



(51) International Patent Classification:

A61K 38/07 (2006.01)

(21) International Application Number:

PCT/IB2018/000622

(22) International Filing Date:

22 May 2018 (22.05.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

PCT/EP17/25141 22 May 2017 (22.05.2017) EP

PCT/EP17/25298 05 October 2017 (05.10.2017) EP

(71) Applicant: NATIONAL HELLENIC RESEARCH FOUNDATION [GR/GR]; 48, Vassileos Constantinou Avenue, GR- 11635 Athens (GR).

(72) Inventor: SKRETAS, Georgios; C/o National Hellenic Research Foundation, 48, Vassileos Constantinou Avenue, GR-11635 Athens (GR).

(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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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:

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

(54) Title: MACROCYCLIC MODULATORS OF DISEASE ASSOCIATED PROTEIN MISFOLDING AND AGGREGATION

(57) Abstract: Aspects of the present invention disclose compounds that modulate the aggregation of amyloidogenic proteins or peptides. In some aspects, disclosed compounds modulate the aggregation of disease-associated proteins and natural β -amyloid peptides. In a preferred embodiment, the compounds can inhibit natural amyloid aggregation. Pharmaceutical compositions comprising the compounds of the embodiments, and diagnostic and treatment methods for diseases (e.g., amyloidogenic diseases) using the compounds, are also disclosed. In addition, there is provided an integrated bacterial platform for the discovery of rescuers of disease-associated protein misfolding.



DESCRIPTION

MACROCYCLIC MODULATORS OF DISEASE ASSOCIATED PROTEIN MISFOLDING AND AGGREGATION

This application claims the benefit of International Patent Application No. PCT/EP2017/025141, entitled “MACROCYCLIC MODULATORS OF β -AMYLOID MISFOLDING AND AGGREGATION” filed on 22 May 2017, and of International Patent Application No. PCT/EP2017/025298, entitled “MACROCYCLIC RESCUERS FOR DISEASE-ASSOCIATED PROTEIN MISFOLDING” filed on 5 October 2017, each of which is incorporated herein by reference in its entirety.

TECHNICAL FIELD OF THE INVENTION

Aspects of the present invention relate to a generalizable bacterial platform for the discovery of chemical modulators of the problematic folding and aggregation of disease-associated, misfolding-prone proteins. More particularly, studies herein demonstrate the applicability of this platform to biosynthetically produce large combinatorial libraries of macrocyclic compounds in *Escherichia coli* cells and to simultaneously screen these libraries in order to identify the bioactive compounds with the ability to rescue the problematic folding and aggregation of mutant Cu/Zn superoxide dismutase using a high-throughput genetic assay. Furthermore, studies herein demonstrate the wider applicability of this platform to identify the bioactive compounds with the ability to rescue the problematic folding and aggregation of additional misfolding-prone polypeptides, such as the β -amyloid peptide. In some aspects, compounds of the invention preferably rescue the misfolding and modulate the aggregation of human Cu/Zn superoxide dismutase and of its variants. In further aspects, of the invention rescue misfolding and modulate the aggregation of natural β -amyloid peptides. In a preferred embodiment, the compounds can inhibit the aggregation of Cu/Zn superoxide dismutase and of its variants. In another preferred embodiment, compounds of the present invention can inhibit natural β -amyloid peptide aggregation. In another preferred embodiment, the Cu/Zn superoxide dismutase modulator compounds of the invention are head-to-tail cyclic oligopeptides, or variants thereof carrying specific modifications, such that the compound alters the aggregation or inhibits the neurotoxicity of Cu/Zn superoxide dismutase and of its variants when contacted with the peptides. In another preferred embodiment, the β -amyloid modulator compounds of the invention are head-to-tail cyclic oligopeptides, or variants thereof carrying specific modifications, such that the compound alters the aggregation or inhibits the

neurotoxicity of natural β -amyloid peptides when contacted with the peptides. Pharmaceutical compositions comprising the compounds of the invention, and diagnostic and treatment methods for amyloidogenic diseases, such as amyotrophic lateral sclerosis, using the compounds of the invention, are also disclosed.

5

BACKGROUND OF THE INVENTION

Protein misfolding is currently linked to more than 50 diseases including Alzheimer's disease, Parkinson's disease, Huntington's disease, type 2 diabetes, cystic fibrosis, amyotrophic lateral sclerosis, Gaucher's disease, nephrogenic diabetes insipidus, and
10 Creutzfeldt-Jakob disease. These disorders are collectively termed "conformational diseases" or "protein misfolding diseases" (PMDs). There are two ways that misfolded prone proteins (MisPs) lead to disease; one is when they lose their ability to execute their physiological function (loss-of-function) and the other when they acquire a new harmful property (gain-of-function). Cellular or environmental factors such as changes in pH, oxidative stress, exposure
15 to high concentrations of metal ions and other chemicals, as well as the presence of a mutation or mutations in amino acid sequences of particular proteins, can play a critical role in protein misfolding. Protein misfolding diseases are becoming more common as the population ages, as many of them are age-related.

PMDs include very serious disorders with high incidence rates and a severe impact on
20 the well-being of the human population, and anti-PMD therapeutics are in enormous demand. One of the most promising approaches for identifying potential anti-PMD therapeutics is the discovery of chemical rescuers of protein misfolding. Such molecules have already been identified for a number of MisPs. For example, linear peptides with homology to certain regions of the β -amyloid peptide ($A\beta$) and small molecules, such as scyllo-inositol,
25 tramiprosate, methylene blue and bexarotene, have been found to modulate $A\beta$ aggregation and inhibit its neurotoxicity *in vitro* and *in vivo*, and some of them have subsequently advanced to clinical studies. Similarly, peptides with homology to the unstructured central hydrophobic region of the PD-related protein α -synuclein (α syn) and natural products, such as baicalein and (2)-epigallocatechin-3-gallate have exhibited similar effects on α syn. Indeed, the small
30 molecule tafamidis, which is capable of rescuing the misfolding of the carrier protein transthyretin, has recently been approved for the treatment of familial amyloidotic polyneuropathy in Europe and Japan and is currently marketed under the name Vyndaqel[®] (Pfizer). The compound and its use for the treatment transthyretin amyloid disease have been disclosed in the European patent EP1587821.

Macrocycles have been characterized as a particularly promising class of compounds of potential therapeutics, which remain underexplored (Driggers, E.M., Hale, S.P., Lee, J. & Terrett, N.K. The exploration of macrocycles for drug discovery--an underexploited structural class. *Nat. Rev. Drug Discov.* 7, 608-624 (2008)). Macrocycles occupy the space between small molecules and larger biologicals and often exhibit the advantages of both classes of molecules i.e., the high bioavailability of small molecules combined with the high specificity and the fewer side-effects of biologicals. Furthermore, their typically larger size and more complex structure makes the macrocycles particularly suitable for targeting currently undruggable targets, such as ones involved in protein-protein interactions. Since many PMDs are characterized by protein aggregation, a process that is dependent on productive protein-protein interactions, macrocycles can be expected to be particularly active modulators for this class of disorders. Their therapeutic potential is slowly beginning to rise in a wide variety of diseases and shown in US patent US9308236 wherein macrocycles are shown to inhibit PD-1/PD-L1 (Programmed Death 1) and CD80/PD-L1 protein/protein interactions, and are thus useful in ameliorating various diseases including cancer and infectious diseases.

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder that affects the motor neurons of the spinal cord, brain stem, and cortex of adults most frequently between 50 and 60 years of age. The disease is ultimately fatal with an average survival time of 3-5 years. Its causes remain both enigmatic and controversial. The majority of cases (90-95%) have no known genetic link and are termed sporadic. For the rest of the 5-10% of cases, there is typically a family history of ALS, the disease is inherited (familial ALS, fALS), and it is caused by genetic mutations present in specific chromosomal loci. Approximately one quarter of all cases of the familial disease are associated with missense mutations mapped onto *SOD1*, the gene encoding for the enzyme Cu/Zn superoxide dismutase (SOD1).

Superoxide dismutases (SODs) belong to the family of isoenzymes involved in the scavenging of O_2 radicals. All mammalian cells possess three isoforms of superoxide dismutase enzymes; the cytosolic copper-zinc dimeric form, known as SOD1, the mitochondrial tetrameric manganese superoxide dismutase or SOD2 and the extracellular tetrameric Cu, Zn superoxide dismutase or SOD3. All enzymes catalyze the same reaction converting the oxygen radical in molecular oxygen and hydrogen peroxide H_2O_2 through the alternate reduction and reoxidation of Cu^{2+} for SOD1 and SOD3 and Mn for SOD2; the H_2O_2 is then enzymatically converted by catalase and glutathione peroxidase in molecular oxygen and H_2O . In physiological conditions, the superoxide dismutases, together with the non-enzymatic reactive

oxygen species (ROS) scavengers vitamins E, A, and C, maintain a steady state between oxidant and anti-oxidant systems.

To date, more than 150 mutations in *SOD1* have been found to be associated with fALS (<http://alsod.iop.kcl.ac.uk/home.aspx>). These result in amino acid substitutions, C-terminal truncations and other modifications in the amino acid sequence of SOD1. It is now well established that these changes in the sequence of SOD1 do not cause ALS due to loss or decrease of enzymatic activity. The main pieces of evidence supporting this are that: (i) *SOD1*-knockout mice do not develop ALS phenotypes, (ii) the onset and duration of motor neuron disease in transgenic mice carrying fALS-associated *SOD1* alleles is similar irrespective of the presence or absence also of the wild-type allele in the animal, and (iii) many fALS-associated SOD1 variants (SOD1*) retain wild type-like levels of dismutase activity. Instead, it has been proposed that fALS-linked mutations introduce a toxic-gain-of-function property in SOD1 by causing protein misfolding and aggregation, and the formation of oligomeric/aggregated SOD1 species which are highly toxic for motor neurons. Gradual accumulation of such toxic oligomers/aggregates of mutated SOD1 initiates motor neuron degeneration and the development of fALS. Indeed, many observations support this theory: (i) biochemical studies of a variety of fALS-associated SOD1 variants have been found to be less stable, more prone to misfolding, and with higher aggregation propensity compared to wild-type SOD1, (ii) prominent SOD1 aggregates have been found in the cytosolic space of cultured motor neurons, in motor neurons and in neighboring astrocytes of SOD1* transgenic mice, and of fALS patients, (iii) SOD1* aggregated species have been found to be toxic for motor neurons, (iv) SOD1*-induced motor neuron toxicity can be suppressed by up-regulating actors that assist SOD1* folding and inhibit its aggregation, such as the heat shock response regulator Hsf1, and (v) the combination of aggregation propensity and loss of stability in fALS-associated SOD1* variants has been found to be a good predictor of disease severity. The theory that SOD1*-linked fALS is a conformational disorder/protein misfolding disease is in agreement with the prevalent theories concerning the molecular origin of other major neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease and prion disorders (1. Chiti, F. & Dobson, C.M. Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem. 75, 333-366 (2006)).

Despite the major advances in identifying genes and mechanisms contributing to ALS pathogenesis, there exist only two currently approved therapeutics: the glutamate antagonist Riluzole and the free radical scavenger that is believed to relieve the effects of oxidative stress, Radicava. However, both treatments only extend the life or the time to mechanical ventilation

of the patient by two to three months. Therefore, there still remains a need for developing a novel and cost-effective treatment approach to ALS that will overcome the obstacles and side-effects of the current treatment regime that only result in a minor delay of the outcome of the disease.

5 Alzheimer's disease (AD) is the most common progressive neurodegenerative disease that causes dementia in aged humans. This condition affects more than 7 million people in Europe and 35 million people worldwide. In financial terms, these numbers translate to an annual AD treatment cost of \$ 818 billion only for the United States (2015 World Alzheimer Report). Due to the aging of the human population, the incidence of AD has been rising steadily
10 in recent years and is projected to increase by 200% within the next 20 years.

AD is neuropathologically characterized by intraneuronal neurofibrillary tangles consisting of abnormally hyperphosphorylated tau, extracellular accumulation of fibrillar amyloid- β (A β) peptide in senile plaques, and the build-up of soluble A β oligomers in the brain. The amyloid cascade hypothesis, which states that the formation of oligomeric and/or fibrillar
15 A β in the brain is neurotoxic and that their accumulation results in neuronal degeneration and death, and the development of AD is the prevalent theory regarding the molecular sources behind the pathology of AD.

A β is produced by the cleavage of the amyloid precursor protein (APP) as a result of the action of the proteases β - and γ -secretase. Due to the broad amino acid recognition
20 specificity of γ -secretase, its proteolytic activity on APP leads to the formation of different forms of A β peptides. The main product of APP cleavage consists of 40 amino acids (A β ₄₀), while one of the secondary products yields a 42-amino-acid-long peptide (A β ₄₂). A β is an amphipathic peptide that includes a hydrophilic (amino acids 1-16) and a hydrophobic region (amino acids 17-40/42). A β ₄₂ contains two additional amino acids at its C terminus and exhibits
25 higher hydrophobicity and higher tendency for aggregation than A β ₄₀. Due to the distribution of the hydrophobic amino acids in its sequence, A β is unable to adopt a well-defined conformation. As a result, certain hydrophobic regions of the A β sequence remain exposed to the aqueous environment of the cell and the protein tends to form oligomers and higher-order aggregates. As mentioned above, these soluble oligomers and/or higher-order aggregates are
30 thought to be the neurotoxic A β species that initiate AD. AD can thus be viewed as a proteinopathy, which is initiated due to the peculiar biochemical/biophysical properties of A β and its associated problematic folding (misfolding), thus categorizing AD in the large group of disorders called "conformational diseases" or "protein misfolding diseases" as described above.

Despite the huge socioeconomic impact of AD and intense research efforts for decades, preventive or therapeutic treatments against AD do not exist currently. Anti-AD therapies could be developed by identifying inhibitors of anyone step of the pathway that is initiated by the biosynthesis of A β and results in neuron degeneration and the development of dementia. For example, molecules with the ability to decrease the concentration of circulating A β could act as agents that down-regulate the formation of neurotoxic A β species and the onset of the disease. This could be achieved, for instance, by using antibodies against A β that lead to sequestration, degradation, and/or clearance of the peptide, or by using β - and γ -secretase inhibitors which inhibit APP cleavage and decrease the rate of biosynthesis of A β . Such compounds have already been discovered but they haven't yet demonstrated the desired therapeutic profiles, primarily due to unwanted side effects. For example, γ -secretase inhibitors have exhibited undesired actions in other physiological pathways, such as Notch signaling, which cause serious side effects in mice (gastrointestinal tract symptoms etc.).

Compounds which can bind to A β , correct its problematic folding, and prevent the formation of A β oligomers/aggregates have the potential to function as inhibitors of A β -induced neurotoxicity and become effective drugs against AD. Small molecules, such as scyllo-inositol, tramiprosate, methylene blue and bexarotene, have been found to modulate A β aggregation and inhibit its neurotoxicity *in vitro* and *in vivo*. In addition, linear peptides with homology to certain regions of A β have exhibited similar properties. These and other research efforts have led to the identification of a number of promising compounds which have been or are being tested in clinical trials. Until now, however, no compound has demonstrated the desired preventive or therapeutic properties against AD. Irrespective of whether it is the oligomeric or more aggregated forms of A β which are primarily responsible for its neurotoxicity and other pathogenic effects, the discovery of chemical modulators of the natural A β oligomerization/aggregation process is considered a very promising approach, since A β oligomerization/aggregation is viewed as a purely pathogenic process, which is not involved in other physiological functions in the cell. Such molecules are thus expected to exhibit a safer pharmacological profile.

DESCRIPTION OF FIGURES

FIGS. 1A-1B: (A) (Left) Schematic of the pSICLOPPS-NuX₁X₂X₃-X₅ vector library encoding the combinatorial oligopeptide library cyclo-NuX₁X₂X₃-X₅. Nu: Cys (C), Ser (S), or Thr (T); X: any of the 20 natural amino acids; NNS: randomized codons, where N=A, T, C or

G and S=G or C; I_C: C-terminal fragment of the split Ssp DnaE intein; I_N: N-terminal fragment of the split Ssp DnaE intein; CBD: chitin-binding domain. (Right) Intein-mediated peptide cyclization using SICLOPPS. The tetra-partite fusion undergoes intein splicing upon intein fragment re-association, leading to peptide cyclization and the production of the cyclo-NuX₁X₂X₃-X₅ library. (B) Theoretical and actual diversity of the constructed combinatorial cyclo-NuX₁X₂X₃-X₅ oligopeptide library.

FIGS. 2A-2C: (A) Relative fluorescence of *E. coli* BL21(DE3) (left) and Origami 2(DE3) (right) cells overexpressing chimeric SOD1-GFP fusions from the corresponding pETSOD1-GFP vectors. Mean values $\pm$ s.e.m. are reported ($n=3$ independent experiments, each one performed in replica triplicates). (B) Solubility analysis of overexpressed SOD1-GFP fusions by SDS-PAGE/western blotting using an anti-polyHis antibody. Representative data from $n=2$ independent experiments are presented. (C) Solubility analysis of SOD1 variants overexpressed as in (A, right) by SDS-PAGE/western blotting using an anti-polyHis (left) or an anti-FLAG (right) antibody. Representative data from $n=2$ independent experiments are presented. The assay was performed using *E. coli* Origami 2(DE3) cells overexpressing GFP-free SOD1 from the corresponding pETSOD1 vectors by the addition of 0.01 mM IPTG at 37 °C for 2 h.

FIGS. 3A-3G: (A) Schematic representation of the utilized bacterial platform for the discovery of cyclic oligopeptide rescuers of MisP misfolding and aggregation. pMisP-GFP: plasmid encoding a MisP-GFP fusion; pSICLOPPS-NuX₁X₂X₃-X₅: vector library encoding the combinatorial oligopeptide library cyclo-NuX₁X₂X₃-X₅. Nu: Cys (C), Ser (S), or Thr (T); X: any of the twenty natural amino acids; NNS: randomized codons, where N=A, T, C or G and S=G or C; P: sorting gate. (B) FACS screening of *E. coli* BL21(DE3) overexpressing SOD1(A4V)-GFP and the combined cyclo-NuX₁X₂X₃-X₅ oligopeptide library. M: mean GFP fluorescence in arbitrary units. (C) Fluorescence of *E. coli* Origami 2(DE3) cells co-expressing SOD1(A4V)-GFP and four individual cyclic peptide sequences isolated after the fourth round of FACS sorting shown in (B) by utilizing wild-type Ssp DnaE intein or the splicing-deficient Ssp DnaE(H24L/F26A) intein. Mean values $\pm$ s.e.m. are reported ($n=4$ independent experiments, each one performed in replica triplicates). (D) Western blot analysis using an anti-CBD antibody of the four individual selected clones investigated in (C). The upper band of ~25 kDa corresponds to the I_C-peptide sequence-I_N-CBD precursor, while the lower band of ~20 kDa corresponds to the processed I_N-CBD product, whose appearance is an indication of

successful intein processing and cyclic peptide formation. CBD: chitin-binding domain. (E) Fluorescence of *E. coli* Origami 2(DE3) cells co-expressing SOD1(A4V) or A β ₄₂-GFP from the vectors pETSOD1(A4V)-GFP or pETA β ₄₂-GFP, respectively, together with the cyclic peptides encoded by the selected clones 1-4 investigated in (C). The SOD1(A4V)-GFP fluorescence of the cell population producing a random cyclic peptide was arbitrarily set to 100. Experiments were carried out in replica triplicates ($n=1$ independent experiments) and the reported data correspond to the mean value $\pm$ s.e.m. (F) Solubility analysis of SOD1(A4V)-GFP overexpressed with/without the four selected cyclic peptide sequences shown in (C) by SDS-PAGE/western blotting using an anti-polyHis antibody. (G) DNA sequences of the peptide-encoding regions of the pSICLOPPS vectors contained in the selected clones tested in (C) along with the predicted amino acid sequences of the cyclic oligopeptides that they encode.

FIGS. 4A-4F: (A) Chemical structure of the selected cyclic pentapeptide SOD1C5-4. (B) Circular dichroism spectra of SOD1(A4V) incubated with/without the selected cyclic pentapeptides SOD1C5-4, A β C5-34 or A β C5-116 at room temperature for 90 d (1:1 and 5:1 indicate cyclic peptide:SOD1(A4V) molar ratios). Representative spectra from $n=2$ independent experiments are presented. (C) Dynamic light scattering analysis of SOD1(A4V) incubated with/without the selected cyclic pentapeptides at room temperature for 60 d. Representative data from $n=2$ independent experiments are presented. (D) Thioflavin T (ThT) fluorescence of SOD1(A4V) incubated with/without the selected cyclic pentapeptides at room temperature for 90 d. Representative data from $n=2$ independent experiments are presented. (E) Filter retardation assay of SOD1(A4V) incubated with/without the selected cyclic pentapeptides at 37 °C for 25 d. Representative images from $n=2$ independent experiments performed in replica duplicates are presented. (F) Differential scanning fluorimetry analysis of isolated SOD1(A4V) in the presence or absence of the selected cyclic pentapeptide SOD1C5-4 (5:1 peptide to protein ratio) using the conformation-sensitive dye SYPRO Orange. Data corresponding to $n=2$ independent experiments each one performed in three replicates $\pm$ s.e.m. are presented.

FIGS. 5A-5C: (A) HEK293 cells transiently expressing SOD1-GFP (top row) or SOD1(A4V)-GFP (middle and bottom rows) in the absence and presence of the selected cyclic pentapeptide SOD1C5-4 and visualized by confocal microscopy. (B) Relative number of cells containing SOD1 aggregates in the cultures described in (A). Aggregate-positive cells are presented as percentage of the total viable and GFP-positive cells. (C) Relative viability of cells

in the cultures described in (A). The viability of cells expressing wild-type SOD1 was arbitrarily set to 100.

FIGS. 6A-6E: (A) Distribution of the different types of selected cyclic oligopeptides among the bacterial clones selected for enhanced SOD1(A4V)-GFP fluorescence after the fourth round of FACS sorting (Fig. 3B). (B) Heat maps depicting the amino acid distribution in the sequences of the selected TXSXW pentapeptides after the fourth round of FACS sorting (Fig. 3B) as revealed by deep sequencing analysis. (C) Frequency of appearance of codons corresponding to the twenty natural amino acids at positions 2 and 4 of the peptide-encoding region of the pSICLOPPS-NuX₁X₂X₃-X₅ vectors contained in the bacterial clones after the fourth round of FACS sorting (Fig. 3B) that encoded for TXSXW cyclic pentapeptides (3,939,406 reads corresponding to TXSXW cyclic pentapeptides out of 4,243,704 total reads that appeared more than 50 times in the sorted peptide pool). (D) Frequency of appearance of the twenty natural amino acids at positions 2 and 4 of the unique TXSXW cyclic pentapeptides encoded by the pSICLOPPS-NuX₁X₂X₃-X₅ vectors contained in the bacterial clones isolated after the fourth round of FACS sorting (Fig. 3B) (46 unique peptide sequences corresponding to TXSXW cyclic pentapeptides out of 367 total unique selected peptide sequences that appeared more than 50 times in the sorted peptide pool). (E) Fluorescence of *E. coli* Origami 2(DE3) cells co-expressing SOD1-GFP, containing either wild-type SOD1 wt) or the ALS-associated variants SOD1(G37R), G(85R) or (G93A) from the corresponding pETSOD1-GFP vectors, together with the indicated selected cyclic peptides. Experiments were carried out in replica triplicates ($n=1$ independent experiments) and the reported data correspond to the mean value $\pm$ s.d.

FIG. 7: Sequences and frequency of appearance of the selected cyclic TXSXW pentapeptides as determined by high-throughput sequencing of the isolated pSICLOPPS-NuX₁X₂X₃-X₅ vectors after the fourth round of bacterial sorting for enhanced SOD1(A4V)-GFP fluorescence.

FIGS. 8A-8G: Depiction of the molecular evolution process with which macrocyclic rescuers of pathogenic A β misfolding and aggregation are identified.

FIGS. 9A-9E: Selected cyclic oligopeptides interfere with the normal A β aggregation process.

FIGS. 10A-10D: Selected cyclic oligopeptides inhibit A β -induced neurotoxicity *in vitro*.

FIGS. 11A-11D: Selected cyclic oligopeptides inhibit A β -induced neurotoxicity *in vivo*.

FIGS. 12A-12H: Next-generation sequencing and site-directed mutagenesis analyses may be used in order to identify all bioactive cyclic oligopeptide A β modulators contained in the tested cyclo-NuX₁X₂X₃-X₅ library and to facilitate structure-activity analyses of the isolated sequences.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide short, drug-like peptide macrocycles with the ability to rescue the misfolding, aggregation and associated pathogenic effects of a prominent aggregation-prone and disease-associated protein target. Particularly, the object of the present invention is to identify cyclic oligopeptides that rescue the misfolding and modulate the natural aggregation process of SOD1 and of its variants, which are implicated in pathogenicity of ALS and fALS. Various tetra- and penta- and hexapeptides with these properties are described in the present invention. More particularly the inventors have identified the general formula cyclo-NuX₁X₂...X_N, wherein Nu=C, S, or T; X is any of the twenty natural amino acid and N=3-5 as a very rich source of chemical rescuers of SOD1 misfolding and modulators of its aggregation.

Another aspect of the present invention is the identification of pentapeptide macrocycles with the general formula cyclo-TX₁SX₂W, wherein the first amino acid is Threonine, the third amino acid is a Serine, the last amino acid is Tryptophan and X is any amino acid, as effective and preferred misfolding rescuers and modulators of the natural process of SOD1. More preferred misfolding rescuers and modulators of SOD1 aggregation are cyclic oligopeptide sequences exhibiting the cyclo-T Ψ_1 S Ψ_2 W motif, where Ψ_1 = any amino acid excluding isoleucine (I), asparagine (N), glutamine (Q), methionine (M), glutamic acid (E), histidine (H), and lysine (K); and Ψ_2 = any amino acid excluding isoleucine (I), asparagine (N), glutamine (Q), cysteine (C), aspartic acid (D), glutamic acid (E), lysine (K) and proline (P). Even more preferred misfolding rescuers and modulators of SOD1 aggregation are cyclic oligopeptide sequences exhibiting the cyclo-T(Φ_1 ,S)S(Φ_2 ,M,H)W motif, where Φ_1 is preferably one of the hydrophobic (Φ) amino acids A, W or F, while Φ_2 is preferably V, W or

F. A small group of three cyclic pentapeptide rescuers with this general formula $T(\Phi_1, S)S(\Phi_2, M, H)W$ are analyzed further.

In the present invention, isolated cyclic oligopeptides are also provided, which comprise the amino acid sequences set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, up to SEQ ID NO:46. Nucleic acid sequences encoding a polypeptide of the invention are also provided. Vectors containing such nucleic acids, and cells containing such vectors, are also provided.

Another object of the present invention is to provide sufficient evidence that the selected peptide macrocycles are successful in inhibiting the misfolding and aggregation of SOD1. More particularly the inventors have studied the SOD1(A4V), a fALS-associated variant, whose misfolding and aggregation causes a very aggressive form of the disease with an average survival time of only 1.2 years after diagnosis. Indeed the effects of three selected oligopeptide macrocycles in inhibiting the misfolding and aggregation of SOD1 are shown using appropriate biochemical and/or biophysical assays.

Another object of the present invention is to provide sufficient evidence that the selected peptide macrocycles are successful in inhibiting the aggregation of SOD1* and the neurotoxicity caused by SOD1* aggregation *in vitro*. Indeed the effects of one selected pentapeptide macrocycle in inhibiting the aggregation and toxicity of SOD1(A4V) are shown in cultured mammalian cells.

Another objective of the present invention is to provide hybrid polypeptides that comprise a peptide motif that specifically interacts with the target polypeptide, which is then inserted into an appropriate protein scaffold. The polypeptide motif is inserted into the scaffold such that any desired function of the scaffold is retained and the inserted motif as presented retains its ability to specifically bind to the target and/or modulate the natural aggregation process of the target polypeptide. The scaffold can include, for example, neuroprotective agents to make SOD1 aggregates less toxic, aggregate-destroying molecules to eliminate amyloid SOD1 species, reagents that prevent SOD1 aggregate formation, or reagents useful for specifically imaging SOD1 aggregates in brain tissue.

It is also an object of the present invention to provide an integrated bacterial platform for the discovery of chemical/biological modulators of the problematic folding and aggregation of SOD1. In this system, large combinatorial libraries of genetically encoded macrocycles are biosynthesized in *E. coli* cells and are simultaneously screened for their ability to modulate the problematic folding and aggregation of SOD1 using a high-throughput genetic screen, which is based on the detection of enhanced fluorescence of chimeric SOD1-GFP fusions.

It is an object of the present invention to provide short, drug-like peptide macrocycles with the ability to rescue the misfolding, aggregation and associated pathogenic effects of a second prominent aggregation-prone and disease-associated protein target. Particularly, the object of the present invention is to identify cyclic oligopeptides that rescue the misfolding and modulate the natural aggregation process of the β -amyloid peptide, which is implicated in pathogenicity of Alzheimer's disease. Various tetra-, penta- and hexapeptides with these properties are described in the present invention. More particularly the inventors have identified the general formula cyclo-NuX₁X₂...X_N, wherein Nu=C, S, or T; X is any of the twenty natural amino acid and N=3-5 as a very rich source of chemical rescuers of A β misfolding and modulators of its aggregation.

Another aspect of the present invention is the identification of pentapeptide macrocycles with the general formula cyclo-TXXXR, wherein the first amino acid is Threonine, the last amino acid is Arginine and X is any amino acid, as effective and preferred modulators of the natural process of A β aggregation. More preferred modulators of A β aggregation are cyclic oligopeptide sequences exhibiting the cyclo-T Φ Z Π R motif, where Φ = any amino acid excluding phenylalanine (F), tryptophan (W), tyrosine (Y), glutamine (Q), aspartic acid (D), glutamic acid (E), lysine (K) and proline (P); Z is any amino acid excluding E; and Π is any natural amino acid excluding Q, methionine (M) and lysine (K). Even more preferred are cyclic oligopeptide sequences exhibiting the cyclo-T(T,A,V) Ψ (A,D,W)R motif, where Ψ is a non-negatively charged amino acid. A small group of six cyclic pentapeptide rescuers with this general formula cyclo-TXXXR are analyzed further.

Another aspect of the present invention is the identification of tetrapeptide macrocycles with the general formula cyclo-TXXR, wherein the first amino acid is Threonine, the last amino acid is Arginine and X is any amino acid, as effective and preferred modulators of the natural process of A β aggregation. More preferred modulators of A β aggregation are cyclic oligopeptide sequences exhibiting the cyclo-T Θ Λ R motif, where Θ = T, R, D, L, F or A, and Λ = C, R, S, G, Q, I, W, D, or F. A small group of two cyclic tetrapeptide rescuers with this general formula cyclo-TXXR are analyzed further.

Another aspect of the present invention is the identification of hexapeptide macrocycles as effective and preferred modulators of the natural process of A β aggregation. Preferred modulators of A β aggregation are cyclic oligopeptide sequences exhibiting the cyclo-NuX₁X₂X₃X₄X₅ motif for the use in rescuing A β misfolding and modulating aggregation, wherein Nu is T; wherein X₁ is any amino acid selected from I, L, V, C, S, K or P, and is more preferably P, V or L; wherein X₂ is selected from A, I, L, V, F, W, C, S, T, D, E, H, K, P, or G

and is more preferably V or A; wherein X_3 is selected from I, L, V, F, W, Y, E or R, and is more preferably W; wherein X_4 is selected from L, V, F, Y, S or R and is more preferably F; wherein X_5 is selected from W, M, N, D or E and is more preferably D. The cyclic hexapeptide comprising the sequence TPVWFD is analyzed further.

5 In the present invention, isolated cyclic oligopeptides are also provided, which comprise the amino acid sequences set forth in SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, up to SEQ ID NO:205. The inventors also identified the pentapeptide macrocycles cyclo-SASPT (SEQ ID NO:206), cyclo-SHSPT (SEQ ID NO:207), cyclo-SICPT (SEQ ID NO:208) and cyclo-SITPT (SEQ ID NO:209) as effective and preferred modulators
10 of the natural process of A β aggregation. Furthermore the inventors also identified the tetrapeptide macrocycles cyclo-TTCR (SEQ ID NO:210), cyclo-TTRR (SEQ ID NO:211), cyclo-TTSR (SEQ ID NO:212), cyclo-TTGR (SEQ ID NO:213), cyclo-TRGR (SEQ ID NO:214), cyclo-TRRR (SEQ ID NO:215), cyclo-TDQR (SEQ ID NO:216), cyclo-TLIR (SEQ ID NO:217), cyclo-TLWR (SEQ ID NO:218), cyclo-TLGR (SEQ ID NO:219), cyclo-TFDR
15 (SEQ ID NO:220), and cyclo-TAFR (SEQ ID NO:221) as effective and preferred modulators of the natural process of A β aggregation. Finally, the inventors also identified the hexapeptide macrocycles comprising the amino acid sequences set forth in SEQ ID NO:222, SEQ ID NO:223, ..., up to SEQ ID NO:255 as effective and preferred modulators of the natural process of A β aggregation. Nucleic acid sequences encoding a polypeptide of the invention are also
20 provided. Vectors containing such nucleic acids, and cells containing such vectors, are also provided.

Another object of the present invention is to provide sufficient evidence that the selected peptide macrocycles are successful in modulating A β aggregation. Indeed the effects of ten selected oligopeptide macrocycles in modulating A β aggregation are shown using
25 appropriate biochemical and/or biophysical assays.

Another object of the present invention is to provide sufficient evidence that the selected peptide macrocycles are successful in inhibiting the neurotoxicity caused by A β aggregation *in vitro*. Indeed the effects of two selected pentapeptide macrocycles in inhibiting neurotoxicity of A β aggregation are shown in cultured primary hippocampal neurons.

30 A main objective of the present invention is to show the successful *in vivo* rescuing of the misfolding and aggregation of β -amyloid peptide by the selected macrocycles produced by the technique described in the present invention. The protective effects of the selected cyclic peptides against A β aggregation and toxicity *in vivo* are demonstrated in established AD animal models in the nematode *Caenorhabditis elegans* (*C. elegans*).

Another objective of the present invention is to provide hybrid polypeptides that comprise a peptide motif that specifically interacts with the target polypeptide, which is then inserted into an appropriate protein scaffold. The polypeptide motif is inserted into the scaffold such that any desired function of the scaffold is retained and the inserted motif as presented
5 retains its ability to specifically bind to the target and/or modulate the natural aggregation process of the target polypeptide. The scaffold can include, for example, neuroprotective agents to make amyloid plaques less toxic, amyloid destroying molecules to eliminate plaques, reagents that prevent amyloid plaque formation, or reagents useful for specifically imaging amyloid plaques in brain tissue.

10 It is also an object of the present invention to provide an integrated bacterial platform for the discovery of chemical/biological modulators of the problematic folding and aggregation of A β . In this system, large combinatorial libraries of genetically encoded macrocycles are biosynthesized in *E. coli* cells and are simultaneously screened for their ability to modulate the problematic folding and aggregation of A β using a high-throughput genetic screen, which is
15 based on the detection of enhanced fluorescence of chimeric A β -GFP fusions.

It is also an object of the present invention to provide a generalizable integrated bacterial platform for the discovery of chemical/biological modulators of the problematic folding and aggregation of disease-associated, misfolding-prone proteins (MisPs). In this system, large combinatorial libraries of genetically encoded macrocycles are biosynthesized in
20 *E. coli* cells and are simultaneously screened for their ability to modulate the problematic folding and aggregation of the target MisP using a high-throughput genetic screen, which is based on the detection of enhanced fluorescence of chimeric MisP-GFP fusions.

In some embodiments, the present disclosure provides a method of identifying modulators of a misfolding-prone protein associated with a protein misfolding disease, comprising: (A) obtaining a population of transformed bacterial cells that co-express: (a) a
25 nucleic acid encoding a library of peptide macrocycles, operably linked to a promoter and (b) a bipartite nucleic acid comprising a sequence for a gene encoding a misfolding-prone protein associated with a protein misfolding disease (MisP) and a sequence encoding a protein reporter; (B) identifying bacterial cells of step (A) that exhibit enhanced levels of protein reporter activity; and (C) identifying the bioactive peptide macrocycles in the library that modulate
30 MisP misfolding. In some aspects, the protein reporter is a fluorescent protein (FP) reporter, and step (B) comprises identifying bacterial cells that exhibit enhanced levels of MisP-FP fluorescence. In some aspects, the protein reporter is an enzyme. In some aspects, the identification in step (B) comprises selection. In some aspects, step (C) comprises sequencing

the nucleic acid of step (Aa). In certain aspects, the nucleic acids of (a) and (b) are encoded on the same vector. In particular aspects, the vector is a plasmid.

In aspects of the embodiments, said MisP is selected from β -amyloid peptide, SOD1, tau, α -synuclein, polyglutaminated huntingtin, polyglutaminated ataxin-1, polyglutaminated ataxin-2, polyglutaminated ataxin-3, prion protein, islet amyloid polypeptide (amylin), β 2-microglobulin, fragments of immunoglobulin light chain, fragments of immunoglobulin heavy chain, serum amyloid A, ABri peptide, ADan peptide, transthyretin, apolipoprotein A1, gelsolin, transthyretin, lysozyme, phenylalanine hydroxylase, apolipoprotein A-I, calcitonin, prolactin, TDP-43, FUS/TLS; insulin, hemoglobin, α 1-antitrypsin, p53, or variants thereof. In some aspects, said peptide macrocycle can be a ribosomally synthesized as a head-to-tail cyclic peptide, side-chain-to-tail cyclic peptide, bicyclic peptide, lanthipeptide, linaridin, proteusin, cyanobactin, thiopeptide, bottromycin, microcin, lasso peptide, microviridin, amatoxin, phalloxin, θ -defensin, orbitide, or cyclotide. In some aspects, the disease is selected from amyotrophic lateral sclerosis, Alzheimer's disease, amyotrophic lateral sclerosis, Parkinson's disease, Huntington's disease, Creutzfeldt-Jakob disease, cancer, phenylketonuria, type 2 diabetes, senile systemic amyloidosis, familial amyloidotic polyneuropathy, familial amyloid cardiomyopathy, leptomeningeal amyloidosis, systemic amyloidosis, familial British dementia, familial Danish dementia, light chain amyloidosis, heavy chain amyloidosis, serum amyloid A amyloidosis, lysozyme amyloidosis, dialysis-related amyloidosis, ApoAI amyloidosis, Finnish type familial amyloidosis, hereditary cerebral hemorrhage with amyloidosis (Icelandic type), medullary carcinoma of the thyroid, pituitary prolactinoma, injection-localized amyloidosis, frontotemporal dementia, spinocerebellar ataxia 1, spinocerebellar ataxia 2, spinocerebellar ataxia 3, α 1-antitrypsin deficiency, sickle-cell anemia, or transmissible spongiform encephalopathy. In further aspects, the method comprises recombinantly producing or chemically synthesizing the identified bioactive peptide macrocycle.

In some aspects, there is provided a method of treatment, prevention or diagnosis of a protein misfolding disease comprising administering to a subject a therapeutically effective amount of the bioactive peptide macrocycle identified by the embodiments and aspects of the invention provided herein. In another aspect, there is provided a pharmaceutical composition comprising the bioactive peptide macrocycle according to the embodiments or aspects of the invention provided herein, and a pharmaceutically acceptable carrier. In some aspects, the pharmaceutical composition may be used for the treatment or prevention of a protein misfolding disease.

In some aspects, there is provided a hybrid molecule comprising: a) a peptide macrocycle identified or produced according to the method of any of the embodiments or aspects of the invention provided herein, and b) a scaffold molecule linked to the peptide macrocycle. In some aspects, the scaffold molecule is a diagnostic or a therapeutic reagent. In particular aspects, the therapeutic reagent is a cytoprotective agent that renders the aggregates of the target protein less toxic or inhibits target protein aggregate formation. In some aspects, the scaffold molecule comprises all or a sufficient portion of a protein selected from the group consisting of antibodies, enzymes, chromogenic proteins, and fluorescent proteins. In specific aspects, the diagnostic reagent specifically targets protein aggregates in diseased or healthy tissue.

In some aspects, there is provided a method of treating or diagnosing a protein misfolding disease associated with aberrant aggregate formation, the method comprising administering a hybrid molecule according to the embodiments and aspects provided herein, wherein the peptide macrocycle of the hybrid molecule specifically interacts with the amyloid or non-amyloid form of the target MisP.

In some embodiments, there is provided a peptide comprising the amino acid sequence $NuX_1X_2...X_N$, wherein: Nu is T; $N=4$; X_1 is any amino acid excluding I, N, Q, M, E, H, and K; $X_2=S$; X_3 is any amino acid excluding I, N, Q, C, D, E, K and P; and $X_4=W$, wherein the specifically interacts with the amyloid or non-amyloid form of SOD1 and/or mutant SOD1.

In some aspects, the X_1 is A, L, V, F, W, Y, C, S, T, D, R, P or G. In particular aspects, the X_1 is S, A, F or W. In some aspects, the X_3 is A, L, V, F, W, Y, M, S, T, R, H or G. In certain aspects, the X_3 is V, F, W, M, or H. In specific aspects, the X_3 is W. In some aspects, Nu is T; $N=4$; X_1 is A, L, V, F, W, Y, C, S, T, D, R, P or G; $X_2=S$; X_3 is A, L, V, F, W, Y, M, S, T, R, H or G; and $X_4=W$. In particular aspects, Nu is T; $N=4$; X_1 is S, A, F or W; $X_2=S$; X_3 is V, F, W, M, or H; and $X_4=W$.

In some embodiments, there is provided a peptide comprising the amino acid sequence set forth in any one of SEQ ID NO:1-46, wherein the peptide prevents misfolding and aggregation of SOD1 and/or mutant SOD1. In some aspects, the peptide comprises an amino acid sequence selected from TWSVW, TASFW, and TFMSW. In some aspects, said peptide prevents misfolding and aggregation of SOD1 and/or mutant SOD1. In some aspects, at least one position of the peptide is a D amino acid. In certain aspects, the peptide is a cyclic peptide. In other aspects, the peptide is a linear peptide.

In some aspects, there is provided a hybrid molecule comprising: a) a peptide set forth in the embodiments and aspects provided herein, and b) a scaffold molecule. In some aspects,

the scaffolding molecule comprises a cell penetrating peptide. In specific aspects, the scaffold molecule comprises a diagnostic or therapeutic reagent. In certain aspects, the scaffold molecule comprises a polypeptide, small molecule or compound. In some aspects, the scaffold molecule comprises all or a sufficient portion of a protein selected from the group consisting of antibodies, enzymes, chromogenic proteins, fluorescent proteins and fragments thereof. In some aspects, the therapeutic agent is a neuroprotective agent that renders SOD1 aggregates less toxic or inhibits SOD1 aggregate formation. In some aspects, the diagnostic reagent specifically images SOD1 aggregates in neuronal tissue.

In some aspects of the embodiments provided herein, there is provided a peptide or a molecule to inhibit protein misfolding and aggregation wherein the peptide prevents misfolding and aggregation of SOD1 and/or mutant SOD1. In particular, there is provided a method of treatment, prevention or diagnosis of amyotrophic lateral sclerosis comprising administering to a subject a therapeutically effective amount of a peptide or hybrid molecule according to any one of the embodiments or aspects provided herein.

In some aspects, there is provided a pharmaceutical composition comprising a peptide or hybrid molecule according to any of the aspects or embodiments provided herein, and a pharmaceutically acceptable carrier. In some aspects, the pharmaceutical composition is used for the treatment or prevention of amyotrophic lateral sclerosis. In some aspects, there is provided an isolated nucleic acid sequence encoding the peptide of the aspects or embodiments provided herein. In some aspects, there is a vector comprising said nucleic acid sequence. In some aspects, the vector is an expression vector. In certain aspects, there is provided a host cell comprising said vector. In certain aspects, the host cell is a prokaryotic or eukaryotic cell.

In some embodiments, there is provided a peptide wherein the peptide comprises the amino acid sequence NuX₁X₂...X_N, wherein: (A) Nu=T; N=3; X₁ is selected from T, R, D, L, F or A; X₂ is selected from C, R, S, G, Q, I, W, D, or F; and X₃=R; (B) Nu=T; N=4; X₁ is any amino acid excluding F, W, Y, Q, D, E, K and P; X₂ is any amino acid excluding E; X₃ is any amino acid excluding Q, M and K; and X₄ is R; or (C) Nu=T; N=5; X₁ is I, L, V, C, S, K or P; X₂ is any amino acid excluding Y, N, Q, M, and R; X₃ is selected from I, L, V, F, W, Y, E, or R; X₄ is selected from L, V, F, Y, S or R; and X₅ is selected from W, M, N, D, or E, wherein the peptide specifically interacts with a monomeric, oligomeric and/or amyloid form of the A β peptide. In some aspects, Nu=T; N=3; X₁ is selected from T, R, D, L, F or A; X₂ is selected from C, R, S, G, Q, I, W, D, or F; and X₃=R. In certain aspects, Nu=T; N=4; X₁ is any amino acid excluding F, W, Y, Q, D, E, K and P; X₂ is any amino acid excluding E; X₃ is any amino acid excluding Q, M and K; and X₄ is R. In specific aspects, X₁ is selected from A, I, L, V, N,

C, M, S, T, R, H, or G. In particular aspects, X₁ is T, V or A. In some aspects, X₂ is selected from A, I, L, V, F, W, Y, N, Q, C, M, S, T, D, R, H, K, P or G. In certain aspects, X₂ is I, L, V, F, Y or T. In specific aspects, X₃ is A, I, L, V, F, W, Y, N, C, S, T, D, E, R, H, P or G. In particular aspects, X₃ is A, W, or D. In some aspects, Nu=T; N=4; X₁ is selected from A, I, L, V, N, C, M, S, T, R, H, or G; X₂ is selected from A, I, L, V, F, W, Y, N, Q, C, M, S, T, D, R, H, K, P or G; X₃ is A, I, L, V, F, W, Y, N, C, S, T, D, E, R, H, P or G; and X₄ is R. In certain aspects, Nu=T; N=4; X₁ is T, V or A; X₂ is I, L, V, F, Y or T; X₃ is A, W, or D; and X₄ is R. In specific aspects, Nu=T; N=5; X₁ is I, L, V, C, S, K or P; X₂ is any amino acid excluding Y, N, Q, M, and R; X₃ is selected from I, L, V, F, W, Y, E, or R; X₄ is selected from L, V, F, Y, S or R; and X₅ is selected from W, M, N, D, or E. In particular aspects, X₁ is selected from I, L, V, C, S, K or P. In specific aspects, the X₁ is P, V or L. In some aspects, X₂ is selected from A, I, L, V, F, W, C, S, T, D, E, H, K, P, or G. In certain aspects, X₂ is V or A. In some aspects, X₃ is selected from I, L, V, F, W, Y, E or R. In specific aspects, X₃ is W. In some aspects, X₄ is selected from L, V, F, Y, S or R. In specific aspects, X₄ is F. In some aspects, X₅ is selected from W, M, N, D, or E. In specific aspects, X₅ is D. In some aspects, Nu=T; N=5; X₁ is P, V or L; X₂ is V or A; X₃ is selected from I, F, or W; X₄ is selected from L, F, or R; and X₅ is selected from W, N, or D. In certain aspects, Nu=T; N=5; X₁ is P, V or L; X₂ is V or A; X₃ is W; X₄ is F; and X₅ is D. In some aspects, the peptide comprises the sequence TTTCR, TTRR, TTTSR, TRGR, TTGR, TRRR, TDQR, TLIR, TLWR, TLGR, TFDR, or TAFR (SEQ ID NOs 210-221).

In some embodiments, there is provided a peptide comprising the amino acid sequence set forth in any one of SEQ ID NO:47-209, wherein, the peptide is cyclic and specifically interacts with a monomeric, oligomeric and/or amyloid form of the A β peptide. In some aspects, the peptide comprises an amino acid sequence selected from TAFDR, TAWCR, TTWCR, TTVDR, TTYAR, TTTAR or SASPT. In some aspects, the the peptide comprises an amino acid sequence selected from the amino acid sequences set forth in SEQ ID NO:176-209. In some aspects, the peptide comprises an amino acid sequence of TPVWFD (SEQ ID NO:176) or TPAWFD (SEQ ID NO:177). In some aspects, at least one position of the peptide is a D amino acid. IN some aspects, the peptide is a cyclic peptide. In other aspects, the peptide is a linear peptide.

In some aspects, there is provided a hybrid molecule comprising: a) a peptide set forth in any one of the embodiments or aspects, that specifically interacts with a monomeric, oligomeric and/or amyloid form of the A β peptide; and b) a scaffold molecule. In some aspects, the scaffolding molecule comprises a cell penetrating peptide. In certain aspects, the scaffold

molecule comprises a diagnostic or therapeutic reagent. In particular aspects, the scaffold molecule comprises a polypeptide, small molecule or compound. In specific aspects, the polypeptide comprises all or a sufficient portion of a protein selected from the group consisting of antibodies, enzymes, chromogenic proteins, or a fluorescent protein. In some aspects, the therapeutic agent is a neuroprotective agent that renders amyloid plaques less toxic or inhibits plaque formation. In some aspects, the diagnostic reagent specifically images oligomers and/or amyloid aggregates in neuronal tissue.

In some aspects of the invention, there is provided methods for the use of a peptide or molecule according to any of the embodiments or aspects provided herein, to inhibit protein misfolding and aggregation. In some aspects, said peptide prevents misfolding and aggregation of the β -amyloid peptide.

In some embodiments, there is provided a method of treatment, prevention or diagnosis of a disease related to protein misfolding and aggregation, comprising administering to a subject a therapeutically effective amount of a peptide or molecule according to any of the embodiments or aspects provided herein, wherein the disease is selected from Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, Creutzfeldt-Jakob disease, type 2 diabetes, familial amyloidotic polyneuropathy, systemic amyloidosis, and transmissible spongiform encephalopathy. In some aspects, there is provided a method of treatment, prevention or diagnosis of Alzheimer's disease comprising administering to a subject a therapeutically effective amount of a peptide according to any one of the aspects or embodiments provided herein. In some aspects, there is provided a pharmaceutical composition comprising a peptide according to any one of the aspects or embodiments provided herein and a pharmaceutically acceptable carrier. In some aspects, there is provided a nucleic acid encoding any of said peptides. In some aspects, there is provided a vector comprising said nucleic acid. In some aspects, the vector is an expression vector. In some aspects, there is provided a host cell comprising said vector. In some cases, the host cell is a prokaryotic or eukaryotic cell.

Other objects and advantages of the present invention will become apparent to those skilled in the art in view of the following detailed description.

DETAILED DESCRIPTION OF THE INVENTION

I. General

It is now well established that fALS-linked amino acid substitutions in SOD1 introduce a toxic-gain-of-function property in SOD1 by causing protein misfolding and aggregation, and

the formation of oligomeric/aggregated SOD1 species, which are highly toxic for motor neurons. Gradual accumulation of such toxic oligomers/aggregates of mutated SOD1 initiates motor neuron degeneration and the development of fALS. Also, the accumulation of misfolded and aggregated wild-type SOD1 has been implicated in sporadic forms of ALS. This invention
5 pertains to compounds, and pharmaceutical compositions thereof, that can modulate the aggregation of amyloidogenic proteins and peptides, in particular compounds that can rescue the misfolding and inhibit the aggregation of SOD1 or of its fALS-associated variants and inhibit the neurotoxicity of these aggregated SOD1 species. A compound of the invention that modulates aggregation of SOD1, referred to herein interchangeably as a SOD1 modulator
10 compound, a SOD1 modulator or simply a modulator, alters the aggregation of SOD1 or of its fALS-associated variants when the modulator is contacted with SOD1 or of its fALS-associated variants. Thus, a compound of the invention acts to alter the natural aggregation process or rate of SOD1 or of its fALS-associated variants, thereby disrupting the normal course this process. A modulator which inhibits SOD1 and/or mutant SOD1 aggregation (an
15 “inhibitory modulator compound”) can be used to prevent or delay the onset of the deposition of SOD1 and/or mutant SOD1 aggregates. Moreover, inhibitory modulator compounds of the invention inhibit the formation and/or activity of neurotoxic aggregates of SOD1 or of its fALS-associated variants (i.e., the inhibitory compounds can be used to inhibit the neurotoxicity of SOD1 or of its fALS-associated variants).

20 Alternatively, in another embodiment, a modulator compound of the invention promotes the aggregation of SOD1 and/or mutant SOD1. The various forms of the term “promotion” refer to an increase in the amount and/or rate of SOD1 and/or mutant SOD1 aggregation in the presence of the modulator, as compared to the amount and/or rate of SOD1 and/or mutant SOD1 aggregation in the absence of the modulator. Such a compound which
25 promotes SOD1 and/or mutant SOD1 aggregation is referred to as a stimulatory modulator compound. Stimulatory modulator compounds may be useful, for example, in decreasing the amounts of neurotoxic SOD1 and/or mutant SOD1 oligomeric species by driving the natural SOD1 aggregation process towards the (possibly) less neurotoxic higher-order SOD1 and/or mutant SOD1 aggregates.

30 Compounds of the present invention may inhibit SOD1 and/or mutant SOD1 aggregation and/or oligomerization. In particular, preferred modulator compounds of the invention comprise cyclic oligopeptides with the general formula cyclo-NuX₁X₂...X_N, where X is any one of the twenty natural amino acids, N=3-5 and Nu = cysteine (Cys or C), serine (Ser or S) or threonine (Thr or T), which is sufficient to alter (and preferably inhibit) the natural

aggregation process or rate of SOD1 and/or mutant SOD1. This SOD1 and/or mutant SOD1 modulator can comprise as few as four amino acid residues (or derivative, analogues or mimetics thereof).

According to the prevalent amyloid cascade hypothesis, the high tendency of A β for
5 misfolding and aggregation results in the formation of neurotoxic A β oligomers/aggregates, whose accumulation ultimately leads to neuron degeneration and the development of the disease. This invention pertains to compounds, and pharmaceutical compositions thereof, that can modulate the aggregation of amyloidogenic proteins and peptides, in particular compounds that can modulate the aggregation of the β -amyloid peptide (A β) and inhibit the neurotoxicity
10 of A β . A compound of the invention that modulates aggregation of A β , referred to herein interchangeably as a A β modulator compound, a A β modulator or simply a modulator, alters the aggregation of natural A β when the modulator is contacted with natural A β . Thus, a compound of the invention acts to alter the natural aggregation process or rate of A β , thereby disrupting the normal course this process. A modulator which inhibits A β aggregation (an
15 “inhibitory modulator compound”) can be used to prevent or delay the onset of β -amyloid deposition. Moreover, inhibitory modulator compounds of the invention inhibit the formation and/or activity of neurotoxic aggregates of natural A β peptide (i.e., the inhibitory compounds can be used to inhibit the neurotoxicity of A β).

Alternatively, in another embodiment, a modulator compound of the invention
20 promotes the aggregation of natural A β peptides. The various forms of the term “promotion” refer to an increase in the amount and/or rate of A β aggregation in the presence of the modulator, as compared to the amount and/or rate of A β aggregation in the absence of the modulator. Such a compound which promotes A β aggregation is referred to as a stimulatory modulator compound. Stimulatory modulator compounds may be useful, for example, in
25 decreasing the amounts of neurotoxic A β oligomeric species by driving the natural A β aggregation process towards the (generally) less neurotoxic higher-order A β aggregates.

Compounds of the present invention may inhibit A β aggregation and/or oligomerization. In particular, preferred modulator compounds of the invention comprise cyclic oligopeptides with the general formula cyclo-NuX₁X₂...X_N, where X is any one of the
30 twenty natural amino acids, N=3-5 and Nu = cysteine (Cys or C), serine (Ser or S) or threonine (Thr or T), which is sufficient to alter (and preferably inhibit) the natural aggregation process or rate A β . This A β modulator can comprise as few as four amino acid residues (or derivative, analogues or mimetics thereof).

II. Discovery of peptide SOD1 and A β modulators

The present application describes the invention of a generalizable bacterial platform for the discovery of macrocyclic peptide rescuers of the misfolding of disease-associated, misfolding-prone proteins (MisPs). The inventors demonstrate the generalizability of this integrated bacterial platform by discovering macrocyclic peptides that modulate the problematic folding and aggregation of SOD1, mutant SOD1, or A β .

This approach offers a number of important advantages. First, it allows the screening of molecular libraries with expanded diversities. Here, a library with diversity >10 million different macrocycles has been investigated.

In another embodiment, the present invention can be applied to construct and screen macrocyclic libraries with diversities up to 10^{10} different molecules. Importantly, *E. coli* can support the *in vivo* biosynthesis not only of head-to-tail cyclic peptides like the ones investigated here, but also of other macrocyclic structures, such as side-chain-to-tail cyclic peptides, bicyclic peptides, cyclotides, macrolides and other macrocyclic structures, which can accommodate not only naturally occurring amino acids, but a large variety of artificial ones as well.

In addition, the analysis of these large libraries is carried out using a very high-throughput genetic screen, which enables the identification of bioactive molecules simply by isolating compounds that enhance the fluorescence of *E. coli* cells expressing MisP-GFP, such as SOD1*-GFP or A β -GFP fusions, by flow cytometric sorting (FACS). Compared to affinity-based approaches for screening DNA-encoded chemical libraries, such as phage and mRNA display, the herein described approach does not detect mere target MisP, such as SOD1* or A β , binding, but selects directly the bioactive compounds with the ability to rescue MisP, such as SOD1* or A β , misfolding, without requiring the availability of purified MisP, such as SOD1* or A β .

Moreover, synthesis of the studied compounds and their screening for bioactivity are carried out *in vivo* as part of a single-step process, without the need for laborious organic synthesis and product isolation steps. Importantly, screening for bioactivity is carried out in a fully unbiased manner without requiring a priori knowledge of the structures of the MisP, such as SOD1* or A β , monomers, oligomers, or higher-order aggregates, specific assumptions about possible binding sites, or prior preparation of specific MisP, such as SOD1* or A β , oligomerization states.

More particularly, combinatorial libraries of random cyclic tetra-, penta-, and hexapeptides have been created and the most prominent targets have been selected to study as potential rescuers of SOD1 and A β misfolding. The technology described in the present

invention utilizes a technique termed split intein circular ligation of peptides and proteins (SICLOPPS) for producing peptide libraries in *E. coli*. SICLOPPS uses split inteins, i.e. self-splicing protein elements for performing N- to C- terminal peptide cyclization and biosynthesize cyclic peptides as short as four amino acids long. The only requirement for the intein splicing reaction and peptide cyclization to occur is the presence of a nucleophilic amino acids cysteine (C), serine (S), or threonine (T) as the first amino acid of the extein following the C-terminus of the intein.

According to the present invention, peptides belonging to the general formula $NuX_1X_2...X_N$, can be used for rescuing protein misfolding and modulating protein aggregation; wherein X is any one of the twenty natural amino acids, $N = 3-5$ and Nu is selected from cysteine (Cys or C), serine (Ser or S) or threonine (Thr or T). According to the preferred embodiment of the present invention the peptide is a cyclic peptide. According to the preferred embodiment of the present invention Nu is T.

In a preferred embodiment of the present invention the peptide with the general formula $cyclo-NuX_1X_2...X_N$, for the use in protein misfolding and aggregation, has the following specifications wherein N is 3, wherein Nu is T; wherein X_1 is selected from T, R, D, L, F or A, wherein X_2 is selected from C, R, S, G, Q, I, W, D, or F; wherein X_3 is R. According to the above specification the preferred cyclic tetrapeptide is selected from cyclo-TTCR (SEQ ID NO:164), cyclo-TTRR (SEQ ID NO:165), cyclo-TTSR (SEQ ID NO:166), cyclo-TTGR (SEQ ID NO:167), cyclo-TRGR (SEQ ID NO:168), cyclo-TRRR (SEQ ID NO:169), cyclo-TDQR (SEQ ID NO:170), cyclo-TLIR (SEQ ID NO:171), cyclo-TLWR (SEQ ID NO:172), cyclo-TLGR (SEQ ID NO:173), cyclo-TFDR (SEQ ID NO:174), and cyclo-TAFR (SEQ ID NO:175) as effective and preferred modulators of the natural process of A β aggregation.

In another preferred embodiment of the present invention the peptide with the general formula $NuX_1X_2...X_N$, for the use in rescuing protein misfolding and modulating, has the following specifications wherein N is 4, wherein Nu is preferably T; wherein X_1 is any amino acid excluding I, N, Q, M, E, H, and K, and more preferably it is S, A, W, or F; wherein X_2 is preferably S; wherein X_3 is any amino acid excluding I, N, Q, C, D, E, K and P, and is more preferably selected from V, W, F, M, or H; wherein X_4 is preferably W. According to the above specification the preferred pentapeptide is selected from the amino acid sequences set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, ..., up to SEQ ID NO:46. According to the above specification the more preferred pentapeptide is selected from cyclo-TASFW (SEQ ID NO:2), cyclo-TWSVW (SEQ ID NO:4), and cyclo-TFSMW (SEQ ID NO:6).

In another preferred embodiment of the present invention the peptide with the general formula $NuX_1X_2...X_N$, for the use in rescuing protein misfolding and modulating aggregation, has the following specifications wherein N is 4, wherein Nu is T or S; wherein X_1 is any amino acid excluding F, W, Y, Q, D, E, K and P, preferably it is S, H, T, V or A, and more preferably it is T, V or A; wherein X_2 is any amino acid excluding E, preferably a non-negatively charged amino acid, and more preferably it is selected from I, L, V, F, W, Y, M, S, T, R, H, or G; wherein X_3 is any amino acid excluding Q, M and K, and is more preferably selected from A, V, F, W, C, S, T, D, C, R, H, P or G; wherein X_4 is preferably R or T. According to the above specification the preferred pentapeptide is selected from the amino acid sequences set forth in SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, ..., up to SEQ ID NO: 205. According to the above specification the more preferred pentapeptide is selected from cyclo-TAFDR (SEQ ID NO:86), cyclo-TAWCR (SEQ ID NO:63), cyclo-TTWCR (SEQ ID NO:60), cyclo-TTVDR (SEQ ID NO:48), cyclo-TTYAR (SEQ ID NO:47), cyclo-TTTAR (SEQ ID NO:56), and cyclo-SASPT (SEQ ID NO:206).

In another preferred embodiment of the present invention the peptide with the general formula $NuX_1X_2...X_N$, for the use in rescuing protein misfolding and modulating aggregation, has the following specifications wherein N is 5, wherein Nu is T; wherein X_1 is any amino acid selected from I, L, V, C, S, K or P, and is more preferably P, V or L; wherein X_2 is selected from A, I, L, V, F, W, C, S, T, D, E, H, K, P, or G and is more preferably V or A; wherein X_3 is selected from I, L, V, F, W, Y, E or R, and is more preferably W; wherein X_4 is selected from L, V, F, Y, S or R and is more preferably F; wherein X_5 is selected from W, M, N, D or E and is more preferably D. The hexapeptide according to the above specifications is selected from the amino acid sequences set forth in SEQ ID NO:222, SEQ ID NO:223, ..., up to SEQ ID NO:255, and is most preferably TPVWFD (SEQ ID NO:222) or TPAWFD (SEQ ID NO:223).

The maximum theoretical diversity of the combined cyclo- $NuX_1X_2X_3-X_5$ library investigated here was > 10 million different sequences. The libraries of genes encoding this combinatorial library of random cyclic oligopeptides were constructed using degenerate codons. The inventors constructed the high diversity pSICLOPPS- $NuX_1X_2X_3-X_5$ vector library which is expected to be encoding the vast majority of the theoretically possible designed cyclic tetra-, penta-, and hexapeptide cyclo- $NuX_1X_2X_3-X_5$ sequences using molecular biology techniques already known and used in the art.

The invention provided herein can be used as a method of treatment, prevention or diagnosis of all diseases related to protein misfolding and aggregation, including but not limited

to amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Parkinson's disease, Huntington's disease, Creutzfeldt-Jakob disease, type 2 diabetes, familial amyloidotic polyneuropathy, systemic amyloidosis, and transmissible spongiform encephalopathy comprising administering to a subject a therapeutically effective amount of a peptide. Preferably the invention presented
5 herein can be used as a method of treatment, prevention or diagnosis of amyotrophic lateral sclerosis or Alzheimer's disease.

To identify cyclic oligopeptide sequences with the ability to interfere with the problematic folding of A β and modulate its oligomerization/aggregation, a bacterial high-throughput genetic screen was utilized. This system monitors A β misfolding and aggregation
10 by measuring the fluorescence of *E. coli* cells overexpressing a chimeric fusion of the human A β_{42} with GFP. It has been demonstrated previously that due to the high aggregation propensity of A β , *E. coli* cells overexpressing A β -GFP fusions produce misfolded fusion protein that accumulates into insoluble inclusion bodies that lack fluorescence, despite the fact that they express these fusions at high levels. Mutations in the coding sequence of A β or the addition of
15 compounds that inhibit A β aggregation, however, result in the formation of soluble and fluorescent A β -GFP, and bacterial cells expressing A β -GFP under these conditions acquire a fluorescent phenotype. The inventors of the present invention adapted this system to perform screening for aggregation-inhibitory macrocycles in a very high-throughput fashion by isolating cyclic oligopeptide-producing bacterial clones that exhibit enhanced levels of A β_{42} -
20 GFP fluorescence using fluorescence-activated cell sorting (FACS).

Herein, the inventors describe that the integrated bacterial platform for the discovery of macrocyclic rescuers that modulate the problematic folding and aggregation of A β as described in the present invention, is also generalizable, i.e., it can be more generally applied for the discovery of macrocyclic peptide rescuers of the misfolding of other disease-associated,
25 misfolding-prone proteins (MisPs) as well. To demonstrate this generalizability, the inventors have used the same system to discover macrocyclic peptides that modulate the problematic folding and aggregation of SOD1 and/or mutant SOD1.

It has been demonstrated previously that the fluorescence of *E. coli* cells expressing a recombinant protein whose C terminus is fused to GFP correlates well with the amount of
30 soluble and folded protein (Waldo GS, Standish BM, Berendzen J, Terwilliger TC. Nat Biotechnol. 1999 Jul;17(7):691-5). Based on this, it was reasoned that the fluorescence of MisP-GFP fusions can serve as a reliable reporter for the identification of chemical rescuers of MisP misfolding for a number of disease-associated MisPs, including SOD1. In order to test this hypothesis, the inventors generated fusions of SOD1 variants, whose misfolding and

aggregation have been linked with the pathology of familial forms of ALS (fALS), with GFP. Expression of these fusions in *E. coli*, yielded levels of cellular fluorescence, which were significantly decreased compared to that of the generally non-pathogenic, wild-type SOD1. Western blot analysis indicated that this occurs because the accumulation of soluble SOD1-GFP is decreased in the presence of misfolding-inducing amino acid substitutions, which in turn takes place due to enhanced misfolding/aggregation of fusion-free SOD1. Thus, as in the case of A β , the fluorescence of *E. coli* cells overexpressing SOD1-GFP fusions appears to be a good indicator of SOD1 folding and misfolding.

To identify rescuers of disease-associated SOD1 misfolding, the inventors screened for cyclic oligopeptides that inhibit the aggregation of SOD1(A4V), a fALS-associated variant, whose misfolding and aggregation causes a very aggressive form of the disease with an average survival time of only 1.2 years after diagnosis.

E. coli BL21(DE3) cells producing the combined cyclo-NuX₁X₂X₃-X₅ library, while simultaneously overexpressing the SOD1(A4V)-GFP reporter, were subjected to FACS sorting for the isolation of clones exhibiting enhanced SOD1(A4V)-GFP fluorescence. This selection yielded an *E. coli* population with ~10-fold increased fluorescence after four rounds of sorting. Among twenty individual clones tested, four exhibited the highest levels of SOD1(A4V)-GFP fluorescence compared to cells expressing the same SOD1(A4V)-GFP fusion in the presence of random cyclic peptide sequences picked from the initial (unselected) cyclo-NuX₁X₂X₃-X₅ library and were selected for further analyses. Furthermore, the observed phenotypic effects were dependent on the ability of the Ssp DnaE intein (utilized for peptide cyclization as part of SICLOPPS) to perform protein splicing, as the double amino acid substitution H24L/F26A in the C-terminal half of the Ssp DnaE intein, which is known to abolish asparagine cyclization at the Ic/extein junction, and prevent extein splicing and peptide cyclization, was found to reduce SOD1(A4V)-GFP fluorescence back to wild-type levels. Finally, the observed increases in fluorescence were found to be SOD1-specific, as the isolated pSICLOPPS-NuX₁X₂X₃-X₅ vectors from these selected clones did not enhance the levels of cellular green fluorescence when the sequence of SOD1(A4V) in the SOD1(A4V)-GFP reporter was replaced with that of the human β -amyloid peptide (A β). On the contrary, the selected pSICLOPPS-NuX₁X₂X₃-X₅ vectors were efficient in enhancing the fluorescence of SOD1-GFP containing wild-type SOD1, as well as three additional SOD1 variants, SOD1(G37R), SOD1(G85R), and SOD1(G93A), all of which are associated with familial forms of ALS. Western blot analysis indicated that this enhanced SOD1(A4V)-GFP fluorescence phenotype occurs due to accumulation of enhanced amounts of soluble SOD1(A4V) in these clones.

The inventors further analyzed the selected peptides by DNA sequencing of the peptide-encoding regions of the four selected clones. This revealed three distinct putative SOD1(A4V) misfolding-rescuing and aggregation-inhibitory cyclic peptide sequences, all of which encoded cyclic pentapeptides with sequences TASFW, TWSVW, and TFSMW, thus indicating a dominant TXSXW bioactive motif. Interestingly, the Ser residue at position 3, encountered among all selected pentapeptides, was encoded by two different codons in the selected pSICLOPPS plasmids, thus suggesting that its dominance among the isolated clones was not coincidental.

As depicted in Examples 4 and 5, the inventors chose the peptide cyclo-TWSVW for further analysis. This cyclic pentapeptide is hereafter referred to as SOD1C5-4 and was produced in mg quantities by solid-phase synthesis.

Isolated SOD1(A4V) was utilized to assess the effect of the selected cyclic pentapeptide SOD1C5-4 on its aggregation process. CD spectroscopy indicated that SOD1C5-4 –but not the A β -targeting cyclic peptides A β C5-34 or A β C5-116– interacts with SOD1(A4V), and that the time-dependent conformational transition that is indicative of SOD1(A4V) aggregation is significantly delayed in the presence of SOD1C5-4. Moreover, analysis by dynamic light scattering (DLS) revealed that SOD1C5-4 addition results in the time-dependent formation of oligomeric/aggregated SOD1(A4V) species with markedly smaller sizes. Detection of large, amyloid-like SOD1(A4V) aggregates by ThT staining and a filter retardation assay indicated that the formation of such species was dramatically decreased in the presence of SOD1C5-4. Finally, staining of SOD1(A4V) with the conformation-sensitive dye SYPRO Orange under heat-induced denaturation conditions, suggested that the aggregation-inhibitory action of SOD1C5-4 may be occurring due to its ability to decrease the propensity of SOD1(A4V) to expose hydrophobic surfaces, a feature which has been proposed to be a molecular determinant of the pathogenesis of fALS-associated SOD1 variants. Taken together, these results demonstrate that SOD1C5-4 is an efficient and specific rescuer of SOD1(A4V) misfolding and aggregation.

The protective effects of SOD1C5-4 in mammalian cells were evaluated in human embryonic kidney 293 (HEK293) cells transiently expressing SOD1(A4V)-GFP. Cells treated with SOD1C5-4 exhibited higher fluorescence, fewer inclusions comprising aggregated SOD1(A4V)-GFP, and higher viability compared to untreated cells.

To identify all bioactive cyclic oligopeptide SOD1 modulators contained in the tested cyclo-NuX₁X₂X₃-X₅ library and to facilitate structure-activity analyses of the isolated sequences, To determine structure-activity relationships for the identified mutant SOD1-

targeting cyclic oligopeptides, the sequences of the peptide-encoding regions from ~5.3 million clones selected after the fourth round of FACS sorting were determined by deep sequencing. 367 distinct oligopeptide sequences appeared more than 50 times among the selected clones and were selected for subsequent analysis, which revealed the following. First, pentapeptides were the dominant peptide species within the sorted pool, with 197 of the distinct oligopeptide sequences selected corresponding to pentapeptides (54%), 148 to hexapeptides (40%) and 22 corresponding to tetrapeptides (6%). Second, the vast majority of the selected peptides exhibited the cyclo-TXSXW motif of SOD1C5-4 (~92% of all selected clones and ~97% of the selected pentapeptide-encoding clones. Third, among the selected cyclo-TXSXW pentapeptides, I, N, Q, M, E, H, and K residues were excluded at position 2, and were preferably S, A, W or F. At position 4, I, N, Q, C, D, E, K and P residues were excluded, and were preferably V, W, F, M, or H. Taken together, these results indicate that the most bioactive macrocyclic structures against SOD1(A4V) misfolding and aggregation in the library are cyclic pentapeptides of the cyclo-T(Φ_1 ,S)S(Φ_2 ,M,H)W motif, where Φ_1 is preferably one of the hydrophobic (Φ) amino acids A, W or F, while Φ_2 is preferably V, W or F.

E. coli BL21(DE3) cells producing the combined cyclo-NuX₁X₂X₃-X₅ library, while simultaneously overexpressing the A β ₄₂-GFP reporter, were subjected to FACS sorting for the isolation of clones exhibiting enhanced A β ₄₂-GFP fluorescence. Increase in the mean fluorescence was measured and ten random clones were picked from the sorted population. A β ₄₂-GFP fluorescence of the isolated peptide-expressing clones was found to be dramatically increased compared to cells expressing the same A β ₄₂-GFP fusion in the presence of random cyclic peptide sequences picked from the initial (unselected) cyclo-NuX₁X₂X₃-X₅ library. Furthermore, the observed phenotypic effects were dependent on the ability of the Ssp DnaE intein (utilized for peptide cyclization as part of SICLOPPS) to perform protein splicing, as the double amino acid substitution H24L/F26A in the C-terminal half of the Ssp DnaE intein, which is known to abolish asparagine cyclization at the Ic/extein junction, and prevent extein splicing and peptide cyclization, was found to reduce A β ₄₂-GFP fluorescence back to wild-type levels. Finally, the observed increases in fluorescence were found to be A β -specific, as the isolated pSICLOPPS-NuX₁X₂X₃-X₅ vectors from these selected clones did not enhance the levels of cellular green fluorescence when the sequence of A β in the A β ₄₂-GFP reporter was replaced with that of each one of two unrelated disease-associated MisPs, the DNA-binding (core) domain of the human p53 containing a Tyr220Cys substitution (p53C(Y220C)) and an Ala4Val substitution of human Cu/Zn superoxide dismutase 1 (SOD1(A4V)). On the contrary, the selected pSICLOPPS-NuX₁X₂X₃-X₅ vectors were efficient in enhancing the fluorescence

of A β -GFP containing two additional A β variants, A β ₄₀ and the E22G (arctic) variant of A β ₄₂, which is associated with familial forms of AD.

Analysis of the expressed A β ₄₂-GFP fusions by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and western blotting revealed that the bacterial clones
5 expressing the selected cyclic peptides produce markedly increased levels of soluble A β ₄₂-GFP compared to random cyclic peptide sequences. Furthermore, when the same cell lysates were analyzed by native PAGE and western blotting, it was observed that co-expression of the selected cyclic peptides reduced the accumulation of higher-order A β ₄₂-GFP aggregates, which could not enter the gel, and increased the abundance of species with higher electrophoretic
10 mobility.

The inventors further analyzed the selected peptides by DNA sequencing of the peptide-encoding regions of ten isolated clones. This revealed eight distinct putative A β aggregation-inhibitory cyclic peptide sequences: one corresponded to a hexapeptide (TPVWFD; present twice among the sequenced clones) and seven pentapeptides (TAFDR, TAWCR, TTWCR,
15 TTVDR, TTYAR (present twice), TTTAR, and SASPT). Interestingly, the Arg residue at position 5, frequently encountered among the selected pentapeptides, was encoded by three different codons in the selected pSICLOPPS plasmids, thus suggesting that its dominance among the isolated clones was not coincidental.

As depicted in Examples 7-10, the inventors chose two cyclic oligopeptide sequences
20 for further analysis. These were cyclo-TAFDR and cyclo-SASPT, hereafter referred to as A β C5-116 and A β C5-34 (*A β -targeting cyclic 5-peptide number 116 and 34*), and were produced by solid-phase synthesis in mg quantities. To further analyze the results the inventors of the present invention chose to focus on pentapeptides, as this was the type of peptide most frequently present among the ones selected from the genetic screen. The inventors decided to
25 further study the sequence A β C5-116 since the TXXXXR motif was particularly dominant among the selected pentapeptides, while A β C5-34 was chosen because it was the only selected pentapeptide whose sequence appeared to deviate from this motif.

Circular dichroism (CD) spectroscopy was first used to assess the effect of the selected pentapeptides on the aggregation process of A β ₄₀ and A β ₄₂. Addition of A β C5-116 was found
30 to strongly inhibit the aggregation of A β ₄₀, which remained at a random coil conformation for extended periods of time. The addition of A β C5-34 did not have the same effect and resulted in the appearance of a low-intensity negative peak. When the same solutions were subjected to a ThT dye-binding assay detecting amyloid fibrils, A β ₄₀ fibril formation was reduced in the presence of A β C5-116, while it remained unaffected by A β C5-34. In the case of A β ₄₂, both

selected cyclic pentapeptides affected its normal aggregation pathway strongly and stabilized β -sheet-like structures. ThT staining of the same samples revealed that the extent of amyloid fibril formation was greatly reduced in both cases. When the cyclic peptides were added at a higher ratio, similar CD patterns were observed, however the negative peaks were much more pronounced and fibril formation was completely prevented. The addition of the control cyclic pentapeptide SOD1C5-4 targeting another protein and of randomly selected cyclic control peptides did not have any effect on $A\beta_{40}$ and $A\beta_{42}$ aggregation. Finally, the inventors also performed Transmission electron microscopy (TEM) to verify the above findings. Taken together, these results demonstrate that the selected cyclic oligopeptides interfere with the normal aggregation process of $A\beta$.

The protective effects of $A\beta$ C5-34 and $A\beta$ C5-116 on $A\beta_{40}$ - and $A\beta_{42}$ -induced toxicity were evaluated in primary mouse hippocampal neurons and in glioblastoma cell lines. The addition of $A\beta$ C5-34 and $A\beta$ C5-116 was found to markedly inhibit the neurotoxicity of both $A\beta_{40}$ and $A\beta_{42}$ in a dose-responsive manner. The inventors also studied the effect of $A\beta$ C5-34 and $A\beta$ C5-116 on the morphology of $A\beta$ -exposed neuronal cells by phase-contrast microscopy. In the presence of pre-aggregated $A\beta$, the population of attached cells was greatly reduced compared to the control, with many detached rounded-up cells floating in the supernatant, while hallmarks of degenerating neurons, such as cell shrinkage, membrane blebbings, fragmented neurites and ill-developed axons were obvious in the preparations. This phenotype was reversed with the addition of the selected cyclic peptides.

To further evaluate the protective effects of the selected cyclic peptides against $A\beta$ aggregation and toxicity *in vivo*, the inventors employed three established models of AD in the nematode worm *Caenorhabditis elegans*. The conservation of genetic and metabolic pathways between *C. elegans* and mammals, in combination with its completely characterized nervous and muscular system, its easy visualization and simple manipulation, has nominated *C. elegans* as an excellent model for neurodegenerative diseases including AD, while chemical screening against $A\beta$ -induced toxicity in *C. elegans* is increasingly used in AD drug discovery. A paralysis assay was performed in the *C. elegans* strain CL4176, where human $A\beta_{42}$ is expressed in the animals' body wall muscle cells under the control of a heat-inducible promoter and $A\beta$ aggregate formation is accompanied by the emergence of a paralysis phenotype. When chemically synthesized $A\beta$ C5-34 (10 μ M) and $A\beta$ C5-116 (5 μ M) were supplied to CL4176 worms, the emergence of the characteristic paralysis phenotype upon temperature up-shift was significantly decelerated compared to the untreated animals. The strain CL2331, which expresses a $A\beta_{(3-42)}$ -GFP fusion again in its body wall muscle cells upon temperature up-shift

was also used, and treatment with either one of the selected peptides resulted in a significant reduction of A β deposits, which was further shown with biochemical analysis of the accumulation levels of both total and oligomeric A β levels in CL4176 animals.

To identify the functionally important residues within the isolated peptides, the inventors performed position 1 substitutions with the other two nucleophilic amino acids present in the initial libraries, as well as alanine scanning mutagenesis at positions 3-5 of the A β C5-34 and A β C5-116 pentapeptides. As judged by the ability of the generated variants to enhance the fluorescence of *E. coli* cells overexpressing A β ₄₂-GFP, A β C5-116 was found to be much more tolerant to substitutions compared to A β C5-34. All tested sequence alterations within A β C5-34, apart from the SIT substitution, were found to be deleterious for its A β aggregation-inhibitory. On the contrary, only the initial Thr and the ultimate Arg were found to be absolutely necessary for the bioactivity of A β C5-116, whereas residues at positions 3 and 4 could be substituted by Ala without significant loss of activity.

To identify all bioactive cyclic oligopeptide A β modulators contained in the tested cyclo-NuX₁X₂X₃-X₅ library and to facilitate structure-activity analyses of the isolated sequences, the peptide sequences isolated from the genetic screen were analyzed by next-generation sequencing. This analysis revealed the following. First, pentapeptides were the dominant peptide species within the sorted pool. Second, the most prevalent motif among the selected pentapeptide sequences were cyclo-TXXXXR pentapeptides (~47% of the selected pentapeptide-encoding pSICLOPPS plasmids; ~42% of the unique selected pentapeptide sequences), in accordance with previous observations. On the contrary, only three pentapeptide sequence was found to have high similarity with A β C5-34. Third, for the selected peptides corresponding to the cyclo-TXXXXR motif, residues at positions 3 and 4 were highly variable and included the majority of natural amino acids, with position 3 exhibiting the highest diversity. At position 2, Thr, Ala, and Val were preferred, while aromatic residues (Phe, Trp, Tyr) were completely excluded from the selected cyclo-TXXXXR peptide pool, in full agreement with the aforementioned site-directed mutagenesis studies. At the highly variable position 3, the complete absence of the negatively charged amino acids Glu and Asp among the selected sequences was notable. In general, both negatively (Glu and Asp) and positively charged residues (Lys, His, and Arg) were found to be disfavored among the selected cyclo-TXXXXR sequences at positions 2 and 3. At position 4, Ala, Asp, and Trp were found to be the preferred residues. It is noteworthy, that Lys and Gln residues were practically absent from all positions, while the β sheet-breaking amino acid Pro that is typically included in designed peptide-based inhibitors of amyloid aggregation appeared with strikingly low frequencies.

Thus, preferred A β modulators are cyclic oligopeptide sequences exhibiting the cyclo-TXXXXR motif, where X is any natural amino acid. More preferred are cyclic oligopeptide sequences exhibiting the cyclo-T Φ Z Π R motif, where Φ = any amino acid excluding F, W, Y, Q, D, E, K and P; Z is any amino acid excluding E; and Π is any natural amino acid excluding Q, M, K.
5 Even more preferred are cyclic oligopeptide sequences exhibiting the cyclo-T(T,A,V) Ψ (A,D,W)R motif, where Ψ is a non-negatively charged amino acid.

The high residue variability observed at position 3 of the selected TXXXXR peptides prompted the inventors to investigate whether A β C5-116 could be further minimized. Indeed, production of truncated variants of A β C5-116, from which Ala2 or Asp4 had been deleted,
10 resulted in a respective two- and three-fold enhancement in the fluorescence of bacterially expressed A β ₄₂-GFP. In accordance with this, a total of ten distinct cyclic tetrapeptide sequences belonging to the TXXR motif were identified among the selected peptide pool. Taken together, these results indicate that the minimal bioactive entity against A β aggregation among this peptide family is a TXXR cyclic tetrapeptide.

15 The present invention describes a versatile and generally applicable method for identifying macrocyclic chemical rescuers of the misfolding of misfolding-prone proteins associated with a protein misfolding disease (MisP), wherein said MisP is selected from SOD1, β -amyloid peptide, tau, α -synuclein, polyglutaminated huntingtin, polyglutaminated ataxin-1, polyglutaminated ataxin-2, polyglutaminated ataxin-3, prion protein, islet amyloid polypeptide
20 (amylin), β 2-microglobulin, fragments of immunoglobulin light chain, fragments of immunoglobulin heavy chain, serum amyloid A, ABri peptide, ADan peptide, transthyretin, apolipoprotein A1, gelsolin, transthyretin, lysozyme, phenylalanine hydroxylase, apolipoprotein A-I, calcitonin, prolactin, TDP-43, FUS/TLS; insulin, hemoglobin, α 1-antitrypsin, p53; or variants thereof.

25 Preferably the invention presented herein can be used as a method for the identification of chemical agents for the treatment, prevention or diagnosis of diseases related to protein misfolding and aggregation, including amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, Creutzfeldt-Jakob disease, cancer, phenylketonuria, type II diabetes, senile systemic amyloidosis, familial amyloidotic polyneuropathy, familial
30 amyloid cardiomyopathy, leptomeningeal amyloidosis, systemic amyloidosis, familial British dementia, familial Danish dementia, light chain amyloidosis, heavy chain amyloidosis, serum amyloid A amyloidosis, lysozyme amyloidosis, dialysis-related amyloidosis, ApoAI amyloidosis, Finnish type familial amyloidosis, hereditary cerebral hemorrhage with amyloidosis (Icelandic type), medullary carcinoma of the thyroid, pituitary prolactinoma,

injection-localized amyloidosis, frontotemporal dementia, spinocerebellar ataxia 1, spinocerebellar ataxia 2, spinocerebellar ataxia 3, α 1-antitrypsin deficiency, sickle-cell anemia, and transmissible spongiform encephalopathy. Most preferably the invention presented herein can be used as a method of treatment, prevention or diagnosis of amyotrophic lateral sclerosis or Alzheimer's disease.

Table 1. Sequences and frequency of appearance of the selected cyclo-TXXXR pentapeptides as determined by high-throughput sequencing of the isolated pSICLOPPS-NuX₁X₂X₃-X₅ vectors after the second round of bacterial sorting for enhanced A β ₄₂-GFP fluorescence.

seq ID NO	Peptide name	Aminoacid sequence					Number of reads	Reads/ Total TXXXR reads (%)	Reads/ Total pentapeptide reads (%)	Reads/ Total peptide reads (%)	Peptide-encoding nucleotide sequence
47	A β C5-2	T	T	Y	A	R	304,753	16.023	7.506	6.727	ACCACGTACGCCAGG
48	A β C5-3	T	T	V	D	R	214,461	11.276	5.282	4.734	ACCACCGTGGACCGG
49	A β C5-5	T	T	T	W	R	175,510	9.228	4.323	3.874	ACCACGACCTGGAGG
50	A β C5-7	T	T	L	H	R	134,018	7.046	3.301	2.958	ACCACGCTGCACCGG
51	A β C5-8	T	T	F	A	R	96,700	5.084	2.382	2.134	ACCACCTTCGCCCGG
52	A β C5-9	T	V	L	D	R	89,669	4.715	2.209	1.979	ACCGTCTTGGACCGG
53	A β C5-12	T	T	W	A	R	65,929	3.466	1.624	1.455	ACCACGTGGGCCAGG
54	A β C5-13	T	A	L	D	R	62,792	3.301	1.547	1.386	ACCGCGCTGGACCGG
55	A β C5-15	T	A	N	V	R	47,855	2.516	1.179	1.056	ACCGCGAAGCTGAGG
56	A β C5-17	T	T	T	A	R	40,135	2.110	0.989	0.886	ACCACCACGGCCCGG
57	A β C5-18	T	T	I	A	R	37,150	1.953	0.915	0.820	ACCACATCGCCCGG
58	A β C5-19	T	V	W	D	R	37,091	1.950	0.914	0.819	ACCGTGTGGGACCGG
59	A β C5-20	T	T	I	S	R	37,044	1.948	0.912	0.818	ACCACATCAGCCCGG
60	A β C5-21	T	T	W	C	R	36,295	1.908	0.894	0.801	ACCACCTGGTGCCCGG
61	A β C5-22	T	V	L	W	R	35,820	1.883	0.882	0.791	ACCGTCCTGTGGAGG
62	A β C5-25	T	T	L	A	R	28,989	1.524	0.714	0.640	ACCACCTTGGCGAGG
63	A β C5-26	T	A	W	C	R	28,391	1.493	0.699	0.627	ACCGCGTGGTGCCGC
64	A β C5-27	T	T	S	A	R	28,188	1.482	0.694	0.622	ACCACGAGCGCCCGC
65	A β C5-29	T	T	L	E	R	27,514	1.447	0.678	0.607	ACCACCCTCGAGAGG
66	A β C5-30	T	S	T	A	R	27,456	1.444	0.676	0.606	ACCTCGACGGCGCGG
67	A β C5-35	T	V	R	D	R	25,428	1.337	0.626	0.561	ACCGTCCGGGACCGG
68	A β C5-41	T	G	W	A	R	21,784	1.145	0.537	0.481	ACCGGCTGGGCGAGG
69	A β C5-44	T	A	W	A	R	20,807	1.094	0.512	0.459	ACCGCTGGGCGAGG
70	A β C5-45	T	T	W	V	R	20,798	1.094	0.512	0.459	ACCACCTGGGTGCGG
71	A β C5-46	T	L	L	W	R	19,957	1.049	0.492	0.440	ACCCTATTGTGGCGG
72	A β C5-47	T	T	I	D	R	19,735	1.038	0.486	0.436	ACCACGATCGACAGG
73	A β C5-50	T	A	L	A	R	19,433	1.022	0.479	0.429	ACCGCGCTCGCGCGC
74	A β C5-51	T	S	V	D	R	19,249	1.012	0.474	0.425	ACCAGCGTGGACAGG
75	A β C5-53	T	T	V	W	R	18,669	0.982	0.460	0.412	ACCACCGTGTGGCGC
76	A β C5-66	T	T	H	W	R	14,304	0.752	0.352	0.316	ACCACGCACTGGCGG
77	A β C5-67	T	A	R	D	R	14,213	0.747	0.350	0.314	ACCGCGAGGGACCGG
78	A β C5-73	T	T	R	D	R	12,894	0.678	0.318	0.285	ACCACGCGGGACCGG
79	A β C5-80	T	S	V	H	R	10,181	0.535	0.251	0.225	ACCAGCGTGCACCGG
80	A β C5-82	T	A	V	W	R	9,781	0.514	0.241	0.216	ACCGCCGTCTGGCGG

81	AβC5-83	T	T	G	C	R	9,362	0.492	0.231	0.207	ACCACGGGGTGCCGG
82	AβC5-89	T	A	T	D	R	7,984	0.420	0.197	0.176	ACCGCCACCGACAGG
83	AβC5-94	T	V	L	F	R	7,442	0.391	0.183	0.164	ACCGTCTTGTTCCGC
84	AβC5-102	T	T	Y	N	R	6,067	0.319	0.149	0.134	ACCACCTACAACCGC
85	AβC5-105	T	V	R	W	R	5,450	0.287	0.134	0.120	ACCGTGCCTGGCGC
86	AβC5-116	T	A	F	D	R	4,243	0.223	0.105	0.094	ACCGCGTTCGACCGG
87	AβC5-117	T	T	R	C	R	4,237	0.223	0.104	0.094	ACCACGCGGTGCAGG
88	AβC5-118	T	T	F	W	R	4,216	0.222	0.104	0.093	ACCACCTTCTGGCGG
89	AβC5-121	T	I	K	D	R	3,970	0.209	0.098	0.088	ACCATCAAGGACCGG
90	AβC5-123	T	T	V	H	R	3,371	0.177	0.083	0.074	ACCACCGTCCACCGG
91	AβC5-126	T	T	L	L	R	3,016	0.159	0.074	0.067	ACCACGCTCCTCAGG
92	AβC5-129	T	T	L	F	R	2,630	0.138	0.065	0.058	ACCACGCTCTTCCGG
93	AβC5-130	T	A	Y	H	R	2,594	0.136	0.064	0.057	ACCGCGTACCACCGG
94	AβC5-136	T	A	L	H	R	2,026	0.107	0.050	0.045	ACCGCGTTCGACCGG
95	AβC5-139	T	T	S	P	R	1,904	0.100	0.047	0.042	ACCACCTCGCCCCGG
96	AβC5-146	T	T	W	S	R	1,612	0.085	0.040	0.036	ACCACCTGGTCGCGG
97	AβC5-147	T	A	M	H	R	1,611	0.085	0.040	0.036	ACCGCCATGCACAGG
98	AβC5-155	T	S	L	D	R	1,251	0.066	0.031	0.028	ACCTCGCTCGACAGG
99	AβC5-158	T	T	G	A	R	1,172	0.062	0.029	0.026	ACCACGGGGGCGCGC
100	AβC5-162	T	S	V	W	R	1,094	0.058	0.027	0.024	ACCTCGGTGTGGAGG
101	AβC5-173	T	T	H	A	R	953	0.050	0.023	0.021	ACCACGCACGCCAGG
102	AβC5-176	T	A	G	W	R	945	0.050	0.023	0.021	ACCGCGGGCTGGAGG
103	AβC5-177	T	A	T	A	R	925	0.049	0.023	0.020	ACCGCCACCGCGAGG
104	AβC5-184	T	V	L	A	R	818	0.043	0.020	0.018	ACCGTGCTCGCGCGG
105	AβC5-185	T	T	F	N	R	800	0.042	0.020	0.018	ACCACGTTCAACAGG
106	AβC5-188	T	G	M	R	R	768	0.040	0.019	0.017	ACCGGGATGAGGCGG
107	AβC5-189	T	T	V	A	R	757	0.040	0.019	0.017	ACCACCGTCGCCAGG
108	AβC5-190	T	L	C	L	R	739	0.039	0.018	0.016	TGCTTGCGCACGCTG
109	AβC5-192	T	G	L	A	R	720	0.038	0.018	0.016	ACCGGGCTGGCGCGG
110	AβC5-198	T	S	W	C	R	679	0.036	0.017	0.015	ACCAGCTGGTGCAGG
111	AβC5-209	T	T	R	A	R	580	0.030	0.014	0.013	ACCACCAGGGCGCGG
112	AβC5-215	T	T	P	W	R	524	0.028	0.013	0.012	ACCACGCCCTGGAGG
113	AβC5-218	T	V	L	H	R	497	0.026	0.012	0.011	ACCGTCTTGACAGG
114	AβC5-223	T	G	L	D	R	464	0.024	0.011	0.010	ACCGGCCTCGACAGG
115	AβC5-230	T	T	S	D	R	442	0.023	0.011	0.010	ACCACGTCGGACCGG
116	AβC5-239	T	T	M	H	R	384	0.020	0.009	0.008	ACCACGATGCACCGC
117	AβC5-242	T	T	S	T	R	376	0.020	0.009	0.008	ACCACCTCGACCCGG
118	AβC5-244	T	T	R	V	R	366	0.019	0.009	0.008	ACCACGCGCGTGAGG
119	AβC5-245	T	T	R	F	R	364	0.019	0.009	0.008	ACCACCCGGTTCCGG
120	AβC5-248	T	T	T	H	R	339	0.018	0.008	0.007	ACCACGACGCACCGG
121	AβC5-250	T	H	A	W	R	334	0.018	0.008	0.007	ACCCACGCCTGGAGG
122	AβC5-252	T	V	I	W	R	331	0.017	0.008	0.007	ACCGTGATCTGGCGC
123	AβC5-253	T	T	W	F	R	327	0.017	0.008	0.007	ACCACGTGGTTCCGG
124	AβC5-255	T	T	S	R	R	325	0.017	0.008	0.007	ACCACCTCGAGACGG
125	AβC5-258	T	T	S	C	R	301	0.016	0.007	0.007	ACCACGTCGTGCCGG
126	AβC5-260	T	T	W	T	R	295	0.016	0.007	0.007	ACCACCTGGACCCGG
127	AβC5-262	T	T	S	S	R	286	0.015	0.007	0.006	ACCACCTCGAGCCGG
128	AβC5-263	T	H	L	A	R	284	0.015	0.007	0.006	ACCCACCTCGCCCCGG
129	AβC5-264	T	S	G	A	R	282	0.015	0.007	0.006	ACCAGCGGGGCCCGG
130	AβC5-266	T	T	L	R	R	274	0.014	0.007	0.006	ACCACGCTGCGCCGG
131	AβC5-270	T	A	T	W	R	266	0.014	0.007	0.006	ACCGCGACCTGGAGG
132	AβC5-272	T	C	M	W	R	254	0.013	0.006	0.006	ACCTGCATGTGGCGC

133	AβC5-275	T	A	H	V	R	249	0.013	0.006	0.005	ACCGCGCACGTGCGC
134	AβC5-276	T	S	W	A	R	249	0.013	0.006	0.005	ACCTCGTGGGCGCGG
135	AβC5-278	T	T	W	L	R	241	0.013	0.006	0.005	ACCACGTGGCTCAGG
136	AβC5-291	T	T	L	D	R	213	0.011	0.005	0.005	ACCACCCTGGACCGG
137	AβC5-294	T	T	P	H	R	207	0.011	0.005	0.005	ACCACGCCTCACCGG
138	AβC5-298	T	T	R	G	R	201	0.011	0.005	0.004	ACCACCCGTGGCCGG
139	AβC5-299	T	T	V	G	R	200	0.011	0.005	0.004	ACCACCGTGGGCCGG
140	AβC5-301	T	T	T	R	R	191	0.010	0.005	0.004	ACCACGACGCGCCGC
141	AβC5-304	T	S	I	N	R	182	0.010	0.004	0.004	ACCTCGATCAACAGG
142	AβC5-305	T	T	A	D	R	181	0.010	0.004	0.004	ACCACCGCGGACCGG
143	AβC5-315	T	T	S	E	R	158	0.008	0.004	0.003	ACCACCTCCGAGAGG
144	AβC5-316	T	T	C	A	R	157	0.008	0.004	0.003	ACCACGTGCGCCAGG
145	AβC5-317	T	T	A	W	R	156	0.008	0.004	0.003	ACCACGGCCTGGAGG
146	AβC5-320	T	T	V	E	R	150	0.008	0.004	0.003	ACCACCGTCGAGCGG
147	AβC5-321	T	T	T	F	R	148	0.008	0.004	0.003	ACCACGACGTTTCAGG
148	AβC5-323	T	A	V	D	R	147	0.008	0.004	0.003	ACCGCCGTGGACCGG
149	AβC5-325	T	V	W	I	R	144	0.008	0.004	0.003	ACCGTGTGGATCAGG
150	AβC5-329	T	T	V	R	R	141	0.007	0.003	0.003	ACCACCGTACGCAGG
151	AβC5-333	T	H	V	R	R	137	0.007	0.003	0.003	ACCCACGTACGCAGG
152	AβC5-343	T	N	L	D	R	125	0.007	0.003	0.003	ACCAACCTGGACCGG
153	AβC5-344	T	T	P	G	R	125	0.007	0.003	0.003	ACCACGCCTGGACCGG
154	AβC5-348	T	T	L	T	R	119	0.006	0.003	0.003	ACCACGCTCACCCGG
155	AβC5-355	T	A	T	V	R	115	0.006	0.003	0.003	ACCGCGACGGTGCGC
156	AβC5-359	T	A	M	W	R	110	0.006	0.003	0.002	ACCGCCATGTGGCGG
157	AβC5-361	T	T	K	W	R	108	0.006	0.003	0.002	ACCACGAAGTGGAGG
158	AβC5-362	T	T	W	D	R	107	0.006	0.003	0.002	ACCACCTGGGACCGG
159	AβC5-364	T	T	M	A	R	106	0.006	0.003	0.002	ACCACCATGGCCCGG
160	AβC5-365	T	T	G	G	R	106	0.006	0.003	0.002	ACCACCGGTGGCCGG
161	AβC5-366	T	T	M	V	R	105	0.006	0.003	0.002	ACCACGATGGTGCGG
162	AβC5-375	T	N	L	A	R	97	0.005	0.002	0.002	ACCAACCTCGCCCGG
163	AβC5-376	T	I	R	D	R	96	0.005	0.002	0.002	ACCATCAGGGACCGG
164	AβC5-378	T	T	T	G	R	96	0.005	0.002	0.002	ACCACGACTGGTAGG
165	AβC5-379	T	R	L	G	R	95	0.005	0.002	0.002	ACCCGTCTTGGCAGG
166	AβC5-381	T	T	H	T	R	93	0.005	0.002	0.002	ACCACGCACACCAGG
167	AβC5-382	T	T	I	T	R	92	0.005	0.002	0.002	ACCACCATCACCCGG
168	AβC5-384	T	T	Y	T	R	90	0.005	0.002	0.002	ACCACGTACACCAGG
169	AβC5-385	T	T	L	Y	R	90	0.005	0.002	0.002	ACCACGCTGTACCGG
170	AβC5-389	T	H	L	D	R	89	0.005	0.002	0.002	ACCCACCTGGACCGG
171	AβC5-391	T	L	L	I	R	88	0.005	0.002	0.002	ACCTTGTGTATCAGG
172	AβC5-392	T	T	C	D	R	87	0.005	0.002	0.002	ACCACGTGCGACCGG
173	AβC5-393	T	T	G	R	R	87	0.005	0.002	0.002	ACCACGGGTGCGCCGG
174	AβC5-394	T	T	V	S	R	86	0.005	0.002	0.002	ACCACCGTGAGCCGG
175	AβC5-395	T	T	Q	H	R	85	0.004	0.002	0.002	ACCACGCAGCACCGG
176	AβC5-396	T	T	T	P	R	84	0.004	0.002	0.002	ACCACTACGCCCGG
177	AβC5-399	T	A	F	A	R	82	0.004	0.002	0.002	ACCGCCTTCGCCCGG
178	AβC5-405	T	T	S	H	R	78	0.004	0.002	0.002	ACCACGTACACCCGG
179	AβC5-410	T	V	L	G	R	76	0.004	0.002	0.002	ACCGTCTTGGGCCGG
180	AβC5-411	T	T	Q	R	R	75	0.004	0.002	0.002	ACCACGCAGCGCAGG
181	AβC5-413	T	S	H	A	R	74	0.004	0.002	0.002	ACCAGTCACGCCAGG
182	AβC5-415	T	T	T	C	R	74	0.004	0.002	0.002	ACCACGACGTGCCGG
183	AβC5-422	T	A	W	R	R	72	0.004	0.002	0.002	ACCGCGTGGCGCCGC
184	AβC5-428	T	T	C	G	R	69	0.004	0.002	0.002	ACCACGTGTGGCCGG

185	AβC5-434	T	T	S	G	R	65	0.003	0.002	0.001	ACCACCTCTGGCCGG
186	AβC5-438	T	T	T	S	R	62	0.003	0.002	0.001	ACCACGACGTCGAGG
187	AβC5-440	T	A	T	G	R	61	0.003	0.002	0.001	ACCGCGACTGGACGG
188	AβC5-441	T	A	W	D	R	61	0.003	0.002	0.001	ACCGCGTGGGACCGG
189	AβC5-443	T	T	H	H	R	60	0.003	0.001	0.001	ACCACGCATCACC GG
190	AβC5-448	T	A	Y	A	R	58	0.003	0.001	0.001	ACCGCGTACGCCAGG
191	AβC5-449	T	A	N	A	R	58	0.003	0.001	0.001	ACCGCGAACGCGAGG
192	AβC5-450	T	R	D	V	R	58	0.003	0.001	0.001	ACCCGCGACGTGAGG
193	AβC5-452	T	H	V	D	R	58	0.003	0.001	0.001	ACCCACGTCGACAGG
194	AβC5-453	T	L	F	W	R	57	0.003	0.001	0.001	ACCCTATTCTGGCGG
195	AβC5-459	T	T	A	A	R	55	0.003	0.001	0.001	ACCACCGCGGCCCGG
196	AβC5-463	T	V	V	D	R	54	0.003	0.001	0.001	ACCGTCGTGGACCGG
197	AβC5-464	T	T	P	A	R	54	0.003	0.001	0.001	ACCACTCCGGCCCGG
198	AβC5-469	T	T	I	G	R	53	0.003	0.001	0.001	ACCACGATCGGCAGG
199	AβC5-472	T	M	Y	A	R	51	0.003	0.001	0.001	ACCATGTACGCCAGG
200	AβC5-473	T	H	V	A	R	51	0.003	0.001	0.001	ACCCACGTGGCCAGG
201	AβC5-474	T	T	W	P	R	51	0.003	0.001	0.001	ACCACCTGGCCGCGG
202	AβC5-475	T	T	G	D	R	51	0.003	0.001	0.001	ACCACCGGTGACCGG
203	AβC5-479	T	T	T	V	R	50	0.003	0.001	0.001	ACCACGACCGTGCGG
204	AβC5-481	T	V	F	G	R	50	0.003	0.001	0.001	ACCGTCTTTGGCAGG
205	AβC5-483	T	R	V	G	R	50	0.003	0.001	0.001	ACCCGTGTGGGCCGG
Sum							1,901,945	100.000	46.847	41.980	

Table 2. Sequences and frequency of appearance of the selected cyclic pentapeptides resembling AβC5-34 as determined by high-throughput sequencing of the isolated pSICLOPPS-NuX₁X₂X₃-X₅ vectors after the second round of bacterial sorting for enhanced Aβ₄₂-GFP fluorescence.

SEQ ID NO	Peptide name	Amino acid sequence					Number of reads	Reads/ Total SASPT-like reads (%)	Reads/ Total pentapeptide reads (%)	Reads/ Total peptide reads (%)
206	AβC5-34	S	A	S	P	T	25673	97.349	0.632	0.567
207	AβC5-216	S	I	C	P	T	516	1.957	0.013	0.011
208	AβC5-380	S	I	T	P	T	94	0.356	0.002	0.002
209	AβC5-387	S	H	S	P	T	89	0.337	0.002	0.002
Sum							26,372	100	0.645	0.578

Table 3. Sequences and frequency of appearance of the selected cyclo-TXXR tetrapeptides as determined by high-throughput sequencing of the isolated pSICLOPPS-Nu X₁X₂X₃-X₅ vectors for enhanced A β ₄₂-GFP fluorescence.

SEQ ID NO.	PEPTIDE NAME	CYCLIC PEPTIDE SEQUENCE	Peptide-encoding nucleotide sequence	Normalized read number (%)
210	A β C4-9	TTCR	ACCACGTGCCGG	1.247884
211	A β C4-11	TTRR	ACCACTCGCCGG	1.199516
212	A β C4-31	TTSR	ACCACGTGCGGG	0.324063
213	A β C4-34	TRGR	ACACGTGGACGG	0.304716
214	A β C4-35	TTGR	ACCACTGGCCGG	0.295042
215	A β C4-41	TRRR	ACACGTGCGAGG	0.246675
216	A β C4B-9	TDQR	ACCGACCAGCGG	2.090359
217	A β C4B-41	TLIR	ACCCTGATCCGC	0.774951
218	A β C4B-80	TLWR	ACCCTGTGGCGG	0.256828
219	A β C4B-86	TLGR	ACCTTGGGCCGG	0.16973
220	A β C5(Δ A2)	TFDR	ACCTTCGACCGG	-
221	A β C5(Δ D4)	TAFR	ACCGCGTTCCGG	-

5 **Table 4.** Sequences and frequency of appearance of the selected cyclic hexapeptides as determined by high-throughput sequencing of the isolated pSICLOPPS-NuX₁X₂X₃-X₅ vectors after the second round of bacterial sorting for enhanced A β ₄₂-GFP fluorescence.

SEQ ID NO	Peptide name	Aminoacid sequence						Number of reads	Reads/ Total hexapeptide reads (%)	Reads/ Total peptide reads (%)	Peptide-encoding nucleotide sequence
222	A β C6-1	T	P	V	W	F	D	131,935	29.151	2.912	ACCCCGGTCTGGTTCGAC
223	A β C6-2	T	P	A	W	F	D	111,132	24.555	2.453	ACCCCGGCCTGGTTCGAC
224	A β C6-4	T	L	E	F	F	D	27,057	5.978	0.597	ACCTTGGAGTTCCTTCGAC
225	A β C6-6	T	V	T	W	F	D	17,100	3.778	0.377	ACCGTCACGTGGTTCGAC
226	A β C6-8	T	L	L	I	R	W	13,135	2.902	0.290	ACCTTGTTGATCAGGTGG
227	A β C6-10	T	L	K	W	L	N	11,016	2.434	0.243	ACCCTCAAGTGGCTGAAC
228	A β C6-21	T	K	E	Y	F	D	1,231	0.272	0.027	ACCAAGGAGTACTTCGAC
229	A β C6-26	T	L	H	W	F	E	647	0.143	0.014	ACCCTCCACTGGTTCGAG
230	A β C6-27	T	C	S	W	F	D	623	0.138	0.014	ACCTGCTCGTGGTTCGAC
231	A β C6-28	T	L	E	Y	F	M	556	0.123	0.012	ACCCTCGAGTACTTCATG
232	A β C6-32	T	L	C	W	L	N	455	0.101	0.010	ACCCTGTGCTGGCTCAAC
233	A β C6-36	T	P	I	V	F	D	384	0.085	0.008	ACCCCGATCGTGTTCGAC
234	A β C6-37	T	L	W	V	F	D	355	0.078	0.008	ACCCTGTGGGTCTTCGAC
235	A β C6-40	T	P	L	W	F	N	316	0.070	0.007	ACCCCTTGTGGTTCAAC
236	A β C6-41	T	S	V	E	Y	E	307	0.068	0.007	ACCTCGGTCGAGTACGAG
237	A β C6-42	T	L	G	W	L	D	307	0.068	0.007	ACCCTGGGCTGGTTGGAC

238	AβC6-44	T	P	P	W	F	D	289	0.064	0.006	ACCCCGCCCTGGTTCGAC
239	AβC6-46	T	P	C	W	F	D	252	0.056	0.006	ACCCCGTGCTGGTTCGAC
240	AβC6-47	T	L	S	W	Y	D	239	0.053	0.005	ACCTTGTCTTGGTACGAC
241	AβC6-48	T	P	V	L	V	D	236	0.052	0.005	ACCCCGGTCCTGGTCGAC
242	AβC6-49	T	L	E	Y	L	W	233	0.051	0.005	ACCCTCGAGTACTTGTGG
243	AβC6-50	T	I	F	W	F	D	227	0.050	0.005	ACCATCTTCTGGTTCGAC
244	AβC6-53	T	P	A	L	V	D	208	0.046	0.005	ACCCCGGCCCTGGTCGAC
245	AβC6-55	T	P	G	W	F	D	180	0.040	0.004	ACCCCGGCTGGTTCGAC
246	AβC6-57	T	L	S	V	F	D	176	0.039	0.004	ACCTTGTCCGTCTTCGAC
247	AβC6-58	T	P	G	L	V	D	142	0.031	0.003	ACCCCGGTCTGGTCGAC
248	AβC6-59	T	L	S	W	F	N	141	0.031	0.003	ACCCTCTCCTGGTTCAAC
249	AβC6-63	T	L	D	F	F	D	114	0.025	0.003	ACCTTGGACTTCTTCGAC
250	AβC6-65	T	P	S	W	F	D	105	0.023	0.002	ACCCCGTCCTGGTTCGAC
251	AβC6-68	T	P	A	L	F	D	101	0.022	0.002	ACCCCGGCCCTGTTTCGAC
252	AβC6-69	T	P	A	W	S	D	86	0.019	0.002	ACCCCGGCCTGGTCCGAC
253	AβC6-78	T	P	A	R	F	D	55	0.012	0.001	ACCCCGGCCCGGTTTCGAC
254	AβC6-79	T	P	A	W	L	D	55	0.012	0.001	ACCCCGGCCTGGCTCGAC
255	AβC6-80	T	P	V	W	L	D	55	0.012	0.001	ACCCCGGTCTGGCTCGAC
Sum								319,553	70.606	7.053	

III. Peptide modifications

Peptides and polypeptides of the invention include those corresponding to linearized versions of the described cyclic oligopeptide SOD1 modulators, i.e., sequences where a break in the amino acid backbone chain of a described cyclic oligopeptide modulator has been introduced and which thereafter contains a free N-terminal NH₂ amino group and a free C-terminal –COOH carboxyl group. For example, for the cyclic pentapeptide SOD1 modulator SOD1C5-4 with amino acid sequence cyclo-TWSVW (SEQ ID NO: 4), a preferred peptide SOD1 modulator of the present invention is also a linearized version of SOD1C5-4, namely the oligopeptide NH₂-TWSVW-COOH. In addition, since the herein described oligopeptide SOD1 modulators are cyclic in nature, they do not possess a “starting point” (e.g. N terminus) or “end point” (e.g. C terminus). Thus, all circular permutants, e.g., linear variants resulting from cleavage of an existing peptide bond to introduce new termini elsewhere in the peptide sequence, of the described cyclic oligopeptide SOD1 modulators are also preferred cyclic oligopeptide SOD1 modulators of the present invention. For example, for the cyclic pentapeptide SOD1 modulator SOD1C5-4 with amino acid sequence cyclo-TWSVW (SEQ ID NO: 4), preferred peptide SOD1 modulators of the present invention are also all equivalent circular permutants of SOD1C5-4, namely the oligopeptides cyclo-WSVWT, cyclo-SVWTW, cyclo-VWTWS, and cyclo-WTWSV. Similarly, preferred peptide SOD1 modulators of the

present invention are the linearized versions of all described cyclic oligopeptide SOD1 modulators and of all of their equivalent circular permutants. For example, for SOD1C5-4, apart from the modification mentioned above, preferred peptide SOD1 modulator of the present invention are linearized versions of all circular permutants equivalent SOD1C5-4, namely the
5 oligopeptides NH₂-WSVWT-COOH, NH₂-SVWTW-COOH, NH₂-VWTWS -COOH, and NH₂-WTWSV-COOH.

Similarly, for the cyclic pentapeptide A β modulator A β C5-116 with amino acid sequence cyclo-TAFDR (SEQ ID NO: 86), a preferred peptide A β modulator of the present invention is also a linearized version of A β C5-116, namely the oligopeptide NH₂-TAFDR-
10 COOH. In addition, since the herein described oligopeptide A β modulators are cyclic in nature, they do not possess a “starting point” (e.g. N terminus) or “end point” (e.g. C terminus). Thus, all circular permutants, e.g., linear variants resulting from cleavage of an existing peptide bond to introduce new termini elsewhere in the peptide sequence, of the described cyclic oligopeptide A β modulators are also preferred cyclic oligopeptide A β modulators of the present
15 invention. For example, for the cyclic pentapeptide A β modulator A β C5-116 with amino acid sequence cyclo-TAFDR (SEQ ID NO: 86), preferred peptide A β modulators of the present invention are also all equivalent circular permutants of A β C5-116, namely the oligopeptides cyclo-AFDRT, cyclo-FDRTA and NH₂-FDRTA-COOH, cyclo-DRTAF, and cyclo-RTAFD. Similarly, preferred peptide A β modulators of the present invention are the linearized versions
20 of all described cyclic oligopeptide A β modulators and of all of their equivalent circular permutants. For example, for A β C5-116, apart from the modification mentioned above, preferred peptide A β modulator of the present invention are linearized versions of all circular permutants equivalent A β C5-116, namely the oligopeptides NH₂-AFDRT-COOH, NH₂-FDRTA-COOH, NH₂-DRTAF-COOH, and NH₂-RTAFD-COOH.

25 Peptides and polypeptides of the invention include those containing conservative amino acid substitutions. Such peptides and polypeptides are encompassed by the invention provided the peptide or polypeptide can bind to SOD1 or A β . As used herein, suitable conservative substitutions of amino acids are known to those of skill in this art and can be made generally without altering the biological activity of the resulting molecule. Those of skill in this art
30 recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, e.g., Watson et al. Molecular Biology of the Gene, 4th Edition, 1987, The Benjamin/Cummings Pub. co., p. 224). Such substitutions can be made in accordance with those set forth as follows: Ala (A) Gly; Ser Arg (R) Lys Asn (N) Gln; His Cys (C) Ser Gln (Q) Asn Glu (E) Asp Gly (G) Ala; Pro His (H) Asn;

Gln Ile (I) Leu; Val Leu (L.) Ile; Val Lys (K) Arg; Gln; Glu Met (N) Leu; Tyr; Ile Phe (F) Met; Leu; Tyr Ser (S) Thr Thr (T) Ser Trp (W) Tyr Tyr (Y) Trp; Phe Val (V) Ile; Leu. Other substitutions also are permissible and can be determined empirically or in accord with known conservative substitutions.

5 The peptidic compounds of the present invention can inhibit protein misfolding and aggregation, wherein said peptidic compounds can be a head-to-tail cyclic peptide, side-chain-to-tail cyclic peptide, bicyclic peptide, lanthipeptide, linaridin, proteusin, cyanobactin, thiopeptide, bottromycin, microcin, lasso peptide, microviridin, amatoxin, phallotoxin, θ -defensin, orbitide, or cyclotide.

10 The peptidic compounds of the present invention also serve as structural models for non-peptidic molecules or “mimetics” with similar biological activity. One can also readily modify peptides by phosphorylation, and other methods for making peptide derivatives of the compounds of the present invention are described in Hruby, et al. *Biochim. J.* 268(2): 249-262, 1990, incorporated herewith by reference. Thus, the peptide compounds of the invention also
15 serve as structural models for non-peptidic compounds with similar biological activity. Those of skill in the art recognize that a variety of techniques are available for constructing compounds with the same or similar desired biological activity as the lead peptide compound but with more favorable activity than the lead with respect to solubility, stability, and susceptibility to hydrolysis and proteolysis. See Morgan and Gainor, *Ann. Rep. Med. Chem.*
20 24:243-252, 1989, incorporated herein by reference. These techniques include replacing the peptide backbone with a backbone composed of phosphonates, amidates, carbamates, sulfonamides, and secondary amines.

 The term mimetic, and in particular, peptidomimetic, is intended to include isosteres. The term “isostere” as used herein is intended to include a chemical structure that can be
25 substituted for a second chemical structure because the steric conformation of the first structure fits a binding site specific for the second structure. The term specifically includes peptide backbone modifications (i.e., amide bond mimetics) well known to those skilled in the art. Such modifications include modifications of the amide nitrogen, the α -carbon, amide carbonyl, complete replacement of the amide bond, extensions, deletions or backbone crosslinks. Several
30 peptide backbone modifications are known, including $\psi[\text{CH}_2\text{S}]$, $\psi[\text{CH}_2\text{NH}]$, $\psi[\text{CSNH}_2]$, $\psi[\text{NHCO}]$, $\psi[\text{COCH}_2]$, and $\psi[(\text{E})\text{or}(\text{Z})\text{CH}=\text{CH}]$. In the nomenclature used above, ψ indicates the absence of an amide bond. The structure that replaces the amide group is specified within the brackets. Other examples of isosteres include peptides substituted with one or more benzodiazepine molecules (see e.g., James, G. L. et al. (1993) *Science* 260:1937—1942).

Other possible modifications include an N-alkyl (or aryl) substitution ($\psi[\text{CONR}]$), backbone crosslinking to construct lactams and other cyclic structures, substitution of all D-amino acids for all L-amino acids Within the compound (“inverso” compounds) or retro-inverso amino acid incorporation ($\psi[\text{NHCO}]$). By “inverso” is meant replacing L-amino acids of a sequence With D-amino acids, and by “retro-inverso” or “enantio-retro” is meant reversing the sequence of the amino acids (“retro”) and replacing the L-amino acids With D-amino acids. For example, if the parent peptide is Thr-Ala-Tyr, the retro modified form is Tyr-Ala-Thr, the inverso form is thr-ala-tyr, and the retro inverso form is tyr-ala-thr (lower case letters refer to D-amino acids). Compared to the parent peptide, a retro-inverso peptide has a reversed backbone while retaining substantially the original spatial conformation of the side chains, resulting in a retro-inverso isomer with a topology that closely resembles the parent peptide. See Goodman et al. “Perspectives in Peptide Chemistry” pp. 283—294 (1981). See also US. Pat. No. 4,522,752 by Sisto for further description of “retro-inverso” peptides.

Preferably, the modulator compound inhibits aggregation of SOD1 or natural β -amyloid peptides when contacted with SOD1 or the natural β -amyloid peptides, and/or inhibits SOD1 or A β neurotoxicity. Alternatively, the modulator compound can promote aggregation of SOD1 or natural β -amyloid peptides when contacted with the SOD1 or natural β -amyloid peptides. The type and number of modifying groups coupled to the modulator are selected such that the compound alters (and preferably inhibits) aggregation of SOD1 or natural β -amyloid peptides when contacted with SOD1 or the natural β -amyloid peptides. A single modifying group can be coupled to the modulator or, alternatively, multiple modifying groups can be coupled to the modulator.

Within a modulator compound of the invention, a peptidic structure (such as a cyclic oligopeptide SOD1 or A β modulator or an amino acid sequence corresponding to a rearranged or modified cyclic oligopeptide SOD1 or A β modulator) is coupled directly or indirectly to at least one modifying group. The term “modifying group” is intended to include structures that are directly attached to the peptidic structure (e.g., by covalent coupling), as well as those that are indirectly attached to the peptidic structure (e.g., by a stable non-covalent association or by covalent coupling to additional amino acid residues, or mimetics, analogues or derivatives thereof, which may flank the cyclic oligopeptide SOD1 or A β modulator). For example, the modifying group can be coupled to a side chain of at least one amino acid residue of a cyclic oligopeptide SOD1 or A β modulator, or to a peptidic or peptidomimetic region flanking the cyclic oligopeptide SOD1 or A β modulator (e.g., through the epsilon amino group of a lysyl residue(s), through the carboxyl group of an aspartic acid residue(s) or a glutamic acid

residue(s), through a hydroxy group of a tyrosyl residue(s), a serine residue(s) or a threonine residue(s) or other suitable reactive group on an amino acid side chain). Modifying groups covalently coupled to the peptidic structure can be attached by means and using methods well known in the art for linking chemical structures, including, for example, amide, alkylamino, carbamate or urea bonds.

The modifying group(s) is selected such that the modulator compound alters, and preferably inhibits, SOD1 or β -amyloid peptides aggregation when contacted with SOD1 or the β -amyloid peptides or inhibits the neurotoxicity of SOD1 or the β -amyloid peptides when contacted by them. Although not intending to be limited by mechanism, the modifying group(s) of the modulator compounds of the invention is thought to function as a key pharmacophore which is important for conferring on the modulator the ability to disrupt SOD1 or A β aggregation.

In one embodiment, the modifying group is a "biotinyl structure", which includes biotinyl groups and analogues and derivatives thereof (such as a 2-iminobiotinyl group). In another embodiment, the modifying group can comprise a "fluorescein-containing group", such as a group derived from reacting a SOD1- or an A β -derived peptidic structure with 5-(and 6-)carboxyfluorescein, succinimidyl ester or fluorescein isothiocyanate. In various other embodiments, the modifying group(s) can comprise an N-acetylneuraminyl group, a trans-4-cotininecarboxyl group, a 2-imino-1-imidazolidineacetyl group, an (S)-(-)-indoline-2-carboxyl group, a (-)-menthoxyacetyl group, a 2-norbornaneacetyl group, a γ -oxo-5-acenaphthenebutyryl, a (-)-2-oxo-4-thiazolidinecarboxyl group, a tetrahydro-3-furoyl group, a 2-iminobiotinyl group, a diethylenetriaminepentaacetyl group, a 4-morpholinecarbonyl group, a 2-thiopheneacetyl group or a 2-thiophenesulfonyl group.

Preferred modifying groups include groups comprising cholyl structures, biotinyl structures, fluorescein-containing groups, a diethylene-triaminepentaacetyl group, a (-)-menthoxyacetyl group, and a N-acetylneuraminyl group. More preferred modifying groups those comprising a cholyl structure or an iminiobiotinyl group. Yet another type of modifying group is a compound that contains a non-natural amino acid.

SOD1 or β -amyloid modulator compounds of the invention can be further modified to alter the specific properties of the compound while retaining the ability of the compound to alter SOD1 or A β aggregation and inhibit SOD1 or A β neurotoxicity. For example, in one embodiment, the compound is further modified to alter a pharmacokinetic property of the compound, such as *in vivo* stability or half-life. In another embodiment, the compound is further modified to label the compound with a detectable substance. In yet another

embodiment, the compound is further modified to couple the compound to an additional therapeutic moiety. To further chemically modify the compound, such as to alter the pharmacokinetic properties of the compound, reactive groups can be derivatized.

A modulator compound can be further modified to label the compound by reacting the compound with a detectable substance. Suitable detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; and examples of suitable radioactive material include ^{14}C , ^{123}I , ^{124}I , ^{125}I , ^{131}I , $^{99\text{m}}\text{Tc}$, ^{35}S or ^3H . In a preferred embodiment, a modulator compound is radioactively labeled with ^{14}C , either by incorporation of ^{14}C into the modifying group or one or more amino acid structures in the modulator compound. Labeled modulator compounds can be used to assess the *in vivo* pharmacokinetics of the compounds, as well as to detect SOD1 or A β aggregation, for example for diagnostic purposes. SOD1 or A β aggregation can be detected using a labeled modulator compound either *in vivo* or in an *in vitro* sample derived from a subject.

Preferably, for use as an *in vivo* diagnostic agent, a modulator compound of the invention is labeled with radioactive technetium or iodine. Accordingly, in one embodiment, the invention provides a modulator compound labeled with technetium, preferably $^{99\text{m}}\text{Tc}$. Methods for labeling peptide compounds with technetium are known in the art (see e.g., U.S. Pat. Nos. 5,443,815, 5,225,180 and 5,405,597, all by Dean et al.; Stepniak-Biniakiewicz, D., et al. (1992) J. Med. Chem. 35:274-279; Fritzberg, A. R., et al. (1988) Proc. Natl. Acad. Sci. USA 85:4025-4029; Baidoo, K. E., et al. (1990) Cancer Res. Suppl. 50:799s-803s; and Regan, L. and Smith, C. K. (1995) Science 270:980-982).

Furthermore, an additional modification of a modulator compound of the invention can serve to confer an additional therapeutic property on the compound. That is, the additional chemical modification can comprise an additional functional moiety. For example, a functional moiety which serves to break down or dissolve amyloid plaques can be coupled to the modulator compound. In this form, the modified modulator serves to target the compound to SOD1 or A β peptides and disrupt their polymerization, whereas the additional functional moiety serves to break down or dissolve SOD1 aggregates or amyloid plaques after the compound has been targeted to these sites.

In an alternative chemical modification, a SOD1 or β -amyloid compound of the invention is prepared in a "prodrug" form, wherein the compound itself does not modulate aggregation, but rather is capable of being transformed, upon metabolism *in vivo*, into a SOD1 or β -amyloid modulator compound as defined herein. For example, in this type of compound, the modulating group can be present in a prodrug form that is capable of being converted upon metabolism into the form of an active modulating group. Such a prodrug form of a modifying group is referred to herein as a "secondary modifying group." A variety of strategies are known in the art for preparing peptide prodrugs that limit metabolism in order to optimize delivery of the active form of the peptide-based drug (see e.g., Moss, J. (1995) in Peptide-Based Drug Design: Controlling Transport and Metabolism, Taylor, M. D. and Amidon, G. L. (eds), Chapter 18. Additionally strategies have been specifically tailored to achieving CNS delivery based on "sequential metabolism" (see e.g., Bodor, N., et al. (1992) Science 257:1698-1700; Prokai, L., et al. (1994) J. Am. Chem. Soc. 116:2643-2644; Bodor, N. and Prokai, L. (1995) in Peptide-Based Drug Design: Controlling Transport and Metabolism, Taylor, M. D. and Amidon, G. L. (eds), Chapter 14. In one embodiment of a prodrug form of a modulator of the invention, the modifying group comprises an alkyl ester to facilitate blood-brain barrier permeability.

Modulator compounds of the invention can be prepared by chemical synthesis using standard techniques known in the art. The peptide component of a modulator composed, at least in part, of a peptide, can be synthesized using standard techniques such as those described in Bodansky, M. Principles of Peptide Synthesis, Springer Verlag, Berlin (1993) and Grant, G. A. (ed.). Synthetic Peptides: A User's Guide, W. H. Freeman and Company, New York (1992). Automated peptide synthesizers are commercially available (e.g., Advanced ChemTech Model 396; Milligen/Bioscience 9600). Additionally, one or more modulating groups can be attached to the SOD1 or A β modulator (e.g., a SOD1 or an A β aggregation core domain) by standard methods, for example using methods for reaction through an amino group, a carboxyl group, a hydroxyl group (e.g., on a tyrosine, serine or threonine residue) or other suitable reactive group on an amino acid side chain (see e.g., Greene, T. W. and Wuts, P. G. M. Protective Groups in Organic Synthesis, John Wiley and Sons, Inc., New York (1991)).

Alternatively, modulator compounds of the invention can be prepared biosynthetically and isolated in pure or enriched form from a recombinant production host, such a bacterial, yeast, plant, or mammalian cell (see, e.g., Scott CP, Abel-Santos E, Jones AD, Benkovic SJ, Structural requirements for the biosynthesis of backbone cyclic peptide libraries. Chem Biol.

2001 Aug;8(8):801-15, as an example of recombinant production of cyclic oligopeptides in bacterial cells).

Alternatively, modulator compounds of the invention can be prepared biosynthetically from a recombinant production host, such a bacterial, yeast, plant, or mammalian cell, but may not be isolated in pure or enriched form, and instead be provided to the diseased organism as part of recombinant production host, such a bacterial, yeast, plant, or mammalian cell producing the specific modulator compound recombinantly in the form of a probiotic. By the term “probiotic”, we mean living microorganisms or other cultured cells that may provide health benefits when administered and consumed in adequate amounts (see, e.g., O'Toole PW, Marchesi JR, Hill C, Next-generation probiotics: the spectrum from probiotics to live biotherapeutics, *Nat Microbiol.* 2017 Apr 25;2:17057).

IV. Hybrid modulators

A hybrid molecule of the invention includes a peptide or polypeptide that binds to the amyloid or non-amyloid form of SOD1 or amyloid form of A β , and a scaffold molecule. The scaffold molecule can include a diagnostic or therapeutic reagent. The therapeutic or diagnostic reagent can be a polypeptide, small molecule or compound.

In particular, provided herein are hybrid molecules, such as hybrid polypeptides, that include a peptide or polypeptide provided herein, and additional amino acid residues (typically, 5, 10, 15, 20, 30, 40, 50, 100 or more) such that the resulting hybrid molecule specifically interacts with SOD1 or A β . The motif can be modified, such as by replacing certain amino acids or by directed and random evolution methods, to produce motifs with greater affinity. As used herein, a hybrid polypeptide refers to a polypeptide that includes regions from at least two sources, such as from an antibody or enzyme or other scaffold that can be a recipient, and a binding motif, such as a polypeptide or peptide that binds to an amyloid or non-amyloid form of SOD1 or the amyloid form of the A β peptide.

Thus, among the hybrid molecules provided herein are hybrid molecules, particularly hybrid polypeptides that are produced by grafting a binding motif (e.g., peptide) from one molecule into a scaffold, such as an antibody or fragment thereof or an enzyme or other reporter molecule. The hybrid polypeptides provided herein, even the hybrid immunoglobulins, are not antibodies per se, but are polypeptides that are hybrid molecules containing a selected motif (e.g., a peptide that binds to the amyloid or non-amyloid form of SOD1 or the amyloid form of the A β peptide) inserted into another polypeptide such that the motif retains or obtains the ability to bind to a protein involved in disease of protein aggregation. The hybrid polypeptides can include portions of antibodies or other scaffolds, but they also include a non-

immunoglobulin or non-scaffold portion grafted therein. The non-immunoglobulin portion is identified by its ability to specifically bind to a targeted polypeptide isoform. The hybrid polypeptide can specifically bind to the targeted infectious or disease-related or a selected isoform of a polypeptide as monomer with sufficient affinity to detect the resulting complex or to precipitate the targeted polypeptide.

The scaffold is selected so that insertion of the motif therein does not substantially alter (i.e., retains) the desired binding specificity of the motif. The scaffold additionally can be selected for its properties, such as its ability to act as a reporter.

Methods for production of hybrid molecules that specifically interact with a one form of a conformer of a protein associated with a disease of protein conformation or involving protein aggregation are provided. In these methods a polypeptide motif from the protein is inserted into a scaffold such that the resulting molecule exhibits specific binding to one conformer compared to other conformers. In particular, the hybrid molecule can exhibit specific binding to the amyloid or non-amyloid form of SOD1 or the amyloid form of the A β peptide.

Peptides of the invention have been shown to bind to SOD1 or A β *in vitro* and *in vivo*. The peptides can be incorporated into a scaffold that comprises additional amino acid sequences and/or compounds. The hybrid molecule can then be used to label or treat the aggregates associated with SOD1 or plaques associated with A β amyloid. The polypeptides, nucleic acids encoding the polypeptides, and methods of using the polypeptides or nucleic acids can be used to identify, diagnose and/or treat disorders associated with plaque formation in brain tissue.

Any molecule, such as a polypeptide, into which the selected polypeptide motif is inserted (or linked) such that the resulting hybrid polypeptide has the desired binding specificity, is contemplated for use as part of the hybrid molecules herein. The polypeptides can be inserted into any sequence of amino acids that at least contains a sufficient number (10, 20, 30, 50, 100 or more amino acids) to properly present the motif for binding to the targeted amyloid plaque. The purpose of the scaffold is to present the motif to the targeted polypeptide in a form that binds thereto. The scaffold can be designed or chosen to have additional properties, such as the ability to serve as a detectable marker or label or to have additional binding specificity to permit or aid in its use in assays to detect particular isoforms of a target protein (e.g., the amyloid or non-amyloid form of SOD1 or the amyloid form of the A β peptide) or for screening for therapeutics or other assays and methods.

The scaffolds include reporter molecules, such as fluorescent proteins and enzymes or fragments thereof, and binding molecules, such as antibodies or fragments thereof. Selected scaffolds include all or portions of antibodies, enzymes, such as luciferases, alkaline phosphatases, β -galactosidases and other signal-generating enzymes, chemiluminescence generators, such as horseradish peroxidase; fluorescent proteins, such as red, green and blue fluorescent proteins, which are well known; and chromogenic proteins.

The peptide motif is inserted into the scaffold in a region that does not disturb any desired activity. The scaffolds can include other functional domains, such as an additional binding site, such as one specific for a second moiety for detection.

V. Nucleic Acid Molecules

Nucleic acid molecules encoding any of the peptides, polypeptides or hybrid polypeptides provided herein are provided in the general experimental procedures. Such molecules can be introduced into plasmids and vectors for expression in suitable host cells. As used herein, the term “nucleic acid” refers to single-stranded and/or double-stranded polynucleotides such as deoxyribonucleic acid (DNA), and ribonucleic acid (RNA) as well as analogs or derivatives of either RNA or DNA. Also included in the term “nucleic acid” are analogs of nucleic acids such as peptide nucleic acid (PNA), phosphorothioate DNA, and other such analogs and derivatives or combinations thereof. The term should be understood to include, as equivalents, derivatives, variants and analogs of either RNA or DNA made from nucleotide analogs, single (sense or antisense) and double-stranded polynucleotides. Deoxyribonucleotides include deoxyadenosine, deoxycytidine, deoxyguanosine and deoxythymidine. For RNA, the uracil base is uridine.

Plasmids and vectors containing the nucleic acid molecules also are provided in the general experimental procedures. Cells containing the vectors, including cells that express the encoded proteins are also provided. The cell can be a bacterial cell, a yeast cell, a fungal cell, a plant cell, an insect cell or an animal cell. Methods for producing a cyclic oligopeptide or a hybrid polypeptide, for example, growing the cell under conditions whereby the encoded polypeptide is expressed by the cell, and recovering the expressed protein, are provided herein. The cells are used for expression of the cyclic oligopeptide or the protein, which can be secreted or expressed in the cytoplasm. The hybrid polypeptides also can be chemically synthesized using standard methods of protein synthesis known in the art.

VI. Pharmaceutical compositions

It is envisioned that one would use the modulators of the present invention as an Alzheimer's disease or amyotrophic lateral sclerosis therapeutic. If the modulator were peptide in nature, one could use a gene therapy technique to deliver DNA constructs encoding the modulator to the affected sites. For drug formulations, one would expect that the formulations reach and be effective at the affected site. These modulators would more likely be carbohydrate and peptide mixtures, especially mixtures capable of overcoming the blood brain barrier. For examples, see Tamai, et al., Adv. Drug Delivery Review 19:401-424, 1996, hereby incorporated by reference. In these cases, the disrupting element of the modulators would also facilitate transport across the blood-brain barrier.

Thus, the present invention encompasses methods for therapeutic treatments of amyotrophic lateral sclerosis and Alzheimer's disease, comprising administering a compound of the invention in amounts sufficient to modulate the natural course of SOD1 or A β aggregation. Accordingly, the present invention includes pharmaceutical compositions comprising, as an active ingredient, at least one of the peptides or other compounds of the invention in association with a pharmaceutical carrier or diluent. The compounds of the invention can be administered by oral, parenteral (intramuscular, intraperitoneal, intravenous (IV) or subcutaneous injection), transdermal, nasal, vaginal, rectal, or sublingual routes of administration.

Solid dosage forms for oral administration include capsules, tablets, pills, powders, and granules. In such solid dosage forms, the active compound is admixed with at least one inert pharmaceutically acceptable carrier such as sucrose, lactose, or starch. Such dosage forms can also comprise, as is normal practice, additional substances other than inert diluents, e.g., lubricating agents such as magnesium stearate. In the case of capsules, tablets, and pills, the dosage forms may also comprise buffering agents. Tablets and pills can additionally be prepared with enteric coatings.

Liquid dosage forms for oral administration include pharmaceutically acceptable emulsions, solutions, suspensions, syrups, with the elixirs containing inert diluents commonly used in the art, such as water. Besides such inert diluents, compositions can also include adjuvants, such as wetting agents, emulsifying and suspending agents, and sweetening, flavoring, and perfuming agents.

Preparations according to this invention for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, or emulsions. Examples of non-aqueous solvents or vehicles are propylene glycol, polyethylene glycol, vegetable oils, such as olive oil

and corn oil, gelatin, and injectable organic esters such as ethyl oleate. Such dosage forms may also contain adjuvants such as preserving, wetting, emulsifying, and dispersing agents. They may be sterilized by, for example, filtration through bacteria-retaining filters, by incorporating sterilizing agents into the compositions, by irradiating the compositions, or by heating the compositions. They can also be manufactured using sterile water, or some other sterile injectable medium, immediately before use.

Compositions for rectal or vaginal administration are preferably suppositories which may contain, in addition to the active substance, excipients such as cocoa butter or a suppository wax. Compositions for nasal or sublingual administration area are also prepared with standard excipients well known in the art.

The dosage of active ingredient in the compositions of this invention may be varied; however, it is necessary that the amount of the active ingredient shall be such that a suitable dosage form is obtained. The selected dosage depends upon the desired therapeutic effect, on the route of administration, and on the duration of the treatment desired.

The following examples illustrate aspects of the present invention including the construction and screening of a peptide macrocycle library; the identification of macrocyclic peptide rescuers of the misfolding and aggregation of the prominent PMD-associated protein target SOD1 and fALS-associated variants thereof, as well as as a second prominent PMD-associated protein target, A β , and finally their use in rescuing of SOD1 and A β aggregation and toxicity *in vitro* and *in vivo*. The Examples are not in any way limiting the scope of invention.

EXAMPLES

Example 1

Combinatorial libraries of random cyclic tetra-, penta-, and hexapeptides have been selected to be studied as potential rescuers of SOD1 and mutant SOD1 misfolding and pathogenic aggregation. A technique named split intein circular ligation of peptides and proteins (SICLOPPS) (US 7354756 B1 "Intein-mediated cyclization of peptides") for producing peptide libraries in *E. coli* is being used. SICLOPPS uses split inteins, i.e. self-splicing protein elements, for performing N- to C- terminal peptide cyclization and biosynthesize cyclic peptides as short as four amino acids long. The only requirement for the intein splicing reaction and peptide cyclization to occur is the presence of a nucleophilic amino acids cysteine (C), serine (S), or threonine (T) as the first amino acid of the extein following the C-terminus of the intein.

In order for the inventors to maximize the diversity of the libraries, they chose to study peptides with the general formula cyclo-NuX₁X₂...X_N, where Nu=C, S or T; X is any one of the twenty natural amino acids and N=3-5 (FIG. 1A). The maximum theoretical diversity of the combined cyclo-NuX₁X₂X₃-X₅ library is > 10 million different sequences (FIG. 1B). The libraries of genes encoding this combinatorial library of random cyclic oligopeptides were constructed using degenerate codons. The inventors constructed the high diversity pSICLOPPS-NuX₁X₂X₃-X₅ vector library which is expected to be encoding all of the theoretically possible designed cyclic tetra-, penta-, and hexapeptide NuX₁X₂X₃-X₅ sequences using molecular biology techniques already known and used in the art.

It has been demonstrated previously that the fluorescence of *E. coli* cells expressing a recombinant protein whose C terminus is fused to GFP correlates well with the amount of soluble and folded protein (Waldo GS, Standish BM, Berendzen J, Terwilliger TC. Nat Biotechnol. 1999 Jul;17(7):691-5). Based on this, the inventors reasoned that the fluorescence of MisP-GFP fusions can serve as a reliable reporter for the identification of chemical rescuers of MisP misfolding for a number of disease-associated MisPs, including SOD1. In order to test this hypothesis, the inventors generated fusions of SOD1 variants, whose misfolding and aggregation have been linked with the pathology of familial forms of ALS (fALS), with GFP. Expression of these fusions in *E. coli*, yielded levels of cellular fluorescence, which were significantly decreased compared to that of the generally non-pathogenic, wild-type SOD1 (FIG. 2A). Western blot analysis indicated that this occurs because the accumulation of soluble SOD1-GFP is decreased in the presence of misfolding-inducing amino acid substitutions, which in turn takes place due to enhanced misfolding/aggregation of GFP-fused, as well as fusion-free SOD1 (FIG. 2B, 2C). Thus, the fluorescence of *E. coli* cells overexpressing SOD1-GFP fusions appears to be a good indicator of SOD1 folding and misfolding.

Example 2

To test whether the bacterial platform can be utilized to identify chemical rescuers of disease-associated SOD1, the inventors screened for cyclic oligopeptides that inhibit the aggregation of SOD1(A4V), a fALS-associated variant, whose misfolding and aggregation causes a very aggressive form of the disease with an average survival time of only 1.2 years after diagnosis. FACS screening of the cyclo-NuX₁X₂X₃-X₅ oligopeptide library for bacterial clones exhibiting enhanced levels of SOD1(A4V)-GFP fluorescence yielded an *E. coli* population with about 10-fold increased fluorescence after four rounds of sorting (FIGS. 3A, 3B). Twenty randomly selected clones from the isolated population exhibited up to 10-fold enhanced fluorescence compared to *E. coli* cells producing randomly selected cyclic

oligopeptides from the initial library. Four of the isolated clones exhibited the highest levels of cellular SOD1(A4V)-GFP fluorescence (FIG. 3C), and were selected for further analysis. These clones (i) expressed tetra-partite I_C-peptide-I_N-CBD fusions, which could undergo splicing (FIG. 3D), (ii) exhibited splicing-activity-dependent enhanced SOD1(A4V)-GFP
5 fluorescence (FIG. 3C), and (iii) exhibited SOD1-specific enhancement of bacterial fluorescence (FIG. 3E). Western blot analysis indicated that this enhanced SOD1(A4V)-GFP fluorescence phenotype occurs due to accumulation of enhanced amounts of soluble SOD1(A4V) in these clones (FIG. 3F). Sequencing of the peptide-encoding region of the pSICLOPPS vector contained in these clones revealed that they all encode cyclic pentapeptides
10 with sequences TASFW, TWSVW, and TFSMW (FIG. 3G), thus indicating a dominant cyclo-TXSXW bioactive motif.

Example 3

The peptide cyclo-TWSVW, hereafter referred to as SOD1C5-4 (FIG. 4A), which was present twice among the four selected clones, was selected for further analysis and was
15 produced in mg quantities by solid-phase synthesis. Isolated SOD1(A4V) was utilized to assess the effect of the selected cyclic pentapeptide SOD1C5-4 on its aggregation process. CD spectroscopy indicated that SOD1C5-4 –but not the control A β -targeting cyclic pentapeptides A β C5-34 or A β C5-116– interacts with SOD1(A4V), and that the time-dependent conformational transition that is indicative of SOD1(A4V) aggregation is significantly delayed
20 in the presence of SOD1C5-4 (FIG. 4B). Moreover, dynamic light scattering (DLS) analysis revealed that SOD1C5-4 addition results in the time-dependent formation of oligomeric/aggregated SOD1(A4V) species with markedly smaller sizes (FIG. 4C). Detection of large, amyloid-like SOD1(A4V) aggregates by ThT staining and a filter retardation assay indicated that the formation of such species was dramatically decreased in the presence of
25 SOD1C5-4 (FIGS. 4D and 4E). Finally, staining of SOD1(A4V) with the conformation-sensitive dye SYPRO Orange under heat-induced denaturation conditions, suggested that the aggregation-inhibitory action of SOD1C5-4 may be occurring due to its ability to decrease the propensity of SOD1(A4V) to expose hydrophobic surfaces (FIG. 4F), a feature which has been proposed to be a molecular determinant of the pathogenesis of fALS-associated SOD1 variants
30 (Münch C, Bertolotti A. J Mol Biol. 2010; 399(3):512-25). Taken together, these results demonstrate that SOD1C5-4 is an efficient and specific rescuer of SOD1(A4V) misfolding and aggregation.

Example 4

The protective effects of SOD1C5-4 in mammalian cells were evaluated in human embryonic kidney 293 (HEK293) cells transiently expressing SOD1(A4V)-GFP. Cells treated with SOD1C5-4 exhibited higher fluorescence, fewer inclusions comprising aggregated SOD1(A4V)-GFP, and higher viability compared to untreated cells (FIGS. 5A-5C).

5 Example 5

To determine structure-activity relationships for the identified mutant SOD1-targeting cyclic oligopeptides, the sequences of the peptide-encoding regions from ~5.3 million clones selected after the fourth round of FACS sorting (FIG. 3B) were determined by deep sequencing. 367 distinct oligopeptide sequences appeared more than 50 times among the selected clones and were selected for subsequent analysis, which revealed the following. First, pentapeptides were the dominant peptide species within the sorted pool, with 197 of the distinct oligopeptide sequences selected corresponding to pentapeptides (54%), 148 to hexapeptides (40%) and 22 corresponding to tetrapeptides (6%) (FIG. 6A). Second, the vast majority of the selected peptides exhibited the cyclo-TXSXW motif of SOD1C5-4 (~92% of all selected clones and ~97% of the selected pentapeptide-encoding clones (FIG. 7). Third, among the selected cyclo-TXSXW pentapeptides, I, N, Q, M, E, H, and K residues were excluded at position 2, and were preferably S, A, W or F. At position 4, I, N, Q, C, D, E, K and P residues were excluded, and were preferably V, W, F, M, or H (FIGS. 6B-6D). Taken together, these results indicate that the most bioactive macrocyclic structures against SOD1(A4V) misfolding and aggregation in the library are cyclic pentapeptides of the cyclo-T(Φ_1 ,S)S(Φ_2 ,M,H)W motif, where Φ_1 is preferably one of the hydrophobic (Φ) amino acids A, W or F, while Φ_2 is preferably V, W or F. Interestingly, selected cyclic pentapeptides belonging to this functional motif were found to be efficient in enhancing the fluorescence of SOD1-GFP containing wild-type SOD1, as well as three additional SOD1 variants, SOD1(G37R), SOD1(G85R), and SOD1(G93A), all of which are associated with familial forms of ALS, thus indicating that these peptide macrocycles are effective rescuers of the misfolding of not only SOD1(A4V), but also of other ALS-related SOD1 variants, as well as wild-type SOD1 (FIG. 6E).

25 Example 6

Combinatorial libraries of random cyclic tetra-, penta-, and hexapeptides have been selected to be studied as potential rescuers of A β misfolding and pathogenic aggregation.

To identify cyclic oligopeptide sequences with the ability to interfere with the problematic folding of A β and inhibit its oligomerization/aggregation, the inventors utilized a bacterial high-throughput genetic screen. This system monitors A β misfolding and aggregation by measuring the fluorescence of *E. coli* cells overexpressing a chimeric fusion of the human

A β (1-42) peptide (A β ₄₂) with GFP (US20070077552A1 “High throughput screen for inhibitors of polypeptide aggregation”). It has been demonstrated previously that due to the high aggregation propensity of A β , *E. coli* cells overexpressing A β -GFP fusions produce misfolded fusion protein that accumulates into insoluble inclusion bodies that lack fluorescence, despite the fact that they express these fusions at high levels. Mutations in the coding sequence of A β or the addition of compounds that inhibit A β aggregation, however, result in the formation of soluble and fluorescent A β -GFP, and bacterial cells expressing A β -GFP under these conditions acquire a fluorescent phenotype. The inventors of the present invention adapted this system to perform screening for aggregation-inhibitory macrocycles in a very high-throughput fashion by isolating cyclic oligopeptide-producing bacterial clones that exhibit enhanced levels of A β ₄₂-GFP fluorescence using fluorescence-activated cell sorting (FACS) as also performed in a similar manner and demonstrated for SOD1 in FIG. 3A.

E. coli BL21(DE3) cells producing the combined cyclo-NuX₁X₂X₃-X₅ library, while simultaneously overexpressing the A β ₄₂-GFP reporter, were subjected to FACS sorting for the isolation of clones exhibiting enhanced A β ₄₂-GFP fluorescence. After two rounds of sorting, the mean fluorescence of the bacterial population increased by almost three-fold compared to that of the initial library (FIG. 8A). Ten individual clones were randomly picked from the sorted population and their A β ₄₂-GFP fluorescence was measured using a plate reader. A β ₄₂-GFP fluorescence of the isolated peptide-expressing clones was found to be dramatically increased compared to cells expressing the same A β ₄₂-GFP fusion in the presence of random cyclic peptide sequences picked from the initial (unselected) cyclo-NuX₁X₂X₃-X₅ library (FIG. 8B). Furthermore, the observed phenotypic effects were dependent on the ability of the Ssp DnaE intein to perform protein splicing, as the double amino acid substitution H24L/F26A in the C-terminal half of the Ssp DnaE intein, which is known to abolish asparagine cyclization at the I_C/extein junction, and prevent extein splicing and peptide cyclization, was found to reduce A β ₄₂-GFP fluorescence back to wild-type levels (FIG. 8B). Finally, the observed increases in fluorescence were found to be A β -specific, as the isolated pSICLOPPS-NuX₁X₂X₃-X₅ vectors did not enhance the levels of cellular green fluorescence when the sequence of A β in the A β ₄₂-GFP reporter was replaced with those of two unrelated disease-associated MisPs, the DNA-binding (core) domain of the human p53 containing a tyrosine to cysteine substitution at position 220 (p53C(Y220C)) and an alanine to valine substitution at position 4 of human Cu/Zn superoxide dismutase 1 (SOD1(A4V)) (FIG. 8C). On the contrary, the selected pSICLOPPS-NuX₁X₂X₃-X₅ vectors were efficient in enhancing the fluorescence of A β -GFP containing two

additional A β variants, A β ₄₀ and the E22G (arctic) variant of A β ₄₂, which is associated with familial forms of AD (FIG. 8D).

Analysis of the expressed A β ₄₂-GFP fusions by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and western blotting revealed that the bacterial clones
5 expressing the selected cyclic peptides produce markedly increased levels of soluble A β ₄₂-GFP compared to random cyclic peptide sequences, despite the fact that accumulation of total A β ₄₂-GFP protein remained at similar levels in all cases (FIG. 8E). Furthermore, when the same cell lysates were analyzed by native PAGE and western blotting, it was observed that co-expression
10 of the selected cyclic peptides reduced the accumulation of higher-order A β ₄₂-GFP aggregates, which could not enter the gel, and increased the abundance of species with higher electrophoretic mobility (FIG. 8F, left). As revealed by in-gel fluorescence analysis, these higher electrophoretic mobility species correspond to the fraction of the total A β ₄₂-GFP that exhibits fluorescence (FIG. 8F, right). Since the solubility and fluorescence of bacterially expressed A β -GFP has been found to be inversely proportional to the aggregation propensity
15 of A β , these results suggest that A β aggregation is significantly decreased in the presence of the selected cyclic peptides.

DNA sequencing of the peptide-encoding regions of ten isolated clones from the selected pool revealed eight distinct putative A β aggregation-inhibitory cyclic peptide sequences: one corresponded to a hexapeptide (TPVWFD; present twice among the sequenced
20 clones) and seven pentapeptides (TAFDR, TAWCR, TTWCR, TTVDR, TTYAR (present twice), TTTAR, and SASPT) (FIG. 8G). Interestingly, the Arg residue at position 5, frequently encountered among the selected pentapeptides, was encoded by three different codons in the selected pSICLOPPS plasmids, thus suggesting that its dominance among the isolated clones was not coincidental.

25 Example 7

Two of the selected cyclic peptide sequences, cyclo-TAFDR and cyclo-SASPT, hereafter referred to as A β C5-116 and A β C5-34 (A β -targeting cyclic 5-peptide number 116 and 34), respectively, were chosen for subsequent analysis and were produced by solid-phase synthesis in mg quantities (FIG. 9A). The inventors of the present invention chose to focus on
30 pentapeptides, as this was the type of peptide most frequently present among the ones selected from the genetic screen. The inventors decided to further study the sequence A β C5-116 since the cyclo-TXXXXR motif was particularly dominant among the selected pentapeptides, while A β C5-34 was chosen because it was the only selected pentapeptide whose sequence appeared to deviate from this motif (FIG. 8G).

Circular dichroism (CD) spectroscopy was first used to assess the effect of the selected pentapeptides on the aggregation process of A β ₄₀ and A β ₄₂. Addition of A β C5-116 was found to strongly inhibit the aggregation of A β ₄₀, which remained at a random coil conformation in the presence of this cyclic peptide for extended periods of time (FIG. 9B). The addition of A β C5-34 did not have the same effect and resulted in the appearance of a low-intensity negative peak (FIG. 9B). When the same CD solutions were subjected to a thioflavin T (ThT) dye-binding assay that detects amyloid fibrils, we observed that A β ₄₀ fibril formation was reduced in the presence of A β C5-116, while it remained almost unaffected by A β C5-34 (FIG. 9C).

In the case of A β ₄₂, both selected cyclic pentapeptides affected its normal aggregation pathway strongly and stabilized β -sheet-like structures (FIG. 9B). ThT staining of the same samples revealed that the extent of amyloid fibril formation was greatly reduced in both cases (FIG. 9C). When the cyclic peptides were added at a higher ratio, similar CD patterns were observed, however the negative peaks were much more pronounced and fibril formation was completely prevented (FIG. 9C, bottom; FIG. 9D). The addition of two control peptides, a randomly designed cyclic pentapeptide sequence and a cyclic pentapeptide (SODC5-4) targeting a different protein did not have any effect on the aggregation process of A β ₄₀ and A β ₄₂ (FIGS. 9B-9C, and data not shown), thus demonstrating that cyclic peptides are not general inhibitors of A β aggregation and that the A β aggregation-modulating effect of the selected sequences relies on their cyclic nature.

Transmission electron microscopy (TEM) images of solutions of A β ₄₂ incubated without/with A β C5-34 and A β C5-116 are presented in FIG. 9E. The A β ₄₂ samples were the same as those employed in the CD and ThT studies to allow for direct correlation of findings. A β ₄₂ incubated alone presented the typical dense network of intertwined fibrils. In the presence of the selected cyclic peptides, however, the fibrils were notably fewer, shorter and ill-developed, and the dense fibrillary network observed in their absence was not detected anywhere on the TEM grid, in agreement with the ThT data. Taken together, these results indicate that the selected cyclic oligopeptides modulate the normal aggregation process of A β , and their presence likely stabilizes the formation of species, which cannot develop into larger fibril-like structures.

Example 8

The effects of A β C5-34 and A β C5-116 on A β ₄₀- and A β ₄₂-induced toxicity were evaluated in primary mouse hippocampal neurons. The addition of A β C5-34 and A β C5-116 was found to markedly inhibit the neurotoxicity of both A β ₄₀ and A β ₄₂ in a dose-responsive manner (FIG. 10A). Similarly, A β C5-34 and A β C5-116 exhibited toxicity-suppressing effects

also in the glioblastoma cell line U87MG (FIG. 10B). On their own, A β C5-34 and A β C5-116 did not exhibit general growth-promoting effects or considerable cytotoxicity (FIGS. 10A and 10B). Control SOD1-targeting cyclic peptides previously found not to interfere with A β aggregation (FIGS. 9B and 9C and data not shown), were also found ineffective in rescuing A β -induced cytotoxicity (FIG. 10C).

The effect of A β C5-34 and A β C5-116 on the morphology of A β -exposed neuronal cells was assessed by phase-contrast microscopy. In the presence of pre-aggregated A β , the population of attached cells was greatly reduced compared to the control, with many detached rounded-up cells floating in the supernatant, while hallmarks of degenerating neurons, such as cell shrinkage, membrane blebbings, fragmented neurites and ill-developed axons were obvious in the preparations (FIG. 10D). The addition of the selected cyclic peptides, however, mitigated the effects of A β toxicity and a marked recovery of the A β -induced alterations was recorded (FIG. 10D).

Example 9

To evaluate the protective effects of the selected cyclic peptides against A β aggregation and toxicity *in vivo*, the inventors employed two established models of AD in the nematode worm *Caenorhabditis elegans*. The conservation of genetic and metabolic pathways between *C. elegans* and mammals, in combination with its completely characterized nervous and muscular system, its easy visualization and simple manipulation, has nominated *C. elegans* as an excellent model for neurodegenerative diseases including AD, while chemical screening against A β -induced toxicity in *C. elegans* is increasingly used in AD drug discovery. The inventors performed initially a paralysis assay in CL2006, a strain where human A β ₄₂ is constitutively expressed in the body wall muscle cells of the animals and A β aggregate formation is accompanied by adult-onset paralysis. Animals fed throughout their lifespan with *E. coli* OP50 cells producing A β C5-34 or A β C5-116 biosynthetically from their corresponding pSICLOPPS vectors, exhibited a significant delay in the appearance of the characteristic paralysis phenotype (FIG. 11A). Similar protective effects were observed in a dose-responsive fashion when synthetic A β C5-34 or A β C5-116 were supplied to CL4176, a strain conditionally expressing human A β ₄₂ under the control of a heat-inducible promoter. When chemically synthesized A β C5-34 (10 μ M) and A β C5-116 (5 μ M) were supplied to CL4176 animals, a significant delay in the appearance of the characteristic paralysis phenotype was recorded indicating protective effects against A β aggregation and toxicity (FIG. 11B). To evaluate the state of A β aggregation *in vivo*, we utilized the strain CL2331, which expresses an A β ₃₋₄₂-GFP

fusion in its body wall muscle cells upon temperature up-shift. Treatment with either one of the selected peptides resulted in a significant reduction of A β deposits (FIG. 11C). Biochemical analysis of the accumulation levels of total and oligomeric A β levels in CL4176 worms, revealed a significant reduction of both A β species upon treatment with A β C5-34 and A β C5-116 (FIG. 11D), an effect coinciding with the observed decelerated paralysis. Taken together, our results demonstrate that A β C5-34 and A β C5-116 exert a protective role against A β

Example 10

To identify the functionally important residues within the isolated peptides, the inventors performed position 1 substitutions with the other two nucleophilic amino acids present in the initial libraries, as well as alanine scanning mutagenesis at positions 3-5 of the A β C5-34 and A β C5-116 pentapeptides. As judged by the ability of the generated variants to enhance the fluorescence of *E. coli* cells overexpressing A β ₄₂-GFP, A β C5-116 was found to be much more tolerant to substitutions compared to A β C5-34. All tested sequence alterations within A β C5-34 were found to be deleterious for its A β aggregation-inhibitory effects (FIG. 12A). On the contrary, only the initial Thr and the ultimate Arg were found to be absolutely necessary for the bioactivity of A β C5-116, whereas residues at positions 3 and 4 could be substituted by Ala without significant loss of activity (FIG. 12B). These observations are in line with the high frequency of initial Thr and ultimate Arg residues in the sequences of the isolated pentapeptides, as well as with the high amino acid variabilities at the corresponding positions 3 and 4, and indicates that all isolated sequences with this pattern may belong to the same consensus motif. In order to investigate this hypothesis, the inventors performed semi-saturation mutagenesis of the Ala residue of A β C5-116 with representative amino acids from all categories. Among the tested amino acids, only Thr and Ser could be tolerated at position 2 (FIG. 12C), in agreement with the fact that four out of six identified pentapeptides containing the cyclo-TXXXXR motif included a Thr at position 2.

The results presented in the invention indicated that there should be a significant number of pentapeptide sequences with the ability to modulate A β aggregation that resemble A β C5-116. On the other hand, very few bioactive sequences resembling A β C5-34 should exist. To test this hypothesis and to identify all the additional bioactive cyclic oligopeptide sequences with the ability to modulate A β oligomerization/aggregation, the inventors turned back to the selected bacterial population exhibiting high A β ₄₂-GFP fluorescence (FIG. 8A). The peptide-encoding vectors contained in these clones were isolated and the peptide-encoding region of approximately 5.6 million of these plasmids was sequenced using an Ion Torrent high-

throughput sequencing platform. 605 distinct oligopeptide sequences appeared more than 50 times within the analyzed population, suggesting that their presence in the isolated pool is not coincidental. Indeed, cloning of four randomly chosen sequences appearing in the sorted pool only with very low frequencies, revealed that they are also efficient in increasing the fluorescence of bacterially expressed A β ₄₂-GFP (FIG. 12F; Table 1). Analysis of the peptide sequences isolated from the genetic screen, and after considering all circular permutants thereof, revealed the following. First, pentapeptides were the dominant peptide species within the sorted pool (FIG. 12D), in agreement with previous observations (Fig. 8G). Second, the most prevalent motif among the selected pentapeptide sequences were TXXXR pentapeptides (~47% of the selected pentapeptide-encoding pSICLOPPS plasmids; ~42% of the unique selected pentapeptide sequences) (Table 1), in accordance with previous observations (FIG. 8G). On the contrary, only three pentapeptide sequences were found to have high similarity with A β C5-34 (Table 2). Third, for the selected peptides corresponding to the TXXXR motif, residues at positions 3 and 4 were highly variable and included the majority of natural amino acids, with position 3 exhibiting the highest diversity (FIG. 12E). At position 2, Thr, Ala, and Val were preferred, while aromatic residues (Phe, Trp, Tyr) were completely excluded from the selected TXXXR peptide pool, in full agreement with our site-directed mutagenesis studies. At the highly variable position 3, the complete absence of the negatively charged amino acids Glu and Asp among the selected sequences was notable (FIG. 12E). In general, both negatively (Glu and Asp) and positively charged residues (Lys, His, and Arg) were found to be strongly disfavored among the selected TXXXR sequences at positions 2 and 3. At position 4, Ala, Asp, and Trp were found to be the preferred residues. It is noteworthy, that Lys and Gln residues were practically absent from all positions, while the β sheet-breaking amino acid Pro that is typically included in designed peptide-based inhibitors of amyloid aggregation appeared with strikingly low frequencies (FIG. 12E). The motif cyclo-T(T,A,V) Ψ (A,D,W)R, where Ψ is anyone of the twenty natural amino acids excluding negatively charged ones, was found to be the most bioactive motif against A β in the investigated macrocycle library.

The high residue variability observed at position 3 of the selected TXXXR peptides prompted the inventors to investigate whether A β C5-116 could be further minimized. Indeed, production of truncated variants of A β C5-116, from which Ala₂ or Asp₄ had been deleted, resulted in a respective two- and three-fold enhancement in the fluorescence of bacterially expressed A β ₄₂-GFP (FIG. 12G). In accordance with this, a total of ten distinct cyclic tetrapeptide sequences belonging to the TXXR motif were identified among the selected peptide pool (Table 3). Taken together, our results indicate that the minimal bioactive entity

against A β aggregation among this peptide family is a TXXR cyclic tetrapeptide, albeit with significantly reduced efficiency compared to the more privileged cyclic pentapeptide scaffold.

In terms of the selected cyclic hexapeptides, sequences with an initial Threonine (T) and an ultimate Aspartic acid (D) were highly dominant among the selected pool (FIG. 12H; Table 4). As in the case of the selected A β -targeting pentapeptides, charged amino acids were strongly disfavored among the selected sequences, with the exception of the dominant ultimate D residue. It is striking that aromatic amino acids were completely (or almost completely) absent at positions 2 and 3 of the selected hexapeptides, but highly dominant at positions 4 and 5. This sequence analysis revealed the motif cyclo-T(P,L)(V,A)WFD as the most bioactive hexapeptide motif against A β in the investigated macrocycle library.

Example 13

Materials

Synthetic human amyloid peptides A β ₄₀ and A β ₄₂ were purchased from Eurogentec, Belgium (> 95% pure). A β C5-34 and A β C5-116 were synthesized by and purchased from Genscript (USA), while SOD1C5-4 was synthesized and purchased from CPC Scientific (USA). All DNA-processing enzymes were purchased from New England Biolabs (USA) apart from alkaline phosphatase FastAP, which was purchased from ThermoFisher Scientific (USA). Recombinant plasmids were purified using NucleoSpin Plasmid from Macherey-Nagel (Germany) or Plasmid Midi kits from Qiagen (Germany). PCR products and DNA extracted from agarose gels were purified using Nucleospin Gel and PCR Clean-up kits from Macherey-Nagel (Germany), respectively. All chemicals were purchased from Sigma-Aldrich (USA), unless otherwise stated. Isopropyl- β -D-thiogalactoside (IPTG) was purchased from MP Biomedicals (Germany). Stock solutions of the synthetic cyclic peptides were as follows: 32.5 mM in water for A β C5-34, 10 mM in 40% DMSO for A β C5-116 and 30 mM in 40% DMSO for SOD1C5-4.

Cyclic oligopeptide library construction and initial characterization

Initially, nine distinct combinatorial cyclic peptide sub-libraries were constructed: the cyclo-CysX₁X₂X₃, cyclo-SerX₁X₂X₃, and cyclo-ThrX₁X₂X₃ tetrapeptide sub-libraries (pSICLOPPS-CysX₁X₂X₃, pSICLOPPS-SerX₁X₂X₃, and pSICLOPPS-ThrX₁X₂X₃ vector sub-libraries), the cyclo-CysX₁X₂X₃X₄, cyclo-SerX₁X₂X₃X₄, and cyclo-ThrX₁X₂X₃X₄ cyclic pentapeptide sub-libraries (pSICLOPPS-CysX₁X₂X₃X₄, pSICLOPPS-SerX₁X₂X₃X₄, and pSICLOPPS-ThrX₁X₂X₃X₄ vector sub-libraries) and the cyclo-CysX₁X₂X₃X₄X₅, cyclo-SerX₁X₂X₃X₄X₅, and cyclo-ThrX₁X₂X₃X₄X₅ cyclic hexapeptide sub-libraries (pSICLOPPS-

CysX₁X₂X₃X₄X₅, pSICLOPPS-SerX₁X₂X₃X₄X₅, and pSICLOPPS-ThrX₁X₂X₃X₄X₅ vector sub-libraries). These vectors express libraries of fusion proteins comprising four parts: (i) the C-terminal domain of the split Ssp DnaE intein (I_C), (ii) a tetra-, penta-, or hexapeptide sequence, (iii) the N-terminal domain of the split Ssp DnaE intein (I_N), and (iv) a chitin-binding domain (CBD) under the control of the P_{BAD} promoter and its inducer L(+)-arabinose (FIG. 1A). The libraries of genes encoding these combinatorial libraries of random cyclic oligopeptides were constructed using degenerate primers. Cys, Ser, and Thr were encoded in these primers by the codons UGC, AGC, and ACC, respectively, which are the most frequently utilized ones for these amino acids in *E. coli*, while the randomized amino acids (X) were encoded using random NNS codons, where N=A, T, G, or C and S=G or C. A second PCR reaction was conducted in each case to eliminate mismatches. The resulting PCR products were digested with BglII and HindIII for 5 h and inserted into the similarly digested and dephosphorylated auxiliary vector pSICLOPPSKanR (see below). The ligation reactions were optimised at a 12:1 insert:vector molar ratio and performed for 4 h at 16 °C. Approximately 0.35, 0.7 and 3.5 µg of the pSICLOPPSKanR vector were used for each one of the tetra-, penta- and hexapeptide libraries, respectively. The ligated DNA was then purified using spin columns (Macherey-Nagel, Germany), transformed into electro-competent MC1061 cells prepared in-house, plated onto LB agar plates containing 25 µg/mL chloramphenicol and incubated at 37 °C for 14-16 h. This procedure resulted in the construction of the combined pSICLOPPS-NuX₁X₂X₃-X₅ library with a total diversity of about 31,240,000 independent transformants, as judged by plating experiments after serial dilutions.

Colony PCR of 124 randomly selected clones with intein-specific primers revealed that 88 of them (~71%) contained the correct insert. Overexpression of the tetra-partite fusion in 150 randomly selected clones using 0.002% arabinose and monitoring of the production of this fusion protein by western blotting using a mouse anti-CBD primary antibody (New England Biolabs, USA; 1:100,000 dilution) and a goat anti-mouse HRP-conjugated secondary antibody (Bio-Rad, USA; 1:4,000 dilution), showed that 99 of them (~66%) produced high yields of the tetra-partite fusion protein. Among these 99 clones that produced precursor fusion protein (molecular mass ~25 kDa), 82 clones (~55% of total clones tested) also yielded a lower molecular weight band (molecular mass ~20 kDa), which corresponds to one of the splicing reaction products, the N-terminal domain of the Ssp DnaE intein fused to CBD (I_N-CBD), after intein splicing and cyclic peptide formation takes place. Therefore, according to these results, the generated bacterial libraries encoding for cyclic tetra-, penta- and hexapeptide contain

approximately 20,760,000 clones, which express tetra-partite peptide fusions at high levels and which are capable of undergoing splicing and potentially yielding cyclic peptide products. This diversity covers fully the theoretical diversity of our combined cyclo-NuX₁X₂X₃, NuX₁X₂X₃X₄ and NuX₁X₂X₃X₄X₅ libraries ($3 \times 20^3 + 3 \times 20^4 + 3 \times 20^5 = 10,104,000$) by more than two-fold (FIG. 1B).

Expression vector construction

For the construction of pETSOD1-GFP, the human *SOD1* cDNA was generated by PCR-mediated gene assembly. The assembled gene was further amplified by PCR and the resulting product was digested with NdeI and BamHI, and inserted into similarly digested pA β ₄₂-GFP vector GFP (Wurth C, Guimard NK, Hecht MH., J Mol Biol. 2002; 319(5):1279-90), in the place of A β ₄₂. For pETSOD1(A4V)-GFP, *SOD1* was amplified by PCR from the pETSOD1-GFP vector using the mutagenic forward primer GS059 and the reverse primer GS060. The resulting PCR product was then digested with NdeI and BamHI, and inserted into similarly digested pETA β ₄₂-GFP. For pETSOD1(G37R)-GFP, pETSOD1(G85R)-GFP and pETSOD1(G93A)-GFP construction, *SOD1* was mutated by overlap extension PCR starting from pETSOD1-GFP as a template. All *SOD1* PCR products were then digested with NdeI and BamHI, and inserted into similarly digested pETA β ₄₂-GFP vector. For the construction of pETSOD1, pETSOD1(G37R), pETSOD1(G85R) and pETSOD1(G93A), the corresponding *SOD1* genes were amplified by PCR from pETSOD1-GFP, pETSOD1(G37R)-GFP, pETSOD1(G85R)-GFP and pETSOD1(G93A)-GFP, respectively. For the construction of pETSOD1(A4V), *SOD1* was amplified from pETSOD1(A4V)-GFP. All *SOD1* PCR products were digested with XbaI and BamHI, and cloned into similarly digested pET28a(+) (Novagen).

For the construction of the pSICLOPPS vectors encoding for variants of the selected A β C5-34 and A β C5-116 peptides, the auxiliary pSICLOPPSKanR vector was generated initially. pSICLOPPSKanR was constructed by PCR amplification of the gene encoding aminoglycoside 3'-phosphotransferase (KanR - the enzyme conferring resistance to the antibiotic kanamycin) from pET28a(+), digestion with BglII and HindIII and insertion into similarly digested pSICLOPPS. For the construction of the vectors pSICLOPPS-A β C5-34(S1C), pSICLOPPS-A β C5-34(S1T), pSICLOPPS-A β C5-34(S3A), pSICLOPPS-A β C5-34(P4A) and pSICLOPPS-A β C5-34(T5A), mutagenic PCR was carried out starting from pSICLOPPS-A β C5-34, followed by digestion of the generated product with BglII and HindIII and insertion into similarly digested pSICLOPPSKanR. The vectors pSICLOPPS-A β C5-

116(T1C), pSICLOPPS-A β C5-116(T1S), pSICLOPPS-A β C5-116(F3A), pSICLOPPS-A β C5-116(D4A), pSICLOPPS-A β C5-116(R5A), pSICLOPPS-A β C5-116(A2F), pSICLOPPS-A β C5-116(A2S), pSICLOPPS-A β C5-116(A2P), pSICLOPPS-A β C5-116(A2T), pSICLOPPS-A β C5-116(A2Y), pSICLOPPS-A β C5-116(A2H), pSICLOPPS-A β C5-116(A2K),
 5 pSICLOPPS-A β C5-116(A2E), pSICLOPPS-A β C5-116(A2W), pSICLOPPS-A β C5-116(A2R), pSICLOPPS-A β C5-116(A2del), pSICLOPPS-A β C5-116(F3del) and pSICLOPPS-A β C5-116(D4del) were generated in a similar fashion.

Cyclic oligopeptide library screening

Electrocompetent *E. coli* BL21(DE3) cells (Novagen, USA) carrying either the
 10 expression vector pETSOD1(A4V)-GFP, which produces SOD1(A4V)-GFP under control of the strong bacteriophage T7 promoter, or pETA β ₄₂-GFP, which produces A β ₄₂-GFP under control of the T7 promoter, were co-transformed with the combined pSICLOPPS-NuX₁X₂X₃-X₅ vector library. Approximately 10⁸ transformants carrying both the vector library and either pETSOD1(A4V)-GFP or pETA β ₄₂-GFP vectors were harvested, pooled together, and grown
 15 in Luria-Bertani (LB) liquid medium containing either 0.005% (pETSOD1(A4F)-GFP) or 0.002% (pETA β ₄₂-GFP) L-arabinose - the inducer of cyclic peptide production - at 37 °C with shaking. When the optical density at 600 nm (OD₆₀₀) of the bacterial culture was about 0.5, 0.01 (pETSOD1(A4F)-GFP) or 0.1 (pETA β ₄₂-GFP) mM isopropyl- β -D-thiogalactoside (IPTG) was added to the medium to induce overexpression of the reporter. After about two
 20 hours at 37 °C, ~10⁸ cells were screened and the population exhibiting the top 1-3% fluorescence was isolated using FACS (BD FACSAria, BD Biosciences, USA). The isolated cells were re-grown and screened for additional rounds in an identical manner until the desired enrichment in high-fluorescence clones was achieved.

Protein/cyclic peptide production in liquid cultures

25 *E. coli* cells freshly transformed with the appropriate expression vector(s) were used for protein production experiments in all cases. Single bacterial colonies were used to inoculate overnight liquid LB cultures containing the appropriate antibiotics for plasmid maintenance (100 μ g/mL ampicillin, 40 μ g/mL chloramphenicol (Sigma, USA)) at 37 °C. These cultures were used with a 1:100 dilution to inoculate fresh LB cultures in all cases.

30 For SOD1 or SOD1-GFP production, BL21(DE3) (Novagen, USA) or Origami 2(DE3) cells (Novagen, USA) were transformed with the corresponding SOD1- or SOD1-GFP-

encoding vector, either with the appropriate pSICLOPPS vector or alone. Cells were grown in 5 mL liquid LB cultures containing 50 µg/mL kanamycin (or 100 µg/mL ampicillin for pASK75-based vectors), 40 µg/mL chloramphenicol (for cell cultures carrying also a pSICLOPPS vector), 200 µM CuCl₂, 200 µM ZnCl₂ and 0.005% arabinose (for cell cultures carrying also a pSICLOPPS vector) at 37 °C to an OD₆₀₀ of ~0.3-0.5 with shaking, at which point SOD1 or SOD1-GFP production was induced by the addition of 0.01 mM IPTG (0.2 µg/mL anhydrotetracycline (aTc) for pASK-based vectors) for 2-3 h.

For Aβ₄₂-GFP production, BL21(DE3) cells were transformed with pETAβ₄₂-GFP and the appropriate pSICLOPPS vector. Cells were grown in 5 mL liquid LB cultures containing 50 µg/mL kanamycin, 40 µg/mL chloramphenicol and 0.02% arabinose at 37 °C to an OD₆₀₀ of ~0.3-0.5 with shaking, at which point Aβ₄₂-GFP production was induced by the addition of 0.1 mM IPTG for 2-3 h.

Bacterial cell fluorescence

Bacterial cells corresponding to 1 mL culture with OD₆₀₀=1 were harvested by centrifugation and re-suspended in 100 µL phosphate-buffered saline (PBS), transferred to a 96-well FLUOTRAC 200 plate (Greiner Bio One International, Austria), and their fluorescence was measured using a TECAN Safire II-Basic plate reader (Tecan, Austria). Excitation was set at 488 nm and emission was measured at 510 nm.

High-throughput sequencing analysis

For the characterization of the initial libraries, a combined pSICLOPPS-NuX₁X₂X₃-X₅ vector library was prepared containing approximately equal amounts of each one of the tetra-, penta- and hexapeptide sub-libraries. These samples were digested with NcoI and BsrGI and the resulting ~250 bp product that contained the variable peptide-encoding region was isolated. High-throughput sequencing analysis was performed using an Ion Torrent high-throughput sequencing platform. From the obtained data, all the sequences with mismatches outside of the variable peptide-encoding region were removed, and only the 12-, 15- or 18-bp-long peptide-encoding sequences were subjected to further analysis. The libraries of the selected cyclic peptides that enhance either SOD1(A4V)-GFP or Aβ₄₂-GFP fluorescence were sequenced in a similar manner, with the only exception being that all sequences including stop codons were discarded from subsequent analysis.

Protein electrophoresis and western blot analysis

Bacterial cells corresponding to 1 mL culture with $OD_{600}=1$ were harvested by centrifugation and re-suspended in 200 μ L PBS. Samples were lysed by brief sonication for 10 s on ice twice. These lysates (total lysate fraction) were then centrifuged at 13,000 x g for 10 min, the supernatant was collected (soluble fraction) and the pellet was re-suspended in 200 μ L PBS (insoluble fraction). For analysis by SDS-PAGE, samples were boiled for 5 min and 10 μ L of each sample were loaded onto 12% or 15% gels. For analysis by native PAGE, 10-20 μ L of each sample were loaded onto SDS-free 10% gels without prior boiling. In-gel fluorescence was analyzed on a ChemiDoc-It² Imaging System equipped with a CCD camera and a GFP filter (UVP, UK), after exposure for 3-5 sec. For western blotting, proteins were transferred to polyvinylidene fluoride (PVDF) membranes (Merck, Germany) for 50 min at 12 V on a semi-dry blotter (Thermo Fisher, USA). Membranes were blocked with 5% non-fat dry milk in Tris-buffered saline containing 0.1% Tween-20 (TBST) for 1 h at room temperature. After washing with TBST three times, membranes were incubated with the appropriate antibody dilution in TBST containing 0.5% non-fat dried milk at room temperature for 1 h. The utilized antibodies are described in SI Materials and Methods. The proteins were visualized using a ChemiDoc-It² Imaging System (UVP, UK). The utilized antibodies were a mouse monoclonal, horseradish peroxidase (HRP)-conjugated anti-polyhistidine antibody (Sigma, USA) at 1:2,500 dilution, a mouse monoclonal anti-FLAG (Sigma, USA) at 1:1,000 dilution, a mouse anti-GFP at 1:20,000 dilution (Clontech, USA), a mouse anti-A β (6E10) (Covance, USA) at 1:2,000 dilution, a mouse anti-CBD (New England Biolabs, USA) at 1:25,000 or 1:100,000 dilution, and a HRP-conjugated goat anti-mouse antibody (Bio-Rad, USA) at 1:4,000.

Preparation of SOD1 stocks and solutions

SOD1 or mutants thereof were overexpressed from the appropriate pET-SOD1 or pASK-SOD1 vectors in *E. coli* Origami 2(DE3) cells in LB medium containing 50 μ g/mL kanamycin (for pET-SOD1) or 100 μ g/mL ampicillin (for pASK-SOD1), 200 μ M CuCl₂, and 200 μ M ZnCl₂ by the addition of 0.01 mM IPTG (for pET-SOD1) or 0.2 μ g/mL anhydrotetracycline (aTc) (for pASK-SOD1), either at 37 °C for 2-3 h or at 18 °C for about 16 h. Origami 2(DE3) cells were utilized in order to provide an oxidizing cytoplasmic environment in order to promote correct formation of disulfide bonds, which are required for proper SOD1 folding and function. Under these conditions, bacterially produced SOD1 is produced in dimeric and enzymatically active form, while it simultaneously co-exists with misfolded, soluble and insoluble SOD1 oligomeric/aggregated species (FIG. 3C). Thus, the

acquired protein is found in a state that resembles the conditions encountered in human cells under stressful or pathogenic conditions. The appearance of misfolded SOD1 oligomers/aggregates is enhanced with increasing incubation temperatures. Thus, for assays that are more appropriate for monitoring the early steps of SOD1 oligomerization/aggregation, such as dynamic light scattering (DLS), we utilized SOD1 produced at 18 °C, whereas for assays that are more appropriate for monitoring the later steps of SOD1 aggregation, such as filter retardation, ThT staining and CD spectroscopy, we utilized SOD1 produced at 37 °C.

Preparation of A β stocks and solutions

Synthetic A β ₄₀ and A β ₄₂ peptides were gently dissolved without vortexing in doubly deionized water to a final concentration of 100 μ M. These solutions were then diluted by PBS addition (10 mM, pH 7.33) to achieve a final A β concentration of 50 μ M.

Circular dichroism

Appropriate amounts of synthetic cyclic peptides were added to either 40 μ M SOD1(A4V) or 50 μ M A β solutions at the desired cyclic peptide:target protein molar ratio. SOD1(A4V) structural changes were monitored for 90 d at 25 °C, under quiescent conditions. A β structural changes were monitored for 30 d at 33 °C under quiescent conditions. CD spectra in the range 190-260 nm were recorded on a JASCO J-715 spectropolarimeter (Jasco Co., Japan) using quartz cuvettes with 1 mm path length. Each reported spectrum is the average of three scans at a rate of 100 nm·min⁻¹ and a resolution of 0.5 nm.

Dynamic light scattering

The sizes of the SOD1 particles were measured using a Zetasizer NanoZS90 (Malvern) instrument. After a 2-min temperature-equilibration step at 37 °C, eighteen consecutive 10-s measurements, per sample, were averaged to produce the particle size (Z average) distributions.

Thioflavin T staining

40 μ M SOD1(A4V) solutions, aged for 90 d at 25 °C, with or without the selected synthetic peptides, were diluted to 10 μ M with PBS. 5 μ L from a stock solution of ThT (Sigma-Aldrich, USA) in PBS (10 mM, pH 7.33) was added to these SOD1(A4V) solutions to achieve a final ThT concentration of 10 μ M. The mixture was agitated adequately by pipetting and immediately thereafter, fluorescence was monitored with excitation at 440 nm (EM slit = 2.5 nm, PMT Voltage 700 V, response 0.4 s) using a HITACHI F-2500 (Japan) spectrofluorometer.

For A β ThT staining, 100 μ L of the 30-d aged 50 μ M CD solutions were diluted in PBS (10 mM, pH 7.33) to form a 25 μ M A β solution with 200 μ L final volume. 2.5 μ L from a stock solution of ThT (Sigma-Aldrich, USA) in PBS (10 mM, pH 7.33) was added to the prepared A β solutions to achieve a final ThT concentration of 5 μ M. The mixture was agitated
5 adequately by pipetting and immediately thereafter, fluorescence was monitored with excitation at 440 nm (EM slit = 2.5 nm, PMT Voltage 700 V, response 0.4 s) using a HITACHI F-2500 (Japan) spectrofluorometer.

Filter retardation assay

SOD1(A4V) solutions (10 μ M), incubated in the presence or absence of the selected
10 cyclic peptides for 25 d at 37 °C, were mixed with a stock solution of SDS to achieve a final SDS concentration of 2% and then boiled for 10 min. These samples were subsequently applied under vacuum on a 0.2 μ m-pore size PVDF membrane (Merck), which had been previously equilibrated with transfer buffer containing 0.1% SDS, and then washed twice with 100 μ L TBS under vacuum. The membrane was blocked with 5% non-fat dry milk in TBST for 1 h at room
15 temperature and then stained with a HRP-conjugated anti-polyHis antibody at a 1:2,500 dilution (Sigma-Aldrich) overnight at 4 °C.

SOD1 aggregation and viability measurements in HEK293 cells

Human embryonic kidney (HEK) 293 cells were transfected using a Nucleofector (Amaxa) following the manufacturer's protocol. 6 μ g DNA (SOD1 or SOD1(A4V) cloned into
20 the pEGFP-N3 plasmid vector) were used per 2×10^6 cells and 5 μ M synthetic SOD1C5-4 was added, where appropriate, before plating. Transfected cells were sorted 18 h later on a FACSaria to isolate GFP-positive clones. 4',6-Diamidino-2-phenylindole dihydrochloride (DAPI) dye was used to exclude dead cells. ~28% of the SOD1 and ~15% of the SOD1(A4V) total cells were found to be GFP-positive. Collected cells were plated onto a 24-well plate at a
25 density of 50,000 cells/well. Microscopy analysis was performed under an inverted microscope on day 1 and day 5 in culture after sorting. Cell counts are the average number of viable GFP-fluorescing cells of two areas per triplicate of wells of 24-well plates (magnification 20x). Cell counts are presented as percentage of viability of SOD1-overexpressing cell. As aggregate-positive cells are counted the fluorescing inclusion body-positive cells. Again, two areas per
30 triplicate of wells of 24-well plate are averaged (magnification 20x). Aggregate-positive cells are presented as percentage of the total viable GFP-fluorescing cells.

Transmission Electron Microscopy (TEM)

For TEM analysis, the 30-d aged 50 μ M CD solutions of A β 42 (with 100 μ M of the selected peptides or without) was mixed well by pipetting. 2 μ L of this solution were placed in a carbon-coated film on 200-mesh copper grids (Agar Scientific, UK) for 5 min. After adsorption, grids were washed in deionized water and negatively stained by applying a 2- μ L drop of freshly prepared 1% (w/v) uranyl acetate (Sigma-Aldrich, USA) in Milli-Q water for 5 min. Excess fluid was blotted off, and grids were washed in deionized water and dried in air. Images were recorded using a FEI CM20 electron microscope (FEI, USA) with a Gatan GIF200 imaging filter (Gatan, USA), equipped with a Peltier-cooled slow-scan CCD camera.

10 Neuronal cell cultures

The media/agents for primary neuronal cell cultures were purchased from Thermo Fisher Scientific (USA). Hippocampal neuronal cultures were obtained from postnatal day 1 female pups of C57BL/6 mice. Briefly, after being dissected, the hippocampus was incubated with 0.25% trypsin for 15 min at 37 °C. The hippocampi were then rinsed in 10 mL of Hibernate containing 10% (v/v) heat-inactivated fetal bovine serum (FBS). Cultures were maintained in Neurobasal-A medium containing 2% B-27 supplement, 0.5 mM Gluta-MAX and 1% penicillin/streptomycin at 37 °C and 5% CO₂. Half of the medium was replaced twice a week. Neuronal hippocampal cells were plated at a density of approximately 2×10^4 per well in 96-well plates and 5×10^5 per well in 24-well plates for MTT and induced cell death assays, respectively. After seven days of incubation in culture well plates, the primary hippocampal neurons were used for the cell viability measurements.

The utilized U87MG cells (human glioblastoma-astrocytoma, epithelial-like cell line) were kind a gift from Dr. Maria Paravatou-Petsotas, Radiobiology Laboratory, Institute of Nuclear & Radiological Sciences & Technology, Energy & Safety, NCSR “Demokritos”. The utilized media/agents for U87MG cell cultures were obtained from Biochrom AG (Germany) and PAA Laboratories (USA). U87MG cells were grown in Dulbecco's modified Eagle medium (DMEM), supplemented with 10% fetal bovine serum (FBS), 2.5 mM L-glutamine, 1% penicillin/streptomycin at 37 °C and 5% CO₂. For MTT cytotoxicity studies, cells were plated at a density of 2×10^4 cells per well in 96-well plates and incubated at 37 °C for 24 h to allow cells to attach. The medium was subsequently removed and *cells were rendered quiescent* by incubation in serum-free medium for 24 h. For cell viability measurements cells were

subsequently treated with the indicated concentrations of A β in the presence or absence of synthetic peptides, as described in Materials and Methods.

Cell viability measurements

Solutions of synthetic A β ₄₀ or A β ₄₂ (10 μ M) in PBS, preincubated at 37 °C (3 d for A β ₄₀ solutions and 1 d for A β ₄₂ solutions) in the presence or absence of synthetic cyclic peptides (1:1 and 2:1 ratio of peptides:A β), were diluted with fresh medium and transferred into wells at a 1 μ M final A β concentration. Cell viability was determined using the MTT assay. MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl tetrazolium bromide) was purchased from AppliChem (Germany). After 24 h of exposure to A β solutions, 100 μ L of a 0.5 mg/mL stock solution of MTT in Neurobasal-A was added to each well of primary hippocampal neurons followed by a 3 h incubation at 37 °C, while 100 μ L of a 1 mg/mL stock solution of MTT in DMEM complete medium was added to each well of U87MG cells followed by a 4 h incubation at 37 °C. The medium was then removed and the cells were diluted in DMSO. The relative formazan concentration was measured by determination of the absorbance at 540 nm using a plate reader (Tecan, Austria). Results were expressed as the percentage of MTT reduction, assuming that the absorbance of control (untreated) cells was 100%, and represent the mean of three independent experiments with six replicate wells for each condition. Induced cell death was also qualitatively examined by phase-contrast microscopy (Carl Zeiss, Axiovert 25 CFL, Germany) using the above solutions. In each run, the effect of solutions of plain synthetic peptides and plain A β ₄₀ or A β ₄₂ was independently checked to serve as internal control.

In vivo assays in *C. elegans*

Strains

We followed standard procedures for *C. elegans* strains maintenance at 16 °C. The following strains were used: CL2179: dvIs179 [myo-3p::GFP::3'UTR(long)+rol-6(su1006)] (available on the world wide web at: cgc.cbs.umn.edu/strain.php?id=26134); CL2331: dvIs37 [myo-3p::GFP::A β (3-42)+rol-6(su1006)] (available on the world wide web at: cgc.cbs.umn.edu/strain.php?id=26135); CL4176: smg-1(cc546) I; dvIs27 [myo-3::A β (1-42)-let 3'UTR(pAF29); pRF4 (rol-6(su1006)] (available on the world wide web at: cgc.cbs.umn.edu/strain.php?id=7663).

Treatment with cyclic peptides

For treatments with synthetic cyclic peptides, nematodes were exposed to the indicated A β C5-34 and A β C5-116 concentrations per NGM plate. Stock solutions of the two chemically synthesized pure peptides were obtained after dissolution in DMSO and stored at -20 °C. The appropriate amount of compound or DMSO (control cultures) was added onto an *E. coli* OP50 bacterial lawn. Synchronized offspring were randomly distributed to treatment plates to avoid systematic differences in egg lay batches. Treatment and control plates were handled, scored and assayed in parallel.

Paralysis assay

Synchronized CL4176 animals (150-300 animals per condition) were transferred to NGM plates containing synthetic A β C5-34, A β C5-116 or 0.26% DMSO at 16 °C for 48 h before transgene induction via temperature up-shift to 25 °C. Synchronized offspring were randomly distributed to treatment plates to avoid systematic differences in egg lay batches. Treatment and control plates were handled, scored and assayed in parallel. Scoring of paralyzed animals was initiated 24 h after temperature up-shift for the CL4176 strain. Nematodes were scored as paralyzed upon failure to move their half end-body upon prodding. Animals that died were excluded. Plates were indexed as 1, 2, 3 etc by an independent person and were given to the observer for scoring in random order. The index was revealed only after scoring.

Dot blot Analysis

CL4176 animals were allowed to lay eggs for 3 h on NGM plates containing either synthetic peptides or 0.26% DMSO. Paralysis was induced upon temperature up-shift and the progeny were exposed to either pure peptides or 0.26% DMSO until 50% of the control population was paralyzed. The animals were then collected and boiled in non-reducing Laemmli buffer. For dot blot analysis, 1–5 μ g of protein lysates were spotted onto 0.2 μ m nitrocellulose membranes (Bio-Rad, USA) after soaking into TBS pre-heated at 80 °C. Immunoblotting was performed using the anti-A β antibody 6E10 (recognizes total A β) and the anti-amyloid protein, oligomer-specific antibody AB9234 (Merck Millipore, Germany). Actin was used as a loading control. Blots were developed with chemiluminescence by using the Clarity™ Western ECL substrate (Bio-Rad, USA). Quantification of the ratio of each detected protein to actin using the anti-actin antibody sc-1615 (Santa Cruz, Germany), and normalization to control appears next to each representative blot.

Confocal microscopy analysis

For A β ₃₋₄₂ deposit measurements, synchronized (at the L4 larval stage) CL2331 and CL2179 (control strain) animals exposed to solvent (0.26% DMSO), 10 μ M A β C5-34 or 5 μ M A β C5-116 and grown at 20 °C (to induce aggregation) until day 2 of adulthood were collected. Animals were mounted onto 2% agarose pads on glass slides, anesthetized with 10 mM
5 levamisole and observed at RT using a Leica TCS SPE confocal laser scanning microscope (Leica Lasertechnik GmbH, Germany). The LAS AF software was used for image acquisition. At least twenty animals/condition in three independent experiments were processed. Images of whole worms and focused images in the posterior area of nematodes were acquired with 10 $\times$ 0.45 and 20 $\times$ 0.70 numerical aperture, respectively.

10

While the invention has been described with respect to specific embodiments, it is apparent that modifications are possible without departing from the scope of the invention

WHAT IS CLAIMED IS:

1. A method of identifying modulators of a misfolding-prone protein associated with a protein misfolding disease, comprising (A) a transformed bacterial cell or transformed bacterial cells that co-expresses/co-express (a) a nucleic acid encoding the recombinant production of peptide macrocycle or a library of peptide macrocycles, operably linked to a promoter followed by expressing the nucleic acid of step (a), thereby producing the recombinant peptide macrocycle or library of peptide macrocycles in the transformed cell, and (b) a bipartite nucleic acid comprising a sequence for a gene encoding a misfolding-prone protein associated with a protein misfolding disease (MisP) and a sequence encoding a fluorescent protein (FP) reporter, operably linked to a promoter followed by expressing the nucleic acid of step (b), thereby producing the recombinant MisP-FP polypeptide fusion in the transformed cell; (B) identification and selection of the bacterial cells of step (A) that exhibit enhanced levels of MisP-FP fluorescence compared to the bacterial cells not containing the nucleic acid of step (Aa); and (C) identification of the bioactive peptide macrocycle(s) that modulate(s) MisP misfolding by determining the nucleotide sequence of step (Aa) that encodes the peptide macrocycle in the selected bacterial cell of step (B).
2. A method of identifying a peptide macrocycle modulating the misfolding of misfolding-prone proteins associated with a protein misfolding disease (MisP) according to claim 1, wherein said MisP is selected from β -amyloid peptide, SOD1, tau, α -synuclein, polyglutaminated huntingtin, polyglutaminated ataxin-1, polyglutaminated ataxin-2, polyglutaminated ataxin-3, prion protein, islet amyloid polypeptide (amylin), β 2-microglobulin, fragments of immunoglobulin light chain, fragments of immunoglobulin heavy chain, serum amyloid A, ABri peptide, ADan peptide, transthyretin, apolipoprotein A1, gelsolin, transthyretin, lysozyme, phenylalanine hydroxylase, apolipoprotein A-I, calcitonin, prolactin, TDP-43, FUS/TLS; insulin, hemoglobin, α 1-antitrypsin, p53; or variants thereof.
3. The peptide macrocycle of claim 1 or 2, to inhibit protein misfolding and aggregation, wherein said peptide macrocycle can be a ribosomally synthesized head-to-tail cyclic peptide, side-chain-to-tail cyclic peptide, bicyclic peptide, lanthipeptide, linaridin, proteusin, cyanobactin, thiopeptide, bottromycin, microcin, lasso peptide, microviridin, amatoxin, phalloxin, θ -defensin, orbitide, or cyclotide.
4. Use of the peptide macrocycle of claims 1 to 3 to inhibit protein misfolding and aggregation associated with a protein misfolding disease, wherein the disease is selected from amyotrophic lateral sclerosis, Alzheimer's disease, amyotrophic lateral sclerosis, Parkinson's

disease, Huntington's disease, Creutzfeldt-Jakob disease, cancer, phenylketonuria, type 2 diabetes, senile systemic amyloidosis, familial amyloidotic polyneuropathy, familial amyloid cardiomyopathy, leptomeningeal amyloidosis, systemic amyloidosis, familial British dementia, familial Danish dementia, light chain amyloidosis, heavy chain amyloidosis, serum amyloid A amyloidosis, lysozyme amyloidosis, dialysis-related amyloidosis, ApoAI amyloidosis, Finnish type familial amyloidosis, hereditary cerebral hemorrhage with amyloidosis (Icelandic type), medullary carcinoma of the thyroid, pituitary prolactinoma, injection-localized amyloidosis, frontotemporal dementia, spinocerebellar ataxia 1, spinocerebellar ataxia 2, spinocerebellar ataxia 3, α 1-antitrypsin deficiency, sickle-cell anemia, and transmissible spongiform encephalopathy.

5. A method of treatment, prevention or diagnosis of a protein misfolding disease comprising administering to a subject a therapeutically effective amount of a peptide macrocycle according to any one of the claims 1 to 3.

6. A pharmaceutical composition comprising a peptide macrocycle according to claims 1 to 3 and a pharmaceutically acceptable carrier.

7. A pharmaceutical composition according to claim 6 used for the treatment or prevention of a protein misfolding disease.

8. A hybrid molecule comprising: a) a peptide macrocycle identified or produced according to the method of claims 1 to 3, and b) a scaffold molecule linked to the peptide macrocycle; which can be selected from a diagnostic or a therapeutic reagent.

9. The hybrid molecule of claim 8, wherein the therapeutic reagent is a cytoprotective agent that renders the aggregates of the target protein less toxic or inhibits target protein aggregate formation.

10. The hybrid molecule according to any of claims 8 and 9, wherein the scaffold molecule comprises all or a sufficient portion of a protein selected from the group consisting of antibodies, enzymes, chromogenic proteins, fluorescent proteins and fragments thereof.

11. The hybrid molecule according to any of the claims 8 to 10, wherein the diagnostic reagent specifically target protein aggregates in diseased or healthy tissue.

12. A method of treating or diagnosing a protein misfolding disease associated with aberrant aggregate formation, the method comprising administering a hybrid molecule according to claim 8, wherein the peptide macrocycle of the hybrid molecule specifically interacts with the amyloid or non-amyloid form of the target MisP.

13. A peptide to inhibit protein misfolding and aggregation wherein the peptide comprises the amino acid sequence NuX₁X₂...X_N, wherein X is any one of the twenty natural amino acids,

N=3-5 and Nu=C, S or T, and wherein said peptide prevents misfolding and aggregation of SOD1 and/or mutant SOD1.

14. A peptide to inhibit protein misfolding and aggregation according to claim 13, wherein the peptide has a cyclic structure.

15. A peptide to inhibit protein misfolding and aggregation according to claim 13, wherein Nu=T.

16. A peptide to inhibit protein misfolding and aggregation according to claim 13, wherein N=4, X₂=S, and X₄=W.

17. A peptide to inhibit protein misfolding and aggregation according to claim 13, wherein the X₁ is any amino acid excluding I, N, Q, M, E, H, and K, preferably it is A, L, V, F, W, Y, C, S, T, D, R, P or G, and more preferably it is S, A, F, W.

18. A peptide to inhibit protein misfolding and aggregation according to claim 13, wherein the X₃ is any amino acid excluding I, N, Q, C, D, E, K and P, preferably it is A, L, V, F, W, Y, M, S, T, R, H or G, and more preferably it is V, F, W, M, or H.

19. Use of a peptide to inhibit protein misfolding and aggregation according to claim 13, wherein the X₃ is preferably W.

20. Use of a peptide to inhibit protein misfolding and aggregation according to claims 13 to 19, wherein the peptide is preferably selected from any one of the amino acid sequences set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, up to SEQ ID NO:46, and more preferably from the amino acid sequences TWSVW, TASFW, and TFMSW.

21. Use of a peptide to inhibit protein misfolding and aggregation according to claims 13 and 14; wherein said peptide is a linearized version of the cyclic peptide.

22. A hybrid molecule comprising: a) a peptide set forth in any one of the claims 13 to 21, that specifically interacts with the amyloid or non-amyloid form of SOD1 and/or mutant SOD1; and b) a scaffold molecule comprising a diagnostic or therapeutic reagent.

23. The hybrid molecule of claim 22, wherein the diagnostic or therapeutic reagent comprises a polypeptide, small molecule or compound.

24. The hybrid molecule of claim 22, wherein the scaffold molecule comprises all or a sufficient portion of a protein selected from the group consisting of antibodies, enzymes, chromogenic proteins, fluorescent proteins and fragments thereof.

25. The hybrid molecule of claim 22, wherein the therapeutic agent is a neuroprotective agent that renders SOD1 aggregates less toxic or inhibits SOD1 aggregate formation.

26. The hybrid molecule of claim 22, wherein the diagnostic reagent specifically images SOD1 aggregates in neuronal tissue.

27. A method of treatment, prevention or diagnosis of amyotrophic lateral sclerosis comprising administering to a subject a therapeutically effective amount of a peptide or hybrid molecule according to any one of the claims 13 to 26.
28. A pharmaceutical composition comprising a peptide or hybrid molecule according to claims 13 to 26 and a pharmaceutically acceptable carrier.
29. A pharmaceutical composition according to claim 28 used for the treatment or prevention of amyotrophic lateral sclerosis.
30. An isolated nucleic acid sequence encoding the peptide of claims 13 to 26.
31. A vector comprising the nucleic acid sequence of claim 30.
32. The vector of claim 31, wherein the vector is an expression vector.
33. A host cell comprising the vector of claim 32.
34. The host cell of claim 33, wherein the host cell is a prokaryotic or eukaryotic cell.
35. Use of a peptide to inhibit protein misfolding and aggregation, wherein the peptide comprises the amino acid sequence NuX₁X₂...X_N, wherein X is any one of the twenty natural amino acids, N=3-5 and Nu=C, S or T, and wherein said peptide prevents misfolding and aggregation of β -amyloid peptides.
36. Use of a peptide to inhibit protein misfolding and aggregation according to claim 35, wherein the peptide has a cyclic structure.
37. Use of a peptide to inhibit protein misfolding and aggregation according to any of the preceding claims, wherein Nu=T.
38. Use of a peptide to inhibit protein misfolding and aggregation according to any of the preceding claims, wherein N=3 and X₄=R.
39. Use of a peptide to inhibit protein misfolding and aggregation according to claim 38, wherein the X₁ is selected from T, R, D, L, F or A.
40. Use of a peptide to inhibit protein misfolding and aggregation according to claim 38, wherein the X₂ is selected from C, R, S, G, Q, I, W, D, or F.
41. Use of a peptide to inhibit protein misfolding and aggregation according to any of claims 38 to 40, wherein the peptide is selected from TTCR, TTRR, TTSR, TRGR, TTGR, TRRR, TDQR, TLIR, TLWR, TLGR, TFDR, TAFR (SEQ ID NOs 210-221).
42. Use of a peptide to inhibit protein misfolding and aggregation according to any of claims 35 to 37, wherein N=4.
43. Use of a peptide to inhibit protein misfolding and aggregation according to claim 42, wherein the X₁ is any amino acid excluding F, W, Y, Q, D, E, K and P, preferably it is S, H, T, V or A, and more preferably it is T, V or A.

44. Use of a peptide to inhibit protein misfolding and aggregation according to claim 43, wherein the X_2 is any amino acid excluding E, preferably a non-negatively charged amino acid, and more preferably it is selected from I, L, V, F, W, Y, M, S, T, R, H, or G.
45. Use of a peptide to inhibit protein misfolding and aggregation according to claim 43, wherein the X_3 is any amino acid excluding Q, M and K, is preferably a non-negatively charged amino acid, and is more preferably selected from A, V, F, W, C, S, T, D, C, R, H, P or G.
46. Use of a peptide to inhibit protein misfolding and aggregation according to claim 43, wherein the X_4 is preferably R or T.
47. Use of a peptide to inhibit protein misfolding and aggregation according to any of claims 43 to 46, wherein the peptide is preferably selected from any one of the amino acid sequences set forth in SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, ..., up to SEQ ID NO:209, and more preferably from the amino acid sequences TAFDR, TAWCR, TTWCR, TTVDR, TTYAR, TTTAR, SASPT.
48. Use of a peptide to inhibit protein misfolding and aggregation according to any of claims 35 to 37, wherein $N=5$.
49. Use of a peptide to inhibit protein misfolding and aggregation according to claim 48, wherein the X_1 is selected from I, L, V, C, S, K or P, and is more preferably P, V or L.
50. Use of a peptide to inhibit protein misfolding and aggregation according to claim 48, wherein the X_2 is selected from A, I, L, V, F, W, C, S, T, D, E, H, K, P, or G and is more preferably V or A.
51. Use of a peptide to inhibit protein misfolding and aggregation according to claim 48, wherein the X_3 is selected from I, L, V, F, W, Y, E or R, and is more preferably W.
52. Use of a peptide to inhibit protein misfolding and aggregation according to claim 48, wherein the X_4 is selected from L, V, F, Y, S or R and is more preferably F.
53. Use of a peptide to inhibit protein misfolding and aggregation according to claim 48, wherein the X_5 is selected from W, M, N, D, or E and is more preferably D.
54. Use of a peptide to inhibit protein misfolding and aggregation according to any of claims 48 to 53 wherein the peptide is selected from the amino acid sequences set forth in SEQ ID NO:176, SEQ ID NO:177, ..., up to SEQ ID NO:209, and is most preferably TPVWFD (SEQ ID NO:176) or TPAWFD (SEQ ID NO:177).
55. Use of a peptide to inhibit protein misfolding and aggregation according to claim 36; wherein said peptide is a linearized version of the cyclic peptide.

56. Use of a peptide to inhibit protein misfolding and aggregation according to any of the preceding claims, wherein said peptide prevents misfolding and aggregation of the β -amyloid peptide.
57. A method of treatment, prevention or diagnosis of a disease related to protein misfolding and aggregation, comprising administering to a subject a therapeutically effective amount of a peptide according to any one of the claims 35 to 55, wherein the disease is selected from Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, Creutzfeldt-Jakob disease, type II diabetes, familial amyloidotic polyneuropathy, systemic amyloidosis, and transmissible spongiform encephalopathy.
58. A method of treatment, prevention or diagnosis of Alzheimer's disease comprising administering to a subject a therapeutically effective amount of a peptide according to any one of the claims 35 to 55.
59. A pharmaceutical composition comprising a peptide according to any one of the claims 35 to 55 and a pharmaceutically acceptable carrier.
60. An isolated nucleic acid sequence encoding the peptide of claims 35 to 55.
61. A vector comprising the nucleic acid sequence of claim 60.
62. The vector of claim 61, wherein the vector is an expression vector.
63. A host cell comprising the vector of claim 62.
64. The host cell of claim 63, wherein the host cell is a prokaryotic or eukaryotic cell.
65. A hybrid molecule comprising: a) a peptide set forth in any one of the claims 35 to 55, that specifically interacts with the amyloid form of the A β peptide; and b) a scaffold molecule comprising a diagnostic or therapeutic reagent.
66. The hybrid molecule of claim 65, wherein the diagnostic or therapeutic reagent comprises a polypeptide, small molecule or compound.
67. The hybrid molecule of claim 66, wherein the polypeptide comprises all or a sufficient portion of a protein selected from the group consisting of antibodies, enzymes, chromogenic proteins, fluorescent proteins and fragments thereof.
68. The hybrid molecule of claim 66, wherein the therapeutic agent is a neuroprotective agent that renders amyloid plaques less toxic or inhibits plaque formation.
69. The hybrid molecule of claim 66, wherein the diagnostic reagent specifically images amyloid aggregates in neuronal tissue.
70. A method of treating or diagnosing a neurodegenerative disease associated with aberrant aggregate formation, the method comprising administering a hybrid molecule of claim 66 to a subject having, or predisposed to having, the disease.

71. A method of identifying modulators of a misfolding-prone protein associated with a protein misfolding disease, comprising:

(A) obtaining a population of transformed bacterial cells that co-express:

(a) a nucleic acid encoding a library of peptide macrocycles, operably linked to a promoter

and;

(b) a bipartite nucleic acid comprising a sequence for a gene encoding a misfolding-prone protein associated with a protein misfolding disease (MisP) and a sequence encoding a protein reporter;

(B) identifying bacterial cells of step (A) that exhibit enhanced levels of protein reporter activity; and

(C) identifying of the bioactive peptide macrocycles in the library that modulate MisP misfolding.

72. The method of claim 71, wherein the protein reporter is a fluorescent protein (FP) reporter, and step (B) comprises identifying bacterial cells that exhibit enhanced levels of MisP-FP fluorescence.

73. The method of claim 71, wherein the protein reporter is an enzyme.

74. The method of claim 71, wherein the identification of step (B) comprises selection.

75. The method of claim 71, wherein step (C) comprises sequencing the nucleic acid of step (Aa).

76. The method of claim 71, wherein the nucleic acids of (a) and (b) are encoded on the same vector.

77. The method of claim 71, wherein the vector is a plasmid.

78. The method of claim 71, wherein said MisP is selected from β -amyloid peptide, SOD1, tau, α -synuclein, polyglutaminated huntingtin, polyglutaminated ataxin-1, polyglutaminated ataxin-2, polyglutaminated ataxin-3, prion protein, islet amyloid polypeptide (amylin), β 2-microglobulin, fragments of immunoglobulin light chain, fragments of immunoglobulin heavy chain, serum amyloid A, ABri peptide, ADan peptide, transthyretin, apolipoprotein A1, gelsolin, transthyretin, lysozyme, phenylalanine hydroxylase, apolipoprotein A-I, calcitonin, prolactin, TDP-43, FUS/TLS; insulin, hemoglobin, α 1-antitrypsin, p53 or variants thereof.

79. The method of claim 71, wherein said peptide macrocycle can be a ribosomally synthesized as a head-to-tail cyclic peptide, side-chain-to-tail cyclic peptide, bicyclic peptide, lanthipeptide, linaridin, proteusin, cyanobactin, thiopeptide, bottromycin, microcin, lasso peptide, microviridin, amatoxin, phallotoxin, θ -defensin, orbitide, or cyclotide.

80. The method of claim 71, wherein the disease is selected from amyotrophic lateral sclerosis, Alzheimer's disease, amyotrophic lateral sclerosis, Parkinson's disease, Huntington's disease, Creutzfeldt-Jakob disease, cancer, phenylketonuria, type 2 diabetes, senile systemic amyloidosis, familial amyloidotic polyneuropathy, familial amyloid cardiomyopathy, leptomeningeal amyloidosis, systemic amyloidosis, familial British dementia, familial Danish dementia, light chain amyloidosis, heavy chain amyloidosis, serum amyloid A amyloidosis, lysozyme amyloidosis, dialysis-related amyloidosis, ApoAI amyloidosis, Finnish type familial amyloidosis, hereditary cerebral hemorrhage with amyloidosis (Icelandic type), medullary carcinoma of the thyroid, pituitary prolactinoma, injection-localized amyloidosis, frontotemporal dementia, spinocerebellar ataxia 1, spinocerebellar ataxia 2, spinocerebellar ataxia 3, α 1-antitrypsin deficiency, sickle-cell anemia, or transmissible spongiform encephalopathy.

81. The method of any one of claims 71-80, further comprising recombinantly producing or chemically synthesizing the identified bioactive peptide macrocycle.

82. A method of treatment, prevention or diagnosis of a protein misfolding disease comprising administering to a subject a therapeutically effective amount of the bioactive peptide macrocycle of claim 81 or a bioactive peptide macrocycle identified by any one of the methods of claims 71-80.

83. A pharmaceutical composition comprising the bioactive peptide macrocycle according to claim 81 or a bioactive peptide macrocycle identified by any one of the methods of claims 71-80; and a pharmaceutically acceptable carrier.

84. A pharmaceutical composition according to claim 83 used for the treatment or prevention of a protein misfolding disease.

85. A hybrid molecule comprising: a) a peptide macrocycle identified or produced according to the method of any one of claims 71-81, and b) a scaffold molecule linked to the peptide macrocycle.

86. The molecule of claim 85, wherein the scaffold molecule is a diagnostic or a therapeutic reagent.

87. The molecule of claim 86, wherein the therapeutic reagent is a cytoprotective agent that renders the aggregates of the target protein less toxic or inhibits target protein aggregate formation.

88. The molecule of claim 85, wherein the scaffold molecule comprises all or a sufficient portion of a protein selected from the group consisting of antibodies, enzymes, chromogenic proteins, and fluorescent proteins.

89. The molecule of claim 86, wherein the diagnostic reagent specifically targets protein aggregates in diseased or healthy tissue.
90. A method of treating or diagnosing a protein misfolding disease associated with aberrant aggregate formation, the method comprising administering a hybrid molecule according to claim 85, wherein the peptide is a cyclic peptide, and wherein the peptide macrocycle of the hybrid molecule specifically interacts with the amyloid or non-amyloid form of the target MisP.
91. A peptide comprising the amino acid sequence NuX₁X₂...X_N, wherein: Nu is T; N=4; X₁ is any amino acid excluding I, N, Q, M, E, H, and K; X₂=S; X₃ is any amino acid excluding I, N, Q, C, D, E, K and P; and X₄=W, wherein the specifically interacts with the amyloid or non-amyloid form of SOD1 and/or mutant SOD1.
92. The peptide according to claim 91, wherein the X₁ is A, L, V, F, W, Y, C, S, T, D, R, P or G.
93. The peptide according to claim 92, wherein the X₁ is S, A, F or W.
94. The peptide according to claim 91, wherein the X₃ is A, L, V, F, W, Y, M, S, T, R, H or G.
95. The peptide according to claim 94, wherein the X₃ is V, F, W, M, or H.
96. The peptide of claim 91, wherein: Nu is T; N=4; X₁ is A, L, V, F, W, Y, C, S, T, D, R, P or G; X₂=S; X₃ is A, L, V, F, W, Y, M, S, T, R, H or G; and X₄=W.
97. The peptide of claim 91, wherein: Nu is T; N=4; X₁ is S, A, F or W; X₂=S; X₃ is V, F, W, M, or H; and X₄=W.
98. A peptide comprising the amino acid sequence set forth in any one of SEQ ID NO:1-46, wherein the peptide prevents misfolding and aggregation of SOD1 and/or mutant SOD1.
99. The peptide according to any one of claims 91-98, wherein the peptide comprises an amino acid sequence selected from TWSVW, TASFW, and TFSMW.
100. The peptide according to any one of claims 91-99, wherein said peptide prevents misfolding and aggregation of SOD1 and/or mutant SOD1.
101. The peptide according to any one of claims 91-100, wherein at least one position of the peptide is a D amino acid.
102. The peptide according to any one of claims 91-100, wherein the peptide is in a linearized form.
103. A hybrid molecule comprising: a) a peptide set forth in any one of the claims 91-102, and b) a scaffold molecule.

104. The molecule of claim 103, wherein the scaffolding molecule comprises a cell penetrating peptide.
105. The molecule of claim 103, wherein the scaffold molecule comprises a diagnostic or therapeutic reagent.
106. The hybrid molecule of claim 103, wherein the scaffold molecule comprises a polypeptide, small molecule or compound.
107. The hybrid molecule of claim 103, wherein the scaffold molecule comprises all or a sufficient portion of a protein selected from the group consisting of antibodies, enzymes, chromogenic proteins, fluorescent proteins and fragments thereof.
108. The hybrid molecule of claim 105, wherein the therapeutic agent is a neuroprotective agent that renders SOD1 aggregates less toxic or inhibits SOD1 aggregate formation.
109. The hybrid molecule of claim 105, wherein the diagnostic reagent specifically images SOD1 aggregates in neuronal tissue.
110. Use of a peptide of any one of claims 91-101 or a molecule of any one of claims 103-109, to inhibit protein misfolding and aggregation wherein the peptide prevents misfolding and aggregation of SOD1 and/or mutant SOD1.
111. A method of treatment, prevention or diagnosis of amyotrophic lateral sclerosis comprising administering to a subject a therapeutically effective amount of a peptide or hybrid molecule according to any one of the claims 91 to 109.
112. A pharmaceutical composition comprising a peptide or hybrid molecule according to any one of claims 91 to 109 and a pharmaceutically acceptable carrier.
113. A pharmaceutical composition according to claim 112 used for the treatment or prevention of amyotrophic lateral sclerosis.
114. An isolated nucleic acid sequence encoding the peptide of claims 91 to 109.
115. A vector comprising the nucleic acid sequence of claim 114.
116. The vector of claim 115, wherein the vector is an expression vector.
117. A host cell comprising the vector of claim 116.
118. The host cell of claim 117, wherein the host cell is a prokaryotic or eukaryotic cell.
119. A peptide, wherein the peptide comprises the amino acid sequence NuX₁X₂...X_N, wherein: (A) Nu=T; N=3; X₁ is selected from T, R, D, L, F or A; X₂ is selected from C, R, S, G, Q, I, W, D, or F; and X₃=R;
- (B) Nu=T; N=4; X₁ is any amino acid excluding F, W, Y, Q, D, E, K and P; X₂ is any amino acid excluding E; X₃ is any amino acid excluding Q, M and K; and X₄ is R; or

(C) Nu=T; N=5; X₁ is I, L, V, C, S, K or P; X₂ is any amino acid excluding Y, N, Q, M, and R; X₃ is selected from I, L, V, F, W, Y, E, or R; X₄ is selected from L, V, F, Y, S or R; and X₅ is selected from W, M, N, D, or E,

wherein the peptide is a cyclic peptide and wherein the peptide specifically interacts with a monomeric, oligomeric and/or amyloid form of the A β peptide.

120. The peptide of claim 119, wherein Nu=T; N=3; X₁ is selected from T, R, D, L, F or A; X₂ is selected from C, R, S, G, Q, I, W, D, or F; and X₃=R.

121. The peptide of claim 119, wherein Nu=T; N=4; X₁ is any amino acid excluding F, W, Y, Q, D, E, K and P; X₂ is any amino acid excluding E; X₃ is any amino acid excluding Q, M and K; and X₄ is R.

122. The peptide of claim 121, wherein X₁ is selected from A, I, L, V, N, C, M, S, T, R, H, or G.

123. The peptide of claim 122, wherein X₁ is T, V or A.

124. The peptide of claim 121, wherein X₂ is selected from A, I, L, V, F, W, Y, N, Q, C, M, S, T, D, R, H, K, P or G.

125. The peptide of claim 124, wherein X₂ is I, L, V, F, Y or T.

126. The peptide of claim 121, wherein X₃ is A, I, L, V, F, W, Y, N, C, S, T, D, E, R, H, P or G.

127. The peptide of claim 126, wherein X₃ is A, W, or D.

128. The peptide of claim 121, wherein Nu=T; N=4; X₁ is selected from A, I, L, V, N, C, M, S, T, R, H, or G; X₂ is selected from A, I, L, V, F, W, Y, N, Q, C, M, S, T, D, R, H, K, P or G; X₃ is A, I, L, V, F, W, Y, N, C, S, T, D, E, R, H, P or G; and X₄ is R.

129. The peptide of claim 121, wherein Nu=T; N=4; X₁ is T, V or A; X₂ is I, L, V, F, Y or T; X₃ is A, W, or D; and X₄ is R.

130. The peptide of claim 119, wherein Nu=T; N=5; X₁ is I, L, V, C, S, K or P; X₂ is any amino acid excluding Y, N, Q, M, and R; X₃ is selected from I, L, V, F, W, Y, E, or R; X₄ is selected from L, V, F, Y, S or R; and X₅ is selected from W, M, N, D, or E

131. The peptide of claim 130, wherein the X₁ is selected from I, L, V, C, S, K or P.

132. The peptide of claim 131, wherein the X₁ is P, V or L.

133. The peptide of claim 130, wherein the X₂ is selected from A, I, L, V, F, W, C, S, T, D, E, H, K, P, or G.

134. The peptide of claim 133, wherein the X₂ is V or A.

135. The peptide of claim 130, wherein the X₃ is selected from I, L, V, F, W, Y, E or R.

136. The peptide of claim 135, wherein the X₃ is W.

137. The peptide of claim 130, wherein the X₄ is selected from L, V, F, Y, S or R.
138. The peptide of claim 137, wherein the X₄ is F.
139. The peptide of claim 130, wherein the X₅ is selected from W, M, N, D, or E.
140. The peptide of claim 139, wherein the X₅ is D. 118. The peptide of claim 119, wherein Nu=T; N=5; X₁ is P, V or L; X₂ is V or A; X₃ is selected from I, F, or W; X₄ is selected from L, F, or R; and X₅ is selected from W, N, or D.
141. The peptide of claim 119, wherein Nu=T; N=5; X₁ is P, V or L; X₂ is V or A; X₃ is W; X₄ is F; and X₅ is D.
142. The peptide of claim 120, wherein the peptide comprises the sequence TTCR, TTRR, TTSR, TRGR, TTGR, TRRR, TDQR, TLIR, TLWR, TLGR, TFDR, or TAFR (SEQ ID NOs 210-221).
143. A peptide comprising the amino acid sequences set forth in any one of SEQ ID NO:47-209, wherein, the peptide is cyclic and specifically interacts with a monomeric, oligomeric and/or amyloid form of the A β peptide..
144. The peptide of claim 128, wherein the peptide comprises an amino acid sequence selected from TAFDR, TAWCR, TTWCR, TTVDR, TTYAR, TTTAR or SASPT.
145. The peptide of claim 129, wherein the peptide comprises an amino acid sequence selected from the amino acid sequences set forth in SEQ ID NO:176-209.
146. The peptide of claim 129 or 145, wherein the peptide comprises an amino acid sequence of TPVWFD (SEQ ID NO:176) or TPAWFD (SEQ ID NO:177).
147. The peptide according to any one of claims 119-147, wherein at least one position of the peptide is a D amino acid.
148. The peptide according to any one of claims 119-147, wherein the peptide is in a linearized form.
149. A hybrid molecule comprising: a) a peptide set forth in any one of the claims 119-147, that specifically interacts with a monomeric, oligomeric and/or amyloid form of the A β peptide; and b) a scaffold molecule.
150. The molecule of claim 149, wherein the scaffolding molecule comprises a cell penetrating peptide.
151. The molecule of claim 149, wherein the scaffold molecule comprises a diagnostic or therapeutic reagent.
152. The hybrid molecule of claim 149, wherein the scaffold molecule comprises a polypeptide, small molecule or compound.

153. The hybrid molecule of claim 152, wherein the polypeptide comprises all or a sufficient portion of a protein selected from the group consisting of antibodies, enzymes, chromogenic proteins, or a fluorescent protein.
154. The hybrid molecule of claim 151, wherein the therapeutic agent is a neuroprotective agent that renders amyloid plaques less toxic or inhibits plaque formation.
155. The hybrid molecule of claim 151, wherein the diagnostic reagent specifically images oligomers and/or amyloid aggregates in neuronal tissue.
156. Use of a peptide according to anyone of claims 119-147 or a molecule of claims 156-161, to inhibit protein misfolding and aggregation.
157. The use of claim 156, wherein said peptide prevents misfolding and aggregation of the β -amyloid peptide.
158. A method of treatment, prevention or diagnosis of a disease related to protein misfolding and aggregation, comprising administering to a subject a therapeutically effective amount of a peptide according to any one of the claims 119 to 147 or a molecule of claims 156-161, wherein the disease is selected from Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, Creutzfeldt-Jakob disease, type 2 diabetes, familial amyloidotic polyneuropathy, systemic amyloidosis, and transmissible spongiform encephalopathy.
159. A method of treatment, prevention or diagnosis of Alzheimer's disease comprising administering to a subject a therapeutically effective amount of a peptide according to any one of the claims 119 to 147.
160. A pharmaceutical composition comprising a peptide according to any one of the claims 119 to 147 and a pharmaceutically acceptable carrier.
161. An isolated nucleic acid sequence encoding the peptide of claims 119 to 147.
162. A vector comprising the nucleic acid sequence of claim 161.
163. The vector of claim 162, wherein the vector is an expression vector.
164. A host cell comprising the vector of claim 163.
165. The host cell of claim 164, wherein the host cell is a prokaryotic or eukaryotic cell.

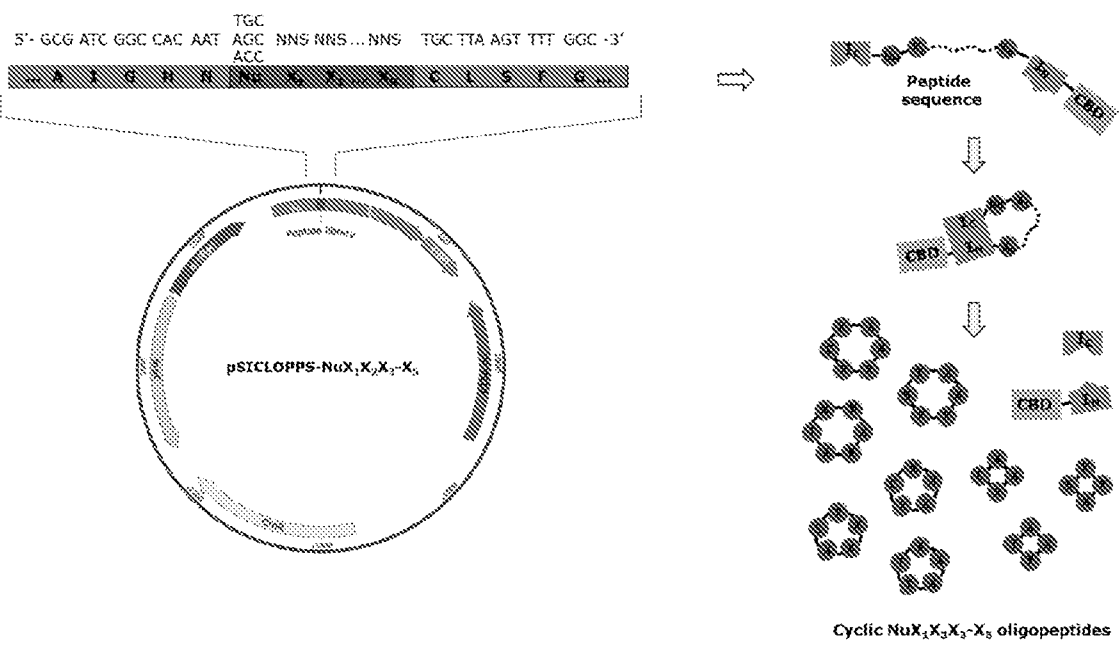


FIG. 1A

Peptide type	General formula ^a	Theoretical diversity	Actual library coverage
Tetrapeptides	cyclo-NuX ₁ X ₂ X ₃	3×20 ³ = 24,000	
Pentapeptides	cyclo-NuX ₁ X ₂ X ₃ X ₄	3×20 ⁴ = 480,000	
Hexapeptides	cyclo-NuX ₁ X ₂ X ₃ X ₄ X ₅	3×20 ⁵ = 9,600,000	
Combined library	cyclo-NuX ₁ X ₂ X ₃ -X ₅	10,104,000	×2

^aNu=C, S, or T; X=anyone of the twenty natural amino acids

FIG. 1B

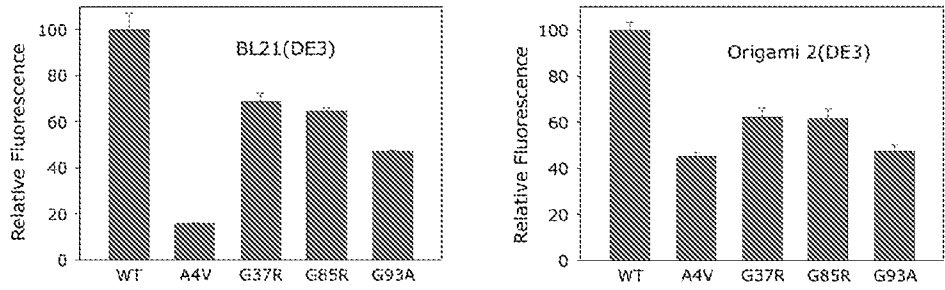


FIG. 2A

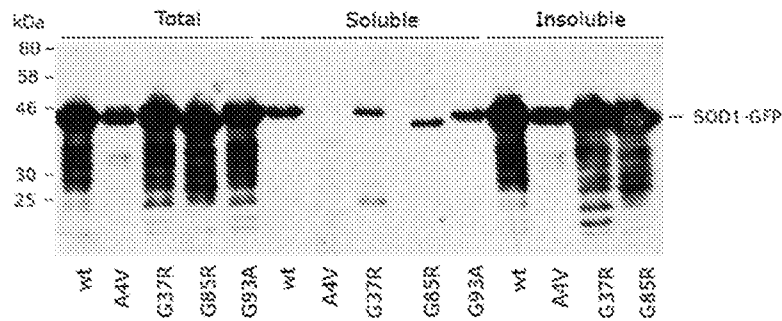


FIG. 2B

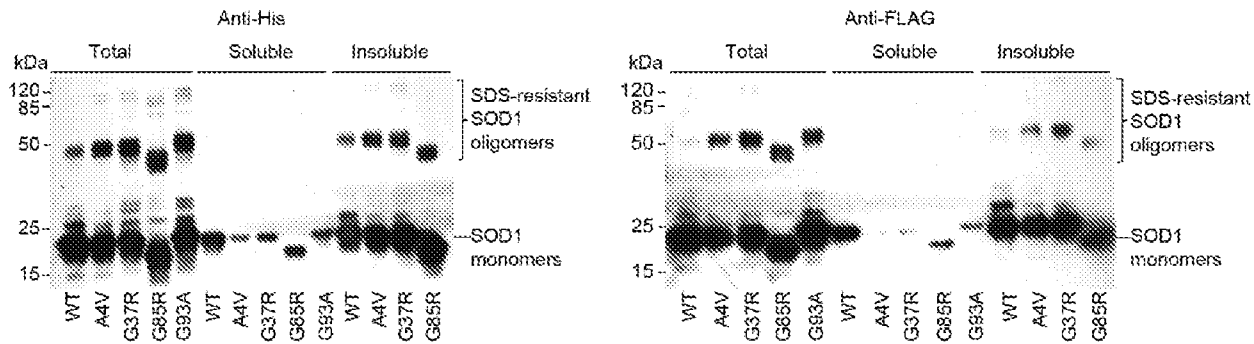


FIG. 2C

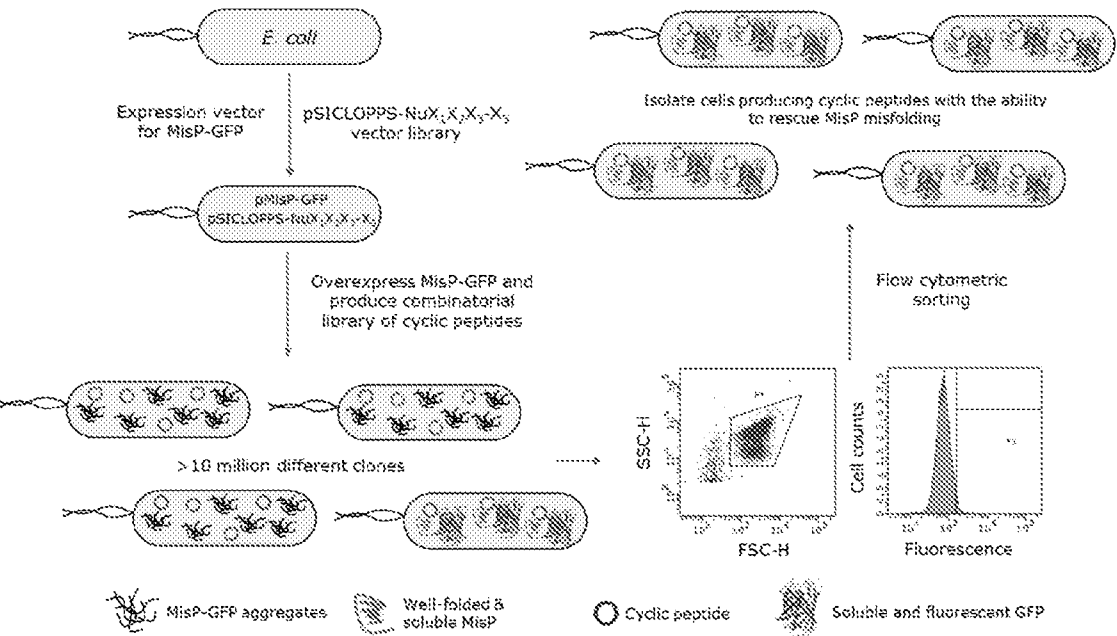


FIG. 3A

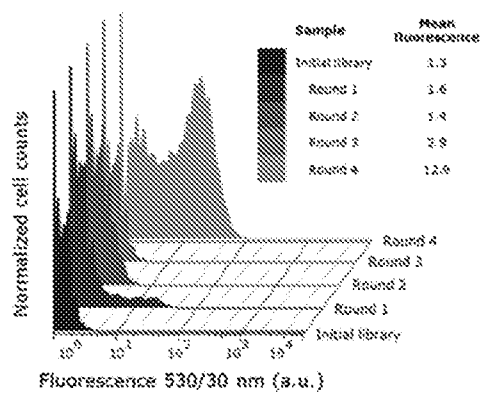


FIG. 3B

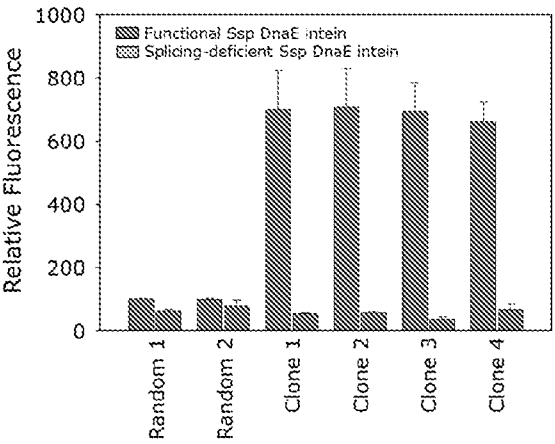


FIG. 3C

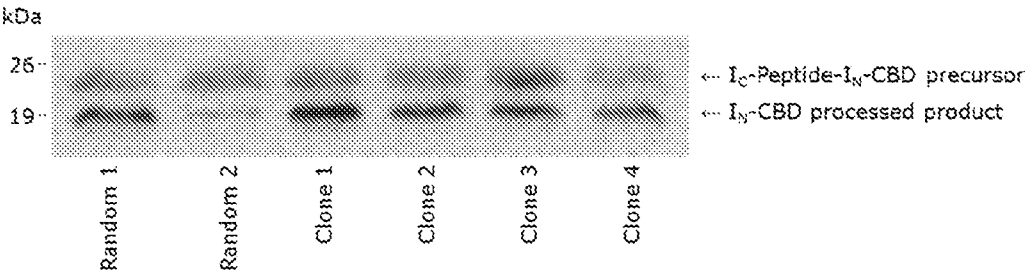


FIG. 3D

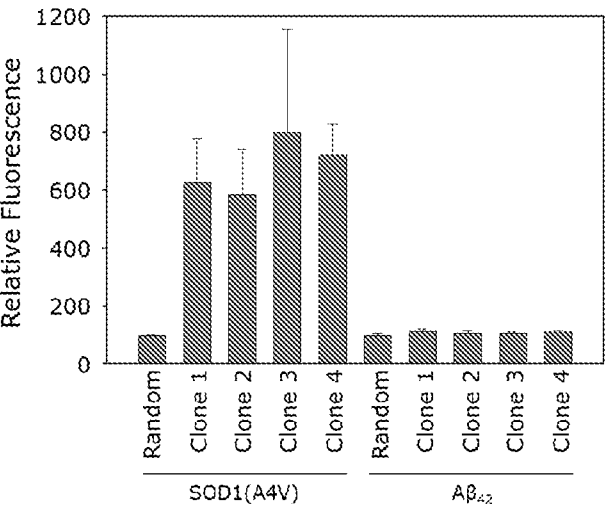
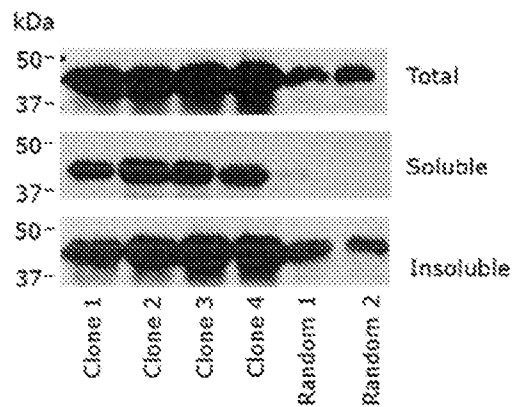
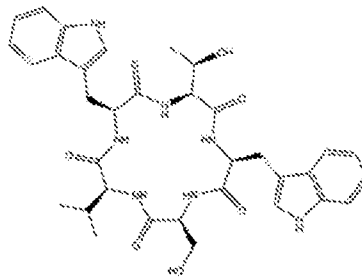


FIG. 3E

**FIG. 3F**

Isolated clone #	DNA sequence of the peptide-encoding region	Peptide type	Amino acid sequence
1	ACC GCG AGC TTC TGG	Pentapeptide	cyclo-TASFW
2	ACC TGG TCC GTG TGG	Pentapeptide	cyclo-TWSVW
3	ACC TTC AGC ATG TGG	Pentapeptide	cyclo-TFSMW
4	ACC TGG TCC GTG TGG	Pentapeptide	cyclo-TWSVW

FIG. 3G



SOD1C5-4

FIG. 4A

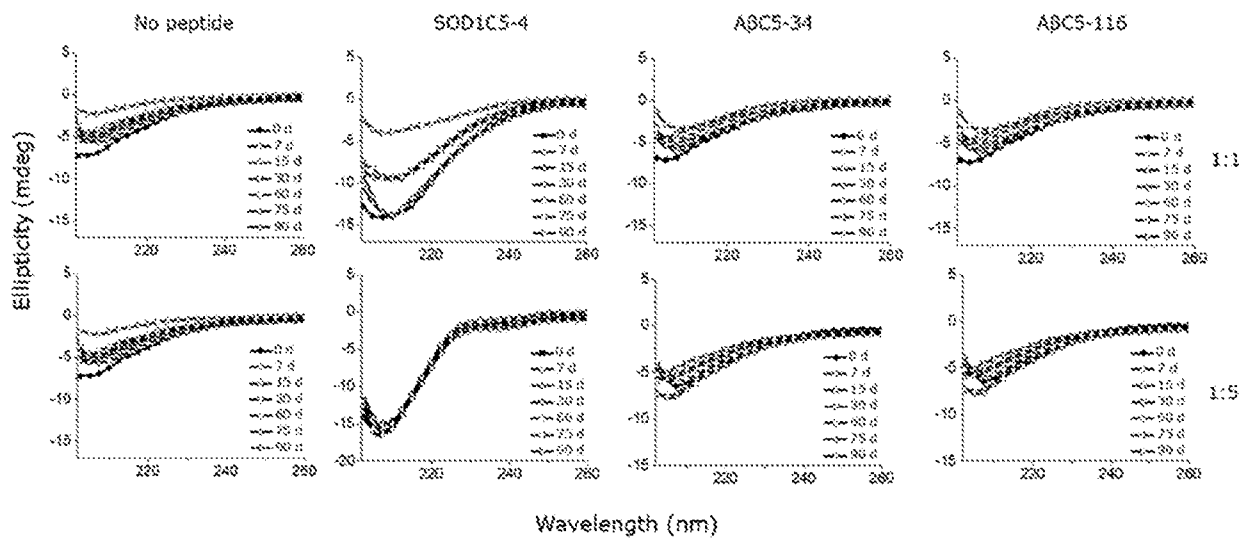


FIG. 4B

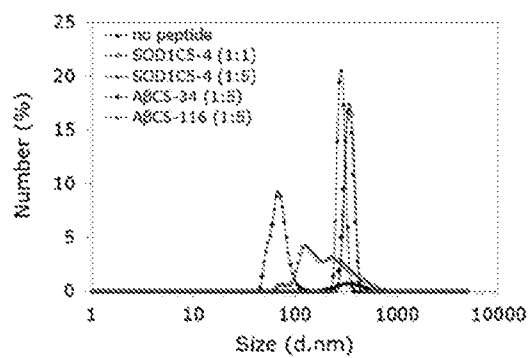


FIG. 4C

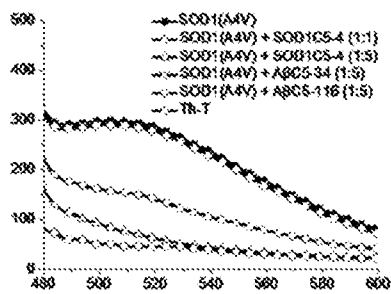


FIG. 4D

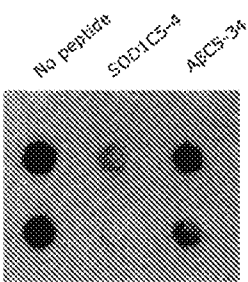


FIG. 4E

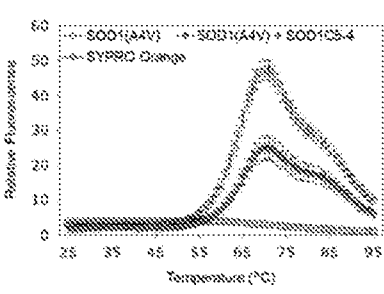


FIG. 4F

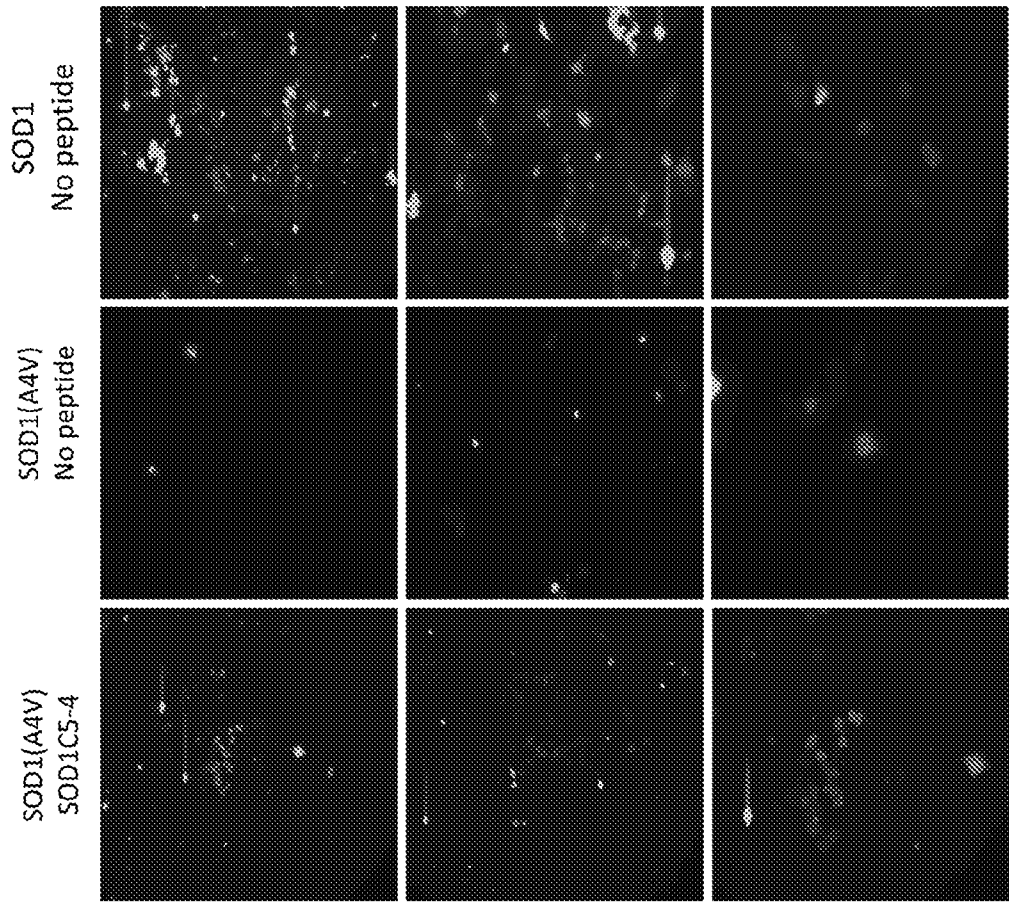


FIG. 5A

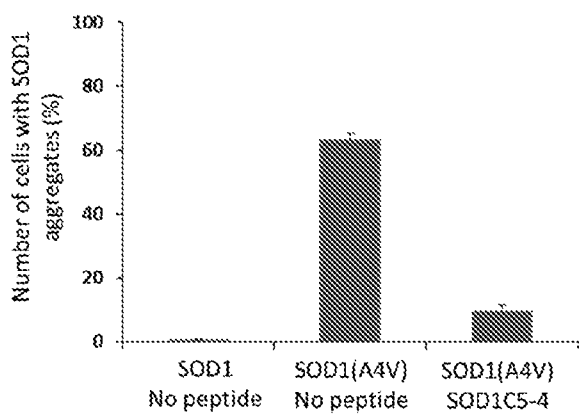


FIG. 5B

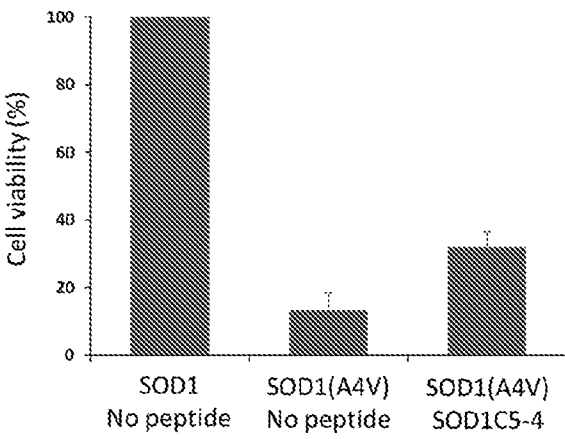


FIG. 5C

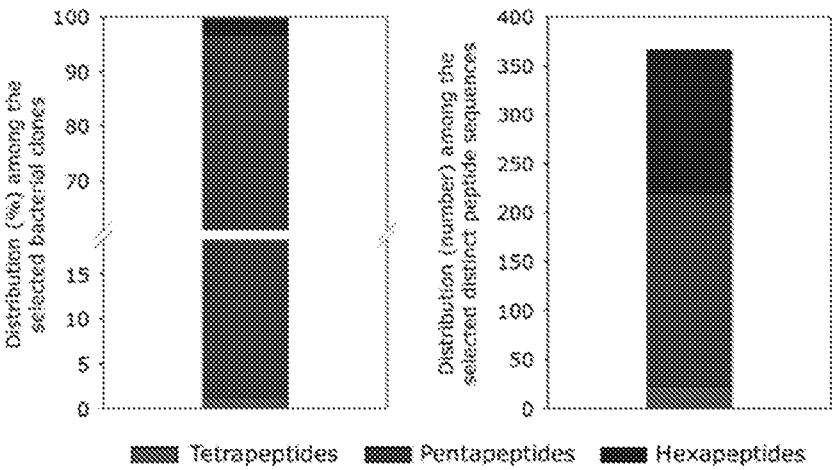


FIG. 6A

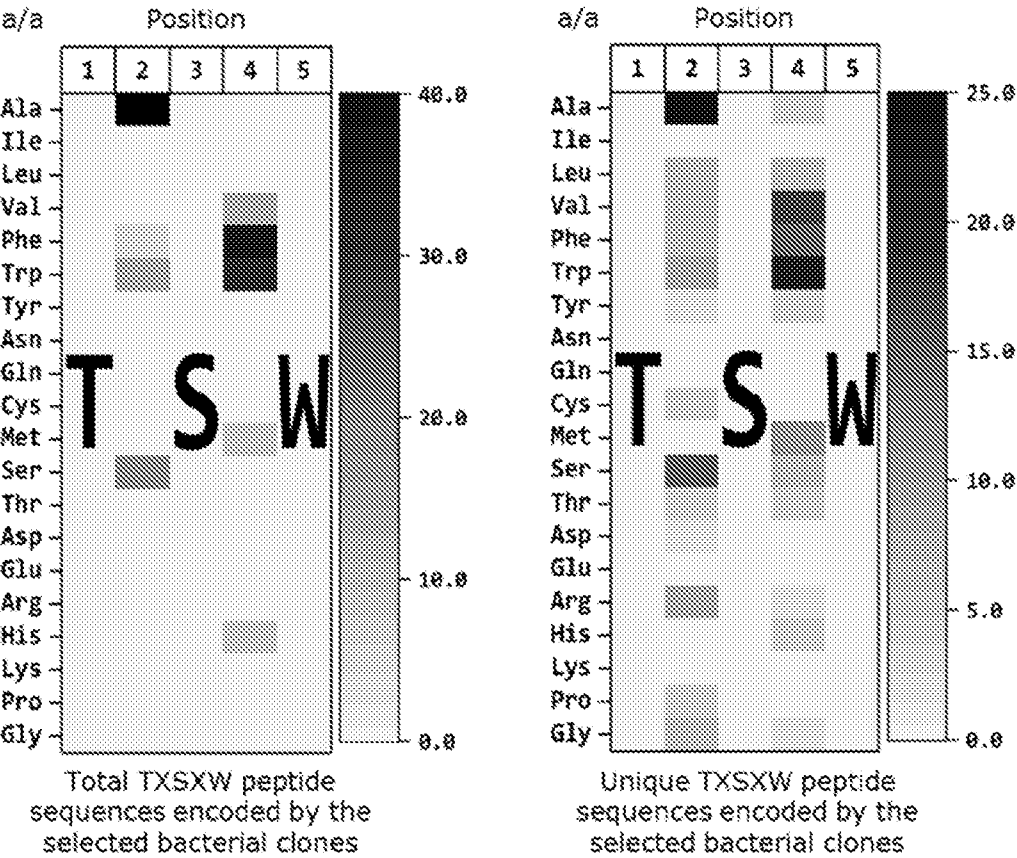


FIG. 6B

Total DNA reads corresponding to TXSXW pentapeptides

Position	Frequency of appearance (number)					Frequency of appearance (%)				
	1	2	3	4	5	1	2	3	4	5
A		2,423,975		629			62.09		0.02	
I		0		0			0.00		0.00	
L		3,760		2,169			0.10		0.06	
V		31,522		548,099			0.80		13.86	
F		209,112		1,479,547			5.31		37.56	
W		545,357		1,257,478	3,939,406		13.84		31.92	100.00
Y		71		128			0.00		0.00	
N		0		0			0.00		0.00	
Q		0		0			0.00		0.00	
C		268		0			0.01		0.00	
M		0		318,343			0.00		8.08	
S		702,153	3,939,406	2,484			17.62	100.00	0.06	
T	3,939,406	498		532		100.00	0.01		0.01	
D		136		0			0.00		0.00	
E		0		0			0.00		0.00	
R		797		1,485			0.02		0.04	
H		0		330,416			0.00		8.39	
K		0		0			0.00		0.00	
P		290		0			0.01		0.00	
G		1,367		96			0.03		0.00	
Sum			3,939,406					100.00		

FIG. 6C

Unique TXSXW pentapeptide sequences selected

Position	Frequency of appearance (number)					Frequency of appearance (%)				
	1	2	3	4	5	1	2	3	4	5
A		11		2			23.91		4.35	
I		0		0			0.00		0.00	
L		3		3			6.52		6.52	
V		3		8			6.52		17.39	
F		3		7			6.52		15.22	
W		4		10	46		8.70		21.74	100.00
Y		1		2			2.17		4.35	
N		0		0			0.00		0.00	
Q		0		0			0.00		0.00	
C		2		0			4.35		0.00	
M		0		5			0.00		10.87	
S		7	46	3			15.22	100.00	6.52	
T	46	2		2		100.00	4.35		4.35	
D		1		0			2.17		0.00	
E		0		0			0.00		0.00	
R		4		1			8.70		2.17	
H		0		2			0.00		4.35	
K		0		0			0.00		0.00	
P		2		0			4.35		0.00	
G		3		1			6.52		2.17	
Sum			46					100.00		

FIG. 6D

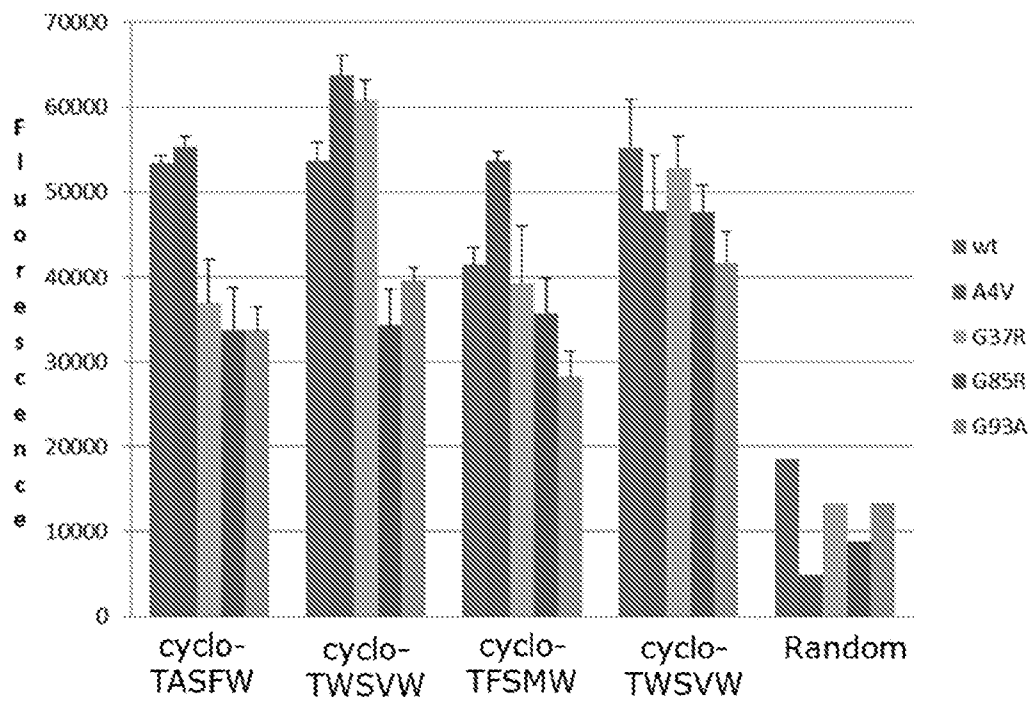


FIG. 6E

FIG. 7

SEQ ID NO	Peptide name	Amino acid sequence				Number of reads	Reads/ Total TXSXW reads (%)	Reads/ Total peptide reads (%)	Reads/ Total peptide reads (%)	Peptide-encoding nucleotide sequence
1	SOD1C5-1	T	A	S	W	W	31.877	30.963	29.591	ACCGCTCGTGGTGG
2	SOD1C5-2	T	A	S	F	W	18.902	18.360	17.547	ACCGCGAGCTTCTGG
3	SOD1C5-3	T	S	S	F	W	17.770	17.261	16.496	ACCTCGTCGTTCTGG
4	SOD1C5-4	T	W	S	V	W	13.809	13.413	12.819	ACCTGGTCCGTGTGG
5	SOD1C5-5	T	A	S	H	W	8.386	8.146	7.785	ACCGCCAGCCACTGG
6	SOD1C5-6	T	F	S	M	W	5.302	5.150	4.922	ACCTTCAGCATGTGG
7	SOD1C5-7	T	A	S	M	W	2.756	2.677	2.559	ACCGCTCGATGTGG
8	SOD1C5-9	T	V	S	F	W	0.795	0.772	0.738	ACCGTCTCGTTCTGG
9	SOD1C5-11	T	L	S	F	W	0.078	0.076	0.072	ACCTCTCCTTCTGG
10	SOD1C5-13	T	A	S	R	W	0.038	0.037	0.035	ACGCCAGCCGCTGG
11	SOD1C5-14	T	A	S	S	W	0.037	0.036	0.034	ACCGCGAGCTCGTGG
12	SOD1C5-18	T	A	S	L	W	0.027	0.026	0.025	ACCGGAGCCTCTGG
13	SOD1C5-20	T	S	S	S	W	0.025	0.024	0.023	ACCTCGTCGTCCTGG
14	SOD1C5-23	T	G	S	V	W	0.019	0.019	0.018	ACCGCTCCGTGTGG
15	SOD1C5-25	T	W	S	L	W	0.017	0.017	0.016	ACCTGGTCCCTGTGG
16	SOD1C5-27	T	L	S	M	W	0.016	0.015	0.015	ACCTCAGCATGTGG
17	SOD1C5-31	T	W	S	A	W	0.015	0.014	0.014	ACCTGGTCCGCGTGG
18	SOD1C5-32	T	G	S	W	W	0.014	0.014	0.013	ACCGCTCGTGGTGG
19	SOD1C5-33	T	R	S	V	W	0.014	0.014	0.013	ACCGGTCCGTGTGG
20	SOD1C5-39	T	S	S	L	W	0.011	0.011	0.010	ACCTCGTCGCTCTGG
21	SOD1C5-44	T	A	S	T	W	0.009	0.009	0.009	ACGCCAGCACCTGG
22	SOD1C5-46	T	S	S	V	W	0.009	0.009	0.008	ACCTCGTCCGTCTGG

23	SOD1C5-53	T	T	S	W	W	295	0.007	0.007	0.007	0.007	ACCACCTCGTGGTGG
24	SOD1C5-65	T	A	S	V	W	245	0.006	0.006	0.006	0.006	ACCGCGAGCGTCTGG
25	SOD1C5-74	T	C	S	W	W	208	0.005	0.005	0.005	0.005	ACCTGCTCGTGGTGG
26	SOD1C5-75	T	P	S	F	W	208	0.005	0.005	0.005	0.005	ACCCCGTCGTTCTGG
27	SOD1C5-76	T	T	S	F	W	203	0.005	0.005	0.005	0.005	ACCACGAGCTTCTGG
28	SOD1C5-80	T	F	S	T	W	171	0.004	0.004	0.004	0.004	ACCTTCAGCACGTGG
29	SOD1C5-82	T	S	S	M	W	164	0.004	0.004	0.004	0.004	ACCTCGAGCATGTGG
30	SOD1C5-89	T	V	S	W	W	144	0.004	0.004	0.004	0.003	ACCGTCTCGTGGTGG
31	SOD1C5-93	T	D	S	W	W	136	0.003	0.003	0.003	0.003	ACCGACTCGTGGTGG
32	SOD1C5-105	T	S	S	W	W	112	0.003	0.003	0.003	0.003	ACCTCGTCC TGGTGG
33	SOD1C5-106	T	R	S	W	W	106	0.003	0.003	0.003	0.002	ACCCGCTCGTGGTGG
34	SOD1C5-109	T	W	S	M	W	99	0.003	0.002	0.002	0.002	ACCTGGTCCATGTGG
35	SOD1C5-116	T	A	S	G	W	96	0.002	0.002	0.002	0.002	ACCGCCTCTGGGTGG
36	SOD1C5-130	T	P	S	W	W	82	0.002	0.002	0.002	0.002	ACGCCCTCGTGGTGG
37	SOD1C5-135	T	R	S	F	W	79	0.002	0.002	0.002	0.002	ACGCGGAGCTTCTGG
38	SOD1C5-140	T	S	S	Y	W	76	0.002	0.002	0.002	0.002	ACCTCGTCTACTGG
39	SOD1C5-143	T	L	S	V	W	72	0.002	0.002	0.002	0.002	ACCTTGAGCGTGTGG
40	SOD1C5-148	T	Y	S	W	W	71	0.002	0.002	0.002	0.002	ACCTACTCATGGTGG
41	SOD1C5-160	T	F	S	V	W	62	0.002	0.002	0.002	0.001	ACCTTCAGCGTGTGG
42	SOD1C5-164	T	C	S	V	W	60	0.002	0.001	0.001	0.001	ACCTGCTCCGTGTGG
43	SOD1C5-167	T	V	S	S	W	59	0.001	0.001	0.001	0.001	ACCGTCTCGTCGTGG
44	SOD1C5-168	T	R	S	H	W	58	0.001	0.001	0.001	0.001	ACCCGCAGCCACTGG
45	SOD1C5-182	T	G	S	A	W	53	0.001	0.001	0.001	0.001	ACCGGCAGCGCGTGG
46	SOD1C5-188	T	A	S	Y	W	52	0.001	0.001	0.001	0.001	ACGCCAGCTACTGG
							3,939,406	100	97.134	92.829		
							Sum					

FIG. 7(cont.)

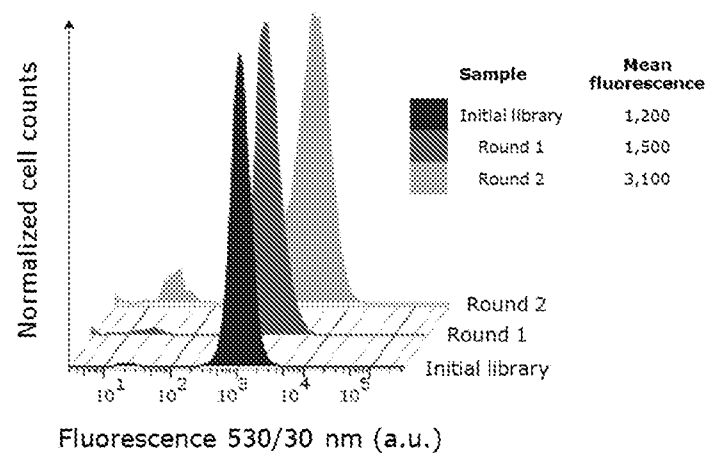


FIG. 8A

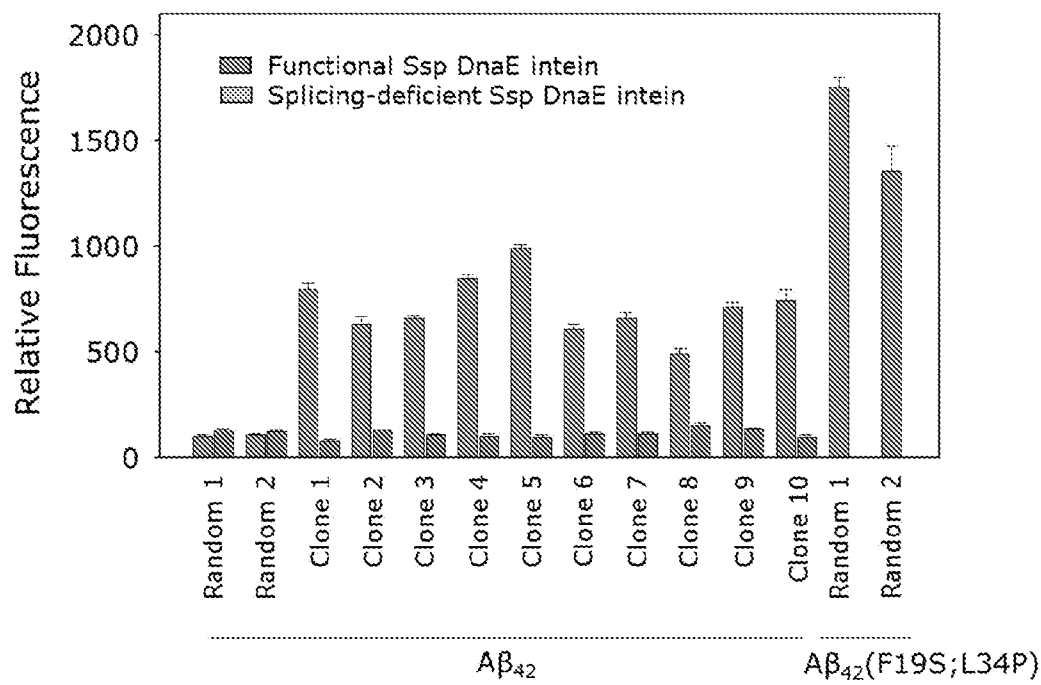


FIG. 8B

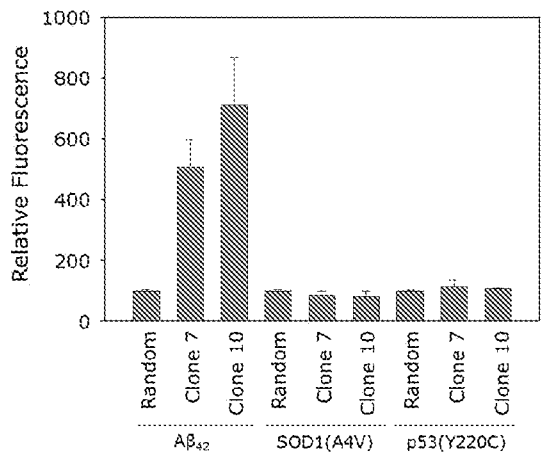


FIG. 8C

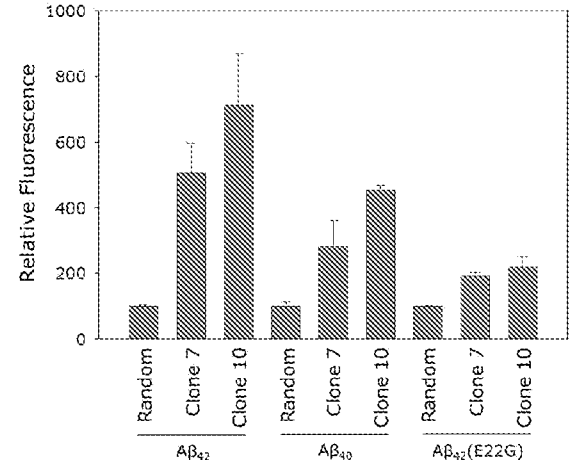


FIG. 8D

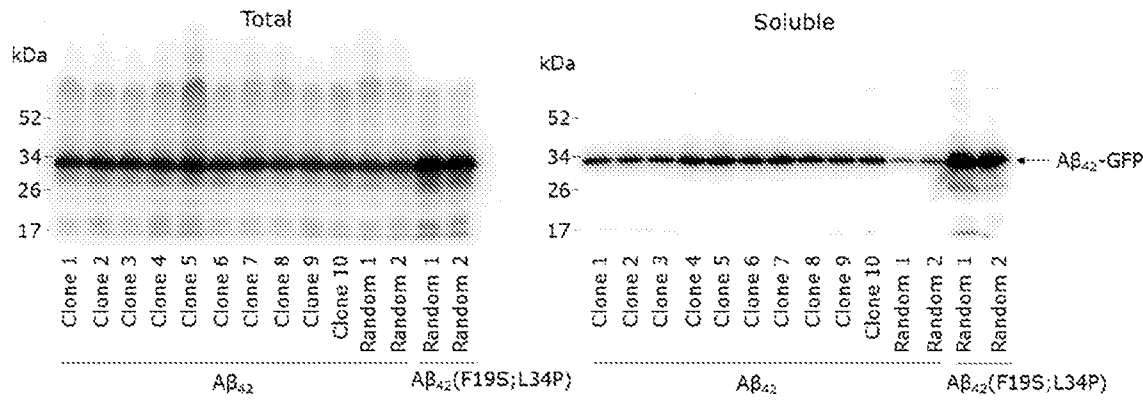


FIG. 8E

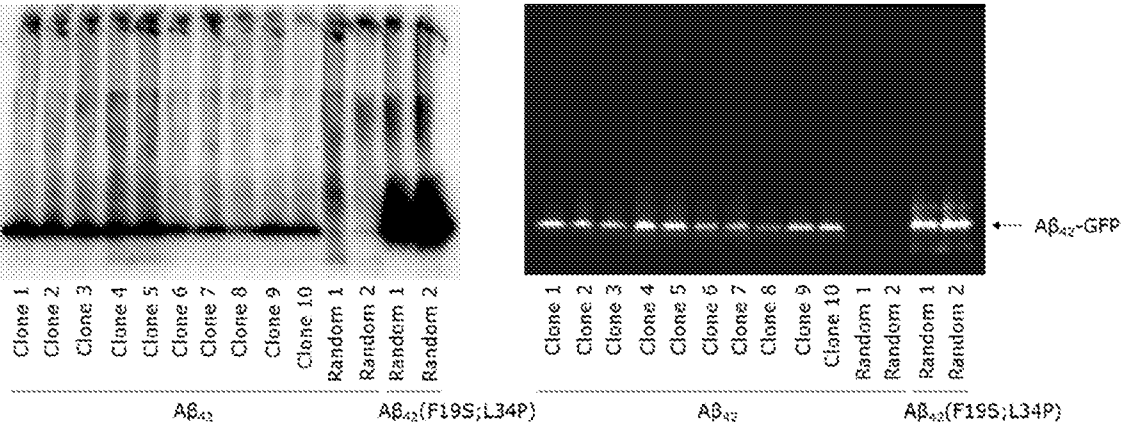


FIG. 8F

Isolated clone #	DNA sequence of the peptide-encoding region	Peptide type	Amino acid sequence
1	ACC ACC GTG GAC CGG	Pentapeptide	cyclo-TTVDR
2	ACC ACG TAC GCC AGG	Pentapeptide	cyclo-TTYAR
3	ACC ACC ACG GCC CGG	Pentapeptide	cyclo-TTTAR
4	ACC CCG GTC TGG TTC GAC	Hexapeptide	cyclo-TPVWFD
5	ACC CCG GTC TGG TTC GAC	Hexapeptide	cyclo-TPVWFD
6	ACC ACG TAC GCC AGG	Pentapeptide	cyclo-TTYAR
7	AGC GCC TCG CCG ACG	Pentapeptide	cyclo-SASPT
8	ACC GCG TGG TGC CGC	Pentapeptide	cyclo-TAWCR
9	ACC ACC TGG TGC CGG	Pentapeptide	cyclo-TTWCR
10	ACC GCG TTC GAC CGG	Pentapeptide	cyclo-TAFDR

FIG. 8G

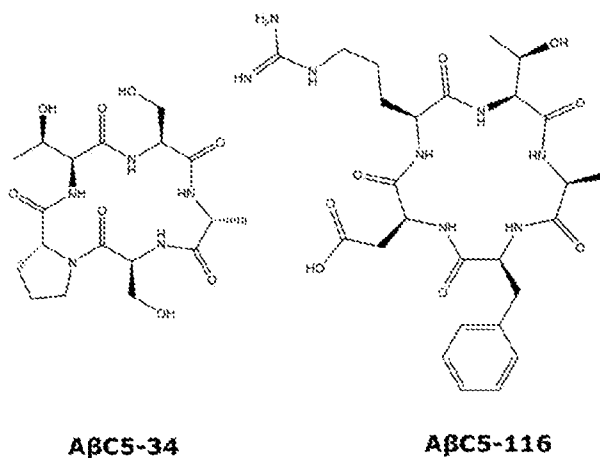


FIG. 9A

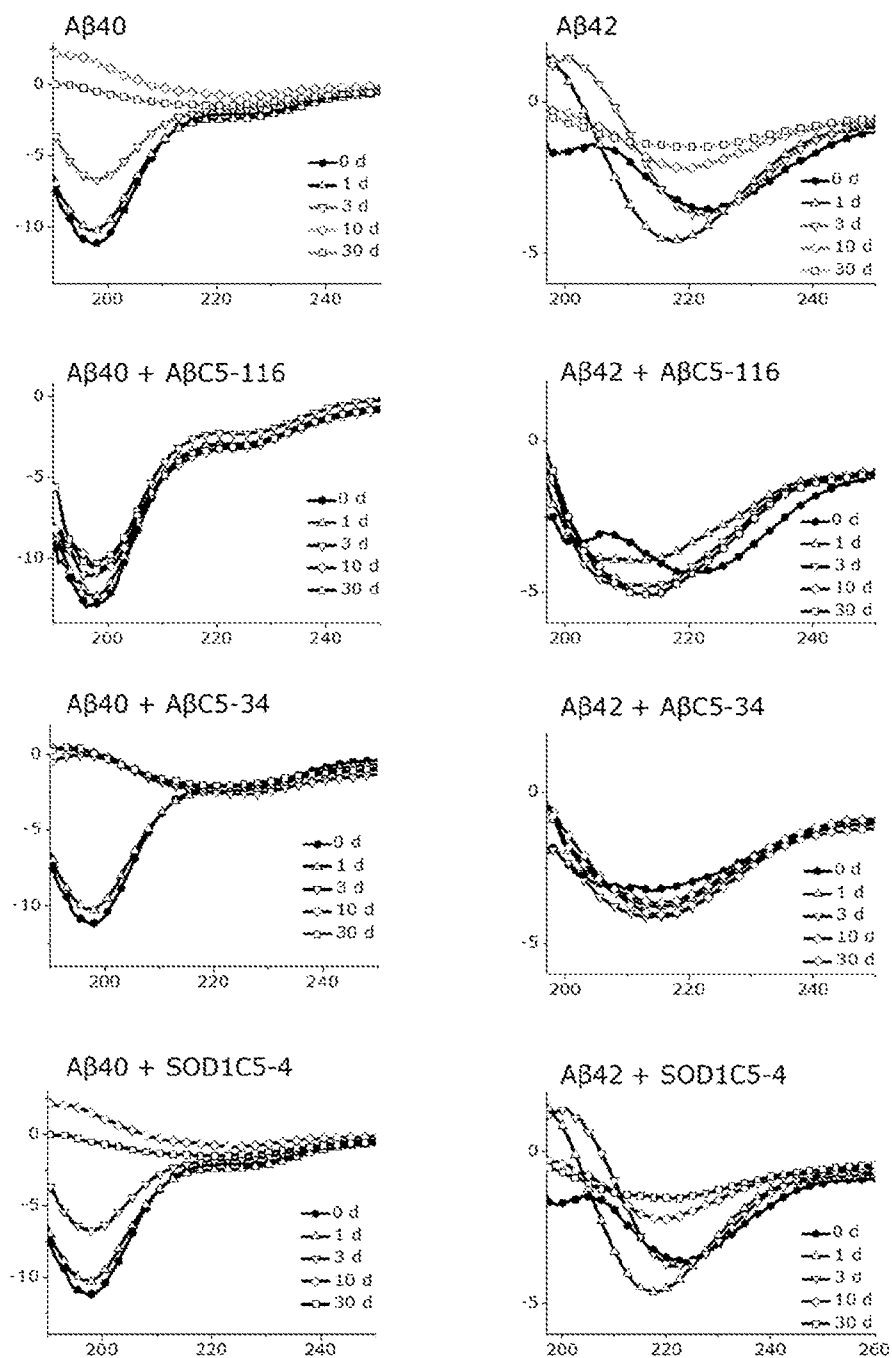


FIG. 9B

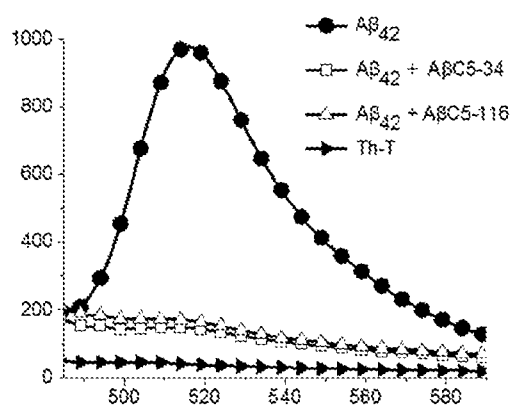
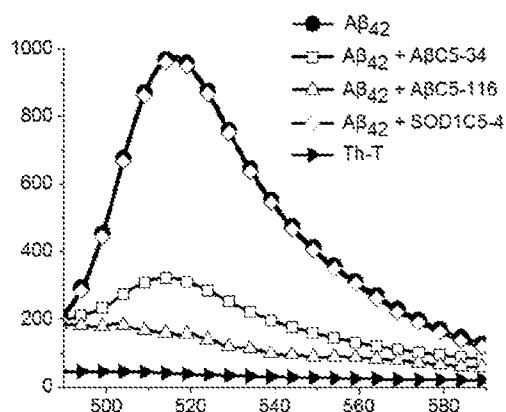
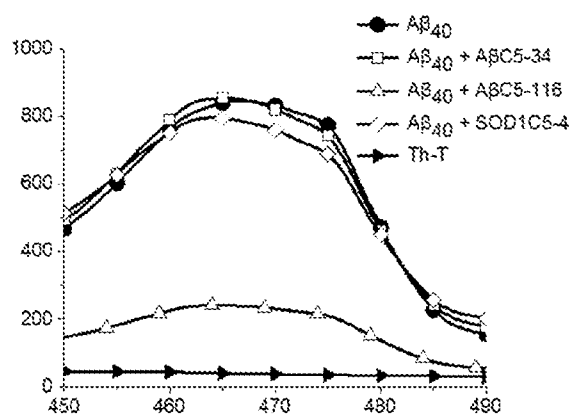


FIG. 9C

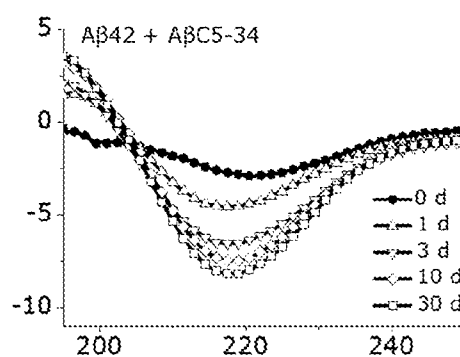
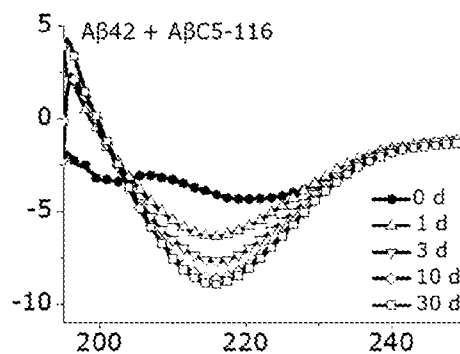
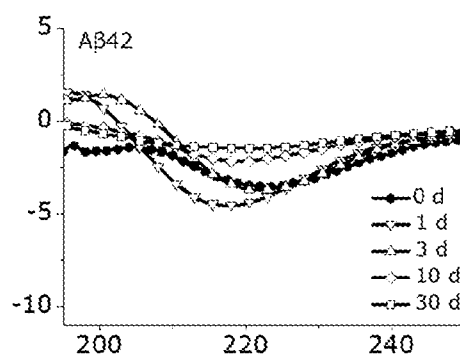


FIG. 9D

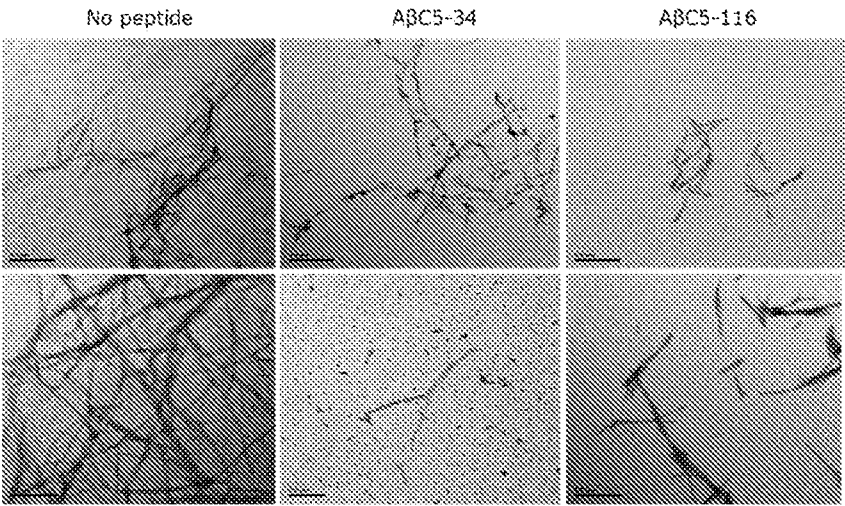


FIG. 9E

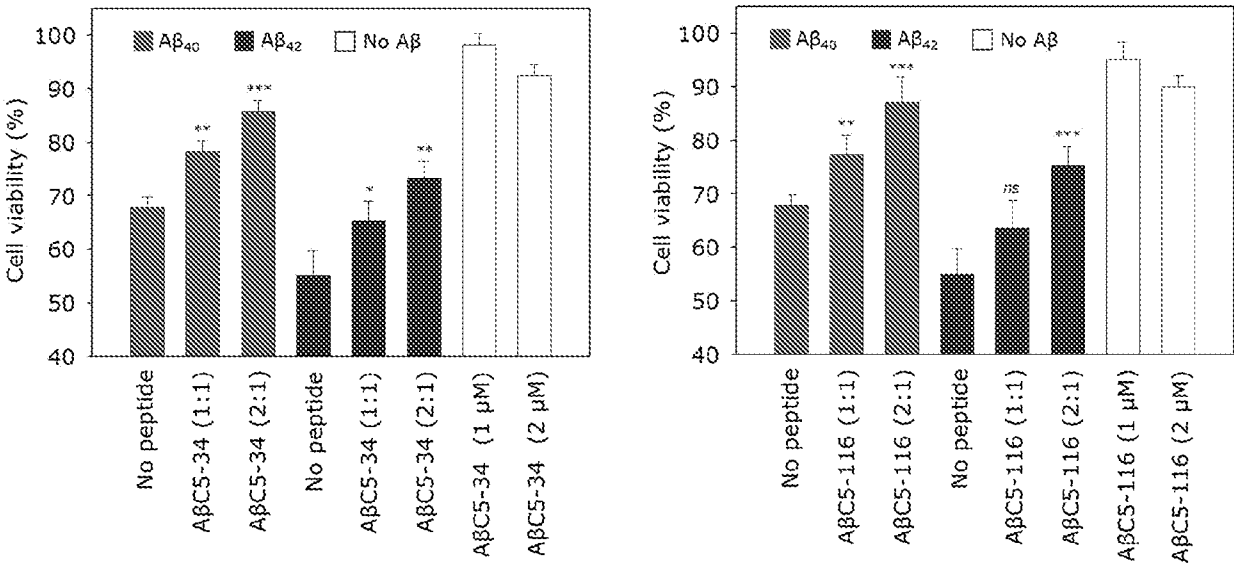


FIG. 10A

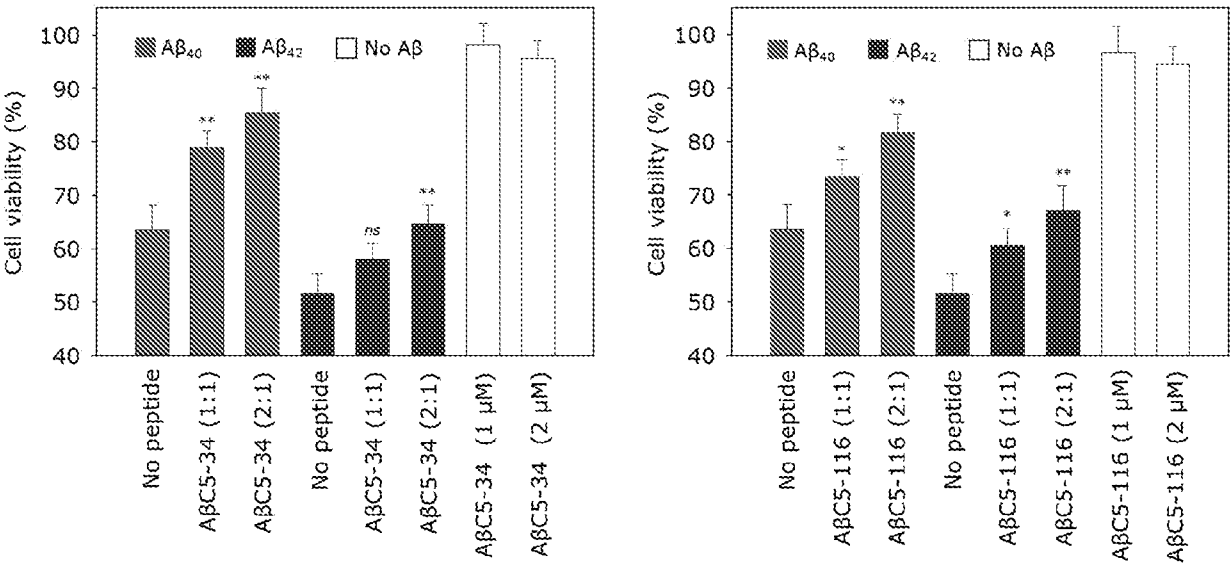


FIG. 10B

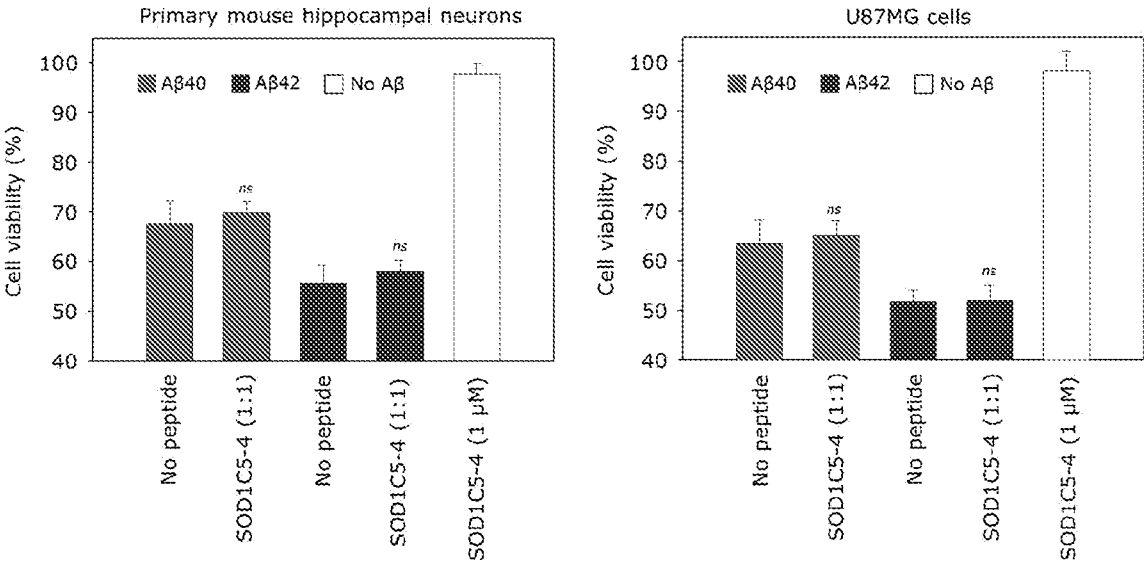


FIG. 10C

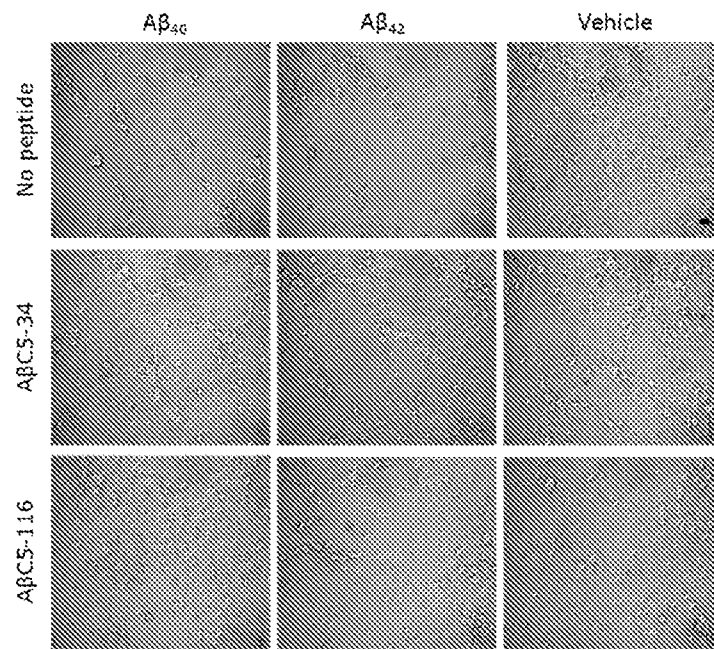


FIG. 10D

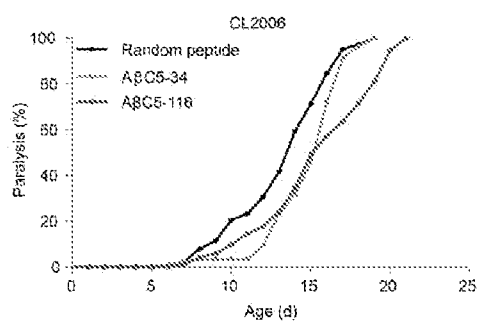


FIG. 11A

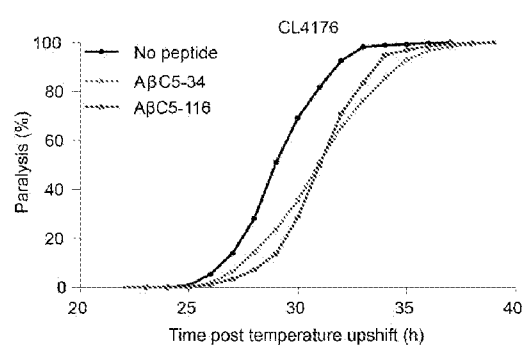


FIG. 11B

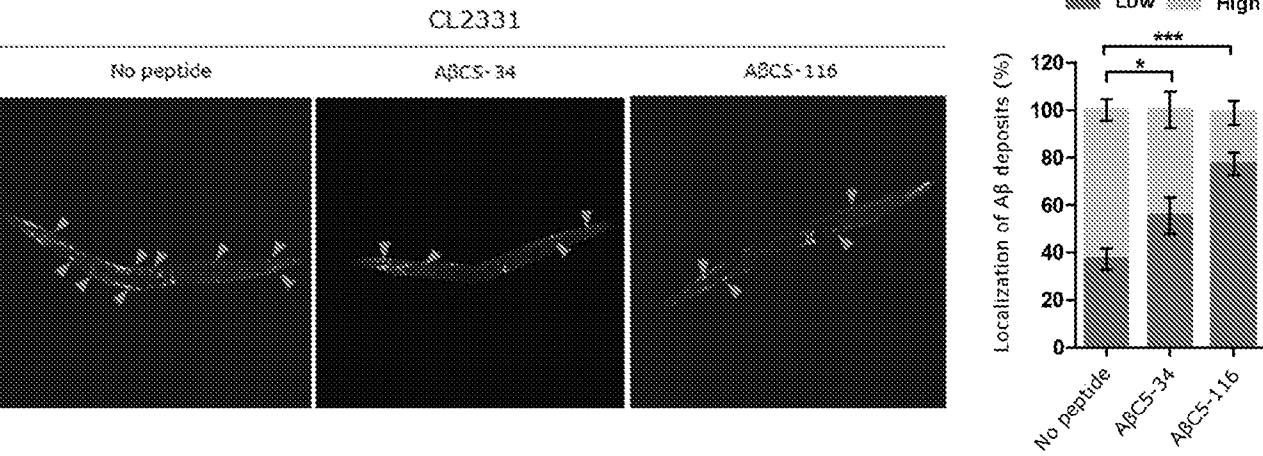


FIG. 11C

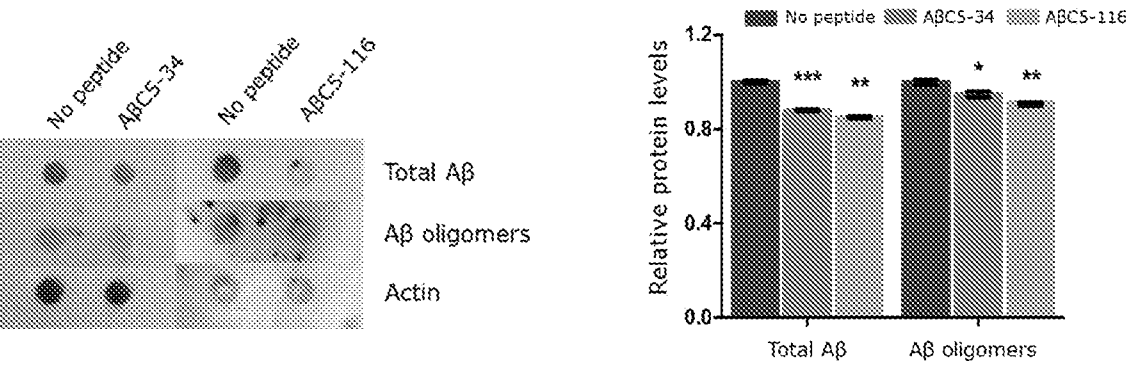


FIG. 11D

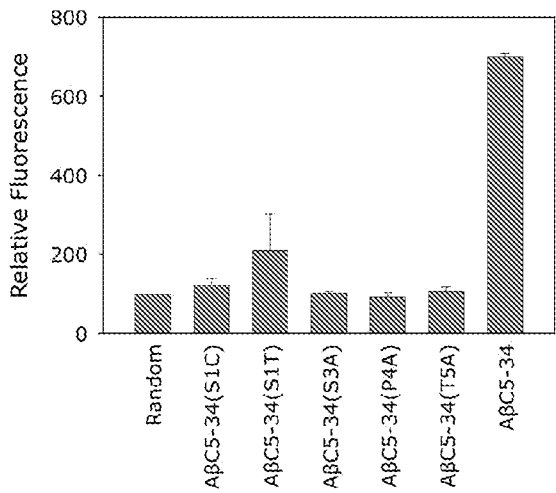


FIG. 12A

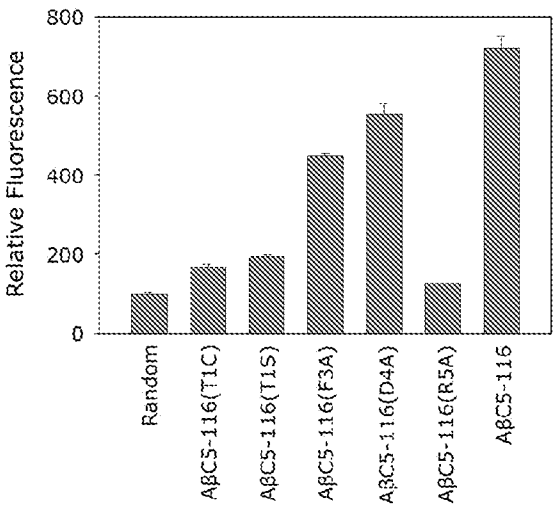


FIG. 12B

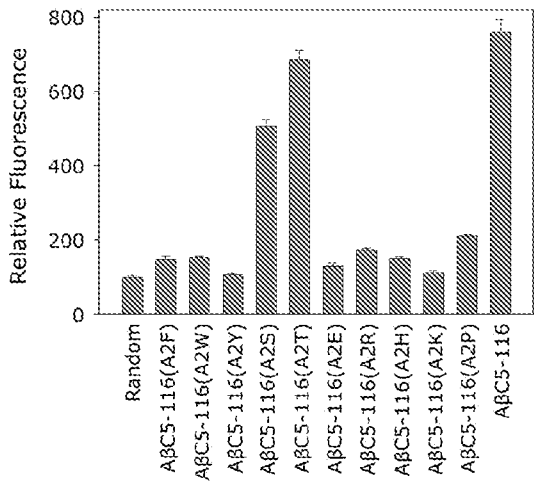


FIG. 12C

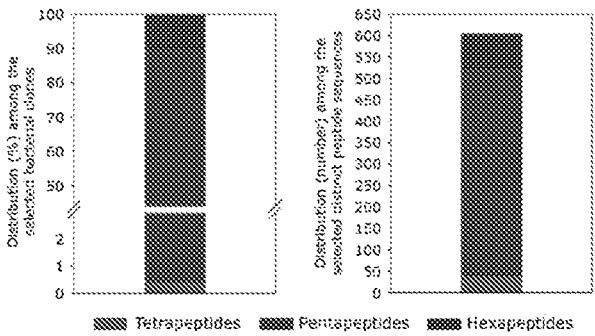


FIG. 12D

Total DNA reads corresponding to TXXXR pentapeptides

Position	Frequency of appearance (number)					Frequency of appearance (%)				
	1	2	3	4	5	1	2	3	4	5
Amino acid	A	224,879	726	668,907		11.82	0.04	36.75		
	I	4,066	64,587	232		0.21	4.97	0.01		
	L	20,641	438,606	3,996		1.10	23.08	0.21		
	V	202,870	278,637	69,596		10.67	14.85	3.66		
	F	0	108,148	10,911		0.00	5.58	0.57		
	W	0	234,933	287,886		0.00	12.35	15.14		
	Y	0	313,613	90		0.00	18.49	0.00		
	N	222	47,913	7,049		0.01	2.52	0.37		
	Q	0	160	0		0.00	0.01	0.00		
	C	254	1,052	79,338		0.01	0.06	4.17		
	M	51	3,338	0		0.00	0.18	0.00		
	S	60,697	32,123	39,090		3.19	1.69	2.06		
	T	1,901,945	303,173	253,496	1,065	100.00	71.67	13.33	0.06	
	D	0	58	515,155		0.00	0.00	27.09		
	E	0	0	27,822		0.00	0.00	1.46		
	R	203	63,829	2,070	1,901,945	0.01	3.36	0.11	100.00	
	H	953	15,733	155,451		0.05	0.83	8.17		
	K	0	4,078	0		0.00	0.21	0.00		
	P	0	910	2,039		0.00	0.05	0.11		
	G	23,736	12,005	1,247		1.25	0.63	0.07		
	Sum		1,901,945				100.00			

Unique TXXXR pentapeptide sequences selected

Position	Frequency of appearance (number)					Frequency of appearance (%)				
	1	2	3	4	5	1	2	3	4	5
Amino acid	A	25	4	32		15.72	2.52	20.13		
	I	2	7	2		1.26	4.40	1.26		
	L	4	28	3		2.52	17.81	1.89		
	V	13	18	8		8.18	11.32	5.03		
	F	0	7	5		0.00	4.40	3.14		
	W	0	18	19		0.00	11.32	11.95		
	Y	0	6	1		0.00	3.77	0.63		
	N	2	2	3		1.26	1.26	1.89		
	Q	0	2	0		0.00	1.26	0.00		
	C	1	4	7		0.63	2.52	4.40		
	M	1	7	0		0.63	4.40	0.00		
	S	10	10	5		6.29	6.29	3.14		
	T	159	67	6		100.00	55.35	10.06	3.77	
	D	0	1	27		0.00	0.63	18.98		
	E	0	0	3		0.00	0.00	1.89		
	R	3	10	9	159	1.89	6.29	5.88	100.00	
	H	6	8	13		3.77	3.77	8.18		
	K	0	2	0		0.00	1.26	0.00		
	P	0	4	3		0.00	2.52	1.89		
	G	4	7	13		2.52	4.40	8.18		
	Sum		159				100.00			

FIG. 12E

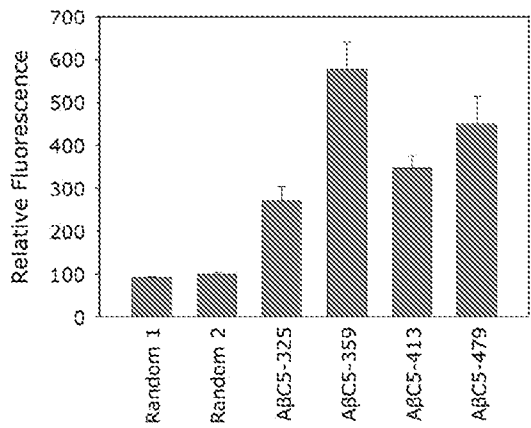


FIG. 12F

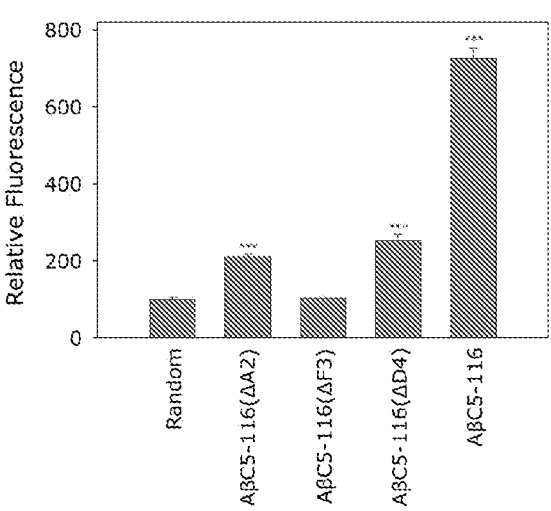


FIG. 12G

Sequence analysis of the total cyclic hexapeptide sequences encoded by the selected bacterial clones

Position	Frequency of appearance (number)						Frequency of appearance (%)					
	1	2	3	4	5	6	1	2	3	4	5	6
A	0	0	111,637	0	0	0	0.00	0.00	34.95	0.00	0.00	0.00
I	0	227	384	13,135	0	0	0.00	0.07	0.12	4.11	0.00	0.00
L	0	54,431	13,451	687	12,121	0	0.00	17.04	4.21	0.22	3.79	0.00
V	0	17,100	132,533	915	586	0	0.00	5.35	41.49	0.29	0.18	0.00
F	0	0	227	27,171	292,976	0	0.00	0.00	0.07	8.51	91.71	0.00
W	0	0	355	275,180	0	13,368	0.00	0.00	0.11	86.14	0.00	4.18
Y	0	0	0	2,020	546	0	0.00	0.00	0.00	0.63	0.17	0.00
N	0	0	0	0	0	11,928	0.00	0.00	0.00	0.00	0.00	3.73
Q	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00
C	0	623	767	0	0	0	0.00	0.20	0.22	0.00	0.00	0.00
M	0	0	0	0	0	556	0.00	0.00	0.00	0.00	0.00	0.17
S	0	307	1,284	0	86	0	0.00	0.10	0.40	0.00	0.03	0.00
T	319,133	0	17,100	0	0	0	100.00	0.00	5.35	0.00	0.00	0.00
D	0	0	114	0	0	292,644	0.00	0.00	0.04	0.00	0.00	91.61
E	0	0	29,077	307	0	954	0.00	0.00	9.10	0.10	0.00	0.30
R	0	0	0	55	13,135	0	0.00	0.00	0.00	0.02	4.11	0.00
H	0	0	647	0	0	0	0.00	0.00	0.20	0.00	0.00	0.00
K	0	1,231	11,016	0	0	0	0.00	0.39	3.45	0.00	0.00	0.00
P	0	245,531	289	0	0	0	0.00	76.86	0.09	0.00	0.00	0.00
G	0	0	629	0	0	0	0.00	0.00	0.20	0.00	0.00	0.00
Sum			319,450						100.00			

Sequence analysis of the unique selected cyclic hexapeptides

Position	Frequency of appearance (number)						Frequency of appearance (%)					
	1	2	3	4	5	6	1	2	3	4	5	6
A	0	0	6	0	0	0	0.00	0.00	17.65	0.00	0.00	0.00
I	0	1	1	1	0	0	0.00	2.94	2.94	2.94	0.00	0.00
L	0	13	2	4	6	0	0.00	38.24	5.88	11.76	17.65	0.00
V	0	1	4	3	3	0	0.00	2.94	11.76	8.82	8.82	0.00
F	0	0	1	2	21	0	0.00	0.00	2.94	5.88	61.76	0.00
W	0	0	1	19	0	2	0.00	0.00	2.94	55.88	0.00	5.88
Y	0	0	0	3	2	0	0.00	0.00	0.00	8.82	5.88	0.00
N	0	0	0	0	0	4	0.00	0.00	0.00	0.00	0.00	11.76
Q	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00
C	0	1	2	0	0	0	0.00	2.94	5.88	0.00	0.00	0.00
M	0	0	0	0	0	1	0.00	0.00	0.00	0.00	0.00	2.94
S	0	1	5	0	1	0	0.00	2.94	14.71	0.00	2.94	0.00
T	34	0	1	0	0	0	100.00	0.00	2.94	0.00	0.00	0.00
D	0	0	1	0	0	23	0.00	0.00	2.94	0.00	0.00	73.53
E	0	0	4	1	0	2	0.00	0.00	11.76	2.94	0.00	5.88
R	0	0	0	1	1	0	0.00	0.00	0.00	2.94	2.94	0.00
H	0	0	1	0	0	0	0.00	0.00	2.94	0.00	0.00	0.00
K	0	1	1	0	0	0	0.00	2.94	2.94	0.00	0.00	0.00
P	0	16	1	0	0	0	0.00	47.06	2.94	0.00	0.00	0.00
G	0	0	3	0	0	0	0.00	0.00	8.82	0.00	0.00	0.00
Sum			34						100.00			

FIG. 12H