand, the amino acid sequence of (a), i.e. that of the peptide with amyloidogenic potential, may have any origin, and may originate from *E. coli* or any other heterologous proteome, whether eukaryotic or prokaryotic, for example, without being limited thereto, bacteria, fungi or yeasts, animals, whether wild or domestic, mammals (including humans, rodents, ruminants, felines, equines or canids) or vegetables. In another preferred embodiment, the amino acid sequence of (a), i.e. that of the peptide with amyloidogenic potential, is of human, non-human animal, vegetable, fungal or bacterial origin; more preferably, human. The amino acid sequence of (a) may be a peptide or protein fragment, or a complete protein or polypeptide, whether natural, synthetic or recombinant.

In the present invention, “peptide with amyloidogenic potential” is understood to mean any peptide, protein fragment, complete protein or polypeptide suspected to present the capacity to form amyloid aggregates, i.e. the capacity to be assembled as amyloid structures. “Amyloid” is understood to mean that insoluble, proteolysis-resistant protein material whose main components are amyloid fibrils, formed by several proto-filaments whose axis is co-axial with the beta-sheets.

In another preferred embodiment, the peptide with amyloidogenic potential is selected from the peptides Aβ, Tau, α-synuclein, prion protein, huntingtin, superoxide dismutase 1, Fus, RepA-WH1 or TDP-43, all known to persons skilled in the art as they are described in the literature.

In a more preferred embodiment, the peptide with amyloidogenic potential originates from the RepA-WH1 protein. This protein is a bacterial prionoid comprising a domain called "Winged-Helix" domain, which integrates a DNA-binding motif of the helix-turn-helix type. The binding of WH1 to DNA promotes the assembly of the protein in the form of amyloid fibrils *in vitro*. The peptide that confers the amyloidogenic properties to WH1 is described in the literature, for example, in Giraldo, 2007, Proc Natl Acad Sci USA 104, 17388-17393. More preferably, this peptide is that with SEQ ID NO: 1.

In another preferred embodiment of the fusion protein of the invention, the amino acid sequence of (b) is bound to 2, 3 or 4 copies of the amino acid sequence of (a). Preferably, those repeats are tandem repeats.

In a more preferred embodiment of the fusion protein of the invention, the copies of the amino acid sequence of (a) are bound to one another by means of a linker or spacer. Preferably, such linker is a Gly (G) amino acid.

In another preferred embodiment of the fusion protein of the invention, the amino acid sequence of (a) is bound to the amino acid sequence of (b) by means of a linker or spacer. Preferably, this linker is the GRSGSSGSSG sequence (SEQ ID NO: 20).

In the most preferred embodiment, the fusion protein of the invention comprises the (complete) amino acid sequence of the RF1 translation termination factor with the N-terminal end bound to at least two, preferably three and, more preferably, four, copies of the amino acid sequence of the peptide with amyloidogenic potential present in the WH1 domain, preferably the peptide with SEQ ID NO: 1. In the present invention, this preferred fusion protein will be referred to as WH1(R_{1-4})-RF1.

In an even more preferred embodiment, the amino acid sequence of (b) is bound to the amino acid sequence of (a), wherein the latter is SEQ ID NO: 18 (L_{26}VLCAVSLI$_{34}$).

Another aspect of the invention relates to a nucleotide sequence that encodes the fusion protein of the invention. Hereinafter, this nucleotide sequence will be referred to as "nucleotide sequence of the invention".

Due to the degeneracy of the genetic code, wherein various nucleotide triplets result in the same amino acid, there are various nucleotide sequences that result in the same amino acid sequence. The terms "nucleotide sequence", "nucleic acid" and "polynucleotide" are used interchangeably herein, and refer to a polymeric form of nucleotides of any length which may or may not be chemically or biochemically modified. Therefore, they refer to any polyribonucleotide or polydeoxyribonucleotide, both single-chain and double-stranded. The polynucleotide of the invention may be artificially obtained by means of conventional cloning and selection methods, or by means of chemical synthesis. In addition to the encoding sequence, the polynucleotide may have other elements, such as, for example, without being limited thereto, non-encoding sequences at the 5' or 3' ends, ribosomal binding sites or stabilising sequences. These polynucleotides may further include sequences that encode additional amino acids which may be useful, for example, without being limited thereto, to increase the stability of the fusion protein generated from it or allow for a better purification thereof.

For example, the polynucleotide of the invention may be introduced into a cloning vector or an expression vector in order to allow for the replication and/or the expression thereof. Preferably, such vector is an appropriate vector for the expression and purification of the fusion protein of the invention.

For this reason, another aspect of the invention relates to an expression vector, hereinafter "first expression vector of the invention", that comprises the nucleotide sequence of the invention.

Another aspect of the invention relates to an expression vector, hereinafter "second expression vector of the invention", that comprises the *lacZ* gene, where said gene comprises a UAG premature termination codon, recognisable by RF1, or a UGA premature termination codon, recognisable by RF2, in its encoding sequence. In a preferred embodiment, the premature termination codon is UAG. In another preferred embodiment, the premature termination codon is located at position A515 of said gene.

In the present invention, this construct that comprises the *lacZ* gene and includes the UAG premature stop codon at position A515 is called *lacZ-amber*.

The "expression vector" referred to in the present invention is selected from the group that comprises, without any limitation whatsoever, plasmids, phages, cosmids, phagemids, bacterial artificial chromosomes (BACs) or any other type of DNA molecule with the capacity to replicate inside a cell, preferably a prokaryotic cell. An expression vector is a DNA molecule wherein another DNA fragment may be integrated without losing its self-replicating capacity. In the expression vector, the nucleic acid that has been cloned therein and which is to be expressed is generally operatively bound to control sequences.

In the preferred embodiment, the first and/or the second expression vector of the invention is a plasmid.

More preferably, the plasmid that comprises the *lacZ* gene, where said gene comprises a UAG premature termination codon at position A515 of its encoding sequence is the plasmid called pFus-lacZ-amber in the present invention, and it is constructed as described in the examples shown further below.

The "*lacZ* gene" is the *E. coli* gene that encodes the enzyme beta-galactosidase, described in the GenBank under access number 945006.

In the present invention, the term "premature termination codon" refers to a UGA or a UAG termination codon, preferably UAG, located inside the encoding sequence of the gene, at a position such that the recognition thereof by RF1 (UAG) or RF2 (UGA) mediates the premature translation termination of the protein by the ribosome, thereby preventing the synthesis of a complete, functional protein.

The first and the second vector of the invention may further comprise other necessary elements for the expression of the nucleotide sequences of interest included therein, such as, for example, promoters, terminators, control sequences, regulators, etc. More preferably, said promoters are inducible; even more preferably, they are arabinose- or IPTG-inducible. Preferably, the first vector of the invention comprises an

IPTG-inducible promoter and the second vector of the invention comprises an arabinose-inducible promoter.

As used herein, the term "promoter" refers to a DNA region, generally located upstream from the starting site of transcription, which is capable of initiating transcription in a cell. This term includes, for example, without being limited thereto, constitutive promoters, specific promoters for a cell or tissue type, or inducible or repressible promoters. The control sequences are dependent upon the origin of the cell wherein the nucleic acid is to be expressed. Examples of prokaryotic promoters include, for example, without being limited thereto, the promoters for the *trp*, *recA*, *lacZ*, *lacI*, *tet*, *gal*, *trc*, or *tac* genes of *E. coli*. The expression of a nucleic acid in a prokaryotic cell also requires the presence of a ribosomal binding site located upstream from (before) the encoding sequence.

The first and the second vector of the invention may be introduced into a bacterial cell in such a way that aforementioned vector is maintained as a chromosomal component or as a self-replicating extrachromosomal vector, for example, a plasmid, a minichromosome or an artificial chromosome. The vector may contain any means to ensure self-replication. Alternatively, the vector may be such that, when introduced into the host cell, it becomes integrated into the genome and replicated jointly with the chromosome(s) wherein it has been integrated. Moreover, a single vector or plasmid or two or more vectors or plasmids that jointly contain the total DNA to be introduced into the genome of the host cell, or a transposon, may be used.

Another aspect of the invention relates to a bacterial cell, hereinafter "bacterial cell of the invention", that expresses the fusion protein of the invention, where said cell further comprises a reporter gene that comprises a UAG premature termination codon (if the fusion protein of the invention comprises the amino acid sequence of RF1) or a UGA premature termination codon (if the fusion protein of the invention comprises the amino acid sequence of RF2), preferably UAG (recognised by RF1), in its encoding sequence.

The bacterial cell of the invention expresses the fusion protein of the invention either because it comprises the first expression vector of the invention, in self-

replicating form or integrated into its chromosome, or because it comprises the nucleotide sequence of the invention integrated into its chromosome.

When, in the fusion protein of the invention expressed in the bacterial cell of the invention, the amino acid sequence of (b) is that of the bacterial RF2 translation termination factor, the cell of the invention comprises a reporter gene which comprises the UGA premature termination codon in its encoding sequence. When, in the fusion protein of the invention expressed in the bacterial cell of the invention, the amino acid sequence of (b) is that of the bacterial RF1 translation termination factor, the cell of the invention comprises a reporter gene which comprises the UAG premature termination codon in its encoding sequence. In a preferred embodiment, the reporter gene comprises a UAG premature termination codon in its encoding sequence.

The "reporter gene" referred to in the present invention may be any gene, preferably of bacterial origin, more preferably from *E. coli*, the expression whereof results in a detectable phenotypic trait, for example, without being limited thereto, a change in colour in the cell that expresses it (for example, beta-galactosidase), the emission of fluorescence (for example, green fluorescent protein, GFP, or similar, such as mCherry), the emission of luminescence (such as luciferase), or metabolic auxotrophy complementation (deficiency in the synthesis of an essential amino acid or metabolite, or the lack whereof manifests itself in a phenotype).

In a more preferred embodiment, the reporter gene is the *lacZ* gene that encodes beta-galactosidase, which more preferably comprises the premature termination codon, even more preferably UAG, at position A515 of its nucleotide sequence.

The reporter gene may be integrated into the bacterial chromosome or may be present in the bacterial cell of the invention in the form of an autonomously replicating plasmid or a plasmid integrated into the bacterial chromosome. Thus, in another preferred embodiment, the reporter gene is comprised in an expression vector, which may further comprise, without being limited thereto, an inducible promoter for the expression of said reporter gene, more preferably an arabinose-inducible promoter. In a more preferred embodiment, this expression vector is the second expression vector of the invention. Even more preferably, this expression vector is pFus-lacZ-amber.

In an even more preferred embodiment, the bacterial cell of the invention further presents inactivation, preferably by means of deletion, of said wild-type reporter gene, preferably *lacZ-WT*, in its chromosome.

In another, even more preferred embodiment, the bacterial cell of the invention further presents inactivation of the wild-type gene that encodes the bacterial RF1 translation termination factor, if the fusion protein comprises the amino acid sequence of RF1, or the bacterial RF2 translation termination factor, if the fusion protein comprises the amino acid sequence of RF2, in its chromosome. I.e. said cell preferably comprises inactivation of the wild-type *prfA* gene, if the fusion protein comprises the amino acid sequence of RF1, or the wild-type *prfB* gene, if the fusion protein comprises the amino acid sequence of RF2, in its chromosome.

In the present invention, "inactivation" is understood to mean total inhibition of the native reporter gene, *prfA* gene or *prfB* gene in the chromosome of the bacterial cell of the invention, which leads to a total suppression of the expression of the gene and, therefore, an elimination of the activity of the protein encoded by it. The inactivation may be performed, for example, without being limited thereto, by means of knockout, transgenesis, introduction of random mutations, interfering RNA, etc.

In another preferred embodiment, the bacterial cell of the invention is *E. coli*. In a more preferred embodiment, the bacterial cell of the invention is *E. coli* strain MRA8ΔlacZ (*prfA1*^{ts}), which comprises the wild-type *lacZ-WT* gene deleted from the chromosome and, moreover, the *prfA* gene that encodes RF1 carries a mutation that makes this protein soluble and functional at 30°C, but not at 42°C. This MRA8ΔlacZ strain (*prfA1*^{ts}) may be designed according to the detailed description provided in the examples shown further below. More preferably, this strain is modified to express the WH1(R_{1-4})-RF1 fusion protein comprised in an expression vector, and further comprises another expression vector that comprises a reporter gene which comprises a UAG premature termination codon in its encoding sequence; even more preferably, this expression vector is the plasmid pFus-lacZ-amber.

The term "expression" includes any step involved in the production of the fusion protein of the invention, which includes, without being limited thereto, transcription,

post-transcriptional modification, translation, post-translational modification and secretion.

Another aspect of the invention relates to a kit that comprises the first expression vector of the invention and a bacterial cell that may be transformed using that vector, where the aforementioned cell further comprises a reporter gene that comprises a UAG premature termination codon (if the fusion protein of the invention comprises the amino acid sequence of RF1) or a UGA premature termination codon (if the fusion protein of the invention comprises the amino acid sequence of RF2), more preferably UAG, in its encoding sequence. The host cell further has the wild-type reporter gene inactivated in its chromosome. More preferably, this bacterial cell is *E. coli* strain MRA8ΔlacZ (*prfA1*^{ts}) or, alternatively, a strain of *E. coli* that, in addition to that of the wild-type reporter gene, also presents inactivation of the native *prfA* gene (if the fusion protein of the invention comprises the amino acid sequence of RF1) or the native *prfB* gene (if the fusion protein of the invention comprises the amino acid sequence of RF2), more preferably *prfA.*

Another aspect of the invention relates to a kit that comprises the first and the second expression vectors of the invention and a bacterial cell, preferably *E. coli*, that may be transformed using those vectors and which further has the *lacZ-WT* reporter gene and the *prfA* gene inactivated in its chromosome. More preferably, this bacterial cell is *E. coli* strain MRA8ΔlacZ (*prfA1*^{ts}).

Another aspect of the invention relates to the use of the bacterial cell of the invention for the identification of amyloidogenic peptides.

Another aspect of the invention relates to the use of the bacterial cell of the invention for the identification of inhibitory compounds of amyloid peptide aggregation. These “inhibitors” may be total or partial inhibitors of protein or peptide amyloid aggregation process, and may be any type of molecule, substance, compound, composition, medicament or similar, whether natural or synthetic.

Another aspect of the invention relates to a method for identifying amyloidogenic peptides, hereinafter “first method of the invention”, which comprises:

a. Culturing the bacterial cell of the invention under conditions that allow for the expression of the fusion protein of the invention and the reporter gene that comprises a UAG premature termination codon (if the fusion protein of the invention comprises the amino acid sequence of RF1) or a UGA premature termination codon (if the fusion protein of the invention comprises the amino acid sequence of RF2), preferably UAG, in its encoding sequence,
b. Detecting the phenotypic changes produced in the bacterium as a consequence of the expression of the reporter gene, and
c. Classifying the peptide with amyloidogenic potential as an amyloidogenic peptide when phenotypic changes associated with the expression of the reporter gene are visualised in step (b).

If the peptide with amyloidogenic potential present in the fusion protein of the invention causes aggregation of the RF1 or the RF2 factors, these factors are inactivated and, therefore, they do not recognise the UAG premature stop codon, in the case of RF1, or the UGA premature stop codon, in the case of RF2, included in the reporter gene. This entails read-through of the stop codon by the ribosome, to generate a complete, functional reporter protein the expression or activity whereof may be detected and even quantified.

Another advantage of the present invention lies in the fact that, even in the presence of the fusion protein of the invention in the bacterial cell of the invention, there is a certain residual level of solubility and activity in the translation termination factor, which allows for translation termination in essential genes of the host cell and, consequently, for maintaining cell viability. Therefore, the methods described herein allow for the *in vivo* evaluation, in a viable cell, of the amyloid aggregation potential of peptides of any origin and the modulatory capacity of inhibitory compounds of amyloid aggregation.

When the reporter gene is the *lacZ* gene, read-through of the premature stop codon entails the translation of a functional enzyme beta-galactosidase detectable by the blue colouration of the bacteria grown in the presence of X-Gal. Thus, the presence of blue-coloured bacterial colonies indicates that the peptide with amyloidogenic potential present in the fusion protein of the invention expressed in the bacterial cell of

the invention is an amyloidogenic peptide, i.e. one with the capacity to form amyloid aggregates. On the contrary, if the peptide under test does not have this capacity, it will not cause the aggregation of the RF1 or the RF2 factor included in the fusion protein of the invention. Such factor will then be functional and, therefore, it will recognise the premature stop codon located in the reporter gene, thereby causing the translation to stop. As a consequence, a functional enzyme beta-galactosidase will not be produced and the colouration of the bacteria will be the normal one (white-pale yellow in the case of *E. coli*) in the presence of X-Gal. Therefore, in a preferred embodiment of the first method of the invention, the reporter gene is the *lacZ* gene that encodes beta-galactosidase; in step (b), the presence or absence of blue colouration in the bacteria is detected; and, in step (c), the peptide with amyloidogenic potential is classified as an "amyloidogenic peptide" when a blue colouration is visualised in step (b).

Another aspect of the invention relates to a method for identifying inhibitory compounds of amyloid peptide aggregation, hereinafter "second method of the invention", which comprises:

a. Culturing the bacterial cell of the invention in the presence of the compound under study under conditions that allow for the expression of the fusion protein of the invention and the reporter gene that comprises a UAG premature termination codon (if the fusion protein of the invention comprises the amino acid sequence of RF1) or a UGA premature termination codon (if the fusion protein of the invention comprises the amino acid sequence of RF2), preferably UAG, in its encoding sequence,
b. Detecting the phenotypic changes produced in the bacterium as a consequence of the expression of the reporter gene, and
c. Classifying the compound under study as an inhibitor of amyloid peptide aggregation when no phenotypic changes associated with the expression of the reporter gene are visualised in step (b).

When the compound under study presents the capacity to inhibit amyloid protein aggregation, the aggregates of the factors RF1 or RF2 included in the fusion protein of the invention produced as a consequence of the amyloidogenic peptide will be solubilised or modified; consequently, said factors will be functional, and the premature stop codon located in the reporter gene will be recognised. Thus, a

complete, functional reporter protein will not be produced, and its activity or expression will not be detected. When said reporter gene is the *lacZ* gene, the colour of the bacterial colonies will be the normal one for them (white-pale yellow in the case of *E. coli*), since a functional enzyme beta-galactosidase is not produced in the presence of an inhibitor of amyloid aggregation.

In a preferred embodiment of the second method of the invention, the reporter gene is the *lacZ* gene that encodes beta-galactosidase; in step (b), the presence or absence of blue colouration in the bacteria is detected; and, in step (c), the compound under study is classified as an inhibitor of amyloid peptide aggregation when the blue colouration is not visualised in step (b).

The compound under study in the second method of the invention may be any molecule, drug, compound, substance, composition and similar, of both synthetic and natural origin.

In a more preferred embodiment of the second method of the invention, the peptide with amyloidogenic potential comprised in the fusion protein expressed in the bacterial cell of the invention is selected from the peptides Aβ, Tau, α-synuclein, prion protein, huntingtin, superoxide dismutase 1, Fus, RepA-WH1 or TDP-43; even more preferably, it is the amyloidogenic peptide of the WH1 domain of RepA; even more preferably, it is SEQ ID NO: 1 or SEQ ID NO: 18.

Culturing of the bacterial cell of the invention in the first and/or the second method of the invention may be performed, for example, without being limited thereto, in a liquid medium (for example, in multi-well plates) or a solid medium (for example, in agar plates). The culture media and conditions for the growth of bacterial cells are widely known to persons skilled in the art.

In a preferred embodiment, the conditions that allow for the expression of the fusion protein of the invention and the reporter gene are the presence de X-Gal and, more preferably, the additional presence of arabinose, glucose and IPTG, even more preferably, 0.02 mM IPTG, in the medium. Even more preferably, said conditions comprise culturing the cell at 42°C when the cell is *E. coli* strain MRA8ΔlacZ (*prfA1*^{ts}).

On the other hand, the detection of phenotypic changes in the bacterium, as indicated in step (b) of the first and the second method of the invention, refers to both the detection and/or quantification of changes in colour, fluorescence, luminescence, etc. produced in the bacterium as a consequence of the expression of the reporter protein, and the detection and/or quantification of the activity of the protein encoded by the reporter gene.

Measurement of protein activity may be performed directly or indirectly. Direct measurement refers to the measurement of the quantity or concentration of protein on the basis of a signal directly obtained from said protein, which correlates directly with the number of protein molecules present in the bacterium. The output signal – which may also be referred to as intensity signal – may be obtained, for example, by measuring a value of intensity for a chemical or physical property of the protein. Indirect measurement includes the measurement obtained from a secondary component (for example, a component other than the protein) or a biological measurement system (for example, the measurement of cellular responses, ligands, “tags” or enzymatic reaction products).

According to the present invention, the detection of the quantity of reporter protein in the bacterial cell of the invention may be performed by any method for determining the quantity of proteins known to persons skilled in the art. In a preferred embodiment, the detection of the quantity of reporter protein is performed by incubation with a specific antibody, in assays such as Western blot, electrophoresis in gels, immunoprecipitation, protein arrays, immunofluorescence, immunohistochemistry, ELISA or any other enzymatic method linked to the generation of colour, fluorescence or luminescence; by means of incubation with a specific ligand; by means of NMR or any other diagnostic imaging technique; or, for example, by means of chromatographic techniques combined with mass spectrometry. The detection and/or quantification of the reporter protein may be performed by means of any of the aforementioned techniques or any combination thereof. The protein may be detected by evaluating its presence or absence. The detection may be performed by means of the specific recognition of any protein fragment by means of any probe and/or any antibody.

Throughout the description and the claims, the word "comprises" and variants thereof are not intended to exclude other technical characteristics, additives,

components or steps. For persons skilled in the art, other objects, advantages and characteristics of the invention will arise, partly from the description and partly from the practice of the invention. The following examples and figures are provided for illustrative purposes, and are not intended to limit the scope of the present invention.

DESCRIPTION OF THE FIGURES

FIG. 1. **(A)** Diagram of the WH1(R_{0-4})-RF1 chimeras. **(B)** It shows the amyloid potential of each of the repeats of the RepA amyloidogenic peptide (SEQ ID NO: 1, LVLCAVSLI), as predicted by the WALTZ algorithm. GRLVL (SEQ ID NO: 19) is the protein sequence, originating from the WH1 peptide, generated upon binding the DNA fragments that encode the WH1 repeats to the vector. This sequence is the same in all the constructs generated. In the DGR sequence shown in this figure, the Gly (G) amino acid is the spacer or linker, Asp (D) is the C-terminal residue of the preceding repeat and Arg (R) is the N-terminal residue of the following repeat.

FIG. 2. Construction of the MRA8Δ*lacZ* strain. Two homologous recombination steps were performed: the first by means of the Km-*parE* module flanked by the 5' and 3' ends of the *lacZ* gene **(A)**, and the second by the fusion of the flanking regions of the *lacZ* gene **(B and C)**.

FIG. 3. **(A)** Diagram of the *lacZ*-wt (top) and *lacZ*-amber (bottom) reporter gene constructs. **(B)** The induction of these constructs with 0.001% arabinose and 0.003% glucose results in blue colonies (in the figure, in the indicated grey scale) and white colonies, respectively, at 30ºC.

FIG. 4. Growth of serial dilutions of *E.coli* MRA8 Δ*lacZ* cells on plates in the presence (30ºC) or absence (42ºC) of cellular RF1. The complementation capacity of the WH1(R0) and the WH1(R3)-RF1 chimera was evaluated using different concentrations of the inducers thereof (IPTG) at 42ºC. In the case of the WH1(R0)-RF1 chimera, the basal escape of the P_{tac} promoter complemented the deficiency in cellular RF1. In the case of the WH1(R3)-RF1 chimera, inducer concentrations (IPTG) ranging between 0.01 and 0.05 mM were required to obtain a degree of complementation comparable to that of the physiological condition, and the quantity of inducer was set at 0.02 mM for the rest of the experiments.

FIG. 5. **(A**, top**)** Diagram of the function of RF1 according to the starting hypothesis: At 30ºC, since cellular RF1 is synthesised, it would be functional and efficient for translation termination at the UAG premature codon, thereby generating white colonies. At 42ºC, the amyloid peptide repeats would promote its aggregation, thereby preventing efficient translation termination at the stop codon, and resulting in the appearance of blue colonies (in the figure, in the indicated grey scale), as shown by the read-through assays in solid medium **(A**, bottom**)** and liquid medium **(B),** and the detected levels of β-galactosidase activity **(C)**.

FIG. 6. Levels of β-galactosidase activity, expressed in Miller units, of *E. coli* MRA8ΔlacZ cells at 3h **(A)** and 6h **(C)** post-induction of the WH1(R_{0-4})-RF1 chimeras and the lacZ-wt and amber reporter genes. **(B)** Kinetics of β-galactosidase activity in the WH1(R2)-RF1 chimera combined with the *lacZ*-wt reporter.

FIG. 7. **(A)** Western-blot: α-His (1/500) of *E. coli* MRA8Δ*lacZ* cells (0.2 OD units) induced with 1 mM of IPTG for 3 h. **(B)** Semi-denaturing agarose gel electrophoresis in the presence of detergent (SDD-AGE): α-His (1/500) of lysates (10 µl) of the same cells (25 ml at OD = 2). (* R0, denatured by boiling; α, anti).

FIG. 8. *In vivo* assay, on *E. coli* MRA8Δ*lacZ* cells, of the action of known inhibitors of amyloid aggregation. The presence of E3G (epigallocatechin-3-gallate) and, especially, resveratrol during the expression of the WH1(R2-4)-RF1 chimeras, combined with the *lacZ*-amber reporter, reverts the characteristic blue colouration phenotype (in the figure, in the indicated grey scale) (row 2) such that the cells become practically white (rows 5 and 6). Row 1 shows, as a control, MRA8Δ*lacZ* cells that express the different WH1(Rn)-RF1 chimeras combined with the *lacZ*-wt reporter, all of which are blue.

FIG. 9. Assay of polyphenolic inhibitors of amyloidosis: *lacZ*-WT controls. **(A)** Colouration in β-galactosidase assays in multi-well plates. **(B)** The levels of β-galactosidase activity (in the figure, in the indicated grey scale) were higher for those cells that expressed the WH1(R3)-RF1 chimera treated with resveratrol, which indicates a specific beneficial effect of resveratrol on those chimeras wherein the repeats of the RepA amyloidogenic sequence cause aggregation of RF1.

FIG. 10. **(A)** Assay of the β-galactosidase activity (in the figure, in the indicated grey scale) of the WH1(R0-4)-RF1 chimeras combined with the *lacZ*-amber reporter in the presence and absence of resveratrol. Treatment with resveratrol decreases β-galactosidase activity in the WH1(R2-4)-RF1 chimeras, which suggests a specific solubilising/anti-aggregating effect of resveratrol on those chimeras wherein the repeats of the RepA amyloidogenic sequence cause aggregation of RF1. **(B)** The results obtained correlate with those observed in phenotypic read-through assays on premature termination codons, where treatment with resveratrol recovers the colourless phenotype of those cells that express the chimeras combined with the *lacz*-amber reporter. **(C)** Growth curves of *E. coli* MDS42 cells at 37°C and 42°C in the presence and absence of resveratrol (100 µM), in order to evaluate the toxicity of the compound. Treatment with resveratrol does not affect the viability of *E. coli*. **(D)** SDD-AGE on MRA8Δ*lacZ* cell lysates expressing the WH1(R3)-RF1 chimera (1 mM IPTG, 3 h) in the presence and absence of resveratrol (500 µM, 1 mM). Treatment with the compound disaggregates the high-molecular-weight oligomers, whilst promoting an increase in the low-molecular-weight monomeric and oligomeric fractions.

EXAMPLES

Below we illustrate the invention by means of assays performed by the inventors, which demonstrate the effectiveness of the bacterial system for identifying amyloidogenic peptides and inhibitors of amyloidosis described in the present invention.

1. MATERIALS

1.1. Strains

For the cloning of bacterial genes, the *Fusion Blue* (*Clontech*) and MDS42 strains were used. Both lack the *recA* gene, which makes them homologous-recombination deficient, and this favours the stability of repeated sequences such as those included in some of the constructs used. The *Fusion Blue* strain presents the genotype *end*A1 *hsd*R17 (r_{K12}^- m_{K12}^+) *sup*E44 *thi*-1 *rec*A1 *gyr*A96, *rel*A1, lacF'[$proA^+B^+$, $lacI^q$ ZΔM15::Tn10(tet^R)]. The MDS42 strain presents the genotype MG1655 recA.

The experiments on growth complementation and read-through of stop codons were performed on the MRA8 and MRA8Δ*lacZ* strains, respectively, both of which are thermosensitive for the *prfA* gene that encodes the bacterial RF1 (release factor-1) translation termination factor. The choice of this strain was influenced by the fact that *prfA*/RF1 is essential in *E.coli* (Gerdes et al., 2003, *J Bacteriol 185*, 5673-5684). The MRA8 strain presents the genotype MG1655 $prfA^{ts}$. The MRA8Δ*lacZ* strain presents the genotype MG1655 $prfA^{ts}$ Δ*lacZ*. Both were used for the expression of the WH1(Rn)-RF1 prionoid chimera.

1.2. Culture media

LB: 10 g/l bactotryptone, 5 g/l yeast extract, 5 g/l NaCl (pH 7.4). It was routinely supplemented with thymine (2 g/l). In order to prepare the solid medium, it was supplemented with 1.5% (w/v) bacteriological agar.

M9+ CAA: 1/10 M9 10x (176.5 g/l $Na_2HPO_4 \cdot 12H_2O$, 30 g/l KH_2PO_4, 5 g/l NaCl, 10 g/l NH_4Cl), 1% Ca Mg (0.01 M $CaCl_2 \cdot 2\ H_2O$, 0.1 M $MgSO_4 \cdot 7H_2O$), 1 mg/ml vitamin B1, 2 g/l CAA, 0.4% glucose.

1.3. DNA

Plasmids

Name	Relevant genotype	Use
pFus	K^R, P_{BAD}, *araC*	Cloning vector *lacZ*
pMLM132	Cm^R, P_{parD}::*lacZ*	Amplification *lacZ*
pELI02	Am^R,P_{BAD}::*prfA*	Amplification *prfA*
pRG-SD1	Am^R, SD-1	Cloning vector *prfA*
pRG-WH1(R0-4)-RF1	Am^R, P_{tac}-SD-1-WH1(R0-4)-*prfA*	Amplification Ptac-SD-1-WH1(R0-4)-*prfA*
pRK2-WH1(R0-4)-RF1	Am^R, P_{tac}-SD-1-WH1(R0-4)-*prfA*	Expression WH1(R0-4)-RF1

pUKC1620-R0+WH1(R1-4)	His3, WH1(R1-4)	Extraction WH1(R1-4) by restriction
pKD46	Am^R, p_{ARA} λ red	Expression λ red
pKD267	Km^R, Km-*parE*	Amplification Km-*parE* module

Table 1. Plasmids used in the invention.

Oligonucleotides

The oligonucleotides used (Table 2) were prepared by solid-phase synthesis by means of phosphoramidite chemistry at the Protein Chemistry Service of the CIB [Biological Research Centre]. The synthesis of a 300-base-pair (bp) fragment with 3 repeats of the RepA-WH1(A31V) amyloidogenic sequence was commissioned from the company *ATG:biosynthetics*.

DNA sequencing

The sequencing was performed by the company *Secugen*, by means of the fluorescent chain termination process coupled with capillary electrophoresis. To this end, universal or specific oligonucleotide primers were designed and used (Table 2).

Name	Sequence	Use
RF1 5'	SEQ ID NO: 2 cgc**ccgcgg**cccgggcgatcgggaagctctggt tcatccggaaagccttctatcgttgcca	Amplification *prfA* gene
RF1 3'	SEQ ID NO: 3 cgc**ggatcc**ttattcctgctcggacaacgc	Amplification *prfA* gene
SpeI 5'	SEQ ID NO: 4 gct**actagt**tgacaattaatcatcggctcg	Amplification Ptac-WH1(R0)-RF1
LacZ 5' SpeI	SEQ ID NO: 5 ccc**actagt**atgaccatgattacggattcact	Amplification *lacZ* gene
LacZ 3' SmaI	SEQ ID NO: 6 cgc**cccggg**ttatttttgacaccagaccaact	Amplification *lacZ* gene
Delta LacZ 5'	SEQ ID NO: 7	Amplification of the Km-ParE

	atagtacataatggatttccttacgcgaaatacggg cagacatggcctgctctctacgccggacgcatcgt g	module with homologous ends to the *lacZ* gene starting from pKD267
Delta LacZ 3'	SEQ ID NO: 8 tatgttgtgtggaattgtgagcggataacaatttcac acaggaaacagctactgatcagtgataagctgtc	Amplification of the Km-ParE module with homologous ends to the *lacZ* gene starting from pKD267
Mut seq LacZ 5'	SEQ ID NO: 9 ggtggttgaactgcacaccg	Sequencing UAG mutation in *lacZ* gene
Mut UAA-UGA LacZ 5'	SEQ ID NO: 10 ggtctggtgtcaaaaat**g**acccggggatcctctag	Mutagenesis UAA termination codon of the *lacZ* gene by UGA
Mut UAA-UGA LacZ 3'	SEQ ID NO: 11 ctagaggatccccgggt**c**atttttgacaccagacc	Mutagenesis UAA termination codon of the *lacZ* gene by UGA
Amber LacZ 5'	SEQ ID NO: 12 gaccagcccttcccg**tag**gtgccgaaatggtcc	Mutagenesis UAG premature amber codon in the *lacZ* gene
Amber LacZ 3'	SEQ ID NO: 13 ggaccatttcggcac**cta**cgggaagggctggtc	Mutagenesis UAG premature amber codon in the *lacZ* gene
Mut PvuI RF1 5'	SEQ ID NO: 14 gctggggagtggcga**c**cgcagcgaccgtaac	Mutagenesis elimination PvuI target in the prfA gene
Mut PvuI RF1 3'	SEQ ID NO: 15 gttacggtcgctgcg**g**tcgccactccccagc	Mutagenesis elimination PvuI target in the prfA gene
Linker SmaI ½ R 5'	SEQ ID NO: 16 **ggg**cgcctagtgcta	Cloning WH1 half-repeat amyloid
Linker SmaI ½ R 3'	SEQ ID NO: 17 tagcactaggcg**ccc**	Cloning WH1 half-repeat amyloid

Table 2. Oligonucleotides used in the invention. The restriction enzyme targets and the bases changed by means of directed mutagenesis are shown in bold letters.

1.4. Proteins

For cell lysis, the glass matrix (1.0 mm φ, *Lysing Matrix* C) from *MP Biomedicals* was used.

Sequi-Blot PVDF filters and membranes from *BioRad* were used for protein

transfer. The molecular weight markers used were *Broad Range* (2-212 kDa), from *New England Biolabs*, and *Precision Blue Protein Standards* (10-250 kDa), from *BioRad.*

For the immunodetection of proteins, *ECL* luminescence kits *(Plus, Prime and Advanced)* from *GE Healthcare* were used. *AGFA Curix RP2* films were used for the developing.

1.5. Others

An EDTA-free protease inhibitor cocktail (1 tablet per 10 ml) from *Roche* was used. Moreover, the following reagents from the company *Sigma* were used: ONPG (4 mg/ml; 200 µl per β-galactosidase reaction), X-Gal (40 µg/ml), resveratrol (100 µM), curcumin (100 µM), quercetin (100 µM) and myricetin (37.5 µM).

2. METHODS

2.1. Construction of the WH1(R_{1-4})-RF1 chimeras

The *prfA* gene was amplified from the pELI02 plasmid (Table 1), with the internal PvuI target in *prfA* having been previously eliminated by means of directed mutagenesis, which does not alter the sequence of the encoded protein, RF1, using the oligonucleotides indicated in Table 2. In the case of the 5' end, in addition to the SacII site, which is necessary to clone the amplified *prfA* fragment, the SmaI and PvuI targets, which are necessary for the subsequent introduction of the tandem repeats of the RepA-WH1(A31V) amyloidogenic peptide (SEQ ID NO: 18, L_{26}VLCAVSLI_{34}), were included. The vector used was pRG-SD1, which carries a suboptimal translation initiation sequence (5 bp between the Shine-Dalgarno sequence and the ATG initiation codon) and expresses six-histidine fusion proteins under the control of the IPTG-inducible P_{tac} promoter. This vector was digested with the SacII and BamHI targets, which maintains the N-terminal His6 tag. The amplified fragment, previously digested with the same enzymes, was bound to the vector in order to obtain pRG-WH1(R0)-RF1. Starting from this plasmid, the Ptac-His6-R0-RF1 module was amplified with SpeI (5') and BamHI (3') ends (Table 2) and cloned into a plasmid with the RK2 replication origin, to obtain pRK2-WH1(R0)-RF1.

For the construction of the rest of the chimeras, a spacer DNA fragment that encodes a half repeat of the RepA-WH1(A31V) amyloidogenic peptide with SmaI (5') PvuI (3') ends was inserted. Given the difficulty inherent in the PCR amplification of

sequence repeats, the inserts with the tandem repeats of the amyloidogenic peptide (Figure 1) were obtained by means of enzymatic digestion (SmaI, EcoRV) and subsequent extraction of the pUKC1620 R1-R4 plasmids from polyacrylamide gels (10%), and bound to pRG-WH1(R0)-RF1 digested with SmaI, to generate the pRG-WH1(R1-4)-RF1 plasmids. Finally, once again by means of enzymatic digestion, the SmaI-BamHI inserts containing the R1, R2, R3 and R4 repeats of the RepA amyloidogenic peptide fused to the *prfA* gene were extracted from the pRG plasmids and bound to the pRK2-WH1(R0)-RF1 vector digested with SmaI and BamHI, to generate the series of pRK2-WH1(R1-4)RF1 vectors.

2.2. Construction of the pFus-*lacZ* (WT/amber) vectors

The *lacZ* gene was amplified from the pMLM132 plasmid (donated by Dr. Díaz-Orejas' laboratory) using specific nucleotides with SpeI (5') and SmaI (3') ends. Both the PCR product and the pFus vector (donated by Damián Lobato, from Dr. Díaz-Orejas' laboratory), a derivative of pBR322, were digested with said enzymes and both fragments were bound, to construct the pFus-*lacZ*-WT plasmid, which carried the *lacZ* gene under the control of the arabinose-inducible P_{ARA} promoter.

The pFus-*lacz*-amber plasmid was generated by means of directed mutagenesis using Pfu Turbo (*Stratagene)* and specific oligonucleotides to introduce a premature amber termination codon (UAG) at position A515 of the *lacZ* gene (A515*).

2.3. Construction of the MRA8Δ*lacZ* strain

Starting from a culture of *E.coli* MRA8 previously transformed by means of the pKD46 plasmid (Table 1), grown for 16 h at 30°C in liquid medium (LB supplemented with 50 μg/ml ampicillin), 2 ml of fresh culture (LB 50 μg/ml ampicillin) were inoculated, and it was incubated for 3 h at 30°C and 1100 rpm (*Thermo Mixer Compact*). Subsequently, the expression of the recombinase of the λ phage (λ red) encoded in pKD46 was induced, by adding 0.2% of arabinose, and grown for 2 h under the same conditions. The cells were centrifuged for 2 minutes at 11000 rpm and 4°C, and subjected to two washings, one with sterile H_2O and the other with sterile 10% glycerol. Following the second washing, they were resuspended in 20-30 μl of sterile 10% glycerol and transferred to an electroporation cuvette (*BioRad*, with a distance of 0.2 cm between the electrodes) and the PCR product was added thereto. The latter

contained the Km-*parE* module with 50-bp ends homologous to the 5' and 3' ends of the *lacZ* gene (Figure 2A). The "Ec2" micropulse was programmed (2.5 kV) in an electroporator (*MicroPulser, BioRad*). Following the electroporation, the cells were recovered by growing them in LB for 3 h at 30°C and 1100 rpm, and seeded in LB agar plates supplemented with kanamycin (50 μg/ml). The MRA8 Δ*lacZ*::Km-*parE* strain was thus obtained. Two 500-bp flanking regions of the *lacZ* gene were separately amplified by means of PCR. Both fragments were bound and the ligation product was amplified, to generate a 1-kbp fragment containing the preceding and the following region of the *E.coli* genome of the *lacZ* gene fused (Figure 2B). Thus, a second recombination step was performed on the MRA8 Δ*lacZ*::Km-*parE* strain, previously transformed by means of the pKD46 plasmid, following the same protocol described above, but using electroporation with the 1-kbp fragment (Figure 2C). In order to select the recombinant cells, it was plated in M9 agar supplemented with rhamnose (0.5%). Rhamnose induces the parE toxin; therefore, only those cells which have lost the Km-*parE* module would grow under those conditions, to generate the MRA8 Δ*lacZ* strain.

2.4. Engineering of the WH1(R_{1-4})-RF1 prionoid chimera.

2.4.1. Bacterial growth

All the assays were performed by expressing the WH1(R_{1-4})-RF1 chimeras in *E.coli* MRA8Δ*lacZ* (*prfA1* ^{ts}) cells, a strain wherein the *prfA* gene that encodes RF1 is soluble and functional at 30°C, but not at 42°C. To this end, cultures were grown in LB or M9+CAA liquid medium supplemented with ampicillin (100 μg/ml) and kanamycin (50 μg/ml) overnight at 30°C. On the following day, they were inoculated (1/100) into fresh medium and allowed to grow at 30°C until they reached an OD_{600nm} of 0.3.

2.4.2. Growth complementation assays in MRA8Δ*lacZ*

After growing the cultures at 30°C, serial dilutions thereof were prepared (10^{-1}-10^{-5}). Subsequently, 7-μl drops of each dilution were inoculated into LB plates with ampicillin, kanamycin and, only in the case of the plates to be grown at 42°C (to induce the factor present in the chimeras), different concentrations of IPTG. The plates were incubated at 30°C and 42°C overnight in order to evaluate the growth capacity of the MRA8Δ*lacZ* (*prfA* ^{ts}) cells expressing the different chimeras in the presence or absence of cellular RF1 (30°C and 42°C, respectively).

2.4.3. Premature termination codon read-through assays

For the assays performed on agar plates, after growing the cultures, serial dilutions thereof were prepared (10^{-1}-10^{-5}). Subsequently, 7-μl drops of each dilution were added to LB plates with ampicillin (100 μg/ml), kanamycin (50 μg/ml), IPTG (0.02 mM), arabinose (0.001%), glucose (0.003%) and X-Gal (40 μg/ml), and grown at 30°C and 42°C overnight in order to evaluate the read-through capacity of the different chimeras, which is determined by the appearance or non-appearance of colonies with a blue colouration.

For the assays in liquid medium, M9-CAA, a colourless medium, was used. The different cultures were grown at 30°C until they reached an OD_{600nm} of 0.3; at this time, 500 μl of each culture were added to a p24 multi-well plate (Falcon). Each well was supplemented with IPTG (0.02 mM), arabinose (0.001%), glucose (0.003%) and X-Gal (40 μg/ml), and the plates were grown at 42°C and 300 rpm (*Thermo Mixer Compact Eppendorf*) for 24 h. In the case of the assays with inhibitors of amyloid aggregation, the same protocol was used and the following inhibitors were added to the multi-well plate: curcumin (50 μg/ml), quercetin (100 μg/ml), epigallocatechin-3-gallate (100 μg/ml), resveratrol (100 μg/ml) and myricetin (37.5 μM), all of them from a 25 mM stock in DMSO.

2.4.4. β-galactosidase activity assays

After growing the cultures at 30°C, they were distributed into 8 aliquots and the expression of the enzyme β-galactosidase was induced (0.001% arabinose, 0.003% glucose); 4 of them were allowed to grow at 30°C and 4 of them at 42°C in the presence of 0.02 mM of IPTG. The β–galactosidase activity was evaluated by means of colorimetry (degradation of ONPG) at 28°C, at 3 and 6 h after the induction. This method quantifies the β-galactosidase activity using the following formula:

$$\text{Miller units} = 1000 \times ((A_{420} - (1.75 \times A_{550})) / (t(\text{min}) \times V(\text{ml}) \times OD_{600}))$$

A_{420}: absorption of the ONPG degradation compounds; (1.75 x A_{550}): light scattering correction factor at 420 nm; OD_{600}: number of cells.

At least 3 independent experiments were performed for each chimera. For the assays at 3 and 6 h performed in LB, the following volumes and reaction times were

used: for the chimeras combined with the *lacZ*-WT reporter, 200 μl of culture taken to a final volume of 1 ml with Z buffer and a reaction time of 5 minutes; in the case of the chimeras combined with the *lacZ*-amber reporter, 8 ml of cells were used, which were resuspended in 1 ml of Z buffer for a reaction time of 15 minutes. In the case of the assays in M9+CAA medium, 200 μl of culture and a reaction time of 5 minutes were used in the combinations with *lacZ*-WT, and 300 μl of culture and a reaction time of 15 minutes were used in the combinations with the *lacZ*-amber reporter.

In the case of the activity measurements for the cells recovered from the multi-well plate, the evaluation was performed at 24 hours post-induction, in order to correlate it with the appearance or non-appearance of a blue colouration. At least 6 replicas of each experiment were performed.

2.4.5. Detection of the chimeras: Western blot and SDD-AGE.

The biochemical determination of the synthesis of the WH1(R_{1-4})-RF1 fusion proteins was performed by means of Western blot, using the His6 tag located at the N-terminal end of each of the chimeras as the epitope, after inducing the expression thereof for 3 h with 1 mM of IPTG. Given the tendency of RF1 to form complexes with the ribosomal proteins, the electrophoresis (SDS-PAGE) was performed in an 8% acrylamide gel, adding 6 M of urea to the loading buffer. The gel was transferred to a PVDF membrane by means of wet transfer (*Mini Trans-blot*, *BioRad*). For the detection by means of Western blot, an α-His antibody from Sigma (1/1000) and an α-mouse secondary antibody (1/10000) were used.

The presence of SDS-resistant amyloid oligomers was evaluated by means of semi-denaturing agarose gel electrophoresis in the presence of detergent (SDD-AGE). For the detection, an α-His antibody from Sigma (1/500) and an α-mouse secondary antibody (1/5000) were used.

EXAMPLE

3.1. The WH1(R_{1-4})-RF1 chimeras complement the *prfA* deficiency in *E. coli* MRA8Δ*lacZ*.

A completely synthetic reporter system was developed using the RF1 protein fused to tandem repeats of the RepA-WH1 amyloidogenic peptide (Figure 1).

In order to evaluate the amyloid potential of said fusions in *E. coli*, the *lacZ* gene under the control of an arabinose-inducible P_{ara} promoter was used as the reporter gene. The wild-type version (WT) of the gene was cloned, or it carried a premature amber termination codon (UAG) (Figure 3A), to generate blue or white colonies, respectively, in the presence of X-Gal, under the induction conditions used in this work (0.001% arabinose; 0.003% glucose) (Figure 3B).

The complementation assays were performed on the MRA8Δ*lacZ* strain of *E.coli*, which is thermosensitive to the *prfA* gene, which encodes RF1. Therefore, this strain is not viable at 42ºC in the absence of RF1 complementation. Thus, the growth complementation capacity when expressing the WH1(R_{1-4})-RF1 chimeras at 42ºC was evaluated using different concentrations of inducer (IPTG: 0; 0.01; 0.05; 0.1; 0.15; 0.2; 0.25 mM). The basal escape of the P_{tac} promoter was sufficient to complement the deficiency of cellular RF1 in the case of the WH1(R0)-RF1 chimera, which indicates that the basal levels of RF1 are low inside the cell at 30ºC. In the case of the WH1(R3)-RF1 chimera, inducer concentrations ranging between 0.01 and 0.05 mM were required to obtain a degree of complementation comparable to that of the physiological state, and the induction conditions were finally set at 0.02 mM (Figure 4).

These results suggest the use of low inducer concentrations to obtain expression levels of RF1 comparable to physiological ones. The *in vivo* levels of RF1 per cell were determined by Adamski et al., being estimated at between 1200 and 4900 RF1 molecules per cell, five times less than in the case of RF2 (5900-24900 molecules per cell).

3.2. The expression of the WH1(R_{2-4})-RF1 chimeras results in the read-through of a UAG premature termination codon.

The appearance of colonies with a blue colouration when expressing the WH1(R_{1-4})-RF1 chimeras and the *lacZ* reporter genes (WT/amber) was qualitatively evaluated, both in solid medium (LB agar) and in liquid medium (M9+CAA), in the presence of X-Gal. Both approaches show that, whereas the expression of the WH1(R0)-RF1 and WH1(R1)-RF1 chimeras combined with the *lacZ*-amber reporter at 42ºC produces colonies with a white colouration, the expression of WH1(R_{2-4})-RF1 at

the same temperature generates colonies with a blue colouration, which indicates a defect in the translation termination capacity of the latter (Figure 5A,B).

However, the same does not occur in the case of the expression of the chimeras combined with the *lacZ*-amber reporter at 30°C. In this case, in which translation termination is not only dependent upon the WH1(Rn)-RF1 chimera, since functional RF1 is being produced from the chromosome, all the colonies have a white colouration. As a control, we show the expression of the chimeras combined with the *lacZ*-WT reporter, which generates colonies with a blue colouration at both 30°C and 42°C (Figure 5A).

Subsequently, the read-through capacity was quantitatively evaluated by measuring β-galactosidase activity, a product of the *lacZ* gene. Thus, after determining the β-galactosidase activity obtained upon combining each of the chimeras with the *lacZ*-WT reporter and the *lacZ*-amber reporter at 42°C, the percentage of β-galactosidase activity obtained for the amber reporter as compared to the WT reporter was calculated for each construct (considering 100% of translation termination to be the one measured for each of the chimeras combined with the *lacZ*-WT reporter) (Figure 5C). Once again, this result suggests a defect in translation termination at the UAG premature codon of the *lacZ*-amber reporter, since the relative levels of enzymatic activity varied between 5% and 15% for the WH1(R_{2-4})-RF1 chimeras, depending on the case, whereas, in the case of WH1(R0) and WH1(R1)-RF1, practically no β-galactosidase activity was detected; this allows us to deduce that these chimeras are capable of terminating the translation in a similar manner to the wild-type RF1 protein.

These results indicate that the capacity to promote the read-through of the UAG termination codon, present in the *lacZ*-amber reporter, is dependent upon the presence of repeats of the RepA-WH1 amyloidogenic peptide, 2 being the minimum number of sequence repeats required in this case to obtain a factor aggregation phenotype.

The values obtained for β-galactosidase activity in MRA8Δ*lacZ* at 42°C for the WH1(R_{0-1})-RF1 chimeras were less than 1%, whereas, in the presence of two or more repeats of the RepA-WH1 amyloidogenic peptide, the levels varied between 5%-15%, which suggests that the effect on the functionality of RF1 is dependent upon the

presence of two or more repeats of the amyloid sequence. Thus, the system developed in this invention (the WH1(Rn)-RF1 prionoid chimeras) constitutes the first synthetic system, based on the modulation of aggregation-dependent protein translation termination, for the *in vivo* identification of sequences with amyloid potential in *E.coli*.

3.3. The expression of WH1(R_{2-4})-RF1 produces metabolic slowdown in *E. coli* MRA8 Δ*lacZ* at 42ºC.

The measurements of β-galactosidase activity were initially performed 3 h after the induction of the expression of the WH1(Rn)-RF1 chimeras and the *lacZ*-WT/amber reporters, at both 30ºC and 42ºC. The results obtained are shown in Figure 6A. For the WH1(R_{0-1})-RF1 chimeras combined with the *lacZ*-WT reporter, the levels of activity were similar at both 30ºC and 42ºC, which suggests that the range of inducer concentration for the chimeras (0.02 mM) is the suitable one, since similar levels to those obtained with cellular RF1 were measured (30ºC). However, the same did not occur in the case of WH1(R_{2-4})-RF1, for which the levels of activity dropped at 42ºC.

In view of this result, we considered the possibility that a second read-through event might be taking place at the natural termination codon of the *lacZ* reporter, since it is a UAA codon that is also recognised, although not exclusively, by RF1 as a translation terminator.

In order to evaluate this possibility, the UAA termination codon of the *lacZ* reporters was mutated with UGA, a termination codon that is only recognised by release factor 2 (RF2), and the β-galactosidase activity was assessed; the same result was obtained (data not shown).

Since the defect in the functionality of the WH1(R_{2-4})-RF1 chimeras not only affects the reporter gene, but all the cell proteins with the UAG termination codon as well, the possibility exists that the cell metabolism slows down, thereby affecting protein synthesis. In order to evaluate this alternative, we performed a kinetic analysis of β-galactosidase activity using the WH1(R2)-RF1 chimera and the *lacZ*-WT reporter gene at both 30ºC and 42ºC. The results obtained show that, whereas, in the case of expression at 30ºC, the levels of activity rapidly increase at 3 h of induction, to gradually decrease thereafter, expression at 42ºC produced a slower, more

progressive accumulation of the enzyme with time, and the levels of activity detected at 30°C and 42°C became the same at 6 h post-induction (Figure 6B).

Finally, we evaluated the β-galactosidase activity at 6 h post-induction, and similar values were obtained for each of the chimeras, at both 30°C and 42°C (Figure 6C). The percentage of relative activity for the amber reporter as compared to the WT reporter was calculated at both 3 and 6 h post-induction, and is shown in Figure 5C.

The kinetic analysis of β-galactosidase activity performed with the WH1(R2)-RF1 chimera combined with the *lacZ*-WT reporter (Figure 6B) showed a delay in the synthesis and degradation of the enzyme at 42°C. This effect may be partly due to growth at 42°C, since temperature affects the protein synthesis, folding and aggregation process that may be essential for cell viability. However, the result observed must be a consequence of a combination of the temperature and the presence of repeats of the RepA-WH1(A31V) amyloidogenic peptide, since the decrease in β-galactosidase activity obtained (in all cases at 42°C) was not comparable for the WH1(R_{0-1})-RF1 and the WH1(R_{2-4})-RF1 chimeras combined with the *lacZ*-WT reporter (Figure 6A). This suggests an effect on the protein synthesis and degradation machinery that is dependent upon the presence of repeats. Aggregation of the factor and, consequently, the decrease in the levels of available RF1 would induce a larger number of ribosomes to pause on the mRNAs with the UAG termination codon. In the genome of *E. coli*, it has been estimated that said genes constitute 7% of the total. The incorrect pause on termination codons is recognised by the tmRNA-SmpB regulatory machinery, a protein synthesis quality control system; in order to recycle paused ribosomes, it labels those proteins that have not been completely synthesised such that they may undergo proteolytic degradation through a process called *trans*-translation.

3.4. The read-through of the UAG premature termination codon is due to the aggregation of the WH1(R_{2-4})-RF1 chimeras.

The biochemical determination of the WH1(R_{0-4})-RF1 fusion protein synthesis levels was performed by means of Western blot, using the His6 tag located at the N-terminal end of each of the chimeras as the epitope. All the chimeras were detected, as shown in Figure 7A. In addition to the bands pertaining to the size (approximately 50 kDa) of the fusion protein monomers, a proteolysis band and an aggregated fraction, retained close to the wells, were detected at 25 kDa. The aggregation may be a

consequence of contacts between RF1 molecules through the amyloid segment and/or interactions of RF1 with ribosomal proteins, since the aggregated fraction is also detected in the WH1(R0)-RF1 chimera. The interaction of RF1 and RF2 with different ribosomal proteins (L2, L7/L12, L11, L16, S3, S4, S5, S10 and S18) throughout the translation termination process has already been described. It has also been described that RF1 has a greater affinity for the ribosomal A site than RF2, which would explain why RF1 co-purifies with the ribosomes, which does not occur in the case of RF2. In fact, the protocols described for the identification of RF1 from total lysed cells by means of Western blot includes the use of chaotropic agents, such as urea, to ensure the complete denaturation of the protein, given its tendency towards aggregation. Moreover, purification protocols involve a first step designed to isolate the ribosomal fraction prior to the chromatographic purification of RF1.

Taking into consideration the difference in the read-through capacity in the phenotypes observed, the aggregates formed by RF1 in the absence of repeats should be structurally and biochemically different from those generated in the WH1(R_{2-4})-RF1 chimeras. The presence of SDS-resistant amyloid oligomers was evaluated by means of semi-denaturing agarose gel electrophoresis in the presence of detergent (SDD-AGE) (Figure 7B). Once again, also by means of this technique, the monomers of the different WH1(Rn)-RF1 chimeras were detected, as were three different types of oligomers: some lower-molecular-weight ones that were absent in the WH1(R0)-RF1 chimera; others, specific to RF1, that were present in all the chimeras; and, finally, oligomers with a higher molecular weight that were only detectable in the WH1(R_{1-4})-RF1 chimeras, but not in the WH(R0)-RF1 construct. The intensity of the latter increased with the number of repeats of the amyloidogenic sequence present in the chimeras (Figure 7B).

3.5. Treatment with Resveratrol recovers the translation termination function in the WH1(R_{2-4})-RF1 chimeras.

Once the screening system for the detection of amyloid sequences based on the WH1(Rn)-RF1 chimeras was developed, we decided to assay different polyphenolic compounds previously described as *in vitro* inhibitors of amyloid protein aggregation. The molecules assayed were: quercetin, curcumin, epigallocatechin-3-gallate (E3G), resveratrol and myricetin. The assays were performed in liquid medium

and multi-well plates, at 42°C, as described in the Methods section, 2.4.3. The results obtained are shown in Figure 8.

Of all the compounds used, only E3G and resveratrol seemed to have an influence on the colouration phenotype: the MRA8Δ*lacZ* cells that expressed the WH1(R_{2-4})-RF1 chimeras combined with the *lacZ*-amber reporter went from having a blue colouration, when grown in the presence of X-Gal and in the absence of inhibitors, to having a reduced colouration intensity (in the case of E3G) or even becoming practically white (in the presence of resveratrol). No effect whatsoever was observed with myricetin. Although curcumin seemed to revert the aggregation phenotype in the WH1(R2)-RF1 chimera, its use was discarded due to its intense yellow colouration and its limited solubility in the medium used (M9+CAA). In the case of quercetin, the assay repetitions showed a great variability for this compound, which was capable of reducing the colouration intensity to a point comparable to that of E3G depending on the assay (data not shown); for this reason, it was included in the following screening step: the *lacZ*-WT reporter controls.

In order to discard the possibility that the effect on the colouration was caused by a direct action on the enzyme β-galactosidase, and not on the aggregation of the chimeras, controls were performed that expressed the WH1(R0) and WH1(R3)-RF1 chimeras, combined with the *lacZ*-WT reporter, in the presence and absence of E3G, resveratrol and quercetin. Moreover, the β-galactosidase activity in the cells was evaluated after 24 h of induction.

The results obtained showed both a decrease in the levels of β-galactosidase activity (Figure 9B) and a reduction in colouration (Figure 9A) in those cells that expressed the WH1(R0) and WH1(R3)-RF1 chimeras combined with the *lacZ*-WT reporter in the presence of quercetin and E3G. These results suggest a direct inhibition of β-galactosidase activity by both polyphenols. However, the presence of resveratrol during the induction of the chimeras did not produce a decrease in colouration, which suggests that the effect observed in the WH1(R_{2-4}) chimeras with the *lacZ*-amber reporter was not produced on the enzyme β-galactosidase, but on the aggregation of the chimeras (Figure 9A). Moreover, in those cells treated with resveratrol, the β-galactosidase activity measurements showed a reduction in activity for the WH1(R0)-RF1 chimera, but not for the WH1(R3)-RF1 chimera, for which the values were even

slightly higher than those of the control treated only with 0.4% DMSO (the solvent used for the inhibitors) (Figure 9B). This result may indicate a specific beneficial effect of resveratrol on those chimeras wherein the repeats of the RepA amyloidogenic sequence were causing the aggregation of RF1, thereby allowing for termination at the UAA natural stop codon as compared to the control, but not in those chimeras wherein aggregation was not taking place, such as WH1(R0), for which the treatment with resveratrol seemed to present a toxicity of indeterminate nature.

In order to confirm this hypothesis, the same experiment was performed, expressing the different chimeras combined with the *lacZ*-amber reporter in the presence and absence of resveratrol (Figure 10A,B). The β-galactosidase activity was normalised for each construct and, in each case, 100% activity was considered to be that observed in the absence of resveratrol. The results obtained show that, in the case of the WH1(R0) and WH1(R1)-RF1 chimeras, wherein no change in colouration from white to blue is observed in the premature termination codon read-through assays, the levels of β-galactosidase activity increase in the presence of resveratrol, which may indicate a greater toxicity of the compound on these chimeras through the promotion of an increased aggregation thereof (Figure 10). However, given the low absolute levels of enzymatic activity detected (< 1 Miller unit), we might be within the margins of error for the assay (Table 3). However, in the case of the WH1(R_{2-4})-RF1 chimeras, the treatment with resveratrol produced a decrease in β-galactosidase activity of about 56%, 50% and 30%, respectively (with respect to each control without the treatment), which suggests a beneficial effect of the compound, probably by promoting the solubilisation of RF1, which increases the translation termination capacity at the UAG premature codon.

	β-galactosidase activity (Miller units)				
	R0	**R1**	**R2**	**R3**	**R4**
DMSO	0.24±0.08	0.36±0.21	11.87±1.35	4.88±0.40	5.50±0.80
Resveratrol (100 μM)	0.47±0.09	0.60±0.22	5.20±1.31	2.44±0.35	3.80±0.55

Table 3: Absolute values of β-galactosidase activity in MRA8Δ*lacZ* cells expressing the WH1(Rn)-RF1 chimeras combined with the *lacZ*-amber reporter at 24 h post-induction

in the presence and absence of resveratrol. The activity values for the WH1(R_{2-4})-RF1 chimeras decrease by 56%, 50% and 30%, respectively. However, although the levels increase for the WH1(R_{0-1})-RF1 chimeras, they are still low enough (less than 1 Miller unit) so as not to be considered significant.

In order to clarify the effect exerted by resveratrol on the WH1(R_{0-1})-RF1 chimeras, growth curves were performed for the *E.coli* strain MDS42 in the presence and absence of resveratrol (100 μM), so as to evaluate the toxicity thereof on their growth. The growth curves shown in Figure 10C indicate that there were no differences in the generation time of MDS42 in the presence or absence of resveratrol at 37°C or at 42°C; therefore, the toxic effect observed on the MRA8Δ*lacZ* strain that expresses the WH1(R0)-RF1 chimera (Figure 9B) seems to be dependent upon that genetic background, such that resveratrol may contribute to increasing the intrinsic aggregation of RF1, rather than being a general effect on the metabolism of *E.coli* K12, since a reduction in biomass is not produced. However, taking into consideration the synthesis and degradation kinetics of β-galactosidase shown in Figure 6B, it is possible that the decrease in enzymatic activity in the WH1(R0)-RF1 chimera combined with the *lacZ*-WT reporter is a consequence of a generalised increase in the solubility of the proteins at 42°C, arising from the action of resveratrol, which promotes enzyme degradation more efficiently than when the compound is absent.

Moreover, in order to corroborate the anti-aggregating effect of resveratrol on the WH1(R_{2-4})-RF1 chimeras, an SDD-AGE biochemical analysis was performed from cell lysates that expressed the WH1(R3)-RF1 chimeras grown in the presence and absence of resveratrol (500 μM and 1 mM), and induced with 1 mM of IPTG. The results obtained show a significant reduction in high-molecular-weight oligomers, as well as an increase in monomeric and oligomeric fractions of a smaller size, depending on the concentration of resveratrol used. However, the oligomeric fraction specific to RF1 was not altered (Figure 10D). This result confirms that the effect on the colouration phenotype observed is in fact due to a solubilising effect of the polyphenol on those chimeras wherein the repeats of the RepA-WH1 amyloidogenic peptide promote the aggregation of RF1.

The results obtained in this invention demonstrate, *in vivo*, the capacity of the polyphenol to solubilise and/or remodel oligomers on the WH1(R_{2-4})-RF1 chimeras, in a

minimalist model system, *E. coli*. The SDD-AGE assays performed in the presence and absence of resveratrol confirm that the compound presents a solubilising effect on high-molecular-weight oligomers, thereby contributing to increase the low-molecular-weight monomeric and oligomeric fractions. Therefore, this result proves that the system developed herein not only monitors the amyloid potential of a given peptide sequence *in vivo*, but, moreover, is useful to evaluate inhibitory compounds of amyloid aggregation, such as resveratrol. Furthermore, the system makes it possible to discard false positives, since it allows not only for a colorimetric evaluation of the reporter gene, but also an enzymatic evaluation, as well as for a quick, simple, economical *in vivo* characterisation of the dose-dependent toxicity of the different compounds.

CLAIMS

1. A fusion protein that comprises:

   a. at least one copy of an amino acid sequence of a peptide with amyloidogenic potential, and
   b. the amino acid sequence of the bacterial RF1 translation termination factor,

   where the C-terminal end of the amino acid sequence of (a) is bound to the N-terminal end of the amino acid sequence of (b).

2. The fusion protein according to claim 1, wherein the amino acid sequence of (a) is of human, non-human animal, vegetable, fungal or bacterial origin.

3. The fusion protein according to any one of claims 1 or 2, wherein the peptide with amyloidogenic potential is selected from peptides Aβ, Tau, α-synuclein, prion protein, huntingtin, superoxide dismutase 1, Fus, RepA-WH1 or TDP-43.

4. The fusion protein according to any one of claims 1 to 3, wherein the amino acid sequence of (b) is bound to 2, 3 or 4 copies of the amino acid sequence of (a).

5. A nucleotide sequence that encodes the fusion protein according to any one of claims 1 to 4.

6. An expression vector that comprises the nucleotide sequence according to claim 5.

7. The expression vector according to claim 6, wherein said vector is a plasmid.

8. A bacterial cell that expresses the fusion protein according to any one of claims 1 to 4, wherein said cell further comprises a reporter gene that comprises a UAG premature termination codon in its encoding sequence.

9. The bacterial cell according to claim 8, wherein the reporter gene is the *lacZ* gene that encodes beta-galactosidase.

10. The bacterial cell according to any one of claims 8 or 9, wherein the reporter gene is comprised in an expression vector.

11. The bacterial cell according to any one of claims 8 to 10, wherein said cell further presents inactivation of said wild-type reporter gene in its chromosome.

12. The bacterial cell according to any one of claims 8 to 11, wherein said cell further presents inactivation of the wild-type gene that encodes the bacterial RF1 translation termination factor in its chromosome.

13. The bacterial cell according to any one of claims 8 to 12, wherein said cell is *Escherichia coli*.

14. Use of the bacterial cell according to any one of claims 8 to 13 for the identification of amyloidogenic peptides.

15. Use of the bacterial cell according to any one of claims 8 to 13 for the identification of inhibitory compounds of amyloid peptide aggregation.

16. A method for identifying amyloidogenic peptides, which comprises:

a. Culturing the bacterial cell according to any one of claims 8 to 13 under conditions that allow for the expression of the fusion protein and the reporter gene that comprises a UAG premature termination codon in its encoding sequence,
b. Detecting the phenotypic changes produced in the bacterium as a consequence of the expression of the reporter gene, and
c. Classifying the peptide with amyloidogenic potential as an amyloidogenic peptide when phenotypic changes associated with the expression of the reporter gene are visualised in step (b).

17. The method according to claim 16, wherein the reporter gene is the *lacZ* gene that encodes beta-galactosidase; in step (b), the presence or absence of a blue colouration in the bacteria is detected; and, in step (c), the peptide with

amyloidogenic potential is classified as an amyloidogenic peptide when said blue colouration is visualised in step (b).

18. A method for identifying inhibitory compounds of amyloid peptide aggregation, which comprises:

    a. Culturing the bacterial cell according to any one of claims 8 to 13 in the presence of the compound under study under conditions that allow for the expression of the fusion protein and the reporter gene that comprises a UAG premature termination codon in its encoding sequence,
    b. Detecting the phenotypic changes produced in the bacterium as a consequence of the expression of the reporter gene, and
    c. Classifying the compound under study as an inhibitor of amyloid peptide aggregation when no phenotypic changes associated with the expression of the reporter gene are observed in step (b).

19. The method according to claim 18, wherein the reporter gene is the *lacZ* gene that encodes beta-galactosidase; in step (b), the presence or absence of a blue colouration in the bacteria is detected; and, in step (c), the compound under study is classified as an inhibitor of amyloid peptide aggregation when said blue colouration is not visualised in step (b).

20. The method according to any one of claims 18 or 19, wherein the peptide with amyloidogenic potential comprised in the fusion protein expressed in the bacterial cell is selected from the peptides Aβ, Tau, α-synuclein, prion protein, huntingtin, superoxide dismutase 1, Fus, RepA-WH1 or TDP-43.

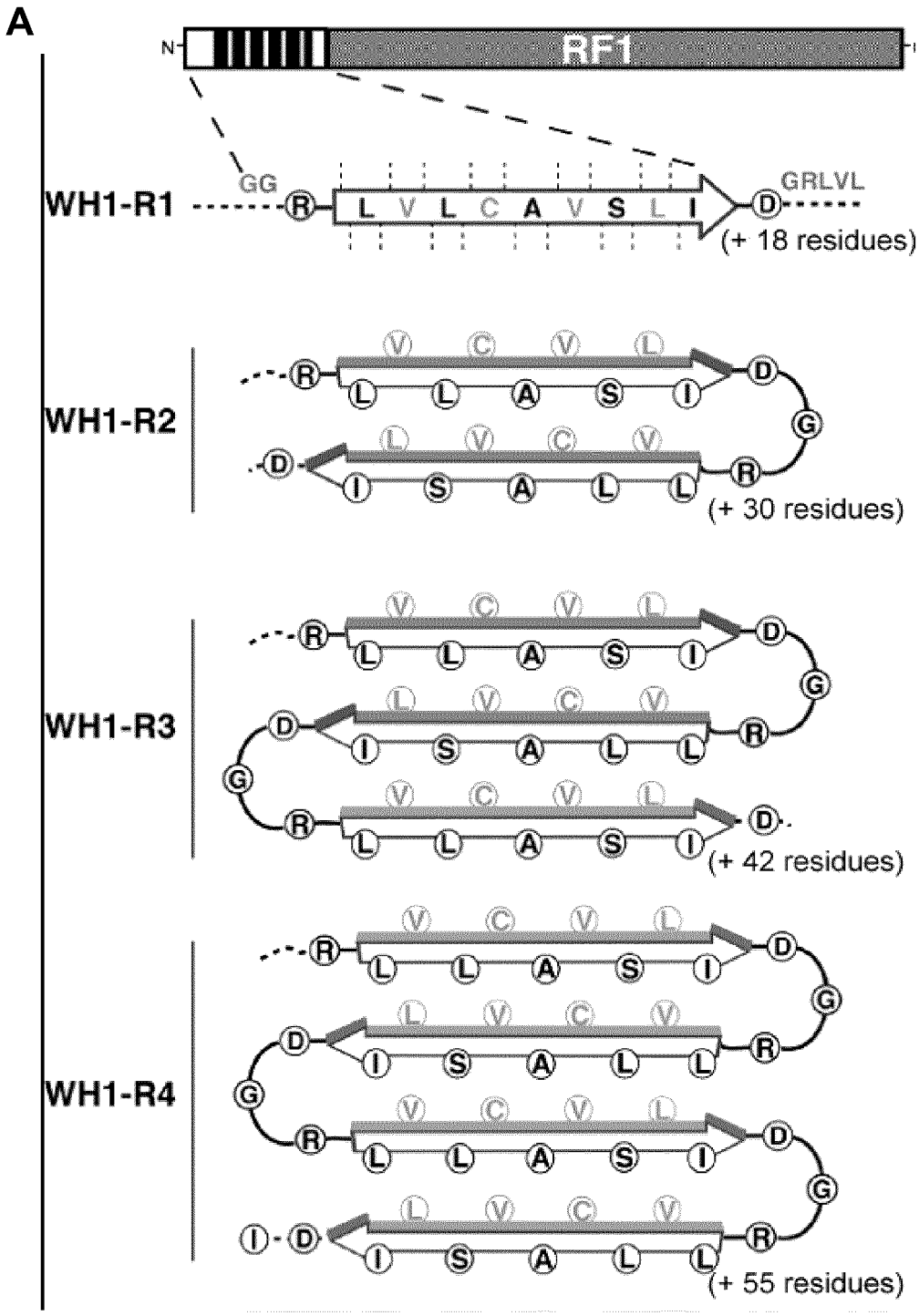

A
N
RF1
C
WH1-R1
GG
R
L V L C A V S L I
D
GRLVL
(+ 18 residues)
WH1-R2
(+ 30 residues)
WH1-R3
(+ 42 residues)
WH1-R4
(+ 55 residues)

FIG. 1

B

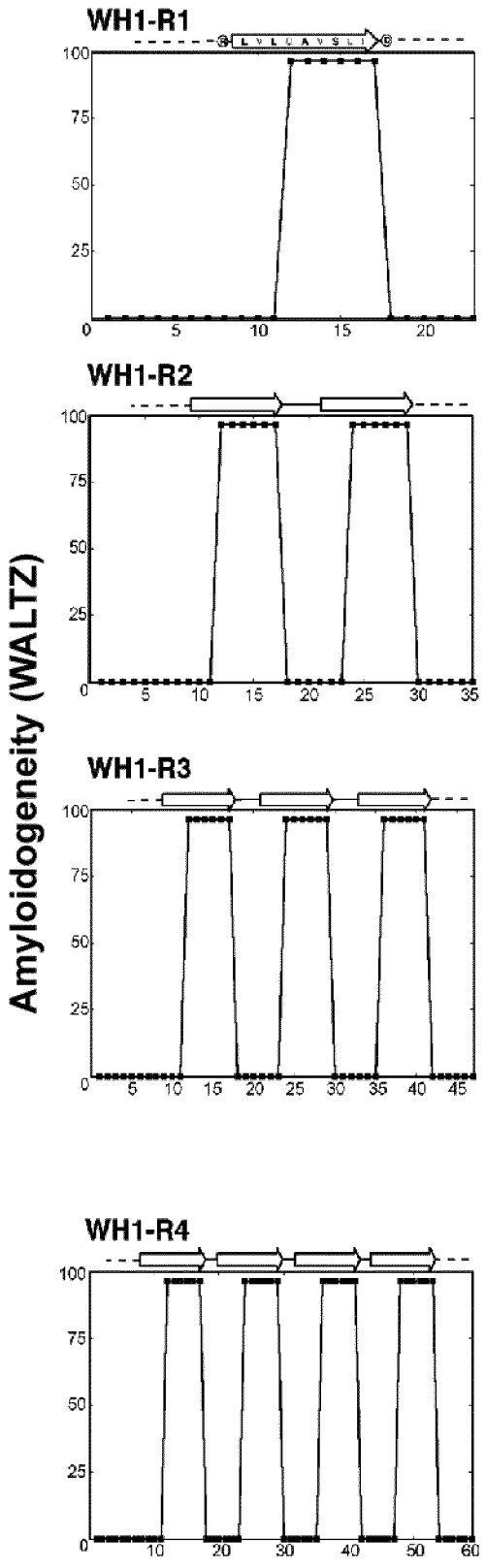


FIG. 1 (cont.)

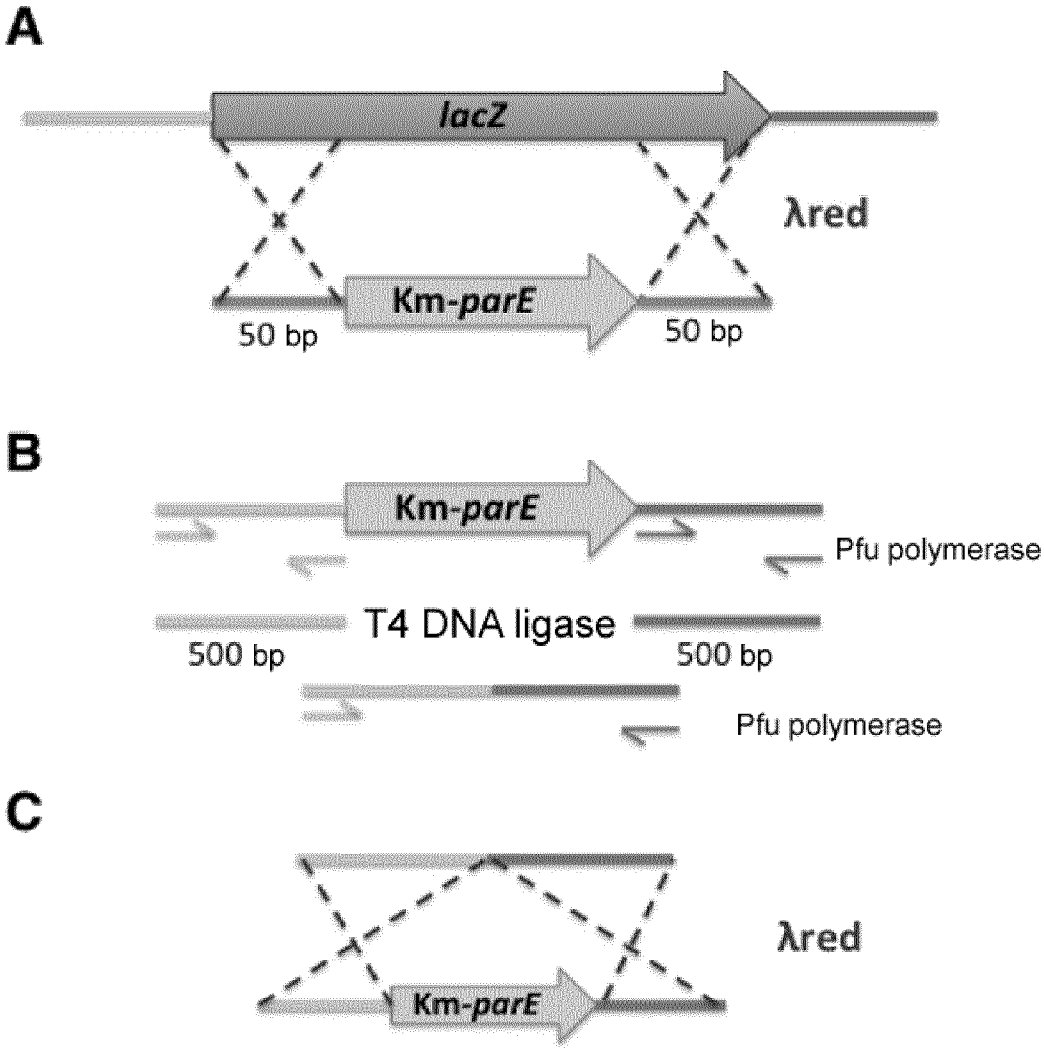

A
lacZ
λred
Km-parE
50 bp
50 bp
B
Km-parE
Pfu polymerase
T4 DNA ligase
500 bp
500 bp
Pfu polymerase
C
λred
Km-parE

FIG. 2

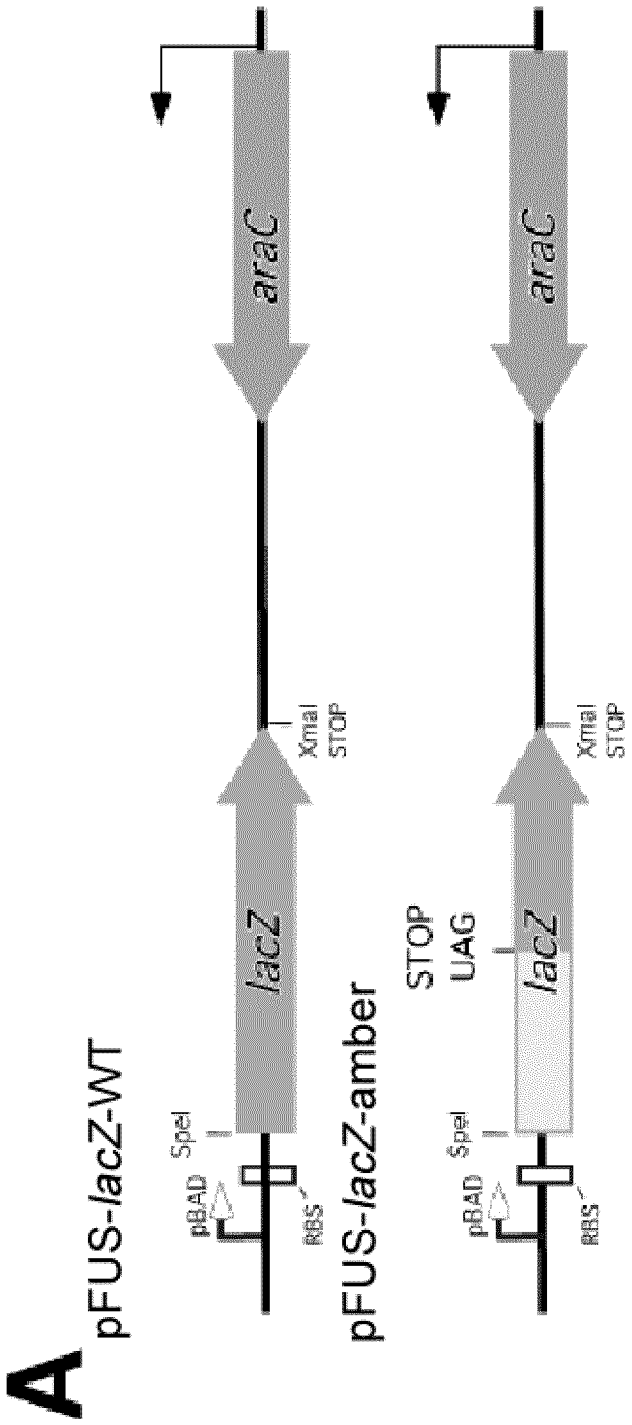

A
pFUS-lacZ-WT
araC
lacZ
SpeI
pBAD
RBS
XmaI
STOP
pFUS-lacZ-amber
STOP
UAG
araC
lacZ
SpeI
pBAD
RBS
XmaI
STOP

FIG. 3

B

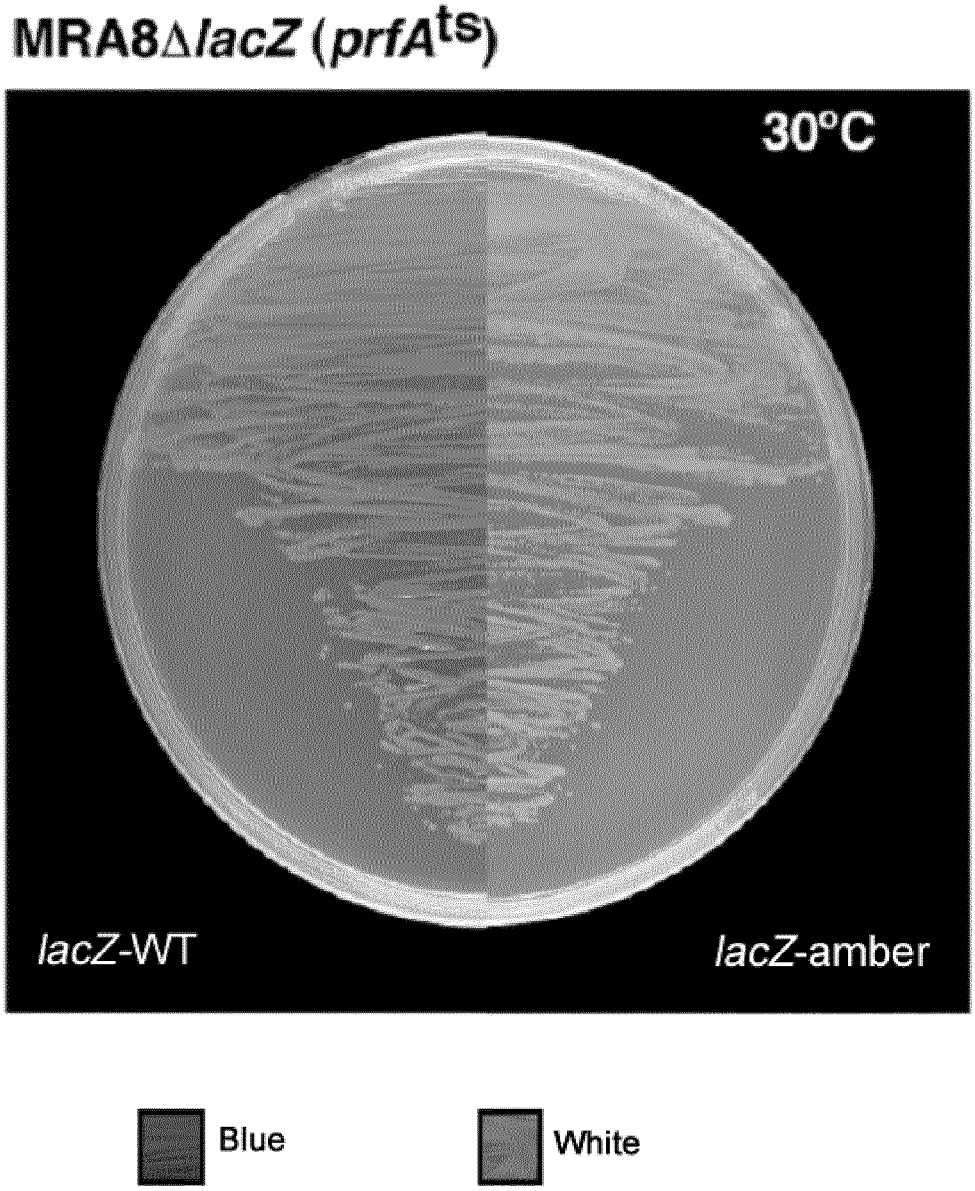


FIG. 3 (cont.)

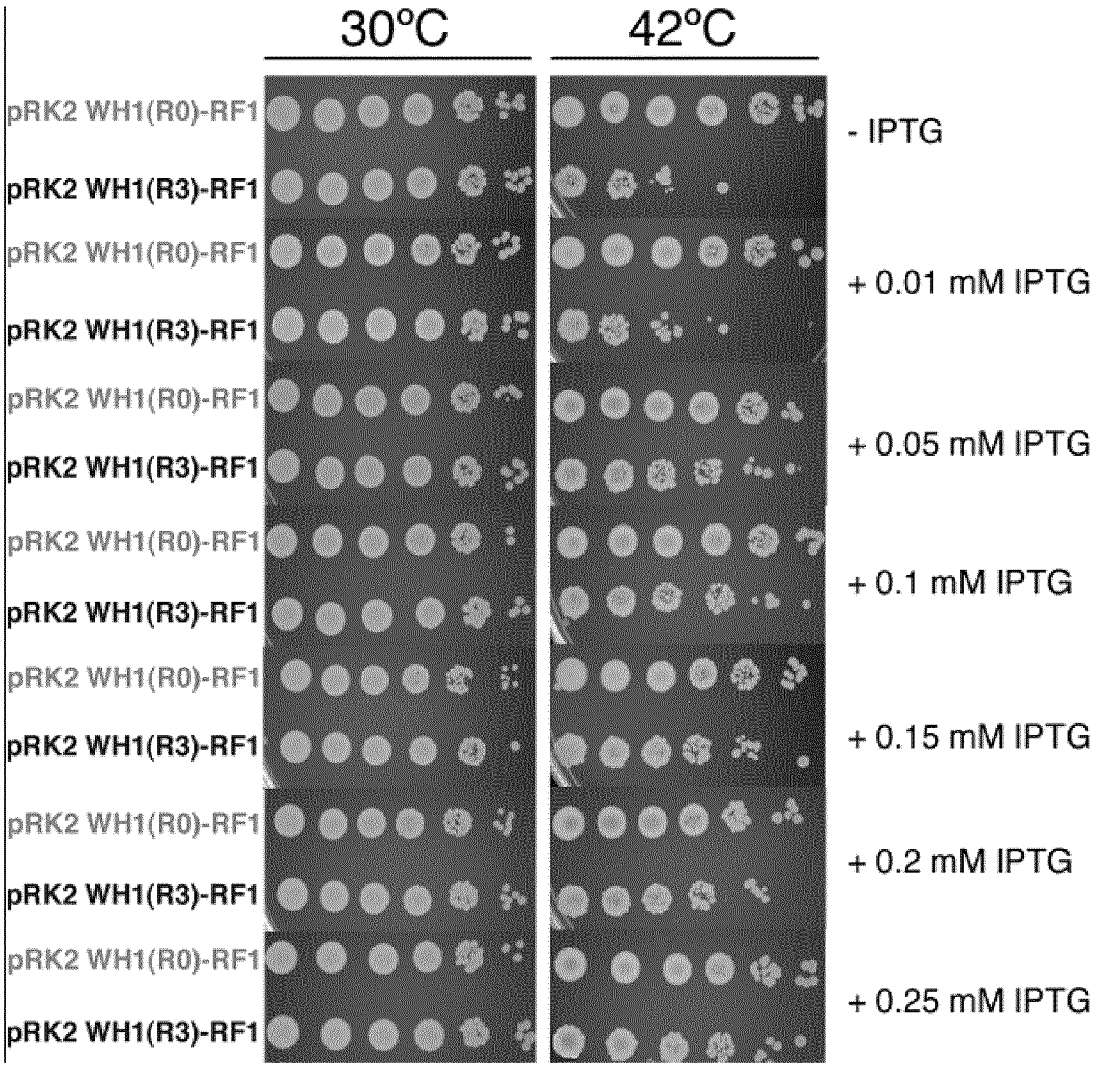

30°C
42°C
pRK2 WH1(R0)-RF1
pRK2 WH1(R3)-RF1
- IPTG
+ 0.01 mM IPTG
+ 0.05 mM IPTG
+ 0.1 mM IPTG
+ 0.15 mM IPTG
+ 0.2 mM IPTG
+ 0.25 mM IPTG

FIG. 4

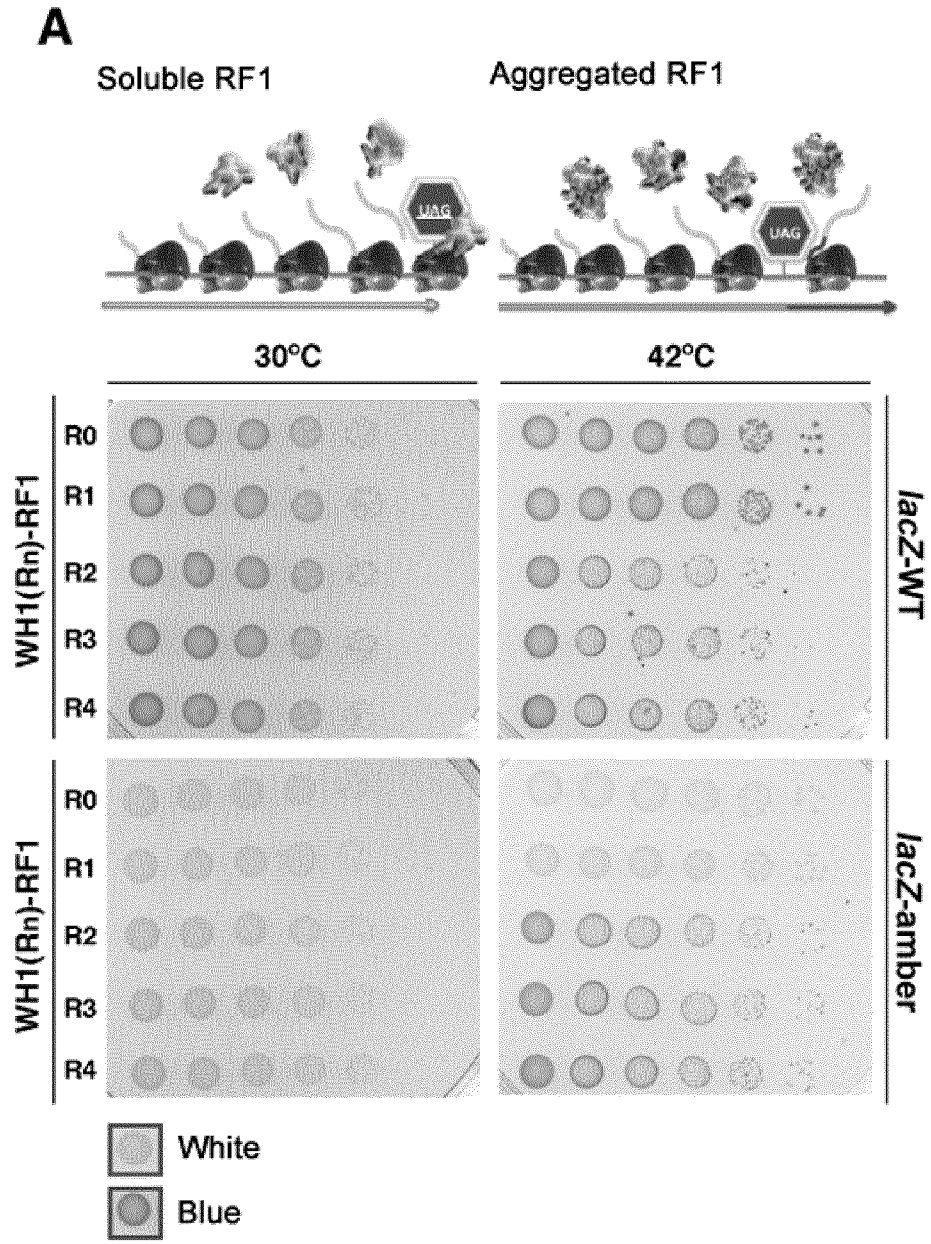

A
Soluble RF1
Aggregated RF1
UAG
UAG
30°C
42°C
WH1(Rn)-RF1
R0
R1
R2
R3
R4
lacZ-WT
WH1(Rn)-RF1
R0
R1
R2
R3
R4
lacZ-amber
White
Blue

FIG. 5

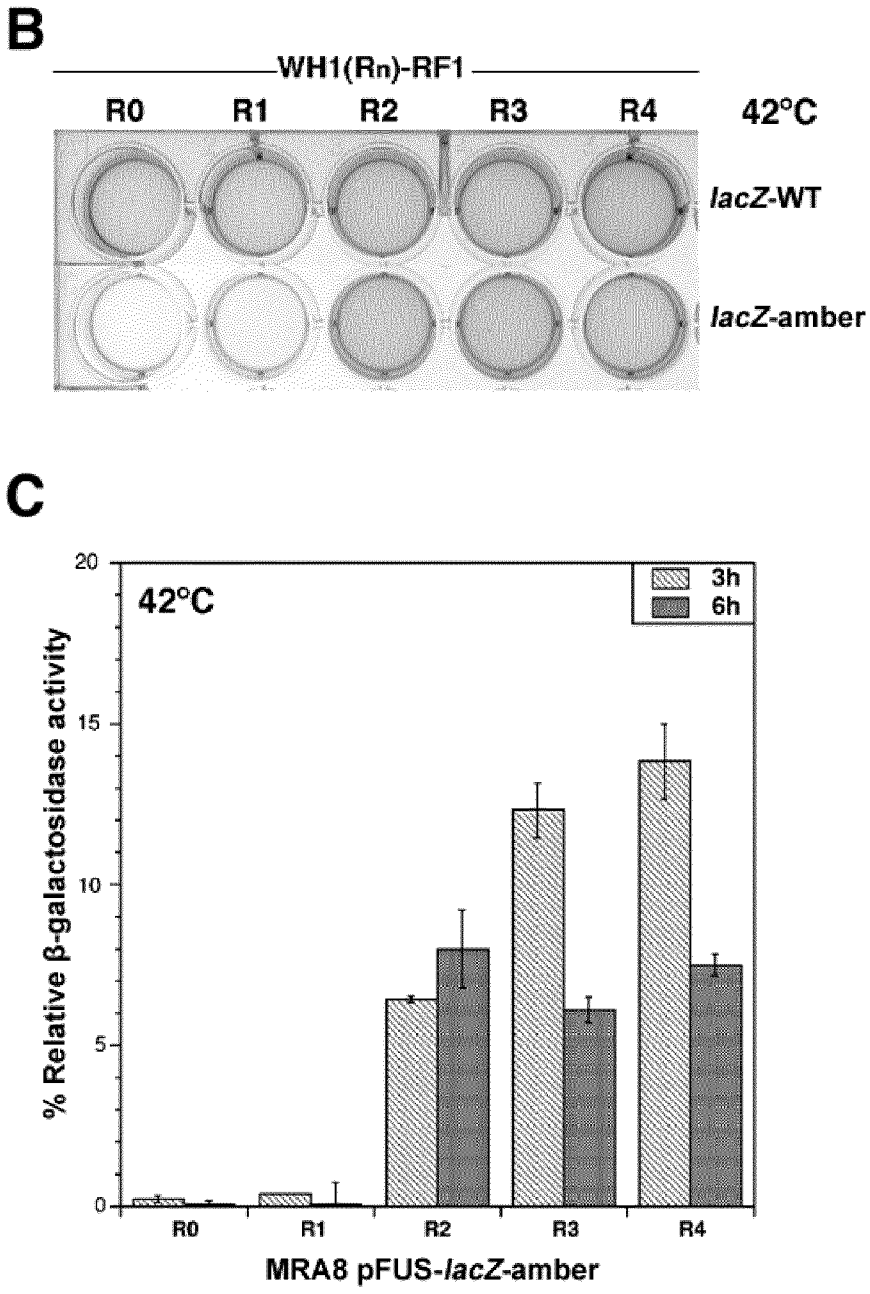

B
WH1(Rn)-RF1
R0
R1
R2
R3
R4
42°C
lacZ-WT
lacZ-amber
C
42°C
3h
6h
% Relative β-galactosidase activity
20
15
10
5
0
R0
R1
R2
R3
R4
MRA8 pFUS-lacZ-amber

FIG. 5 (cont.)

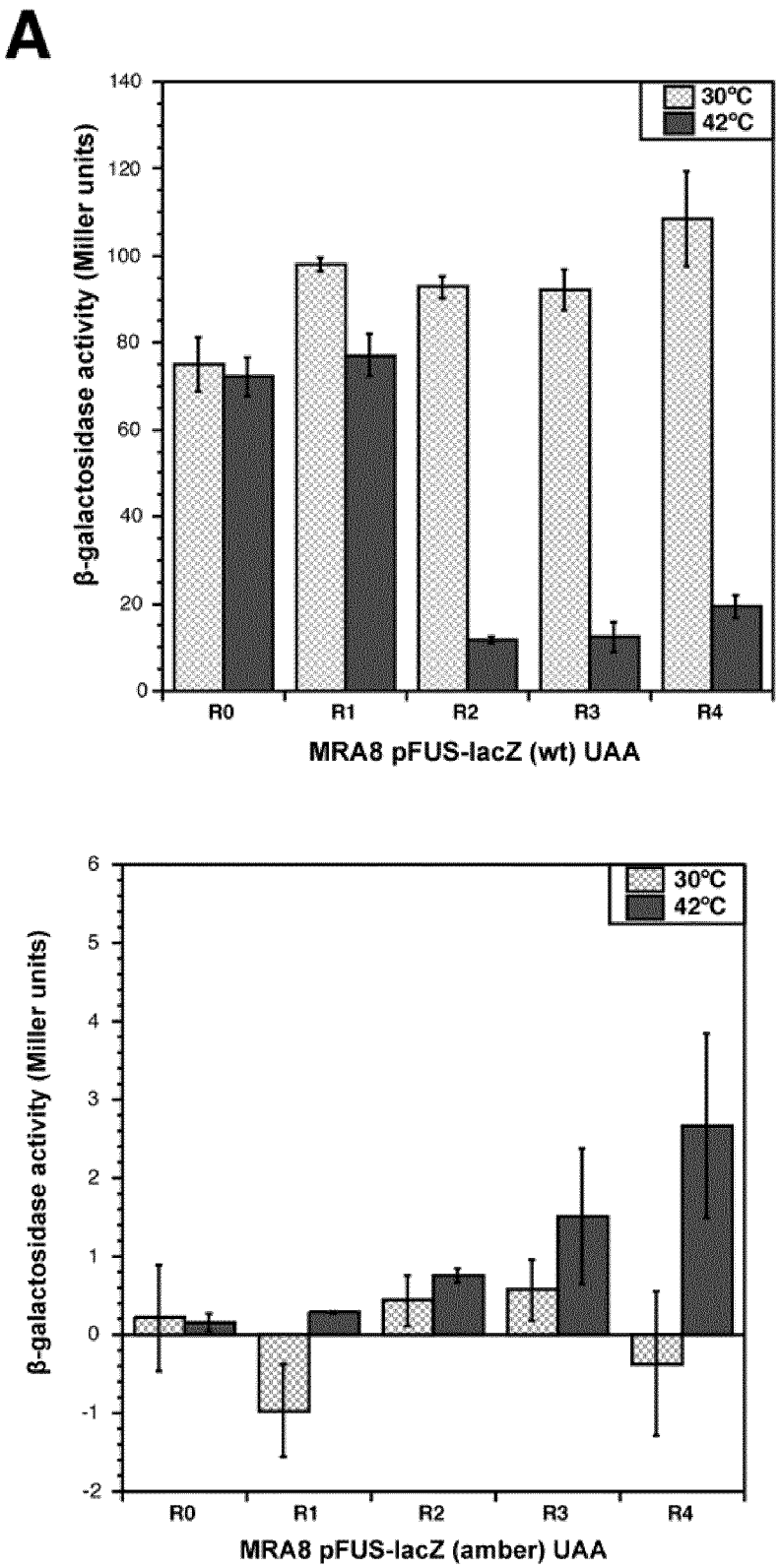

A
β-galactosidase activity (Miller units)
140
120
100
80
60
40
20
0
30°C
42°C
R0
R1
R2
R3
R4
MRA8 pFUS-lacZ (wt) UAA
β-galactosidase activity (Miller units)
6
5
4
3
2
1
0
-1
-2
30°C
42°C
R0
R1
R2
R3
R4
MRA8 pFUS-lacZ (amber) UAA

FIG. 6

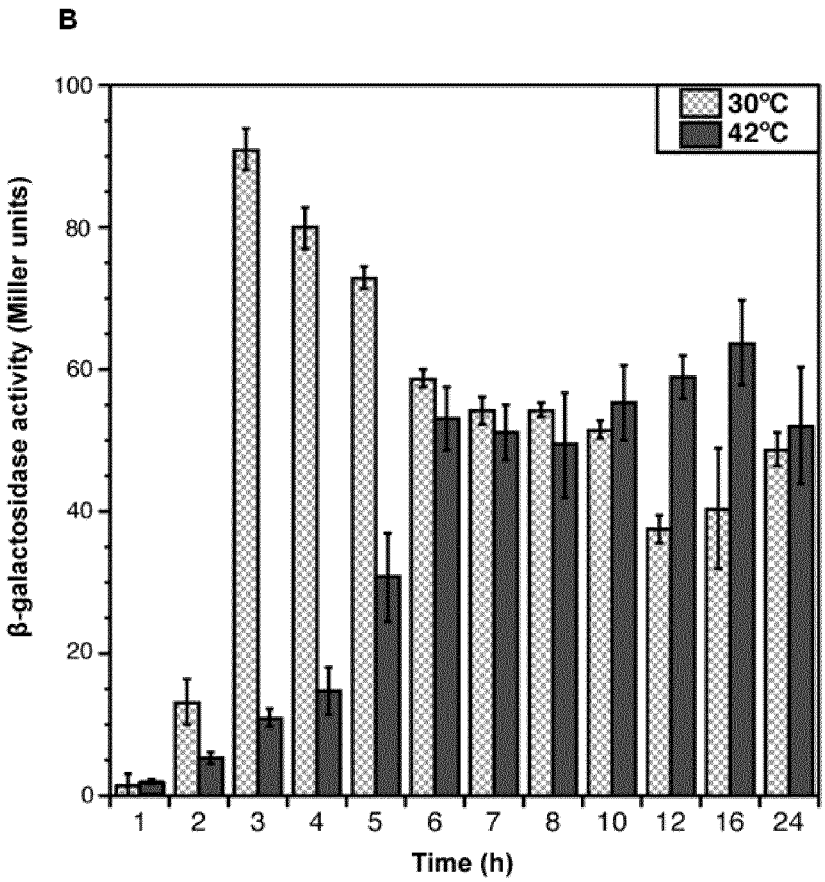
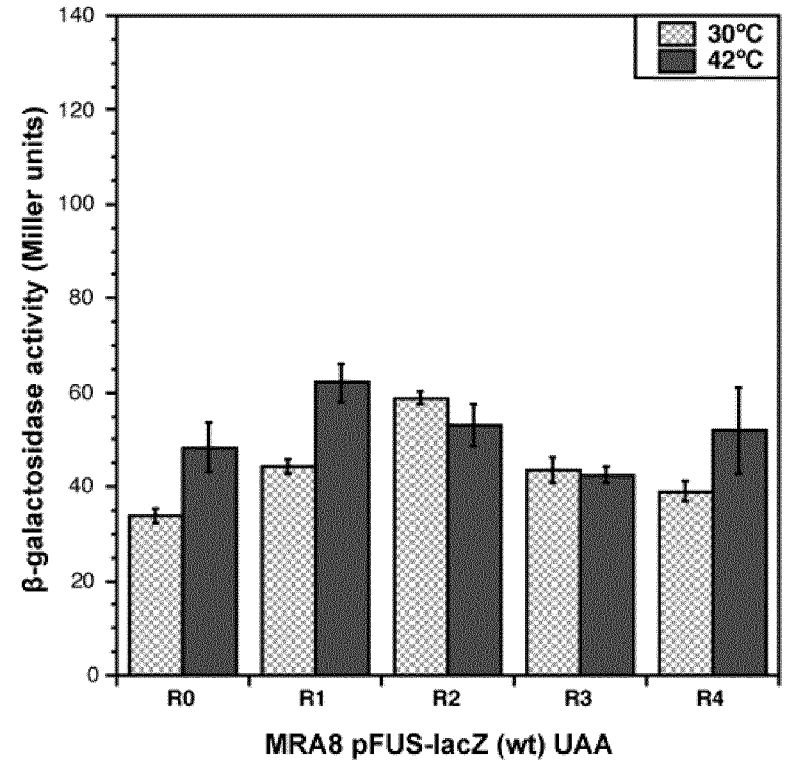

B
β-galactosidase activity (Miller units)
100
80
60
40
20
0
30°C
42°C
1
2
3
4
5
6
7
8
10
12
16
24
Time (h)

FIG. 6 (cont.)

C


30°C
42°C
β-galactosidase activity (Miller units)
0
20
40
60
80
100
120
140
R0
R1
R2
R3
R4
MRA8 pFUS-lacZ (wt) UAA

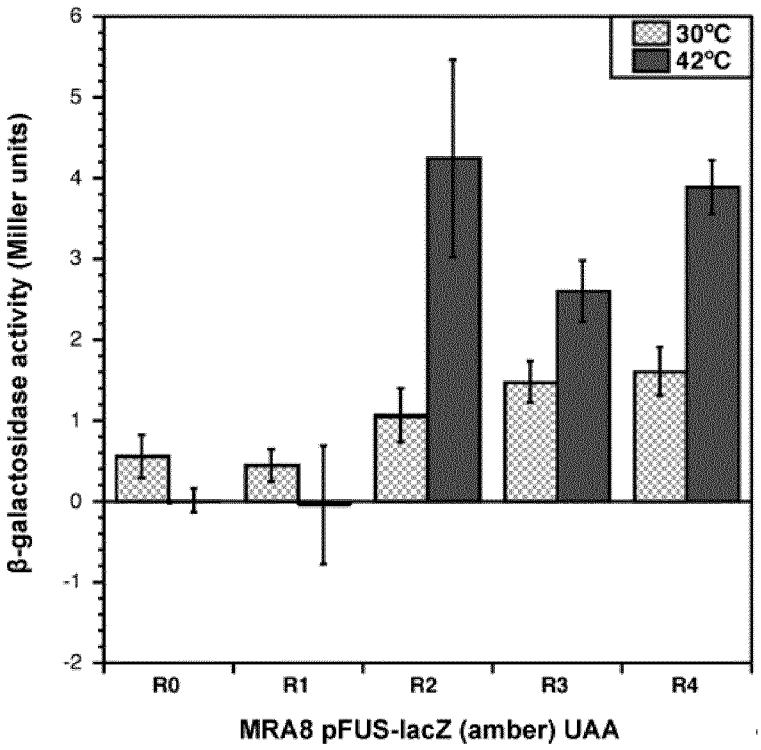

30°C
42°C
β-galactosidase activity (Miller units)
-2
-1
0
1
2
3
4
5
6
R0
R1
R2
R3
R4
MRA8 pFUS-lacZ (amber) UAA

FIG. 6 (cont.)

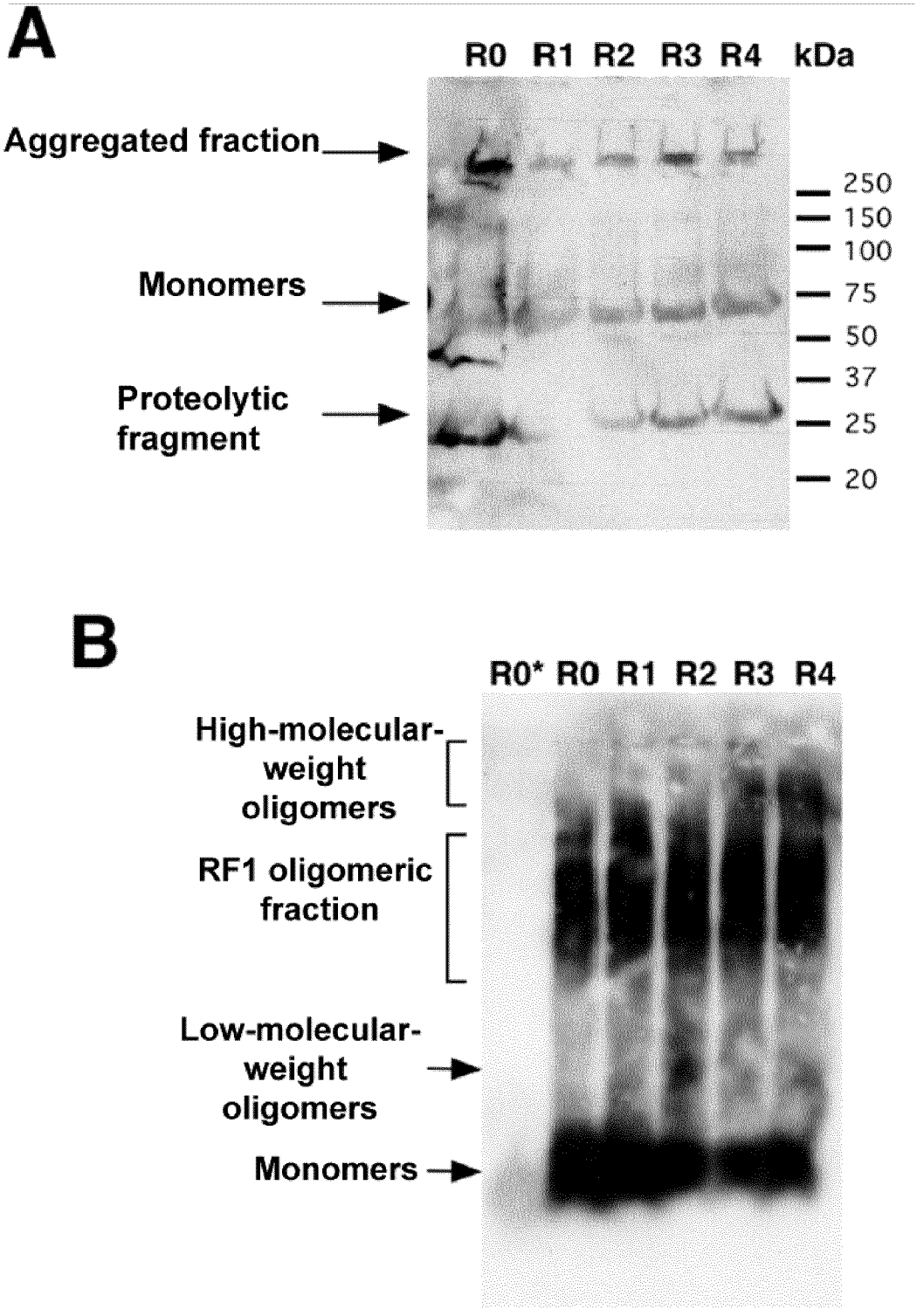

A
R0 R1 R2 R3 R4 kDa
Aggregated fraction
Monomers
Proteolytic fragment
250
150
100
75
50
37
25
20
B
R0* R0 R1 R2 R3 R4
High-molecular-weight oligomers
RF1 oligomeric fraction
Low-molecular-weight oligomers
Monomers

FIG. 7

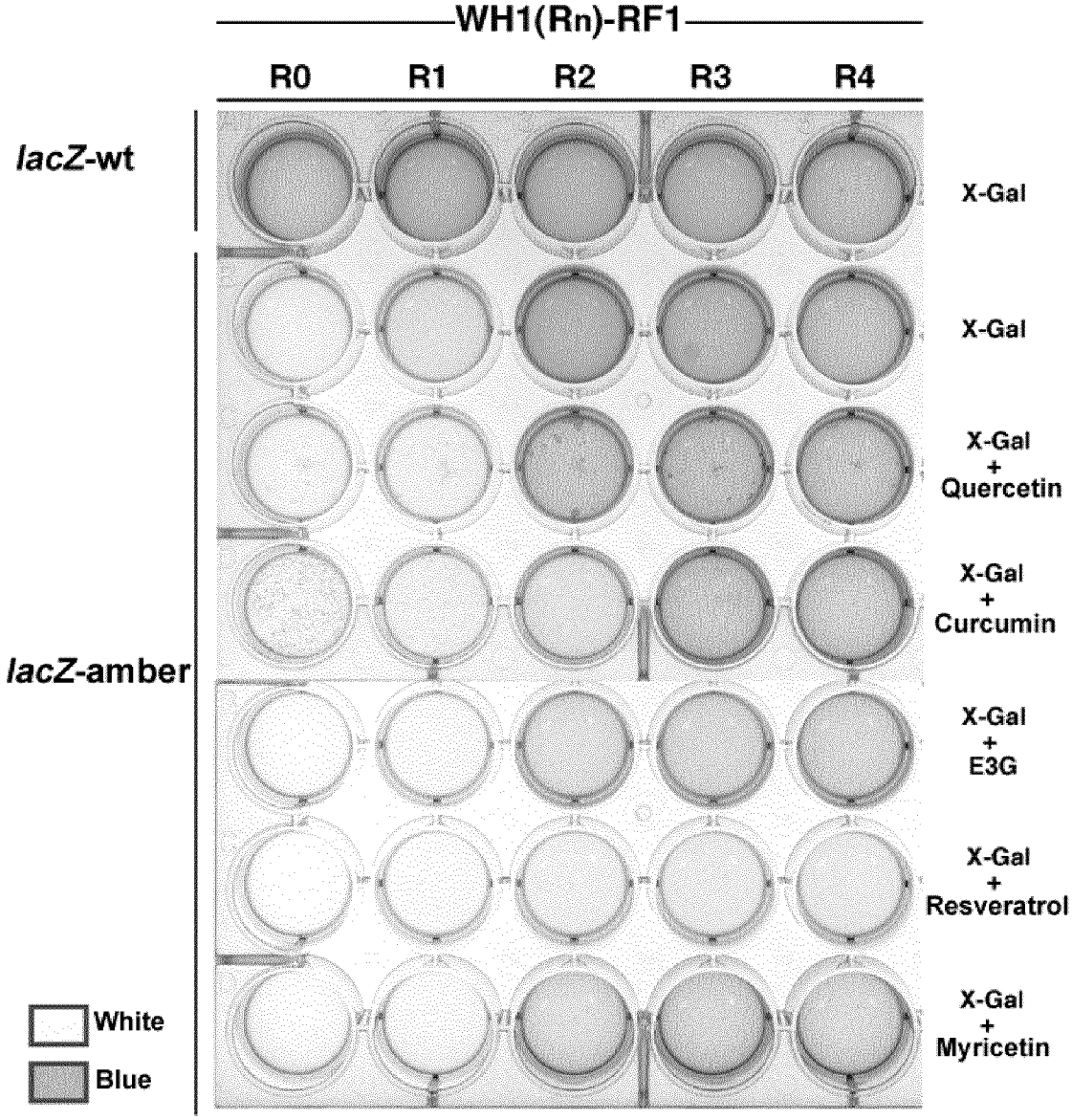

WH1(Rn)-RF1
R0
R1
R2
R3
R4
lacZ-wt
lacZ-amber
X-Gal
X-Gal
X-Gal + Quercetin
X-Gal + Curcumin
X-Gal + E3G
X-Gal + Resveratrol
X-Gal + Myricetin
White
Blue

FIG. 8

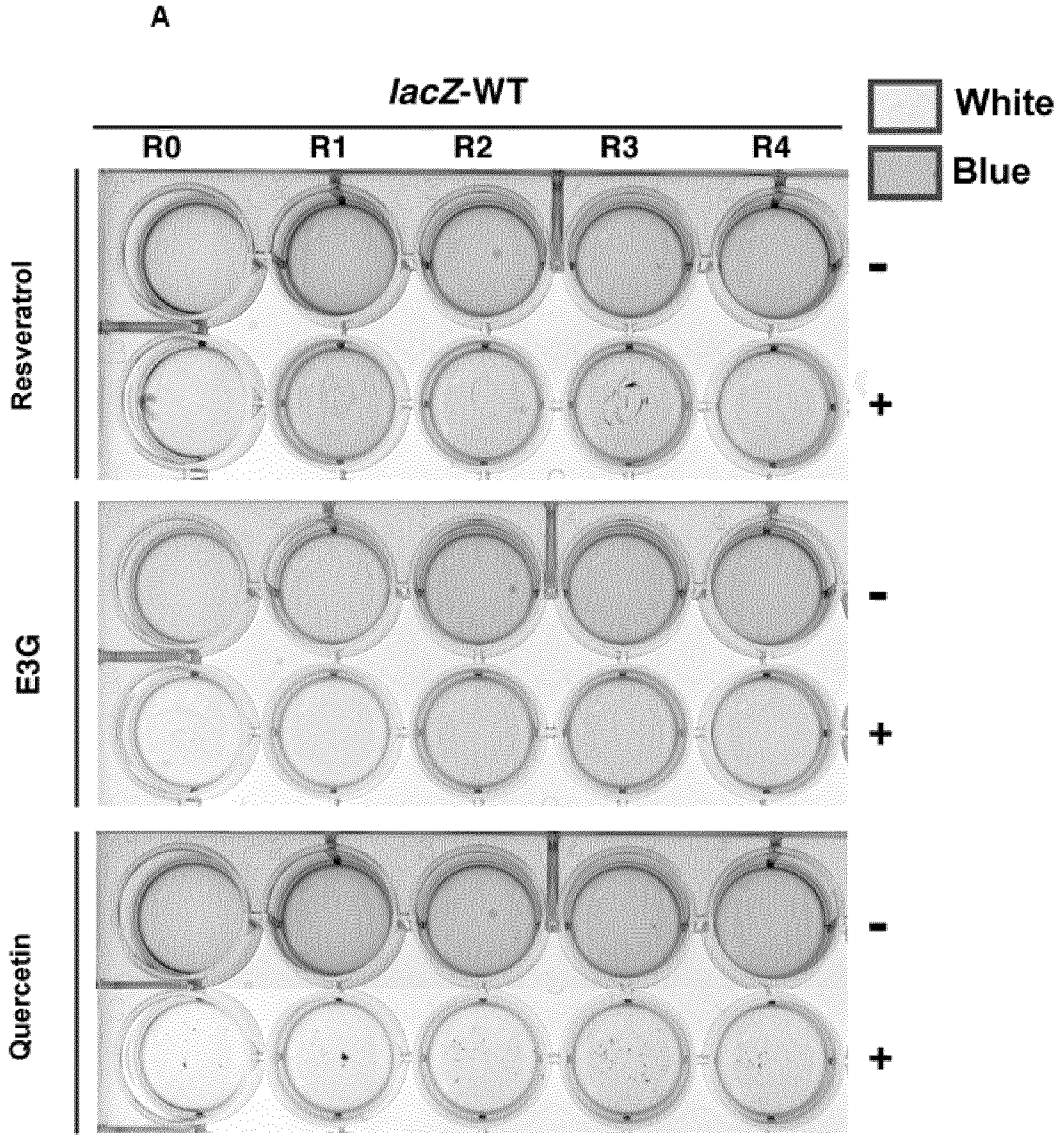


FIG. 9

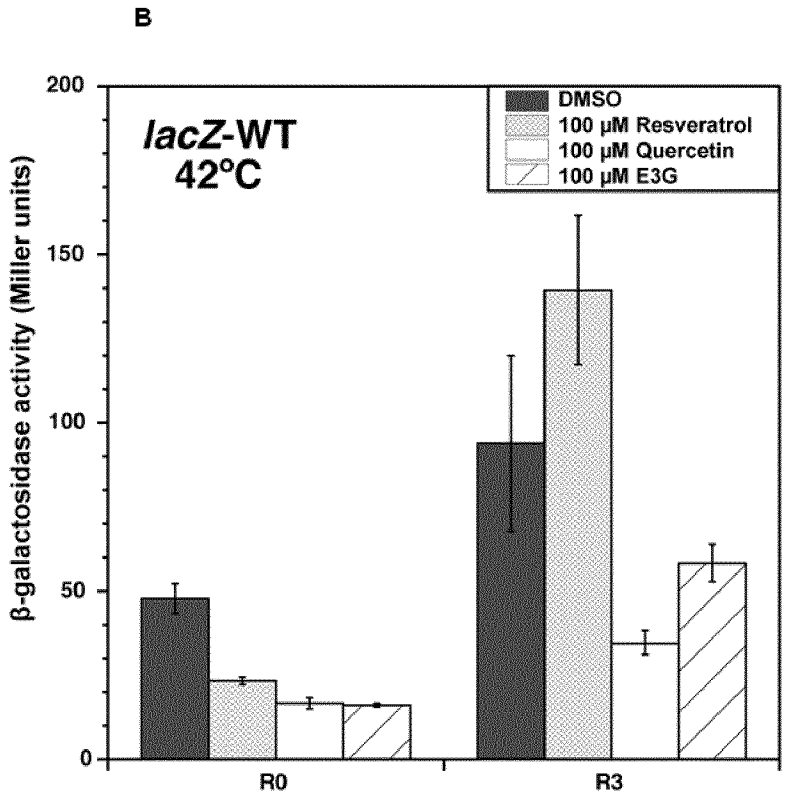

B
lacZ-WT
42°C
DMSO
100 µM Resveratrol
100 µM Quercetin
100 µM E3G
β-galactosidase activity (Miller units)
200
150
100
50
0
R0
R3

FIG. 9 (cont.)

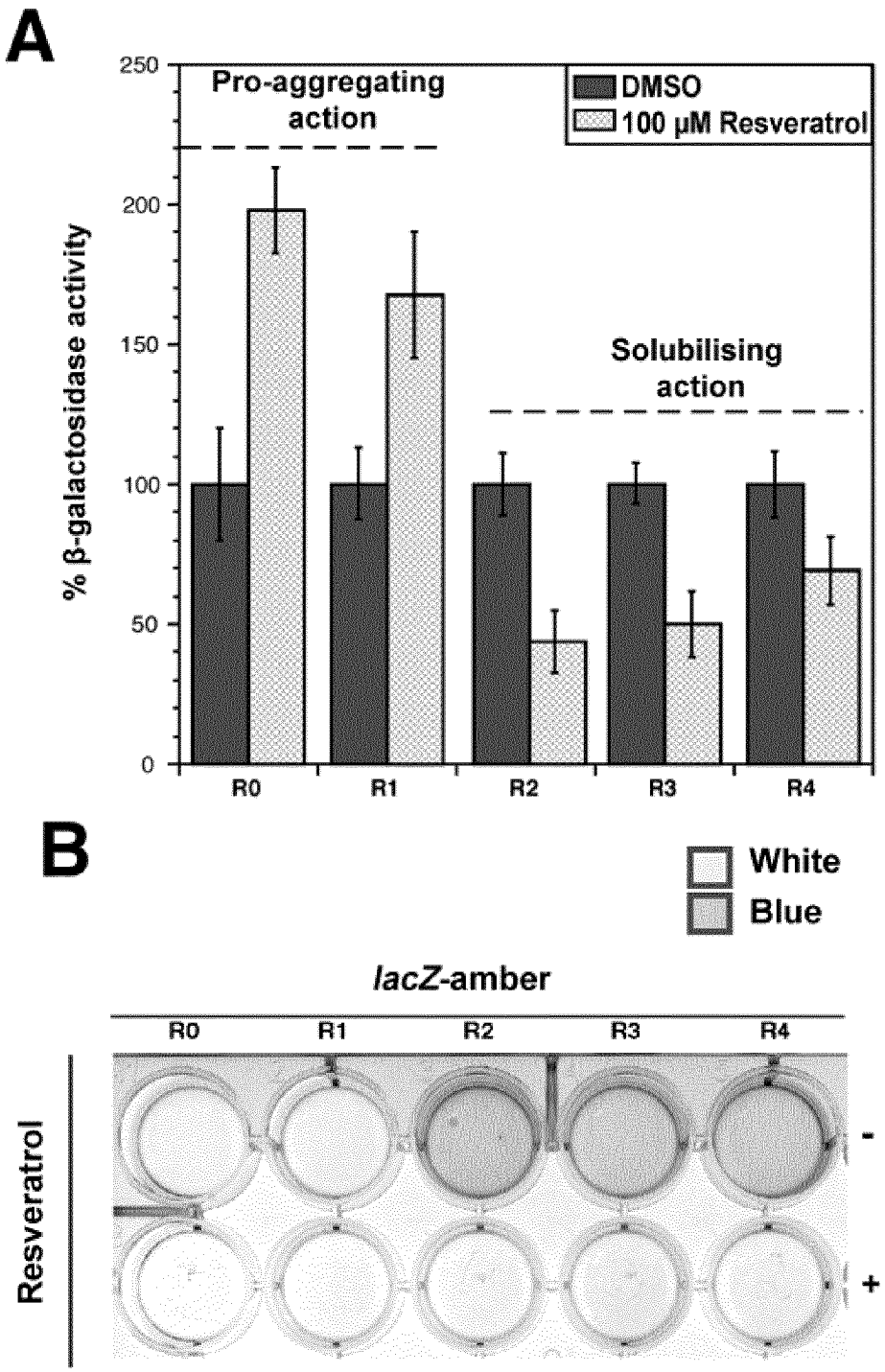

A
250
200
150
100
50
0
% β-galactosidase activity
Pro-aggregating action
Solubilising action
DMSO
100 µM Resveratrol
R0
R1
R2
R3
R4
B
White
Blue
lacZ-amber
R0
R1
R2
R3
R4
Resveratrol
-
+

FIG. 10

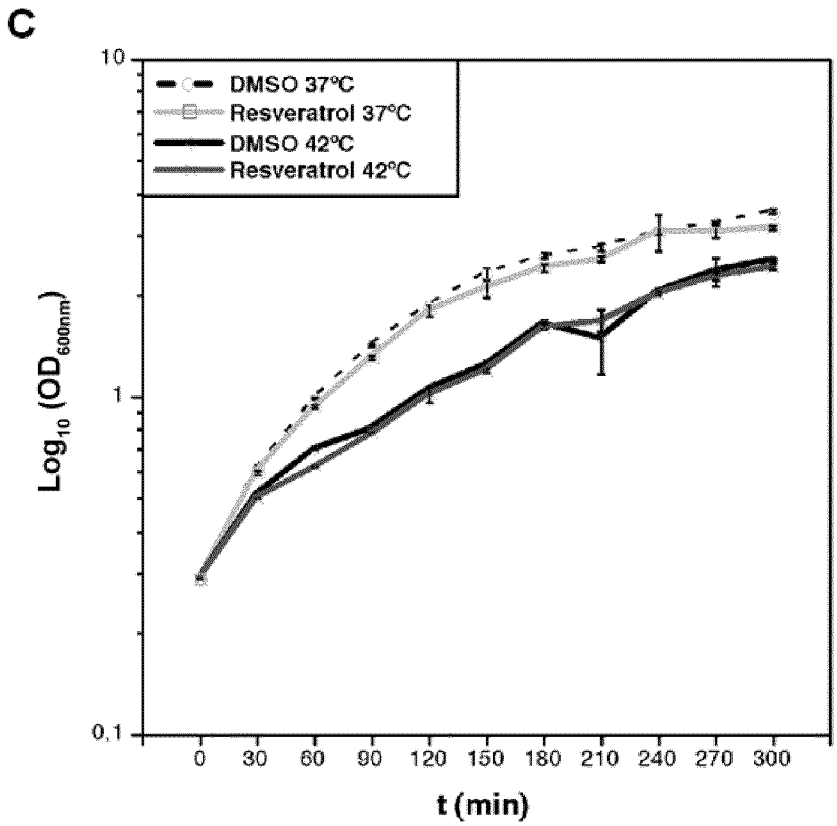

C
DMSO 37°C
Resveratrol 37°C
DMSO 42°C
Resveratrol 42°C
Log10 (OD600nm)
10
1
0,1
0
30
60
90
120
150
180
210
240
270
300
t (min)

FIG. 10 (cont.)

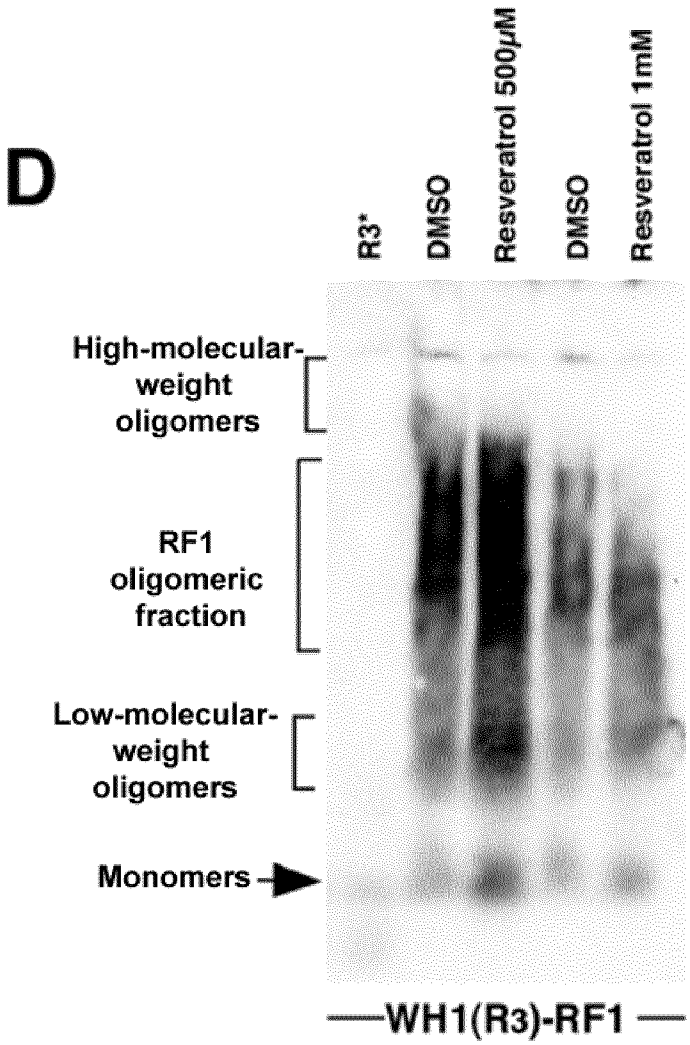

D
R3*
DMSO
Resveratrol 500μM
DMSO
Resveratrol 1mM
High-molecular-weight oligomers
RF1 oligomeric fraction
Low-molecular-weight oligomers
Monomers
WH1(R3)-RF1

FIG. 10 (cont.)

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/057543

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/245 C07K14/47 G01N33/68
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Meng Sun: "DEVELOPMENT OF THE NEW YEAST-BASED ASSAYS FOR PRION PROPERTIES", , 1 December 2011 (2011-12-01), XP055277754, Retrieved from the Internet: URL:https://smartech.gatech.edu/bitstream/handle/1853/45831/sun_meng_201112_phd.pdf?sequence=1 [retrieved on 2016-06-03] page 16 - page 20 ----- -/--	1-20

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 June 2016	22/06/2016

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Vollbach, Silke

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/057543

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	L. MOLINA-GARCIA ET AL: "Aggregation Interplay between Variants of the RepA-WH1 Prionoid in Escherichia coli", JOURNAL OF BACTERIOLOGY, vol. 196, no. 14, 15 July 2014 (2014-07-15), pages 2536-2542, XP055278001, US ISSN: 0021-9193, DOI: 10.1128/JB.01527-14 abstract	1-20
A	US 2007/077552 A1 (HECHT MICHAEL [US] ET AL) 5 April 2007 (2007-04-05) figure 1	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2016/057543

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2007077552 A1	05-04-2007	NONE	