

(51) International Patent Classification:
Not classified

(21) International Application Number:
PCT/EP2024/071389

(22) International Filing Date:
26 July 2024 (26.07.2024)

(25) Filing Language:
English

(26) Publication Language:
English

(30) Priority Data:
2311501.7 26 July 2023 (26.07.2023) GB

(71) Applicant: **IMPERIAL COLLEGE INNOVATIONS LTD** [GB/GB]; Level 1 Faculty Building, c/o Imperial College, Exhibition Rd, London SW7 2AZ (GB).

(72) Inventors: **ISALAN, Mark**; 204 Tudor Drive, Kingston upon Thames KT2 5QH (GB). **MIELCAREK, Michal**; 21 Lorne Court, London SW15 6RX (GB).

(74) Agent: **KELTIE LLP**; No. 1 London Bridge, London SE1 9BA (GB).

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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).

(54) Title: ZINC FINGER PEPTIDES, ENCODED NUCLEIC ACIDS, METHODS AND USES

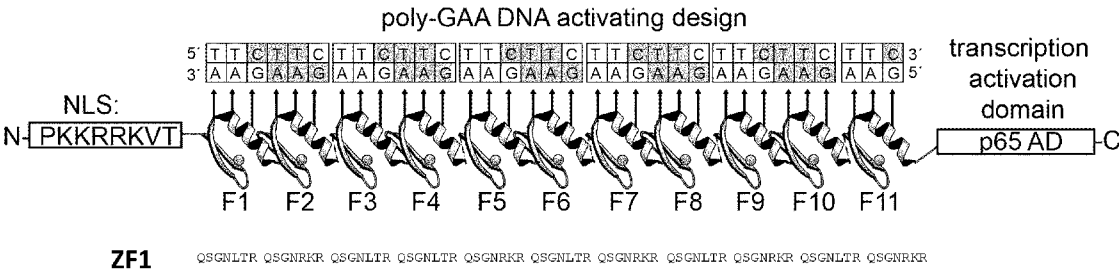


Figure 1A

(57) Abstract: Disclosed herein are polypeptides for use in treating diseases associated with pathogenic genomic repeat sequences, such as neurological disorders. Also disclosed are nucleic acid molecules and vectors that encode such polypeptides. Therapeutic uses and methods for treating such diseases are also disclosed; in particular, therapeutic uses and methods for treating Friedreich's ataxia (FRDA). Also disclosed is a method and associated peptides and nucleic acids for active, long-term delivery of therapeutic molecules to target cells *in vivo* or *in vitro*.

Published:

- *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*
- *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

ZINC FINGER PEPTIDES, ENCODED NUCLEIC ACIDS, METHODS AND USES**FIELD OF THE INVENTION**

5 This invention relates to novel zinc finger peptides and nucleic acids having desirable properties, and to methods and uses for such peptides and nucleic acids. In particular, the invention relates to novel zinc finger encoding nucleic acids or zinc finger peptides for therapeutic uses and methods in the treatment of Friedreich's ataxia (FRDA).

10 BACKGROUND OF THE INVENTION

Neurological disorders are diseases that affect the central nervous system (brain and spinal cord), the peripheral nervous system (peripheral nerves and cranial nerves), and the autonomic nervous system (parts of which are located in both central and peripheral nervous systems). More than 600 neurological
15 diseases have been identified in humans, which together affect all functions of the body, including coordination, communication, memory, learning, eating, and in some cases mortality.

Although many tissues and organs in animals are capable of self-repair, generally the neurological system is not. Therefore, neurological disorders are often characterised by a progressive worsening of
20 symptoms, beginning with minor problems that allow detection and diagnosis, but becoming steadily more severe – often resulting in the death of the affected individual.

While the exact causes or triggers of many neurological disorders are still unknown, for others the causes are well documented and researched. For some of these diseases there are 'effective'
25 treatments, which alleviate symptoms and/or prolong survival. However, despite intense research efforts, for most neurological disorders, and particularly for the most serious diseases, there are still no cures. Hence, there is a need for new therapeutics and treatments for neurological disorders.

Current knowledge of neurological disorders suggests that they can be caused by many different
30 factors, including (but not limited to): inherited genetic abnormalities, problems in the immune system, injury to the brain or nervous system, or diabetes. One known cause of neurological disorder is a genetic abnormality leading to the pathological expansion of nucleic acid repeats sequences, such as: CAG repeats in the *htt* gene that leads to Huntington's disease (HD) (Walker (2007) *Lancet* 369(9557): 218-228; and Kumar *et al. Pharmacol. Rep.* 62(1): 1-14), GGGGCC repeats in the C9ORF72 gene in
35 Amyotrophic lateral sclerosis (ALS) or Frontotemporal dementia (FTD) (DeJesus-Hernandez *et al.* (2011), *Neuron*, 72: 245-56); and GAA repeats in intron 1 of the frataxin gene in Friedreich's ataxia (PMID: 30905359).

Friedreich's ataxia (FRDA) is an autosomal recessive disorder manifesting with progressive ataxia,
40 impaired speech, hearing and vision, cardiomyopathy, diabetes, and skeletal muscle abnormalities (PMID: 29053830). It is a rare, inherited genetic disorder with a prevalence of approx. 1 in 40,000 and there is currently no cure (PMID: 30905359). The disease is caused by the expansion of the nucleotide

triplet GAA, in up to 900 repeating units, within intron 1 of the frataxin locus. Frataxin is important for several cellular processes, and the GAA expansion significantly reduces frataxin mRNA transcripts levels (by up to 90%), which, in turn, leads to down-regulation of the fraxin gene product. Amongst other things, the reduction in frataxin expression causes nerve fibres in the spinal cord and peripheral nervous system to degenerate, becoming thinner and resulting in critical illnesses (PMID: 30905359). While homozygous FRDA patients are fully symptomatic, heterozygous carriers are non-symptomatic and comprise up to 2% of the general population (PMID: 8596916).

To date, available treatments for these and similar diseases, have focussed on trying to control the symptoms, rather than reversing the causes of illness. For example, physical and/or speech therapy can help with normal body functions and activities; walking aids may be used, and surgery may be required to correct physical conditions, such as spinal curvature for foot deformities. Medications are commonly used to treat the heart disease and diabetes that typically accompanies the disease.

In 2023, omaveloxolone (Skyclarys) was approved by the U.S. Food and Drug Administration (FDA) as the first treatment for Friedreich's ataxia. However, it would be highly desirable to have alternative and/or more effective therapeutic molecules and treatments for diseases such as FRDA and related disorders or conditions caused by expanded GAA repeats.

Thus, the present invention seeks to overcome or at least alleviate one or more of the problems found in the prior art.

SUMMARY OF THE INVENTION

The present inventors have identified that by up-regulating repressed mutant gene alleles responsible to onset of disease symptoms, an improved or normal / wild-type function may be restored.

Thus, in general terms, the present invention provides new zinc finger peptides and encoding nucleic acid molecules that can be used for the modulation of gene expression *in vitro* and/or *in vivo*. The new zinc finger peptides of the invention may be particularly useful in the modulation of target genes associated with expanded GAA trinucleotide repeats, and more specifically the targeted activation of such mutant genes.

Beneficially, the invention relates to a new class of zinc finger 'gene switch' to activate the loss-of-function mutation causing Friedreich's ataxia. To our knowledge, zinc finger gene therapies with targeted activators have not previously been developed successfully. For example, the inventors have previously developed a zinc finger technology platform for gene therapy that allows to silence specific gene loci at the DNA level. This technology has yielded very promising results both *in vitro* and in animal models. For example, up to 80% repression of target disease genes in treated organs has been achieved, with repression effects lasting for months (or even years) after a single AAV injection. However, the present invention requires that expression of the disease genes is activated, rather than repressed, restoring a loss-of-function in the pathological state. Accordingly, the present invention

exploits the zinc finger approach to activate the endogenous locus of the frataxin gene, in order to increase frataxin protein levels towards those in wild type cells. Significantly, such gene activation cannot be achieved by any other gene-targeting / binding approach, e.g. antisense oligonucleotides, RNAi, etc, which are repressive.

5

In aspects of the present invention, the inventors have designed zinc finger peptides (ZFPs) to target the GAA-expansion, which may be useful for targeting pathogenic frataxin gene locuses therapeutically. Zinc fingers are DNA-binding proteins that may be reengineered to bind to user-defined DNA-sequences (*Nat. Biotechnol.*, (2001) 19, 656-60). The GAA trinucleotide repeat-targeting zinc finger peptide sequences of the present invention are designed to function in a long single-chain poly-zinc finger protein that may suitably be tuned to bind longer GAA expansions, preferentially using designed binding-destabilising mutations and/or linkers. Second, these constraints are applied within the further constraint of minimising potential epitopes and non-host (mouse, human) residues, in order to increase immunocompatibility in a therapeutic application. The inventors have accordingly devised a formula to

10

15

In various aspects and embodiments of the invention, the new zinc finger peptides (ZFPs) of the invention beneficially bind to expanded GAA trinucleotide repeats associated with mutated pathogenic gene sequences more effectively / efficaciously (e.g. with greater specificity and affinity) than to wild-type GAA trinucleotide repeat sequences associated with non-pathogenic, normal genes. As a consequence, the possibility of more specific gene targeting is envisaged, which may be particularly useful for the modulation of gene expression within the genome and/or for distinguishing between similar nucleic acid sequences of differing lengths. Such ZFPs may particularly up-regulate / activate the expression of target pathogenic genes. In some embodiments, non-target non-pathogenic (WT) genes are not up-regulated / activated or are activated to a much lesser extent than the mutant pathogenic genes. As such, the present invention may provide ZFP activator peptides that discriminate in their binding strength or affinity between WT and pathogenic sequences.

20

25

Thus, the invention relates to therapeutic molecules, molecular combinations and compositions for use in methods for treating neurological diseases, such as Friedreich's ataxia (FRDA). In some aspects and embodiments, the invention is directed to methods and therapeutic treatment regimes for treating patients affected by or diagnosed with FRDA. For example, the therapeutic molecules of the invention may be used in medical treatments in isolation, in combination with other medicaments and in combination with each other. In particular, aspects and embodiments of the invention relate to combination therapies comprising one or more ZFP that up-regulates / activates the expression of target pathogenic gene sequences (a ZFP activator) in combination with one or more additional therapeutic molecule or composition. According to some aspects and embodiments of the invention, the ZFP activator protein binds to and targets the same trinucleotide repeat sequence (GAA) associated with pathogenic alleles, which may also be associated with WT gene alleles; or may bind to specific sequences of the frataxin promoter region, which are common to both wild-type and pathogenic genes.

30

35

40

In any of the aspects and embodiments disclosed herein, the poly-zinc finger peptide is non-natural or may be considered to be engineered / artificial / synthetic; i.e. it is not a naturally occurring poly-zinc finger protein.

- 5 The peptides / proteins of the invention may be useful *in vitro* and/or *in vivo*. In particular, the peptides of the invention may be useful in disease therapy, such as gene therapy; e.g. for delaying the onset of symptoms, and/or for treating or alleviating the symptoms of a disease or diseases; and/or for reducing the severity of or preventing the progression of a disease or diseases. Particular diseases include Friedreich's ataxia (FRDA).

10

In aspects and embodiments, the invention is directed towards novel zinc finger peptides (ZFP) that may exhibit prolonged, mid- to long-term, expression in target organisms *in vivo*, so as to be useful in medical treatments that may require long-term activity of the therapeutic agent. The ZFP sequences of the invention, in some embodiments, are adapted / optimised to closely match endogenous / wild-type peptide sequences expressed in the target organism so as to have reduced toxicity and immunogenicity. Cells expressing the zinc finger peptides of the invention may therefore be protected from the immune response of the target organism so as to prolong expression of the heterologous peptide in these cells.

15

- 20 In aspects and embodiments, therefore, there is provided a polynucleotide encoding a polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within the frataxin gene promoter of SEQ ID NO: 67, or a sequence complementary thereto, and a transcriptional activation domain.

- 25 In aspects and embodiments, there is provided a polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within the frataxin gene promoter of SEQ ID NO: 67, or a sequence complementary thereto, and a transcriptional activation domain.

- In aspects and embodiments, there is provided a polynucleotide encoding a polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within a 5'-GAA-3' trinucleotide repeat sequence, frameshift variants therefore (i.e. 5'-AGA-3' or 5'-AAG-3') or a nucleic acid sequence complementary thereto, and a transcriptional activation domain.

30

- In aspects and embodiments, there is provided a polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within a 5'-GAA-3' trinucleotide repeat sequence, frameshift variants therefore (i.e. 5'-AGA-3' or 5'-AAG-3') or a nucleic acid sequence complementary thereto, and a transcriptional activation domain.

35

In aspects and embodiments, there is provided a vector comprising a polynucleotide of the disclosure.

40

In aspects and embodiments, there is provided a vector capable of expressing a polypeptide of the disclosure.

In aspects and embodiments, there is provided a pharmaceutical composition comprising a polynucleotide, vector or polypeptide according to the disclosure, and a pharmaceutically acceptable carrier.

In aspects and embodiments, there is provided a polynucleotide, vector, polypeptide or pharmaceutical composition for use in treating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal. For example, the disease, disorder or condition may be Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

Polypeptides of the invention may comprise sequences having at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to any of the polypeptides of SEQ ID NOs: 97 to 106, 116 to 121, 147 and 148.

As indicated above, the invention is directed to polynucleotide (or nucleic acid) molecules that encode the zinc finger peptides and polypeptides of the invention. Particularly, isolated polynucleotides are encompassed. In addition, the polynucleotides (or nucleic acid molecules) of the invention may be expression constructs for the expression of the peptide or polypeptide of the invention *in vitro* and/or *in vivo*. The nucleic acids of the invention may be adapted for expression in any desired system or organism, but preferred organisms are mouse – in which therapeutic effects for diseases targeted by the therapeutic polypeptides of the invention may be tested, and humans – which will likely be the ultimate recipients or any potential therapy.

For expression of polypeptides, nucleic acid molecules are conveniently inserted into a vector or plasmid. Vectors and plasmids may be adapted for replication (e.g. to produce large quantities of its own nucleic acid sequence in host cells), or may be adapted for protein expression (e.g. to produce large or suitable quantities of zinc finger-containing protein in host cells). Any vector may be used, but preferred are polypeptide expression vectors so that the encoded polypeptide is expressed in host cells (e.g. for purposes of therapeutic treatment). Advantageously, the vector comprises a beneficial long acting, tissue specific and/or (very) strong promoter / enhancer sequence such as pNSE, pHsp90, CBh, EF1 α -1 or synapsin, as described herein.

Viral vectors are particularly useful for potential use in therapeutic applications due to their ability to target and/or infect specific cell types. Suitable viral vectors may include those derived from retroviruses (such as influenza, SIV, HIV, lentivirus, and Moloney murine leukaemia); adenoviruses; adeno-associated viruses (AAV); herpes simplex virus (HSV); and chimeric viruses. Adeno-associated virus (AAV) vectors are considered particularly useful for targeting therapeutic peptides to the central and peripheral nervous systems and to the brain. A preferred viral vector delivery system is based on the AAV2/1 and AAV2/9 viral subtypes.

Thus, the invention is particularly directed to an adeno-associated virus (AAV) vector comprising a nucleic acid expression construct capable of expressing at least one polypeptide comprising a zinc

finger peptide, wherein the polypeptide and the zinc finger peptide are defined as disclosed herein. The invention is also, therefore, directed to a gene therapy method; as well as to methods for treating diseases; particularly neurological diseases, such as FRDA.

5 In some embodiments of the methods and therapeutic uses of the invention more than one (e.g. two) nucleic acid construct may be administered sequentially, simultaneously or separately to a cell or patient to be treated. Each nucleic acid construct may encode one or more ZFP according to the invention, so as to cause two or more complementary ZFPs to be expressed, advantageously within the same cell.

10 The invention relates to polypeptides comprising zinc finger peptides as defined herein. Typically, the polypeptides of the invention include a zinc finger portion comprising a plurality of zinc finger domains and one or more beneficial auxiliary sequences, such as effector domains. Effector domains include nuclear localisation sequences and transcriptional activation domains as described elsewhere herein. In embodiments, other effector domains may also be used. Thus, in aspects and embodiments, the
15 invention relates to chimeric or fusion proteins comprising the zinc finger peptides of the invention conjugated to one or more non-zinc finger domain, such as an effector domain.

Conveniently, the ZFPs according to various aspects and embodiments of the invention bind double-stranded trinucleotide repeat sequences comprising GAA-repeat, AGA-repeat and/or AAG-repeat
20 sequences. In preferred embodiments, the ZFPs of the invention target and bind to 5'-GAA-3' or 5'-AAG-3' repeat sequences.

In some aspects and embodiments, ZFPs according to the invention bind double-stranded GAA-repeat sequences containing at least 66 trinucleotide repeats, at least 100 trinucleotide repeats, at least 200
25 trinucleotide repeats, at least 300 trinucleotide repeats, at least 400 trinucleotide repeats, at least 500 trinucleotide repeats, at least 600 trinucleotide repeats, at least 700 trinucleotide repeats, at least 800 trinucleotide repeats, at least 900 trinucleotide repeats, or at least 1000 trinucleotide repeats. In embodiments, ZFPs according to these aspects and embodiments of the invention preferentially bind double-stranded trinucleotide repeat sequences containing between about 66 and 2000 trinucleotide
30 repeats, between about 100 and 1,600 trinucleotide repeats, or between about 300 and 1,200 hexanucleotide repeats. Such nucleic acid sequences are beneficially bound with a binding dissociation constant (Kd) of less than about 1 μ M, less than about 100 nM, less than about 10 nM, or less than about 1 nM. Suitably, ZFPs according to these embodiments of the invention bind to such double-stranded trinucleotide repeat sequences preferentially over double-stranded hexanucleotide repeat
35 sequences containing less than 40 trinucleotide repeats, less than 30 hexanucleotide repeats, and particularly over double-stranded trinucleotide repeat sequences containing up to 12 trinucleotide repeats.

Polypeptides of the invention may also be administered to an individual or patient in need thereof.
40 Suitably, the polypeptides of the invention are to treat neurodegenerative diseases; particularly diseases associated with expanded trinucleotide repeat sequences, such as FRDA.

A gene therapy method according to the invention may comprise administering to a person in need thereof or to cells previously removed from a person, a nucleic acid encoding a ZFP of the invention, and causing the polypeptide to be expressed in cells of the person / subject. In this way, the gene therapy method may be useful for treating a neurodegenerative disease; and particularly diseases associated with expanded trinucleotide repeat sequences, such as FRDA. Suitably, the ZFP is a ZFP activator protein. In embodiments of these aspects of the invention, the method comprises administering more than one nucleic acid expression construct, each encoding a ZFP of the invention, and causing the ZFPs to be expressed in cells of the subject to be treated. The ZFPs may comprise a complementary pair of ZFPs, one of which is a ZFP activator that targets and binds to a pathogenic GAA-trinucleotide repeat sequence associated with the target gene, and one of which is a ZFP activator which targets and binds to a wild-type promoter sequence region of the target gene. In such embodiments, the two activator proteins of the pair suitably cooperate in order to enhance expression of the target gene, so as to return transcript (and/or protein) levels of the target gene in a pathogenic cell to at least 50% of the expression level in a wild-type cell. In embodiments, the expression level of the target gene (e.g. frataxin) may be increased to at least 60%, at least 70%, at least 80%, at least 90%, or at least 100% of the expression level in a wild-type, non-diseased cell. In some embodiments, the expression level of the target transcript or protein may be over 100% of the level in a wild-type cell. In some embodiments, the method or therapeutic use comprises administering one nucleic acid encoding two (or more) ZFPs according to the invention. In other methods of therapeutic uses more than one nucleic acid / expression construct of the invention may be used. In such embodiments, the two or more nucleic acids or polypeptides may be administered simultaneous, sequentially or separately.

It will be appreciated that the invention encompasses any polypeptides that may be encoded by the nucleic acid molecules defined herein; and any nucleic acid molecules encoding a polypeptide as defined herein.

Pharmaceutical composition of the invention may comprise nucleic acid molecules (such as vectors) and/or polypeptides as defined herein. It is envisaged that the pharmaceutical compositions of the invention may be used in a method of combination therapy with one or more additional therapeutic agent, may be used on their own, or may be used in combination with other compositions of the invention and optionally one or more additional therapeutic agent.

Some aspects and embodiments of the invention include formulations, medicaments and pharmaceutical compositions comprising the zinc finger peptides. In some embodiments, the invention relates to a zinc finger peptide for use in medicine. More specifically, the zinc finger peptides and therapeutics of the invention may be used for modulating the expression of a target gene in a cell. In some embodiments the target gene is the frataxin gene in Friedreich's ataxia (FRDA). Particularly, in these aspects and embodiments the invention relates to the treatment of diseases or conditions associated with the expanded GAA trinucleotide repeat and/or expression of gene products associated with such repeat sequences. Treatment may also include preventative as well as therapeutic treatments and alleviation of a disease or condition.

Beneficially, nucleic acid expression constructs according to the invention are suitable for sustained constitutive expression of ZFPs. Accordingly, nucleic acid sequences encoding ZFPs may be operably linked / associated with promoter sequences suitable for such sustained expression *in vivo*. Sustained expression is beneficially for a period of at least 3 weeks, at least 6 weeks, at least 12 weeks or at least 24 weeks. In the context of this invention, 'promoter' sequences may encompass both transcriptional promoter and enhancer elements within a nucleic acid sequence which have the effect of enabling, causing and/or enhancing transcription of an associated gene / nucleic acid construct. In other words, the use of the term 'promoter' does not exclude the possibility that the nucleic acid sequence concerned may also encompass other elements associated with transcription, such as enhancer elements.

Gene therapy methods are also disclosed, comprising administering to a subject in need thereof or to cells previously removed from the subject, a nucleic acid encoding one or more ZFP of the invention under the control of natural or synthetic promoter-enhancer sequences, and causing the polypeptide to be expressed in cells of the subject.

Thus, in embodiments there is provided a gene therapy method comprising administering to a subject in need thereof, or to cells previously removed from the subject, a vector comprising a pNSE, pHsp90, CBh, EF1 α -1 or synapsin promoter-enhancer construct. In embodiments, the methods comprise administering to the subject to be treated (or to cells of the subject) a vector according to the invention with neuronal targeting specificity in combination with a promiscuous vector according to the invention. The method may comprise administering to the subject to be treated an AAV2/1 subtype adeno-associated virus (AAV) vector according to the invention in combination with an AAV2/9 subtype adeno-associated virus (AAV) vector according to the invention. The administering 'in combination' may be simultaneous, separate or sequential, as appropriate. Therapeutic uses of the constructs and viral vectors of the invention are also encompassed. The methods and constructs of the invention may be for treating a neurological disease or condition; particularly Friedreich's ataxia (FRDA).

It will be appreciated that any features of one aspect or embodiment of the invention may be combined with any combination of features in any other aspect or embodiment of the invention, unless otherwise stated, and such combinations fall within the scope of the claimed invention.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention is further illustrated by the accompanying drawings in which:

Figure 1 (A) Schematic illustration of an 11-zinc finger activator protein according to the invention, showing recognition helices from the adjacent zinc finger domains contacting 5'-GAA-3' trinucleotide repeats on the lower DNA strand. A nuclear localisation signal (NLS) is provided at the N-terminus and a transcription activator domain is located at the C-terminus. The amino acid sequences of representative DNA recognition helices from each of the zinc finger domains are displayed below the zinc finger arrays, showing that the primary, specific nucleic acid-binding amino acids at the -1, 3 and 6 positions of each zinc finger alpha-helix domain are Q, N and R, respectively for optimal interaction

and binding specificity to the 5-GAA-3' trinucleotide repeat sequence. Similar arrays comprising from 6 to 32 zinc fingers, for example, 6, 8, 10, 12 and 18 zinc finger domains can be built. (B) Schematic illustration of an 11-zinc finger activator protein according to the invention, showing recognition helices from the adjacent zinc finger domains contacting 5'-AAG-3' trinucleotide repeats (of an expanded GAA repeat sequence) on the lower DNA strand. A nuclear localisation signal (NLS) is provided at the N-terminus and a transcription activator domain is located at the C-terminus. The amino acid sequences of representative DNA recognition helices from each of the zinc finger domains are displayed below the zinc finger arrays, showing that the primary, specific nucleic acid-binding amino acids at the -1, 3 and 6 positions of each zinc finger alpha-helix domain are R, N and Y/Q, respectively for optimal interaction and binding specificity to the 5-AAG-3' trinucleotide repeat sequence. Similar arrays comprising from 6 to 32 zinc fingers, for example, 6, 8, 9, 10, 12 and 18 zinc finger domains can be built. (C) Schematic illustration of an 9-zinc finger activator protein (ZF3) according to the invention for binding to the wild-type promoter region of human frataxin, showing recognition helices from the adjacent zinc finger domains contacting the 5'-GGGAGGCAGAGCTTGCACTGAGCCGAG-3' sequence on the lower DNA strand. A nuclear localisation signal (NLS) is provided at the N-terminus and a transcription activator domain is located at the C-terminus. The amino acid sequences of representative DNA recognition helices from each of the zinc finger domains are displayed below the zinc finger arrays. Similar arrays comprising from 6 to 32 zinc fingers, for example, 6, 8, 10, 11, 12 and 18 zinc finger domains can be built. (D) Schematic illustration of a 6-zinc finger activator protein (ZF4) according to the invention for binding to the wild-type promoter region of human frataxin, showing recognition helices from the adjacent zinc finger domains contacting the 5'-AGCTGGGTGTGGTGGTGC-3' sequence on the lower DNA strand. A nuclear localisation signal (NLS) is provided at the N-terminus and a transcription activator domain is located at the C-terminus. The amino acid sequences of representative DNA recognition helices from each of the zinc finger domains are displayed below the zinc finger arrays. Similar arrays comprising from 8 to 32 zinc fingers, for example, 8, 9, 10, 11, 12 and 18 zinc finger domains can be built. In the embodiments of (A) to (D), for optimal use in human cells, the NLS is a human NLS, e.g. PKKRRKVT (human protein KIAA2022, SEQ ID NO: 108), and the transcriptional activator domain is from a human activator protein, e.g. from human p65. Alternatively, in mice, a mouse primase NLS may be used, such as RIRKKLR (mouse primase p58 NLS9, SEQ ID NO: 109).

Figure 2 Graph showing zinc finger activator peptide mediated upregulation of a luciferase reporter construct in an *in vitro* assay. A reporter plasmid was constructed containing the luciferase gene operably linked downstream of a 2.5 kb section of the wild-type frataxin gene promoter sequence (i.e. from 0 to 2,500 bases upstream of the frataxin transcription start site). Zinc finger activator peptides containing either the ZF3 or ZF4 zinc finger domain arrays – engineered to bind to the wild-type frataxin gene promoter – were expressed in cells containing the luciferase reporter construct and luciferase expression was measured at 48 hours and 72 hours post-transfection. Data for 72 hours is shown. Data at 48 hours was comparable. Columns from left to right as follows: 'negative control' = no zinc finger peptide expression vector, no transfection reagents; 'MOCK' = transfection reagents only; '2.5kb promoter only' =, reporter construct containing 2.5kb frataxin promoter sequence with luciferase gene; 'ZF3p65' = ZF3-p65 expression vector only; 'ZF3P64' = ZF3-VP64 expression vector only; 'ZF4p65' = ZF4-p65 expression vector only; 'ZF4VP64' = ZF4-VP64 expression vector only; 'ZF3p65/2.5kb

promoter' = ZF3-p65 expression vector and luciferase vector operably linked to 2.5kb frataxin promoter sequence; 'ZF3P64/2.5kb promoter' = ZF3-P64 expression vector and luciferase vector operably linked to 2.5kb frataxin promoter sequence; 'ZF4p65/2.5kb reporter' = ZF4-p65 expression vector and luciferase vector operably linked to 2.5kb frataxin promoter sequence; 'ZF4VP64/2.5kb promoter' = ZF4-VP64 expression vector and luciferase vector operably linked to 2.5kb frataxin promoter sequence. Results are based on 8 transfections per condition; 10,000 cells were seeded for each assay. Data demonstrate that ZF3 activator peptides linked to either the P65 or VP64 activation domains are able to activate luciferase gene expression to similar levels in the *in vitro* reporter assays.

Figure 3 Graph showing ZF1 and ZF2-based zinc finger activator peptide mediated activation of the pathogenic human frataxin (FXN) gene locus in each of two FRDA-model human fibroblast cell lines, GM03816 (which contained 300 and 500 GAA repeats, respectively, on two alleles) and GM04078 (which contained 600 and 850 GAA repeats, respectively, on two alleles). Measurements based on RNA transcript levels. Control cells contained wild-type frataxin gene; zinc finger activator peptides were tested containing either the p65 or the VP64 transcriptional activation domain. Columns from left to right as follows: 'control WT/MOCK' = transfection reagents, no zinc finger activator peptide, wild-type cell line; 'control WT/ZF1p65' = ZF1-p65 zinc finger activator peptide expression vector, wild-type cell line; 'control WT/ZF1VP64' = ZF1-VP64 zinc finger activator peptide expression vector, wild-type cell line;; 'control WT/ZF2p65' = ZF2-p65 zinc finger activator peptide expression vector, wild-type cell line;; 'control WT/ZF2VP64' = ZF2-VP64 zinc finger activator peptide expression vector, wild-type cell line;; 'GM03816/MOCK' = GM03816 FRDA pathogenic FXN disease model, transfection reagents without zinc finger activator peptide; 'GM03816/ZF1p65' = GM03816 FRDA pathogenic FXN disease model, ZF1-p65 zinc finger activator peptide; 'GM03816/ZF1VP64' = GM03816 FRDA pathogenic FXN disease model, ZF1-VP64 zinc finger activator peptide; 'GM03816/ZF2p65' = GM03816 FRDA pathogenic FXN disease model, ZF2-p65 zinc finger activator peptide; 'GM03816/ZF2VP64' = GM03816 FRDA pathogenic FXN disease model, ZF2-VP64 zinc finger activator peptide; 'GM04078/MOCK' = GM04078 FRDA pathogenic FXN disease model, transfection reagents without zinc finger activator peptide; 'GM04078/ZF1p65' = GM04078 FRDA pathogenic FXN disease model, ZF1-p65 zinc finger activator peptide; 'GM04078/ZF1VP64' = GM04078 FRDA pathogenic FXN disease model, ZF1-VP64 zinc finger activator peptide; 'GM04078/ZF2p65' = GM04078 FRDA pathogenic FXN disease model, ZF2-p65 zinc finger activator peptide; 'GM04078/ZF2VP64' = GM04078 FRDA pathogenic FXN disease model, ZF2-VP64 zinc finger activator peptide. Data were normalised to 18S, GAPDH, HPRT. Results are an average of 5 repeats. Data demonstrate that ZF2 activator peptides linked to either the P65 or VP64 activation domains are able to upregulate frataxin gene expression from pathogenic FRDA models to above normal levels of frataxin gene transcript.

Figure 4 Graph showing ZF3 and ZF4-based zinc finger activator peptide mediated activation of the pathogenic human frataxin (FXN) gene locus in each of two FRDA-model human fibroblast cell lines, GM03816 (which contained 300 and 500 GAA repeats, respectively, on the two alleles) and GM04078 (which contained 600 and 850 GAA repeats, respectively, on the two alleles). Measurements based on RNA transcript levels. Control cells contained wild-type frataxin gene; zinc finger activator peptides were tested containing either the p65 or the VP64 transcriptional activation domain. Columns

from left to right as follows: 'control WT/MOCK' = transfection reagents, wild-type cell line;; no zinc finger activator peptide; 'control WT/ZF3p65' = ZF3-p65 zinc finger activator peptide expression vector, wild-type cell line;; 'control WT/ZF3VP64' = ZF3-VP64 zinc finger activator peptide expression vector, wild-type cell line;; 'control WT/ZF4p65' = ZF4-p65 zinc finger activator peptide expression vector, wild-type cell line;; 'control WT/ZF4VP64' = ZF4-VP64 zinc finger activator peptide expression vector, wild-type cell line;; 'GM03816/MOCK' = GM03816 FRDA pathogenic FXN disease model, transfection reagents without zinc finger activator peptide; 'GM03816/ZF3p65' = GM03816 FRDA pathogenic FXN disease model, ZF3-p65 zinc finger activator peptide; 'GM03816/ZF3VP64' = GM03816 FRDA pathogenic FXN disease model, ZF3-VP64 zinc finger activator peptide; 'GM03816/ZF4p65' = GM03816 FRDA pathogenic FXN disease model, ZF4-p65 zinc finger activator peptide; 'GM03816/ZF4VP64' = GM03816 FRDA pathogenic FXN disease model, ZF4-VP64 zinc finger activator peptide; 'GM04078/MOCK' = GM04078 FRDA pathogenic FXN disease model, transfection reagents without zinc finger activator peptide; 'GM04078/ZF3p65' = GM04078 FRDA pathogenic FXN disease model, ZF3-p65 zinc finger activator peptide; 'GM04078/ZF3VP64' = GM04078 FRDA pathogenic FXN disease model, ZF3-VP64 zinc finger activator peptide; 'GM04078/ZF4p65' = GM04078 FRDA pathogenic FXN disease model, ZF4-p65 zinc finger activator peptide; 'GM04078/ZF4VP64' = GM04078 FRDA pathogenic FXN disease model, ZF4-VP64 zinc finger activator peptide. Data were normalised to 18S, GAPDH, HPRT. Results are an average of 5 repeats. Data demonstrate that ZF3 activator peptides linked to either the P65 or VP64 activation domains are able to upregulate frataxin gene expression from pathogenic FRDA models to normal levels of frataxin gene transcript. ZF4 activator peptides linked to either the P65 or VP64 activation domains are able to upregulate frataxin gene expression from pathogenic FRDA models by approx. 20% in these assays.

Figure 5 Graphs demonstrating that ZF2a activates the faulty frataxin locus *in vivo* in a Friedreich's ataxia mouse model (YG8-800). 8 weeks old mice were injected with AAVs carrying a gene expressing ZF2a either by intra-theal injection (A-C; AAV2/9, measured 3 weeks after injection) or tail vein injection (D-F; AAV-PHP.eB measured 8 days after injection). Frataxin protein levels were significantly reactivated in all tested tissues: **(A)** cerebellum; **(B)** spinal cord; **(C)** cerebrum; **(D)** heart; **(E)** cerebellum; and **(F)** quadriceps. Transcript levels were measured by Taq-man qPCR while an ELISA assay specific to human frataxin protein was employed for protein quantification. RNA transcripts or protein assays are indicated on y-axes: ZFP samples (dark grey); vehicle (Veh) samples (light grey). Taqman qPCR values were normalised to the geometric mean of three housekeeping genes: 18S rRNA, GAPDH and HPRT. Error bars are $\pm$ SEM (n = 5). Student's t-test: ***p<0.001. IT – intra-theal injection; IV intra venous (tail vein) injections.

Figure 6 Graphs demonstrating that ZF2a does not activate inflammatory or immunoresponse markers *in vivo*, in a Friedreich's ataxia mouse model (YG8-800). 8 weeks old mice were injected with AAVs carrying a gene expressing ZF2a either by intra-theal injection (A-C AAV2/9, measured 3 weeks after injection) or tail vein injection (D-F; AAV-PHP.eB measured 8 days after injection). A number of marker genes were not altered in the mice treated with ZF2a in comparison to vehicle (Veh) treated mice: **(A)** cerebellum; **(B)** spinal cord; **(C)** cerebrum; **(D)** heart; **(E)** cerebellum; and **(F)** quadriceps. Transcript levels were measured by Taq-man qPCR. All Taq-man qPCR values were normalised to the

geometric mean of three housekeeping genes: 18S rRNA, GAPDH and HPRT. Error bars are $\pm$ SEM (n = 5). Student's t-test: all n.s.; IT – intra-thecal injection; IV intra-venous (tail vein) injections. *GFAP* – Glial Fibrillary acidic protein; *CXCL10* – C-X-C motif chemokine ligand 10; *IL1B* – Interleukin 1 beta; *DDX58* – DEAD (Asp-Glu-Ala-Asp) box polypeptide 58; *TNF α* – Tumor necrosis factor alpha.

Figure 7 (A) Schematic representation showing the model of active delivery in an *in vivo* system – e.g. in the target brain (1). A therapeutic peptide is delivered to a first population of target cells (2) using a suitable delivery system (e.g. such as a viral delivery vector); therapeutic peptide is expressed and secreted from the first population of target cells; and secreted therapeutic peptide diffuses within the in vivo system coming into contact with a second population of target cells (3); cell penetration of the secreted therapeutic peptide allows the therapeutic effect to take effect in both the first (2) and second (3) populations of target cells. In this way, delivery of therapeutic peptide to relatively small first populations of target cells (2, 4) can enable therapeutic effect in a relatively larger population of target cells (3, 5) (Key: 1 = target brain; 2 = therapeutic viral delivery site A: therapeutic peptide expressed in viral-infected cells; 3 = diffusion volume of secretable therapeutic peptide expressed at site A; 4 = therapeutic viral delivery site B: secretable therapeutic peptide expressed in viral-infected cells; and 5 = diffusion volume of secretable therapeutic peptide expressed at site B); and (B) schematic illustration showing hypothetical deliver of therapeutic peptide via 'active delivery' in neuronal cells: step (1) infection with AAV-ZF; step (2) ZF secretion; (3) ZF cell penetration (Key: 6 = microglia; 7 = oligodendrocytes; 8 = myelin sheath; 9 = neuron; 10 = dendrite; 11 = synapse; 12 = axon; 13 = astrocyte).

DETAILED DESCRIPTION OF THE INVENTION

All references cited herein are incorporated by reference in their entirety. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs (e.g. in cell culture, molecular genetics, nucleic acid chemistry and biochemistry).

Unless otherwise indicated, the practice of the present invention employs conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA technology, chemical methods, pharmaceutical formulations and delivery and treatment of animals, which are within the capabilities of a person of ordinary skill in the art. Such techniques are also explained in the literature, for example, J. Sambrook, E. F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Books 1-3, Cold Spring Harbor Laboratory Press; Ausubel, F. M. *et al.* (1995 and periodic supplements; *Current Protocols in Molecular Biology*, ch. 9, 13, and 16, John Wiley & Sons, New York, N. Y.); B. Roe, J. Crabtree, and A. Kahn, 1996, *DNA Isolation and Sequencing: Essential Techniques*, John Wiley & Sons; J. M. Polak and James O'D. McGee, 1990, *In Situ Hybridisation: Principles and Practice*, Oxford University Press; M. J. Gait (Editor), 1984, *Oligonucleotide Synthesis: A Practical Approach*, IRL Press; and D. M. J. Lilley and J. E. Dahlberg, 1992, *Methods of Enzymology: DNA Structure Part A: Synthesis and Physical Analysis of DNA Methods in Enzymology*, Academic Press. Each of these general texts is herein incorporated by reference.

In order to assist with the understanding of the invention several terms are defined herein.

The term 'amino acid' in the context of the present invention is used in its broadest sense and is meant to include naturally occurring L α -amino acids or residues. The commonly used one and three letter abbreviations for naturally occurring amino acids are used herein: A=Ala; C=Cys; D=Asp; E=Glu; F=Phe; G=Gly; H=His; I=Ile; K=Lys; L=Leu; M=Met; N=Asn; P=Pro; Q=Gln; R=Arg; S=Ser; T=Thr; V=Val; W=Trp; and Y=Tyr (Lehninger, A. L., (1975) *Biochemistry*, 2d ed., pp. 71-92, Worth Publishers, New York). The general term 'amino acid' further includes D-amino acids, retro-inverso amino acids as well as chemically modified amino acids such as amino acid analogues, naturally occurring amino acids that are not usually incorporated into proteins such as norleucine, and chemically synthesised compounds having properties known in the art to be characteristic of an amino acid, such as β -amino acids. For example, analogues or mimetics of phenylalanine or proline, which allow the same conformational restriction of the peptide compounds as do natural Phe or Pro, are included within the definition of amino acid. Such analogues and mimetics are referred to herein as 'functional equivalents' of the respective amino acid. Other examples of amino acids are listed by Roberts and Vellaccio, *The Peptides: Analysis, Synthesis, Biology*, Gross and Meiehofer, eds., Vol. 5 p. 341, Academic Press, Inc., N.Y. 1983, which is incorporated herein by reference.

The term 'peptide' as used herein (e.g. in the context of a zinc finger peptide (ZFP) or framework) refers to a plurality of amino acids joined together in a linear or circular chain. term oligopeptide is typically used to describe peptides having between 2 and about 50 or more amino acids. Peptides larger than about 50 amino acids are often referred to as polypeptides or proteins. For purposes of the present invention, however, the term 'peptide' is not limited to any particular number of amino acids, and is used interchangeably with the terms 'polypeptide' and 'protein'.

As used herein, the term 'zinc finger domain' refers to an individual 'finger', which comprises a $\beta\beta\alpha$ -fold stabilised by a zinc ion (as described elsewhere herein). Each zinc finger domain typically includes approximately 30 amino acids. The term 'domain' (or 'module'), according to its ordinary usage in the art, refers to a discrete continuous part of the amino acid sequence of a polypeptide that can be equated with a particular function. Zinc finger domains are largely structurally independent and may retain their structure and function in different environments. Typically, a zinc finger domain binds a triplet or (overlapping) quadruplet nucleotide sequence. Adjacent zinc finger domains arranged in tandem are joined together by linker sequences. A zinc finger peptide of the invention is composed of a plurality of 'zinc finger domains', which in combination do not exist in nature. Therefore, they may be considered to be artificial, synthetic or engineered zinc finger peptides.

The terms 'nucleic acid', 'polynucleotide', and 'oligonucleotide' are used interchangeably and refer to a deoxyribonucleotide (DNA) or ribonucleotide (RNA) polymer, in linear or circular conformation, and in either single- or double-stranded form. For the purposes of the present invention such DNA or RNA polymers may include natural nucleotides, non-natural or synthetic nucleotides, and mixtures thereof. Non-natural nucleotides may include analogues of natural nucleotides, as well as nucleotides that are

modified in the base, sugar and/or phosphate moieties (e.g. phosphorothioate backbones). Examples of modified nucleic acids are PNAs and morpholino nucleic acids. Generally, an analogue of a particular nucleotide has the same base-pairing specificity, i.e. an analogue of G will base-pair with C. For the purposes of the invention, these terms are not to be considered limiting with respect to the length of a polymer.

A 'gene', as used herein, is the segment of nucleic acid (typically DNA) that is involved in producing a polypeptide or ribonucleic acid gene product. It includes regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons). Conveniently, this term also includes the necessary control sequences for gene expression (e.g. enhancers, silencers, promoters, terminators etc.), which may be adjacent to or distant to the relevant coding sequence, as well as the coding and/or transcribed regions encoding the gene product. Preferred genes in accordance with the present invention are those associated with neurological disease conditions; particularly those exhibiting aberrant hexanucleotide repeat sequences, such as mutant *C9Orf72* genes.

As used herein the term 'modulation', in relation to the expression of a gene refers to a change in the gene's activity. Modulation includes both activation (i.e. increase in activity or expression level) and repression (i.e. reduction or inhibition) of gene activity. In some embodiments of the invention, the therapeutic molecules (e.g. peptides) of the invention are repressors of gene expression or activity; in some embodiments of the invention, the therapeutic molecules (e.g. peptides) of the invention are activators of gene expression or activity.

A nucleic acid 'target', 'target site' or 'target sequence', as used herein, is a nucleic acid sequence to which a zinc finger peptide of the invention will bind, provided that conditions of the binding reaction are not prohibitive. A target site may be a nucleic acid molecule or a portion of a larger polynucleotide. Particularly suitable target sites comprise repetitive nucleic acid sequences; especially trinucleotide or hexanucleotide repeat sequences. Preferred target sequences in accordance with the invention include those defined by GAA-repeat sequence (e.g. GAAGAAGAA...; AGAAGAAGA...; and AAGAAGAAG...), and their complementary sequences. In accordance with the invention, a target sequence for a poly-zinc finger peptide of the invention may comprise a single contiguous nucleic acid sequence, or more than one non-contiguous nucleic acid sequence (e.g. two separate contiguous sequences, each representing a partial target site), which are separated by one or more intervening nucleotide or sequence of nucleotides. These terms may also be substituted or supplemented with the terms 'binding site', 'binding sequence', or 'recognition site', which are used interchangeably.

As used herein, 'binding' in the context of the present invention refers to a non-covalent interaction between macromolecules (e.g. between a zinc finger peptide and a nucleic acid molecule containing an appropriate target site). In some cases, binding will be sequence-specific, such as between one or more specific nucleotides (or base pairs) and one or more specific amino acids. It will be appreciated, however, that not all components of a binding interaction need be sequence-specific (e.g. non-covalent interactions with phosphate residues in a DNA backbone). Binding interactions between a nucleic acid

sequence and a zinc finger peptide of the invention may be characterised by binding affinity and/or dissociation constant (K_d). A suitable dissociation constant for a zinc finger peptide of the invention binding to its target site may be in the order of 1 μ M or lower, 1 nM or lower, or 1 pM or lower, as described elsewhere herein. 'Affinity' refers to the strength of binding, such that increased binding affinity correlates with a lower K_d value. Zinc finger peptides may have DNA-binding activity, RNA-binding activity, and/or even protein-binding activity. Generally, the zinc finger peptides of the invention are designed or selected to have sequence specific nucleic acid-binding activity, especially to dsDNA. Typically, the target site for a particular zinc finger peptide is a sequence to which the zinc finger peptide concerned is capable of nucleotide-specific binding. It will be appreciated, however, that depending on the amino acid sequence of a zinc finger peptide it may bind to or recognise more than one target sequence, although typically one sequence will be bound in preference to any other recognised sequences, depending on the relative specificity of the individual non-covalent interactions. In some beneficial embodiments, zinc finger activator peptides of the invention may target unique nucleic acid sequences, e.g. in the frataxin promoter region. In such cases, high, specific binding affinity for the unique target sequence is advantageous. Generally, specific binding is preferably achieved with a dissociation constant (K_d) of 1 μ M or lower, 1 nM or lower, 100 pM or lower; or 10 pM or lower. In some embodiments, particularly as regards zinc finger activator proteins of the invention that target pathogenic repeat sequences, binding affinity for a target site may be deliberately weakened (reduced) such that a zinc finger activator protein of the invention may bind preferentially to expanded, pathogenic-repeat sequences in FRDA, e.g. comprising 300 or more, 500 or more, 600 or more, or 850 or more repeat sequences; as compared to shorter trinucleotide repeat sequences, e.g. comprising less than 100, less than 50, less than 20 or between 3 and 12 trinucleotide repeat sequences. In some embodiments, therefore, a zinc finger peptide of the invention may bind a target sequence with a dissociation constant that is weaker than about 100 pM, weaker than 1 nM, weaker than 10 nM, or weaker than 100 nM.

By 'non-target' it is meant that the nucleic acid sequence concerned is not appreciably bound by the relevant zinc finger peptide. In some embodiments, it may be considered that, where a zinc finger peptide of the invention has a known sequence-specific target sequence, essentially all other nucleic acid sequences may be considered to be non-target. From a practical perspective it can be convenient to define an interaction between a non-target sequence and a particular zinc finger peptide as being sub-physiological (i.e. not capable of creating a physiological response under physiological target sequence / zinc finger peptide concentrations). For example, if any binding can be measured between the zinc finger peptide and the non-target sequence, the dissociation constant (K_d) is typically weaker than 1 μ M, such as 10 μ M or weaker, 100 μ M or weaker, or at least 1 mM.

Zinc Finger Peptides

A 'zinc finger' is a relatively small polypeptide domain comprising approximately 30 amino acids, which folds to form a secondary structure including an α -helix adjacent an antiparallel β -sheet (known as a $\beta\beta\alpha$ -fold). The fold is stabilised by the co-ordination of a zinc ion between four largely invariant (depending on zinc finger framework type) Cys and/or His residues, as described further below. Natural

zinc finger domains have been well studied and described in the literature, see for example, Miller *et al.*, (1985) *EMBO J.* 4: 1609-1614; Berg (1988) *Proc. Natl. Acad. Sci. USA* 85: 99-102; and Lee *et al.*, (1989) *Science* 245: 635-637. A zinc finger domain typically recognises and binds to a nucleic acid triplet, or an overlapping quadruplet (as explained below), in a double-stranded DNA target sequence.

- 5 However, zinc fingers are also known to bind RNA and proteins (Clemens, K. R. *et al.* (1993) *Science* 260: 530-533; Bogenhagen, D. F. (1993) *Mol. Cell. Biol.* 13: 5149-5158; Searles, M. A. *et al.* (2000) *J. Mol. Biol.* 301: 47-60; Mackay, J. P. & Crossley, M. (1998) *Trends Biochem. Sci.* 23: 1-4).

- 10 Zinc finger proteins generally contain strings or chains of zinc finger domains (or modules). Thus, a natural zinc finger protein may include two or more zinc finger domains, which may be directly adjacent one another, e.g. separated by a short (canonical) or canonical-like linker sequence; or a longer, flexible or structured polypeptide sequence. Adjacent zinc finger domains linked by short canonical or canonical-like linker sequences of 5, 6 to 7 amino acids are expected to bind to contiguous nucleic acid sequences, i.e. they typically bind to adjacent trinucleotides / triplets; or protein structures. In some cases, cross-binding may also occur between adjacent zinc fingers and their respective target triplets, which helps to strengthen or enhance the recognition of the target sequence, and leads to the binding of overlapping quadruplet sequences (Isalan *et al.*, (1997) *Proc. Natl. Acad. Sci. USA*, 94: 5617-5621). By comparison, distant zinc finger domains within the same poly-zinc finger protein may recognise (or bind to) non-contiguous nucleic acid sequences or even to different molecules (e.g. protein rather than nucleic acid). Indeed, naturally occurring zinc finger-containing proteins may include both zinc finger domains for binding to protein structures as well as zinc finger domains for binding to nucleic acid sequences.

- 25 In accordance with the invention, adjacent zinc finger domains of the same zinc finger peptide may be separated by relatively long, flexible linker sequences. Such adjacent zinc fingers can readily bind to non-contiguous nucleic acid sequences, although it is also possible for them to bind to contiguous sequences. In such embodiments, the relative binding location of the pairs of zinc finger domains separated by long linker sequences may be determined by the sequence context, i.e. by dominant binding interactions from other zinc finger domains within the peptide.

- 30 The majority of the amino acid side chains in a zinc finger domain that are important for dsDNA base recognition are located on the α -helix of the finger. Conveniently, therefore, the amino acid positions in a zinc finger domain are numbered from the first residue in the α -helix, which is given the number (+)1; and the helix is generally considered to end at the final zinc-coordinating Cys or His residue, which is typically position +11. Thus, "-1" refers to the residue in the framework structure immediately preceding the first residue of the α -helix. With respect to a particular zinc finger domain, as used herein, residues referred to as "++" are located in the immediately adjacent (C-terminal) zinc finger domain. Generally, nucleic acid recognition by a zinc finger module is achieved primarily by the amino acid side chains at positions -1, +3, +6 and ++2; although other amino acid positions (especially of the α -helix) may sometimes contribute to binding between the zinc finger and the target molecule. Since the vast majority of base-specific interactions between dsDNA and a zinc finger domain come from this relatively short stretch of amino acids, it is convenient to define the sequence of the zinc finger domain from -1 to +6

(i.e. residues -1, 1, 2, 3, 4, 5 and 6) as a zinc finger 'recognition sequence'. For ease of understanding, it is worth noting that the first invariant histidine residue that coordinates the zinc ion is position (+)7 of the zinc finger domain. The remainder of the zinc finger domain (i.e. either side of the recognition sequence) can be referred to as the zinc finger peptide 'scaffold', since it functions to support the recognition α -helix in a position suitable for interacting with bases of the nucleic acid target site.

When binding to a nucleic acid sequence, the zinc finger recognition sequence primarily interacts with one strand of a double-stranded nucleic acid molecule (the primary strand or sequence). However, there can be subsidiary interactions between amino acids of a zinc finger domain and the complementary (or secondary) strand of the double-stranded nucleic acid molecule. For example, the amino acid residue at the ++2 position typically may interact with a nucleic acid residue in the secondary strand.

During binding, the α -helix of the zinc finger domain almost invariably lies within the major groove of dsDNA and aligns antiparallel to the target nucleic acid strand. Accordingly, the primary nucleic acid sequence is arranged 3' to 5' in order to correspond with the N-terminal to C-terminal sequence of the zinc finger peptide. Since nucleic acid sequences are conventionally written 5' to 3', and amino acid sequences N-terminus to C-terminus, when a target nucleic acid sequence and a zinc finger peptide are aligned according to convention, the primary interaction of the zinc finger peptide is with the complementary (or minus) strand of the nucleic acid sequence, since it is this strand which is aligned 3' to 5' (see also Figures 1A to 1D). These conventions are followed in the nomenclature used herein.

Zinc finger peptides according to the invention are non-natural and suitably contain 3 or more, for example, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 24 or more (e.g. up to approximately 30 or 32) zinc finger domains arranged adjacent one another in tandem. Such peptides may also be referred to herein as 'poly-zinc finger peptides'.

In aspects and embodiments, zinc finger peptides of the invention include at least 6 zinc finger domains, preferably at least 8, at least 9, at least 10, at least 11, or at least 12, and in some cases at least 18 zinc finger domains. Preferably, the zinc finger peptides in these aspects and embodiments of the invention have from 6 to 18, from 9 to 18 or from 9 to 12 zinc finger domains arranged in tandem (e.g. 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or 16). Particularly beneficial zinc finger peptides have 6, 9 or 11 zinc finger domains arranged in tandem; and especially 9 or 11 zinc finger domains.

In some aspects and embodiments, there may be two poly-zinc finger peptides, which may differ in the number of zinc finger domains arranged in tandem. For example, one poly-zinc finger peptide in these aspects and embodiments has 9 or fewer zinc finger domains arranged in tandem and the other poly-zinc finger peptide has 9 or more zinc finger domains arranged in tandem. For example, one zinc finger peptide may have from 6 to 9 (e.g. 6, 7, 8 or 9) zinc finger domains arranged in tandem; and the other zinc finger peptide may have from 9 to 12 (e.g. 9, 10, 11 or 12) zinc finger domains arranged in tandem. In one particular embodiment one zinc finger peptide of a pair has 9 zinc finger domains in tandem and the other zinc finger peptide has 11 zinc finger domains in tandem. In another embodiment one zinc

finger peptide of a pair has 6 zinc finger domains in tandem and the other zinc finger peptide has 11 zinc finger domains in tandem. In one particular embodiment one zinc finger peptide of a pair has 6 zinc finger domains in tandem and the other zinc finger peptide has 9 zinc finger domains in tandem.

- 5 As already noted, the zinc finger peptides of the invention may bind to non-contiguous or contiguous nucleic acid binding sites. When targeted to non-contiguous binding sites, each sub-site (or half-site where there are two non-contiguous sequences) is suitably at least approximately 18 bases long, but may alternatively be approximately 12, 15 or 24 bases long. Preferred 11 zinc finger peptides of the invention recognise and bind to nucleic acid target sequences which are approximately 33 nucleotides
10 long, but which may contain two subsites, e.g. of 18 and 15 nucleotides arranged directly adjacent to one another to form a contiguous sequence, or which subsites are separated by intervening nucleotides to create a non-contiguous target site. Typically, a 12 zinc finger peptide binds a full-length nucleic acid target sequence which is approximately 36 nucleotides long, but which may be formed of two suitable subsites, e.g. each of 18 nucleotides that are arranged directly adjacent to one another to form a
15 contiguous sequence, or may be separated by intervening nucleotides as in the case of a non-contiguous target site. Preferred 9 zinc finger peptides of the invention bind to full-length nucleic acid target sequence which is approximately 27 nucleotides long, but which may contain two subsites of suitable length (e.g. 12 and 15 bases) arranged directly adjacent to one another to form a contiguous sequence, or which are separated by intervening nucleotides to create a non-contiguous target site.
- 20 Preferred 6 zinc finger peptides of the invention bind to full-length nucleic acid sequences which are approximately 18 nucleotides long, but which may contain two subsites, of e.g. 9 nucleotides, arranged directly adjacent to one another to form a contiguous sequence, or which are separated by intervening nucleotides to create a non-contiguous target site.
- 25 In (poly-)zinc finger peptides of the present invention, adjacent zinc finger domains are joined to one another by 'linker sequences' that may be canonical, canonical-like, flexible or structured, as described, for example, in WO 01/53480 (Moore *et al.*, (2001) *Proc. Natl. Acad. Sci. USA* 98: 1437-1441); WO2012/049332; WO2017/077329 and WO2022/003361. Generally, a natural zinc finger linker sequence lacks secondary structure in the free form of the peptide. However, when the protein is bound
30 to its target site a canonical linker is typically in an extended, linear conformation, and amino acid side chains within the linker may form local interactions with the adjacent nucleic acid. In a tandem array of zinc finger domains, the linker sequence is the amino acid sequence that lies between the last residue of the α -helix in an N-terminal zinc finger and the first residue of the β -sheet in the next (i.e. C-terminal adjacent) zinc finger. For the purposes of the present invention, the last amino acid of the α -helix in a
35 zinc finger is considered to be the final zinc coordinating histidine (or cysteine) residue, while the first amino acid of the following finger is generally a tyrosine, phenylalanine or other hydrophobic residue.

- It is desirable that the zinc finger peptides of the invention bind relatively specifically to their target sequence. It will be appreciated, however, that 'specificity' to a highly repetitive sequence is not a
40 straightforward concept in the sense that relatively shorter and relatively longer repetitive sequences may both be targeted and bound with good affinity. In accordance with some embodiments of the invention (and as described elsewhere herein), the zinc finger peptides of the invention may beneficially

exhibit preferential binding to relatively longer repeat sequences over relatively shorter repeat sequences.

Binding affinity (e.g. dissociation constant, K_d) is one way to assess the binding interaction between a zinc finger peptide and a potential target nucleic acid sequence. The binding affinity of a zinc finger peptide for its selected / potential target sequence can be measured using techniques known to the person of skill in the art, such as surface plasmon resonance, or biolayer interferometry. Biosensor approaches are reviewed by Rich *et al.* (2009), "A global benchmark study using affinity-based biosensors", *Anal. Biochem.*, 386:194-216. Alternatively, real-time binding assays between a zinc finger peptide and target site may be performed using biolayer interferometry with an Octet Red system (Fortebio, Menlo Park, CA). It can be useful to measure binding affinity of the zinc finger peptides of the invention to ensure that each achieves the desired binding strength. In addition, where zinc finger peptides of the invention are modified, e.g. to lower potential immunogenicity for host-optimisation, it can be useful to measure the binding affinity so ensure that those modifications – especially those in the recognition sequence region – have not adversely affected nucleic acid binding affinity.

Zinc finger peptides of the invention typically have μM or higher binding affinity for a target nucleic acid sequence. Suitably, in some embodiments a zinc finger peptide of the invention has nM or sub-nM binding affinity for its specific target sequence; for example, 10^{-9} M, 10^{-10} M, 10^{-11} M, or 10^{-12} M or less. In some particularly preferred embodiments, the affinity of a zinc finger peptide of the invention for its target sequence is in the pM range or below, for example, in the range of 10^{-13} M, 10^{-14} M, or 10^{-15} M or less. In other embodiments a zinc finger peptide of the invention has weaker than nM or sub-nM binding affinity for its specific target sequence; for example, 10^{-9} M, 10^{-8} M, 10^{-7} M, or 10^{-6} M or less.

Binding affinity between a zinc finger peptide of the invention and a target nucleic acid sequence can conveniently be assessed using an ELISA assay, as is known to the person of skill in the art.

The present invention relates to non-naturally occurring poly-zinc finger peptides for binding to sequences of the frataxin gene (or associated sequences), such as the repetitive, pathogenic trinucleotide repeat sequences (particularly to GAA-repeats) or any off-frame repeat variants, as may be found in naturally-occurring genomic DNA sequences. In other embodiments the zinc finger peptide of the invention may bind to non-repetitive, unique target sequences. The invention also relates to the use of such poly-zinc finger peptides as therapeutic molecules and to related methods of treatment: for example, for treating diseases associated with loss of function / loss of expression of frataxin, e.g. as may be caused by expanded GAA-repeat sequences; in particular, Friedreich's ataxia (FRDA). Desirably, in some embodiments poly-zinc finger peptides of the invention bind to expanded GAA-repeats or the related frame-shift variant repeat sequences, which are associated with mutated, pathogenic gene sequences in preference to and/or selectively over the shorter GAA-repeat sequences, which may be found in normal, non-pathogenic genes. For example, the binding affinity of a zinc finger peptide of the invention for a pathogenic nucleotide repeat sequence may be at least 2-fold higher, at least 10-fold higher, or at least 100-fold higher than for a wild-type / non-pathogenic nucleotide repeat sequence for the respective gene. In FRDA embodiments, the binding affinity of zinc

finger peptides of the invention for sequences of 300 or more GAA repeats may be at least 2-fold higher, at least 5-fold, or at least 10-fold higher than it is for sequences of 12 or less GAA repeats. Suitably, the affinity of such zinc finger peptides of the invention for DNA sequences having at least 300 GAA repeats is at least 5-fold, at least 10-fold or at least 20-fold higher than for sequences having 12 or less GAA repeats. In some particularly advantageous embodiments, the affinity of zinc finger peptides of the invention for DNA sequences having at least 600 GAA repeats is at least 5-fold, at least 10-fold or at least 20-fold higher than for sequences having up to 20 or up to 12 GAA repeats.

Zinc Finger Peptide Frameworks and Derivatives

Zinc finger peptides have proven to be extremely versatile scaffolds for engineering novel DNA-binding domains (e.g. Rebar & Pabo (1994) *Science* 263: 671-673; Jamieson *et al.*, (1994) *Biochemistry* 33: 5689-5695; Choo & Klug (1994) *Proc. Natl. Acad. Sci. USA.* 91: 11163-11167; Choo *et al.*, (1994) *Nature* 372: 642-645; Isalan & Choo (2000) *J. Mol. Biol.* 295: 471-477; and many others).

There are a number of natural zinc finger frameworks known in the art, and any of these frameworks may be suitable for use in the zinc finger peptide frameworks of the invention. In general, a natural zinc finger framework has the sequence, Formula 1: $X_{0-2} C X_{1-5} C X_{9-14} H X_{3-6}^{H/C}$; or Formula 2: $X_{0-2} C X_{1-5} C X_{2-7} X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6} H X_{3-6}^{H/C}$ where X is any amino acid, the numbers in subscript indicate the possible numbers of residues represented by X, the numbers in superscript indicate the position of the amino acid in the α -helix, and H/C means that the amino acid in that position can be either H or C (for a review explaining this terminology, see: Encyclopaedia of Biological Chemistry III (Third Edition) Volume 3, 2021, Pages 506-516 <https://doi.org/10.1016/B978-0-12-809633-8.21266-1>). In embodiments of the invention, the zinc finger peptide framework is based on an array of zinc finger domains of Formula 1 or 2. Alternatively, the zinc finger motif may be represented by the general sequence, Formula 3: $X_2 C X_{2,4} C X_{12} H X_{3,4,5}^{H/C}$; or Formula 4: $X_2 C X_{2,4} C X_5 X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6} H X_{3,4,5}^{H/C}$. Still more preferably the zinc finger motif may be represented by the general sequence, Formula 5: $X_2 C X_2 C X_{12} H X_3 H$; or Formula 6: $X_2 C X_2 C X_5 X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6} H X_3 H$. Accordingly, an extended zinc finger peptide framework of the invention may be based on zinc finger domains of Formulas 1 to 6, or combinations of Formulas 1 to 6, joined together in an array using the linker sequences described herein.

In these formulas, the fixed C and H residues coordinate the zinc ion to stabilise the zinc finger structure: the first H residue is position +7 of the α -helix. Particularly preferred positions for diversification within the zinc finger domain frameworks of the invention, in order to direct binding to a desired target, are those within or adjacent the α -helix, for example, positions -1, 2, 3 and 6. It can be beneficial to minimise these diversifications, particularly with respect to residues of the α -helix outside of these positions, where the zinc finger framework is otherwise native to the biological system in which the zinc finger peptides of the invention may be used *in vivo*, so as to reduce host-immune reactions.

Preferred zinc finger peptide arrays of the invention have a sequence and framework (excluding the recognition sequences, which are described elsewhere herein) according to one or more of Structures

I, II, III and IV as defined in our earlier patent applications, WO 2012/049332 and WO 2017/077329, which teaching of said zinc finger peptide frameworks (i.e. Structures I, II, III and IV) is explicitly incorporated herein by reference in its entirety, including any preferred and optional features thereof.

5 In some aspects and embodiments of the invention, the extended zinc finger peptide framework comprises 6 or more zinc finger domains of one of Formulas 1 to 6, joined together by linker sequences, i.e. Structure V: [(Formula 1-6) – linker]_n – (Formula 1-6)], where n is ≥5, such as between 5 and 31. As indicated, in Structure V any combination of Formulas 1 to 6 may be used. In another embodiment the extended zinc finger peptide framework comprises between 6 and 18 (e.g. 6 to 12, 6 to 11, 8 to 11 or 9
10 to 11) zinc finger domains of the above formulae. Suitably, therefore, n in Structure V is 5 to 17 (i.e. one less than the number of zinc finger domains); particularly n is 5 to 11; suitably n is 5, 6, 7, 8, 9, 10 or 11; more suitably n is 5, 8 or 10; and preferably n is 8 or 10.

As already described, adjacent zinc finger domains are joined together by linker sequences. In a natural
15 zinc finger protein, threonine is often the first residue in the linker, and proline is often the last residue of the linker. On the basis of sequence homology, the canonical natural linker sequence is considered to be -TGEKP- (Linker 1 or L1; SEQ ID NO: 76). However, natural linkers can vary greatly in terms of amino acid sequence and length. Therefore, a common consensus sequence based on natural linker sequences may be represented by -TG^{E/Q}K/RP- (Linker 2 or L2, where E/Q means that either E or Q
20 can be present, and K/R means that either K or R can be present; SEQ ID NO: 77), and this sequence is preferred for use as a 'canonical' (or 'canonical-like') linker in accordance with the invention. Thus, another useful canonical linker sequence is -TGQKP- (SEQ ID NO: 78).

However, in extended zinc finger arrays of e.g. 4 or more zinc finger domains, it has been shown that
25 it can be beneficial to periodically disrupt the canonical linker sequence, when used between adjacent zinc fingers in an array, by adding one or more amino acid residue (e.g. Gly and/or Ser), to create groups of 2 or 3 zinc finger domains within the array (Moore *et al.*, (2001) *Proc. Natl. Acad. Sci. USA* 98: 1437-1441; and WO 01/53480). Therefore, suitable linker sequences for use in accordance with the invention include canonical linker sequences of 5 amino acids (e.g. Linker 1 or Linker 2, above), and
30 related canonical-like linker sequences of 6 or 7 amino acids.

Canonical-like linkers for use in accordance with the invention may suitably be based on the sequence, -TG^{G/S}E/Q^{K/R}P- (Linker 3 or L3; SEQ ID NO: 79). Preferred canonical-like linkers thus include the specific sequences: TGGERP (SEQ ID NO: 80), TGSERP (SEQ ID NO: 81), TGGQRP (SEQ ID NO: 82),
35 TGSQRP (SEQ ID NO: 83), TGGEKP (SEQ ID NO: 84), TGSEKP (SEQ ID NO: 85), TGGQKP (SEQ ID NO: 86), or TGSQKP (SEQ ID NO: 87). A particularly preferred canonical-like linker is TGSERP (Linker 4 or L4; SEQ ID NO: 81). Another particularly preferred canonical-like linker is TGSQKP (Linker 5 or L5; SEQ ID NO: 87). However, other linker sequences may also be used between one or more pairs of zinc finger domains, for example, linkers of the sequence -TG^(G/S)0-2E/Q^{K/R}P- (SEQ ID NO: 88)
40 or -T^(G/S)0-2G^{E/Q}K/RP- (Linker 6 or L6; SEQ ID NO: 89).

In some embodiments still longer flexible linkers of 8 or more amino acids may be used, as previously described. Linkers of 8 amino acids include the sequences -TG^(G/S)₃G^{E/Q}K^RP- (SEQ ID NO: 90) and -T^(G/S)₃G^{E/Q}K^RP- (L12; SEQ ID NO: 91). Alternative long flexible linkers are: LRQKD(GGGGS)₁₋₄QLVGTAERP (Linker 7 or L7; SEQ ID NO: 92) and LRQKD(GGGGS)₁₋₄QKP (Linker 8 or L8; SEQ ID NO: 93). Preferred long flexible linkers for use in the zinc finger peptides of the invention are, LRQKDGGGGSGGGSGGGSQLVGTAERP (Linker 9 or L9; SEQ ID NO: 94), and LRQKDGGGGSGGGSGGGSQKP (Linker 10 or L10; SEQ ID NO: 95).

A. Extended Poly-Zinc Finger Proteins

For specific biological functionality and therapeutic use, particularly *in vivo* (e.g. in gene therapy and transgenic animals), it is generally desirable that a poly-zinc finger peptide of the invention is able to target unique or virtually unique sites (or clusters) within any genome. For complex genomes, such as in humans, it is generally considered that an address of at least 16 bps is required to specify a potentially unique DNA sequence. Shorter DNA sequences have a significant probability of appearing several times in a genome, which increases the possibility of obtaining undesirable non-specific gene targeting and biological effects. Since individual zinc fingers generally bind to three consecutive nucleotides, 6-zinc finger domains with an 18 bp binding site could, in theory, be used for the specific recognition of a unique target sequence within any genome. Accordingly, a great deal of research has been carried out into so-called 'designer transcription factors' for targeted gene regulation, which typically involve 4, 6 or, in the inventors previous work, 11 or 12-zinc finger domains that may be arranged in tandem or in dimerisable groups (e.g. of two or three-finger units).

In some aspects and embodiments, the present invention relates to targeting of long arrays of nucleotide repeat sequences, and so there will be considerably more than one identical target site within the genome. Nevertheless, effective targeting (e.g. for therapy) of a desired sequence can be difficult taking into account the potential for yet more identical sequences associated with non-pathogenic, wild-type genes.

The inventors have previously shown (WO 2012/049332 and WO 2017/077329) that by selecting appropriate linker sequences and suitable combinations of linker sequences within an array of zinc fingers, extended arrays of zinc finger peptides of at least 8 or 10 zinc fingers (such as 10, 11, 12 or 18) can be synthesised, expressed and can have selective gene targeting activity. Extended arrays of zinc finger peptides of the invention are conveniently arranged in tandem. By way of example, 9- or 11-zinc finger peptides can recognise and specifically bind 27 or 33 nucleic acid residues, respectively. In this way, the extended zinc finger peptides of the invention can be targeted to preferred genomic sequences, e.g. expanded GAA trinucleotide repeat sequences.

In accordance with aspects and embodiments of the invention, target nucleic acid binding sites are selected as described elsewhere herein such that the poly-zinc finger peptide binds effectively to the intended target sequences, such as pathogenic GAA-repeat nucleic acid sequences, while reducing,

minimising or preventing binding to non-pathogenic (off-target), wild-type GAA-repeat sequences in the intended subject / host (e.g. mouse or human).

The inventor's earlier work (e.g. WO 2012/049332; WO 2017/077329, each of which are incorporated herein by reference in their entirety) was the first to demonstrate that tandem arrays of more than 6 zinc finger domains, such as 8, 9, 10, 11, 12, 18 or more zinc fingers can be synthesised and expressed; and, more significantly, that such long arrays of non-natural zinc finger domains can have *in vitro* or *in vivo* (specific) nucleic acid binding activity. In this earlier work we also reported that such extended arrays of zinc finger peptides were capable of targeting genomic DNA sequences and have gene modulation activity *in vitro* and/or *in vivo*. We have also demonstrated that such extended zinc finger peptide frameworks comprising at least 8, at least 9, at least 11, at least 12, or at least 18 zinc finger domains can preferentially target expanded nucleic acid repeat sequences – e.g. as associated with pathogenic phenotypes preferentially over wild-type shorter repeat sequences.

In embodiments, suitable extended poly-zinc finger peptide frameworks of the invention comprise from 6 to 32 zinc finger domains, from 6 to 28 zinc finger peptides, from 6 to 24 zinc finger peptides, from 6 to 18 zinc finger peptides, or from 6 to 12 zinc finger peptides. Preferred zinc finger peptides according to aspects and embodiments of the invention comprise 6, 9 or 11 zinc finger domains.

The zinc finger peptide frameworks of the invention may comprise directly adjacent zinc finger domains having canonical (or canonical-like) linker sequences between adjacent zinc finger domains, such that they preferentially bind to contiguous nucleic acid sequences. Accordingly, a 6-zinc finger peptide (framework) of the invention is particularly suitable for binding to contiguous stretches. Typically, extended poly-zinc finger peptides, according to the invention are designed to bind nucleic acid sequences which may be arranged as a contiguous stretch or as a non-contiguous stretch comprising two or three subsites. For example, a 9-zinc finger peptide is particularly suitable for binding a target sequence of approximately 27 nucleotides; and an 11-zinc finger peptide is suitable for binding approximately 33 nucleotides. As already described, such target sequences may be arranged contiguously or in non-contiguous subsites especially arranged in subsites of e.g. 9, 12, 15 or 18 nucleotide lengths.

The extended arrays of zinc finger domains in the peptides and polypeptides of the invention typically comprise canonical linker sequences, short flexible (canonical-like) linker sequences and long flexible linker sequences. Thus, in some embodiments, one or more pairs of adjacent zinc finger domains of a zinc finger peptide according to the invention may be separated by short canonical linker sequences (e.g. TGERP, SEQ ID NO: 96; TGEKP, SEQ ID NO: 76; etc.). In some embodiments, one or more pairs of adjacent zinc finger domains may be separated by short flexible linker sequences (e.g. of 6 or 7 amino acids), 'canonical-like' linker sequences, which preferably comprise the amino acid residues of a canonical linker with an additional one or two amino acid residues within, before or after the canonical sequence (preferably within). Adjacent zinc finger domains separated by canonical and canonical-like (short flexible linker sequences) i.e. which are between 5 and 7 amino acids long, typically bind to contiguous nucleic acid target sites. In accordance with the invention, however, one or more pairs of

adjacent zinc finger domains of a zinc finger peptide may be separated by long flexible linker sequences, for example, comprising 8 or more amino acids, such as between 8 and 50 amino acids. Particularly suitable long flexible linkers have between approximately 10 and 40 amino acids, between 15 and 35 amino acids, or between about 20 and 30 amino acids. Preferred long flexible linkers may have 18, 23 or 29 amino acids. Adjacent zinc finger domains separated by long flexible linkers have the capacity to bind to non-contiguous binding sites in addition to the capacity to bind to contiguous binding sites. The length of the flexible linker may influence the length of intervening DNA that may lie between such non-contiguous binding sub-sites. This can be a particular advantage in accordance with the invention, since poly-zinc finger peptides that target extended trinucleotide repeat sequences may then have a number of options for binding to contiguous as well as discontinuous target sequences.

Suitably, the zinc finger peptides / frameworks of the invention may comprise two or more (e.g. 2, 3 or 4) arrays of 3, 4, 5 or 6 directly adjacent zinc finger domains (or any combination thereof) separated by long flexible (or structured) linkers. Preferably, such extended (poly-)zinc finger peptides are arranged in multiple arrays of 3, 4, 5 and/or 6-finger units separated by long flexible linkers.

The inventors have previously shown that such extended zinc finger peptides of more than 6 zinc fingers in total can exhibit specific and high affinity binding to desired target sequences, both *in vitro* and *in vivo*. For example, whereas a 3-finger peptide (with a 9 bp recognition sequence) may bind DNA with nanomolar affinity, a 6-finger peptide might be expected to bind an 18 bp sequence with an affinity of between 10^{-9} and 10^{-18} M, depending on the arrangement and sequence of zinc finger peptides. To optimise both the affinity and specificity of 6-finger peptides, a fusion of three 2-finger domains has been shown to be advantageous (Moore *et al.*, (2001) *Proc. Natl. Acad. Sci. USA* 98: 1437-1441; and WO 01/53480). Therefore, in some embodiments, the zinc finger peptides of the invention comprise a series of 2-finger units arranged in tandem. Zinc finger peptides of the invention may alternatively include or comprise a series of 3-finger units.

However, the inventors have previously reported that extended poly-zinc finger peptides can be 'tuned' to moderate binding affinity for nucleic acid-repeat sequences according to the presence of both pathogenic and non-pathogenic (WT) target sequences within the same target cells. In aspects and embodiments of the invention, therefore, zinc finger activator proteins are tuned to bind preferentially to extended, pathogenic repeat sequences. In this way, expression of wild-type, target gene products may be upregulated, whereas expression of non-target gene products is not upregulated.

Furthermore, it has been demonstrated that the extended zinc finger peptides of the invention can be stably expressed within a target cell, can be non-toxic to the target cell, and can have a specific and desired gene modulation activity. In particular, it has been shown that the zinc finger activator proteins of the invention can have prolonged expression in target cells *in vivo*, without causing toxic side-effects that are often associated with the expression of heterologous / foreign protein sequences *in vivo*.

As noted above, in some aspects and embodiments, the extended zinc finger peptides of the invention are adapted for binding to repeat sequences (i.e. trinucleotide repeats) in target genes. According to

such aspects and embodiments, suitable target sequences in pathogenic frataxin gene sequences may comprise at least 35, at least 90, at least 100, at least 200, at least 300, at least 400, at least 500, or at least 600 trinucleotide repeats. In some embodiments, pathogenic frataxin gene sequences may have at least 700, at least 800, at least 900 or at least 1,700 trinucleotide repeats. Typically, non-pathogenic, wild-type genome sequences may have less than 35 trinucleotide repeats; for example, less than 34, less than 30, less than 20, or less than 10 trinucleotide repeats. In some aspects and embodiments, non-pathogenic frataxin gene sequences may also or alternatively be targeted by zinc finger activator peptides of the invention in order to upregulate expression of frataxin transcripts and protein.

Suitably, for targeting a pathogenic frataxin gene such as in FRDA, the binding site comprises repeats of 5'- GAA -3', which are typically inserted into intron 1 of the pathogenic frataxin gene. However, in view of the repeat register, suitable binding sites may also or alternatively comprise repeats of 5'- AGA -3' or 5'- AAG -3'. Desirably, target sequences for the extended zinc finger peptides of the invention may comprise 90 or more contiguous 5'- GAA -3' repeats, such as at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, or at least 900 5'- GAA -3' repeats.

Poly-zinc finger peptides of 8 or more (e.g. 11) tandem zinc finger domains can exhibit specific and high affinity binding to desired target sequences, both *in vitro* and *in vivo*. The inventors' previous studies, see e.g. WO 2012/049332, were the first to report on the systematic exploration of the binding modes of different-length ZFP to long repetitive DNA tracts. In particular, it has been demonstrated that whereas all poly-zinc finger peptides may bind to expanded (e.g. pathogenic) nucleic acid repeat sequences in preference over shorter (e.g. wild-type) repeat sequences; it appears that longer arrays of zinc fingers may demonstrate more pronounced preference for expanded repeat sequences. It is believed that this may, in part, be due to steric reasons, whereby long arrays of zinc fingers may interfere with each other when trying to bind shorter repeat sequences.

In some aspects and embodiments, therefore, a particular advantage of the zinc finger peptides of the invention is that they bind to longer arrays of GAA-repeat sequences in preference to shorter, non-pathogenic arrays. Accordingly, the GAA trinucleotide repeat-targeting zinc finger peptides of the invention suitably bind more effectively (e.g. with higher affinity or greater gene modulation ability) to expanded, pathogenic nucleotide-repeat sequences compared to similar, shorter wild-type nucleotide-repeat sequences. For targeting / treatment of FRDA, GAA-targeting zinc finger peptides of the invention bind with higher affinity to expanded GAA-repeat sequences containing at least 90 repeats, compared to sequences containing e.g. 34 or less repeats. Similarly, sequences containing at least 90 GAA repeats may be bound preferentially over sequences containing 10 or less repeats; sequences containing at least 200 or 1,000 GAA repeats may be bound preferentially over sequences containing 20 or less repeats (as well as sequences including 10 or less repeats). Similarly, sequences containing at least 300 GAA repeats may be bound preferentially over sequences containing 20 or less repeats (including 10 or less repeats).

In some aspects and embodiments, however, zinc finger activator peptides of the invention target sequences within the frataxin promoter, which may be unique (or almost unique) in the genome of the

organism. Therefore, for such zinc finger peptides that do not target repeat sequences there is preferably only one nucleic acid sequence in the genome that is identical to the desired binding / target site, i.e. the selected sequence within the frataxin promoter.

5 As the skilled person will appreciate, the amino acid sequence of the zinc finger recognition sequence of each zinc finger domain of a poly-zinc finger peptide is suitably determined by the nucleic acid sequence of the target nucleic acid triplet (or staggered quadruplet) to which it is adapted to bind. Therefore, where the zinc finger peptides are designed to target unique, non-repeating sequences (e.g. of the frataxin promoter), the recognition sequences of each zinc finger domain of a poly-zinc finger
10 peptide of these aspects and embodiments is generally different from its neighbouring zinc finger domains. By contrast, where a poly-zinc finger peptide is intended to bind a repeating, e.g. trinucleotide repeat sequence, the recognition helices within the array may have the same or similar, repeating pattern; especially at the main nucleic acid-contacting positions or -1, 3 and 6 of the domain. However, since the main nucleic acid-binding interactions are formed with the residues in the -1, 3 and 6 positions
15 (as previously described), the amino acids at the helix 'framework' positions (outside of the -1, 3 and 6 positions) could potentially be the same, similar or different for adjacent zinc finger domains irrespective of the target binding site sequence.

For example, as described, in embodiments of the invention, the residues at positions +4 and +5 may
20 suitably be, respectively, L and T. In this way host matching for *in vivo* mouse or human applications may be improved. In such embodiments, the residues R and K, respectively, may be used in the 4 and 5 positions. In some embodiments, adjacent zinc finger domains may have alternating (regular or irregular) patterns of amino acids at the 4 and 5 positions; for example, the residues L and T in the 4 and 5 positions may be used interchangeably with the residues R and K, depending – for example – on
25 the overall charge of amino acid residues in the recognition helix of a zinc finger domain, or within an array of zinc finger domains. In this way, the target binding site – and more particularly – the amino acid residues selected to bind the target binding site, may influence the selection of recognition helix framework residues, especially at the 4 and 5 positions.

30 In some embodiments, the first zinc finger of a zinc finger array (or sub-array – where adjacent sub-arrays are separated from each other by long, flexible linkers) has a recognition sequence wherein the residues at positions 4 and 5 are, respectively, L and T. In such embodiments, all remaining recognition sequences of the array (or sub-array) may have the residues R and K, respectively, in the 4 and 5 positions; may have the residues L and T, respectively, in the 4 and 5 positions; or may include a mixture
35 of R and K or L and T residues in the 4 and 5 positions, respectively. In this way, the positive charge of the recognition helix of a zinc finger domain can be adjusted; for example, depending on the residues selected at the -1, 3 and 6 positions; e.g. with the aim of reducing the overall positive charge of the recognition helix of a zinc finger domain.

40 In some embodiments, an S may be substituted for a T residue – particularly at position 5 in the recognition helix, without adversely affecting binding specificity of the zinc finger domain. Likewise, in some embodiments, a K may be substituted for an I residue at the 5 position in the recognition helix.

As such, the 4 position of the recognition helix is typically selected from L or R; whereas the 5 position of the recognition helix may typically be selected from T, K, S or I; and is more typically selected from T or K.

5

B. Poly Zinc Finger Acitvator Proteins for Targeting 5'-GAA-3' or 5'-AAG-3' Repeats

According to aspects and embodiments of the invention, the zinc finger peptides are designed to target repeating GAA / AAG triplets. Accordingly, the recognition sequences of adjacent zinc finger domains of a poly-zinc finger peptide of the invention may be identical, or very similar, along the length of the zinc finger array. However, to optimise one or more feature of the zinc finger peptide, such as cloning, expression, or binding behaviour, it may be beneficial to vary the recognition sequence of one or more zinc finger domains along the length of the zinc finger array. For example, the recognition helix sequence may be varied to control amino acid side chain charges, e.g. to avoid the poly zinc finger peptide from having an undesirably high overall positive or negative charge, which may affect binding of the zinc finger peptide to DNA. For example, too high a positive charge may increase binding affinity for non-target DNA sequences by generating more higher non-specific binding affinity for the generally negatively-charged DNA backbone (WO2017/077329; WO2022/003361).

In some embodiments, the recognition sequences of zinc finger domains in zinc finger peptides for targeting GAA-repeat (or AAG-repeat) sequences may be selected from two or more general formulae, which conveniently may alternate along the zinc finger array.

For clarity, as used herein, the nomenclature 'ZF1' will be used with respect to zinc finger peptides engineered for binding to the 5'- GAA -3' triplet repeat, and the nomenclature 'ZF2' will be used with respect to zinc finger peptides engineered for binding to the 5'- AAG -3' triplet repeat.

In embodiments relating to ZF1 peptides, in order to tune the binding affinity of the poly-zinc fingers peptides towards the 5'- GAA -3' repeat triplet, the residue at the -1 position is preferably Q; the residue at the 3 position is preferably N; and/or the residue at the 6 position is preferably R. In embodiments the residue at the 2 position is preferably G.

Further, in embodiments relating to ZF2 peptides, in order to tune the binding affinity of the poly-zinc fingers of an activator peptide towards the 5'- AAG -3' repeat triplet (of a GAA trinucleotide repeat), the residue at the -1 position is preferably R; the residue at the 3 position is preferably N; and/or the residue at the 6 position may be preferably selected from Y or Q. In embodiments the residue at the 2 position may be preferably selected from S or A.

According to the invention, ZF1 zinc finger recognition sequences (i.e. positions X^{-1} , X^{+1} , X^{+2} , X^{+3} , X^{+4} , X^{+5} and X^{+6} in Formulas 2, 4 and 6, or Structure V above) for binding to a 5'- GAA -3' (3'-AAG-5') repeat triplet may be represented by the amino acid sequences below:

SEQ ID NO: 1: (Q/N)S(G/A/Q)N(L/R)(T/K/G/S)R
 SEQ ID NO: 2: QSGN(L/R)(T/K)R
 SEQ ID NO: 3: (Q/N)SANL(S/G)R
 SEQ ID NO: 4: (Q/N)S(A/Q)NLSR
 5 SEQ ID NO: 5: QSGNLTR
 SEQ ID NO: 6: QSGNRKR
 SEQ ID NO: 7: QSANLGR
 SEQ ID NO: 8: NSANLSR
 SEQ ID NO: 9: QSQNLSR

10

In some embodiments, at least 2 – for example, 2, 3, 4 or 5 of the variable positions in each of SEQ ID NOs: 1 to 4 are selected to be the first residue within each set of parentheses “(...)”. In some embodiments at least 1 – for example, 1, 2, 3 or 4 – of the variable positions in each of SEQ ID NOs: 1 to 4 are selected to be other than the first residue within each set of parentheses “(...)”. In some
 15 embodiments, 2 of the variable positions in each of SEQ ID NOs: 1 to 4 are selected to be the first residue within each set of parentheses “(...)” and 2 of the variable positions in each of SEQ ID NOs: 1 to 4 are selected to be the second residue within each set of parentheses “(...)”.

15

Thus, in embodiments, there is provided an engineered zinc finger (DNA-binding) peptide comprising
 20 at least 6, such as from 6 to 32, 6 to 18 or 6 to 12 zinc finger domains; for example, 6, 9, 11 or 12 zinc finger domains; and more specifically 9 or 11 zinc finger domains, having zinc finger recognition sequences selected from those of SEQ ID NO: 1. In embodiments, the zinc finger domains may have zinc finger recognition sequences selected from those of SEQ ID NO: 2; SEQ ID NO: 3 and/or SEQ ID NO: 4.

25

In some embodiments, the zinc finger domain recognition sequences of a zinc finger peptide according to the invention are selected from SEQ ID NO: 5 and/or SEQ ID NO: 6. In some embodiments, the zinc finger domain recognition sequences of a zinc finger peptide according to the invention are selected from SEQ ID NO: 7 and/or SEQ ID NO: 8. In some embodiments, the zinc finger domain recognition
 30 sequences of a zinc finger peptide according to the invention are selected from SEQ ID NO: 8 and/or SEQ ID NO: 9.

30

In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 5 and zinc finger domains having the recognition
 35 helix sequence according to SEQ ID NO: 6. In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 7 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 8. In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 8 and zinc finger domains having the recognition
 40 helix sequence according to SEQ ID NO: 9. By selecting zinc finger recognition domain sequences of different sequences, despite those domains targeting the same nucleic acid triplet sequence, zinc finger peptides of the invention can beneficially be ‘tuned’ to improve or optimise binding affinity for the

40

intended target site, including controlling overall amino acid side chain charge across the plurality of the zinc finger domains along the length of the engineered zinc finger peptide. This strategy can be particularly beneficial when every (or the majority) of zinc finger domains in a extended zinc finger peptide (e.g. of 6 or more zinc finger domains) are adapted to recognise and bind to the same nucleic acid target sequence (e.g. GAA or AAG), because it can prevent or control the multiplication of like charges that would otherwise result from repeating the same recognition helix sequence across all zinc finger domains of the peptide.

It should be understood, that within the scope of the invention, one or more recognition sequence of SEQ ID NO: 5 may be replaced with the sequence of SEQ ID NO: 6 and *vice versa*, one or more sequence of SEQ ID NO: 7 may be replaced with the sequence of SEQ ID NO: 8 and *vice versa*, and one or more recognition sequence of SEQ ID NO: 8 may be replaced with the sequence of SEQ ID NO: 9 and *vice versa*, in order to tune the zinc finger peptide to have the desired binding characteristics. It is also noted that the features of the 11-zinc finger peptide embodiments set out above apply equally, with appropriate logical adjustments, according to the number of zinc finger domains, to all other extended poly-zinc finger peptides of the invention.

In some embodiments, the two different recognition sequences (e.g. SEQ ID NO: 5 and SEQ ID NO: 6; SEQ ID NO: 7 and SEQ ID NO: 8; or SEQ ID NO: 8 and SEQ ID NO: 9) may alternate along the length of the zinc finger peptide array. Suitably, in any embodiments, the first zinc finger domain of the zinc finger peptide of the invention has an L in position 5 of the sequence. In some embodiments, the first zinc finger domain of the zinc finger peptide of the invention has a T in position 6 of the sequence. In particular embodiments, such zinc finger peptides may comprise, include or consist of 9, 10, 11 or 12 zinc finger domains having such sequences. Beneficially, therefore, the engineered zinc finger peptides of the invention comprise at least 9, 10, 11, 12 or 18 adjacent zinc finger modules. In some embodiments, the zinc finger peptides of the invention comprise more than 9, 10, 11, 12 or 18 zinc finger domains – such as any number between 9 and 32 zinc finger domains, provided that at least 9, 10, 11, 12 or 18 adjacent domains have the specified recognition sequence. In some embodiments all zinc finger domains of a zinc finger peptide of the invention are the recognition sequences as set out herein.

In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 5. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 6. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 5. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 6. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 5. In

embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 6.

In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 7. In
5 In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 8. In
In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 7. In
10 In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 8. In
In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 7. In
In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein
15 fingers F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 8.

In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 8. In
In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein
20 fingers F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 9. In
In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 8. In
In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 9. In
25 In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 8. In
In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 9.

In embodiments, in order to tune the binding affinity of the extended poly-zinc fingers of the invention against the 5'- GAA -3' repeat sequence, at least one residue of SEQ ID NOs: 1 to 4 is A or G. Suitably,
at least one residue of SEQ ID NOs: 1 to 4 is G. In some such embodiments the residue at position 2
is G; in some embodiments the residue at position 6 is G. In some embodiments the residue at position
2 is A. In some embodiments the residue at position 2 is A and the residue at position 6 is G.

Table 1 below summarises preferred recognition sequence arrangements of the extended poly-zinc
finger peptides (e.g. activator peptides) of these aspects and embodiments of the invention. In this
table, as noted above, one or more sequence of SEQ ID NO: 5 may be substituted with the sequence
of SEQ ID NO: 6. Similarly, in the table below, one or more sequence of SEQ ID NO: 7 may be
35 substituted with the sequence of SEQ ID NO: 8. Similarly, in the table below, one or more sequence of
SEQ ID NO: 8 may be substituted with the sequence of SEQ ID NO: 9.

Finger number of zinc finger peptide	Zinc Finger Number (N to C terminal direction)				ZFP Array
	F1, F3	F2	F4, F6, F8, F10 etc.	F5, F7, F9, F11 etc.	
Recognition Sequence	SEQ ID NO: 1	SEQ ID NO: 1	SEQ ID NO: 1	SEQ ID NO: 1	A
	SEQ ID NO: 2	SEQ ID NO: 2	SEQ ID NO: 2	SEQ ID NO: 2	B
	SEQ ID NO: 3	SEQ ID NO: 3	SEQ ID NO: 3	SEQ ID NO: 3	C
	SEQ ID NO: 4	SEQ ID NO: 4	SEQ ID NO: 4	SEQ ID NO: 4	D
	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 6	SEQ ID NO: 5	E
	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 5	SEQ ID NO: 6	F
	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 8	SEQ ID NO: 7	G
	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 7	SEQ ID NO: 8	H
	SEQ ID NO: 8	SEQ ID NO: 9	SEQ ID NO: 9	SEQ ID NO: 8	I
	SEQ ID NO: 8	SEQ ID NO: 9	SEQ ID NO: 8	SEQ ID NO: 9	J

Table 1: Exemplary zinc finger recognition helix arrangements of zinc finger peptides according to the invention for binding 5'- GAA -3' repeat sequences, e.g. for treating Friedreich's ataxia (FRDA). Zinc finger peptides disclosed in this table may have from 8 to 32 fingers, for example, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 or 18 zinc finger domains.

5

Suitably, the zinc finger activator peptides of the invention for binding to 5'-GAA-3' repeats comprise have or consist of 11-zinc finger domains which are arranged in tandem. Exemplary 11-zinc finger peptide sequences of the invention for binding to 5'-GAA-3' repeat sequences in frataxin genes comprise a polypeptide having the sequence of SEQ ID NO: 97 (ZF1 peptide), SEQ ID NO: 98 (ZF1a peptide) or SEQ ID NO: 99 (ZF1b peptide), as shown in Table 9. The invention also encompasses zinc finger activator peptides comprising any of SEQ ID NOs: 97, 98 and 99 fused (or covalently linked as described herein to a suitable transcriptional activator domain – particularly an activator domain compatible with mouse and/or human cell expression; such as human p65 (SEQ ID NOs: 122 or 136) or human VP64 (SEQ ID NO: 125); for example. SEQ ID NOs: 100 and 101, (ZF1-p65, and ZF1-VP64, respectively). In aspects and embodiments, the invention also encompasses polypeptides having 90% or more, 95% or more, such as 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequences of SEQ ID NOs: 97 to 101.

10

15

20

According to the invention, ZF2 zinc finger recognition sequences (i.e. positions X^{-1} , X^{+1} , X^{+2} , X^{+3} , X^{+4} , X^{+5} and X^{+6} in Formulas 2, 4 and 6, or Structure V above) for binding to a 5'- AAG -3' (3'-GAA-5') repeat triplet may be represented by the amino acid sequences below:

SEQ ID NO: 10 R(S/N)(D/S/A)N(L/R)(T/K/S/I)(Q/Y/T)

SEQ ID NO: 11: RSDN(L/R)(T/K)Q

25 SEQ ID NO: 12: RS(S/A)N(L/R)(T/K/S)(Q/Y)

SEQ ID NO: 13: R(S/N)(D/S)N(L/R)(S/I)T

SEQ ID NO: 14: RSDNLTQ

SEQ ID NO: 15: RSDNRKQ

SEQ ID NO: 16: RSSNLTY

30 SEQ ID NO: 17: RSANRKQ

SEQ ID NO: 18: RSSNLSY

SEQ ID NO: 19: RSANLTQ
 SEQ ID NO: 20: RSSNRIQ
 SEQ ID NO: 21: RSANLSY
 SEQ ID NO: 22: RSSNRTQ
 5 SEQ ID NO: 23: RSANLSQ
 SEQ ID NO: 24: RSSNRKY
 SEQ ID NO: 25: RSSNRKQ
 SEQ ID NO: 26: RNDNRIT
 SEQ ID NO: 27: RSSNLST

10

In some embodiments, at least 2 – for example, 2, 3, 4 or 5 of the variable positions in each of SEQ ID NOs: 10 to 13 are selected to be the first residue within each set of parentheses “(...)”. In some embodiments at least 1 – for example, 1, 2, 3 or 4 – of the variable positions in each of SEQ ID NOs: 10 to 13 are selected to be other than the first residue within each set of parentheses “(...)”. In some
 15 embodiments, 2 of the variable positions in each of SEQ ID NOs: 10 to 13 are selected to be the first residue within each set of parentheses “(...)” and 2 of the variable positions in each of SEQ ID NOs: 10 to 13 are selected to be the second residue within each set of parentheses “(...)”.

Thus, in embodiments, there is provided an engineered zinc finger (DNA-binding) peptide comprising
 20 at least 6, such as from 6 to 32, 6 to 18 or 6 to 12 zinc finger domains; for example, 6, 9, 11 or 12 zinc finger domains; and more specifically 9 or 11 zinc finger domains, having zinc finger recognition sequences selected from those of SEQ ID NO: 10. In embodiments, the zinc finger domains may have zinc finger recognition sequences selected from those of SEQ ID NO: 11; SEQ ID NO: 12 and/or SEQ ID NO: 13.

25

In some embodiments, the zinc finger domain recognition sequences of a zinc finger peptide according to the invention are selected from SEQ ID NO: 14 and/or SEQ ID NO: 15. In some embodiments, the zinc finger domain recognition sequences of a zinc finger peptide according to the invention are selected from SEQ ID NOs: 16 to 25. In some embodiments, the zinc finger domain recognition sequences of a
 30 zinc finger peptide according to the invention are selected from two or more of SEQ ID NOs: 16 to 25. In some embodiments, the zinc finger domain recognition sequences of a zinc finger peptide according to the invention are selected from three or more of SEQ ID NOs: 16 to 25. In some embodiments, the zinc finger domain recognition sequences of a zinc finger peptide according to the invention are selected from four or more of SEQ ID NOs: 16 to 25. In some embodiments, the zinc finger domain recognition
 35 sequences of a zinc finger peptide according to the invention are selected from five or more of SEQ ID NOs: 16 to 25. In some embodiments, the zinc finger domain recognition sequences of a zinc finger peptide according to the invention are selected from six or more of SEQ ID NOs: 16 to 25. In some embodiments, the zinc finger domain recognition sequences of a zinc finger peptide according to the invention are selected from seven or more of SEQ ID NOs: 16 to 25. In some embodiments, the zinc
 40 finger domain recognition sequences of a zinc finger peptide according to the invention are selected from eight or more of SEQ ID NOs: 16 to 25. In some embodiments, the zinc finger domain recognition sequences of a zinc finger peptide according to the invention are selected from nine or more of SEQ ID

invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 18 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 23. In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 21 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 20. In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 21 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 22. In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 21 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 24. In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 18 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 25. In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 23 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 20. In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 23 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 22. In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 23 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 24. In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 23 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 25.

In embodiments, zinc finger peptides according to the invention include zinc finger domains having the recognition helix sequence according to SEQ ID NO: 26 and zinc finger domains having the recognition helix sequence according to SEQ ID NO: 27.

As described above with respect to zinc finger peptides for binding to 3'-AAG-5', by selecting zinc finger recognition domain sequences of different sequences within the same poly-zinc finger peptide for binding 3'-GAA-5' repeats, despite that the domains target the same nucleic acid triplet sequences, zinc finger peptides of the invention can beneficially be 'tuned' to improve or optimise binding affinity for the intended target site, including controlling overall amino acid side chain charge across the plurality of the zinc finger domains along the length of the engineered zinc finger peptide.

It should be understood, that within the scope of the invention, one or more recognition sequence of SEQ ID NO: 14 may be replaced with the sequence of SEQ ID NO: 15 and *vice versa*, one or more sequence of SEQ ID NO: 26 may be replaced with the sequence of SEQ ID NO: 27 and *vice versa*. Similarly, one or more recognition sequence of SEQ ID NOs: 16, 18, 20, 22, 24 or 25 may be used interchangeably. Likewise, one or more recognition sequence of SEQ ID NOs: 17, 19, 21 and 23 may be used interchangeably. In this way, the zinc finger peptides can be tuned to have the desired binding characteristics. It is also noted that the features of the 11-zinc finger peptide embodiments set out herein

equally, with appropriate logical adjustments, according to the number of zinc finger domains, to all other extended poly-zinc finger peptides of the invention.

In some embodiments, two different recognition sequences (e.g. SEQ ID NO: 14 and SEQ ID NO: 15; SEQ ID NO: 26 and SEQ ID NO: 27) may alternate along the length of the zinc finger peptide array. In other embodiments, two or more of SEQ ID NOs: 16 to 25 may alternate along the length of the zinc finger peptide array.

Suitably, in any embodiments, the first zinc finger domain of the zinc finger peptide of the invention has an L in position 5 of the sequence. In some embodiments, the first zinc finger domain of the zinc finger peptide of the invention has a T in position 6 of the sequence. In particular embodiments, such zinc finger peptides may comprise, include or consist of 9, 10, 11 or 12 zinc finger domains having such sequences. Beneficially, therefore, the engineered zinc finger peptides of the invention comprise at least 9, 10, 11, 12 or 18 adjacent zinc finger domains. In some embodiments, the zinc finger peptides of the invention comprise more than 9, 10, 11, 12 or 18 zinc finger domains – such as any number between 9 and 32 zinc finger domains, provided that at least 9, 10, 11, 12 or 18 adjacent domains have the specified recognition sequence. In some embodiments all zinc finger domains of a zinc finger peptide of the invention have the recognition sequences as set out herein.

In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 14. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 15. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 14. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 15. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 14. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 15.

In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 26. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 27. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 26. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 27. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 26. In

embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 27.

In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F7, F9 and F11 have recognition sequences selected from one or more of SEQ ID NOs: 16, 18, 20, 22, 24 and 25. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F6, F8 and F10 have recognition sequences selected from one or more of SEQ ID NO: 17, 19, 21 and 23. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 16. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F6, F8 and F10 have recognition sequences selected from one or more of SEQ ID NO: 17, 19, 21 and 23.

In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F4, F6, F8 and F10 have recognition sequences selected from one or more of SEQ ID NOs: 16, 18, 20, 22, 24 and 25. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F5, F7, F9 and F11 have recognition sequences selected from one or more of SEQ ID NO: 17, 19, 21 and 23. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 16. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F7, F9 and F11 have recognition sequences selected from one or more of SEQ ID NO: 17, 19, 21 and 23.

In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F6, F8 and F10 have recognition sequences selected from one or more of SEQ ID NOs: 16, 18, 20, 22, 24 and 25. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F7, F9 and F11 have recognition sequences selected from one or more of SEQ ID NO: 17, 19, 21 and 23. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 16. In embodiments, an extended zinc finger peptide of the invention has 11 zinc finger domains, wherein fingers F2, F4, F7, F9 and F11 have recognition sequences selected from one or more of SEQ ID NO: 17, 19, 21 and 23.

Table 2 below summarises preferred recognition sequence arrangements of the extended poly-zinc finger peptides (e.g. activator peptides) of these aspects and embodiments of the invention. In this table, as noted above, one or more sequence of SEQ ID NO: 14 may be substituted with the sequence of SEQ ID NO: 15. Similarly, in the table below, one or more sequence of SEQ ID NO: 26 may be substituted with the sequence of SEQ ID NO: 27. Similarly, in the table below, any of sequences SEQ ID NOs: 16 to 25 may be substituted with any other of sequences SEQ ID NOs: 16 to 25.

Finger number of	Zinc Finger Number (N to C terminal direction)				ZFP Array
	F1, F3 [#]	F2	F4, F6, F8, F10 etc.	F5, F7, F9, F11 etc.	

zinc finger peptide					
Recognition Sequence	SEQ ID NO: 10	SEQ ID NO: 10	SEQ ID NO: 10	SEQ ID NO: 10	K
	SEQ ID NO: 11	SEQ ID NO: 11	SEQ ID NO: 11	SEQ ID NO: 11	L
	SEQ ID NO: 12	SEQ ID NO: 12	SEQ ID NO: 12	SEQ ID NO: 12	M
	SEQ ID NO: 13	SEQ ID NO: 13	SEQ ID NO: 13	SEQ ID NO: 13	N
	SEQ ID NO: 14	SEQ ID NO: 15	SEQ ID NO: 15	SEQ ID NO: 14	O
	SEQ ID NO: 14	SEQ ID NO: 15	SEQ ID NO: 14	SEQ ID NO: 15	P
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 16	Q
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 16	SEQ ID NO: 17	R
	SEQ ID NO: 16	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 16	S
	SEQ ID NO: 16	SEQ ID NO: 19	SEQ ID NO: 16	SEQ ID NO: 19	T
	SEQ ID NO: 16	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 16	U
	SEQ ID NO: 16	SEQ ID NO: 20	SEQ ID NO: 16	SEQ ID NO: 20	V
	SEQ ID NO: 16	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 16	W
	SEQ ID NO: 16	SEQ ID NO: 22	SEQ ID NO: 16	SEQ ID NO: 22	X
	SEQ ID NO: 16	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 16	Y
	SEQ ID NO: 16	SEQ ID NO: 24	SEQ ID NO: 16	SEQ ID NO: 24	Z
	SEQ ID NO: 16	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 16	AA
	SEQ ID NO: 16	SEQ ID NO: 25	SEQ ID NO: 16	SEQ ID NO: 25	AB
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 18	AC
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 18	SEQ ID NO: 17	AD
	SEQ ID NO: 18	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 18	AE
	SEQ ID NO: 18	SEQ ID NO: 19	SEQ ID NO: 18	SEQ ID NO: 19	AF
	SEQ ID NO: 18	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 18	AG
	SEQ ID NO: 18	SEQ ID NO: 20	SEQ ID NO: 18	SEQ ID NO: 20	AH
	SEQ ID NO: 18	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 18	AI
	SEQ ID NO: 18	SEQ ID NO: 22	SEQ ID NO: 18	SEQ ID NO: 22	AJ
	SEQ ID NO: 18	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 18	AK
	SEQ ID NO: 18	SEQ ID NO: 24	SEQ ID NO: 18	SEQ ID NO: 24	AL
	SEQ ID NO: 18	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 18	AM
	SEQ ID NO: 18	SEQ ID NO: 25	SEQ ID NO: 18	SEQ ID NO: 25	AN
	SEQ ID NO: 21	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 21	AO
	SEQ ID NO: 21	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 17	AP
	SEQ ID NO: 21	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 21	AQ
	SEQ ID NO: 21	SEQ ID NO: 19	SEQ ID NO: 21	SEQ ID NO: 19	AR
	SEQ ID NO: 21	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 21	AS
	SEQ ID NO: 21	SEQ ID NO: 20	SEQ ID NO: 21	SEQ ID NO: 20	AT
	SEQ ID NO: 21	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 21	AU
	SEQ ID NO: 21	SEQ ID NO: 22	SEQ ID NO: 21	SEQ ID NO: 22	AV
	SEQ ID NO: 21	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 21	AW
	SEQ ID NO: 21	SEQ ID NO: 24	SEQ ID NO: 21	SEQ ID NO: 24	AX
	SEQ ID NO: 21	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 21	AY
	SEQ ID NO: 21	SEQ ID NO: 25	SEQ ID NO: 21	SEQ ID NO: 25	AZ
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 20	BA
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 22	BB
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 24	BC
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 25	BD
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 20	BE
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 22	BF
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 24	BG
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 25	BH
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 20	BI
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 22	BJ
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 24	BK
	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 25	BL
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 20	BM

	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 22	BN
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 24	BO
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 25	BP
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 20	BQ
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 22	BR
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 24	BS
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 25	BT
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 20	BU
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 22	BV
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 24	BW
	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 25	BX
	SEQ ID NO: 26	SEQ ID NO: 27	SEQ ID NO: 27	SEQ ID NO: 26	BY
	SEQ ID NO: 26	SEQ ID NO: 27	SEQ ID NO: 26	SEQ ID NO: 27	BZ

Table 2: Exemplary zinc finger recognition helix arrangements of zinc finger peptides according to the invention for binding 5'- AAG -3' repeat sequences, e.g. for treating Friedreich's ataxia (FRDA). Zinc finger peptides disclosed in this table may have from 8 to 32 fingers, for example, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 or 18 zinc finger domains. Any sequence with RK at positions +4 and +5 may be exchanged for RI; and any sequence with LT at positions +4 and +5 may be exchanged for LS and all such combinations are explicitly disclosed herein. Similarly, any sequence with RK or RI at positions +4 and +5 may be exchanged for LT or LS; and any sequence with LT or LS at positions +4 and +5 may be exchanged for RK or RI and all such combinations are explicitly disclosed herein. #All / any SEQ ID NO: 16 sequences at finger 3 (F3) may be exchanged for SEQ ID NO: 18, and all / any SEQ ID NO: 18 sequences at finger 1 (F1) may be exchanged for SEQ ID NO: 16.

Suitably, the zinc finger activator peptides of the invention for binding to 5'-AAG-3' repeats comprise have or consist of 11-zinc finger domains which are arranged in tandem. Exemplary 11-zinc finger peptide sequences of the invention for binding to 5'-AAG-3' repeat sequences in frataxin genes comprise a polypeptide having the sequence of SEQ ID NO: 102 (ZF2 peptide), SEQ ID NO: 103 (ZF2a peptide) or SEQ ID NO: 104 (ZF2b peptide), as shown in Table 9. The invention also encompasses zinc finger activator peptides comprising any of SEQ ID NOs: 102, 103 and 104 fused (or covalently linked as described herein to a suitable transcriptional activator domain – particularly an activator domain compatible with mouse and/or human cell expression; such as p65 (SEQ ID NOs: 122 or 136) or VP64 (SEQ ID NO: 125); for example. SEQ ID NOs: 105 and 106, (ZF2-p65, and ZF2-VP64, respectively). In aspects and embodiments, the invention also encompasses polypeptides having 90% or more, 95% or more, such as 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequences of SEQ ID NOs: 102 to 106.

C. Poly Zinc Finger Activator Proteins for Targeting the Frataxin Promoter

For targeting non-repeat genome sequences, as noted above, a register of approx. 18 nucleotides may provide a unique sequence within a complex genome. However, since promoter regions can include relatively conserved sequence portions (e.g. binding sites for wild-type transcription factors), it can be desirable to design zinc finger peptide arrays to bind to sequences having more than 18 bases. Zinc finger peptide frameworks of the invention for targeting the frataxin promoter may, therefore, have from 6 to 12 zinc finger domains, and particularly from 6 to 9 zinc finger domains, e.g. 6, 7, 8 or 9 zinc finger domains. Preferred zinc finger peptide activators according to these aspects and embodiments of the

invention may comprise, have or consist of 6 or 9 zinc finger domains; and particularly preferred zinc finger peptides of these aspects and embodiments of the invention comprise, have or consist of 9 zinc finger domains.

5 In various embodiments, zinc finger peptide activators according to these aspects and embodiments may be based on the frameworks of Structures I to V as defined above and in our previous publications, WO 2012/049332; WO 2017/077329). Alternatively, such zinc finger peptides may be constructed from 2-finger building blocks, as described, for example, in Moore *et al.* (2001), *Proc. Natl. Acad. Sci. USA*, 98: 1437–1441. Zinc finger activator proteins of the invention may also be constructed from 3-finger
10 building blocks, as is known in the art (Moore *et al.* (2001) *Proc. Natl. Acad. Sci. USA* 98(4): 1437-1441; and Kim & Pabo (1998) *Proc. Natl. Acad. Sci. USA* 95(6): 2812-2817), or from a combination of 2 and 3 finger building blocks, as desired. In some embodiments a 9- or 6-finger binding unit may be provided by multiple 3-zinc finger units, each of which is covalently linked together (e.g. with a canonical-like or flexible linker), or which may be provided with complementary dimerisation domains to form a 6- or 9-
15 zinc finger arrays.

The arrays of zinc finger domains in the zinc finger activator proteins of these aspects and embodiments typically comprise canonical linker sequences, short flexible (canonical-like) linker sequences and, in some embodiments long flexible linker sequences. Thus, as described in relation to the extended poly-
20 zinc finger peptides for binding to GAA- (or AAG-) repeat sequences, in some embodiments one or more pairs of adjacent zinc finger domains may be separated by short canonical linker sequences; and/or one or more pairs of adjacent zinc finger domains may be separated by short flexible linker sequences (e.g. of 6 or 7 amino acids), 'canonical-like' linker sequences. In some embodiments, one or more pairs of adjacent zinc finger domains of a zinc finger peptide may be separated by long flexible
25 linker sequences, for example, comprising 8 or more amino acids, such as between 8 and 50 amino acids as described elsewhere herein. Suitably, the zinc finger activator proteins of these aspects and embodiments may comprise zinc finger domains arranged in tandem and linked to each other by canonical or canonical-like linker sequences only.

30 In some embodiments, the zinc finger activator proteins may comprise two or more sub-arrays of from 2 to 6 directly adjacent zinc finger domains (or any combination thereof) separated by long flexible (or structured) linkers. Preferably, such poly-zinc finger peptides are arranged in two sub-arrays of 3-, 4- or 5-finger units separated by a long flexible linker to provide the desired 6- to 12-zinc finger peptide.

35 As described above, the amino acid sequence of the recognition helix of each zinc finger domain of a poly-zinc finger activator peptide is suitably determined by the nucleic acid sequence of the target nucleic acid triplet (or staggered quadruplet). Therefore, where the zinc finger peptides are designed to target unique, non-repeating sequences (e.g. of the frataxin promoter), the recognition sequences of each zinc finger domain of a poly-zinc finger peptide of these aspects and embodiments is generally
40 different from its neighbouring zinc finger domains. However, since the main nucleic acid-binding interactions are formed with the residues in the -1, 3 and 6 positions (as previously described), the

amino acids at the helix 'framework' positions (outside or the -1, 3 and 6 positions) may be similar or the same for adjacent zinc finger domains.

For example, as described herein, in embodiments of the invention, the residues at positions +4 and +5 may suitably be, respectively, L and T. In this way host matching for *in vivo* mouse or human applications may be improved. In such embodiments, the residues R and K, respectively, may alternatively be used in the 4 and 5 positions. In some embodiments, adjacent zinc finger domains may have alternating (regular or irregular) patterns of amino acids at the 4 and 5 positions; for example, the residues L and T in the 4 and 5 positions may be used interchangeably with the residues R and K, depending – for example – on the overall charge of amino acid residues in the recognition helix of a zinc finger domain, or within an array of zinc finger domains. In some embodiments, as previously described, the T at position 5 of the recognition helix may be substituted for an S residue, without adversely affecting binding specificity of the zinc finger domain; and/or a K at position 5 may be substituted for an I residue.

In some embodiments, the first zinc finger of a zinc finger array (or sub-array – where sub-arrays are separated from each other by long, flexible linkers) has a recognition sequence wherein the residues at positions 4 and 5 are, respectively, L and T. In such embodiments, all remaining recognition sequences of the array (or sub-array) may have the residues R and K, respectively, in the 4 and 5 positions; may have the residues L and T, respectively, in the 4 and 5 positions; or may include a mixture of R and K or L and T residues in the 4 and 5 positions, respectively.

It will be appreciated that various positions of the frataxin promoter region from -1000 to 0 bases upstream of the transcriptional start site (TSS; SEQ ID NO: 67) may be targeted by zinc finger activator proteins in order to increase the expression of frataxin in pathogenic genomes, and such zinc finger activator proteins are encompassed within the scope of this disclosure. For example, nucleic acid target regions for zinc finger activator peptides of the invention may be the frataxin promoter region from -1000 to -300 (i.e. positions 1 to 700 of SEQ ID NO: 67) bases upstream of the TSS; from -1000 to -500 bases upstream of the TSS (i.e. positions 1 to 500 of SEQ ID NO: 67); from -1000 to -600 bases upstream of the TSS (i.e. positions 1 to 400 of SEQ ID NO: 67); and from -980 to -620 bases upstream of the TSS (i.e. positions 20 to 380 of SEQ ID NO: 67). In embodiments, nucleic acid target regions for zinc finger activator peptides of the invention are the frataxin promoter region from -700 to -600 bases upstream of the TSS (SEQ ID NO: 68); or from -660 to -610 bases upstream of the TSS (SEQ ID NO: 69). In other embodiments, nucleic acid target regions for zinc finger activator peptides of the invention are the frataxin promoter region from -1000 to -900 bases upstream of the TSS (SEQ ID NO: 70); or from -990 to -940 bases upstream of the TSS (SEQ ID NO: 71). For example, a suitable target binding site for the zinc finger activator peptides of the invention may have 36, 33, 30, 27, 24, 21 or 18 contiguous nucleotides from any of the frataxin promoter region sequences disclosed herein. Beneficially, the target binding site may have 33, 27 or 18 contiguous nucleotides of the target region; and more beneficially, the target binding site may have 27 or 18 contiguous nucleotides of the selected region of the frataxin promoter.

For example, particularly suitable target / binding sites for zinc finger peptides of the invention include:

- (I) 3'- ACGCTAGAGCCGAGTGACGTTTCGAGACGGAGGGCCAAAG -5'
SEQ ID NO: 72 (5' to 3' orientation); or
- 5 (II) 3'- GAGCCGAGTGACGTTTCGAGACGGAGGG -5'
SEQ ID NO: 73 (5' to 3' orientation); and
- (III) 3'- GATGTCCACACGTGGTGGTGTGGGTCGA -5'
SEQ ID NO: 74 (5' to 3' orientation); and
- (IV) 3'- CGTGGTGGTGTGGGTCGA -5'
10 SEQ ID NO: 75 (5' to 3' orientation).

According to the invention, in various embodiments, zinc finger peptides are engineered to bind to at least 18 bases of a sequence selected from SEQ ID NOs: 67 to 75. In embodiments, zinc finger peptides are engineered to bind to at least 21, or at least 24 bases of a sequence selected from SEQ ID NOs:

15 67 to 74. In some beneficial embodiments, zinc finger peptides of the invention are engineered to bind to 27 bases of a sequence selected from SEQ ID NOs: 67 to 74. In particular embodiments, zinc finger peptides of the invention are engineered to bind to 18 or more bases of SEQ ID NOs: 74 or 75; and especially to SEQ ID NO 75. In other particular embodiments, zinc finger peptides of the invention are engineered to bind to 27 or more bases of SEQ ID NOs: 72 or 73; and especially to SEQ ID NO: 73.

20 For clarity, as used herein, the nomenclature 'ZF3' will be used with respect to zinc finger peptides engineered for binding to promoter region (I) or (II) above, and the nomenclature 'ZF4' will be used with respect to zinc finger peptides engineered for binding to promoter regions (III) or (IV) above.

25 According to the invention, ZF3 zinc finger recognition sequences (i.e. positions X^{-1} , X^{+1} , X^{+2} , X^{+3} , X^{+4} , X^{+5} and X^{+6} in Formulas 2, 4 and 6, or Structure V above) for binding to promoter region (II) may be represented by the amino acid sequences below:

30 Finger 1 (F1) SEQ ID NO: 28 R(S/N)SN(L/R)(T/S/K/I)R
SEQ ID NO: 29 R(S/N)SNL(T/S)R
SEQ ID NO: 30 RSSNLTR

35 Finger 2 (F2) SEQ ID NO: 31 D(S/N)SV(L/R)(T/S/K/I)R
SEQ ID NO: 32 D(S/N)SVL(T/S)R
SEQ ID NO: 33 DSSVLTR

40 Finger 3 (F3) SEQ ID NO: 34 Q(S/N)SH(L/R)(T/S/K/I)K
SEQ ID NO: 35 Q(S/N)SHL(T/S)K
SEQ ID NO: 36 QSSHLTK

Finger 4 (F4) SEQ ID NO: 37 R(S/N)DN(L/R)(T/S/K/I)T
SEQ ID NO: 38 R(S/N)DNL(T/S)T

		SEQ ID NO: 39	RSDNLTT
	Finger 5 (F5)	SEQ ID NO: 40	R(S/N)ST(L/R)(T/S/K/I)T
		SEQ ID NO: 41	R(S/N)STL(T/S)T
5		SEQ ID NO: 42	RSSTLTT
	Finger 6 (F6)	SEQ ID NO: 43	D(S/N)SH(L/R)(T/S/K/I)Q
		SEQ ID NO: 44	D(S/N)SHL(T/S)Q
		SEQ ID NO: 45	DSSHLTQ
10			
	Finger 7 (F7)	SEQ ID NO: 46	R(S/N)SN(L/R)(T/S/K/I)E
		SEQ ID NO: 47	R(S/N)SNL(T/S)E
		SEQ ID NO: 48	RSSNLTE
	Finger 8 (F8)	SEQ ID NO: 49	R(S/N)SH(L/R)(T/S/K/I)Q
15		SEQ ID NO: 50	R(S/N)SHL(T/S)Q
		SEQ ID NO: 51	RSSHLTQ
	Finger 9 (F9)	SEQ ID NO: 52	R(S/N)SH(L/R)(T/S/K/I)R
20		SEQ ID NO: 53	R(S/N)SHL(T/S)R
		SEQ ID NO: 54	RSSHLTR

With respect to ZF3 zinc finger peptides, also included are embodiments comprising, having or consisting of 6, 7 or 8 consecutive zinc fingers domains having sequences selected from those given for F1 to F9 above; for example, poly-zinc finger peptides (and corresponding activators) comprising, having or consisting of fingers: F1 to F6, F2 to F7, F3 to F8, F4 to F9; F1 to F7, F2 to F8, F3 to F9, F1 to F8, F2 to F9 and F1 to F9 above are encompassed within the scope of the invention.

Thus, in embodiments, there is provided an engineered zinc finger (DNA-binding) peptide comprising from 6 to 9, such as 6, 7, 8 or 9 zinc finger domains, and particularly 9 zinc finger domains having the zinc finger recognition sequences set out above for Fingers F1 to F9. Such zinc finger peptides may be comprised within zinc finger peptides of from 7 to 11 zinc finger domains (depending on the length of the array selected).

Exemplary poly-zinc finger activator peptides of these aspects and embodiments include 9 zinc finger domains, wherein fingers F1 to F9 have recognition sequences selected from SEQ ID NOs: 28, 31, 34, 37, 40, 43, 46, 49 and 52, respectively; particularly selected from SEQ ID NOs: 29, 32, 35, 38, 41, 44, 47, 50 and 53, respectively; and according to SEQ ID NOs: 30, 33, 36, 39, 42, 45, 48, 51 and 54, respectively.

Table 3 below summarising beneficial recognition sequence arrangements for poly-zinc finger activator peptides of these aspects and embodiments of the invention.

Finger number of zinc finger peptide	Zinc Finger Number (N to C terminal direction)									ZFP Array
	F1	F2	F3	F4	F5	F6	F7	F8	F9	
Recognition Sequence: SEQ ID NO.	28	31	34	37	40	43				CA
	31	34	37	40	43	46				CB
	34	37	40	43	46	49				CC
	37	40	43	46	49	52				CD
	28	31	34	37	40	43	46			CE
	31	34	37	40	43	46	49			CF
	34	37	40	43	46	49	52			CG
	28	31	34	37	40	43	46	49		CH
	31	34	37	40	43	46	49	52		CI
	28	31	34	37	40	43	46	49	52	CJ
	29	32	35	38	41	44				CK
	32	35	38	41	44	47				CL
	35	38	41	44	47	50				CM
	38	41	44	47	50	53				CN
	29	32	35	38	41	44	47			CO
	32	35	38	41	44	47	50			CP
	35	38	41	44	47	50	53			CQ
	29	32	35	38	41	44	47	50		CR
	32	35	38	41	44	47	50	53		CS
	29	32	35	38	41	44	47	50	53	CT
	30	33	36	39	42	45				CU
	33	36	39	42	45	48				CV
	36	39	42	45	48	51				CW
	39	42	45	48	51	54				CX
	30	33	36	39	42	45	48			CY
	33	36	39	42	45	48	51			CZ
	36	39	42	45	48	51	54			DA
	30	33	36	39	42	45	48	51		DB
	33	36	39	42	45	48	51	54		DC
	30	33	36	39	42	45	48	51	54	DD

Table 3: Exemplary zinc finger recognition helix arrangements of zinc finger activator peptides according to the invention for binding to the frataxin promoter region sequence 5'-GGGAGGCAGAGCTTGAGCCGAG-3' (SEQ ID NO: 73). Zinc finger peptides disclosed in this table may have from 6 to 9 zinc finger domains, as illustrated. Zinc finger peptides according to the invention may comprise, have or consist of zinc finger domain arrays according to any of CB to DE as indicated above; and particularly, may comprise, have or consist of zinc finger domain arrays having the pattern of CK, CU or DE. Zinc finger peptides of the invention poly-zinc finger peptides having more than 6, 7, 8 or 9 zinc finger domains, wherein a zinc finger array having 6, 7, 8 or 9 consecutive zinc finger domains (according to Table 3 above) is present within a longer poly-zinc finger peptide, e.g. of from 7 to 11 zinc finger domains.

Suitably, the zinc finger activator peptides of the invention for binding to the 5'-GGGAGGCAGAGCTTGAGCCGAG-3' target sequence comprise, have or consist of 9-zinc finger domains which are arranged in tandem. An exemplary 9-zinc finger peptide sequences of the invention for binding to this region of the frataxin genes promoter comprises a polypeptide having the sequence of SEQ ID NO: 116 (ZF3 peptide), as shown in Table 9. The invention also encompasses zinc finger activator peptides comprising SEQ ID NO: 116 fused (or covalently linked as described

herein to a suitable transcriptional activator domain – particularly an activator domain compatible with mouse and/or human cell expression; such as p65 (SEQ ID NOs: 122 or 136) or VP64 (SEQ ID NO: 125); for example. SEQ ID NOs: 117 and 118, (ZF3-p65, and ZF3-VP64, respectively). In aspects and embodiments, the invention also encompasses polypeptides having 90% or more, 95% or more, such as 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequences of SEQ ID NOs: 116 to 118.

According to the invention, ZF4 zinc finger recognition sequences (i.e. positions X^{-1} , X^{+1} , X^{+2} , X^{+3} , X^{+4} , X^{+5} and X^{+6} in Formulas 2, 4 and 6, or Structure V above) for binding to promoter region (III) may be represented by the amino acid sequences below:

15	Finger 1 (F1)	SEQ ID NO: 55	D(S/N)SH(L/R)(T/S/K/I)K
		SEQ ID NO: 56	D(S/N)SHL(T/S)K
		SEQ ID NO: 57	DSSHLTK
20	Finger 2 (F2)	SEQ ID NO: 58	R(S/N)DH(L/R)(T/S/K/I)T
		SEQ ID NO: 59	R(S/N)DHL(T/S)T
		SEQ ID NO: 60	RSDHLTT
25	Finger 3 (F3)	SEQ ID NO: 58	R(S/N)DH(L/R)(T/S/K/I)T
		SEQ ID NO: 59	R(S/N)DHL(T/S)T
		SEQ ID NO: 60	RSDHLTT
30	Finger 4 (F4)	SEQ ID NO: 61	R(S/N)DT(L/R)(T/S/K/I)R
		SEQ ID NO: 62	R(S/N)DTL(T/S)R
		SEQ ID NO: 63	RSDLTR
35	Finger 5 (F5)	SEQ ID NO: 58	R(S/N)DH(L/R)(T/S/K/I)T
		SEQ ID NO: 59	R(S/N)DHL(T/S)T
		SEQ ID NO: 60	RSDHLTT
40	Finger 6 (F6)	SEQ ID NO: 64	D(S/N)SH(L/R)(T/S/K/I)Q
		SEQ ID NO: 65	D(S/N)SHR(K/I)Q
		SEQ ID NO: 66	DSSHRKQ

With respect to ZF4 zinc finger peptides of the invention, also included are embodiments comprising, having or consisting of 6 or more, such as 6, 7, 8, 9, 10 or 11 zinc finger domains wherein 6 consecutive zinc finger domains of the peptide have sequences selected from those given for F1 to F6 above.

Exemplary poly-zinc finger activator peptides of these aspects and embodiments have 6 to 11 zinc finger domains, wherein six consecutive zinc fingers have recognition sequences selected from SEQ ID NOs: 55, 58, 58, 61, 58 and 64, respectively; particularly selected from SEQ ID NOs: 56, 59, 59, 62,

59 and 65, respectively; and especially selected from SEQ ID NOs: 57, 60, 60, 63, 60 and 66, respectively.

Table 4 below summarising preferred recognition sequence arrangements of the poly-zinc finger activator peptides of these aspects and embodiments of the invention.

Finger number of zinc finger peptide or of each sub-array of a zinc finger peptide	Zinc Finger Number (N to C terminal direction)						ZFP Array
	F1	F2	F3	F4	F5	F6	
Recognition Sequence:	55	58	58	61	58	64	DE
SEQ ID NO.	56	59	59	62	59	65	DF
	57	60	60	63	60	66	DG

Table 4: Exemplary zinc finger recognition helix arrangements of zinc finger activator peptides according to the invention for binding to the frataxin promoter region sequence 5'-AGCTGGGTGTGGTGGTGC-3' (SEQ ID NO: 75). Zinc finger peptides disclosed in this table may have from 6 to 11 zinc finger domains, as illustrated. Zinc finger peptides according to the invention may comprise, have or consist of zinc finger domain arrays according to any of DF to DH as indicated above; and particularly, may comprise, have or consist of zinc finger domain arrays having the pattern of DH.

Suitably, the zinc finger activator peptides of the invention for binding to the 5'-AGCTGGGTGTGGTGGTGC-3' target sequence comprise, have or consist of 6-zinc finger domains which are arranged in tandem. An exemplary 6-zinc finger peptide sequences of the invention for binding to this region of the frataxin genes promoter comprises a polypeptide having the sequence of SEQ ID NO: 119 (ZF4 peptide), as shown in Table 9. The invention also encompasses zinc finger activator peptides comprising SEQ ID NO: 119 fused (or covalently linked as described herein to a suitable transcriptional activator domain – particularly an activator domain compatible with mouse and/or human cell expression; such as p65 (SEQ ID NOs: 122 or 136) or VP64 (SEQ ID NO: 125); for example. SEQ ID NOs: 120 and 121, (ZF4-p65, and ZF4-VP64, respectively). In aspects and embodiments, the invention also encompasses polypeptides having 90% or more, 95% or more, such as 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequences of SEQ ID NOs: 119 to 121.

D. Zinc Finger Derivatives and Associated Sequences

The invention also encompasses derivatives of the zinc finger peptides of the invention. In this regard, it will be appreciated that modifications, such as amino acid substitutions may be made at one or more positions in the peptide without adversely affecting its physical properties (such as binding specificity or affinity). By 'derivative' of a zinc finger peptide it is meant a peptide sequence that has the desired activity (e.g. binding affinity for a selected target sequence, especially poly GAA-repeat sequences), but that includes one or more mutations or modifications to the primary amino acid sequence having the desired activity. Thus, a derivative of the invention may have one or more (e.g. 1, 2, 3, 4, 5 or more) chemically modified amino acid side chains, such as pegylation, sialylation and glycosylation

modifications. In addition, or alternatively, a derivative may contain one or more (e.g. 1, 2, 3, 4, 5 or more) amino acid mutations, substitutions, deletions or combinations thereof to the primary sequence of a selected poly-zinc finger peptide. Accordingly, the invention encompasses the results of maturation experiments conducted on a selected zinc finger peptide or a zinc finger peptide framework to improve or change one or more characteristics of the initially identified peptide. By way of example, one or more amino acid residues of a selected zinc finger domain may be randomly or specifically mutated (or substituted) using procedures known in the art (e.g. by modifying the encoding DNA or RNA sequence). The resultant library or population of derivatised peptides may further be selected – by any known method in the art – according to predetermined requirements: such as improved specificity against particular target sites; or improved drug properties (e.g. solubility, bioavailability, immunogenicity etc.). One benefit of the invention is improved compatibility with the host / target organism as assessed by sequence similarity to known host peptide sequences and/or immunogenicity / adverse immune response to the heterologous peptide when expressed. Peptides selected to exhibit such additional or improved characteristics and that display the activity for which the peptide was initially selected are derivatives of the zinc finger peptides of the invention and also fall within the scope of the invention.

Zinc finger frameworks of the invention may be diversified at one or more positions in order to improve their compatibility with the host system in which it is intended to express the proteins. In particular, specific amino acid substitutions may be made within the zinc finger peptide sequences and in any additional peptide sequences (such as effector domains) to reduce or eliminate possible immunological responses to the expression of these heterologous peptides *in vivo*. Target amino acid residues for modification or diversification are particularly those that create non-host amino acid sequences or epitopes that might not be recognised by the host organism and, consequently, might elicit an undesirable immune response. In some embodiments the framework is diversified or modified at one or more of amino acids positions -1, 1, 2, 3, 4, 5 and 6 of the recognition sequence. The polypeptide sequence changes may conveniently be achieved by diversifying or mutating the nucleic acid sequence encoding the zinc finger peptide frameworks at the codons for at least one of those positions, so as to encode one or more polypeptide variant. Particular diversification strategies in relation to positions 4 and 5 of the zinc finger domain recognition helix have been described elsewhere herein. All such nucleic acid and polypeptide variants are encompassed within the scope of the invention.

The amino acid residues at each of the selected positions may be non-selectively randomised, i.e. by allowing the amino acid at the position concerned to be any of the 20 common naturally occurring amino acids; or may be selectively randomised or modified, i.e. by allowing the specified amino acid to be any one or more amino acids from a defined sub-group of the 20 naturally occurring amino acids. It will be appreciated that one way of creating a library of mutant peptides with modified amino acids at each selected location, is to specifically mutate or randomise the nucleic acid codon of the corresponding nucleic acid sequence that encodes the selected amino acid. On the other hand, given the knowledge that has now accumulated in relation to the sequence specific binding of zinc finger domains to nucleic acids, in some embodiments it may be convenient to select a specific amino acid (or small sub-group of amino acids) at one or more chosen positions in the zinc finger domain, for example, where it is known that a specific amino acid provides optimal binding to a particular nucleotide residue in a specific

target sequence. In accordance with the invention, a predicted optimal interaction may be introduced when not already present (e.g. to optimise binding affinity in the case of a zinc finger peptide activator); or a predicted optimal interaction may be removed when it is already present and it is desired to reduce the binding affinity of the zinc finger peptide for the target sequence (e.g. in the case of a zinc finger repressor according to the invention). The resultant peptides or frameworks may be considered to be the result of rational or 'intelligent' design. Conveniently the whole of the zinc finger recognition sequence may be selected by intelligent design and inserted / incorporated into an appropriate zinc finger framework both of which, ideally, are derived from the intended host organism, such as mouse or human. The person of skill in the art is well aware of the codon sequences that may be used in order to specify one or more than one particular amino acid residue within a library. Preferably all amino acid positions in each zinc finger domain and in any additional peptide sequences (such as effector domains and leader sequences) are chosen from known wild-type sequences from the host organism in which the protein is intended to be used.

Taking into account that minor modifications to the primary sequence of the peptides / proteins of the invention can be made without substantially altering the scope of the claimed invention, the invention should be considered to encompass, in addition, any polypeptide sequences that are substantially the same as the specific amino acid sequences disclosed herein. For example, the claimed invention encompasses polypeptide sequences that have at least 80% identity to the SEQ ID NOs of the polypeptide sequences disclosed herein; at least 85% identity, at least 90% identity, at least 95% identity, at least 98% identity, at least 99% identity or approx. 100% identity to the polypeptide sequences of the SEQ ID NOs explicitly disclosed herein.

Zinc Finger Peptide Modulators and Effectors

It will be appreciated that the zinc finger peptide framework sequences of the invention may further include optional (N-terminal) leader sequences, such as: amino acids to aid expression (e.g. N-terminal Met-Ala or Met-Gly dipeptide); purification tags (e.g. FLAG-tags); and localisation / targeting sequences (e.g. nuclear localisation sequences (NLS), such as PKKKRKV (SV40 NLS, SEQ ID NO: 107); PKKRRKVT (human protein KIAA2022, SEQ ID NO: 108); or RIRKKLR (mouse primase p58 NLS9, SEQ ID NO: 109). Thus, a suitable leader sequence for use in conjunction with zinc finger peptide sequences of the invention includes MGRIRKKLRLAERP for expression and cellular localisation in mouse (SEQ ID NO: 110) and MGPKKRRKVTGERP for expression and cellular localisation in human cells (SEQ ID NO: 111).

Also, the peptides of the invention may optionally include additional C-terminal sequences, such as linker sequences for fusing zinc finger domains to effector molecules, and the effector molecules themselves. Other sequences may be employed for cloning purposes. The sequences of any N- or C-terminal sequences may be varied, typically without altering the binding activity of the zinc finger peptide framework, and such variants are encompassed within the scope of the invention.

Suitably a zinc finger peptide of the invention for expression and use in mouse or human respectively, does not include purification tags where it is not intended to purify the zinc finger-containing peptide, e.g. where gene regulatory and/or therapeutic activities are intended. Thus, for reason of improved host-matching (reduced toxicity and reduced immunogenicity) the peptides and polypeptides of the invention are preferably devoid of peptide purification tags and the like, which are not found in endogenous, wild-type proteins of a host organism.

Particularly preferred polypeptides of the invention comprise an appropriate nuclear localisation sequence arranged N-terminal of a poly-zinc finger peptide, which is itself arranged N-terminal to an effector domain that may repress expression of a target gene. Effector domains are conveniently attached to the poly-zinc finger peptide covalently, such as by a peptide linker sequence as disclosed elsewhere herein.

While the zinc finger peptides of the invention may have useful biological properties in isolation, they can also be given useful biological functions by the addition of effector domains. Therefore, in some cases it is desirable to conjugate a zinc finger peptide of the invention to one or more non-zinc finger domain, thus creating chimeric or fusion zinc finger peptides. It may also be desirable, in some instances, to create a multimer (e.g. a dimer), of a zinc finger peptide of the invention – for example, to bind more than one target sequence simultaneously, which target sequences may be the same or different.

Thus, having identified a desirable zinc finger peptide, an appropriate effector or functional group may then be attached, conjugated or fused to the zinc finger peptide. The resultant protein of the invention, which comprises at least a zinc finger portion (of more than one zinc finger domain) and a non-zinc finger effector domain, portion or moiety may be termed a 'fusion', 'chimeric' or 'composite' zinc finger peptide. Beneficially, the zinc finger peptide will be linked to the other moiety at a position and/or via a linker that does not interfere with the activity of either moiety.

A 'non-zinc finger domain' (or moiety) as used herein, refers to an entity that does not contain a zinc finger ($\beta\beta\alpha$ -) fold. Thus, non-zinc finger moieties include nucleic acids and other polymers, peptides, proteins, peptide nucleic acids (PNAs), antibodies, antibody fragments, and small molecules, amongst others.

Potential effector domains include transcriptional repressor domains, transcriptional activator domains, transcriptional insulator domains, chromatin remodelling, condensation or decondensation domains, nucleic acid or protein cleavage domains, dimerisation domains, enzymatic domains, signalling / targeting sequences or domains, or any other appropriate biologically functional domain. Other domains that may also be appended to zinc finger peptides include peptide sequences involved in protein transport, localisation sequences (e.g. subcellular localisation sequences, nuclear localisation, protein targeting) or signal sequences. Zinc finger peptides can also be fused to epitope tags, e.g. for use to signal the presence or location of a target nucleotide sequence recognised by the zinc finger peptide. Functional fragments of any such domain may also be used.

Beneficially, zinc finger peptides and fusion proteins / polypeptides of the invention have transcriptional modulatory activity, and preferred biological effector domains include transcriptional activators, as well as their functional fragments. The effector domain can be directly derived from a basal or regulated transcription factor such as, for example, transactivators and proteins that bind to insulator or silencer sequences (see Choo & Klug (1995) *Curr. Opin. Biotech.* 6: 431-436; Choo & Klug (1997) *Curr. Opin. Str. Biol.* 7:117-125; and Goodrich *et al.* (1996) *Cell* 84: 825-830); or from receptors such as nuclear hormone receptors (Kumar & Thompson (1999) *Steroids* 64: 310-319); or co-activators (Ugai *et al.* (1999) *J. Mol. Med.* 77: 481-494).

Other useful functional domains for control of gene expression include, for example, protein-modifying domains such as histone acetyltransferases, kinases, methylases and phosphatases, which can activate genes by modifying DNA structure or the proteins that associate with nucleic acids (Wolffe (1996) *Science* 272: 371-372; and Hassig *et al.*, (1998) *Proc. Natl. Acad. Sci. USA* 95: 3519-3524). Additional useful effector domains include those that modify or rearrange nucleic acid molecules such as methyltransferases, endonucleases, ligases, recombinases, and nucleic acid cleavage domains (see for example, Smith *et al.* (2000) *Nucleic Acids Res.*, 17: 3361-9; WO 2007/139982 and references cited therein).

In embodiments, suitable transcriptional / gene activation domains for fusing to zinc finger peptides in order to produce a zinc finger activator protein of the invention include: the VP64 domain, SEQ ID NO: 125 (see Seipel *et al.*, (1996) *EMBO J.* 11: 4961-4968) and the herpes simplex virus (HSV) VP16 domain, SEQ ID NO: 124 (Hagmann *et al.* (1997) *J. Virol.* 71: 5952-5962; Sadowski *et al.* (1988) *Nature* 335: 563-564), which may be used in place of the VP64 domain in any of the aspects and embodiments disclosed herein; and transactivation domain 1 and/or 2 of the p65 subunit of nuclear factor- κ B (NF κ B; Schmitz *et al.* (1995) *J. Biol. Chem.* 270: 15576-15584; Schmitz and Baeuerle (1991) *EMBO J.* 10(12):3805-17) in human (SEQ ID NOs: 122 or 136). A corresponding mouse p65 activation domain (from RelA) may be used when the protein is intended to be used in mouse cells (SEQ ID NO: 123). Such zinc finger activator proteins of the invention are useful in upregulating the expression of wild-type gene products that are under (or not) expressed in a pathogenic condition.

All known methods of conjugating an effector domain to a peptide sequence are incorporated. The term 'conjugate' is used in its broadest sense to encompass all methods of attachment or joining that are known in the art, and is used interchangeably with the terms such as 'linked', 'bound', 'associated' or 'attached'. The effector domain(s) can be covalently or non-covalently attached to the binding domain: for example, where the effector domain is a polypeptide, it may be directly linked to a zinc finger peptide (e.g. at the C-terminus) by any suitable flexible or structured amino acid (linker) sequence (encoded by the corresponding nucleic acid molecule). Non-limiting suitable linker sequences for joining an effector domain to the C-terminus of a zinc finger peptide include, for example, LRQKDGGGSGGGGSGGGGSQLVSS (SEQ ID NO: 126), LRQKDGGGSGGGGSS (SEQ ID NO: 127), LRQKDGGGSGGGGS (SEQ ID NO: 128) and LRQKDGGGSGGGGS (SEQ ID NO: 129). Alternatively, a synthetic non-amino acid or chemical linker may be used, such as polyethylene glycol,

a maleimide-thiol linkage (useful for linking nucleic acids to amino acids), or a disulphide link. Synthetic linkers are commercially available, and methods of chemical conjugation are known in the art. It will be appreciated, however, that the amino acid sequences of such long, flexible linkers may not be critical, and, for example, the number of G and/or S repeats may be varied as desired, provided the resultant linker does not interfere with the activities of any associated effector domains.

Non-covalent linkages between a zinc finger peptide and an effector domain can be formed using, for example, leucine zipper / coiled coil domains, or other naturally occurring or synthetic dimerisation domains (Luscher & Larsson (1999) *Oncogene* 18: 2955-2966; and Gouldson *et al.* (2000) *Neuropsychopharm.* 23: S60-S77. Other non-covalent means of conjugation may include a biotin-(strept)avidin link or the like. In some cases, antibody (or antibody fragment)-antigen interactions may also be suitably employed, such as the fluorescein-antifluorescein interaction.

To cause a desired biological effect via modulation of gene expression, zinc finger peptides or their corresponding fusion peptides are allowed to interact with, and bind to, one or more target nucleotide sequence associated with the target gene, either *in vivo* or *in vitro* depending to the application. Beneficially, therefore, a nuclear localisation domain is attached to the DNA binding domain to direct the protein to the nucleus. As previously discussed, one useful nuclear localisation sequence is the SV40 NLS (PKKKRKV, SEQ ID NO: 107). Desirably, however, the nuclear localisation sequence is a host-derived sequence, such as the NLS from human protein KIAA2022 NLS (PKKRRKVT; NP_001008537.1, SEQ ID NO: 108) for use in humans; or the NLS from mouse primase p58 (RIRKKLR; GenBank: BAA04203.1, SEQ ID NO: 109) for use in mice.

Thus, preferred zinc finger-containing polypeptides of the invention include a nuclear localisation sequence (NLS), a poly-zinc finger peptide sequence and a transcriptional activator (e.g. p65-RelA activation domain). Particularly preferred poly-zinc finger peptide sequences of the disclosure include SEQ ID NOs: 97 to 106 and 116 to 121, which in embodiments are beneficially operable linked to one or more nuclear localisation sequence (NLS), a transcriptional activator domain (e.g. p65-RelA activation domain) and optionally signal peptide sequences as described herein.

In some embodiments, it may be advantageous to include more than one NLS as described herein; for example, between 2 and 5 NLSs; suitably 2 or 3 NLSs; and particularly 2. When more than one NLS is provided, said NLSs may suitably be arranged in tandem. NLS sequences generally provide a net positive charge, and arranging more than one NLS (e.g. 2, 3, 4 or 5) in tandem can enhance cell-penetration of the zinc finger-containing polypeptide by providing a concentration of positively charged amino acid residues.

In accordance with some beneficial embodiments, as described elsewhere, the zinc finger polypeptides of the invention may further include one or more protein secretion signal (SS) or signal peptide (SP) for promoting secretion of zinc finger polypeptides from the cell in which they are produced. A suitable protein secretion signal for use in human cells is the human BMP10 protein secretion signal, MGSLVLTLCALFCLAAYLVSG (SEQ ID NO: 131). In some such embodiments a nucleic acid or

polypeptide cleavage site may be incorporated between the signal peptide and the zinc finger peptide sequence of the encoded zinc finger polypeptide, for example, so that the signal peptides of some expressed polypeptides may be separated from the transcription factor portion of the zinc finger polypeptide before it is secreted. In this way, at least some expressed zinc finger polypeptide remains inside the cell in which it was expressed. Suitably, the cleavage sequence is the RIRR peptide cleavage site (SEQ ID NO: 132).

DNA regions from which to affect the up-regulation of specific genes may include promoters, enhancers or locus control regions (LCRs). In accordance with the invention, preferred target sequences for activation of pathogenic frataxin genes are found within the frataxin promoter region, e.g. from 0 to 2,000 bases upstream of the transcriptional start point; up to 1,500 bases upstream of the transcriptional start point; up to 1,200 bases upstream of the transcriptional start point; or up to 1,000 bases upstream of the transcriptional start point. In some embodiments, the selected zinc finger activator protein binding site is between 200 and 1,500 bases upstream of the transcriptional start point; between 300 and 1,200 bases upstream of the transcriptional start point; between 400 and 1,100 bases upstream of the transcriptional start point; or between 500 and 1,000 bases upstream of the transcriptional start point. Particularly beneficial binding sites for transcriptional activators of the invention are between 900 and 1,000 bases upstream of the transcriptional start point; and between 600 and 700 bases upstream of the transcriptional start point. Still more specifically, suitable binding sites for zinc finger activator proteins of the invention are found between 620 and 650 bases upstream of the transcriptional start point; and between 950 and 980 bases upstream of the transcriptional start point. Particularly preferred binding site sequences in the frataxin promoter are 5'-GAAACCGGGAGGCAGAGCTTGCAGTGAGCCGAGATCGCA-3' (SEQ ID NO: 72); 5'-GGGAGGCAGAGCTTGCAGTGAGCCGAG-3' (SEQ ID NO: 73); 5'-AGCTGGGTGTGGTGGTGCACACCTGTAG-3' (SEQ ID NO: 74); and 5'-AGCTGGGTGTGGTGGTGC-3' (SEQ ID NO: 75).

Surprisingly, the present inventors have also found that suitable DNA regions from which to affect the up-regulation of specific genes may include introns. In accordance with various aspects and embodiments of the invention, therefore, target sequences for activation of pathogenic frataxin genes are found within intron 1 of the frataxin gene, and particularly within the GAA-trinucleotide repeat sequences, which may include 66 or more, such as 300 or more, 500 or more, 600 or more, or 850 or more such repeats.

Nucleic Acids and Peptide Expression

The zinc finger peptides according to the invention and, where appropriate, the zinc finger peptide modulators (conjugate / effector molecules) of the invention may be produced by recombinant DNA technology and standard protein expression and purification procedures.

The invention also encompasses nucleic acid molecules that encode the peptide sequences of any aspects and embodiments of the invention, including their derivatives. In view of codon redundancy, it

will be appreciated that many slightly different nucleic acid sequences may accurately code for each of the zinc finger peptides of the invention, and each of these variants is encompassed within the scope of the present invention. The skilled person can readily determine suitable nucleic acid sequences for encoding each of the zinc finger peptides of the invention, and may select appropriate codon codes according to the system in which the zinc finger peptide is to be expressed (e.g. mouse or human). For example, any nucleic acid sequences that encode for the peptides of SEQ ID NOs: 97 to 106 and 116 to 121 are encompassed within the invention.

Furthermore, in view of the diversification that is possible in an encoding nucleic acid without even altering the amino acid sequence, the claimed invention encompasses polynucleotide sequences that have at least 70% identity to the polynucleotide SEQ ID NOs disclosed herein; at least 80% identity, at least 85% identity, at least 90% identity, at least 95% identity, at least 98% identity, at least 99% identity or approx. 100% identity to the polynucleotide sequences encoding the SEQ ID NOs explicitly disclosed herein.

The invention further provides nucleic acid constructs, such as expression vectors, that comprise nucleic acid encoding peptides and derivatives according to the invention.

In this regard, the DNA encoding the relevant peptide can be inserted into a suitable expression vector (e.g. pGEM[®], Promega Corp., USA), where it is operably linked to appropriate expression sequences, and transformed into a suitable host cell for protein expression according to conventional techniques (Sambrook J. *et al.*, Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY). Suitable host cells are those that can be grown in culture and are amenable to transformation with exogenous DNA, including bacteria, fungal cells and cells of higher eukaryotic origin, preferably mammalian cells (e.g. particularly mice or human).

To aid in purification, the zinc finger peptides (and corresponding nucleic acids) of the invention may include a purification sequence, such as a His-tag. In addition, or alternatively, the zinc finger peptides may, for example, be grown in fusion with another protein and purified as insoluble inclusion bodies from bacterial cells. This is particularly convenient when the zinc finger peptide or effector moiety may be toxic to the host cell in which it is to be expressed. Alternatively, peptides of the invention may be synthesised *in vitro* using a suitable *in vitro* (transcription and) translation system (e.g. the *E. coli* S30 extract system, Promega corp., USA). The present invention is particularly directed to the expression of zinc finger-containing peptides of the invention in host cells *in vivo* or in host cell for *ex vivo* applications, to modulate the expression of endogenous genes. Preferred peptides of the invention may therefore be devoid of such sequences (e.g. His-tags) that are intended for purification or other *in vitro* based manipulations.

The term 'operably linked', when applied to DNA sequences, for example in an expression vector or construct, indicates that the sequences are arranged so that they function cooperatively in order to achieve their intended purposes, i.e. a promoter sequence allows for initiation of transcription that proceeds through a linked coding sequence as far as the termination sequence.

It will be appreciated that, depending on the application, the zinc finger peptide or fusion protein of the invention may comprise an additional peptide sequence or sequences at the N- and/or C-terminus for ease of protein expression, cloning, and/or peptide or RNA stability, without changing the sequence of any zinc finger domain. For example, suitable N-terminal leader peptide sequences for incorporation into peptides of the invention are MA or MG and ERP. Nuclear localisation sequences (one or more) may be suitably incorporated at the N-terminus of the peptides of the invention to create an N-terminal leader sequence. Preferred host-compatible N-terminal additional sequences are Met-Gly dipeptide for protein expression in humans and mice; human KIAA2022 NLS (PKKRRKVT, SEQ ID NO: 108) or mouse primase p58 NLS9 (RIRKKLR, SEQ ID NO: 109) nuclear localisation sequences for expression in human or mouse respectively; and host-derived effector domain sequences as discussed above. Thus, a particularly useful N-terminal leader sequence for expression and nuclear targeting in human cells is MGPKKRRKVTGERP (SEQ ID NO: 111) or MGPKKRRKVTLAERP (SEQ ID NO: 112), and a useful N-terminal leader sequence for expression and nuclear targeting in mouse cells is MGRIRKKLRLAERP (SEQ ID NO: 110). Another particularly useful nuclear localisation sequence is the SV40 sequence PKKKRKV (SEQ ID NO: 107), which may be used in tandem (e.g. SEQ ID NOs: 113 (prt) or 137 (dna)) to enhance cellular uptake (as well as nuclear localisation). Indeed, it can be beneficial to use other NLS sequences in tandem too, for improved efficacy, for example, a double human KIAA2022 NLS (SEQ ID NO: 114) for use in human cells, or a double mouse primase p58 NLS9 (SEQ ID NOs: 115 (prt) and 138 (dna)) for use in mouse cells.

In some applications it may be desirable to control the expression of zinc finger (fusion) polypeptides of the invention by tissue specific promoter sequences or inducible promoters, which may provide the benefits of organ or tissue specific and/or inducible expression of polypeptides of the invention. These systems may be particularly advantageous for *in vivo* applications and gene therapy *in vivo* or *ex vivo*. Examples of tissue-specific promoters include the human CD2 promoter (for T-cells and thymocytes, Zhumabekov *et al.* (1995) *J. Immunological Methods* 185: 133-140); the alpha-calcium-calmodulin dependent kinase II promoter (for hippocampus and neocortex cells, Tsien *et al.* (1996) *Cell* 87: 1327-1338); the whey acidic protein promoter (mammary gland, Wagner *et al.* (1997) *Nucleic Acids Res.* 25: 4323-4330); the mouse myogenin promoter (skeletal muscle, Grieshammer *et al.* (1998) *Dev. Biol.* 197: 234-247); and many other tissue specific promoters that are known in the art.

It is particularly desirable to express the zinc finger peptides and other zinc finger constructs of the invention, such as zinc finger repressor or zinc finger activator proteins, from vectors suitable for use *in vivo* or *ex vivo*, e.g. for therapeutic applications (gene therapy). Where the therapy involves use of zinc finger nucleic acid constructs for expression of protein *in vivo*, the expression system selected should be capable of expressing protein in the appropriate tissue / cells where the therapy is to take effect. Desirably an expression system for use in accordance with the invention is also capable of targeting the nucleic acid constructs or peptides of the invention to the appropriate region, tissue or cells of the body in which the treatment is intended. A particularly suitable expression and targeting system is based on recombinant adeno-associated virus (AAV), e.g. the AAV2/1 subtype.

For Friedreich's ataxia (FRDA) disease gene therapy it is desirable to infect particular parts of the brain (e.g. the cerebellum), central nervous system (e.g. spinal cord) and/or muscle with therapeutic viral vectors. In some embodiments, AAV2/1 subtype vectors (see e.g. *Molecular Therapy* (2004) 10: 302–317) may be useful for this purpose. Such vectors can be used with a strong AAV promoter or a weak promoter according to preference.

Alternatively, or additionally, a broad-tropism AAV vector may be used in conjunction with a zinc finger activator protein of the invention (to provide relatively large quantities of the extended poly-zinc finger-containing proteins of the invention). Accordingly, instead or in addition to AAV2/1 subtype vectors, other AAV subtype vectors may be used, such as AAV2/9 subtype vectors. The AAV2/1 tropism is more specific for infecting neurons, whereas AAV2/9 infects more widely (*Expert Opin Biol Ther.* 2012 June; 12(6): 757–766.) and certain variants can even be applied intravenously (*Nature Biotech* 34(2): 204–209). In FRDA the heart is a particular target as patients frequently succumb to the disease before exhibiting significant CNS deterioration is detected. Therefore, using the AAV2/9 subtype (alone or in combination with AAV2/1) should advantageously allow targeting of a wider, and potentially highly important variety of cell types. For example, in the context of FRDA, this should allow targeting of other (non-neuron) cell types not only in the brain that may also play a role in disease, such as glia, but also in the heart. Thus, the use of AAV2/9, for example, may advantageously allow targeting to peripheral tissues, such as the heart, muscle or liver which may be advantageous in some embodiments and therapeutic applications.

A promoter for use in AAV2/1 viral vectors and that is suitable for use in humans and mice is the pCAG promoter (CMV early enhancer element and the chicken β -actin promoter). Another useful sequence for inclusion in AAV vectors is the Woodchuck hepatitis virus post-transcriptional regulatory element (WPRE; Garg *et al.*, (2004) *J. Immunol.*, 173: 550–558). More suitably, other promoters that may be advantageous for sustained expression in human and mice / rats *in vivo* include: (i) the pNSE promoter (neuron-specific promoter of the enolase gene), as described in Xu *et al.* (2001), *Gene Ther.*, 8:1323–32 (rat: NCBI NC_005103.4; human: NCBI NC_000012.12); (ii) the pHsp90ab1 promoter, as described in WO 2017/077329 (mouse: NCBI 15516 NC_000083.6; human: NCBI 3326 NC_000006.12); (iii) the CBh promoter (including the CMV enhancer, chicken b-actin promoter and hybrid intron), as described in Gray *et al.*, (2011), *Human Gene Therapy* (2011), 22(9):1143–1153; (iv) the human EF1 α -1 promoter, as described in Zheng and Baum (2014), *Int. J. Med. Sci.*, 11(5):404–408; and (v) the human synapsin promoter, as described in Kügler *et al.* (2003), *Gene Ther.*, 10(4):337–47).

Furthermore, endogenous promoters such as pNSE and pHSP90AB1 (PMID: 37138658) are expressed in neurons and ubiquitously, respectively. NSE is 'very strong' promoter, while HSP90AB1 is a 'strong' promoter. These promoters are typically used for the high-level expression of zinc finger proteins in accordance with the invention. In this regard, the present inventors have previously designed synthetic mouse and human pNSE promoter-enhancers (see e.g. WO 2017/077329, Example 17) comprising a portion of sequence upstream and downstream of the transcription start site of the enolase gene from human and rat: such sequences are explicitly incorporated herein as promoter-enhancer regions, which are minimal where no flanking sequences are also included. Of course, however, any other suitable

endogenous promoter sequence may alternatively be used. As the skilled person will appreciate, the selection of an appropriate endogenous promoter may suitably be construct- and/or application-dependent; e.g. according to the desired expression level of the zinc finger polypeptide concerned. Thus, the selection of endogenous promoter can be used to tune the expression level of the zinc finger polypeptide as desired. Flanking restriction sites may be added to the sequence for cloning into an appropriate vector. Since the pNSE promoter is neuron-specific, it is particularly advantageously used in combination with AAV2/1 or other neuron-specific vectors.

A promoter that may be suitable for use with AAV2/9 viral vectors is the pHSP promoter (promoter of the ubiquitously expressed Hsp90ab1 gene). This promoter may also be suitable for use in humans and mice. Again, as disclosed in the inventors earlier patent application (WO 2017/077329, Example 17), it was found that a synthetic promoter-enhancer design comprising a portion of the sequence upstream and downstream of the transcription start site of the mouse or human Hsp90ab1 gene could be advantageously used to obtain sustained expression of a transgene, such as the zinc finger peptides of the invention. In particular, a 1.7 kb region upstream of the transcription start site of the Hsp90ab1 gene that comprises multiple enhancers and can be advantageously used as a minimal hsp90ab1 constitutive promoter, in combination with a portion of exon 1 of the gene. The sequences of the mouse and human minimal promoters with flanking restriction sites for cloning into a vector are explicitly incorporated herein by reference. Mouse and human minimal promoters without flanking restriction sites are also explicitly incorporated herein by reference. These promoter-enhancer sequences may be operably associated with / linked to nucleic acid sequences encoding the zinc finger peptides and modulators of the invention; and the use / methods of using such constructs for sustained expression of (zinc finger) peptides *in vivo*. Particularly appropriate *in vivo* systems are human and mouse. The present invention therefore encompasses expression constructs and vectors (e.g. AAV2/1 or AAV2/9 viral vectors) comprising these sequences, as well as the use of such promoter sequences for expression of zinc finger activator peptides of the invention.

Suitable medical uses and methods of therapy may, in accordance with the invention, encompass the combined use – either separate, sequential or simultaneous – of the viral vectors AAV2/1 and AAV2/9. In some such embodiments, at least the AAV2/9 vector may comprise a hsp90ab1 constitutive promoter according to Example 17 of WO 2017/077329. Suitably, these medical uses and methods of therapy further comprise such vectors encoding one or more zinc finger peptide / modulator of the invention. Most suitably the medical uses and methods of therapy are directed to the treatment of FRDA in a subject, such as a human; or the study of FRDA in a subject, such as a mouse.

As the person skilled in the art would understand, strict compliance to the sequences provided is not necessary for the function of the promoter, provided that functional elements, e.g. enhancers, and their spatial relationships are essentially maintained. In particular, the promoter sequences provided comprise flanking restriction sites for cloning into a vector. The person skilled in the art would know to adapt these restriction sites to the particular cloning system used, as well as to make any point mutations that may be required in the sequence of the promoter to remove e.g. a cryptic restriction site (see e.g. Example 17 of WO 2017/077329).

Suitable inducible systems may use small molecule induction, such as the tetracycline-controlled systems (tet-on and tet-off), the radiation-inducible early growth response gene-1 (EGR1) promoter, and any other appropriate inducible system known in the art.

5

Therapeutic Compositions

A zinc finger peptide or chimeric modulator of the invention may be incorporated into a pharmaceutical composition for use in treating an animal; preferably a human. A therapeutic peptide of the invention
10 (or derivative thereof) may be used to treat one or more diseases or conditions. Alternatively, a nucleic acid encoding the therapeutic peptide may be inserted into an expression construct / vector and incorporated into pharmaceutical formulations / medicaments for the same purpose.

As will be understood by the person of skill in the art, potential therapeutic molecules, such as zinc
15 finger peptides and modulators of the invention may be tested in an animal model, such as a mouse, before they can be approved for use in human subjects. Accordingly, zinc finger peptide or chimeric modulator proteins of the invention may be expressed *in vivo* in mice or *ex vivo* in mouse cells as well as in humans. In accordance with the invention, appropriate expression cassettes and expression constructs / vectors may be designed for each animal system specifically.

20

Zinc finger peptides and chimeric modulators of the invention typically contain naturally occurring amino acid residues, but in some cases non-naturally occurring amino acid residues may also be present. Therefore, so-called 'peptide mimetics' and 'peptide analogues', which may include non-amino acid chemical structures that mimic the structure of a particular amino acid or peptide, may also be used
25 within the context of the invention. Such mimetics or analogues are characterised generally as exhibiting similar physical characteristics such as size, charge or hydrophobicity, and the appropriate spatial orientation that is found in their natural peptide counterparts. A specific example of a peptide mimetic compound is a compound in which the amide bond between one or more of the amino acids is replaced by, for example, a carbon-carbon bond or other non-amide bond, as is well known in the art (see, for
30 example Sawyer, in *Peptide Based Drug Design*, pp. 378-422, ACS, Washington D.C. 1995). Such modifications may be particularly advantageous for increasing the stability of zinc finger peptide therapeutics and/or for improving or modifying solubility, bioavailability and delivery characteristics (e.g. for *in vivo* applications) when a peptide is to be administered as the therapeutic molecule.

The therapeutic peptides and nucleic acids of the invention may be particularly suitable for the treatment of diseases, conditions and/or infections that can be targeted (and treated) intracellularly, for example, by targeting genetic sequences within an animal cell; and also for *in vitro* and *ex vivo* applications. As used herein, the terms 'therapeutic agent' and 'active agent' encompass both peptides and the nucleic acids that encode a therapeutic zinc finger peptide of the invention. Therapeutic nucleic acids include
40 vectors, viral genomes and modified viruses, such as AAV, which comprise nucleic acid sequences encoding zinc finger peptides and fusion proteins of the invention.

Therapeutic uses and applications for the zinc finger peptides and nucleic acids include any disease, disorder or other medical condition that may be treatable by modulating the expression of a target gene or nucleic acid.

5 A particular target of the present therapies based on poly-zinc finger therapeutic molecules is Friedreich's ataxia (FRDA), which are associated with expanded GAA polynucleotide repeat sequences. As such, zinc finger peptides of the invention are particularly adapted to target and bind to GAA-repeat sequences, or frameshifted GAA-repeat sequences, such as -AGA- and AAG- repeat
10 sequences within human or animal genomes. A preferred target gene is therefore frataxin, which is known to be susceptible to expansion of the GAA-repeat sequences. In this example, a pathogenic gene is typically associated with more than 100 GAA-repeat sequences, and generally between 600 and 1,700 such repeats. On the other hand, normal, non-pathogenic genes comprise less than 50, and typically in the range of 7 to 40 GAA repeat sequences.

15 One or more additional pharmaceutically acceptable carrier (such as diluents, adjuvants, excipients or vehicles) may be combined with the therapeutic peptide(s) of the invention in a pharmaceutical composition. Suitable pharmaceutical carriers are described in "*Remington's Pharmaceutical Sciences*" by E. W. Martin. Pharmaceutical formulations and compositions of the invention are formulated to conform to regulatory standards and can be administered orally, intravenously, topically, or via other
20 standard routes.

In accordance with the invention, the therapeutic peptides or nucleic acids may be manufactured into medicaments or may be formulated into pharmaceutical compositions. When administered to a subject, a therapeutic agent is suitably administered as a component of a composition that comprises a
25 pharmaceutically acceptable vehicle. The molecules, compounds and compositions of the invention may be administered by any convenient route, for example, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, oral, sublingual, intranasal, intravaginal, transdermal, rectally, by inhalation, or topically to the skin. Administration can be systemic or local. Delivery systems that are known also include, for example, encapsulation in microgels, liposomes, microparticles,
30 microcapsules, capsules, etc., and any of these may be used in some embodiments to administer the compounds of the invention. Any other suitable delivery systems known in the art are also envisaged in use of the present invention.

Acceptable pharmaceutical vehicles can be liquids, such as water and oils, including those of petroleum,
35 animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. The pharmaceutical vehicles can be saline, gum acacia, gelatin, starch paste, talc, keratin, colloidal silica, urea, and the like. In addition, auxiliary, stabilising, thickening, lubricating and colouring agents may be used. When administered to a subject, the pharmaceutically acceptable vehicles are preferably sterile. Water is a suitable vehicle particularly when the compound of the invention is administered
40 intravenously. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid vehicles, particularly for injectable solutions. Suitable pharmaceutical vehicles also include excipients such as starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium

stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The present compositions, if desired, can also contain minor amounts of wetting or emulsifying agents, or buffering agents.

- 5 The medicaments and pharmaceutical compositions of the invention can take the form of liquids, solutions, suspensions, lotions, gels, tablets, pills, pellets, powders, modified-release formulations (such as slow or sustained-release), suppositories, emulsions, aerosols, sprays, capsules (for example, capsules containing liquids or powders), liposomes, microparticles or any other suitable formulations known in the art. Other examples of suitable pharmaceutical vehicles are described in *Remington's*
10 *Pharmaceutical Sciences*, Alfonso R. Gennaro ed., Mack Publishing Co. Easton, Pa., 19th ed., 1995, see for example pages 1447-1676.

- In some embodiments the therapeutic compositions or medicaments of the invention are formulated in accordance with routine procedures as a pharmaceutical composition adapted for oral administration
15 (more suitably for human beings). Compositions for oral delivery may be in the form of tablets, lozenges, aqueous or oily suspensions, granules, powders, emulsions, capsules, syrups, or elixirs, for example. Thus, in one embodiment, the pharmaceutically acceptable vehicle is a capsule, tablet or pill.

- Orally administered compositions may contain one or more agents, for example, sweetening agents
20 such as fructose, aspartame or saccharin; flavouring agents such as peppermint, oil of wintergreen, or cherry; colouring agents; and preserving agents, to provide a pharmaceutically palatable preparation. When the composition is in the form of a tablet or pill, the compositions may be coated to delay disintegration and absorption in the gastrointestinal tract, so as to provide a sustained release of active agent over an extended period of time. Selectively permeable membranes surrounding an osmotically
25 active driving compound are also suitable for orally administered compositions. In these dosage forms, fluid from the environment surrounding the capsule is imbibed by the driving compound, which swells to displace the agent or agent composition through an aperture. These dosage forms can provide an essentially zero order delivery profile as opposed to the spiked profiles of immediate release formulations. A time delay material such as glycerol monostearate or glycerol stearate may also be
30 used. Oral compositions can include standard vehicles such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Such vehicles are preferably of pharmaceutical grade. For oral formulations, the location of release may be the stomach, the small intestine (the duodenum, the jejunum, or the ileum), or the large intestine. One skilled in the art is able to prepare formulations that will not dissolve in the stomach, yet will release the material in the
35 duodenum or elsewhere in the intestine. Suitably, the release will avoid the deleterious effects of the stomach environment, either by protection of the peptide (or derivative) or by release of the peptide (or derivative) beyond the stomach environment, such as in the intestine. To ensure full gastric resistance a coating impermeable to at least pH 5.0 would be essential. Examples of the more common inert ingredients that are used as enteric coatings are cellulose acetate trimellitate (CAT),
40 hydroxypropylmethylcellulose phthalate (HPMCP), HPMCP 50, HPMCP 55, polyvinyl acetate phthalate (PVAP), Eudragit L30D, Aquateric, cellulose acetate phthalate (CAP), Eudragit L, Eudragit S, and Shellac, which may be used as mixed films.

To aid dissolution of the therapeutic agent or nucleic acid (or derivative) into the aqueous environment a surfactant might be added as a wetting agent. Surfactants may include anionic detergents such as sodium lauryl sulfate, dioctyl sodium sulfosuccinate and dioctyl sodium sulfonate. Cationic detergents might be used and could include benzalkonium chloride or benzethonium chloride. Potential nonionic detergents that could be included in the formulation as surfactants include: laurumacrogol 400, polyoxyl 40 stearate, polyoxyethylene hydrogenated castor oil 10, 50 and 60, glycerol monostearate, polysorbate 20, 40, 60, 65 and 80, sucrose fatty acid ester, methyl cellulose and carboxymethyl cellulose. These surfactants, when used, could be present in the formulation of the peptide or nucleic acid or derivative either alone or as a mixture in different ratios.

Typically, compositions for intravenous administration comprise sterile isotonic aqueous buffer. Where necessary, the compositions may also include a solubilising agent.

Another suitable route of administration for the therapeutic compositions of the invention is via pulmonary or nasal delivery.

Additives may be included to enhance cellular uptake of the therapeutic peptide (or derivative) or nucleic acid of the invention, such as the fatty acids, oleic acid, linoleic acid and linolenic acid.

In one exemplary pharmaceutical composition of the invention, one or more zinc finger peptide or nucleic acid of the invention (and optionally any associated non-zinc finger moiety, e.g. a modulator of gene expression and/or targeting moiety) may be mixed with a population of liposomes (i.e. a lipid vesicle or other artificial membrane-encapsulated compartment), to create a therapeutic population of liposomes that contain the therapeutic agent and optionally the modulator or effector moiety. The therapeutic population of liposomes can then be administered to a patient by any suitable means, such as by intravenous injection. Where it is necessary for the therapeutic liposome composition to target specifically a particular cell-type, such as a particular microbial species or an infected or abnormal cell, the liposome composition may additionally be formulated with an appropriate antibody domain or the like (e.g. Fab, F(ab)₂, scFv *etc.*) or alternative targeting moiety, which naturally or has been adapted to recognise the target cell-type. Such methods are known to the person of skill in the art.

The therapeutic peptides or nucleic acids of the invention may also be formulated into compositions for topical application to the skin of a subject.

In embodiments of the invention the therapeutic compositions may include only one therapeutic peptide / protein or nucleic acid of the invention; or may include two or more e.g. two complementary therapeutic peptides / proteins or nucleic acids of the invention. For example, a poly-zinc finger repressor protein of the invention may be used alone, or in combination with another zinc-finger peptide or therapeutic agent, e.g. to up-regulate expression of a pathogenic gene target. In other embodiments, two therapeutic zinc finger peptides of the invention may be used in concert; e.g. a zinc finger activator

protein for targeting the GAA-repeat sequence within intron 1 of frataxin in a pathogenic gene, and a second zinc finger activator protein for targeting the promoter region of frataxin.

When two (or more) therapeutic zinc finger peptides are contemplated, the different zinc finger peptides or encoding nucleic acid constructs or viral vectors may be incorporated into the same pharmaceutical composition, or may be manufactured separately. Where two (or more) pharmaceutical compositions are manufactured for administration to the same individual, it will be appreciated that the compositions may be administered simultaneously, sequentially, or separately, as directed / required.

Zinc finger peptides and nucleic acids of the invention may also be useful in non-pharmaceutical applications, such as in diagnostic tests, imaging, as affinity reagents for purification and as delivery vehicles.

Gene Therapy

One aspect of the invention relates to gene therapy treatments utilising zinc finger peptides of the invention for treating diseases.

Gene therapy relates to the use of heterologous genes in a subject, such as the insertion of genes into an individual's cell (e.g. animal or human) and biological tissues to treat disease, for example: by replacing deleterious mutant alleles with functional / corrected versions, by inactivated mutant alleles by removing all or part of the mutant allele, or by inserting an expression cassette for sustained expression of a therapeutic zinc finger construct according to the invention. The most promising target diseases to date are those that are caused by single-gene defects, such as cystic fibrosis, haemophilia, muscular dystrophy, sickle cell anaemia, Huntington's disease (HD), ALS, FTD, FXTAS and FXS. Other common gene therapy targets are aimed at cancer and hereditary diseases linked to a genetic defect, such as expanded nucleotide repeats. The present invention is concerned with the treatment of genes associated with expanded polynucleotide repeats, and in particular, with expanded repeats of the trinucleotide sequence GAA or variants thereof (such as AGA and AAG).

Gene therapy is classified into two types: germ line gene therapy, in which germ cells, (i.e. sperm or eggs), are modified by the introduction of therapeutic genes, which are typically integrated into the genome and have the capacity to be heritable (i.e. passed on to later generations); and somatic gene therapy, in which the therapeutic genes are transferred into somatic cells of a patient, meaning that they may be localised and are not inherited by future generations.

Gene therapy treatments require delivery of the therapeutic gene (or DNA or RNA molecule) into target cells. There are two categories of delivery systems, either viral-based delivery mechanisms or non-viral mechanisms, and both mechanisms are envisaged for use with the present invention.

Viral systems may be based on any suitable virus, such as: retroviruses, which carry RNA (e.g. influenza, SIV, HIV, lentivirus, and Moloney murine leukaemia); adenoviruses, which carry dsDNA;

adeno-associated viruses (AAV), which carry ssDNA; herpes simplex virus (HSV), which carries dsDNA; and chimeric viruses (e.g. where the envelop of the virus has been modified using envelop proteins from another virus).

- 5 A particularly preferred viral delivery system is AAV. AAV is a small virus of the parvovirus family with a genome of single stranded DNA. A key characteristic of wild-type AAV is that it almost invariably inserts its genetic material at a specific site on human chromosome 19. However, recombinant AAV, which contains a therapeutic gene in place of its normal viral genes, may not integrate into the animal genome, and instead may form circular episomal DNA, which is likely to be the primary cause of long-
- 10 term gene expression. Advantages of AAV-based gene therapy vectors include: that the virus is non-pathogenic to humans (and is already carried by most people); most people treated with AAV will not build an immune response to remove either the virus or the cells that have been successfully infected with it (in the absence or heterologous gene expression); it will infect dividing as well as non-dividing (quiescent) cells; and it shows particular promise for gene therapy treatments of muscle, eye, and brain.
- 15 AAV vectors have been used for first- and second-phase clinical trials for the treatment of cystic fibrosis; and first-phase clinical trials have been carried out for the treatment of haemophilia. There have also been encouraging results from phase I clinical trials for Parkinson's disease, which provides hope for treatments requiring delivery to the central nervous system. Gene therapy trials using AAV have also been reported for treatment of Canavan disease, muscular dystrophy and late infantile neuronal ceroid
- 20 lipofuscinosis. HSV, which naturally infects nerve cells in humans, may also offer advantages for gene therapy of diseases involving the nervous system.

Suitably, in accordance with the invention, zinc finger encoding nucleic acid constructs (as described herein) are inserted into an adeno-associated virus (AAV) vector, particularly the AAV2/9 subtype (see

25 e.g. *Molecular Therapy* (2004) 10: 302–317). As described, since this vector has a relatively wide tropism, it can be useful for targeting of certain key tissues, such as the heart, which are affected by FRDA. In this way, the zinc finger encoding nucleic acid constructs of the invention can be delivered to desired target cells, and the zinc finger peptides expressed in order to increase the expression of pathogenic genes associated with GAA repeat sequences, such as mutant frataxin (FXN) genes for

30 treatment of Friedreich's ataxia (FRDA). Zinc finger peptides of the invention may also increase the expression of mutant frataxin (FXN) genes by targeting different nucleic acid sequence regions associated with the pathogenic gene (or non-pathogenic frataxin genes in heterozygous patients), e.g. promoter regions which are common to both pathogenic and non-pathogenic genes.

35 Ubiquitous / promiscuous viral vectors, such as AAV2/9, may also be used alone, for example, where the therapy is targeted at peripheral tissues. In addition, AAV2/9 can beneficially be used systemically and intravenously, and/or delivered to different organs of a subject, e.g. by intramuscular injection. Again, however, intrathecal administration of AAV2/9 therapeutics may be preferred.

40 In other embodiments, zinc finger encoding nucleic acid constructs (as described herein) are inserted into the AAV2/1 subtype. This vector is particularly suitable for injection into and infection of the striatum, in the brain, where the therapeutics of the invention may be particularly useful. Alternatively, the vector

can be injected intrathecally or directly into the cisterna magna or brain. Intrathecally is a preferred mode route for administration of AAV2/1 therapeutics of the present invention.

5 In embodiments, in addition to vectors for targeting the heart or other peripheral tissues / organs, vectors with a more specific tropism, for example, the neuron specific AAV2/1 subtype may be used in combination with a broad tropism vector, such as the AAV2/9 subtype. This may advantageously allow targeting of both neurons and other types of cells present in the brain, such as glial cells.

10 Although FRDA is primarily considered to be a neurological disease, the effects of the diseases are far-reaching throughout the body. Therefore, targeting of tissues other than the central nervous system with the zinc finger peptides / modulators of the invention may prove beneficial. In such applications use of a promiscuous vector (such as AAV2/9) or an organ / tissue specific vector may be particularly useful.

15 In embodiments, the tropism of the viral vector and the specificity of the promoter used for expression of the therapeutic construct can be tailored for targeting of specific populations of cells. For example, neuron-specific viral vectors may be used in combination with neuron-specific promoters. Conversely, promiscuous vectors may be used in combinations with ubiquitous promoters (or tissue specific promoters as desired).

20 In specific embodiments, AAV2/9 viruses may be used in combination with a synthetic pHSP vector, also as described above (see also WO 2017/077329). In other embodiments, AAV2/1 viruses may be used in combination with a synthetic pNSE promoter, as described above (see also WO 2017/077329). In embodiments, combinations of these two types of constructs may be used in order to simultaneously
25 target multiple cell types, e.g. for the treatment of FRDA.

For some applications non-viral based approaches for gene therapy can provide advantages over viral methods, for example, in view of the simple large-scale production and low host immunogenicity. Types of non-viral mechanism include: naked DNA (e.g. plasmids); oligonucleotides (e.g. antisense, siRNA,
30 decoy ds oligodeoxynucleotides, and ssDNA oligonucleotides); lipoplexes (complexes of nucleic acids and liposomes); polyplexes (complexes of nucleic acids and polymers); and dendrimers (highly branched, roughly spherical macromolecules).

35 Accordingly, the zinc finger-encoding nucleic acids of the invention may be used in methods of treating diseases by gene therapy. As already explained, particularly suitable diseases are those of the nervous system (especially motor neurons); and preferably those associated with GAA repeat sequences, such as FRDA.

40 Accordingly, the gene therapy therapeutics and regimes of the invention may provide for the expression of therapeutic zinc fingers in target cells *in vivo* or in *ex vivo* applications for repressing the expression of target genes, such as those having non-wild-type expanded GAA-repeat sequences, and especially the mutant frataxin gene.

Friedreich's Ataxia (FRDA)

As in many other neurological disorders, such as Alzheimer's and Parkinson's diseases, Friedreich's ataxia (FRDA) has a complex disease pathology.

Friedreich's ataxia (FRDA) is an autosomal recessive disorder which typically develops in children and teens and gradually worsens over time. It is a relatively rare inherited genetic disorder with a prevalence of approx. 1 in 40,000 to 50,000 and there is currently no cure (Zhang *et al.*, (2019), Trends Pharmacol. Sci., 40(4):229-233; PMID: 30905359). The condition damages the spinal cord, peripheral nerves, and the cerebellum portion of the brain, and the disease manifests with progressive ataxia, impaired speech, hearing and vision, cardiomyopathy, diabetes, and skeletal muscle abnormalities (Cook & Giunti (2017), Br. Med. Bull., 124(1):19-30; PMID: 29053830). Patients tend to be unsteady, and have awkward movements and a loss of sensation due to nerve injury develop as the disease progresses.

FRDA is typically caused by GAA expansions of up to 1,700 units (although longer expansions are known), within intron 1 of the frataxin locus. This in turn leads to down-regulation of the frataxin mRNA transcripts by up to 90% (depending on severity), and resulting in toxicity and disease phenotypes (PMID: 30905359). Homozygous patients are fully symptomatic while heterozygous carriers are typically non-symptomatic and comprise between 1 and 2% of the general population (Campuzano *et al.*, (1996), Science, 271(5254):1423-1427; PMID: 8596916).

Although its role is not fully understood, frataxin is important for the normal function of mitochondria, the energy-producing centres within cells. In FRDA, the normal GAA-trinucleotide repeat which may be present in intron 1 of the frataxin gene in less than about 40 copies, is multiplied to 66 or more (typically hundreds of) copies, which greatly disrupts the normal production of frataxin. Research suggests that without a normal level of frataxin, certain cells in the body (especially peripheral nerve, spinal cord, brain, and heart muscle cells) produce energy less effectively and may build up toxic byproducts, leading to oxidative stress. Lack of normal levels of frataxin also may lead to increased levels of iron in the mitochondria. When the excess iron reacts with oxygen, free radicals can be produced. Although free radicals are essential molecules in the body metabolism, they can also destroy cells and harm the body.

As with many degenerative diseases of the nervous system, there is currently no cure or effective treatment for FRDA. However, many of the symptoms and accompanying complications can be treated to help individuals maintain optimal functioning as long as possible. A multi-specialty team approach is generally essential to the treatment of someone with FRDA. For example, treatments may be required for diabetes, if present, and various heart problems. Orthopedic problems such as foot deformities and scoliosis can be corrected with braces or surgery. Furthermore, physical therapy may prolong the use of the arms and legs; but swallowing and speech issues can also develop, as can hearing impairment, which can be helped with hearing aids.

Ultimately, FRDA typically results in death of an affected individual within about 20 years from diagnosis, and so there is a desperate need for more effective therapies and treatments for FRDA.

Interestingly, a triplet repeat expansion has been implicated as the cause of several diseases in which a person needs to inherit only one abnormal gene. However, FRDA is the only known genetic disorder that requires inheriting two copies of the abnormal, pathogenic gene to cause the disease. Therefore, restoring expression of frataxin up to only about 50% of normal wild-type values may be expected to be of therapeutic benefit, and may be achievable with the therapeutic zinc finger activator peptides described herein.

Host Organism Toxicity and Immunogenicity

It was proposed that toxicity and immunogenicity (immunotoxicity) of heterologous peptides when expressed in host organisms might be reduced by optimising the primary peptide sequence to match the primary peptide sequence of natural host peptides.

As previously described (Garriga *et al.*, 2012 and in WO 2017/077329), zinc finger peptides based on a generic / universal zinc finger peptide framework, and particularly on the peptide framework of Zif268, which is a natural zinc finger protein having homologues in both mice and humans can be beneficial for reducing host immune reactions. However, in general, the recognition sequences of a zinc finger domain should be based on the perceived best match for the target nucleic acid sequences (i.e. the recognition code for zinc finger-dsDNA interactions) and on binding optimisation studies. Such designs according to the prior art have no regard to the target host organism in which the zinc finger peptides would be ultimately expressed (e.g. mouse or human). Similarly, effector domains, such as transcriptional activator and repressor domains and other effector functions, such as nuclear localisation and purification tags have been previously selected without regard to the host organism. This has been shown to be a potential reason for failure to express exogenous, therapeutic peptides over the long term in a host organism. The inventors' previous work (WO 2017/077329) addressed this problem in the art, and the present invention follows those important teachings.

Thus, zinc finger peptides and modulator peptides of the invention have greater than 50%, greater than 60%, greater than 70% or even greater than 75% identity to endogenous / natural protein sequences in the target, host organism in which they are intended to be expressed for therapeutic use. More suitably, the peptides of the invention have at least 80%, 81%, 82%, 83%, 84% or at least 85% identity to endogenous / natural proteins in the target organism. In some cases, it is desirable to have still greater identity to peptide sequences of the target / host organism, such as between approximately 75% and 98% identity, between 78% and 95% identity, between 80% and 90% identity. At the same time, it will be appreciated that the peptides of the invention are different to known peptide sequences. Thus, the peptides may be up to 50%, up to 40%, up to 30% or up to 25% non-identical to endogenous / natural peptide sequences found in the host organism and/or previously known. It will be appreciated that by 'up to x%', in this context, means greater than 0% and less than x%. Preferably, the peptides of the invention are up to 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11% or 10% non-identical

to endogenous / natural peptide sequences found in the host organism; for example, the peptides of the invention may be between approximately 1% and 25%, between approximately 3% and 20% or between approximately 5% and 15% non-identical to an endogenous peptide sequence of the host organism.

5

Sequence identity can be assessed in any way known to the person of skill in the art, such as using the algorithm described by Lipman & Pearson (1985), *Science* 227, pp1435; or by sequence alignment.

10

As used herein, 'percent identity' means that, when aligned, that percentage of amino acid residues (or bases in the context of nucleic acid sequences) are the same when comparing the two sequences. Amino acid sequences are not identical, where an amino acid is substituted, deleted, or added compared to the reference sequence. In the context of the present invention, since the subject proteins may be considered to be modular, i.e. comprising several different domains or effector and auxiliary sequences (such as NLS sequences, expression peptides, zinc finger modules / domains, and effector domains (e.g. activator peptides)), sequence identity may conveniently be assessed separately for each domain / module of the peptide relative to any homologous endogenous or natural peptide domain / module known in the host organism. This is considered to be an acceptable approach since relatively short peptide fragments (epitopes) of any host-expressed peptides may be responsible for determining immunogenicity through recognition or otherwise of self / non-self peptides when expressed in a host organism *in vivo*. By way of example, a peptide sequence of 100 amino acids comprising a host zinc finger domain directly fused to a host activator domain wherein neither sequence has been modified by mutation would be considered to be 100% identical to host peptide sequences. It does not matter for this assessment whether such zinc finger domain(s) or non-zinc finger domain(s), e.g. an activator domain, is only a fragment from a natural, larger protein expressed in the host. If one of 100 amino acids has been modified from the natural sequence, however, the modified sequence would be considered 99% identical to natural protein sequences of the host; whilst if the same zinc finger domain were linked to the same activator domain by a linker sequence of 10 amino acids and that linker sequence is not naturally found in that context in the host organism, then the resultant sequence would be $(10/110) \times 100\%$ non-identical to host sequences.

30

35

40

Thus, the degree of sequence identity between a query sequence and a reference sequence may, in some embodiments be determined by: (1) aligning the two sequences by any suitable alignment program using the default scoring matrix and default gap penalty; (2) identifying the number of exact matches, where an exact match is where the alignment program has identified an identical amino acid or nucleotide in the two aligned sequences on a given position in the alignment; and (3) dividing the number of exact matches with the length of the reference sequence. In other embodiments, step (3) may involve dividing the number of exact matches with the length of the longest of the two sequences; and in other embodiments, step (3) may involve dividing the number of exact matches with the 'alignment length', where the alignment length is the length of the entire alignment including gaps and overhanging parts of the sequences. As explained above, in this context, the alignment length is the accumulative amino acid length of all peptide domains, modules or fragments that have been used as reference sequences for each respective domain or module of the query peptide.

Sequence identity comparisons can be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. Commercially available computer programs may use complex comparison algorithms to align two or more sequences that best reflect the evolutionary events that might have led to the difference(s) between the two or more sequences. Therefore, these algorithms operate with a scoring system rewarding alignment of identical or similar amino acids and penalising the insertion of gaps, gap extensions and alignment of non-similar amino acids. The scoring system of the comparison algorithms may include one or more and typically all of: (i) assignment of a penalty score each time a gap is inserted (gap penalty score); (ii) assignment of a penalty score each time an existing gap is extended with an extra position (extension penalty score); (iii) assignment of high scores upon alignment of identical amino acids; and (iv) assignment of variable scores upon alignment of non-identical amino acids. Most alignment programs allow the gap penalties to be modified. However, it is preferred to use the default values when using such software for sequence comparisons.

Suitable computer programs for carrying out such an alignment include, but are not limited to, Vector NTI (Invitrogen Corp.) and the ClustalV, ClustalW and ClustalW2 programs (Higgins DG & Sharp PM (1988), Higgins *et al.* (1992), Thompson *et al.* (1994), Larkin *et al.* (2007). A selection of different alignment tools is available from the ExPASy Proteomics server at www.expasy.org. Another example of software that can perform sequence alignment is BLAST (Basic Local Alignment Search Tool), which is available from the webpage of National Center for Biotechnology Information which can currently be found at <http://www.ncbi.nlm.nih.gov/> and which was firstly described in Altschul *et al.* (1990), *J. Mol. Biol.* 215; pp 403-410. Examples of programs that perform global alignments are those based on the Needleman-Wunsch algorithm, e.g. the EMBOSS Needle and EMBOSS Stretcher programs. In one embodiment, it is preferred to use the ClustalW software for performing sequence alignments. ClustalW2 is for example made available on the internet by the European Bioinformatics Institute at the EMBL-EBI webpage www.ebi.ac.uk under tools - sequence analysis - ClustalW2.

Importantly, since Zif268 has homologues in mouse and human cells, and the zinc finger scaffold framework of Zif268 is almost identical in mice and humans (WO2012/049332; WO2017/077329 and WO2022/003361), the inventors have previously shown that a single appropriately modified host-optimised zinc finger peptide sequence of the invention may be suitable for use in both mouse and human cells without resulting in adverse immunogenic effects. In this regard, it has previously been shown that improved host-optimisation can be achieved by modifying the originally designed recognition helices and zinc finger linkers in order to match them as closely as possible to the human (or mouse respectively) Zif268 transcription factor sequences. Thus, for example, the first zinc finger recognition sequence in a zinc finger array may have the amino acid sequence LT in the +4 and +5 positions, respectively, of the alpha-helix, rather than the amino acid sequence RK, which is found in the third recognition sequence of Zif268.

Differences between the mouse and human variants of the zinc finger peptides of this invention may lie in the selected of activation or other effector domains, which may be selected from mouse or human

variants, as required. For example, the nuclear localisation signal (NLS), may suitably be derived from a human variant peptide for use in humans (e.g. human protein KIAA2022 NLS), or a mouse peptide for use in mouse, as described elsewhere herein. Similarly, the activation domain of zinc finger activator peptides of the invention may be the p65 RelA activation domain derived from the human variant for use in humans or from the mouse variant for use in mice (*EMBO J.* (1991) 10(12):3805-17), or VP16 / VP64 activation domains may be used as appropriate.

It has thus been found that several design variants of zinc finger peptide sequences can be synthesised to retain desired poly-GAA binding characteristics, while improving / maximising host matching properties and minimising toxicity *in vivo*. Surprisingly, such design variants can include a relatively high number of modifications within zinc finger alpha-helical recognition sequences and within zinc finger linker sequences, both of which might be expected to affect (e.g. reduce) target nucleic acid binding affinity and specificity, without adversely affecting the efficacy of the potential therapeutic for use *in vivo*. Moreover, by beneficially reducing immunogenicity and toxicity effects *in vivo*, mid to long-term activity of the therapeutic peptides of the invention are significantly increased.

Active Delivery of Therapeutic Zinc Finger Peptides

Efficient long-term delivery of gene regulatory factors to somatic cells has great potential in medicine: especially for cases where one wishes to reprogram genetic networks or to control gene expression at will.

In recent years, there have been reported in the art many examples of designer gene-specific transcription factors being used to up- or down-regulate target disease genes. However, in most cases long-term treatment (from a single therapeutic administration) is impossible. Against this background, the inventors have developed a universal method for enhanced control of gene expression *in vitro* and, advantageously, *in vivo* with artificial gene-regulatory transcription factors, such as zinc finger peptides. This new method provides a means for significantly increasing the ability to artificially control somatic gene expression, based on the concept of 'active delivery' of therapeutic peptides, such as transcription factors (e.g. zinc finger peptides), to cells. The process of active delivery involves the general steps of: expression of a therapeutic peptide in a first cell; secretion of the therapeutic peptide from the first cell; diffusion of the therapeutic peptide from the first cell to a neighbouring (second) cell; cell-penetration of the neighbouring cell by the secreted therapeutic peptide; and therapeutic peptide targeting, such that the therapeutic peptide delivers its therapeutic effect to a desired location within the neighbouring cell. The therapeutic peptide is desirably a designer transcription factor, such as one or more of the zinc finger peptides described herein.

Thus, the present disclosure also relates to methods and peptide / nucleic acid constructs for prolonged and/or enhanced therapy. In this regard, the inventors have previously reported (WO 2022/003363) that 'active delivery' of therapeutic zinc finger peptides to diseased cells can be achieved *in vitro* and *in vivo*, and that such active delivery can improve the efficacy of a therapeutic treatment. In particular, active delivery of therapeutic peptides to pathogenic cells which have not been directly contacted with or

transduced by a gene therapy vector (such as an AAV vector) can enhance a single therapeutic treatment, by delivering therapeutic peptides to diseased cells that would otherwise be unaffected by the treatment. In addition, active delivery of therapeutic peptides can continue to deliver therapeutic peptides to diseased cells which previously had been treated with a gene therapy or therapeutic peptide, in circumstances where the gene therapy has been silenced or has otherwise become ineffective.

Indeed, the inventors have previously shown that ZFP therapies are currently limited by long-term expression efficiency: for example, for treatment of Huntington's disease, despite that long term expression of therapeutic ZFP transcription factors was achieved by, *inter alia*, host-matching of therapeutic peptide sequences; target gene repression was limited to approximately 25% in the whole brain after 6 months (Agustín-Pavón *et al.* (2016) *Mol. Neurodegener.*, 11(1):64). Therefore, while expression of a therapeutic peptide in a proportion of target cells may be effective for a short time period, the therapeutic benefit to the host organism may be rapidly diminished due to the initial failure to deliver the therapeutic transgene into every desirable target cell, followed by the loss of expression of therapeutic transgenes in cells that were initially successfully targeted. Having regard to the prior art, a transgene expression profile after 6 months of 25% of target cells is currently a positive result, but this significantly reduces the effectiveness of any therapy such that further treatments will be necessary to maintain a therapeutic effect in the mid- to long-term.

The inventors have also shown that active delivery constructs can improve long-term therapeutic effects by continuing to provide (e.g. to 'drip-feed') secreted cell-penetrating therapeutic zinc finger transcription factors to bystander / neighbouring cells in the brain and other tissues, which would not otherwise be exposed to the therapeutic molecules (see Figures 7A and 7B).

As exemplified in Figure 7A, therapeutic delivery agents, e.g. viral vectors (or other delivery systems, such as naked nucleic acids) may conveniently be used to deliver nucleic acid expression constructs to target cells within a host organ(ism). Direct injection of the therapeutic delivery agent is one convenient means for delivering the agent to a desired region of a subject organism. However, whilst such therapeutic delivery agents may infect / enter a plurality of target cells, complete delivery of agent to every target cell is impossible and, even if the delivery were complete or almost complete, it is known that the effectiveness of a gene therapy treatment (e.g. by expression of an exogenous therapeutic peptide agent), is typically limited by gene silencing or vector / transgene loss within the short- or medium-term (e.g. between a few days and a few months). As shown in Figure 7A, a (first) population of target cells at sites of administration / injection A and B receive a therapeutic transgene (in this example from a viral vector delivery agent), and successfully express the therapeutic peptide. Expressed therapeutic peptides are adapted to be secretable from targeted cells by way of an expressed protein secretion signal (SS) or signal peptide (SP), which causes at least a proportion of the expressed therapeutic peptide to be secreted from the targeted cells that express the peptide. Secreted therapeutic peptides may then diffuse away from the cell in which they were expressed into a 'diffusion volume' (e.g. a surrounding region within the host organism), and may come into contact with a multitude more cells of similar type (i.e. a second population of target cells) within the diffusion volume. For example, as depicted in Figure 7B, infected neuronal cells may express and secrete therapeutic

peptides, which diffuse away from the cell in which they were expressed and come into contact with non-treated cells, such as astrocytes and other neuronal cells. Furthermore, the secreted therapeutic peptides are advantageously adapted for cell penetration, for example, by way of one or more expressed nuclear localisation signal (NLS), which provides a net positive charge, enhancing the ability of the peptide to penetrate cells. Once inside a 'neighbouring' cell, the therapeutic peptide may be targeted to the nucleus (for example), in order to provide a beneficial therapeutic effect in the new cell.

In this way, less than total delivery and expression of a trans / exogenous gene in target cells can be supplemented by exposure of neighbouring cells to the resultant, expressed therapeutic peptide. Such a mechanism can greatly increase the effectiveness of a therapeutic treatment by increasing both the proportion of target cells that receive therapeutic agents and the length of time over which target cells are exposed to therapeutic peptides / agents.

This approach is particularly beneficial in conjunction with the zinc finger peptides described herein, because the process of cell penetration positively exploits the intrinsic cell penetrating properties of zinc finger peptides (Gaj *et al.*, (2012) *Nat. Methods*, 9, 805-7; Gaj *et al.*, (2014) *ACS Chem. Biol.*, 9, 1662-7; Liu *et al.*, (2015) *Mol. Ther. Nucleic Acids*, 4, e232; Mino *et al.*, (2013) *PLoS One*, 8, e56633). These cell-penetration properties have not been coupled before to secretion *in vivo*, nor to gene therapy processes based on delivery of an agent with AAVs.

Active delivery can be achieved within a population of cells *in vitro* or, more advantageously, *in vivo*: for example, in mouse or humans, using AAV-based vectors to deliver expression constructs encoding therapeutic peptides capable of secretion from and penetration into target cells. It will be appreciated, however, than any other suitable delivery agent / virus could be used, as could any other appropriately modified therapeutic peptide / agent.

It is generally desired that a delivery vector for use in 'active delivery' should be capable of cell / tissue-type specific expression and/or long-term expression and/or strong expression of therapeutic peptides. Thus, delivery vectors according to this disclosure may beneficially comprise a promoter / enhancer sequence such as pCMV, pNSE, pHsp90, CBh, EF1 α -1, synapsin or pCAG, which may also be depending on the target organism (e.g. human, mouse, rat etc.). Preferred promoter / enhancer sequences are pNSE, pHsp90, CBh, EF1 α -1 and synapsin; especially pNSE and pHsp90, as described herein.

As explained above, a therapeutic peptide for 'active delivery' (at least *in vivo*) must be capable of secretion from the cell in which it is expressed. Multiple cell secretion methods are known to the person skilled in the art and may potentially be employed in accordance with the invention. In particular, cell secretion peptide signal sequences are known and are convenient for use in conjunction with an expressed peptide therapeutic. Thus, the therapeutic peptide may suitably comprise at least one protein secretion signal (SS) or signal peptide (SP), which is expressed as a fusion with the therapeutic peptide. A convenient protein secretion signal is the sequence from human BMP10 protein, which has the sequence MGSLVLTLCALFCLAAYLVSG (SEQ ID NO: 131). However, any secretion signal with

downstream cleavage site may alternatively be used (see e.g. Hegde *et al.* (2006) *Trends Biochem Sci.*, 31(10), 563-71; <http://www.signalpeptide.de> for examples of possible sequences). Preferably, the SS / SP is host-matched: e.g. human signals would preferably be used for use in humans. Following cell secretion, the therapeutic peptide must be capable of penetrating a cell, and, if the therapeutic peptide is a transcription factor or other DNA-interacting molecule, targeting the nucleus of a cell. Thus, it is convenient that the therapeutic peptide further comprises at least one nuclear localisation sequence (NLS). A suitable NLS sequence is the SV40 NLS (PKKKRKV, SEQ ID NO: 107). However, the nuclear localisation sequence could be a host-derived sequence, such as the NLS from human protein KIAA2022 NLS (PKKRRKVT; NP_001008537.1, SEQ ID NO: 108) for use in humans; or the NLS from mouse primase p58 (RIRKKLR; GenBank: BAA04203.1, SEQ ID NO: 109) for use in mice. In other embodiments, any other suitable NLS known to the person of skill in the art could also be used; e.g. human or mouse NLSs from NLSdb (Nair *et al.* (2003) *Nucleic Acids Res.* 31(1): 397–399). In any of these embodiments, in order to enhance cellular uptake, it may be advantageous to combine more than one NLS sequence in tandem; for example, up to 6 NLS, such as 2 (SEQ ID NOs: 113, 114 and 115), 3, 4 or 5.

The expression construct may further be designed / adapted to place a peptide cleavage site between the SS or SP sequence and the therapeutic peptide effector domain (e.g. such as a zinc finger peptide). Peptide cleavage at the cleavage site separates the therapeutic peptide sequence from the SS or SP sequence and, hence, cleaved therapeutic peptide sequences may remain inside the cell in which they were expressed (or may remain inside the cell in which it eventually penetrates), such that a therapeutic effect may be experienced in the cell that expressed the therapeutic peptide, or the cell in which the therapeutic peptide is delivered to. In preferred embodiments, the gene encoding the therapeutic peptide for active delivery may be constructed such that the NLS sequence or sequences are N-terminal to the therapeutic peptide / zinc finger peptide sequence when expressed. Suitably, also, the secretion signal (SS) or signal peptide (SP) may be arranged N-terminal to the zinc finger peptide sequence. In some particularly beneficial embodiments, the SS or SP sequence is N-terminal to the one or more NLS. Accordingly, cleaved therapeutic peptide advantageously retains the NLS in combination with the therapeutic effector molecule and, thus, the ability to target the nucleus via the NLS or NLSs. It will be appreciated that any suitable peptide cleavage sequence may be employed in conjunction with embodiments of the invention. One convenient cleavage site is the RIRR peptidase cleavage site. In alternative embodiments, where the therapeutic effect is to be delivered by targeting an organelle other than the nucleus, it will be appreciated that the therapeutic peptide may not comprise an NLS; and may instead include an alternative, appropriate, targeting / cell localisation sequence.

In summary, a therapeutic peptide or designer transcription factor secretion / cell-penetration system according to the invention may advantageously enable bystander cells (neighbouring cells that have not been directly transduced by the therapeutic peptide / transcription factor construct) to receive a steady flow of freshly-expressed therapeutic protein / transcription factor, which may significantly enhance the percentage of a target tissue / organ that can be treated (e.g. by gene regulation). For example, if only 25% of cells would continue expressing a non-secreted therapeutic peptide / artificial transcription factor at 6 months after transduction, then such a treatment could only have a maximum

efficacy of 25%. By contrast, if that first population of 25% of the target cells continue to express the therapeutic peptide and the expressed peptide is capable of secretion and subsequent cell-penetration, those 25% of expressing cells may deliver the therapeutic agent to a second population of the target cells, and thereby produce a much more effective functional signal to a much higher percentage of target cells (see Figure 7B).

The active delivery platform described here is particularly beneficial in conjunction with gene expression construct delivery in patients, and is amenable for a variety of monogenic diseases where targeted genes need to be switched on or off. The approach is especially amenable to direct, injectable therapies.

EXAMPLES

The invention will now be further illustrated by way of the following non-limiting examples.

Unless otherwise indicated, commercially available reagents and standard techniques in molecular biological and biochemistry were used.

Materials and Methods

The following procedures used by the Applicant are described in Sambrook, J. *et al.*, 1989 *supra.*: analysis of restriction enzyme digestion products on agarose gels and preparation of phosphate buffered saline. General purpose reagents, oligonucleotides, chemicals and solvents were purchased from Merck (UK). Enzymes and polymerases were obtained from New England Biolabs (NEB, UK).

Vector and Zinc Finger Peptide (ZFP) Construction for Binding GAA Repeats and Fataxin (FXN) Promoter Sequences.

To build a zinc finger peptide (ZFP) frameworks that recognise 3'-AAG-5' or 3'-GAA-5' repeat DNA sequences (which are found within expanded GAA-repeats), or which recognise FXN promoter sequences, a zinc finger scaffold based on the wild-type backbone sequence of the zinc finger region of wild-type human Zif268 was selected. Amino acid residues responsible for DNA target recognition (i.e. the 'recognition sequence', which essentially corresponds to the α -helical region of the framework) were first designed having regard to known zinc finger amino acid-nucleic acid recognition codes (e.g. Isalan *et al.* (1998) *Biochemistry* 37(35): 12026-12033; WO 2012/049332).

For binding to 3'-AAG-5' repeat sequences, the amino acids residues at the -1, 3 and 6 positions of each zinc finger alpha-helix were selected to be Q, N and R, respectively, for optimal interaction and binding specificity to the target trinucleotide repeat sequence (see also Figure 1A).

For binding to 3'-GAA-5' repeat sequences, the amino acids residues at the -1, 3 and 6 positions of each zinc finger alpha-helix were selected to be R, N and Y, or R, N and Q, respectively, for optimal interaction and binding specificity to the target trinucleotide repeat sequence (see also Figure 1B).

Therefore, zinc finger peptides to targeting the 5'- AAG -3' binding site may have one of two or more recognition sequences.

For binding to frataxin promoter sequences, the amino acid residues are again selected having regard to known zinc finger amino acid-nucleic acid recognition codes; and so the amino acid residues selected at each of the -1, 3 and 6 positions are dependent on the target sequence and so the sequences of each zinc finger are expected to vary accordingly.

Suitable target sequences within the frataxin gene promoter (*Homo sapiens* (human); Gene ID: 2395) (SEQ ID NO: 67) were selected on the basis of being within approx. 1 kbp of the transcription start site (TSS) to enable efficient gene activation. Initially two sites were selected, the first at -621 from the TSS, i.e. 3'-GAG CCG AGT GAC GTT CGA GAC GGA GGG -5' (SEQ ID NO: 73 in 5' to 3' orientation), and the second at -959 from the TSS, i.e. 3'- CGT GGT GGT GTG GGT CGA-5' (SEQ ID NO: 75 in 5' to 3' orientation). Zinc finger recognition helix sequences were designed according to the nucleic acid sequence of the promotor region that was selected.

Initially, poly-zinc finger peptides having 6, 9 and 11 zinc finger domains were produced and cloned into a pUC57 vector (Genscript Corporation (Piscataway, NJ), with the names and sequences indicated in Table 1 below. This vector also included a T7 promoter, an N-terminal NLS (PKKRRKVT for use in human cells, SEQ ID NO: 108; and RIRKKLR for use in mouse cells, SEQ ID NO: 109). Subcloning was performed similarly to that previously described in WO 2012/049332. The zinc finger peptides were then subcloned into the mammalian expression vector pCDNA 3.1 (Invitrogen) under a CMV promoter; and for all zinc finger peptides either the p65 RelA (human or mouse) or VP64 (synthetic adapted *Herpes simplex*) transcription activation domain coding sequence was introduced at the C-terminus using specific restriction sites or by Gipsen assembly.

In all cases, a peptide linker sequence based on G and S amino acids was placed between the zinc finger peptide and the effector domain as described in the inventors' earlier work (WO2012/049332; WO2017/077329; WO2022/003361).

For poly-zinc finger peptides having 6, 9 and 11 zinc finger domains, targeting GAA-repeat or FXN promoter sequences, nucleotide sequences were produced as described herein.

Phage ELISA experiments as previously described (Isalan *et al.* (2001), *Nat. Biotechnol.* 19: 656–660), were performed to guide the alpha-helix recognition sequence design to ensure that the zinc finger peptides have an appropriate binding strength and selectivity to target nucleotide sequences.

In Vitro Gel Shift Assays

Based on the pUC57 vector zinc finger constructs, appropriate forward and reverse primers were used to generate PCR products for *in vitro* expression of the ZFP, using the TNT T7 Quick PCR DNA kit (Promega). Double stranded DNA probes with different numbers of GAA repeats were produced by Klenow fill-in as described in WO 2012/049332. 100 ng of double stranded DNA was used in a DIG-

labelling reaction using Gel Shift kit, 2nd generation (Roche), following the manufacturer's instructions. For gel shift assays, 0.005 pmol of DIG-labelled probe were incubated with increasing amounts of TNT-expressed protein in a 20 µl reaction containing 0.1 mg/ml BSA, 0.1 µg/ml polydI:dC, 5% glycerol, 20 mM Bis-Tris Propane, 100 mM NaCl, 5 mM MgCl₂, 50 mg/ml ZnCl₂, 0.1% NonidetP40 and 5 mM DTT for 1 hour at 25°C. Binding reactions were separated in a 7% non-denaturing acrylamide gel for 1 hour at 100 V, transferred to a nylon membrane for 30 min at 400 mA, and visualisation was performed following manufacturer's instructions.

Cell Culture and Gene Delivery

The cell line HEK-293T (ATCC) was cultured in 5% CO₂ at 37°C in DMEM (Gibco) supplemented with 10% FBS (Gibco). Qiagen purified DNA was transfected into cells using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Briefly, cells were plated onto 6 well plates (VWR, UK) to a density of 60% and 300 ng of reporter plasmid, 700 ng of ZFP expression plasmid and 2 µl of Lipofectamine 2000 were mixed and added to the cells. Cells were harvested for analysis 48 or 72 hours later.

Human fibroblast cell lines GM03816 (300, 500 GAA-repeats); GM04078 (600,850 GAA-repeats), and a control WT cell line (negative for any GAA expansion) were obtained from the Coriell Institute (US) and were cultured according to the provider's instructions.

Briefly, cells were cultured in 5% CO₂ at 37°C in DMEM supplemented with 15% FBS (Gibco). DNA was transfected into cells using Lipofectamine LTX. Briefly, cells were plated onto 6-well plates (VWR, UK), to a density of 60% and 1 µg of plasmid containing ZFP under the CMV promoter was mixed with 5 µl of Lipofectamine LTX and added to the cells. Next cells were harvested for analysis 48 or 72 hours post-transfection.

Flow Cytometry Analysis

Cells were harvested 48 hours post-transfection and analysed in a BD FACS Canto Flow cytometer using BD FACSDiva software.

Western Blot

293T cells were harvested 48 hours post-transfection in 100 µl of 2xSDS loading dye with Complete protease inhibitor (Roche). 20 µl of sample was separated in 4-15% Criterion Tris-HCl ready gels (BioRad) for 2 hours at 100V, transferred to Hybond-C membrane (GE Healthcare) for 1 hour at 100V. Proteins were detected with either the primary antibody anti β-actin (Sigma A1978) at 1:3000 dilution or anti-EGFP (Roche) at 1:1500 dilution and with a peroxidase-conjugated donkey anti-mouse secondary antibody (Jackson ImmunoResearch) at 1:10000 dilution. Visualisation was performed with ECL system (GE Healthcare) using a LAS-3000 imaging system (Fujifilm). Human fibroblast cells were trypsinised and harvested in PBS containing Complete protease inhibitor (Roche). Cells were resuspended in RIPA buffer (1% TritonX-100, 1% sodium deoxycholate, 40 mM Tris-HCl, 150 mM NaCl, 0.2% SDS, Complete), incubated in ice for 15 min, and were centrifuged at 13000 rpm for 15 min. The supernatant was collected and protein concentration was determined using BioRad's D_c protein assay.

60 µg of protein was separated in a 5% Criterion Tris-HCl ready gel (BioRad) for 2 hours at 100V, transferred using iBlot Dry Blotting System (Invitrogen) for 8 min and endogenous FXN protein was detected with anti-FXN primary antibody (Invitrogen FXN Monoclonal Antibody 18A5DB1) at a 1:1000 dilution or anti-FXN rabbit monoclonal antibody (abcam, Ab219414).

5

Luciferase activity

The HEK-293T cells were transfected with ZFP plasmids and a reporter plasmid carrying a luciferase gene under a control of 2.5kb human frataxin promoter (Addgene cat no. 14980) and were harvested 48H or 72H post-transfection. Cells were lysed in the 1 x lysis reagent, spun down and mixed with the Luciferase Assay reagent (Promega, E1500) as per manufacturer's instruction. Signal was detected with a microplate reader.

10

Production of Adeno-Associated Viral Vector

rAAV2/1, rAAV2/9 or rAAV-PHP.eB vectors containing zinc finger peptides / effectors of the invention as described in WO 2017/077329, e.g. containing a pCAG promoter (CMV early enhancer element and the chicken beta-actin promoter) and WPRE (Woodchuck post-translational regulatory element), can be produced, for example, at the Centre for Animal Biotechnology and Gene Therapy of the Universitat Autònoma de Barcelona (CBATEG-UAB; see also Salvetti *et al.* (1998) *Hum. Gene Ther.* 9: 695-706). Recombinant virus can be purified by precipitation with PEG8000 followed by iodixanol gradient ultracentrifugation with a final titre of approx. 10^{12} genome copies/ml.

15

20

Animals - Frataxia Transgenic Mice

For this study we used the transgenic expansion repeat model and wild-type (WT) mice. For example, the Jackson Laboratories (Jax) Frataxia mouse model: YG8 800 Fxnnull::YG8s(GAA)>800. Fxnnull::YG8s(GAA)>800 mice are a human FXN YAC transgenic mouse model harbouring a global null allele of mouse frataxin (Fxnnull) and the human FXN YAC transgene single repeat YG8s with a GAA repeat size of >800. Jax breeding of hemizygous mice with noncarrier mice results in 58% of offspring that are hemizygous, and 80% of those hemizygotes stay in the 775-900 GAA -repeat range. Recently, it has been demonstrated that this model accurately reflects the human disease with a progressive neuromuscular degeneration and heart hypertrophy at 26 to 30 weeks of age (PMID: 36089099). Other mouse models include B6.129-Fxn^{tm1Pand}/J (Jax stock No: 008470), expressing a (GAA)₂₃₀ expansion repeat from the endogenous *Fxn* locus. Homozygotes produce an average of 75% of wild-type levels of frataxin protein. Another strain known as FVB;B6.Tg(FXN); Fxn- (Jax stock no: 018299) harbours the FXN*500GAA transgene (Tg(FXN)1Sars) and a frataxin knockout allele (Fxn^{tm1Mkn}). In practice, any suitable GAA expansion model may be used.

25

30

35

All animal experiments were conducted in accordance with Directive 86/609/EU of the European Commission, the Animals (Scientific Procedures) 1986 Act of the United Kingdom, and following protocols approved by the Ethical Review Body of Imperial College London.

40

Stereotaxic Surgery

Briefly, mice are anaesthetised with isoflurane for any surgical application and fixed on a stereotaxic frame if necessary. Buprenorphine is injected at 8 µg/kg to provide analgesia.

AAVs are injected bilaterally or unilaterally (depending on the study) into various brain regions using a 10 µl Hamilton syringe at a rate of 0.25 µl/min controlled by an Ultramicropump (World Precision Instruments). For each injection, a total volume of 1.5 to 3 µl (approx. 2×10^9 genomic particles) or 1.5 µl PBS is injected. For example, a two-step administration may be performed as follows: 1.5 µl are injected at -3.0 mm DV, the needle is let to stand for 3 minutes in position, and then the other half is injected at -2.5 mm DV, as in case of intra-striatal injections.

In some studies, mice are injected only in one hemisphere with AAV expressing the test protein (either zinc finger or GFP control protein), or with PBS as a negative control.

Alternative AAV delivery routes to stereotaxis are also applied, including but not limited to: intrathecal injection for transducing the CNS (100 µl of approx. 1×10^{11} genomic particles per mouse); intrajugular injection for transducing mainly the heart (100 µl of approx. 1×10^{11} genomic particles per mouse); tail vein injection for all tissues (100 µl of approx. 1×10^{13} genomic particles of AAV-PHP.eB per mouse).

Mice are sacrificed at different ages for posterior analysis by RT-PCR, immunohistochemistry or western blot; typically at 3, 6, 12 or 24 weeks after administration of agent.

Animal Behavioural Tests

Behavioural monitoring typically commences at 4 weeks of age and tests take place bimonthly until at least 26 weeks of age. All the experiments are performed double-blind with respect to the genotype and treatment of the mice.

Examples of behavioural tests that may be performed:

Clasping behaviour is checked by suspending the animal by the tail for 20 seconds. Mice clasping their hindlimbs are given a score of 1, and mice that do not clasp are given a score of 0.

Grip strength is measured by allowing the mice to secure to a grip strength meter, then pulling gently by the tail. The test is repeated three times and the mean and maximum strength recorded.

For the accelerating rotarod test, mice are trained at 4 weeks of age to stay on the rod at a constant speed of 40 rpm until they reach a criterion of 3 consecutive minutes on the rod. In the testing phase, mice are put on the rotarod at 40 rpm and the speed is constantly increased for 2 minutes until 80 rpm - 120 rpm is reached. The assay is repeated twice and the maximum and average latency taken to fall from the rod is recorded. In one example, a mouse treated with a zinc finger activator peptide of the invention was able to reach a speed of 120 rpm before falling off.

For the open field test, mice are put in the centre of a white methacrylate squared open field (70x70 cm), illuminated by a dim light (70 lux) to avoid aversion, and their distance travelled, speed and position

is automatically measured with a video tracking software (SMART system, Panlab, Spain). Other activities, such as rearing, leaning, grooming and number of faeces are monitored *de visu*.

For the hanging wire test, the test measures time of sustained limb tension to oppose the mouse weight.

- 5 The maximum time the mouse remains hanging is 300 seconds per test

For the paw print test, mice hind paws are painted with a non-toxic dye and mice are allowed to walk through a small tunnel (10x10x70 cm) with a clean sheet of white paper on the floor. Footsteps are analysed for three step cycles and three parameters measured: (1) stride length – the average distance between one step to the next; (2) hind-base width – the average distance between left and right hind footprints; and (3) splay length – the diagonal distance between contralateral hind paws as the animal walks.

10

For the evaluation of heart function by echocardiography. This test may be used to assess heart function in mice under anesthesia. Different parameters like left ventricular dimensions, stroke volume, ejection fraction are measured.

15

Examples of molecular analysis:

qRT-PCR

20

For studies of target gene expression *in vivo*, mice are humanely killed by cervical dislocation. As rapidly as possible, they are decapitated and various brain regions are dissected on ice and immediately frozen in liquid nitrogen for later RNA extraction.

RNA is prepared with an RNeasy kit (Qiagen) and reversed transcribed with Superscript III (Invitrogen).

25

Real Time PCR is performed in a qPCR thermocycler (Bio-Rad) using Taqman master mix (Bio-Rad). A specific set of primers and probes is used to assess molecular readouts of disease progression.

Immunohistochemistry

Mice are transcardially perfused with PBS followed by formalin 4% (v/v). Tissues are removed and post-fixed overnight at 4°C in formalin 4% (v/v). Tissues are then cryoprotected in a solution of sucrose 30% (w/v), at 4°C, until they sink. Tissues are then frozen and sliced with a freezing microtome in six parallel coronal series of 40 µm (distance between slices in each parallel series: 240 µm). The indirect ABC procedure is employed for the detection of the neuronal marker Neu-N (1:100, MAB377 Millipore) in the first series; the reactive astroglial marker GFAP (1:500, Dako) in the second series; and the microglial marker Iba1 (1:1000, Wako) in the third series. Briefly, sections are blocked with 2% (v/v) Normal Goat Serum (NGS, Vector Laboratories) in PBS-Triton100 0.3% (v/v) and endogenous peroxidase activity blocked with 1% (v/v) hydrogen peroxide (H₂O₂) in PBS for 30 minutes at room temperature. This is based on similar approach used to assess the therapeutic effect of ZFP on disease progression in HD.

35

40

Subsequently, sections are incubated for 30 minutes at room temperature in: (i) primary antibody (at the concentration indicated above) in PBS with 0.3 % (v/v) Triton X100 and 2% (v/v) NGS; (ii) biotinylated secondary antibody in the same buffer; and (iii) avidin-biotin–peroxidase complex (ABC

Elite kit Vector Laboratories) in PBS-Triton X-100 0.3% (v/v). Sections are washed for 3x10 min in PBS and peroxidase activity is revealed with SIGMAFAST-DAB (3,3'-Diaminobenzidine tetrahydrochloride, Sigma-Aldrich) in PBS for 5 min. Sections are rinsed and mounted onto slides, cleared with HistoClear (Fisher Scientific) and cover-slipped with Eukitt (Fluka).

5

The fourth GFP-injected series is mounted onto slides and covered with Mowiol (Sigma-Aldrich) for fluorescence analysis.

Image Analysis

10 Determination of the volume of injection:

Five coronal slices per GFP-injected hemisphere from bregma 1.5 mm levels, separated by 240 μm , are photographed with a digital camera attached to a macrozoom microscope (Leica). The contours around the GFP-expressing area and dorsal striatum are manually defined and the area is measured with ImageJ software (National Institute of Health, USA). Volume is calculated as area per distance
15 between slices, according to the Cavalieri principle (Oorschot (1996), *J. Comp. Neurol.*; 366: 580–599).

Determination of O.D. for GFAP and Iba1 stainings:

Four coronal slices per mouse and hemisphere covering the specific brain region is captured with a 10x objective using a digital camera attached to a microscope (Leica DMIRBE). To assess FRDA
20 progression / treatment, the cerebellum, brain stem and spinal cord are the most relevant tissue areas. The O.D. of the areas is measured with ImageJ, the mean density per hemisphere calculated, and O.D. for GFAP and Iba1 of control hemispheres is subtracted from the injected hemisphere.

Determination of the neuronal density of the different brain region on striata as an example:

25 Cell density is calculated using an adaptation of the unbiased fractionator method (Oorschot (1996), *J. Comp. Neurol.*; 366: 580–599). Four coronal slices per mouse and hemisphere covering the striatum from bregma 1.5 mm levels are selected, and a region of interest of $447 \times 598 \mu\text{m}^2$ in the middle of the dorsal striatum is captured with a 15x objective, using a digital camera attached to a microscope (Leica DMIRBE). A grid image leaving 16 squares of $35 \times 35 \mu\text{m}^2$ is superimposed onto the pictures, and a
30 person (blinded to sample treatment) counts the number of stained nuclei.

Statistical Analysis

Data are analysed using the StatPlus package for Excel (Microsoft) and IBM SPSS Statistics 22. To test the inflammatory response, the difference of O.D. of the injected hemisphere versus the control
35 hemisphere is calculated, and a Student's t test is performed against the no difference value (0).

For neuronal density, a paired Student's t test of neuronal density in the injected hemisphere, versus the control hemisphere, is performed. Neuronal density is analysed across contralateral hemispheres with ANOVA, followed by post-hoc comparisons with the contralateral hemispheres of the PBS samples.

40 To test repression, the percentage of mutant *gene of interest* in the injected brain is calculated with respect to the control hemisphere, and a one sample Student's t test against the no repression value (100%) is performed. To ensure a fair comparison between injected and contralateral hemispheres,

only mice with <1% ZF expression in the contralateral hemisphere, relative to the injected hemisphere, are used for statistical analyses. To test the correlation between RNA levels of the different genes and ZF expression, a linear regression test is applied. To test expression levels across different times post-injection, a one-way ANOVA is performed. All significance values may, for example, be set at $p=0.05$.

5

Example 1

1A. Design of Zinc Finger Peptide (ZFP) Arrays to Bind GAA Repeats

10 It is known that zinc finger domains can be concatenated to form multi-finger (e.g. 6-finger) chains (Moore *et al.* (2001) *Proc. Natl. Acad. Sci. USA* 98(4): 1437-1441; and Kim & Pabo (1998) *Proc. Natl. Acad. Sci. USA* 95(6): 2812-2817). Our previous study, see WO 2012/049332, was the first to report on the systematic exploration of the binding modes of different-length ZFP to long repetitive DNA tracts of 35 to 160-repeat CAG coding sequence repeats. In this earlier study, rational design was used to
15 construct a zinc finger domain (ZFxHunt) that would bind the 5'- GC(A/T) -3' sequence in double stranded DNA.

This earlier teaching of how to produce extended arrays of poly-zinc finger peptides was adapted in the aspects and embodiments disclosed herein to provide extended arrays of zinc finger binding pairs, to
20 bind the trinucleotide repeat sequences 3'-AAG-5' or 3'-GAA-5' (see Materials and Methods above and Figures 1A and 1B) found in pathogenic frataxin genes. In contrast to the earlier study, the poly-zinc finger peptides of this invention are adapted to bind long 300 to 900-repeat non-coding GAA repeat sequences.

25 In various aspects and embodiments, both 3'-AAG-5' and 3'-GAA-5' repeat sequences were targeted as part of the zinc finger peptide 'tuning' process to understand and manipulate binding interactions between different zinc finger peptides and their respective target sites.

To try to avoid the zinc finger peptides of the invention losing their register with cognate DNA (after 3
30 or more adjacent fingers and 9 contiguous base pairs of double helical DNA), the linker sequences were carefully designed. In particular, the length of the linkers between adjacent zinc fingers in the arrays was modulated. In this way, the register between the longer arrays of zinc finger peptides, especially on binding to dsDNA, could be optimised. Using structural considerations, it was decided to periodically modify the standard canonical linker sequences in the arrays. Therefore, canonical-like linker
35 sequences containing an extra Gly (or Ser) residue were included in long zinc finger arrays (e.g. of 6 or more zinc finger domains) after every 2-zinc fingers, and flexible (up to 29-residue) linker sequences were included in long zinc finger arrays (e.g. of 9 or more zinc finger domains) after every 5 or 6 zinc fingers. In this way, different zinc fingers peptide constructs could be tested for optimal length-dependent discrimination.

40

As described above, zinc finger peptide arrays of 6, 9 and 11-zinc fingers were constructed against the various genomic target sequences.

Sequences of the various zinc finger peptides having 11-zinc finger domains arranged in tandem for binding to the 3'-AAG-5' or 3'-GAA-5' repeat sequence are given in the table below. These zinc finger peptides were designed for use as transcriptional activators in order to increase expression of a frataxin transcripts and protein. The sequences of Table 5 below, therefore, include a linker peptide, for example, LRQKDGGGS (SEQ ID NO: 130) at the C-terminus for attaching to a transcriptional effector domain. It will be appreciated, however, that the linker of SEQ ID NO: 130 may be replaced with any one or the linkers of SEQ ID NOs: 126 to 129, for example, as desired.

ZF1 3'-AAG-5'-binding amino acid sequence (SEQ ID NO: 97 + SEQ ID NO: 130):

YACPVESCDRRFSQSGNLTRHIRIH TGSQKP
 FQCRICMRNFSQSGNRKRHIRTH TGEKP
 FACDICGRKFAQSGNLTRHTKIH TGSQKP
 FQCRICMRNFSQSGNLTRHIRTH TGEKP
 FACDICGRKFAQSGNRKRHTKIH LRQKDGGGSGGGGSGGGGSQKP
 FQCRICMRNFSQSGNLTRHIRTH TGEKP
 FACDICGRKFAQSGNRKRHTKIH TGSQKP
 FQCRICMRNFSQSGNLTRHIRTH TGEKP
 FACDICGRKFAQSGNRKRHTKIH TGSQKP
 FQCRICMRNFSQSGNLTRHIRTH TGEKP
 FACDICGRKFAQSGNRKRHTKIH LRQKDGGGS

ZF1a 3'-AAG-5'-binding amino acid sequence (SEQ ID NO: 98 + SEQ ID NO: 130):

YACPVESCDRRFSQSANLGRHIRIH TGSQKP
 FQCRICMRNFSNSANLSRHIRTH TGEKP
 FACDICGRKFAQSANLGRHTKIH TGSQKP
 FQCRICMRNFSNSANLSRHIRTH TGEKP
 FACDICGRKFAQSANLGRHTKIH LRQKDGGGSGGGGSGGGGSQKP
 FQCRICMRNFSNSANLSRHIRTH TGEKP
 FACDICGRKFAQSANLGRHTKIH TGSQKP
 FQCRICMRNFSNSANLSRHIRTH TGEKP
 FACDICGRKFAQSANLGRHTKIH TGSQKP
 FQCRICMRNFSNSANLSRHIRTH TGEKP
 FACDICGRKFAQSANLGRHTKIH LRQKDGGGS

ZF1b 3'-AAG-5'-binding amino acid sequence (SEQ ID NO: 99 + SEQ ID NO: 130):

YACPVESCDRRFSNSANLSRHIRIH TGSQKP
 FQCRICMRNFSQSQNLSRHIRTH TGEKP
 FACDICGRKFANSANLSRHTKIH TGSQKP
 FQCRICMRNFSQSQNLSRHIRTH TGEKP
 FACDICGRKFANSANLSRHTKIH LRQKDGGGSGGGGSGGGGSQKP
 FQCRICMRNFSQSQNLSRHIRTH TGEKP
 FACDICGRKFANSANLSRHTKIH TGSQKP
 FQCRICMRNFSQSQNLSRHIRTH TGEKP
 FACDICGRKFANSANLSRHTKIH TGSQKP
 FQCRICMRNFSQSQNLSRHIRTH TGEKP
 FACDICGRKFANSANLSRHTKIH LRQKDGGGS

ZF2 3'-GAA-5'-binding amino acid sequence (SEQ ID NO: 102 + SEQ ID NO: 130):

YACPVESCDRRFSRSDNLTQHIRIH TGSQKP
 FQCRICMRNFSRSDNRKQHIRTH TGEKP
 FACDICGRKFARSDNLTQHTKIH TGSQKP
 FQCRICMRNFSRSDNLTQHIRTH TGEKP
 FACDICGRKFARSDNRKQHTKIH LRQKDGGGSGGGGSGGGGSQKP
 FQCRICMRNFSRSDNLTQHIRTH TGEKP
 FACDICGRKFARSDNRKQHTKIH TGSQKP

FQCRICMRNFSRSDNLTQHIRTH **TGEKP**
 FACDICGRKFARSDNRKQHTKIH **TGSQKP**
 FQCRICMRNFSRSDNLTQHIRTH **TGEKP**
 FACDICGRKFARSDNRKQHTKIH **LRQKDGGGS**

ZF2a 3'-GAA-5'-binding amino acid sequence (SEQ ID NO: 103 + SEQ ID NO: 130):

YACPVESCDRRFSRSSLTYHIRIH **TGSQKP**
 FQCRICMRNFSRSANRQHIRTH **TGEKP**
 FACDICGRKFARSSNLSYHTKIH **TGSQKP**
 FQCRICMRNFSRSANLTQHIRTH **TGEKP**
 FACDICGRKFARSSNRQHTKIH **LRQKDGGGSGGGGSGGGGSQKP**
 FQCRICMRNFSRSANLSYHIRTH **TGEKP**
 FACDICGRKFARSSNRTQHTKIH **TGSQKP**
 FQCRICMRNFSRSANLSQHIRTH **TGEKP**
 FACDICGRKFARSSNRKYHTKIH **TGSQKP**
 FQCRICMRNFSRSANLSQHIRTH **TGEKP**
 FACDICGRKFARSSNRKQHTKIH **LRQKDGGGS**

ZF2b 3'-GAA-5'-binding amino acid sequence (SEQ ID NO: 104 + SEQ ID NO: 130):

YACPVESCDRRFSRNDNRITHIRIH **TGSQKP**
 FQCRICMRNFSRSSLSTHIRTH **TGEKP**
 FACDICGRKFARNDNRITHTKIH **TGSQKP**
 FQCRICMRNFSRSSLSTHIRTH **TGEKP**
 FACDICGRKFARNDNRITHTKIH **LRQKDGGGSGGGGSGGGGSQKP**
 FQCRICMRNFSRSSLSTHIRTH **TGEKP**
 FACDICGRKFARNDNRITHTKIH **TGSQKP**
 FQCRICMRNFSRSSLSTHIRTH **TGEKP**
 FACDICGRKFARNDNRITHTKIH **TGSQKP**
 FQCRICMRNFSRSSLSTHIRTH **TGEKP**
 FACDICGRKFARNDNRITHTKIH **LRQKDGGGS**

Table 5: Zinc finger peptide framework amino acid sequences of humanised or mousified 11-zinc finger peptides of the invention for binding to 3'-AAG-5' or 3'-GAA-5' repeat nucleic acid sequences. Nucleic acid-binding recognition sequences are underlined and zinc finger linker sequences are shown in bold. C-terminal linker sequence (SEQ ID NO: 130) is shown in italics.

1B. Design of Zinc Finger Peptide (ZFP) Arrays to Bind FXN Promoter Sequences

Following on from the construction of poly-zinc finger peptides for binding to GAA trinucleotide repeat sequences above, poly-zinc finger peptides were similarly designed and adapted to bind to targets in the human FXN promoter (see Materials and Methods above and Figure 1C and 1D). The same structural considerations were taken into account, and the various zinc finger peptides synthesised, having 6- and 9-zinc finger domains arranged in tandem, are indicated in the Table 6 below. Again, these zinc finger peptides were designed for use as transcriptional activators in order to increase expression of a frataxin transcripts and protein. Therefore, the sequences of Table 6 include a linker peptide, for example, LRQKDGGGS (SEQ ID NO: 130), at the C-terminus, for attaching to a transcriptional effector domain. It will be appreciated, however, that the linker of SEQ ID NO: 130 may be replaced with any one or the linkers of SEQ ID NOs: 126 to 129, for example, as desired.

ZF3 3'-GAG CCG AGT GAC GTT CGA GAC GGA GGG-5'-binding amino acid sequence (SEQ ID NO: 116 + SEQ ID NO: 130):

YACPVESCDRRFSRSSLTRHIRIH **TGSQKP**
 FQCRICMRNFSDSVLTRHIRTH **TGEKP**

	FACDICGRKFAQSSHLTKHTKIH	TGSQKP
	FQCRICMRNFSRSDNLTTHIRTH	TGEKP
	FACDICGRKFARSSLTTHTKIH	LRQKDGGGSGGGGSGGGGSQKP
	FQCRICMRNFS <u>DSSHLTQH</u> IRTH	TGEKP
5	FACDICGRKFARSSNLTEHTKIH	TGSQKP
	FQCRICMRNFSRSSHLTQHIRTH	TGEKP
	FACDICGRKFARSSHLTRHTKIH	LRQKDGGGS
10	ZF4 3'-CGT GGT GGT GTG GGT CGA-5'-binding amino acid sequence (SEQ ID NO: 119 + SEQ ID NO: 130):	
	FQCRICMRNFS <u>DSSHLTKH</u> IRTH	TGEKP
	FACDICGRKFARSDHLTTHTKIH	TGSQKP
	FQCRICMRNFSRSDHLTTHIRTH	TGEKP
	FACDICGRKFARSDTLTRHTKIH	TGSQKP
15	FQCRICMRNFSRSDHLTTHIRTH	TGEKP
	FACDICGRKFADSSHRKQHHTKIH	LRQKDGGGS

Table 6: Zinc finger peptide framework amino acid sequences of humanised or mousified 6- and 9-zinc finger peptides of the invention for binding to nucleic acid sequences in the frataxin (FXN) promoter (3'-GAG CCG AGT GAC GTT CGA GAC GGA GGG C-5' or 3'-A CGT GGT GGT GTG GGT CGA-5'). Nucleic acid-binding recognition sequences are underlined and zinc finger linker sequences are shown in bold. C-terminal linker sequence (SEQ ID NO: 130) is shown in italics.

Example 2

Binding of Zinc Finger Peptides to DNA Target Sequences and Activation of Reporter Gene In Vitro

To show that the zinc finger peptides of Example 1 are capable of binding to GAA-repeat and FXN promoter sequences, *in vitro* gel shift assays were carried out as follows.

First, the zinc finger peptide arrays containing 6-, 9- and 11-zinc finger domains of Examples 1A and 1B were constructed, expressed and tested in gel shift assays for binding to double-stranded DNA sequence probes. All zinc finger peptides of Example 1A and 1B demonstrated the ability to bind their respective, specific DNA probes *in vitro*.

In addition, the zinc finger activator peptides, containing the 6-, 9- and 11-zinc finger domains constructed according to Examples 1A and 1B can be tested for their ability to activate a reporter gene *in vitro*. Zinc finger activator peptides based on the ZF1, ZF2, ZF3 and ZF4 frameworks are expressed *in vitro* in cells containing a luciferase reporter construct operably linked to a promoter sequence containing their respective nucleic acid target binding sites.

The results of an *in vitro* reporter assay for peptide constructs based on the ZF3 and ZF4 frameworks is shown in Figure 2. The positive test reporter construct contained the luciferase gene linked downstream of a 2.5 kb section of the wild-type frataxin gene promoter sequence (i.e. the 2,500 base sequence immediately upstream of the wild-type frataxin gene transcription start site). As demonstrated, the activator peptide comprising the ZF3 zinc finger peptide was able to bind to its nucleic acid target site and cause strong activation and expression of the luciferase reporter in transient

transfections. ZF3 peptides linked to either the P65 or VP64 activation domains were shown to activate luciferase gene expression to similar levels. On the other hand, the ZF4-based activator peptide was unable to significantly activate luciferase gene expression in this assay. Figure 2 shows luciferase gene expression data at 72 hours from the start of zinc finger peptide expression. Data from 48 hours (not shown) followed the same pattern.

It is expected that the longer zinc finger peptides having 11-fingers and designed for optimal binding interactions with the target sites bind most specifically and efficiently to the longer repeat sequence target sites; whereas the shorter zinc finger peptides having 6- or 9-fingers exhibit less affinity but high specificity for their unique target sites. It is unclear why the ZF4-based zinc finger activator peptide did not appreciably activate luciferase reporter expression in this assay.

Example 3

Chromosomal Activation of human Mutant Frataxin (FXN) gene (Gene ID: 2395)

In this Example, a new class of zinc finger peptide for activating the loss-of-function mutation causing Friedreich's ataxia (FRDA) is demonstrated. FRDA is a lethal degenerative disease with no cure that currently affects approx. 30,000 to 40,000 people in Europe alone.

First, a suitable DNA target within or associated with the pathogenic frataxin gene is identified, and a suitable DNA-binding zinc finger peptide array along with an appropriate effector domain fusion is designed and engineered. In this case, since FRDA is a disease caused by loss of function / expression of the affected gene, and for testing in human cells, an appropriate human-compatible transcriptional activation domain was selected.

In this example, four binding sites / target sequences were selected, and novel synthetic zinc finger peptides for specifically binding to each of the different DNA sequence regions of the human frataxin locus were designed (as described above). The four zinc finger peptides are termed ZF1, ZF2, ZF3 and ZF4, herein; wherein ZF1 is an 11-zinc finger peptide designed to bind to the 5'-GAA-3' repeat sequence; ZF2 is an 11-zinc finger peptide designed to bind to the 5'-AAG-3' repeat sequence; ZF3 is a 9-zinc finger peptide designed to bind to the 5'-GGGAGGCAGAGCTTGAGTGCAGCCGAG-3' repeat sequence; and ZF4 is a 6-zinc finger peptide designed to bind to the 5'-AGCTGGGTGTGGTGGTGC-3' repeat sequence. All synthetic ZFPs were fused to and tested separately with two different potential activation domains.

Efficacy of the zinc finger activator peptides of this Example were assessed in two FRDA-model fibroblast cell lines GM03816 (which contained 300 and 500 GAA repeats on different alleles) and GM04078 (which contained 600 and 850 GAA repeats on different alleles). A wild-type control cell line (negative for any GAA-repeat mutation) was also included as a negative control. All cell lines were obtained from the Coriell Institute (US).

Cell culture and transfection:

Human fibroblast cell lines were cultured in DMEM medium, supplemented with 15% fetal bovine serum (FBS, Gibco). Cells were kept in suspension in tissue culture T75 flasks (NUNC, Thermo Scientific) at 37°C in a 5% CO₂ incubator and maintained between 2×10^5 and 8×10^5 cells/ml.

For transfection, cells were passaged at 1×10^5 cells, 24 hours before transfection. GM03816 or GM4078 or control (WT cell lines) cells were transfected with 1 µg of pcDNA 3.1-ZF-Activator plasmid or empty pcDNA3.1 plasmid. Negative control cells received transfection reagents only (mock). Transfections were conducted with the Lipofectamine LTX kit according to the manufacturer's instructions (Invitrogen). After transfection, cells were suspended in medium and incubated overnight under normal cell culture conditions, and then replaced with fresh medium. The cells were pelleted 72 hours post-transfection, washed twice with ice-cold PBS, resuspended in the TRIzol reagent (Ambion) and stored at -80°C for further analysis. Frataxin transcript levels were then measured using the Taqman qPCR method with a primer / probe mix obtained from Life technologies (Hs00175940_m1).

RNA extraction and Taqman real-time PCR expression analysis:

Total RNA from cells was extracted with the mini-RNA kit (Qiagen, UK), according to the manufacturer's instructions. The reverse transcription reaction was performed using MMLV superscript reverse transcriptase (Invitrogen) and random hexamers (Invitrogen). All qPCR reactions were performed with a thermocycler (Bio-Rad). The qPCR reaction was carried out using Taqman Master Mix buffer (Bio-Rad). mRNA copy number was determined in triplicate for each RNA sample by comparison with the geometric mean of three endogenous human housekeeping genes: *Gapdh*, *18S* and *Hprt* (Primer Design, UK). The frataxin transcripts were detected using pre-designed primers and probe mix from Applied biosystems (Hs00175940_m1).

Statistical analysis:

Quantitative real time PCR analysis was carried out using the $2(-\Delta\Delta C(T))$ method. Values were presented as mean $\pm$ SEM. Statistical analysis was performed using paired Student t tests (Excel). A p-value of 0.05 was considered as a significant difference.

Results:

In relation to the zinc finger activator peptide based on the ZF2 design, the results (see Figure 3) showed that the zinc finger peptide designed to bind the 5'-AAG-3' repeat sequence efficiently activated the locus of the frataxin gene and restored frataxin transcripts to wild-type levels.

In relation to the zinc finger activator peptide based on the ZF1 design, the results (also see Figure 3) showed that the zinc finger peptide designed to bind the 5'-GAA-3' repeat sequence achieved a significant (approx. 20%) restoration of frataxin transcript levels.

Both p65 and VP64 activation domains yielded a similar degree of frataxin activation in these assays (Figure 3).

In relation to the zinc finger activator peptide based on the ZF3 design, the results showed that the zinc finger peptide designed to bind the 18 nucleotide target sequence in the frataxin promoter (5'-GGGAGGCAGAGCTTGCAGTGAGCCGAG-3') efficiently activated the locus of the frataxin gene and restored frataxin transcripts to wild-type levels (see Figure 4). By contrast, there was no significant restoration of frataxin transcript levels by the 6-zinc finger activator peptide based on the ZF4 design under these conditions.

Again, both p65 and VP64 activation domains yielded a similar degree of frataxin activation in these assays (Figure 4).

In Friedreich's ataxia (FRDA), homozygous patients are fully symptomatic while heterozygous carriers are non-symptomatic and comprise up to 2% of the general population (PMID: 8596916). Therefore, restoring expression of frataxin by up to 50% of WT values is expected to be of therapeutic benefit and is achievable with synthetic zinc finger activators. As such, our newly developed zinc finger peptide activators, especially those based on the ZF2 and ZF3 zinc finger peptide designs are capable of reversing the frataxin deficiency in FRDA cell models by restoring its transcript levels to the level seen in non-mutant carriers / healthy individuals. Indeed, the restoration of frataxin transcript levels by 20% or more, as achieved in this Example by the zinc finger peptide activator based on the ZF1 design may also provide useful therapeutic results.

As the skilled person will understand, the efficacy of any of the zinc finger activator peptides of this invention can also or alternatively be assayed for activation of the mutant / pathogenic frataxin locus using primary human B lymphocyte or fibroblast cells isolated from various frataxin mutant carriers, as well as mouse models, where necessary or convenient.

Example 4

Cell Toxicity Assay

Since it would be advantageous for a ZFP-activator therapy / therapeutic to have low toxicity, dye-labelling cell viability assays were performed to test the (non-specific) toxicity of the zinc finger peptides of the invention.

HEK-293T cells are transfected with 400 ng of the indicated vector constructs using Lipofectamine2000 and harvested 48 hours after transfection. As a control Lipofectamine2000-only or non-transfected cells (negative) may be used. Cytotoxicity can be analysed using the Guava Cell Toxicity (PCA) Assay according to the manufacturer's instructions, and the results presented as the percentage of dead, mid-apoptotic and viable cells.

It is expected that the data will show that no statistically significant toxicity effects are produced in cells expressing zinc finger peptides of the invention, as compared to control experiments. It is expected that the transcriptional activation properties of the zinc finger peptides of the invention, and their potential

for stable expression, will further reaffirm that the peptides of the invention have significant potential for gene therapeutic applications.

Example 5

Long-Term Activation of FXN DNA Target in vivo

Frataxia / FRDA mouse models, as described in *Materials and Methods* (e.g. YG8 800 Fxnnull::YG8s(GAA)>800), can be used to assay long term repression. Previously, the inventors have used a similar molecular and behavioural approach to assay zinc finger peptide efficacy in Huntington's Disease (HD) mouse models (see e.g. Garriga-Canut *et al.* (2012), *Proc. Natl. Acad. Sci.*, 109, E3136–3145); Agustín-Pavón *et al.* (2016) *Mol. Neurodegener.*, 11(1):64). Briefly, rAAV-encoding zinc finger activator peptides (see *Materials and Methods*) are injected into appropriate mouse models by a variety of well-known routes including stereotaxis, intrathecal, and intravenous (e.g. intrajugular and tail vein).

For example, for stereotaxis, test injections are either performed only in one hemisphere (so that the contralateral hemisphere is left untreated for the purpose of having a baseline comparison) or in whole brains to monitor overall efficiency (*Molecular Neurodegeneration* 11(1):64 (2016)). Brain samples from sacrificed animals are taken at 2, 4, 6 and 24 weeks post-injection, and RNA levels are analysed via quantitative real-time PCR (Garriga-Canut *et al.* (2012), *Proc. Natl. Acad. Sci.*, 109, E3136–3145; Agustín-Pavón *et al.* (2016) *Mol. Neurodegener.*, 11(1):64).

Example 6

Efficacy of zinc finger protein AAV viral delivery to increase Frataxin Levels in an FRDA model mouse

AAV-ZFP vector preparation:

To establish the efficacy of zinc finger (ZF) transcription factors to upregulate the FXN DNA Target *in vivo*, the zinc finger DNA, coding for protein ZF2a (SEQ ID NO: 103 + SEQ ID NO: 130) was first synthesised. The ZF2a protein is designed for 3'-GAA-5'-binding, which is the repeat expansion found in Friedreich's Ataxia patients. The ZF2a-coding DNA sequence was placed under a CMV promoter (SEQ ID NO: 139) and the full sequence (SEQ ID NO: 147) was cloned into an AAV2/9 vector. Alternatively, this ZF2a-coding DNA sequence was placed under a CBH promoter with an intron (SEQ ID NO: 141), and the full sequence (SEQ ID NO: 148) was cloned into an AAV-PHP.eB vector. The AAV vectors contained p65 activation domains (p65AD; SEQ ID NO: 122). The constructs also contained combined WPRE and polyA signals (SEQ ID NO: 149), downstream of the coding region, in order to increase expression from mRNA (see sequences in Table 7). AAVs were produced by standard methods at titres of 1e+13 gc/mL (AAV2/9-CMV-ZF2a-p65AD-WPRE-polyA = AAV2/9) and 5e+12 gc/mL (AAV-PHP.eB-CbH-ZF2a-p65AD-WPRE-polyA = AAV-PHP.eB). For transduction, these recombinant AAV2/9 or AAV-PHP.eB viral vectors were used, as previously described (Agustín-Pavón *et al.* (2016) *Mol. Neurodegener.*, 11(1):64).

ZF2a-p65AD cloned under CMV promoter (DNA): (SEQ ID NO: 147)

ccaatgatagatctacataacttacggtaaatggcccgctggctgaccgcccacgacccccgccattgacgt
caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggt
5 aaactgcccacttggcagtagcatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaat
ggcccgctggcattatgcccagtagcatgaccttatgggactttcctacttggcagtagcatctacgtatttagtca
tcgctattaccatgggtgatgcggttttggcagtagcatcaatgggctggatagcggtttgactcacggggatttc
caagtctccacccccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgta
acaactccgcccaccATGGGCAGAATCAGAAAAAGCTGAGACTCGTGAAAGACCCCTACGCCTGTCCTGTTGAGA
10 GCTGCGACAGAAGATTCAGCAGAAGCTCTAACCTGACATACCACATCCGCATCCACACCGGCAGCCAGAAACCTT
TTCAGTGCAGAATCTGTATGCGGAACCTTCTCCAGATCCGCCAACAGGAAACAGCACATCAGAACACACACCGGTG
AGAAGCCTTTTGCCTGCGACATCTGCGGCAGAAAGTTTCGCCAGATCATCCAATCTGTCTTATCACACAAAGATCC
ACACCGGCTCCCAAAAGCCCTTCCAGTGCCGGATATGTATGCGTAACTTCAGCCGGAGCGCCAATCTGACCCAGC
ATATTTCGGACACACACCGGCGAAAAGCCTTTTCGCCTGCGACATCTGCGGCGGAAAATTTCGCTCGGAGCAGCAACA
15 GAATCCAGCACACCAAGATCCACCTGAGACAGAAAGACGGCGGGCGGGCTCAGGCGGGCGGGCAGCGGAGGAG
GCGGCAGCCAAAAGCCTTTCCAGTGCAGAATCTGCATGAGAAACTTCAGCCGGTCTGCCAACCTGAGCTACCATA
TTCGGACACACACCGGCGAGAAGCCTTTTCGCCTGCGATATCTGCGGCAGAAAGTTTCGCCAGATCCAGCAACCGGA
CGCAGCACACAAAGATCCACACAGGATCTCAGAAGCCTTTCCAATGTAGAATCTGTATGAGAACTTCAGCAGAA
20 GCGCCAACCTCTCTCAGCACATCAGAACCCACACCGGCGAAAAGCCTTTTCGCTTGCACATCTGCGGACGGAAGT
TCGCCAGAAGCAGCAATAGGAAGTACCACACCAAGATCCACACAGGAAGCCAGAAACCTTTTCAGTGCCGGATCT
GCATCGGGAACCTTCAGCAGAAGCGCCAACCTGAGCCAGCACATCAGAACCCACACAGGCGAGAAGCCTTTTGCCT
GTGACATATGCGGCAGAAAGTTTGGCAGAAGCAGCAACAGAAAGCAGCACACCAAGATCCATCTGCGGCAGAAAG
ATGGCGGAGGCAGCCCTATGGAATTCCAATACCTGCCCCGACACCGGATGATCGGCACAGAATCGAGGAAAAAAGAA
AGCGGACCTACGAAACCTTCAAGAGCATCATGAAGAAGTCTCCATTTCAGCGGCCCCACCGATCCTAGACCTCCCC
25 CAAGAAGAATCGCCGTGCCATCTAGAAGCTCTGCCAGCGTGCCTAAACCTGCCCCCTCAACCTTACCTTTTACAA
GTAGCCTGTCCACCATCAACTACGACGAGTTTCTACCATGGTGTTCCTTCTGGACAGATCAGCCAAGCTTCTG
CCCTGGCCCCCGCCCCCTCCTCAGGTGCTGCCCCAGGCCCCCTGCCCCGCCCCCGCCCCCGCTATGGTCAGCGCTC
TGGCCCAGGCCCCCTGCCCCAGTGCCTGTGCTGGCCCCCTGGCCCTCCTCAGGCCGTGGCCCCCTCAGCCCCCTAAGC
CCACCCAGGCAGGCGAAGGCACCCCTGAGCGAGGCCCTGCTTCAGCTGCAGTTCGACGACGAGGACCTGGGAGCCT
30 TGCTGGGCAATAGCACCGACCCCGCCGTGTTACCGACCTGGCTTCCGTGGACAACAGCGAGTTCAGCAGCTGC
TGAACCAGGGCATCCCCGTGGCCCCCACCACAAACAGAGCCTATGCTGATGGAATACCCCGAGGCCATTACCAGAC
TGGTGACCGGAGCTCAGAGGCCACCAGATCCTGCACCCGCCCTCTGGGCGCTCCCGGCTGCCAAATGGCCTGC
TGAGCGGTGATGAGGATTTTCAGCAGCATCGCCGACATGGACTTCAGCGCCCTGCTGAGCCAGATCAGCTCCTGA

ZF2a-p65AD cloned under CBH promoter (DNA): (SEQ ID NO: 148)

ccaatgatagatctgttacataacttacggtaaatggcccgctggctgaccgcccacgacccccgccattga
cgtcaatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttgg
cagtagcatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgcttggcatt
gtgccagtagcatgaccttatgggactttcctacttggcagtagcatctacgtatttagtcatcgctattaccatgg
40 tcgaggtgagccccacgttctgtcttcaactctccccatctccccccctccccaccccccaattttgtatttttta
tttttttaattatttttgtgcagcgatggggggcgggggggggggggggggcgcgcgccaggcgggggcgggggcg
agggggcgggggcgggggcgaggcgagaggtgcggcggcagccaatcagagcggcgcgctccgaaagtttcctttta
tggcgaggcgggcgggcgggcgggccctataaaaaagcgaagcgcgcgggcgggcgaggctcgtgcgacgctgcctt
cgccccgtgccccgctccgcgcgcgcctcgcgccgcccggcctctgactgaccgcggttaactcccacaggtg
45 agcgggcggggaacggcccttctcctccgggctgtaattagctgagcaagaggttaagggttaagggtggttgggt
ggtgggggtattaatgtttaattacctggagcaacctgcctgaaatcactttttttcaggttggggccaccATGGGCC
CTAAGAAGCGCAGAAAGGTGACCGGCGAAAGACCATACGCCTGCCAGTGGAAGCTGCGACAGAAGATTTCAGCA
GAAGCAGCAACCTGACCTACCACATTAGAATCCATACTGGAAGCCAAAAGCCTTTCCAATGCAGAATCTGCATGA
50 GAAATTTTAGCAGATCCGCCAATCGGAAACAGCACATCAGGACACACACCGGCGAGAAGCCTTTTGCCTGCGATA
TCTGTGGCCGCAATTTGCTAGAAGCAGCAACCTGAGCTACCATACAAAGATCCACACAGGAAGCCAGAAACCTT
TCCAATGTAGAATCTGCATGCGGAACCTTCAGCAGAAGCGCCAATCTTACACAGCACATCCGGACCCACACCGGCG
AAAAGCCTTTTGCCTGCGACATCTGCGGCAGAAAGTTTCGCCCGGTCCAGCAACAGGATTCAGCACACCAAGATCC
ACCTGAGACAGAAGGACGGCGGGCGGGCTCTGGCGGAGGGGGTTCTGGCGGCGGTGGAAGCCAGAAGCCTTTCC
AGTGTAGAATCTGCATGCGGAACCTTCTCTCGGTCTGCCAACCTGTCTTATCATATCCGGACACACACCGGAGAGA
55 AGCCTTTTGCCTGCGACATCTGTGGCAGAAAGTTTCGCTCGGTCCAGCAACCGGACCCAAACACACCAAAATCCACA
CCGGATCTCAAAAACCTTTTCAGTGAGAATCTGCATGAGAACTTCAGCCGCGAGCGCCAACCTGAGCCAGCACAC
TCAGAACCCACACCGGAGAAAAGCCTTTTGCCTGTGATATCTGCGGCAGAAAGTTTCGCCGAGAAGCAGCAATAGAA
AGTACCACACAAAGATTACACAGGCAGCCAGAAACCATTTTCAGTGCCGGATCTGCATGAGGAACCTTCAGCAGAA
GCGCCAACCTGAGCCAGCACATCCGGACCCACACAGGCGAGAAAGCCCTTTGCTTGTGATATCTGCGGCAGAAAGT

87

TGCGCCAGAAGCAGCAACCGGAAGCAGCATACCAAGATCCACCTGCGGCAGAAAGACGGCGGCGGAAGCCCTATGG
AATTCAGTACCTGCCTGACACCGATGATAGACACAGAATCGAGGAAAAGAGGAAGCGGACATACGAGACCTTCA
AGAGCATCATGAAAAAGTCCCCCTTCTCCGGCCCCACAGACCCAGACCTCCTCCTCGGCGGATCGCCGTGCCAA
GCAGAAAGTTCTGCTAGCGTGCCCCAAGCCCGCACCCCCAGCCTTACCCCTTCACAAAGCAGCCTGTCTACCATCAACT
ACGACGAGTTCCCTACAATGGTGTTCCTCCAGCGGCCAGATCAGCCAAGCCAGCGCCCTGGCTCCTGCCCCACCAC
AGGTGCTGCCCCAGGCCCCCTGCCCCCTGCCCCCGCCCCCTGCCATGGTGTCCGCCCTGGCCCCAGGCCCCCGCCCCAG
TCCCCGTGCTGGCCCCAGGCCCCCTCTCAAGCTGTTGCCCTCCTGCTCCTAAGCCTACCCAGGCCGGCGAGGGCA
CCCTGCTCCGAGGCCCTGTCTGACGCTGCACTGTGCAGACGAGGACCTGGCGCCCTGCTGTGGGCAACAGCACCAGTAC
CTGCAGTGTTTACCAGACCTGGCCTCTGTGGACAACAGCGAGTTCGACGAGCTGCTGAACAGGGCATCCCCGTGG
CCCCCACAGCACCAGCCATATGCTGATGGAATACCTTGAGGCCATCAGCAGACTGGTGACCCGGCGCTCAGAGAC
CTCCAGACCCCGCTCCTGCCCCCTGTGGGCGCCCCCGGCCCTGCCTAACGGCCTGCTGAGCGGCACGAAGATTTC
GCTCCATCGCCGACATGGACTTCAGCGCCCTGCTCAGCCAGATCAGCAGCTGA

WPRe-polyA sequence placed immediately downstream of ZF2a coding sequence (DNA): (SEQ ID NO: 149)

gaattcgatatcaagcttatcgataatcaacctctggattacaaaatttgtgaaagattgactggatttcttaac
tatgttgctccttttacgctatgtggatacgctgctttaatgcctttgtatcatgctattgcttcccgtatggct
ttcattttctcctccttgataaaatcctgggtgctgtctctttatgaggagtgtggcccggtgtcaggcaacgt
ggcgtgggtgtgcactgtgtttgctgacgcaacccccactggttggggcattgccaccacctgtcagctcctttcc
gggactttcgctttccccctccctattgccacggcggaactcatcgccgcctgccttgcccgcgtgtggacaggg
gctcggctgttgggactgacaattcctgggtgttgcggggaaatcatcgtcctttccttggctgctcgctgt
gttgccacctggattctgcgcgggacgtccttctgctacgtcccttcggccctcaatccagcggaccttccttcc
cgcggcctgtgcgggctctgcggcctcttccgcgtcttcgccttcgccctcagacgagtcggatctcccttttg
gccgcctccccgcatacgataccgtcgacctcgaggggggggcccggtaccgccgaatcaattcactcctcaggtgc
aggctgcctatcagaagggtggctgtgtgtggcgaatgccttggctcacaaatcaccactgagatctttttccct
ctgccaaaaattatggggacatcatgaagccccttgagcatctgacttctggctaataaaggaaattttttca
ttgcaatag

Table 7: Zinc finger DNA expression sequences for cloning into AAV vectors. Promoter sequences (CMV, CBH) are indicated in lower case. Protein coding sequences (NLS, zinc finger ZF2a and p65 activation domain - p65AD) are uppercase. Functional sequences to increase RNA expression (WPRE and polyA signal) are in lowercase italics.

Mouse injections and analysis of FXN upregulation by qRT-PCR: Fxn^{em2.1}Lutz

Tg(FXN)YG8Pook/800J(Fxnnnull::YG8s(GAA)>800 = YG8s) male and female mice (JAX stock# 030395) were bred and genotyped. Wild-type mice (C57BL/6J; JAX Stock# 000664) were used as controls. At 6 weeks of age, mixed sex mice (5 per group) were randomised into 6 groups as detailed in Table 8 below:

Group	Genotype	N	Sex	Treatment	Dose	Dosing Route	Dosing	Necropsy
1	WT	5	3M 2F	Placebo	10 μ l, 1E+11 gc	IT	Once at 8 weeks of age	3 Weeks post dose
2	YG8s	5	3M 2F	Placebo	10 μ l, 1E+11 gc	IT	Once at 8 weeks of age	3 Weeks post dose
3	YG8s	5	3M 2F	AAV2/9	10 μ l, 1E+11 gc	IT	Once at 8 weeks of age	3 Weeks post dose
4	WT	5	3M 2F	Placebo	5E+13 gc/kg	IV	Once at 8 weeks of age	8 days post dose
5	YG8s	5	3M 2F	Placebo	5E+13 gc/kg	IV	Once at 8 weeks of age	8 days post dose

6	YG8s	5	3M 2F	AAV- PHP.eB	5E+13 gc/kg	IV	Once at 8 weeks of age	8 days post dose
---	------	---	-------	----------------	----------------	----	------------------------------	---------------------

Table 8: Study design. M=male; F=female; Placebo = PBS vehicle; IT = intrathecal; IV = intravenous (adult tail vein).

5 Mice were injected at 8 weeks of age by standard injection methods, including intrathecal injection (10 µl, 1E+11gc) or intravenous injection (10 ml/kg, 5E+13 gc/kg viral preparation), as previously described (see Materials and Methods). Mice were euthanised at the times indicated (Table 8) and tissues were harvested and assayed for restored levels of Frataxin gene expression by qRT-PCR (Taqman; see Materials and Methods), and FXN ELISA (Human Frataxin ELISA Kit, Abcam Ltd). Thus, Frataxin gene expression was assayed both at the transcript and protein levels. The data demonstrates that the zinc
10 finger peptide of the disclosure (ZF2a) works efficiently *in vivo* and is capable of restoring frataxin transcript levels to within the therapeutic window (Figure 5; RNA). Importantly, frataxin protein levels were also restored (Figure 5; PROTEIN).

15 The study in the FRDA mouse model also showed that the zinc finger peptide of the disclosure (ZF2a) was well tolerated for up to 3 weeks – i.e. until the time when the study was terminated. To obtain this data, Taq-man analysis of transcripts known to be associated with inflammation and immunological responses was performed (using the same RNA samples as in Figure 5). It was found that none of these validated control markers were significantly deregulated in mice treated with the therapeutic zinc
20 finger peptide, indicating that the therapeutic zinc finger peptide is well tolerated (Figure 6).

Taken together, the results demonstrate that both intrathecal (IT) and intravenous (IV) injections of either AAV preparation (AAV2/9, AAV-PHP.eB) are capable of rescuing / restoring Frataxin levels in Frataxin mouse models based on intronic poly-GAA expansions, via the designed zinc finger activator peptides, and that the formulations are well-tolerated and long-lasting.

25 **Example 7**

Active delivery of zinc finger activator peptides in vivo enhances gene regulation when compared with standard delivery

30 The inventors have previously shown that zinc finger peptide (ZFP) therapies are currently limited by long-term expression efficiency. For instance, for the treatment of Huntington's disease, it was found that target mutant gene repression by zinc finger transcription factors was limited to only approx. 25% in the whole brain after 6 months (Agustín-Pavón *et al.* (2016) *Mol. Neurodegener.*, 11(1):64). The
35 concept of 'active delivery' could improve this situation by continuing to 'drip-feed' secreted cell-penetrating factors to neighbouring / bystander cells in the brain and other tissues (Figures 7A and 7B).

In this Example, a method for achieving enhanced control of gene activation *in vivo* in mouse and human cells / systems with artificial gene-regulatory transcription factors is demonstrated. This method
40 is based on the concept of 'active delivery' of zinc finger peptides (ZFPs) by a combination of gene

expression, secretion and cell-penetration of engineered transcription factors, such as the zinc finger activator peptides of the invention. Beneficially, this approach exploits the intrinsic cell penetrating properties of zinc finger peptides (Gaj *et al.* (2012), *Nat. Methods*, 9(8):805-807; Gaj *et al.* (2014), *ACS Chem. Bio.*, 9(8):1662-1667; Liu *et al.* (2015), *Mol. Ther. Nucleic Acids*, 10;4:e232; and Lee *et al.* (1997), *Virus Research*, 52(1):97-108.

Two engineered zinc finger activator peptides were used in this Example: (i) an 11-zinc finger peptide that demonstrates preferential binding to mutant GAA trinucleotide repeat sequences (e.g. as found in FRDA - Friedreich's Ataxia); and (ii) a 9-zinc finger peptide that demonstrates preferential binding to the frataxin promoter (as described elsewhere herein).

Method steps:

1. In the first step, expression cassettes were engineered to contain (in 5' to 3' / N- to C- direction): the constitutive promoter / enhancer CMV; a protein secretion signal (SS) from human BMP10 protein (also known as a signal peptide (SP); SEQ ID NOs: 131 (prt) and 133 (dna)); a tandem array of two human Nuclear Localisation Signals (NLSs; PKRRKVT/TPKRRKVT (SEQ ID NO: 114 (prt) and 134 (dna)) to enhance cell-penetration by providing a net positive charge; an 11-zinc finger peptide fused to a transcription activation domain (from human p65 RelA). The pCMV-IRES-GFP vector backbone (Clontech) was used as the template for the construct, where the GFP can be used to monitor transfection efficiency. In this construct an RIRR (SEQ ID NO: 132 (prt) and 135 (dna)) peptide cleavage site was placed between the SP and the NLS. Two zinc finger activator peptides were tested: one 11-zinc finger peptide previously shown by the inventors to successfully target the 3'-GAA-5' repeat associated with FRDA (SEQ ID NO: 105; ZF2a-GAA-p65); and one 9-zinc finger peptide shown herein to target the sequence 3'-GAG CCG AGT GAC GTT CGA GAC GGA GGG-5' in the FXN promoter (SEQ ID NO: 117; ZF3-promoter-p65).

2. HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM) + 1 g/L D-glucose and pyruvate supplemented with 10% (v/v) foetal bovine serum (FBS; Life Technologies, UK) without antibiotics, at sub-confluent cell density, in an incubator at 5% CO₂ and 37°C. Cells were passaged every two days, using 0.05% trypsin-EDTA (Life Technologies, UK). Cells were transfected at 50-60% confluency, using 5 µl of Lipofectamine LTX (Invitrogen) and 1 µg of plasmid DNA (pCMV-SS-2NLS-ZFP-KOX-IRES-GFP or pCMV-IRES-GFP) per 10 cm plate using the manufacturer's protocol. 24 hours post transfection, transfection efficiency was checked using a fluorescence microscope and cells reached on average 90% transfection efficiency. Next, medium was replaced with fresh serum-free culture medium. Cells were cultured for a further 96 hours without medium replacement. Next, enriched medium containing secreted ZFP was harvested and centrifuged for 5 minutes at 800 x g at 4°C in order to remove cell debris. The supernatant fraction was retained.

3. The following cell lines were used as zinc finger peptide receivers: human fibroblast cell lines GM03816 (300, 500 GAA-repeats); GM04078 (600, 850 GAA-repeats), and a control WT cell line (negative for any GAA mutation) were obtained from the Coriell Institute (US) and were cultured according to the provider's instructions. Briefly, cells were cultured in 5% CO₂ at 37°C in DMEM

supplemented with 15% FBS (Gibco). DNA was transfected into cells using Lipofectamine LTX, as per manufacturer's instructions.

4. SF medium containing secreted zinc finger peptide from Step 2 was diluted in fresh medium to provide 0%, 50% or 100% v/v mixtures of zinc finger peptide medium to fresh medium; and this was added to separate samples of cell receivers from Step 3 and incubated for 96h. Next, all three sample lines were washed with PBS and harvested by a direct application of 1 ml of TRIZOL reagent (Invitrogen). Cell lysates were immediately frozen and stored at -80°C. The next day, cell lysates were incubated at 37°C for 2-3 minutes and placed on ice. 200 µl of chloroform was applied per 1 ml of cell lysate following by centrifugation at 8,000 x g at 4°C for 15 minutes. The upper aqueous fraction was then transferred into new tubes (approximately 400 µl) and an RNeasy Mini Kit (QIAGEN, UK) was used to extract total RNA following the manufacturer's instructions.

5. RNA samples (1 µg of total RNA) were treated with RNase-free DNase I (Promega, US) at 37°C for 1h, followed by deactivation at 65°C for 20 min. 1 µg of total RNA sample was reverse-transcribed using SuperScript III First - Strand Synthesis Kit (Invitrogen) according to manufacturer's instructions.

6. RT and Taqman qPCR: All qPCR reactions were performed using Light Cycler 480 Real Time Thermal Block Cycler in 384-well plates (Roche). Typically, 3 µl of approximately 5 ng/µl cDNA were used per reaction. For each biological replicate, three technical replicates were used. Sigma water was used as a negative control. qPCR cycling parameters were as follows: denaturation at 95°C for 20s, followed by 45 cycles of amplification at 95°C for 1min, and subsequently cooling at 40°C for 30s. Double Delta CT (cycle threshold) analysis was used for relative quantification, according to the equation $\text{Expression fold change} = 2^{(-\Delta\Delta Ct)}$.

Wild-type and mutant target mRNAs were analysed by Taqman qPCR. Values were normalized to the housekeeping gene human 18S. Error bars are SEM (n = 3). Student's t-test: *p < 0.05; **p < 0.01. Zinc finger peptide secretion leading to cell penetration and target gene repression are thus demonstrated *in vitro* in mouse and human cells.

The zinc finger peptide supernatant from HeLa cells (i.e. cell medium including secreted 11-zinc finger transcriptional activator peptide) can specifically activate FXN gene targets, as expected (data not shown). The target gene activation level is proportional to the concentration of zinc finger peptide in the medium to which the target cells are exposed. Activation is demonstrated in both whole brain and peripheral tissue (muscle, heart). Similar results were obtained for each zinc finger activator peptide against its target pathogenic sequence, showing in all cases that the zinc finger transcriptional activator peptides were able to specifically upregulate target disease gene sequences while leaving non-target gene expression essentially at normal, expected levels.

7. For active delivery *in vivo*, the desired gene construct or constructs (e.g. SEQ ID NO: 145) is/are subcloned into a suitable vector to give a full AAV Vector sequence (e.g. SEQ ID NO: 146) together with a suitable promoter-enhancer, for example CBh (SEQ ID NO: 140 or SEQ ID NO: 141). Other

suitable promoter-enhancers include the CMV promoter (SEQ ID NO: 139), the synthetic rat 1.8 kb eno2 (pNSE) promoter (SEQ ID NO: 142), the synthetic mouse 1.8 kb hsp90ab1 promoter (SEQ ID NO: 143) and the synthetic human 1.8 kb hsp90ab1 promoter (SEQ ID NO: 144). For brain, heart and muscle transduction, a recombinant AAV2/9 or AAV-PhP.eB viral vector was used, as previously described (Agustín-Pavón *et al.* (2016) *Mol. Neurodegener.*, 11(1):64). Delivery of viral vector was achieved by standard injection methods, including stereotaxis (2 µl viral preparation per hemisphere) and intrathecal or intravenous injection (100 µl viral preparation) as previously described (see Materials and Methods). For mouse model studies, the injections were carried out in FRDA models (e.g. Jackson Laboratories (Jax) Frataxin mouse model: YG8 800 Fxnnull::YG8s(GAA)>800); tissues were harvested as for direct delivery and were assayed for restored levels of Frataxin gene expression by qRT-PCR and Western Blotting (see Materials and Methods). As expected, the active delivery constructs showed enhanced gene regulation effects both in terms of more cells being upregulated and for a greater period of time than with standard delivery.

Discussion

In these Examples, zinc finger peptides have been designed that are able to recognise and bind GAA-repeats and FXN promoter sequences which are found in FRDA Frataxin patients; and it has been shown that such proteins are able to induce transcriptional activation of target genes both *in vitro* and *in vivo*, in order to increase the amount of frataxin transcript.

Fusing the p65-ReIA or VP64 activation domain to the poly-zinc finger peptides of the invention was found to increase the expression of targeted genes.

In addition to the studies reported in the Examples above, activation of the mutant / pathogenic frataxin locus can also be assessed using primary human B lymphocyte or fibroblast cells isolated from various frataxin mutant carriers: a collection of 58 cell lines is available from the Coriell Institute, US. Furthermore, a repository of >80 fibroblast lines from Friedreich's ataxia (FRDA) patients, carriers and healthy individuals has been established at the University of Alabama at Birmingham. Primary cultures isolated from one or more frataxin mouse models; e.g. B6.129-Fxntm1Pand/J (Jax stock No: 008470), expressing a (GAA)₂₃₀ expansion repeat from the endogenous Fxn locus, wherein homozygotes cells produce an average of 75% of wild-type levels of frataxin protein; Fxn^{em2.1Lutzy} Tg(FXN)YG8Pook/800J (Fxnnull::YG8s(GAA)>800 = YG8s) (JAX stock no: 030395); and another strain known as FVB;B6.Tg(FXN); Fxn⁻ (Jax stock no: 018299), which harbours the FXN*500GAA transgene (Tg(FXN)1Sars); and a frataxin knockout allele (Fxntm1Mkn) could be used in further studies, for instance, to assess the effectiveness of additional zinc finger peptides, and, for example, zinc finger peptides based on the ZF4 design.

Toxicity effects of therapeutic molecules, especially for use in gene therapy and other similar strategies that require mid- or long-term expression of a heterologous protein, is a particular issue. Indeed, studies have previously shown that non-self proteins can elicit immune responses *in vivo* that are severe enough to cause widespread cell death.

In order to improve the mid- to long-term effects of zinc finger peptide expression in target organisms, especially in the brain, the inventors have previously developed strategies to reduce the toxicity and immunogenicity of the potentially therapeutic zinc finger peptides and repressor proteins of the invention (see e.g. WO 2017/077329). Thus, in first aspects and embodiments, the present disclosure also provides zinc finger peptides and nucleic acid sequences that are suitable for activation of, and/or increase or restore expression levels of the frataxin gene, and its products, *in vivo* and/or *ex vivo* in both mouse and human cells. Likewise, the zinc finger peptides disclosed herein are suitable for the targeting and modulation of other genes – especially those containing long GAA-trinucleotide or promoter sequences of genes which are pathologically under-expressed.

As demonstrated in the Examples, both intrathecal (IT) and intravenous (IV) injections of AAV preparations expressing zinc finger activator peptides of the disclosure are capable of restoring Frataxin levels in Frataxin mouse models, and that the AAV-zinc finger peptide formulations are well-tolerated and long-lasting.

Gene therapy is an attractive therapeutic strategy for various neurodegenerative diseases. For example, lentiviral vectors have been used to mediate the widespread and long-term expression of transgenes in non-dividing cells such as mature neurons (Dreyer, *Methods Mol. Biol.* 614: 3-35). Additionally, further benefits are associated with the use of the ubiquitous promoter, pHSP (based on Hsp90) characterised in our earlier patent application, WO 2017/077329. In particular, these benefits of the invention are enhanced when the promoter is used in combination with rAAV2/9 vectors, based on a virus that infects a wide variety of cell types. Alternatively, the ubiquitous CBh promoter can be used with good results. Additionally, the neuron-specific promoter (pNSE) has been shown to provide comparable results. Similar effects can be expected in animal (human) subjects using either the mouse promoter or the human equivalent of the synthetic pHSP promoter used in some of these studies.

The benefits of the zinc finger activator peptides of the invention may be further enhanced when used in combination with the humanised 'active delivery' system for gene activators disclosed herein. In this regard, by creating zinc finger peptide constructs that comprise a combination of secretion and cell-penetration signal sequences / peptides, therapeutic peptides are created that are capable of directing their own secretion from the cell in which they were expressed, and their subsequent penetration of a neighbouring cell which they come into contact with, e.g. by diffusion. Once inside such a neighbouring cell, the zinc finger peptide of the invention can, for example, be targeted to the cell nucleus (e.g. by way of a nuclear localisation sequence) so that it can deliver its intended therapeutic effect within that neighbouring cell.

Accordingly, the active delivery system of the invention may provide one or both of prolonged therapeutic activity – by potentially continuing to deliver therapeutic peptides to cells that previously expressed, but no longer express, the therapeutic peptide (for example, a result of gene silencing); and broader / enhanced therapeutic effect – by delivery of active, therapeutic peptides to cells that were not initially infected / transduced with the therapeutic construct.

Conclusion

This study demonstrates that extended poly-zinc finger activator proteins can be designed and constructed to increase expression of target gene sequences both *in vitro* and *in vivo*, especially in cases where the pathology of a disease or disorder is based on an under-expression of a particular gene. Such zinc finger activator peptides comprise at least 6 zinc finger domains, and suitably at least 9 zinc finger domains. For example, zinc finger activator peptides of 9 and 11 zinc fingers in length have been shown to be particularly useful for the upregulation of pathogenic genes associated with expanded GAA-repeat sequences, such as for the potential treatment of Friedreich's ataxia (FRDA). Friedreich's ataxia (FRDA) is an autosomal recessive disorder manifesting with progressive ataxia, impaired speech, hearing and vision, cardiomyopathy, diabetes, and skeletal muscle abnormalities (PMID: 29053830). FRDA is a rare inherited genetic disorder with a prevalence of ~1 in 40,000 and there is currently no cure (PMID: 30905359). It is caused by GAA expansions of up to 900 units, within intron 1 of the frataxin locus, which in turn leads to down-regulation of its mRNA transcripts, by up to 90%, resulting in toxicity (PMID: 30905359). Homozygous patients are fully symptomatic while heterozygous carriers are non-symptomatic and comprise up to 2% of the general population (PMID: 8596916). Therefore, restoring expression of frataxin by up to 50% of wild-type values is expected to be of therapeutic benefit and is demonstrated to be achievable with synthetic zinc finger activator peptides described herein.

Accordingly, the Examples provided here demonstrated that a new class of zinc finger peptide activators can overcome the loss-of-function mutation that causes Friedreich's ataxia. Significantly, the results demonstrate that the specific targeting of the pathogenic frataxin gene and associated beneficial therapeutic effects can be achieved by identifying a suitable DNA target binding site in a genomic sequence that is specifically associated with the diseased / pathogenic gene, e.g. within the GAA repeat sequence; or by identifying a suitable genomic target site in the frataxin gene promoter sequence region, which may allow for the condition to be treated by upregulation of either pathogenic or non-pathogenic genes.

By designing a zinc finger activator peptide having both a suitable zinc finger DNA-binding domain and an appropriate human- (or mouse-) compatible transcription activation domain, the present inventors have thus demonstrated the beneficial upregulation of a target gene *in vitro* and *in vivo* for the treatment of a disease.

Moreover, it has been demonstrated that long-term gene therapy treatments involving up-regulation of pathogenic genes is enhanced through 'active delivery' of therapeutic agents to non-transduced target cells; i.e. by delivery of therapeutic peptides from cells in which they are expressed to neighbouring cells in which they are not expressed. In this way, despite a possible reduction in the proportion of cells in a target cell population that express therapeutic peptide over time, a relatively enhanced therapeutic effect can be maintained by secretion and cell penetration of therapeutic peptides from expressing cells into neighbouring, non-expressing target cells. By adapting the therapeutic zinc finger peptides of the

invention for active delivery, as described herein, it is believed that long-term (over 6 months) effective gene therapy treatment can be achieved *in vivo* from a single treatment / administration.

Sequences

SEQ ID NO:	Sequence Type	Sequence
1	Recognition (prt)	(Q/N) S (G/A/Q) N (L/R) (T/K/G/S) R
2	Recognition (prt)	QSGN (L/R) (T/K) R
3	Recognition (prt)	(Q/N) SANL (S/G) R
4	Recognition (prt)	(Q/N) S (A/Q) NLSR
5	Recognition (prt)	QSGNLTR
6	Recognition (prt)	QSGNRKR
7	Recognition (prt)	QSANLGR
8	Recognition (prt)	NSANLSR
9	Recognition (prt)	QSQNLSR
10	Recognition (prt)	R (S/N) (D/S/A) N (L/R) (T/K/S/I) (Q/Y/T)
11	Recognition (prt)	RSDN (L/R) (T/K) Q
12	Recognition (prt)	RS (S/A) N (L/R) (T/K/S) (Q/Y)
13	Recognition (prt)	R (S/N) (D/S) N (L/R) (S/I) T
14	Recognition (prt)	RSDNL TQ
15	Recognition (prt)	RSDNRKQ
16	Recognition (prt)	RSSNLTY
17	Recognition (prt)	RSANRKQ
18	Recognition (prt)	RSSNLSY
19	Recognition (prt)	RSANLTQ
20	Recognition (prt)	RSSNR IQ
21	Recognition (prt)	RSANLSY
22	Recognition (prt)	RSSNRTQ
23	Recognition (prt)	RSANLSQ
24	Recognition (prt)	RSSNRKY
25	Recognition (prt)	RSSNRKQ
26	Recognition (prt)	RNDNRIT
27	Recognition (prt)	RSSNLST
28	Recognition (prt)	R (S/N) SN (L/R) (T/S/K/I) R
29	Recognition (prt)	R (S/N) SNL (T/S) R
30	Recognition (prt)	RSSNLTR
31	Recognition (prt)	D (S/N) SV (L/R) (T/S/K/I) R
32	Recognition (prt)	D (S/N) SVL (T/S) R
33	Recognition (prt)	DSSVLTR
34	Recognition (prt)	Q (S/N) SH (L/R) (T/S/K/I) K
35	Recognition (prt)	Q (S/N) SHL (T/S) K
36	Recognition (prt)	QSSHLTK
37	Recognition (prt)	R (S/N) DN (L/R) (T/S/K/I) T
38	Recognition (prt)	R (S/N) DNL (T/S) T
39	Recognition (prt)	RSDNLTT
40	Recognition (prt)	R (S/N) ST (L/R) (T/S/K/I) T
41	Recognition (prt)	R (S/N) STL (T/S) T
42	Recognition (prt)	RSSTLTT
43	Recognition (prt)	D (S/N) SH (L/R) (T/S/K/I) Q
44	Recognition (prt)	D (S/N) SHL (T/S) Q
45	Recognition (prt)	DSSHLTQ
46	Recognition (prt)	R (S/N) SN (L/R) (T/S/K/I) E
47	Recognition (prt)	R (S/N) SNL (T/S) E
48	Recognition (prt)	RSSNLTE

49	Recognition (prt)	R (S/N) SH (L/R) (T/S/K/I) Q
50	Recognition (prt)	R (S/N) SHL (T/S) Q
51	Recognition (prt)	RSSHLTQ
52	Recognition (prt)	R (S/N) SH (L/R) (T/S/K/I) R
53	Recognition (prt)	R (S/N) SHL (T/S) R
54	Recognition (prt)	RSSHLTR
55	Recognition (prt)	D (S/N) SH (L/R) (T/S/K/I) K
56	Recognition (prt)	D (S/N) SHL (T/S) K
57	Recognition (prt)	DSSHLTK
58	Recognition (prt)	R (S/N) DH (L/R) (T/S/K/I) T
59	Recognition (prt)	R (S/N) DHL (T/S) T
60	Recognition (prt)	RSDHLTT
61	Recognition (prt)	R (S/N) DT (L/R) (T/S/K/I) R
62	Recognition (prt)	R (S/N) DTL (T/S) R
63	Recognition (prt)	RSDTLTR
64	Recognition (prt)	D (S/N) SH (L/R) (T/S/K/I) Q
65	Recognition (prt)	D (S/N) SHR (K/I) Q
66	Recognition (prt)	DSSHRKQ
67	FXN promoter sequence (-1,000 to 0 from transcription start point (5'-3') (dna)	TTTCTACAAAAAATACAAAACT AGCTGGGTGTGGTGGTGC ACACCTGTAGTCCCAGCTACTTGGAAGGCTGAAATGGGAAGACTGCTTGAGCCCCGGGAGGGAGAAAGTTGCAGTAAGCCAGGACCACACCACTGCACTCCAGCCTGGGCAACAGAGTGAGACTCTGTCTCAACAAACAAATAAATGAGGCGGGTGGATCACGAGGTGAGTAGATCGAGACCATCCTGGCTAACACGGTGAAACCCGTCTCTACTAAAAAAAAAAAAAAAAATACAAAAATTAGCCAGGCATGGTGGCGGGCGCCTGTAGTCCCAGTTACTCGGGAGGCTGAGGCAGGAGAATGGCGTGAAACCGGGAGGCAGAGCTTG CAGTGAGCCGAGAT CGCACCACTGCCCTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAATCAATCAATCAATCAATAAAATCTATTAACAATATTTATTGTGCACTTAACAGGAACATGCCCTGTCCAAAAAACTTTACAGGGCTTAACCTATTTTATCCTTACCACAATCCTATGAAGTAGGAACTTTTATAAAACGCATTTTATAAACAAGGCACAGAGAGGTTAATTAACCTGCCCCTCTGGTCACACAGCTAGGAAGTGGGCAGAGTACAGATTTACACAAGGCATCCGTCTCCTGGCCCCACATACCACACTGCTGTAAACCCATACCGGCGGCCAAGCAGCCTCAATTTGTGCATGCACCCACATCCCAGCAAGACAGCAGCTCCCAAGTTCCTCCTGTTTAGAATTTAGAAAGCGGCGGCCACAGGCTGCAGTCTCCCTTGGGTGAGGGTCCCTGGTTGCACTCCGTGCTTTGCACAAAGCAGGCTCTCCATTTTGTAAATGCACGAATAGTGCTAAGCTGGGAAGTTCCTCTGAGGTCTAACCTCTAGCTGCTCCCCACAGAAGAGTGCTGCGGCCAGTGGCCACCAGGGGTCGCCGAGCACCCAGCGCTGG
68	FXN promoter sequence (-700 to -600 from transcription start point (5'-3') (dna)	CGCCTGTAGTCCCAGTTACTCGGGAGGCTGAGGCAGGAGAATGGCGTGAAACCG GGGAGGCAGAGCTTGCACTGAGCCGAGAT CGCACCACTGCCCTCCAG
69	FXN promoter sequence (-660 to -610 from transcription start point (5'-3') (dna)	ATGGCGTGAAACCGGGAGGCAGAGCTTG CAGTGAGCCGAGAT CGCACCCAC
70	FXN promoter sequence (-1000 to -900 from transcription start point (5'-3') (dna)	TTTCTACAAAAAATACAAAACT AGCTGGGTGTGGTGGTGC ACACCTGTAGTCCCAGCTACTTGGAAGGCTGAAATGGGAAGACTGCTTGAGCCCCGGGAG
71	FXN promoter sequence (-990 to -940 from	AAATACAAAACT AGCTGGGTGTGGTGGTGC ACACCTGTAGTCCCAGCTA

	transcription start point (5'-3') (dna)	
72	FXN promoter target site L2	GAAACCGGGAGGCAGAGCTTGCAGTGAGCCGAGATCGCA
73	FXN promoter target site S2	GGGAGGCAGAGCTTGCAGTGAGCCGAG
74	FXN promoter target site L1	AGCTGGGTGTGGTGGTGCACACCTGTAG
75	FXN promoter target site S1	AGCTGGGTGTGGTGGTGC
76	Linker (prt)	TGEKP
77	Linker (prt)	TG (E/Q) (K/R) P
78	Linker (prt)	TGQKP
79	Linker (prt)	TG (G/S) (E/Q) (K/R) P
80	Linker (prt)	TGGERP
81	Linker (prt)	TGSERP
82	Linker (prt)	TGGQRP
83	Linker (prt)	TGSQRP
84	Linker (prt)	TGGEKP
85	Linker (prt)	TGSEKP
86	Linker (prt)	TGGQKP
87	Linker (prt)	TGSQKP
88	Linker (prt)	TG (G/S) ₀₋₂ (E/Q) (K/R) P
89	Linker (prt)	T (G/S) ₀₋₂ G (E/Q) (K/R) P
90	Linker (prt)	TG (G/S) ₃ (E/Q) (K/R) P
91	Linker (prt)	T (G/S) ₃ G (E/Q) (K/R) P
92	Linker (prt)	LRQKD (GGGS) ₁₋₄ QLVGT AERP
93	Linker (prt)	LRQKD (GGGS) ₁₋₄ QKP
94	Linker (prt)	LRQKDGGGSGGGSGGGSGQLVGT AERP
95	Linker (prt)	LRQKDGGGSGGGSGGGSGQKP
96	Linker (prt)	TGERP
97	ZF1 11-finger peptide binding 3'-AAG-5'(prt)	YACPVESCDRRFSQSGNLTRHIRIHTGSQKPFQCRICMRNFSQSGNRKRHIRTHTGE KPFACDICGRKFAQSGNLTRHTKIHTGSQKPFQCRICMRNFSQSGNLTRHIRTHTGE KPFACDICGRKFAQSGNRKRHTKIHLRQKDGGGSGGGSGGGSGQKPFQCRICMRN FSQSGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIHTGSQKPFQCRICMRN FSQSGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIHTGSQKPFQCRICMRN FSQSGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIH
98	ZF1a 11-finger peptide binding 3'-AAG-5'(prt)	YACPVESCDRRFSQSANLGRHIRIHTGSQKPFQCRICMRNFSNSANLSRHIRTHTGE KPFACDICGRKFAQSANLGRHTKIHTGSQKPFQCRICMRNFSNSANLSRHIRTHTGE KPFACDICGRKFAQSANLGRHTKIHLRQKDGGGSGGGSGGGSGQKPFQCRICMRN FSNSANLSRHIRTHTGEKPFACDICGRKFAQSANLGRHTKIHTGSQKPFQCRICMRN FSNSANLSRHIRTHTGEKPFACDICGRKFAQSANLGRHTKIHTGSQKPFQCRICMRN FSNSANLSRHIRTHTGEKPFACDICGRKFAQSANLGRHTKIH
99	ZF1b 11-finger peptide binding 3'-AAG-5'(prt)	YACPVESCDRRFSNSANLSRHIRIHTGSQKPFQCRICMRNFSQSQNLSRHIRTHTGE KPFACDICGRKFANSANLSRHTKIHTGSQKPFQCRICMRNFSQSQNLSRHIRTHTGE KPFACDICGRKFANSANLSRHTKIHLRQKDGGGSGGGSGGGSGQKPFQCRICMRN FSQSQNLSRHIRTHTGEKPFACDICGRKFANSANLSRHTKIHTGSQKPFQCRICMRN FSQSQNLSRHIRTHTGEKPFACDICGRKFANSANLSRHTKIHTGSQKPFQCRICMRN FSQSQNLSRHIRTHTGEKPFACDICGRKFANSANLSRHTKIH
100	ZF1-AAG-p65 (human)	MGPKKRRKVTGERPYACPVESCDRRFSQSGNLTRHIRIHTGSQKPFQCRICMRNFSQ SGNRKRHIRTHTGEKPFACDICGRKFAQSGNLTRHTKIHTGSQKPFQCRICMRNFSQ SGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIHLRQKDGGGSGGGSGGG GSQKPFQCRICMRNFSQSGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIHT GSQKPFQCRICMRNFSQSGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIHT GSQKPFQCRICMRNFSQSGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIHL RQKDGGGSPMEFYLPDTPDRHRIEEKRRTYETFKSIMKKSPFSGPTDPRPPPRRI AVPSRSSASVPKAPQYPFTSSLSTINYDEFPTMVFPSSGQISQASALAPAPPQVLP

		QAPAPAPAPAMVSALAQAAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQF DDEDLGALLGNSTDPVFTDLASVDNSEFQQLLNQGI PVAPHTTEPMLMEYPEAITR LVTGAQRPPDPAPAPLGA PLPGLNGLLSGDEDFSSIADMDFSALLSQISS
101	ZF1-AAG-VP64 (human)	MGPKKRRKVTGERPYACPVESCDRRFSQSGNLTRHIRIHTGSQKPFQCRICMRNFSQ SGNRKRHIRTHTGEKPFACDICGRKFAQSGNLTRHTKIHTGSQKPFQCRICMRNFSQ SGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIHLRQKDGGGSGGGGSGGG GSQKPFQCRICMRNFSQSGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIHT GSQKPFQCRICMRNFSQSGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIHT GSQKPFQCRICMRNFSQSGNLTRHIRTHTGEKPFACDICGRKFAQSGNRKRHTKIHL RQKDGGGSELQLTSDALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDLGS D ALDDFDLMDLSSQLSQ
102	ZF2 11-finger peptide binding 3'-GAA-5'(prt)	YACPVESCDRRFSRSDNLTQHIRIHTGSQKPFQCRICMRNFSRSDNRKQHIRTHTGE KPFACDICGRKFARSDNLTQHTKIHTGSQKPFQCRICMRNFSRSDNLTQHIRTHTGE KPFACDICGRKFARSDNRKQHTKIHLRQKDGGGSGGGGSGGGGSGQKPFQCRICMRN FSRSDNLTQHIRTHTGEKPFACDICGRKFARSDNRKQHTKIHTGSQKPFQCRICMRN FSRSDNLTQHIRTHTGEKPFACDICGRKFARSDNRKQHTKIHTGSQKPFQCRICMRN FSRSDNLTQHIRTHTGEKPFACDICGRKFARSDNRKQHTKIH
103	ZF2a 11-finger peptide binding 3'-GAA-5'(prt)	YACPVESCDRRFSRSSLTYHIRIHTGSQKPFQCRICMRNFSRSANRKQHIRTHTGE KPFACDICGRKFARSSNLSYHTKIHTGSQKPFQCRICMRNFSRSANLTQHIRTHTGE KPFACDICGRKFARSSNRIQHTKIHLRQKDGGGSGGGGSGGGGSGQKPFQCRICMRN FSRSANLSYHIRTHTGEKPFACDICGRKFARSSNRTQHTKIHTGSQKPFQCRICMRN FSRSANLSQHIRTHTGEKPFACDICGRKFARSSNRKYHTKIHTGSQKPFQCRICMRN FSRSANLSQHIRTHTGEKPFACDICGRKFARSSNRKQHTKIH
104	ZF2b 11-finger peptide binding 3'-GAA-5'(prt)	YACPVESCDRRFSRNDNRITHIRIHTGSQKPFQCRICMRNFSRSSLSTHIRTHTGE KPFACDICGRKFARNDNRITHTKIHTGSQKPFQCRICMRNFSRSSLSTHIRTHTGE KPFACDICGRKFARNDNRITHTKIHLRQKDGGGSGGGGSGGGGSGQKPFQCRICMRN FSRSSLSTHIRTHTGEKPFACDICGRKFARNDNRITHTKIHTGSQKPFQCRICMRN FSRSSLSTHIRTHTGEKPFACDICGRKFARNDNRITHTKIHTGSQKPFQCRICMRN FSRSSLSTHIRTHTGEKPFACDICGRKFARNDNRITHTKIH
105	ZF2a-GAA-p65 (human; prt)	MGPKKRRKVTGERPYACPVESCDRRFSRSSLTYHIRIHTGSQKPFQCRICMRNFSR SANRKQHIRTHTGEKPFACDICGRKFARSSNLSYQHTKIHTGSQKPFQCRICMRNFS RSANLTQHIRTHTGEKPFACDICGRKFARSSNRIQHTKIHLRQKDGGGSGGGGSGGG GGSQKPFQCRICMRNFSRSANLSYHIRTHTGEKPFACDICGRKFARSSNRTQHTKIH TGSQKPFQCRICMRNFSRSANLSQHIRTHTGEKPFACDICGRKFARSSNRKYHTKIH TGSQKPFQCRICMRNFSRSANLSQHIRTHTGEKPFACDICGRKFARSSNRKQHTKIH LRQKDGGGSPMEFYLPDTPDRHRIEEKRKRTYETFKSIMKKSFPFGPTDPRPPRR I AVPSRSSASVPKPAPQYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAPPQVL PQAPAPAPAPAMVSALAQAAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQ F DDEDLGALLGNSTDPVFTDLASVDNSEFQQLLNQGI PVAPHTTEPMLMEYPEAIT RLVTGAQRPPDPAPAPLGA PLPGLNGLLSGDEDFSSIADMDFSALLSQISS
106	ZF2a-GAA- VP64 (human; prt)	MGPKKRRKVTGERPYACPVESCDRRFSRSSLTYHIRIHTGSQKPFQCRICMRNFSR SANRKQHIRTHTGEKPFACDICGRKFARSSNLSYQHTKIHTGSQKPFQCRICMRNFS RSANLTQHIRTHTGEKPFACDICGRKFARSSNRIQHTKIHLRQKDGGGSGGGGSGGG GGSQKPFQCRICMRNFSRSANLSYHIRTHTGEKPFACDICGRKFARSSNRTQHTKIH TGSQKPFQCRICMRNFSRSANLSQHIRTHTGEKPFACDICGRKFARSSNRKYHTKIH TGSQKPFQCRICMRNFSRSANLSQHIRTHTGEKPFACDICGRKFARSSNRKQHTKIH LRQKDGGGSELQLTSDALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDLGS DAL DDFDLMDLSSQLSQ
107	SV40 NLS (prt)	PKKKRKV
108	Human KIAA2022 NLS (prt)	PKKKRKVT
109	mouse primase p58 NLS9 (prt)	RIRKKLR
110	Mouse N- terminal leader (prt)	MGRIRKKLR LAERP

111	Human N-terminal leader (prt)	MGPKKRRKVTGERP
112	Human N-terminal leader (prt)	MGPKKRRKVTLAERP
113	Double SV40 NLS (prt)	PKKKRKVPKKRKV
114	Double Human KIAA2022 NLS (prt)	PKKRRKVTPKKRRKVT
115	Double mouse primase p58 NLS9 (prt)	RIRKKLRRIRKKLR
116	ZF3 9-zinc finger peptide (prt)	YACPVESCDRRFSRSSLNLRHIRIHTGSQKPFQCRICMRNFSDSSVLTRHIRTHTGEKPFACDICGRKFAQSSHLTKHTKIHTGSQKPFQCRICMRNFSDNLTTTHIRTHHTGEKPFACDICGRKFARSSSTLTTHTKIHLRQKDG GGGSGGGSGGGSGQKPFQCRICMRNFSDSSHLTQHIRTHTGEKPFACDICGRKFARSSNLTEHTKIHTGSQKPFQCRICMRNFSDSSHLTQHIRTHTGEKPFACDICGRKFARSSHLTRHTKIH
117	ZF3-promoter-p65 (human) (prt)	MGPKKRRKVTGERPYACPVESCDRRFSRSSLNLRHIRIHTGSQKPFQCRICMRNFSDSSVLTRHIRTHTGEKPFACDICGRKFAQSSHLTKHTKIHTGSQKPFQCRICMRNFSDNLTTTHIRTHHTGEKPFACDICGRKFARSSSTLTTHTKIHLRQKDG GGGSGGGSGGGSGQKPFQCRICMRNFSDSSHLTQHIRTHTGEKPFACDICGRKFARSSNLTEHTKIHTGSQKPFQCRICMRNFSDSSHLTQHIRTHTGEKPFACDICGRKFARSSHLTRHTKIH LRLQKDG GGGSGGGSGGGSGPMEFYQLPDTDDRHRIEEKRKRTYETFKSIMKKS PFSGPTDPRPPPRRIAVPSRSSASVPKPAPQYPYFTSSSLSTINYDEFPTMVFP SGQISQASALAPAPPQVLPQAPAPAPAPAMVSALAQAAPAPVPVLAPGPPQAVAP PAPKPTQAGEGTLSEALLQLQFDDDLGALLGNSTDPAVFTDLASVDNSEFQQLLNQ GIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDFSSI ADMDFSALLSQISS
118	ZF3-promoter-VP64 (human) (prt)	MGPKKRRKVTGERPYACPVESCDRRFSRSSLNLRHIRIHTGSQKPFQCRICMRNFSDSSVLTRHIRTHTGEKPFACDICGRKFAQSSHLTKHTKIHTGSQKPFQCRICMRNFSDNLTTTHIRTHHTGEKPFACDICGRKFARSSSTLTTHTKIHLRQKDG GGGSGGGSGGGSGQKPFQCRICMRNFSDSSHLTQHIRTHTGEKPFACDICGRKFARSSNLTEHTKIHTGSQKPFQCRICMRNFSDSSHLTQHIRTHTGEKPFACDICGRKFARSSHLTRHTKIH LRLQKDG GGGSELQLTSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALD
119	ZF4 6-zinc finger peptide (prt)	FQCRICMRNFSDSSHLTKHIRTHTGEKPFACDICGRKFARSDHLTTHTKIHTGSQKPFQCRICMRNFSDHLTTTHIRTHHTGEKPFACDICGRKFARSDTLTRHTKIHTGSQKPFQCRICMRNFSDHLTTTHIRTHHTGEKPFACDICGRKFADSSHRKQHTKIH
120	ZF4-promoter-p65 (human) (prt)	MGPKKRRKVTGERPFQCRICMRNFSDSSHLTKHIRTHTGEKPFACDICGRKFARSDHLTTHTKIHTGSQKPFQCRICMRNFSDHLTTTHIRTHHTGEKPFACDICGRKFARSDTLTRHTKIHTGSQKPFQCRICMRNFSDHLTTTHIRTHHTGEKPFACDICGRKFADSSHRKQHTKIH LRLQKDG GGGSPMEFYQLPDTDDRHRIEEKRKRTYETFKSIMKKS PFSGPTDPRPPPRRIAVPSRSSASVPKPAPQYPYFTSSSLSTINYDEFPTMVFP SGQISQASALAPAPPQVLPQAPAPAPAPAMVSALAQAAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDDLGALLGNSTDPAVFTDLASVDNSEFQQLLNQ GIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDFSSI ADMDFSALLSQISS
121	ZF4-promoter-VP64 (human) (prt)	MGPKKRRKVTGERPFQCRICMRNFSDSSHLTKHIRTHTGEKPFACDICGRKFARSDHLTTHTKIHTGSQKPFQCRICMRNFSDHLTTTHIRTHHTGEKPFACDICGRKFARSDTLTRHTKIHTGSQKPFQCRICMRNFSDHLTTTHIRTHHTGEKPFACDICGRKFADSSHRKQHTKIH LRLQKDG GGGSELQLTSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALD
122	Long Human p65- RelA	PMEFYQLPDTDDRHRIEEKRKRTYETFKSIMKKS PFSGPTDPRPPPRRIAVPSRSSASVPKPAPQYPYFTSSSLSTINYDEFPTMVFP SGQISQASALAPAPPQVLPQAPAPAPAMVSALAQAAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDDLGALLGNSTDPAVFTDLASVDNSEFQQLLNQ GIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDFSSI ADMDFSALLSQISS

	activation domain (prt)	LGNSTDPVFTDLASVDNSEFQQLLNQGI PVAPHTTEPMLMEYPEAITRLVTGAQRP PDPAPAPPLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISS
123	Short Mouse p65-RelA activation domain (prt)	PLGTSGLPNGLSGDEDFSSIADMDFSALLSQISS
124	Herpes simplex virus VP16 domain (prt)	PADALDDFDLMDLGDGDSP
125	VP64 domain (prt)	ELQLTSDALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDLS SQLSQ
126	Linker (prt)	LRQKDGGGSGGGGSGGGGSQLVSS
127	Linker (prt)	LRQKDGGGSGGGGSS
128	Linker (prt)	LRQKDGGGSGGGGS
129	Linker (prt)	LRQKDGGGSGGGGS
130	Linker (prt)	LRQKDGGGS
131	BMP10 secretion signal (prt)	MGSLVLTLCALFCLAAYLVSG
132	Peptide cleavage sequence (prt)	RIRR
133	BMP10 secretion signal (dna)	ATGGGCTCTCTGGTCTGACACTGTGCGCTCTTTTCTGCCTGGCAGCTTACTTGGTT TCTGGC
134	Double-human-NLS (dna)	CCTAAGAAACGACGCAAGGTGACCCCTAAGAAACGACGCAAGGTGACC
135	Peptide cleavage sequence (dna)	CGAATCAGAAGG
136	Short Human p65- RelA activation domain (prt)	PLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISS
137	Double-SV40-NLS (dna)	CCGAAGAAAAACGTAAAGTGCCGAAGAAAAACGTAAAGTG
138	Double-mouse-NLS (dna)	AGGATCAGGAAGAAGCTGAGGAGGATCAGGAAGAAGCTGAGG
139	CMV promoter to Kozak ATG (dna)	acataaacttacggtaaatggcccgctggctgaccgcccacgacccccgcccattg acgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgt caatgggtggagtatattacggtaaaactgccacttggcagtacatcaagtgtatcat atgccaagtacgccccctattgacgtcaatgacggtaaatggcccgctggcattat gcccagtacatgaccttatgggactttcctacttggcagtacatctacgtattagtc atcgctattaccatgggtgatgcggttttggcagtacatcaatgggcgtggatagcgg tttgactcacggggatttccaagtctccacccattgacgtcaatgggagtttggtt tggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccaccatgg
140	CbH promoter	gttacataacttacggtaaatggcccgctggctgaccgcccacgacccccgccc ttgacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattga cgtcaatgggtggagtatattacggtaaaactgccacttggcagtacatcaagtgtat catatgccaaagtacgccccctattgacgtcaatgacggtaaatggcccgctggcat tatgccagtacatgaccttatgggactttcctacttggcagtacatctacgtatta gtcatcgctattaccatgggtcgaggtgagccccacgttctgcttcaactctccccatc tccccccccctccccaccccccaattttgtattttatttttttaattattttgtgca gcgatgggggcggggggggggggggggcgcgccaggcggggcggggcggggcgag gggcggggccaccatgg
141	CbH promoter+intron	acgcgtcggttacataacttacggtaaatggcccgctggctgaccgcccacgaccc cgcgccattgacgtcaatagtaacgccaatagggactttccattgacgtcaatgggt ggagtattttacggtaaaactgccacttggcagtacatcaagtgtatcatatgccaaag

		tacgccccctattgacgtcaatgacggtaaattggcccgctggcattgtgcccagta catgaccttatgggactttcctacttggcagtagacatctacgtattagtcacgctat taccatgggtcgaggtgagccccacgttctgcttactctccccatctccccccccctc cccccccccaattttgtattttatttttttaattattttgtgcagcgatggggggc ggggggggggggggggggcgcgccaggcgggcgggcgggcgagggcgggggcg ggcgagggcgagaggtgcgggcgccagccaatcagagcgggcgctccgaaagtttcc ttttatggcgaggcgggcgggcgggcgccctataaaaaagcgaagcgcgcgggcg gggagtcgctgcgacgctgccttcgccccgtgccccgctccgcccgcgctcgcgcc gcccgcggcggtctgactgaccgcgttactcccacaggtgagcgggcgggacggcc cttctcctccgggctgtaattagctgagcaagaggttaagggtttaagggatggttg ttggtgggggtattaatgtttaattacctggagcacctgcctgaaatcacttttttc agggtggggccaccatgg
142	Synthetic rat 1.8 kb eno2 (pNSE) promoter-enhancer design – with NheI flanking restriction sites (gctagc) – bold	gctagcctcctctgctcgcccaatccttccaacccccctatggtggtatggctgacac agaaaatgtctgctcctgtatgggacatttgccccctcttctccaaatataagacagg atgaggcctagcttttctgctgctccaaagttttaaaagaacacattgcacggcattta gggactctaaaggtggagggaatgaggaattgcatcatgccaaggctggtcct catccatcactgcttccagggccagagtggttccaggaggtattcttaaaaagga agcccgatctgtagctaacactcagagccccattttcctgcgttaacccccctcccgacc tcatatacaggagtaacatgatcagtgacctgggggagctggccaaactgcgggacc tgcccaagctgagggccttgggtgctgctggacaacccccgtgcccgatgagactgact accgccaggagggcctggtgcagatggcacacctagagcgcttagacaaaagagtact atgaggacgaggaccgggcagaagctgaggagatccgacagaggctgaaggaggaaac aggagcaagaactcgacccggaccaagacatggaaccgtacctcccgccaaacttagt ggcacctctagcctgcagggacagtaaaggatggcaggaaggcagcccccgagg tcaaaggctgggcacgcgggaggagagggccagagtcagaggctgcgggtatctcaga tatgaaggaaagatgagagaggctcaggaagaggttaagaaaagacacaagagaccag agaagggagaagaattagagaggggaggcagaggaccgctgtctctacagacatagct ggtagagactgggaggaagggtgaaccctgagcgcatgaagggaaggaggtggctg gtggtatatggaggatgtagctgggcccagggaagatcctgcactaaaaatctgaa gctaaaaataacaggacacgggggtggagaggcgaaaggaggggcagagtgaggcagag agactgagaggcctggggatgtgggcattccggtagggcacacagttcacttgtctt ctctttttccaggaggccaaagatgctgacctcaagaactcataataccccagtggg gaccaccgcattcatagcctgttacaagaagtggagatggttcttttcttccag actggaaatccgtttacatcccagggtcagggttctgtggtgtcatctctgtgtggc ttgttctgtgggctacctaaagtccaaagcacagctctcaagcagatccgaggcga ctaagatgctagtaggggttgtctggagagaagagccgaggaggtgggctgtgatgg atcagttcagctttcaataaaaaggcgtttttatattctgtgtcgagttcgtgaac ccctgtggtgggcttctccatctgtctgggttagtacctgccactatactggaataa ggggacgcctgttccctcgagttggctggacaagggttatgagcatccgtgtactta tgggggttgcagcttgggtcctggatcgcccgggcccttccccaccggttcggttcc ccaccaccaccgcgctcgtagctgcgtctccgctgcagctcttgactcatcgggg cccccggggtcacatgcgctcgctcggtctataggcgccgccccctgcccacCCCC CGCCCGCGCTGGGAGCCGCGAGCCGCCGCTCTGCTCTCTGCGCCGCCCGCGCT CACCACCGCCACCGCCACCGGCTGAGTCTGCAGTCgctagc
143	Synthetic mouse 1.8 kb hsp90ab1 promoter-enhancer design – with NheI flanking restriction sites (gctagc) – bold	gctagcaacacccctagggccttctgagcaatcctacccagtgctcctcatatattg atttctttatgggttcacacacacacacacacacacacacacacacacacacacac acacacagaattaaggagaggctaacagacagtgaggatgggatgataacagacga agtagacagaggcaaggagaaagcaactactgtttaacaatgaatgcacattagaca gactgcaggcaagcaccgggaacaaagggtgtgggcgggtggtgtggggacacaagcca gcatgagctaagatagcagagcactgagtgccatcctctactggagggtcatcag tccaacaagcttccagatgcagccttggaaaaaggcaaggctagattgccagctgaa ggacatggcaggccacctttagaacagagggcactggcacaacttggtttctgtgctc ctggaactgggccaacctgtgaccagcaccttcatgcggatgctagaactccagct tctctgaaaagactgggacctgctcctcttaggtccaaagagctgcatgcagtagg gaagaggctagagaagcgaaaccagcttgagaaacagcttgtgctcacataggagg gcgcacgtacccgcgcgtgtgtacgtgggagaccggggaggctgaggggtgggggag tgttctacccagtagcgcaagctgatagctcggttctctgttactagaagggtgtcc gcagtcactcacccccacagccccgtgcctgtgaccgatccaggtcagctatccc tccctctgcgctccactccccactgttatgtgggcctcttagggccacgcgtggag ggtcggttaacacctggccacggtaggcagacttggggaaaatttcttccagggtta

		agatcaaggtaggggaaaaaaaaaaaaaaaaaagccaccagccaagcggcgacga agacactgccccgcgcagcaggggaggtggagcctaggggggaggggtggagacc gccgagacaggcctagaaactgctggaagaaatcgacagcaccacgctgctgaccc tccgccgcaggccgccaagagtcctaccagcccaggcccgctgcccctcccctcgg ggaaagcggctcccagcctgaagctgtgctgtaccgggaggggtggggatgggggaa tcgggggcctccttaaagttggacaaggaatttatcatccttttctcttgatgtgcg atttgtagggaacattctagtaagatcgggtctggaaatggcagccgagttggccac ctccattctctttcagtcctcctgagttctggactcttgggggggtgggggggtggaag cgctaccttgagttttctgaggcagtcctgtaggggtattcgcccgagatacatccc taattgcatatgcatgctccctgctcatcttgaggggggacatgtcctactcctgca gaaatgggggatgtgcaaaacgatattgaattggccttgactcaggaaccaggccccg gggctcccgctcctccccgccccctccacgatctgctacatgacgtcaaggtgggcg ggcggcggcaggtgctggtggccgcagccactcctttaaggcggagggatccaagggc ggggcccgggctgtgcttcgccttatatagggcggtcggggcggttcgggagctCTC TTGAGTCACCCCGCGCAGCCTAGGCTTGCCGTGCGAGTCGGACTTGGTCCGGGCC ACCACCCTGCTCTGTACTACTACTCGGCTTTCCCGTCAAG gctagc
144	Synthetic human 1.8 kb hsp90ab1 promoter-enhancer design – with NheI flanking restriction sites (gctagc) – bold	gctagc ctgcttcagcctcccaagtagcggggactatagggcgctaccacgcccgg ctaatTTTTgtatTTTTtagtagagacaaggTTTcaccagagtgggccaggctggctctt gaactcctgacctgggtgatccgctggactcggcctcccaaagtgtggaattacagg catgagccaccgcgcccagccccctTTTTTTTTTTTTTTTTTTTTTtagacggagtt tctctTTTTgttggccaggctggagtgcattggcatgggtctcagctcactgcagcctc tgctccagagttcaagcaatcttctcctgcctcagcctcccaagtagctgggattaca ggcatgtgccaccacgcccagctaattTTTTtagtagagacgaggtTTTTcaccatgtt ggccaggctgggttgtaactactgacctcgggtgatccaccacctcggccttccaa agtgcctgggattacaggtatgagccactgtgccagcagagaaagtggatttctat gctgggtgtaacatttgagattttccatttagaaaagtcttcaggaagccgtcgagtc ttgacacagaaaggatccagctcaggggaagcccatccacaggagtgattcctgctt tggatttaggtcctaggggtgagatcacatccctgctccatttctggagcagaaagg gagctagtgcagcaaggactctgagtagacgggccttgggagtcactgaagccactt cctgttgacagaatgtgcacactgactttcatccgctcacaggagagtgaagattt gtgaatgctctctgtgaactggggagcagattTTTTaatgagctctgcttagaaggag aggacacgttaagggtacagctcccactgggttgcggttttggtcaggaatagagc acccaacgcccccttcttatagacttctccagccataggctttttgcagaccataat gaaagggagaaaggatatccaatcgggatgggtgatggcagggcatggcgaagtggaca gggacaaagaagaggaaagcaagggacacccttTgttcttTggtattcaacacatga acaaaacttgtaacgaatcgtagactgggttaacaacagggtgggtgagggaggacg atgtgggtggctgtggggaaactcttcccaagcatcccagggggtgggcccaggaagg gagcagatggacatctacagggcacagagcaggagtggtattctctcctggagggt gttgtgggtccagcaggctttcggtatgagcccttggaaggcaaggacagcttgcca actggagggcaggcaaaacggctgagccaggccagacacggcagctcactggccagg ggcatggcctaggggtctctgggtgccaaggctgaaggggaaacattggccagccgct ctgtggcgcatcgctcagcggtcctccaagagctccggctgcctgcactggttcc cagagactccctccttcccagggtccaaatgggtgcagggagcgaagtgggcggaaaaa aagcgaaccagcttgagaaagggttgacgtgcctgcgtagggagggcgcatgtccc cgtgctccgtgtacgtggcgccgcaggggtAGAGGGGGGTCCCCCGCAGGTAC TCCACTCTCAGTCTGCAAAAGTGTACGCCCGCAGAGCCGCCCCAGGTGCCTGGGTGT TGTGTGATTGACGCGGGGAAGGAGGGGTGAGCCGATCCCTCCCCAACCTCCATCCC ATCCCTGAGGATTGGGCTGGTACCCGCGTCTCTCGGACAG gctagc
145	Active delivery construct (dna): pCbH-SS-2xNLS-ZFP(ZF2a)-p65-AD-STOP-WPRE-polyA	acgcgtcgttacataacttacggtaaatggccgcctggctgaccgcccacgaccc ccgcccattgacgtcaatagtaacgcccaatagggaactttccattgacgtcaatgggt ggagtatttacggtaaaactgccacttgccagtagtacatcaagtgtatcatatgccaag tacgccccctattgacgtcaatgacggtaaatggccgcctggcattgtgcccagta catgaccttatgggactttcctacttggcagtagacatctacgtattagtcacgtat taccatggtcgaggtgagccccagttctgcttactctccccatctccccccctc cccaccccccaattttgtatTTTatTTTatTTTatTTTatTTTatTTTgtgcagcgatgggggc gggggggggggggggggcgcgcgccaggcgggggcgggggcgagggggcgggggcg ggcgagggcgagaggtgcgggcggcagccaatcagagcggcgcgctccgaaagtTtcc TTTtatggcgagggcgggcgggcgggcgccctataaaaagcgaagcgcgcgggcg gggagtcgctgcgacgctgccttcgccccgtgccccgctccgcccgcgctcgcgcc gccgcccccggtctgactgaccgcgttactcccacaggtgagcgggcgggacggcc

		cttctcctccgggctgtaattagctgagcaagaggttaaggggtttaagggatggttggtggtggtggttattaatgtttaattacctggagcacctgcctgaaatcactttttttcagggttgGccaccATGGGCTCTCTGGTCCTGACACTGTGCGCTCTTTTCTGCCTGGCAGCTTACTTGGTTTCTGGCCGAATCAGAAGGGATATGGGGCTAAGAAACGACGCAAAGGTGACCCCTAAGAAACGACGCAAGGTGACCGGCGAAAGGCCATACGCTTGTCTGTGGAGTCATGTGATAGGAGGTTTCAGTCGGTCATCTAACCTGACCTACCACATCAGAATCCACACAGGCTCCCAGAAGCCCTTCCAGTGCCGGATCTGTATGAGAACTTTTCTAGGAGCGCAATCGCAAGCAGCACATCAGGACCCACACAGGCGAGAAGCCCTTTCGCCTGCGATATCTGTGGCCGGAAGTTTGCCAGAAGCTCCAATCTGAGCTATCACACCAAGATTCATACAGGCTCCCAGAAGCCATTCCAGTGCCAGGATCTGTATGCGCAACTTTTCCCCGTCTGCCAATCTGACCCAGCACATCCGCACTCATACAGGCGAGAAGCCCTTCGCCTGCGACATCTGTGGCAGAAAGTTTGCCAGGTCTAGCAACCGCATCCAGCACACAAAGATCCACCTGCGGCAGAGGATGGAGGAGGAGGAGCGGCGGAGGAGGCTCCGGAGGAGGAGGCTCTCAGAAGCCCTTTTCAGTGTAGAATTTGTATGAGAACTTTAGCAGGTCCGCAATCTGAGCTACCACATCAGAACTCATACCGGCGAGAAGCCATTTCGCCTGTGATATCTGTGGCCGCAAGTTTGCCCGGTCTCTAATAGAACTCAGCATACAAAGATCCACACCGGCTCTCAGAAACCTTTTCAGTGTAGAATCTGCATGCGCAACTTTTCTCGGAGCGCCAATCTGAGCCAGCACATCAGAACTCACACCGGAGAAAAACCTTTCGCCTGTGACATTTGTGGCAGAAATTTGCCAGGAGCTCCAACCGCAAGTATCACACCAAGATTACACACGGCTCCCAGAAACCTTTTCAGTGCGAGAATCTGCATGAGGAACTTTTCCAGGTCTGCAATCTGTCTCAGCACATCCGCACTCACACCGGAGAAAAGCCCTTCGCCTGTGATATCTGCGGCCGGAAGTTTGCCAGAAGTTCAAACAGAAAGCAGCATACAAAGATTACCTGAGGCAGAAGGACGGAGGAGGATCCCCCATGGAGTTTCAGTACCTGCCTGATACAGACGATCGGCACAGAATCGAGGAGAAGAGGAAGCGCACCTATGAGACATTCAGTCCATCATGAAGAAGAGCCCTTTTCCGGACCAACCGACCTCGGCCCCCTCCAAGGAGAATCGCCGTGCCTTCTAGAAGCTCCGCCAGCGTGCCAAAGCCAGCACCTCAGCCATACCCCTTCACCTCTAGCCTGTCCACAATCAACTATGATGAGTTCCCAACCATGGTGTTTCCCTCTGGACAGATCTCTCAGGCCAGCGCCCTGGCACCAGCCCCCCTCAGGTGCTGCCACAGGCACCTGCACCAGCACCTGCCATGGTGAGCGCCCTGGCCCAGGCCCCAGCCCCGTGCCTGTGCTGGCCCCCGGCCACCCAGGCCGTGGCCCCCTCCAGCCCCAAAGCCAACCCAGGCAGGAGAGGGCACACTGTCTGAGGCCCTGTGCGAGTGCAGTTGCGACGATGAGGACCTGGGCGCCCTGCTGGGCAATAGCACCGATCCCGCGTGTTCACAGACCTGGCCTCCGTGGATAACTCTGAGTTTCAGCAGCTGCTGAATCAGGGCATCCCAGTGGCCCCCACACCACAGAGCCTATGCTGATGGAGTACCCAGAGGCCATCACCCGGCTGGTGACAGGCGCCAGCGGCCCTGACCCTGCACCAGCACCACTGGGAGCACCTGGCCTGCCAAACGGCCTGCTGAGCGGCGACGAGGATTTCTCTCTATCGCCGACATGGACTTCAGCGCCCTGCTGTCCAGATCAGCTCCTGAgcgggccgcaaGAATTCgata tcaagcttatcgataatcaacctctggattacaaaatgttgaaagattgactggta ttcttaactatgttgctccttttacgctatgtggatacgtgctttaatgcctttgt atcatgctattgttcccgatgtggctttcattttctcctccttgataaatcctgggtgctgtctctttatgaggagttgtggccggtgtgcaggcaacgtggcggtgtgtgca ctgtgtttgtgacgcaacccccactgggtggggcattgccaccacctgtcagctcc tttccgggactttcgctttccccctccctattgccacggcggaactcatcgccgcctgccttgcccgctgctggacaggggctcggtgttgggcaactgacaattccgtgggtgtgtcggggaaATCATcgctcctttccTtggctgctcgctgtgttgccacctggattc tgcgcgggaacgtccttctgctacgtcccttcggccctcaatccagcggaccttccttcccgcggcctgctgcggcctctgcggcctcttcgcgtcttcgccttcgcctcagacgagtcggatctccttttgggcccgcctccccgcacgataccgtcgacctcgagggggggcccgggtaccCGGGAATCAATTCATCCTCAGGTGCAGGCTGCCATCAGAAGGTGGTGGCTGGTGTGGCCAATGCCCTGGCTCACAAATACCACTGAGATCTTTTTCCCTCTGCCAAAAATTATGGGGACATCATGAAGCCCTTGAGCATCTGACTTCTGGCTAATAAAGGAAATTTATTTTCATTGCAATAG
146	Full AAV Vector (dna): for active delivery: pCbH-SS-2xNLS-ZFP(ZF2a)-p65-AD-STOP-WPRE-polyA	CCCGGGCAAAGCCCGGGCGTCCGGGCGACCTTTGGTCCGCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCTTGTAGTTAATGATTAACCCGCCATGCTACTTATCTACTCGACATTGATTATTGAgctagcacgcgtcgt tacataacttacggtaaatggcccgcctggctgaccgcccacgacccccgcccattgacgtcaatagtaacgccaatagggactttccattgacgtcaatgggtggagtattt acggtaaactgcccacttggcagtagcatcaagtgtatcatatgccaaagtacgcccc tattgacgtcaatgacggtaaatggcccgcctggcattgtgcccagtagcatgacctt atgggactttcctacttggcagtagcatctacgtattagtagcatcgctattaccatgggt

		cgaggtgagccccagttctgtttcactctcccatctccccccctccccaccccc aattttgtatttatttattttttaattattttgtgcagcgatgggggcggggggggg gggggggcgcgccagggcgggcgggcgggcgagggcggggcgggcgagggcg gagaggtgcggcgccagccaatcagagcgcgctccgaaagtctttttatggc gaggcgggcgggcgggcgccctataaaaagcgaagcgcgcgggcgggagtcgc tgcgacgctgccttcgccccgtgccccgctccgcccgcctcgcgccgccccccc ggctctgactgacgcgttactcccacaggtgagcgggcgggacggcccttctcctc cgggctgtaattagctgagcaagaggttaagggtttaagggatgggttggtgggg tattaatgtttaattacctggagcacctgcctgaaatcactttttttcaggttggGc caccATGGGCTCTCTGGTCTGACACTGTGCGCTCTTTTCTGCCTGGCAGCTTACTT GGTTTCTGGCCGAATCAGAAGGGATATGGGGCCTAAGAAACGACGCAAGGTGACCCC TAAGAAACGACGCAAGGTGACCGGCGAAAGGCCATACGCTTGTCTGTGGAGTCATG TGATAGGAGGTTCACTCGGTCACTAACCTGACCTACCACATCAGAATCCACACAGG CTCCCAGAAGCCCTTCCAGTGCCGGATCTGTATGAGAACTTTTCTAGGAGCGCCAA TCGCAAGCAGCACATCAGGACCCACACAGGCGAGAAGCCCTTTCGCCTGCGATATCTG TGGCCGGAAGTTTGCCAGAAGCTCCAATCTGAGCTATCACACCAAGATTTCATACAGG CTCCCAGAAGCCATTCCAGTGACAGGATCTGTATGCGCAACTTTTCCCGGTCTGCCAA TCTGACCCAGCACATCCGCACTCATACAGGCGAGAAGCCCTTTCGCCTGCGACATCTG TGGCAGAAAGTTTGCCAGGTCTAGCAACCGCATCCAGCACACAAAGATCCACCTGCG GCAGAAGGATGGAGGAGGAGGAGGAGCGGCGGAGGAGGCTCCGGAGGAGGAGGCTCTCA GAAGCCTTTTTCAGTGATAGAATTTGTATGAGAACTTTAGCAGGTCCGCCAATCTGAG CTACCACATCAGAACTCATACCGGCGAGAAGCCATTTCGCCTGTGATATCTGTGGCCG CAAGTTTGCCCGGTCTCTAATAGAACTCAGCATACAAAGATCCACACCGGCTCTCA GAAACCTTTTTCAGTGATAGAATCTGCATGCGCAACTTTTCTCGGAGCGCCAATCTGAG CCAGCACATCAGAACTCACACCGGAGAAAAACCTTTTCGCCTGTGACATTTGTGGCAG AAAATTTGCCAGGAGCTCCAACCGCAAGTATCACACCAAGATTTCACACCGGCTCCCA GAAACCTTTTTCAGTGACAGAATCTGCATGAGGAACTTTTCCAGGTCTGCCAATCTGTC TCAGCACATCCGCACTCACACCGGAGAAAAGCCCTTTCGCCTGTGATATCTGCGGCCG GAAGTTTGCCAGAAGTTCAAACAGAAAGCAGCATACAAAGATTTCACCTGAGGCAGAA GGACGGAGGAGGATCCCCATGGAGTTTTCAGTACCTGCCTGATACAGACGATCGGCA CAGAATCGAGGAGAAGAGGAAGCGACCTATGAGACATTCAGTCCATCATGAAGAA GAGCCCTTTTTCGGACCAACCGACCCTCGGCCCCCTCCAAGGAGAATCGCCGTGCC TTCTAGAAGCTCCGCCAGCGTGCCAAAGCCAGCACCTCAGCCATAACCCCTTCACCTC TAGCCTGTCCACAATCAACTATGATGAGTTCCCAACCATGGTGTTTCCCTCTGGACA GATCTCTCAGGCCAGCGCCCTGGCACCAGCCCCCCCCCTCAGGTGCTGCCACAGGCACC TGCACCAGCACACGACCTGCCATGGTGAGCGCCCTGGCCAGGCCCCAGCCCCCGT GCCTGTGCTGGCCCCCGGCCACCCAGGCCGTGGCCCCCTCAGCCCCAAAGCCAAC CCAGGCAGGAGAGGGCACACTGTCTGAGGCCCTGCTGCAGCTGCAGTTCGACGATGA GGACCTGGGCGCCCTGCTGGGCAATAGCACCGATCCCGCCGTGTTACAGACCTGGC CTCCGTGGATAACTCTGAGTTTTCAGCAGCTGCTGAATCAGGGCATCCCAGTGGCCCC CCACACCACAGAGCCTATGCTGATGGAGTACCCAGAGGCCATCACCCGGCTGGTGAC AGGCGCCCAGCGGCCCCCTGACCCTGCACCAGCACCACTGGGAGCACCTGGCCTGCC AAACGGCCTGCTGAGCGGCGACGAGGATTTCTCCTCTATCGCCGACATGGACTTCAG CGCCCTGCTGTCCCAGATCAGCTCCTGAGcgggccgcaaGAATTCgatatcaagctta tcgataatcaacctctggattacaaaatttgtgaaagattgactggtattcttaact atggttgctccttttacgctatgtggatacgctgctttaatgcctttgtatcatgcta ttgcttcccgatggctttcattttctcctccttgataaatcctggttgctgtctc tttatgaggagttgtggcccggtgtcaggcaacgtggcggtggtgtgactgtgtttg ctgacgcaacccccactgggttggggcattgccaccacctgtcagctcctttccggga ctttcgctttccccctccctattggcacggcggaactcatcgccgctgcttgccc gctgctggacaggggctcggtgttgggcaactgacaattccgtggtgttgctgggga aATCATcgctcctttccTtggtgctgcgcctgtgttgccacctggattctgcgcggga cgctccttctgctacgtcccttcggccctcaatccagcggaaccttcttcccgcgcc tgetgcggctctgcggcctcttccgctcttgccttcgcccctcagacgagtcgga tctccctttgggcccgcctccccgcacgataccgtcgacctcgaggggggggcccggg accCGGAATCAATTCACCTCCTCAGGTGCAGGCTGCCTATCAGAAGGTGGTGGCTGG TGTGGCCAATGCCCTGGCTCACAAATACCACTGAGATCTTTTCCCTCTGCCAAAAA TTATGGGGACATCATGAAGCCCTTGAGCATCTGACTTCTGGCTAATAAAGGAAATT TATTTTCATTGCAATAGTGTGTTGGAATTTTTTGTGTCTCTCACTCGGAAGGACATA TGGGAGGGCAAATCATTTAAACATCAGAATGAGTATTTGGTTTAGAGTTTGGCAAC
--	--	--

		ATATGCCCATATGCTGGCTGCCATGAACAAAGGTTGGCTATAAAGAGGTCATCAGTA TATGAAACAGCCCCCTGCTGTCCATTCTTATTCCATAGAAAAGCCTTGACTTGAGG TTAGATTTTTTTTATATTTTGTGTTATTTTTTTCTTTAACATCCCTAAAAAT TTCCTTACATGTTTTACTAGCCAGATTTTTCTCTCTCTGACTACTCCAGTCAT AGCTGTCCCTCTTCTCTTATGGAGATCCCTCGACCTGCAGCCCAAGCTGTAGATAAG TAGCATGGCGGGTTAATCATTAACTACAAGGAACCCCTAGTGATGGAGTTGGCCACT CCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGC CCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGGGCCTCAGTG AGCGAGCGAGCGCGCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT TGCGTATTGGGCGCTCTTCCGCTTCTCTGCTCACTGACTCGCTGCGCTCGGTCGTTT GGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT CAGGGGATAACGCAGGAAAGAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATC ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC AGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCCTGCCGCTTA CCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC GCTGTAGGTATCTCAGTTCGGTGAGGTGCTTCCGCTCCAAGCTGGGCTGTGTGCACG AACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACCTATCGTCTTGAGTCCA ACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACGGCT ACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA AAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGA TCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGG TCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTTAAATTAATAATGAAGTT TTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGAC TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTG CAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGC CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCCTCCATCCAGT CTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCA ACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGCTGATGGCTT CATTGAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCA AAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAGTAAGTTGGCCGCAG TGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCG TAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTA TGCGGCGACCGAGTTGCTCTTGCCCGCGTCAATACGGGATAATACCGCGCCACATA GCAGAACTTTAAAGTGCTCATCATTGGAACCGTTCTTCGGGGCGAAACTCTCAA GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGAT CTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAA ATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATACTCATACTCTTCC TTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG TGCCACCTGACGTCTAAGAAACCATATTATCATGACATTAACTATAAAAAATAGGC GTATCACGAGGCCCTTTCGTCTCGCGCTTTCGGTGATGACGGTGAAAACCTCTGAC ACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGAC AAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACCTATG CGGCATCAGAGCAGATTGTAAGTACTGAGAGTGACCATATGCGGTGTGAAATACCGCACA GATGCGTAAGGAGAAAATACCGCATCAGGCGATTCCAACATCCAATAAATCATACAG GCAAGGCAAAGAATTAGCAAAATTAAGCAATAAAGCCTCAGAGCATAAAGCTAAATC GGTTGTACCAAAAACATTATGACCCTGTAATACTTTTTCGGGGAGAAGCCTTTATTTT AACGCAAGGATAAAAATTTTTAGAACCTCATATATTTTAAATGCAATGCCTGAGTA ATGTGTAGGTAAAGATTCAAACGGGTGAGAAAGGCCGAGACAGTCAAATCACCATC AATATGATATTCAACCGTTCTAGCTGATAAATTCATGCCGGAGAGGGTAGCTATTTT TGAGAGGTCTCTACAAAGGCTATCAGGTCAATGCCTGAGAGTCTGGAGCAACAAGA GAATCGATGAACGGTAATCGTAAACTAGCATGTCAATCATATGTACCCCGGTTGAT AATCAGAAAAGCCCCAAAAACAGGAAGATTGTATAAGCAAAATATTTAAATTTGAAGC GTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTTTAAC CAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAAGATAGACCGAGATAGGG TTGAGTGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAAC
--	--	--

		GTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCC TAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCCTAAAGGG AGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAAGGAAGGG AAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGC GTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTACTATGTTGTC TTTGACGAGCACGTATAACGTGCTTTCTCGTTAGAATCAGAGCGGGAGCTAAACAG GAGGCCGATTAAAGGGATTTTAGACAGGAACGGTACGCCAGAATCCTGAGAAAGTGT TTTATAATCAGTGAGGCCACCGAGTAAAAGAGTCTGTCCATCACGCAAATTAACCGT TGTCGCAATACTTCTTTGATTAGTAATAACATCACTTGCCGTGAGTAGAAGAACTCAA ACTATCGGCCTTGCTGGTAATATCCAGAACAAATATTACCGCCAGCCATTGCAACGGA ATCGCCATTGCGCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCT CTTCGCTATTACGCCAGCTGCGCGCTCGCTCGCTCACTGAGGCCG
147	ZF2a-p65AD, CMV promoter (dna)	ccaatgatagatctacataacttacggtaaattggccgcctggctgaccgcccacg acccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaataggga ctttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtac atcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaattggc cgcctggcattatgcccagtcacatgaccttatgggactttcctacttggcagtacat tctacgtattagtcacgtcattaccatgggtgatgcggttttggcagtcacatcaatg ggcgtggatagcgggtttgactcacggggatttccaagtctccaccccattgacgtca atgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtacaacact ccgcccacATGGGCAGAATCAGAAAAAGCTGAGACTCGTGAAAGACCCCTACGCCT GTCCTGTTGAGAGCTGCGACAGAAGATTGAGCAGAAGCTCTAACCTGACATACCACA TCCGCATCCACACCGGCAGCCAGAAACCTTTTCAGTGCAAGATCTGTATGCGGAACCT TCTCCAGATCCGCCAACAGGAAACAGCACATCAGAACACACACCGGTGAGAAAGCCTT TTGCCTGCGACATCTGCGGCAGAAAGTTGCGCCAGATCATCCAATCTGTCTTATCACA CAAAGATCCACACCGGCTCCCAAAAGCCCTTCCAGTGCCGGATATGTATGCGTAACT TCAGCCGGAGCGCAATCTGACCCAGCATATTGCGACACACACCGGCGAAAAAGCCTT TCGCCTGCGACATCTGCGGCCGGAATTCGCTCGGAGCAGCAACAGAATCCAGCACA CCAAGATCCACCTGAGACAGAAAGACGGCGGGCGGGCTCAGGCGGCGGCGGCGAGCG GAGGAGGCGGCAGCCAAAAGCCTTTCCAGTGCAAGATCTGCATGAGAACTTCAGCC GGTCTGCCAACCTGAGCTACCATATTGCGACACACACCGGCGAGAAGCCTTTTCGCCT GCGATATCTGCGGCAGAAAGTTGCGCAGATCCAGCAACCGGACGCGACACACAAAGA TCCACACAGGATCTCAGAAGCCTTTCCAATGTAGAATCTGTATGAGAACTTCAGCA GAAGCGCCAACCTCTCTCAGCACATCAGAACCCACACCGGCGAAAAAGCCTTTTCGCCT GCGACATCTGCGGACGGAAGTTGCGCCAGAAGCAGCAATAGGAAGTACCACACCAAGA TCCACACAGGAAGCCAGAAACCTTTTCAGTGCCGGATCTGCATGCGGAACCTTCAGCA GAAGCGCCAACCTGAGCCAGCACATCAGAACCCACACAGGCGAGAAGCCTTTTGCCT GTGACATATGCGGCAGAAAGTTTGCCAGAAGCAGCAACAGAAAGCAGCACACCAAGA TCCATCTGCGGCAGAAAGGATGGCGGAGGCAGCCCTATGGAATTCCAATACCTGCCCCG ACACCGATGATCGGCACAGAATCGAGGAAAAAAGAAAGCGGACCTACGAAACCTTCA AGAGCATCATGAAGAAGTCTCCATTGAGCGGCCCCACCGATCCTAGACCTCCCCCAA GAAGAATCGCCGTGCCATCTAGAAGCTCTGCCAGCGTGCCCTAAACCTGCCCCCTCAAC CTTACCCCTTTTACAAGTAGCCTGTCCACCATCAACTACGACGAGTTTCTTACCATGG TGTTCCCTTCTGGACAGATCAGCCAAGCTTCTGCCCTGGCCCCCGCCCCCTCCTCAGG TGCTGCCCCAGGCCCCCTGCCCCCGCCCCCGCCCCGCTATGGTCAGCGCTCTGGCCCC AGGCCCCCTGCCCCAGTGCCTGTGCTGGCCCCCTGGCCCTCCTCAGGCCGTGGCCCCCTC CAGCCCCCTAAGCCCACCCAGGCAGGCGAAGGCACCCCTGAGCGAGGCCCTGCTTCAGC TGCAGTTCGACGACGAGGACCTGGGAGCCTTGCTGGGCAATAGCACCAGCCCCGCCG TGTTACCCGACCTGGCTTCCGTGGACAACAGCGAGTTCCAGCAGCTGCTGAACCAGG GCATCCCCGTGGCCCCCACACAACAGAGCCTATGCTGATGGAATACCCCGAGGCCA TTACCAGACTGGTGACCGGAGCTCAGAGGCCACCAGATCCTGCACCCGCCCCCTCTGG GCGCTCCCGGCTGCCAAATGGCCTGCTGAGCGGTGATGAGGATTTTCAGCAGCATCG CCGACATGGACTTCAGCGCCCTGCTGAGCCAGATCAGCTCCTGA
148	ZF2a-p65AD, CBH promoter (dna)	ccaatgatagatctgttacataacttacggtaaattggccgcctggctgaccgccc acgacccccgcccattgacgtcaatagtaacgccaatagggaactttccattgacgtc aatgggtggagtatttacggtaaactgcccacttggcagtcacatcaagtgtatcata tgccaagtacgccccctattgacgtcaatgacggtaaattggccgcctggcattgtg cccagtcacatgaccttatgggactttcctacttggcagtcacatctacgtattagtc tcgctattaccatgggtcgaggtgagccccacgttctgcttcaactctccccatctccc ccccctccccaccccccaattttgtattttattttttaattattttgtgcagcga

		tgggggcggggggggggggggggggcgcgcgccaggcggggcggggcgaggggc ggggcggggcgaggcgagaggtgcggcggcagccaatcagagcgggcgcgctccgaa agtttccttttatggcgaggcgggcgggcgggcgccctataaaaagcgaagcgcg ggcgggcgaggagtcgctgcgacgctgccttcgccccgtgccccgctccgcccgcg tcgcgccgccccgggctctgactgaccggttactcccacaggtgagcgggcg gacggcccttctcctccgggctgtaattagctgagcaagaggttaagggtttaaggga tggttggttggtgggtattaatgtttaattacctggagcacctgcctgaaatcact ttttttcaggttgggccaccATGGGCCCTAAGAAGCGCAGAAAGGTGACCGGCGAAA GACCATACGCCTGCCCAGTGGAAAGCTGCGACAGAAGATTGAGCAGAAGCAGCAACC TGACCTACCACATTAGAAATCCATACTGGAAGCCAAAAGCCTTTCCAATGCAGAATCT GCATGAGAAATTTTAGCAGATCCGCCAATCGGAAACAGCACATCAGGACACACACCG GCGAGAAGCCTTTTGCCTGCGATATCTGTGGCCGCAAATTTGCTAGAAGCAGCAACC TGAGCTACCATACAAAGATCCACACAGGAAGCCAGAAACCTTTCCAATGTAGAATCT GCATGCGGAACCTTACGACAGAAGCGCCAATCTTACACAGCACATCCGGACCCACACCG GCGAAAAGCCTTTTGCCTGCGACATCTGCGGCAGAAAGTTGCCCCGGTCCAGCAACA GGATTACGACACCAAGATCCACCTGAGACAGAAGGACGGCGGCGGGCTCTGGCG GAGGGGGTTCTGGCGGCGGTGGAAGCCAGAAGCCTTTCCAGTGTAAGATCTGCATGC GGAACCTTCTCGGTCTGCCAACCTGTCTTATCACATCCGGACACACACCGGAGAGA AGCCTTTTCGCTGCGACATCTGTGGCAGAAAGTTGCTCGGTCCAGCAACCGGACCC AACACACCAAATCCACACCGGATCTCAAAAACCTTTCCAGTGCAAGATCTGCATGA GAACTTTCAGCCGAGCGCCAACCTGAGCCAGCACATCAGAACCCACACCGGAGAAA AGCCTTTTGCCTGTGATATCTGCGGCAGAAAGTTGCGCAGAAGCAGCAATAGAAAGT ACCACACAAAGATTACACAGGCAGCCAGAAACCTTTCACTGCGCGATCTGCATGA GGAACCTTACGACAGAAGCGCCAACCTGAGCCAGCACATCCGGACCCACACAGGCGAGA AGCCCTTTGCTTGTGATATCTGCGGCAGAAAGTTGCGCAGAAGCAGCAACCGGAAGC AGCATACCAAGATCCACCTGCGGCAGAAAGACGGCGGCGGAAGCCCTATGGAATTCC AGTACCTGCCTGACACCGATGATAGACACAGAATCGAGGAAAAAGAGGAAGCGGACAT ACGAGACCTTCAAGAGCATCATGAAAAAGTCCCCTTTCTCCGGCCCCACAGACCCCA GACCTCCTCCTCGGCGGATCGCCGTGCCAAGCAGAAGTTCTGCTAGCGTGCCCAAGC CCGCACCCACAGCCTTACCCCTTCAAGCAGCCTGTCTACCATCAACTACGACGAGT TCCCTACAATGGTGTTCCCCAGCGGCCAGATCAGCCAAGCAGCGCCCTGGCTCCTG CCCCACCACAGGTGCTGCCCCAGGCCCCCTGCCCTGCCCCCGCCCTGCCATGGTGT CCGCCCTGGCCCCAGGCCCCCGCCCCAGTCCCCGTGCTGGCCCCAGGCCCTCCTCAAG CTGTTGCCCCCTCCTGCTCCTAAGCCTACCCAGGCCGGCGAGGGCACCCCTGTCCGAGG CCCTGCTGCAGCTGCAGTTTCAGACGACGAGGACCTGGGCGCCCTGCTGGGCAACAGCA CCGATCCTGCAGTGTTACCGACCTGGCCTCTGTGGACAACAGCGAGTTCCAGCAGC TGCTGAACCAGGGCATCCCCGTGGCCCCCACACGACCGAGCCTATGCTGATGGAAT ACCTTGAGGCCATCACCAGACTGGTGACCGGCGCTCAGAGACCTCCAGACCCCGCTC CTGCCCCCTCTGGGCGCCCCCGGCCCTGCCTAACGGCCTGCTGAGCGGCGACGAAGATT TCAGCTCCATCGCCGACATGGACTTCAGCGCCCTGCTCAGCCAGATCAGCAGCTGA
149	WPRE-polyA sequence immediately downstream of ZF2a of SEQ ID NOs: 147 and 148	gaattcgatatcaagccttatcgataatcaacctctggattacaaaatttgtgaaaga ttgactgggtattcttaactatgttgctccttttacgctatgtggatacgtgcttta atgcctttgtatcatgctattgcttcccgtatggctttcattttctcctccttgat aaatcctgggttgctgtctctttatgaggagttgtggccggtgtcaggcaacgtggc gtggtgtgcaactgtgtttgctgacgcaacccccactgggtggggcattgccaccacc tgtcagctcctttccgggactttcgctttccccctccctattgccacggcggaactc atcgccgcctgccttgcccgtgctggacaggggctcggtgttgggcactgacaat tccgtgggtgttgctggggaatcatcgtcctttccttggtgctcgctgtgttgcc acctggattctgcgcggaacgtccttctgtacgtcccttcggccctcaatccagcg gaccttccctcccgcgccgtgctgcccgtctgcccgtcttccgctcttccgctt cgccctcagacgagtcggatctccctttgggcccgtcccgcatcgataccgtcga cctcgagggggggcccggtacccgggaatcaattcactcctcaggtgaggctgcct atcagaaggtgggtggctgggtgtggccaatgccctggctcacaaataccactgagatc tttttccctctgccaaaaattatggggacatcatgaagcccccttgagcatctgactt ctggctaataaaggaaatttttttatttgcaatag

Table 9: Peptide and Nucleic Acid Sequences.

CLAUSES

Alternative expressions of the inventive concept are set out in each of the following numbered clauses.

A1. A polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within the frataxin gene promoter of SEQ ID NO: 67, or a sequence complementary thereto, and a transcriptional activation domain.

A2. The polypeptide of Clause A1, wherein the poly-zinc finger peptide is capable of binding to a target sequence within positions 1 to 700 of SEQ ID NO: 67, within positions 1 to 500 of SEQ ID NO: 67, within positions 1 to 400 of SEQ ID NO: 67, or within positions 20 to 380 of SEQ ID NO: 67.

A3. The polypeptide of Clause A1 or A2, wherein the poly-zinc finger peptide is capable of binding to a target sequence within any of SEQ ID NOs: 68 to 75, or a sequence complementary thereto.

A4. The polypeptide of Clause A3, wherein the poly-zinc finger peptide is capable of binding to SEQ ID NOs: 72, 73, 74 or 75.

A5. The polypeptide of any of Clauses A1 to A4, wherein the poly-zinc finger peptide comprises 6 or more zinc finger domains.

A6. The polypeptide of Clause A5, wherein the poly-zinc finger peptide has 6 zinc finger domains, 9 zinc finger domains, or 11 zinc finger domains.

A7. The polypeptide of any of Clauses A1 to A6, wherein the poly-zinc finger peptide comprises a zinc finger peptide having a zinc finger peptide array of any of CA to DD in Table 3.

A8. The polypeptide of any of Clauses A1 to A7, wherein the poly-zinc finger peptide comprises 6 zinc finger domains, and wherein each zinc finger domain (F1 to F6) comprises a recognition helix sequence as follows:

(i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; and F6, SEQ ID NO: 43;

(ii) F1, SEQ ID NO: 31; F2, SEQ ID NO: 34; F3, SEQ ID NO: 37; F4, SEQ ID NO: 40; F5, SEQ ID NO: 43; and F6, SEQ ID NO: 46;

(iii) F1, SEQ ID NO: 34; F2, SEQ ID NO: 37; F3, SEQ ID NO: 40; F4, SEQ ID NO: 43; F5, SEQ ID NO: 46; and F6, SEQ ID NO: 49;

(iv) F1, SEQ ID NO: 37; F2, SEQ ID NO: 40; F3, SEQ ID NO: 43; F4, SEQ ID NO: 46; F5, SEQ ID NO: 49; and F6, SEQ ID NO: 52;

(v) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; and F6, SEQ ID NO: 44;

(vi) F1, SEQ ID NO: 32; F2, SEQ ID NO: 35; F3, SEQ ID NO: 38; F4, SEQ ID NO: 41; F5, SEQ ID NO: 44; and F6, SEQ ID NO: 47;

(vii) F1, SEQ ID NO: 35; F2, SEQ ID NO: 38; F3, SEQ ID NO: 41; F4, SEQ ID NO: 44; F5, SEQ ID NO: 47; and F6, SEQ ID NO: 50;

(viii) F1, SEQ ID NO: 38; F2, SEQ ID NO: 41; F3, SEQ ID NO: 44; F4, SEQ ID NO: 47; F5, SEQ ID NO: 50; and F6, SEQ ID NO: 53;

5 (ix) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; and F6, SEQ ID NO: 45;

(x) F1, SEQ ID NO: 33; F2, SEQ ID NO: 36; F3, SEQ ID NO: 39; F4, SEQ ID NO: 42; F5, SEQ ID NO: 45; and F6, SEQ ID NO: 48;

(xi) F1, SEQ ID NO: 36; F2, SEQ ID NO: 39; F3, SEQ ID NO: 42; F4, SEQ ID NO: 45; F5, SEQ ID NO: 48; and F6, SEQ ID NO: 51; or

(xii) F1, SEQ ID NO: 39; F2, SEQ ID NO: 42; F3, SEQ ID NO: 45; F4, SEQ ID NO: 48; F5, SEQ ID NO: 51; and F6, SEQ ID NO: 54.

15 A9. The polypeptide of any of Clauses A1 to A7, wherein the poly-zinc finger peptide comprises 7 zinc finger domains, and wherein each zinc finger domain (F1 to F7) comprises a recognition helix sequence as follows:

(i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; F6, SEQ ID NO: 43; and F7, SEQ ID NO: 46;

(ii) F1, SEQ ID NO: 31; F2, SEQ ID NO: 34; F3, SEQ ID NO: 37; F4, SEQ ID NO: 40; F5, SEQ ID NO: 43; F6, SEQ ID NO: 46; and F7, SEQ ID NO: 49;

(iii) F1, SEQ ID NO: 34; F2, SEQ ID NO: 37; F3, SEQ ID NO: 40; F4, SEQ ID NO: 43; F5, SEQ ID NO: 46; F6, SEQ ID NO: 49; and F7, SEQ ID NO: 52;

(iv) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; F6, SEQ ID NO: 44; and F7, SEQ ID NO: 47;

25 (v) F1, SEQ ID NO: 32; F2, SEQ ID NO: 35; F3, SEQ ID NO: 38; F4, SEQ ID NO: 41; F5, SEQ ID NO: 44; F6, SEQ ID NO: 47; and F7, SEQ ID NO: 50;

(vi) F1, SEQ ID NO: 35; F2, SEQ ID NO: 38; F3, SEQ ID NO: 41; F4, SEQ ID NO: 44; F5, SEQ ID NO: 47; F6, SEQ ID NO: 50; and F7, SEQ ID NO: 53;

(vii) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; F6, SEQ ID NO: 45; and F7, SEQ ID NO: 48;

(viii) F1, SEQ ID NO: 33; F2, SEQ ID NO: 36; F3, SEQ ID NO: 39; F4, SEQ ID NO: 42; F5, SEQ ID NO: 45; F6, SEQ ID NO: 48; and F7, SEQ ID NO: 51; or

(ix) F1, SEQ ID NO: 36; F2, SEQ ID NO: 39; F3, SEQ ID NO: 42; F4, SEQ ID NO: 45; F5, SEQ ID NO: 48; F6, SEQ ID NO: 51; and F7, SEQ ID NO: 54.

35

A10. The polypeptide of any of Clauses A1 to A7, wherein the poly-zinc finger peptide comprises 8 zinc finger domains, and wherein each zinc finger domain (F1 to F8) comprises a recognition helix sequence as follows:

(i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; F6, SEQ ID NO: 43; F7, SEQ ID NO: 46; and F8, SEQ ID NO: 49;

(ii) F1, SEQ ID NO: 31; F2, SEQ ID NO: 34; F3, SEQ ID NO: 37; F4, SEQ ID NO: 40; F5, SEQ ID NO: 43; F6, SEQ ID NO: 46; F7, SEQ ID NO: 49; and F8, SEQ ID NO: 52;

40

(iii) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; F6, SEQ ID NO: 44; F7, SEQ ID NO: 47; and F8, SEQ ID NO: 50;

(iv) F1, SEQ ID NO: 32; F2, SEQ ID NO: 35; F3, SEQ ID NO: 38; F4, SEQ ID NO: 41; F5, SEQ ID NO: 44; F6, SEQ ID NO: 47; F7, SEQ ID NO: 50; and F8, SEQ ID NO: 53;

5 (v) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; F6, SEQ ID NO: 45; F7, SEQ ID NO: 48; and F8, SEQ ID NO: 51; or

(vi) F1, SEQ ID NO: 33; F2, SEQ ID NO: 36; F3, SEQ ID NO: 39; F4, SEQ ID NO: 42; F5, SEQ ID NO: 45; F6, SEQ ID NO: 48; F7, SEQ ID NO: 51; and F8, SEQ ID NO: 54.

10 A11. The polypeptide of any of Clauses A1 to A7, wherein the poly-zinc finger peptide comprises 9 zinc finger domains, and wherein each zinc finger domain (F1 to F9) comprises a recognition helix sequence as follows:

(i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; F6, SEQ ID NO: 43; F7, SEQ ID NO: 46; F8, SEQ ID NO: 49; and F9, SEQ ID NO: 52;

15 (ii) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; F6, SEQ ID NO: 44; F7, SEQ ID NO: 47; F8, SEQ ID NO: 50; and F9, SEQ ID NO: 53; or

(iii) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; F6, SEQ ID NO: 45; F7, SEQ ID NO: 48; F8, SEQ ID NO: 51; and F9, SEQ ID NO: 54.

20

A12. The polypeptide of any of Clauses A1 to A11, wherein the poly-zinc finger comprises the sequence of SEQ ID NO: 116 or a sequence having at least 90%, at least 95%, or at least 98%, at least 99% identity thereto.

25 A13. A polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within a 5'-GAA-3' trinucleotide repeat sequence, frameshift variants therefore (i.e. 5'-AGA-3' or 5'-AAG-3') or a nucleic acid sequence complementary thereto, and a transcriptional activation domain.

30 A14. The polypeptide of Clause A13, wherein the poly-zinc finger peptide is capable of binding to a nucleic acid target sequence within a 5'-GAA-3' trinucleotide repeat sequence, and wherein the 5'-GAA-3' trinucleotide repeat sequence comprises: at least 9 contiguous 5'-GAA-3' repeats; at least 11 contiguous 5'-GAA-3' repeats; at least 50 contiguous 5'-GAA-3' repeats; at least 66 contiguous 5'-GAA-3' repeats; at least 100 contiguous 5'-GAA-3' repeats; at least 300 contiguous 5'-GAA-3' repeats; at
35 least 600 contiguous 5'-GAA-3' repeats; or at least 900 contiguous 5'-GAA-3' repeats.

A15. The polypeptide of Clauses A13 or Clause A14, wherein the poly-zinc finger peptide comprises from 6 to 32, from 6 to 18 or from 6 to 12 zinc finger domains.

40 A16. The polypeptide of any of Clauses A13 to A15, wherein the poly-zinc finger peptide comprises 6, 9, 11 or 12 zinc finger domains.

A17. The polypeptide of any of Clauses A13 to A16, wherein at least 6 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 1.

5 A18. The polypeptide of any of Clauses A13 to A17, wherein at least 6 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising: SEQ ID NO: 2; SEQ ID NO: 3; or SEQ ID NO: 4.

10 A19. The polypeptide of any of Clauses A13 to A17, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 2.

15 A20. The polypeptide of any of Clauses A13 to A17, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 3.

A21. The polypeptide of any of Clauses A13 to A17, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 4.

20 A22. The polypeptide of any of Clauses A13 to A17, wherein the poly-zinc finger peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to F11) of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 5 or SEQ ID NO: 6, and wherein at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 5, and at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 6.

A23. The polypeptide of Clause A22, wherein:

30 (i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences according to SEQ ID NO: 5, and zinc finger domains F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 6;

(ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 5, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 6; or

35 (iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 5, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 6.

40 A24. The polypeptide of any of Clauses A13 to A17, wherein the poly-zinc finger peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to F11) of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 7 or SEQ ID NO: 8, and wherein at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide

comprise a recognition helix sequence comprising SEQ ID NO: 7, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 8.

5 A25. The polypeptide of Clause A24, wherein:

(i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences according to SEQ ID NO: 7, and zinc finger domains F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 8;

10 (ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 7, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 8; or

(iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 7, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 8.

15

A26. The polypeptide of any of Clauses A13 to A17, wherein the poly-zinc finger peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to F11) of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 8 or SEQ ID NO: 9, and wherein at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide
20 comprise a recognition helix sequence comprising SEQ ID NO: 8, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 9.

A27. The polypeptide of Clause A26, wherein:

25 (i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences according to SEQ ID NO: 8, and zinc finger domains F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 9;

(ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 8, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences
30 according to SEQ ID NO: 9; or

(iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 8, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 9.

35 A28. The polypeptide of any of Clauses A13 to A27, wherein the poly-zinc finger peptide comprises a zinc finger peptide having a zinc finger peptide array of any of A to J in Table 1.

A29. The polypeptide of any of Clauses A13 to A28, wherein the poly-zinc finger peptide comprises 9, 10, 11 or 12 zinc finger domains forming a zinc finger array according to any one of the following
40 patterns:

	F1, F3	F2	F4, F6, F8, F10, F12	F5, F7, F9, F11
A:	SEQ ID NO: 1	SEQ ID NO: 1	SEQ ID NO: 1	SEQ ID NO: 1

B:	SEQ ID NO: 2	SEQ ID NO: 2	SEQ ID NO: 2	SEQ ID NO: 2
C:	SEQ ID NO: 3	SEQ ID NO: 3	SEQ ID NO: 3	SEQ ID NO: 3
D:	SEQ ID NO: 4	SEQ ID NO: 4	SEQ ID NO: 4	SEQ ID NO: 4
E:	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 6	SEQ ID NO: 5
5 F:	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 5	SEQ ID NO: 6
G:	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 8	SEQ ID NO: 7
H:	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 7	SEQ ID NO: 8
I:	SEQ ID NO: 8	SEQ ID NO: 9	SEQ ID NO: 9	SEQ ID NO: 8
J:	SEQ ID NO: 8	SEQ ID NO: 9	SEQ ID NO: 8	SEQ ID NO: 9.

10

A30. The polypeptide of any of Clauses A13 to A29, wherein the poly-zinc finger comprises a sequence selected from one of SEQ ID NOs: 97 to 99 or a sequence having at least 90%, at least 95%, or at least 98%, at least 99% identity thereto.

15

A31. The polypeptide of any of Clauses A13 to A16, wherein at least 6 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 10.

20

A32. The polypeptide of any of Clauses A13 to A16 or A31, wherein at least 6 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising: SEQ ID NO: 11; SEQ ID NO: 12; or SEQ ID NO: 13.

25

A33. The polypeptide of any of Clauses A13 to A16, A31 or A32, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 11.

30

A34. The polypeptide of any of Clauses A13 to A16, A31 or A32, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 12.

35

A35. The polypeptide of any of Clauses A13 to A16, A31 or A32, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 13.

40

A36. The polypeptide of any of Clauses A13 to A16, or A31 to A33, wherein the poly-zinc finger peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to F11) of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 14 or SEQ ID NO: 15, and wherein at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 14, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 15.

A37. The polypeptide of Clause A36, wherein:

(i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences according to SEQ ID NO: 14, and zinc finger domains F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 15;

(ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 14, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 15; or

(iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 14, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 15.

A38. The polypeptide of any of Clauses A13 to A16, A31, A32 or A34, wherein the poly-zinc finger peptide comprises at least 11 zinc finger domains, and each of the at least 11 adjacent zinc finger domains (F1 to F11) of the poly-zinc finger peptide comprise a recognition helix sequence independently selected from any one of SEQ ID NO: 16 to 25, and wherein:

(i) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 17;

(ii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 19;

(iii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21;

(iv) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23;

(v) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 17;

(vi) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 19;

(vii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21;

(viii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23;

5 (ix) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 20;

10 (x) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 22;

15 (xi) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 24;

20 (xii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 25;

(xiii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 20;

25 (xiv) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 22;

30 (xv) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 24;

35 (xvi) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 25.

A39. The polypeptide of Clause A38, wherein the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise recognition helix sequences selected from:

- 40 (i) two of SEQ ID NOs: 16 to 25;
 (ii) three of SEQ ID NOs: 16 to 25;
 (iii) four of SEQ ID NOs: 16 to 25;

- (iv) five of SEQ ID NOs: 16 to 25;
- (v) six of SEQ ID NOs: 16 to 25;
- (vi) seven of SEQ ID NOs: 16 to 25;
- (vii) eight of SEQ ID NOs: 16 to 25;
- 5 (viii) nine of SEQ ID NOs: 16 to 25; or
- (ix) ten of SEQ ID NOs: 16 to 25.

A40. The polypeptide of Clause A38 or A39, wherein:

- (i) zinc finger domains F1, F3, F5, F7, F9 and F11 each has a recognition helix sequence
10 independently selected from any one of SEQ ID NOs: 16, 18, 20, 22, 24 and 25, and zinc finger domains
F2, F4, F6, F8 and F10 each has a recognition helix sequence independently selected from any one of
SEQ ID NO: 17, 19, 21 and 23;
- (ii) zinc finger domains F1, F3, F4, F6, F8 and F10 each has a recognition helix sequence
independently selected from any one of SEQ ID NOs: 16, 18, 20, 22, 24 and 25, and zinc finger domains
15 F2, F5, F7, F9 and F11 each has a recognition helix sequence independently selected from any one of
SEQ ID NO: 17, 19, 21 and 23; or
- (iii) zinc finger domains F1, F3, F5, F6, F8 and F10 each has a recognition helix sequence
independently selected from any one of SEQ ID NOs: 16, 18, 20, 22, 24 and 25, and zinc finger domains
F2, F4, F7, F9 and F11 each has a recognition helix sequence independently selected from any one of
20 SEQ ID NO: 17, 19, 21 and 23.

A41. The polypeptide of any of Clauses A13 to A16, A31, A32 and A35, wherein the poly-zinc finger
peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to
F11) of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 26
25 or SEQ ID NO: 27, and wherein at least one of the at least 11 adjacent zinc finger domains of the poly-
zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 26, and at least one
or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition
helix sequence comprising SEQ ID NO: 27.

30 A42. The polypeptide of Clause A41, wherein:

- (i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences
according to SEQ ID NO: 26, and zinc finger domains F2, F4, F6, F8 and F10 have recognition
sequences according to SEQ ID NO: 27;
- (ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according
35 to SEQ ID NO: 26, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences
according to SEQ ID NO: 27; or
- (iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according
to SEQ ID NO: 26, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences
according to SEQ ID NO: 27.

40

A43. The polypeptide of any of Clauses A13 to A42, wherein the poly-zinc finger peptide comprises
a zinc finger peptide having a zinc finger peptide array of any of K to BZ in Table 2.

A44. The polypeptide of any of Claims A13 to A16 or A31 to A43, wherein the poly-zinc finger peptide comprises 9, 10, 11 or 12 zinc finger domains forming a zinc finger array according to any one of the following patterns:

	F1, F3	F2	F4, F6, F8, F10, F12	F5, F7, F9, F11
5	K: SEQ ID NO: 10	SEQ ID NO: 10	SEQ ID NO: 10	SEQ ID NO: 10
	L: SEQ ID NO: 11	SEQ ID NO: 11	SEQ ID NO: 11	SEQ ID NO: 11
	M: SEQ ID NO: 12	SEQ ID NO: 12	SEQ ID NO: 12	SEQ ID NO: 12
	N: SEQ ID NO: 13	SEQ ID NO: 13	SEQ ID NO: 13	SEQ ID NO: 13
10	O: SEQ ID NO: 14	SEQ ID NO: 15	SEQ ID NO: 15	SEQ ID NO: 14
	P: SEQ ID NO: 14	SEQ ID NO: 15	SEQ ID NO: 14	SEQ ID NO: 15
	Q: SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 16
	R: SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 16	SEQ ID NO: 16
	S: SEQ ID NO: 16	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 16
15	T: SEQ ID NO: 16	SEQ ID NO: 19	SEQ ID NO: 16	SEQ ID NO: 19
	U: SEQ ID NO: 16	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 16
	V: SEQ ID NO: 16	SEQ ID NO: 20	SEQ ID NO: 16	SEQ ID NO: 20
	W: SEQ ID NO: 16	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 16
	X: SEQ ID NO: 16	SEQ ID NO: 22	SEQ ID NO: 16	SEQ ID NO: 22
20	Y: SEQ ID NO: 16	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 16
	Z: SEQ ID NO: 16	SEQ ID NO: 24	SEQ ID NO: 16	SEQ ID NO: 24
	AA: SEQ ID NO: 16	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 16
	AB: SEQ ID NO: 16	SEQ ID NO: 25	SEQ ID NO: 16	SEQ ID NO: 25
	AC: SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 18
25	AD: SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 18	SEQ ID NO: 17
	AE: SEQ ID NO: 18	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 18
	AF: SEQ ID NO: 18	SEQ ID NO: 19	SEQ ID NO: 18	SEQ ID NO: 19
	AG: SEQ ID NO: 18	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 18
	AH: SEQ ID NO: 18	SEQ ID NO: 20	SEQ ID NO: 18	SEQ ID NO: 20
30	AI: SEQ ID NO: 18	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 18
	AJ: SEQ ID NO: 18	SEQ ID NO: 22	SEQ ID NO: 18	SEQ ID NO: 22
	AK: SEQ ID NO: 18	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 18
	AL: SEQ ID NO: 18	SEQ ID NO: 24	SEQ ID NO: 18	SEQ ID NO: 24
	AM: SEQ ID NO: 18	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 18
35	AN: SEQ ID NO: 18	SEQ ID NO: 25	SEQ ID NO: 18	SEQ ID NO: 25
	AO: SEQ ID NO: 21	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 21
	AP: SEQ ID NO: 21	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 17
	AQ: SEQ ID NO: 21	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 21
	AR: SEQ ID NO: 21	SEQ ID NO: 19	SEQ ID NO: 21	SEQ ID NO: 19
40	AS: SEQ ID NO: 21	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 21
	AT: SEQ ID NO: 21	SEQ ID NO: 20	SEQ ID NO: 21	SEQ ID NO: 20
	AU: SEQ ID NO: 21	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 21

	AV:	SEQ ID NO: 21	SEQ ID NO: 22	SEQ ID NO: 21	SEQ ID NO: 22
	AW:	SEQ ID NO: 21	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 21
	AX:	SEQ ID NO: 21	SEQ ID NO: 24	SEQ ID NO: 21	SEQ ID NO: 24
	AY:	SEQ ID NO: 21	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 21
5	AZ:	SEQ ID NO: 21	SEQ ID NO: 25	SEQ ID NO: 21	SEQ ID NO: 25
	BA:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 20
	BB:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 22
	BC:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 24
	BD:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 25
10	BE:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 20
	BF:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 22
	BG:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 24
	BH:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 25
	BI:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 20
15	BJ:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 22
	BK:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 24
	BL:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 25
	BM:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 20
	BN:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 22
20	BO:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 24
	BP:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 25
	BQ:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 20
	BR:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 22
	BS:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 24
25	BT:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 25
	BU:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 20
	BV:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 22
	BW:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 24
	BX:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 25
30	BY:	SEQ ID NO: 26	SEQ ID NO: 27	SEQ ID NO: 27	SEQ ID NO: 26
	BZ:	SEQ ID NO: 26	SEQ ID NO: 27	SEQ ID NO: 26	SEQ ID NO: 27.

A45: The polypeptide of any of Clauses A13 to A16 or A31 to A44, wherein the poly-zinc finger comprises a sequence selected from one of SEQ ID NOs: 102 to 104, or a sequence having at least 90%, at least 95%, or at least 98%, at least 99% identity thereto.

A46. The polypeptide of any of Clauses A1 to A45, wherein the poly-zinc finger peptide has a sequence defined by the formula:

$$\text{N}' - [(\text{Formula 4}) - \text{L}_3]_{\text{n}0} - \{[(\text{Formula 6}) - \text{L}_2 - (\text{Formula 6}) - \text{L}_3]_{\text{n}1} - [(\text{Formula 6}) - \text{L}_2 - (\text{Formula 6}) - \text{X}_\text{L}]\}_{\text{n}2} - [(\text{Formula 4}) - \text{L}_2 - (\text{Formula 6}) - \text{L}_3]_{\text{n}3} - [(\text{Formula 6}) - \text{L}_2 - (\text{Formula 6})] - [\text{L}_3 - (\text{Formula 6})]_{\text{n}4} - \text{C}'$$

wherein n_0 is selected to be 0 or 1, n_1 is an integer selected from 1 to 4, n_2 is an integer selected from 1 or 2, n_3 is an integer selected from 1 to 4, n_4 is selected to be 0 or 1, L_2 is a linker sequence selected from the sequence $-TG^{E/Q/K/R}P-$ (SEQ ID NO: 77), L_3 is a linker sequence selected from the sequence $-TG^{G/S^E/Q^K/R}P-$ (SEQ ID NO: 79), and X_L is a linker sequence of between 8 and 50 amino acids;

Formula 4 is a zinc finger domain having a sequence selected from the formula $X_2 C X_2$ or $4 C X_5 X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6} H X_{3,4 \text{ or } 5} H/C$ and Formula 6 is a zinc finger domain having a sequence selected from the formula $X_2 C X_2 C X_5 X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6} H X_3 H$, wherein X is an amino acid, the number in subscript indicates the number of amino acids X in that position of the sequence and $X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6}$ are the amino acids in positions -1 to 6 of each zinc finger domain recognition helix, wherein the sequence $X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6}$ is defined according to the zinc finger recognition helix SEQ ID NOs given in Clauses A1 to A45.

A47. The polypeptide according to any of Clauses A1 to A46, wherein the polypeptide comprises an activation domain selected from the VP64 domain (SEQ ID NO: 125), the herpes simplex virus (HSV) VP16 domain (SEQ ID NO: 124), or the p65-RelA activation domain (SEQ ID NO: 122 or SEQ ID NO: 123).

A48. The polypeptide according to Clause A47, wherein the activation domain is the human p65-RelA activation domain (SEQ ID NO: 122) or the mouse p65-RelA activation domain (SEQ ID NO: 123).

A49. The polypeptide according to Clause A46 or A47, wherein the activation domain is attached to the C-terminal end of the zinc finger peptide.

A50. The polypeptide according to any of Clauses A46 to A49, wherein the activation domain is attached to the C-terminus of the zinc finger peptide via a linker sequence comprising a sequence selected from any one or SEQ ID NOs: 126 to 130.

A51. The polypeptide according to any of Clauses A1 to A50, wherein the polypeptide comprises a nuclear localisation signal (NLS) sequence.

A52. The polypeptide according to Clause A51, wherein the nuclear localisation signal comprises a nuclear localisation signal from SV40, mouse primase p58, or human protein KIAA2022.

A53. The polypeptide according to Clause A51 or A52, wherein the nuclear localisation signal is selected from the SV40 NLS (SEQ ID NO: 107), the mouse primase p58 NLS (SEQ ID NO: 109), the human protein KIAA2022 NLS (SEQ ID NO: 108); the double SV40 NLS (SEQ ID NO: 113), the double human KIAA2022 NLS (SEQ ID NO: 114) or the double mouse primase p58 NLS (SEQ ID NO: 115).

A54. The polypeptide according to any of Clauses A51 to A53, wherein the nuclear localisation signal is the mouse primase p58 NLS or the double mouse primase NLS.

A55. The polypeptide according to any of Clauses A51 to A53, wherein the nuclear localisation signal is the human KIAA2022 NLS or the double human KIAA2022 NLS.

5 A56. A polypeptide which comprises a poly-zinc finger peptide DNA binding domain and a transcriptional activation domain, which comprises a sequence according to any of SEQ ID NOs: 100, 101, 105, 106, 117 or 118, or a sequence having at least 90%, at least 95%, at least 98%, or at least 99% identity thereto.

10 A57. The polypeptide of any of Clauses A13 to A56, which binds to a 5'-GAA-3' trinucleotide repeat sequence containing at least 66 trinucleotide repeats, with a binding affinity stronger than about 1 μ M, stronger than about 100 nM, stronger than about 10 nM, or stronger than about 1 nM.

15 A58. The polypeptide of any of Clauses A13 to A57, which binds to a 5'-GAA-3' trinucleotide repeat sequence containing at least 66 trinucleotide repeats in preference to or with a higher binding affinity than to a 5'-GAA-3' trinucleotide repeat sequence containing less than 12 trinucleotide repeats.

20 A59. The polypeptide of any of Clauses A13 to A57, which binds to a 5'-GAA-3' trinucleotide repeat sequence containing at least 300 trinucleotide repeats in preference to or with a higher binding affinity than to a 5'-GAA-3' trinucleotide repeat sequence containing less than 20 trinucleotide repeats.

A60. The polypeptide of any of Clauses A13 to A57, which binds to a 5'-GAA-3' trinucleotide repeat sequence containing at least 600 trinucleotide repeats in preference to or with a higher binding affinity than to a 5'-GAA-3' trinucleotide repeat sequence containing less than 30 trinucleotide repeats.

25 A61. An isolated nucleic acid encoding the polypeptide of any of Clauses A1 to A60; or a nucleic acid having the sequence of SEQ ID NO: 147 or 148.

A62. A vector comprising the nucleic acid of Clause A61.

30 A63. The vector according to Clause A61, which is a viral vector derived from retroviruses, such as influenza, SIV, HIV, lentivirus, and Moloney murine leukaemia; adenoviruses; adeno-associated viruses (AAV); herpes simplex virus (HSV); and chimeric viruses.

35 A64. The vector according to Clause A62 or A63, which is an adeno-associated virus (AAV) vector selected from an AAV2/1 subtype vector, or an AAV2/9 subtype vector.

A65. A pharmaceutical composition comprising the polypeptide according to any of Clauses A1 to A60, a nucleic acid according to Clause A61, or a vector according to any of Clauses A62 to A64.

40 A66. The polypeptide according to any of Clauses A1 to A60, the nucleic acid according to Clause A61, the vector according to any of Clauses A62 to A64, or the pharmaceutical composition according

to Clause A65, for use in treating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

5 A67. The polypeptide according to any of Clauses A1 to A60, the nucleic acid according to Clause A61, the vector according to any of Clauses A62 to A64, or the pharmaceutical composition according to Clause A65, for use in preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

10 A68. The polypeptide, nucleic acid, vector or pharmaceutical composition for use according to Clause A66 or A67, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

15 A69. A method of treating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polypeptide according to any of Clauses A1 to A60, the nucleic acid according to Clause A61, the vector according to any of Clauses A62 to A64, or the pharmaceutical composition according to Clause A65.

20 A70. A method of preventing or ameliorating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polypeptide according to any of Clauses A1 to A60, the nucleic acid according to Clause A61, the vector according to any of Clauses A62 to A64, or the pharmaceutical composition according to Clause A65.

25 A71. The method of Clause A69 or A70, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

30 A72. A pharmaceutical composition comprising a polypeptide according to any of Clauses A1 to A12 in combination with a polypeptide according to any of Clauses A13 to A60; or a nucleic acid encoding a polypeptide according to any of Clauses A1 to A12 in combination with a nucleic acid encoding a polypeptide according to any of Clauses A13 to A60.

A73. A polypeptide according to any of Clauses A1 to A12 in combination with a polypeptide according to any of Clauses A13 to A60 for use in treating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

35 A74. A polypeptide according to any of Clauses A1 to A12 in combination with a polypeptide according to any of Clauses A13 to A60 for use in preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

40 A75. The combination of polypeptides for use according to Clause A73 or A74, wherein the polypeptide according to any of Clauses A1 to A12 is administered simultaneously, sequentially or separately from the polypeptide according to any of Clauses A13 to A60.

A76. A nucleic acid encoding a polypeptide according to any of Clauses A1 to A12 in combination with a nucleic acid encoding a polypeptide according to any of Clauses A13 to A60 or a nucleic acid comprising SEQ ID NO: 147 or 148 for use in treating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

5

A77. A nucleic acid encoding a polypeptide according to any of Clauses A1 to A12 in combination with a nucleic acid encoding a polypeptide according to any of Clauses A13 to A60 or a nucleic acid comprising SEQ ID NO: 147 or 148 for use in preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

10

A78. The combination of nucleic acids for use according to Clause A76 or A77, wherein the nucleic acid encoding a polypeptide according to any of Clauses A1 to A12 is administered simultaneously, sequentially or separately from the nucleic acid encoding a polypeptide according to any of Clauses A13 to A60 or the nucleic acid comprising SEQ ID NO: 147 or 148.

15

A79. A method of treating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polypeptide according to any of Clauses A1 to A12 in combination with the polypeptide according to any of Clauses A13 to A60.

20

A80. A method of preventing or ameliorating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polypeptide according to any of Clauses A1 to A12 in combination with the polypeptide according to any of Clauses A13 to A60.

25

A81. A method of treating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject a nucleic acid encoding the polypeptide according to any of Clauses A1 to A12 in combination with a nucleic acid encoding the polypeptide according to any of Clauses A13 to A60 or a nucleic acid comprising SEQ ID NO: 147 or 148.

30

A82. A method of preventing or ameliorating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject a nucleic acid encoding the polypeptide according to any of Clauses A1 to A12 in combination with a nucleic acid encoding the polypeptide according to any of Clauses A13 to A60 or a nucleic acid comprising SEQ ID NO: 147 or 148.

35

A83. The pharmaceutical composition according to Clause A72 for use in treating, preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

40

A84. The combination of polypeptides for use according to Clauses A73 to A75, the combination of nucleic acids for use according to Clauses A76 to A78, the method of any of Clauses A79 to A82 or the pharmaceutical composition for use according to Clause A83, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

A85. A vector comprising a nucleic acid encoding a polypeptide according to any of Clauses A1 to A12 and a nucleic acid encoding a polypeptide according to any of Clauses A13 to A60 or a nucleic acid comprising SEQ ID NO: 147 or 148.

5 A86. The vector according to Clause A85, which is a viral vector derived from retroviruses, such as influenza, SIV, HIV, lentivirus, and Moloney murine leukaemia; adenoviruses; adeno-associated viruses (AAV); herpes simplex virus (HSV); and chimeric viruses.

10 A87. The vector according to Clause A86, which is an adeno-associated virus (AAV) vector selected from an AAV2/1 subtype vector, or an AAV2/9 subtype vector.

15 A88. A gene therapy method comprising administering to a subject in need thereof a vector according to any of Clauses A62 to A64, or A85 to A87, for use in treating, preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

A89. The gene therapy method of Clause A88, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

20 A90. The polypeptides, nucleic acids, vectors or pharmaceutical compositions for use according to Clause A68 or A84, or the methods according to any of Clauses A71, A84 or A89, wherein expression of frataxin transcript in a pathogenic target cell is increased to at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or at least 100% of the level of frataxin transcript in a wild-type, non-pathogenic cell of the host subject or animal; or compared to the level of frataxin transcript in a control
25 subject or animal that does not have Friedreich's ataxia (FRDA).

30 B1. A polynucleotide encoding a polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within the frataxin gene promoter of SEQ ID NO: 67, or a sequence complementary thereto, and a transcriptional activation domain.

B2. The polynucleotide according to Clause B1, wherein the polypeptide is defined according to any of Clauses A2 to A12.

35 B3. A polynucleotide encoding a polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within a 5'-GAA-3' trinucleotide repeat sequence, frameshift variants therefore (i.e. 5'-AGA-3' or 5'-AAG-3') or a nucleic acid sequence complementary thereto, and a transcriptional activation domain.

40 B4. The polypeptide according to Clause B3, wherein the polypeptide is defined according to any of Clauses A14 to A60.

B5. A vector comprising the nucleic acid of any of Clauses B1 to B4.

B6. The vector according to Clause B5, which is a viral vector derived from retroviruses, such as influenza, SIV, HIV, lentivirus, and Moloney murine leukaemia; adenoviruses; adeno-associated viruses (AAV); herpes simplex virus (HSV); and chimeric viruses.

5

B7. The vector according to Clause B6, which is an adeno-associated virus (AAV) vector selected from an AAV2/1 subtype vector, or an AAV2/9 subtype vector.

B8. A pharmaceutical composition comprising a polynucleotide according to any of Clauses B1 to B4, or a vector according to any of Clauses B5 to B7.

10

B9. The polynucleotide according to any of Clauses B1 to B4, the vector according to any of Clauses B5 to B7, or the pharmaceutical composition according to Clause B8, for use in treating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

15

B10. The polynucleotide according to any of Clauses B1 to B4, the vector according to any of Clauses B5 to B7, or the pharmaceutical composition according to Clause B8, for use in preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

20

B11. The polynucleotide, vector or pharmaceutical composition for use according to Clause B9 or B10, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

25

B12. A method of treating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polynucleotide according to any of Clauses B1 to B4, the vector according to any of Clauses B5 to B7, or the pharmaceutical composition according to Clause B8.

30

B13. A method of preventing or ameliorating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polynucleotide according to any of Clauses B1 to B4, the vector according to any of Clauses B5 to B7, or the pharmaceutical composition according to Clause B8.

35

B14. The method of Clause B12 or B13, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

B15. A pharmaceutical composition comprising a polynucleotide according to Clause B1 or B2 in combination with a polynucleotide according to Clause B3 or B4.

40

B16. A polynucleotide according to Clause B1 or B2 in combination with a polynucleotide according to Clause B3 or B4 for use in treating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

5 B17. A polynucleotide according to Clause B1 or B2 in combination with a polynucleotide according to Clause B3 or B4 for use in preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

10 B18. The combination of polynucleotides for use according to Clause B16 or B17, wherein the polynucleotide according to Clause B1 or B2 is administered simultaneously, sequentially or separately from the polynucleotide according to Clause B3 or B4.

15 B19. A method of treating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polynucleotide according to Clause B1 or B2 in combination with the polynucleotide according to Clause B3 or B4.

20 B20. A method of preventing or ameliorating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polynucleotide according to Clause B1 or B2 in combination with the polynucleotide according to Clause B3 or B4.

B21. The pharmaceutical composition according to Clause B15 for use in treating, preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

25 B22. The combination of polynucleotides for use according to Clauses B16 to B18, the method of Clause B19 or B20, or the pharmaceutical composition for use according to Clause B21, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

30 B23. A vector comprising a polynucleotide according to Clause B1 or B2 and a polynucleotide according to Clause B3 or B4.

35 B24. The vector according to Clause B23, which is a viral vector derived from retroviruses, such as influenza, SIV, HIV, lentivirus, and Moloney murine leukaemia; adenoviruses; adeno-associated viruses (AAV); herpes simplex virus (HSV); and chimeric viruses.

B25. The vector according to Clause B24, which is an adeno-associated virus (AAV) vector selected from an AAV2/1 subtype vector, or an AAV2/9 subtype vector.

40 B26. A gene therapy method comprising administering to a subject in need thereof a vector according to any of Clauses B5 to B7 or B23 to B25, for use in treating, preventing or ameliorating a

disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

B27. The gene therapy method of Clause B26, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

B28. The polynucleotides, vectors or pharmaceutical compositions for use according to Clause B11 or B22, or the methods according to any of Clauses B14, B22 or B27, wherein expression of frataxin transcript in a pathogenic target cell is increased to at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or at least 100% of the level of frataxin transcript in a wild-type, non-pathogenic cell of the host subject or animal; or compared to the level of frataxin transcript in a control subject or animal that does not have Friedreich's ataxia (FRDA).

C1. A polypeptide comprising a zinc finger peptide capable of binding to a nucleic acid target sequence within the frataxin gene promoter of SEQ ID NO: 67, or a sequence complementary thereto.

C2. The polypeptide of Clause C1, wherein the zinc finger peptide is capable of binding to a target sequence within positions 1 to 700 of SEQ ID NO: 67, within positions 1 to 500 of SEQ ID NO: 67, within positions 1 to 400 of SEQ ID NO: 67, or within positions 20 to 380 of SEQ ID NO: 67.

C3. The polypeptide of Clause C1 or C2, wherein the zinc finger peptide is capable of binding to a target sequence within any of SEQ ID NOs: 68 to 75, or a sequence complementary thereto.

C4. The polypeptide of Clause C3, wherein the zinc finger peptide is capable of binding to SEQ ID NOs: 72, 73, 74 or 75.

C5. The polypeptide of any of Clauses C1 to C4, wherein the zinc finger peptide comprises 6 or more zinc finger domains.

C6. The polypeptide of Clause C5, wherein the zinc finger peptide has 6 zinc finger domains, 9 zinc finger domains, or 11 zinc finger domains.

C7. The polypeptide of any of Clauses C1 to C6, wherein the zinc finger peptide comprises a zinc finger array according to any of CA to DD in Table 3.

C8. The polypeptide of any of Clauses C1 to C7, wherein the zinc finger peptide comprises 6 zinc finger domains, and wherein each zinc finger domain (F1 to F6) comprises a recognition helix sequence according to the following:

(i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; and F6, SEQ ID NO: 43;

(ii) F1, SEQ ID NO: 31; F2, SEQ ID NO: 34; F3, SEQ ID NO: 37; F4, SEQ ID NO: 40; F5, SEQ ID NO: 43; and F6, SEQ ID NO: 46;

(iii) F1, SEQ ID NO: 34; F2, SEQ ID NO: 37; F3, SEQ ID NO: 40; F4, SEQ ID NO: 43; F5, SEQ ID NO: 46; and F6, SEQ ID NO: 49;

(iv) F1, SEQ ID NO: 37; F2, SEQ ID NO: 40; F3, SEQ ID NO: 43; F4, SEQ ID NO: 46; F5, SEQ ID NO: 49; and F6, SEQ ID NO: 52;

5 (v) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; and F6, SEQ ID NO: 44;

(vi) F1, SEQ ID NO: 32; F2, SEQ ID NO: 35; F3, SEQ ID NO: 38; F4, SEQ ID NO: 41; F5, SEQ ID NO: 44; and F6, SEQ ID NO: 47;

10 (vii) F1, SEQ ID NO: 35; F2, SEQ ID NO: 38; F3, SEQ ID NO: 41; F4, SEQ ID NO: 44; F5, SEQ ID NO: 47; and F6, SEQ ID NO: 50;

(viii) F1, SEQ ID NO: 38; F2, SEQ ID NO: 41; F3, SEQ ID NO: 44; F4, SEQ ID NO: 47; F5, SEQ ID NO: 50; and F6, SEQ ID NO: 53;

(ix) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; and F6, SEQ ID NO: 45;

15 (x) F1, SEQ ID NO: 33; F2, SEQ ID NO: 36; F3, SEQ ID NO: 39; F4, SEQ ID NO: 42; F5, SEQ ID NO: 45; and F6, SEQ ID NO: 48;

(xi) F1, SEQ ID NO: 36; F2, SEQ ID NO: 39; F3, SEQ ID NO: 42; F4, SEQ ID NO: 45; F5, SEQ ID NO: 48; and F6, SEQ ID NO: 51; or

20 (xii) F1, SEQ ID NO: 39; F2, SEQ ID NO: 42; F3, SEQ ID NO: 45; F4, SEQ ID NO: 48; F5, SEQ ID NO: 51; and F6, SEQ ID NO: 54.

C9. The polypeptide of any of Clauses C1 to C7, wherein the zinc finger peptide comprises 7 zinc finger domains, and wherein each zinc finger domain (F1 to F7) comprises a recognition helix sequence according to the following:

25 (i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; F6, SEQ ID NO: 43; and F7, SEQ ID NO: 46;

(ii) F1, SEQ ID NO: 31; F2, SEQ ID NO: 34; F3, SEQ ID NO: 37; F4, SEQ ID NO: 40; F5, SEQ ID NO: 43; F6, SEQ ID NO: 46; and F7, SEQ ID NO: 49;

30 (iii) F1, SEQ ID NO: 34; F2, SEQ ID NO: 37; F3, SEQ ID NO: 40; F4, SEQ ID NO: 43; F5, SEQ ID NO: 46; F6, SEQ ID NO: 49; and F7, SEQ ID NO: 52;

(iv) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; F6, SEQ ID NO: 44; and F7, SEQ ID NO: 47;

(v) F1, SEQ ID NO: 32; F2, SEQ ID NO: 35; F3, SEQ ID NO: 38; F4, SEQ ID NO: 41; F5, SEQ ID NO: 44; F6, SEQ ID NO: 47; and F7, SEQ ID NO: 50;

35 (vi) F1, SEQ ID NO: 35; F2, SEQ ID NO: 38; F3, SEQ ID NO: 41; F4, SEQ ID NO: 44; F5, SEQ ID NO: 47; F6, SEQ ID NO: 50; and F7, SEQ ID NO: 53;

(vii) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; F6, SEQ ID NO: 45; and F7, SEQ ID NO: 48;

40 (viii) F1, SEQ ID NO: 33; F2, SEQ ID NO: 36; F3, SEQ ID NO: 39; F4, SEQ ID NO: 42; F5, SEQ ID NO: 45; F6, SEQ ID NO: 48; and F7, SEQ ID NO: 51; or

(ix) F1, SEQ ID NO: 36; F2, SEQ ID NO: 39; F3, SEQ ID NO: 42; F4, SEQ ID NO: 45; F5, SEQ ID NO: 48; F6, SEQ ID NO: 51; and F7, SEQ ID NO: 54.

C10. The polypeptide of any of Clauses C1 to C7, wherein the zinc finger peptide comprises 8 zinc finger domains, and wherein each zinc finger domain (F1 to F8) comprises a recognition helix sequence according to the following:

- 5 (i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; F6, SEQ ID NO: 43; F7, SEQ ID NO: 46; and F8, SEQ ID NO: 49;
- (ii) F1, SEQ ID NO: 31; F2, SEQ ID NO: 34; F3, SEQ ID NO: 37; F4, SEQ ID NO: 40; F5, SEQ ID NO: 43; F6, SEQ ID NO: 46; F7, SEQ ID NO: 49; and F8, SEQ ID NO: 52;
- 10 (iii) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; F6, SEQ ID NO: 44; F7, SEQ ID NO: 47; and F8, SEQ ID NO: 50;
- (iv) F1, SEQ ID NO: 32; F2, SEQ ID NO: 35; F3, SEQ ID NO: 38; F4, SEQ ID NO: 41; F5, SEQ ID NO: 44; F6, SEQ ID NO: 47; F7, SEQ ID NO: 50; and F8, SEQ ID NO: 53;
- (v) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; F6, SEQ ID NO: 45; F7, SEQ ID NO: 48; and F8, SEQ ID NO: 51; or
- 15 (vi) F1, SEQ ID NO: 33; F2, SEQ ID NO: 36; F3, SEQ ID NO: 39; F4, SEQ ID NO: 42; F5, SEQ ID NO: 45; F6, SEQ ID NO: 48; F7, SEQ ID NO: 51; and F8, SEQ ID NO: 54.

C11. The polypeptide of any of Clauses C1 to C7, wherein the zinc finger peptide comprises 9 zinc finger domains, and wherein each zinc finger domain (F1 to F9) comprises a recognition helix sequence according to the following:

- 20 (i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; F6, SEQ ID NO: 43; F7, SEQ ID NO: 46; F8, SEQ ID NO: 49; and F9, SEQ ID NO: 52;
- (ii) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; F6, SEQ ID NO: 44; F7, SEQ ID NO: 47; F8, SEQ ID NO: 50; and F9, SEQ ID NO: 53;
- 25 or
- (iii) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; F6, SEQ ID NO: 45; F7, SEQ ID NO: 48; F8, SEQ ID NO: 51; and F9, SEQ ID NO: 54.

C12. The polypeptide of any of Clauses C1 to C11, which comprises the sequence of SEQ ID NO: 116 or a sequence having at least 90%, at least 95%, or at least 98%, at least 99% identity thereto.

C13. A polypeptide comprising a zinc finger peptide capable of binding to a nucleic acid target sequence within a 5'-GAA-3' trinucleotide repeat sequence, frameshift variants thereof (i.e. 5'-AGA-3' or 5'-AAG-3') or a nucleic acid sequence complementary thereto.

35 C14. The polypeptide of Clause C13, wherein the zinc finger peptide is capable of binding to a nucleic acid target sequence within a 5'-GAA-3' trinucleotide repeat sequence, and wherein the 5'-GAA-3' trinucleotide repeat sequence comprises: at least 9 contiguous 5'-GAA-3' repeats; at least 11 contiguous 5'-GAA-3' repeats; at least 50 contiguous 5'-GAA-3' repeats; at least 66 contiguous 5'-GAA-3' repeats; at least 100 contiguous 5'-GAA-3' repeats; at least 300 contiguous 5'-GAA-3' repeats; at least 600 contiguous 5'-GAA-3' repeats; or at least 900 contiguous 5'-GAA-3' repeats.

C15. The polypeptide of Clauses C13 or Clause C14, which comprises from 6 to 32, from 6 to 18 or from 6 to 12 zinc finger domains.

5 C16. The polypeptide of any of Clauses C13 to C15, which comprises 6, 9, 11 or 12 zinc finger domains.

C17. The polypeptide of any of Clauses C13 to C16, wherein at least 6 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 1.

10 C18. The polypeptide of any of Clauses C13 to C17, wherein at least 6 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising: SEQ ID NO: 2; SEQ ID NO: 3; or SEQ ID NO: 4.

15 C19. The polypeptide of any of Clauses C13 to C17, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 2.

20 C20. The polypeptide of any of Clauses C13 to C17, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 3.

25 C21. The polypeptide of any of Clauses C13 to C17, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 4.

30 C22. The polypeptide of any of Clauses C13 to C17, wherein the zinc finger peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to F11) of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 5 or SEQ ID NO: 6, and wherein at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 5, and at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 6.

35 C23. The polypeptide of Clause C22, wherein:
(i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences according to SEQ ID NO: 5, and zinc finger domains F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 6;

40 (ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 5, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 6; or

(iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 5, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 6.

5 C24. The polypeptide of any of Clauses C13 to C17, wherein the zinc finger peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to F11) of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 7 or SEQ ID NO: 8, and wherein at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 7, and at least one or the at least 11 adjacent zinc
10 finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 8.

C25. The polypeptide of Clause C24, wherein:

(i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences
15 according to SEQ ID NO: 7, and zinc finger domains F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 8;

(ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 7, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 8; or

(iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 7, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 8.

25 C26. The polypeptide of any of Clauses C13 to C17, wherein the zinc finger peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to F11) of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 8 or SEQ ID NO: 9, and wherein at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 8, and at least one or the at least 11 adjacent zinc
30 finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 9.

C27. The polypeptide of Clause C26, wherein:

(i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences according to SEQ ID NO: 8, and zinc finger domains F2, F4, F6, F8 and F10 have recognition
35 sequences according to SEQ ID NO: 9;

(ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 8, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 9; or

(iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 8, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 9.

C28. The polypeptide of any of Clauses C13 to C27, wherein the zinc finger peptide comprises a zinc finger peptide array according to any of A to J in Table 1.

C29. The polypeptide of any of Clauses C13 to C28, wherein the zinc finger peptide comprises 9, 10, 11 or 12 zinc finger domains and forms a zinc finger array according to any one of the following patterns:

	F1, F3	F2	F4, F6, F8, F10, F12	F5, F7, F9, F11
A:	SEQ ID NO: 1	SEQ ID NO: 1	SEQ ID NO: 1	SEQ ID NO: 1
B:	SEQ ID NO: 2	SEQ ID NO: 2	SEQ ID NO: 2	SEQ ID NO: 2
10 C:	SEQ ID NO: 3	SEQ ID NO: 3	SEQ ID NO: 3	SEQ ID NO: 3
D:	SEQ ID NO: 4	SEQ ID NO: 4	SEQ ID NO: 4	SEQ ID NO: 4
E:	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 6	SEQ ID NO: 5
F:	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 5	SEQ ID NO: 6
G:	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 8	SEQ ID NO: 7
15 H:	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 7	SEQ ID NO: 8
I:	SEQ ID NO: 8	SEQ ID NO: 9	SEQ ID NO: 9	SEQ ID NO: 8
J:	SEQ ID NO: 8	SEQ ID NO: 9	SEQ ID NO: 8	SEQ ID NO: 9.

C30. The polypeptide of any of Clauses C13 to C29, which comprises a sequence selected from one of SEQ ID NOs: 97 to 99 or a sequence having at least 90%, at least 95%, or at least 98%, at least 99% identity thereto.

C31. The polypeptide of any of Clauses C13 to C16, wherein at least 6 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 10.

C32. The polypeptide of any of Clauses C13 to C16 or C31, wherein at least 6 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising: SEQ ID NO: 11; SEQ ID NO: 12; or SEQ ID NO: 13.

C33. The polypeptide of any of Clauses C13 to C16, C31 or C32, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 11.

C34. The polypeptide of any of Clauses C13 to C16, C31 or C32, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 12.

C35. The polypeptide of any of Clauses C13 to C16, C31 or C32, wherein at least 6, at least 9, or at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 13.

C36. The polypeptide of any of Clauses C13 to C16, or C31 to C33, wherein the zinc finger peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to F11) of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 14 or SEQ ID NO: 15, and wherein at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 14, and at least one or the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 15.

C37. The polypeptide of Clause C36, wherein:

(i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences according to SEQ ID NO: 14, and zinc finger domains F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 15;

(ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 14, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 15; or

(iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 14, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 15.

C38. The polypeptide of any of Clauses C13 to C16, C31, C32 or C34, wherein the zinc finger peptide comprises at least 11 zinc finger domains, and each of the at least 11 adjacent zinc finger domains (F1 to F11) of the zinc finger peptide comprise a recognition helix sequence independently selected from any one of SEQ ID NO: 16 to 25, and wherein:

(i) at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 17;

(ii) at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 19;

(iii) at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21;

(iv) at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23;

(v) at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at least 11

(xvi) at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23, and at least one or the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 25.

5

C39. The polypeptide of Clause C38, wherein the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise recognition helix sequences selected from:

- (i) two of SEQ ID NOs: 16 to 25;
- (ii) three of SEQ ID NOs: 16 to 25;
- 10 (iii) four of SEQ ID NOs: 16 to 25;
- (iv) five of SEQ ID NOs: 16 to 25;
- (v) six of SEQ ID NOs: 16 to 25;
- (vi) seven of SEQ ID NOs: 16 to 25;
- (vii) eight of SEQ ID NOs: 16 to 25;
- 15 (viii) nine of SEQ ID NOs: 16 to 25; or
- (ix) ten of SEQ ID NOs: 16 to 25.

C40. The polypeptide of Clause C38 or C39, wherein:

- (i) zinc finger domains F1, F3, F5, F7, F9 and F11 each has a recognition helix sequence independently selected from any one of SEQ ID NOs: 16, 18, 20, 22, 24 and 25, and zinc finger domains F2, F4, F6, F8 and F10 each has a recognition helix sequence independently selected from any one of SEQ ID NO: 17, 19, 21 and 23;
- (ii) zinc finger domains F1, F3, F4, F6, F8 and F10 each has a recognition helix sequence independently selected from any one of SEQ ID NOs: 16, 18, 20, 22, 24 and 25, and zinc finger domains 25 F2, F5, F7, F9 and F11 each has a recognition helix sequence independently selected from any one of SEQ ID NO: 17, 19, 21 and 23; or
- (iii) zinc finger domains F1, F3, F5, F6, F8 and F10 each has a recognition helix sequence independently selected from any one of SEQ ID NOs: 16, 18, 20, 22, 24 and 25, and zinc finger domains 30 F2, F4, F7, F9 and F11 each has a recognition helix sequence independently selected from any one of SEQ ID NO: 17, 19, 21 and 23.

C41. The polypeptide of any of Clauses C13 to C16, C31, C32 and C35, wherein the zinc finger peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to F11) of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 26 or 35 SEQ ID NO: 27, and wherein at least one of the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 26, and at least one or the at least 11 adjacent zinc finger domains of the zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 27.

40 C42. The polypeptide of Clause C41, wherein:

(i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences according to SEQ ID NO: 26, and zinc finger domains F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 27;

(ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 26, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 27; or

(iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 26, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 27.

C43. The polypeptide of any of Clauses C13 to C42, wherein the zinc finger peptide comprises a zinc finger array according to any of K to BZ in Table 2.

C44. The polypeptide of any of Clauses C13 to C16 or C31 to C43, wherein the zinc finger peptide comprises 9, 10, 11 or 12 zinc finger domains forming a zinc finger array according to any one of the following patterns:

	F1, F3	F2	F4, F6, F8, F10, F12	F5, F7, F9, F11
K:	SEQ ID NO: 10	SEQ ID NO: 10	SEQ ID NO: 10	SEQ ID NO: 10
L:	SEQ ID NO: 11	SEQ ID NO: 11	SEQ ID NO: 11	SEQ ID NO: 11
M:	SEQ ID NO: 12	SEQ ID NO: 12	SEQ ID NO: 12	SEQ ID NO: 12
N:	SEQ ID NO: 13	SEQ ID NO: 13	SEQ ID NO: 13	SEQ ID NO: 13
O:	SEQ ID NO: 14	SEQ ID NO: 15	SEQ ID NO: 15	SEQ ID NO: 14
P:	SEQ ID NO: 14	SEQ ID NO: 15	SEQ ID NO: 14	SEQ ID NO: 15
Q:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 16
R:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 16	SEQ ID NO: 16
S:	SEQ ID NO: 16	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 16
T:	SEQ ID NO: 16	SEQ ID NO: 19	SEQ ID NO: 16	SEQ ID NO: 19
U:	SEQ ID NO: 16	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 16
V:	SEQ ID NO: 16	SEQ ID NO: 20	SEQ ID NO: 16	SEQ ID NO: 20
W:	SEQ ID NO: 16	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 16
X:	SEQ ID NO: 16	SEQ ID NO: 22	SEQ ID NO: 16	SEQ ID NO: 22
Y:	SEQ ID NO: 16	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 16
Z:	SEQ ID NO: 16	SEQ ID NO: 24	SEQ ID NO: 16	SEQ ID NO: 24
AA:	SEQ ID NO: 16	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 16
AB:	SEQ ID NO: 16	SEQ ID NO: 25	SEQ ID NO: 16	SEQ ID NO: 25
AC:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 18
AD:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 18	SEQ ID NO: 17
AE:	SEQ ID NO: 18	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 18
AF:	SEQ ID NO: 18	SEQ ID NO: 19	SEQ ID NO: 18	SEQ ID NO: 19
AG:	SEQ ID NO: 18	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 18
AH:	SEQ ID NO: 18	SEQ ID NO: 20	SEQ ID NO: 18	SEQ ID NO: 20
AI:	SEQ ID NO: 18	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 18

	AJ:	SEQ ID NO: 18	SEQ ID NO: 22	SEQ ID NO: 18	SEQ ID NO: 22
	AK:	SEQ ID NO: 18	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 18
	AL:	SEQ ID NO: 18	SEQ ID NO: 24	SEQ ID NO: 18	SEQ ID NO: 24
	AM:	SEQ ID NO: 18	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 18
5	AN:	SEQ ID NO: 18	SEQ ID NO: 25	SEQ ID NO: 18	SEQ ID NO: 25
	AO:	SEQ ID NO: 21	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 21
	AP:	SEQ ID NO: 21	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 17
	AQ:	SEQ ID NO: 21	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 21
	AR:	SEQ ID NO: 21	SEQ ID NO: 19	SEQ ID NO: 21	SEQ ID NO: 19
10	AS:	SEQ ID NO: 21	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 21
	AT:	SEQ ID NO: 21	SEQ ID NO: 20	SEQ ID NO: 21	SEQ ID NO: 20
	AU:	SEQ ID NO: 21	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 21
	AV:	SEQ ID NO: 21	SEQ ID NO: 22	SEQ ID NO: 21	SEQ ID NO: 22
	AW:	SEQ ID NO: 21	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 21
15	AX:	SEQ ID NO: 21	SEQ ID NO: 24	SEQ ID NO: 21	SEQ ID NO: 24
	AY:	SEQ ID NO: 21	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 21
	AZ:	SEQ ID NO: 21	SEQ ID NO: 25	SEQ ID NO: 21	SEQ ID NO: 25
	BA:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 20
	BB:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 22
20	BC:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 24
	BD:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 25
	BE:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 20
	BF:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 22
	BG:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 24
25	BH:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 25
	BI:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 20
	BJ:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 22
	BK:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 24
	BL:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 25
30	BM:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 20
	BN:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 22
	BO:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 24
	BP:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 25
	BQ:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 20
35	BR:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 22
	BS:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 24
	BT:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 25
	BU:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 20
	BV:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 22
40	BW:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 24
	BX:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 25
	BY:	SEQ ID NO: 26	SEQ ID NO: 27	SEQ ID NO: 27	SEQ ID NO: 26

BZ: SEQ ID NO: 26 SEQ ID NO: 27 SEQ ID NO: 26 SEQ ID NO: 27.

C45: The polypeptide of any of Clauses C13 to C16 or C31 to C44, which comprises a sequence selected from one of SEQ ID NOs: 102 to 104, or a sequence having at least 90%, at least 95%, or at least 98%, at least 99% identity thereto.

C46. The polypeptide of any of Clauses C1 to C45, wherein the zinc finger peptide has a sequence defined by the formula:

$$N'-[(\text{Formula 4})-L_3]_{n0}-\{[(\text{Formula 6})-L_2-(\text{Formula 6})-L_3]_{n1}-[(\text{Formula 6})-L_2-(\text{Formula 6})-X_L]_{n2}-[(\text{Formula 4})-L_2-(\text{Formula 6})-L_3]_{n3}-[(\text{Formula 6})-L_2-(\text{Formula 6})]-[L_3-(\text{Formula 6})]_{n4}-C',$$

wherein n_0 is selected to be 0 or 1, n_1 is an integer selected from 1 to 4, n_2 is an integer selected from 1 or 2, n_3 is an integer selected from 1 to 4, n_4 is selected to be 0 or 1, L_2 is a linker sequence selected from the sequence $-TG^{E/Q/K/R}P-$ (SEQ ID NO: 77), L_3 is a linker sequence selected from the sequence $-TG^{G/S^E/Q/K/R}P-$ (SEQ ID NO: 79), and X_L is a linker sequence of between 8 and 50 amino acids;

Formula 4 is a zinc finger domain having a sequence selected from the formula $X_2 C X_{2 \text{ or } 4} C X_5 X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6} H X_{3,4 \text{ or } 5} H/C$ and Formula 6 is a zinc finger domain having a sequence selected from the formula $X_2 C X_2 C X_5 X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6} H X_3 H$, wherein X is an amino acid, the number in subscript indicates the number of amino acids X in that position of the sequence and $X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6}$ are the amino acids in positions -1 to 6 of each zinc finger domain recognition helix, wherein the sequence $X^{-1} X^{+1} X^{+2} X^{+3} X^{+4} X^{+5} X^{+6}$ is defined according to the zinc finger recognition helix SEQ ID NOs given in Clauses A1 to A45.

C47. The polypeptide according to any of Clauses C1 to C46, which further comprises effector domain selected from: nuclear localisation sequences, transcriptional repressor domains, transcriptional activation domains, transcriptional insulator domains, chromatin remodelling domains, chromatin condensation domains, chromatin decondensation domains, nucleic acid cleavage domains, protein cleavage domains, dimerisation domains, enzymatic domains, signalling sequences, and targeting domains.

C48. The polypeptide according to Clause C47, wherein the effector domain is a transcriptional activation domain selected from the VP64 domain (SEQ ID NO: 125), the herpes simplex virus (HSV) VP16 domain (SEQ ID NO: 124), or the p65-RelA activation domain (SEQ ID NO: 122 or SEQ ID NO: 123).

C49. The polypeptide according to Clause C47 or C48, wherein the effector domain is attached to the C-terminal end of the zinc finger peptide.

C50. The polypeptide according to any of Clauses C47 to C49, wherein the effector domain is attached to the C-terminus of the zinc finger peptide via a linker sequence comprising a sequence selected from any one or SEQ ID NOs: 126 to 130.

C51. The polypeptide according to any of Clauses C1 to A50, which comprises a nuclear localisation signal (NLS) sequence.

5 C52. The polypeptide according to Clause C51, wherein the nuclear localisation signal comprises a nuclear localisation signal from SV40, mouse primase p58, or human protein KIAA2022.

C53. The polypeptide according to Clause C51 or C52, wherein the nuclear localisation signal is selected from the SV40 NLS (SEQ ID NO: 107), the mouse primase p58 NLS (SEQ ID NO: 109), the
10 human protein KIAA2022 NLS (SEQ ID NO: 108); the double SV40 NLS (SEQ ID NO: 113), the double human KIAA2022 NLS (SEQ ID NO: 114) or the double mouse primase p58 NLS (SEQ ID NO: 115).

C54. The polypeptide according to any of Clauses C51 to C53, wherein the nuclear localisation signal is the mouse primase p58 NLS or the double mouse primase NLS.

15

C55. The polypeptide according to any of Clauses C51 to C53, wherein the nuclear localisation signal is the human KIAA2022 NLS or the double human KIAA2022 NLS.

C56. A polypeptide of any of Clauses C1 to C55, which comprises a sequence according to any of
20 SEQ ID NOS: 100, 101, 105, 106, 117 or 118, or a sequence having at least 90%, at least 95%, at least 98%, or at least 99% identity thereto.

C57. The polypeptide of any of Clauses C13 to C56, which binds to a 5'-GAA-3' trinucleotide repeat sequence containing at least 66 trinucleotide repeats, with a binding affinity stronger than about 1 μ M,
25 stronger than about 100 nM, stronger than about 10 nM, or stronger than about 1 nM.

C58. The polypeptide of any of Clauses C13 to C57, which binds to a 5'-GAA-3' trinucleotide repeat sequence containing at least 66 trinucleotide repeats in preference to or with a higher binding affinity than to a 5'-GAA-3' trinucleotide repeat sequence containing less than 12 trinucleotide repeats.

30

C59. The polypeptide of any of Clauses C13 to C57, which binds to a 5'-GAA-3' trinucleotide repeat sequence containing at least 300 trinucleotide repeats in preference to or with a higher binding affinity than to a 5'-GAA-3' trinucleotide repeat sequence containing less than 20 trinucleotide repeats.

35 C60. The polypeptide of any of Clauses C13 to C57, which binds to a 5'-GAA-3' trinucleotide repeat sequence containing at least 600 trinucleotide repeats in preference to or with a higher binding affinity than to a 5'-GAA-3' trinucleotide repeat sequence containing less than 30 trinucleotide repeats.

C61. An isolated nucleic acid encoding the polypeptide of any of Clauses C1 to C60 or a nucleic acid
40 comprising SEQ ID NO: 147 or 148.

C62. A vector comprising the nucleic acid of Clause C61.

C63. The vector according to Clause C61, which is a viral vector derived from retroviruses, such as influenza, SIV, HIV, lentivirus, and Moloney murine leukaemia; adenoviruses; adeno-associated viruses (AAV); herpes simplex virus (HSV); and chimeric viruses.

5

C64. The vector according to Clause C62 or C63, which is an adeno-associated virus (AAV) vector selected from an AAV2/1 subtype vector, or an AAV2/9 subtype vector.

10

C65. A pharmaceutical composition comprising the polypeptide according to any of Clauses C1 to C60, a nucleic acid according to Clause C61, or a vector according to any of Clauses C62 to C64.

15

C66. The polypeptide according to any of Clauses C1 to C60, the nucleic acid according to Clause C61, the vector according to any of Clauses C62 to C64, or the pharmaceutical composition according to Clause C65, for use in treating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

20

C67. The polypeptide according to any of Clauses C1 to C60, the nucleic acid according to Clause C61, the vector according to any of Clauses C62 to C64, or the pharmaceutical composition according to Clause C65, for use in preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

25

C68. The polypeptide, nucleic acid, vector or pharmaceutical composition for use according to Clause C66 or C67, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

30

C69. A method of treating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polypeptide according to any of Clauses C1 to C60, the nucleic acid according to Clause C61, the vector according to any of Clauses C62 to C64, or the pharmaceutical composition according to Clause C65.

35

C70. A method of preventing or ameliorating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polypeptide according to any of Clauses C1 to C60, the nucleic acid according to Clause A61, the vector according to any of Clauses C62 to C64, or the pharmaceutical composition according to Clause A65.

40

C71. The method of Clause C69 or C70, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

C72. A pharmaceutical composition comprising a polypeptide according to any of Clauses C1 to C12 in combination with a polypeptide according to any of Clauses C13 to C60; or a nucleic acid encoding a polypeptide according to any of Clauses C1 to C12 in combination with a nucleic acid encoding a polypeptide according to any of Clauses C13 to C60.

C73. A polypeptide according to any of Clauses C1 to C12 in combination with a polypeptide according to any of Clauses C13 to C60 for use in treating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

5

C74. A polypeptide according to any of Clauses C1 to C12 in combination with a polypeptide according to any of Clauses C13 to C60 for use in preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

10 C75. The combination of polypeptides for use according to Clause C73 or C74, wherein the polypeptide according to any of Clauses C1 to C12 is administered simultaneously, sequentially or separately from the polypeptide according to any of Clauses C13 to C60.

15 C76. A nucleic acid encoding a polypeptide according to any of Clauses C1 to C12 in combination with a nucleic acid encoding a polypeptide according to any of Clauses C13 to C60 or a nucleic acid comprising SEQ ID NO: 147 or 148 for use in treating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

20 C77. A nucleic acid encoding a polypeptide according to any of Clauses C1 to C12 in combination with a nucleic acid encoding a polypeptide according to any of Clauses C13 to C60 or a nucleic acid comprising SEQ ID NO: 147 or 148 for use in preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

25 C78. The combination of nucleic acids for use according to Clause C76 or C77, wherein the nucleic acid encoding a polypeptide according to any of Clauses C1 to C12 is administered simultaneously, sequentially or separately from the nucleic acid encoding a polypeptide according to any of Clauses C13 to C60 or a nucleic acid comprising SEQ ID NO: 147 or 148.

30 C79. A method of treating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polypeptide according to any of Clauses C1 to C12 in combination with the polypeptide according to any of Clauses C13 to C60.

35 C80. A method of preventing or ameliorating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject the polypeptide according to any of Clauses C1 to C12 in combination with the polypeptide according to any of Clauses C13 to C60.

40 C81. A method of treating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject a nucleic acid encoding the polypeptide according to any of Clauses C1 to C12 in combination with a nucleic acid encoding the polypeptide according to any of Clauses C13 to C60.

C82. A method of preventing or ameliorating a disease, disorder or condition in a subject in need thereof, the method comprising administering to the subject a nucleic acid encoding the polypeptide according to any of Clauses C1 to C12 in combination with a nucleic acid encoding the polypeptide according to any of Clauses C13 to C60 or a nucleic acid comprising SEQ ID NO: 147 or 148.

5

C83. The pharmaceutical composition according to Clause C72 for use in treating, preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal, such as a human or mouse.

10

C84. The combination of polypeptides for use according to Clauses C73 to C75, the combination of nucleic acids for use according to Clauses C76 to C78, the method of any of Clauses C79 to C82 or the pharmaceutical composition for use according to Clause C83, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

15

C85. A vector comprising a nucleic acid encoding a polypeptide according to any of Clauses C1 to C12 and a nucleic acid encoding a polypeptide according to any of Clauses C13 to C60 or a nucleic acid comprising SEQ ID NO: 147 or 148.

20

C86. The vector according to Clause C85, which is a viral vector derived from retroviruses, such as influenza, SIV, HIV, lentivirus, and Moloney murine leukaemia; adenoviruses; adeno-associated viruses (AAV); herpes simplex virus (HSV); and chimeric viruses.

25

C87. The vector according to Clause C86, which is an adeno-associated virus (AAV) vector selected from an AAV2/1 subtype vector, or an AAV2/9 subtype vector.

30

C88. A gene therapy method comprising administering to a subject in need thereof a vector according to any of Clauses C62 to C64, or C85 to C87, for use in treating, preventing or ameliorating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal.

C89. The gene therapy method according to Clause C88, wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

35

C90. The polypeptides, nucleic acids, vectors or pharmaceutical compositions for use according to Clause C68 or C84, or the methods according to any of Clauses C71, C84 or C89, wherein expression of frataxin transcript in a pathogenic target cell is increased to at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or at least 100% of the level of frataxin transcript in a wild-type, non-pathogenic cell of the host subject or animal; or compared to the level of frataxin transcript in a control subject or animal that does not have Friedreich's ataxia (FRDA).

40

D1. An isolated polynucleotide encoding a polypeptide for delivery of an effector peptide according to any of Clauses A1 to A60 or C1 to C60 to a cell different to the cell in which it was expressed; the polynucleotide comprising:

(a) sequence encoding a polypeptide, the polypeptide comprising:

(i) the effector peptide sequence comprising a polypeptide according to any of

Clauses A1 to A60 or C1 to C60;

(ii) a cell secretion peptide sequence operably linked to the effector peptide sequence;

(iii) a cell penetration peptide sequence operably linked to the effector peptide sequence; and

(b) a polypeptide expression element operable to cause the polypeptide to be expressed in a target cell *in vivo*.

D2. The polynucleotide according to Clause D1, wherein the cell secretion peptide sequence comprises a protein secretion signal (SS) from human BMP10 protein.

D3. The polynucleotide according to Clause D1 or Clause D2, wherein the cell penetration peptide sequence comprises one or more nuclear localisation signals (NLS); optionally wherein the cell penetration peptide sequence has 2, 3, 4 or 5 NLSs arranged in tandem.

D4. The polynucleotide according to any of Clauses D1 to D3, wherein the cell penetration peptide sequence comprises:

(i) the nuclear localisation sequence from SV40 virus (PKKKRKV, SEQ ID NO: 107);

(ii) the nuclear localisation sequence from human protein KIAA2022 (PKKRRKVT; NP_001008537.1, SEQ ID NO: 108); or

(iii) the nuclear localisation sequence from mouse primase p58 (RIRKKLR; GenBank: BAA04203.1, SEQ ID NO: 109).

D5. The polynucleotide according to any of Clauses D1 to D4, wherein the effector peptide comprises a transcriptional activation domain.

D6. The polynucleotide according to Clause D5, wherein the transcriptional activation domain is selected from the VP64 domain (SEQ ID NO: 125), the herpes simplex virus (HSV) VP16 domain (SEQ ID NO: 124), or the p65-RelA activation domain (SEQ ID NO: 122 or SEQ ID NO: 123).

D7. The polynucleotide according to any of Clauses D1 to D6, wherein the polypeptide expression element comprises a strong endogenous constitutive promoter and/or enhancer; preferably, wherein the polypeptide expression element comprises a constitutive promoter / enhancer sequence selected from the group consisting of: CMV (SEQ ID NO: 139), pNSE (SEQ ID NO: 142), PHSP90ab1 (SEQ ID NOs: 143 and 144), Cbh (SEQ ID NO: 140), human EF1 α -1, human synapsin promoter and pCAG-promoter.

D8. The polynucleotide of any of Clauses D1 to D7, wherein the polynucleotide encodes a polypeptide comprising the cell secretion peptide arranged N-terminal to the cell penetration peptide, and the cell penetration peptide arranged N-terminal to the effector peptide.

5 D9. The polynucleotide of Clause D8, which encodes a peptide cleavage sequence arranged between the cell secretion peptide and the cell penetration peptide.

D10. The polynucleotide of Clause D9, wherein the peptide cleavage sequence comprises the RIRR amino acid cleavage site (SEQ ID NO: 132).

10

D11. The polynucleotide of any of Clauses D1 to D10, wherein the cell secretion peptide comprises the amino acid sequence of MGSLVLTLCALFCLAAYLVSG (SEQ ID NO: 131)

15 D12. The polynucleotide of any of Clauses D1 to D11, wherein the cell penetration peptide comprises the amino acid sequence of PKKKRKVPKKRKV (SEQ ID NO: 113).

D13. The polynucleotide of any of Clauses D1 to D12, wherein:

20 (i) the polynucleotide encoding the cell penetration peptide comprises the nucleic acid sequence of CCGAAGAAAAACGTAAAGTGCCGAAGAAAAACGTAAAGTG (SEQ ID NO: 137);

(ii) the polynucleotide encoding the cell secretion peptide comprises the nucleic acid sequence of

ATGGGCTCTCTGGTCCTGACACTGTGCGCTCTTTTCTGCCTGGCAGCTTACTTGGTTTCTGGC (SEQ ID NO: 133); and/or

25 (iii) the polynucleotide encoding the RIRR amino acid cleavage site comprises the nucleic acid sequence of CGAATCAGAAGG (SEQ ID NO: 135).

D14. The polynucleotide of any of Clauses D1 to D13, wherein the effector peptide comprises a peptide according to any of Clauses A1 to A60.

30 D15. The polynucleotide of any of Clauses D1 to D13, wherein the effector peptide comprises a peptide according to any of Clauses C1 to C60.

35 D16. The polynucleotide according to any of Clauses D1 to D15, which encodes a polypeptide comprising the sequence of any of SEQ ID NOs: 97 to 106 or 116 to 121 or a sequence having at least 90%, at least 95%, at least 98%, or at least 99% identity thereto.

D17. A vector comprising the nucleic acid of any of Clauses D1 to D16.

40 D18. The vector according to Clause D17, which is a viral vector derived from retroviruses, such as influenza, SIV, HIV, lentivirus, and Moloney murine leukaemia; adenoviruses; adeno-associated viruses (AAV); herpes simplex virus (HSV); and chimeric viruses.

D19. The vector according to Clause D18, which is an adeno-associated virus (AAV) vector; optionally wherein the AAV vector is an AAV2/1 subtype vector; or an AAV2/9 subtype vector; preferably wherein the AAV vector is an AAV2/1 subtype vector.

5 D20. A polypeptide encoded by the polynucleotide or vector of any of Clauses D1 to D19.

D21. A method for delivery of a biological effector moiety comprising a polypeptide according to any of Clauses A1 to A60 or C1 to C60 to a target cell in which it was not expressed or which cell does not comprise a nucleic acid expression sequence for the biological effector moiety, the method comprising:

10 (i) providing a nucleic acid expression construct encoding an expressible biological effector peptide, the biological effector peptide adapted for cell secretion from a first target cell and cell penetration of a second target cell, wherein the first and second target cells may be of the same type or of different types;

(ii) delivering the nucleic acid expression construct to the first target cell;

15 (iii) expressing the expressible biological effector peptide in the first target cell and allowing it to be secreted from the first target cell;

(iv) bringing the secreted biological effector peptide into contact with a second target cell under conditions that allow the biological effector peptide to penetrate the second target cell;

20 thereby to deliver the biological effector moiety comprising the polypeptide according to any of Clauses A1 to A60 or C1 to C60 to the target cell.

D22. The method of Clause D21, wherein the method is performed *in vivo* or *in vitro*.

25 D23. The method of Clause D21 or Clause D22, wherein the biological effector moiety comprises a polypeptide as defined in any of Clauses A1 to A60.

D24. The method of Clause D21 or Clause D22, wherein the biological effector moiety comprises a polypeptide as defined in any of Clauses C1 to C60.

CLAIMS

1. A polynucleotide encoding a polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within a 5'-GAA-3' trinucleotide repeat sequence, frameshift variants therefore (i.e. 5'-AGA-3' or 5'-AAG-3') or a nucleic acid sequence complementary thereto, and a transcriptional activation domain.
2. A polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within a 5'-GAA-3' trinucleotide repeat sequence, frameshift variants therefore (i.e. 5'-AGA-3' or 5'-AAG-3') or a nucleic acid sequence complementary thereto, and a transcriptional activation domain.
3. The polynucleotide according to Claim 1 or the polypeptide according to Claim 2, wherein the poly-zinc finger peptide comprises from 6 to 32, from 6 to 18, from 6 to 12 zinc finger domains, and/or comprises 6, 9, 11 or 12 zinc finger domains.
4. The polynucleotide according to Claim 1 or Claim 3 or the polypeptide according to Claim 2 or Claim 3, wherein:
- (i) at least 6 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 10;
 - (ii) at least 6 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising: SEQ ID NO: 11; SEQ ID NO: 12; or SEQ ID NO: 13;
 - (iii) at least 6, at least 9, or at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 11;
 - (iv) at least 6, at least 9, or at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 12; or
 - (v) at least 6, at least 9, or at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 13.
5. The polynucleotide according to any of Claims 1, 3 and 4 or the polypeptide according to any of Claims 2 to 4, wherein the poly-zinc finger peptide comprises at least 11 zinc finger domains, and at least 11 adjacent zinc finger domains (F1 to F11) of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 14 or SEQ ID NO: 15, and wherein at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 14, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 15.
6. The polynucleotide according to any of Claims 1 and 3 to 5 or the polypeptide according to any of Claims 2 to 5, wherein:
- (i) zinc finger domains F1, F3, F5, F7, F9 and F11 have recognition helix sequences according to SEQ ID NO: 14, and zinc finger domains F2, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 15;

(ii) zinc finger domains F1, F3, F4, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 14, and zinc finger domains F2, F5, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 15; or

5 (iii) zinc finger domains F1, F3, F5, F6, F8 and F10 have recognition sequences according to SEQ ID NO: 14, and zinc finger domains F2, F4, F7, F9 and F11 have recognition sequences according to SEQ ID NO: 15.

7. The polynucleotide according to any of Claims 1, 3 and 4 or the polypeptide according to any of Claims 2 to 4, wherein the poly-zinc finger peptide comprises at least 11 zinc finger domains, and
10 each of the at least 11 adjacent zinc finger domains (F1 to F11) of the poly-zinc finger peptide comprise a recognition helix sequence independently selected from any one of SEQ ID NO: 16 to 25, and wherein:

(i) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix
15 sequence comprising SEQ ID NO: 17;

(ii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 19;

20 (iii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21;

(iv) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide
25 comprise a recognition helix sequence comprising SEQ ID NO: 16, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23;

(v) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide
30 comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 17;

(vi) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide
35 comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 19;

(vii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide
comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21;

40 (viii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide
comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at

least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23;

(ix) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21, and at least one or the at

least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 20;

(x) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21, and at least one or the at

least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 22;

(xi) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 21, and at least one or the at

least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 24;

(xii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 18, and at least one or the at

least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 25;

(xiii) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23, and at least one or the at

least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 20;

(xiv) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23, and at least one or the at

least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 22;

(xv) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23, and at least one or the at

least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 24;

(xvi) at least one of the at least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 23, and at least one or the at

least 11 adjacent zinc finger domains of the poly-zinc finger peptide comprise a recognition helix sequence comprising SEQ ID NO: 25.

8. The polynucleotide according to any of Claims 1 and 3 to 7 or the polypeptide according to any of Claims 2 to 7, wherein the poly-zinc finger peptide comprises 9, 10, 11 or 12 zinc finger domains forming a zinc finger array according to any one of the following patterns:

	F1, F3	F2	F4, F6, F8, F10, F12	F5, F7, F9, F11
K:	SEQ ID NO: 10	SEQ ID NO: 10	SEQ ID NO: 10	SEQ ID NO: 10
L:	SEQ ID NO: 11	SEQ ID NO: 11	SEQ ID NO: 11	SEQ ID NO: 11
M:	SEQ ID NO: 12	SEQ ID NO: 12	SEQ ID NO: 12	SEQ ID NO: 12

	N:	SEQ ID NO: 13	SEQ ID NO: 13	SEQ ID NO: 13	SEQ ID NO: 13
	O:	SEQ ID NO: 14	SEQ ID NO: 15	SEQ ID NO: 15	SEQ ID NO: 14
	P:	SEQ ID NO: 14	SEQ ID NO: 15	SEQ ID NO: 14	SEQ ID NO: 15
	Q:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 16
5	R:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 16	SEQ ID NO: 16
	S:	SEQ ID NO: 16	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 16
	T:	SEQ ID NO: 16	SEQ ID NO: 19	SEQ ID NO: 16	SEQ ID NO: 19
	U:	SEQ ID NO: 16	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 16
	V:	SEQ ID NO: 16	SEQ ID NO: 20	SEQ ID NO: 16	SEQ ID NO: 20
10	W:	SEQ ID NO: 16	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 16
	X:	SEQ ID NO: 16	SEQ ID NO: 22	SEQ ID NO: 16	SEQ ID NO: 22
	Y:	SEQ ID NO: 16	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 16
	Z:	SEQ ID NO: 16	SEQ ID NO: 24	SEQ ID NO: 16	SEQ ID NO: 24
	AA:	SEQ ID NO: 16	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 16
15	AB:	SEQ ID NO: 16	SEQ ID NO: 25	SEQ ID NO: 16	SEQ ID NO: 25
	AC:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 18
	AD:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 18	SEQ ID NO: 17
	AE:	SEQ ID NO: 18	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 18
	AF:	SEQ ID NO: 18	SEQ ID NO: 19	SEQ ID NO: 18	SEQ ID NO: 19
20	AG:	SEQ ID NO: 18	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 18
	AH:	SEQ ID NO: 18	SEQ ID NO: 20	SEQ ID NO: 18	SEQ ID NO: 20
	AI:	SEQ ID NO: 18	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 18
	AJ:	SEQ ID NO: 18	SEQ ID NO: 22	SEQ ID NO: 18	SEQ ID NO: 22
	AK:	SEQ ID NO: 18	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 18
25	AL:	SEQ ID NO: 18	SEQ ID NO: 24	SEQ ID NO: 18	SEQ ID NO: 24
	AM:	SEQ ID NO: 18	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 18
	AN:	SEQ ID NO: 18	SEQ ID NO: 25	SEQ ID NO: 18	SEQ ID NO: 25
	AO:	SEQ ID NO: 21	SEQ ID NO: 17	SEQ ID NO: 17	SEQ ID NO: 21
	AP:	SEQ ID NO: 21	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 17
30	AQ:	SEQ ID NO: 21	SEQ ID NO: 19	SEQ ID NO: 19	SEQ ID NO: 21
	AR:	SEQ ID NO: 21	SEQ ID NO: 19	SEQ ID NO: 21	SEQ ID NO: 19
	AS:	SEQ ID NO: 21	SEQ ID NO: 20	SEQ ID NO: 20	SEQ ID NO: 21
	AT:	SEQ ID NO: 21	SEQ ID NO: 20	SEQ ID NO: 21	SEQ ID NO: 20
	AU:	SEQ ID NO: 21	SEQ ID NO: 22	SEQ ID NO: 22	SEQ ID NO: 21
35	AV:	SEQ ID NO: 21	SEQ ID NO: 22	SEQ ID NO: 21	SEQ ID NO: 22
	AW:	SEQ ID NO: 21	SEQ ID NO: 24	SEQ ID NO: 24	SEQ ID NO: 21
	AX:	SEQ ID NO: 21	SEQ ID NO: 24	SEQ ID NO: 21	SEQ ID NO: 24
	AY:	SEQ ID NO: 21	SEQ ID NO: 25	SEQ ID NO: 25	SEQ ID NO: 21
	AZ:	SEQ ID NO: 21	SEQ ID NO: 25	SEQ ID NO: 21	SEQ ID NO: 25
40	BA:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 20
	BB:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 22
	BC:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 24

	BD:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 25
	BE:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 20
	BF:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 22
	BG:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 24
5	BH:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 25
	BI:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 20
	BJ:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 22
	BK:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 24
	BL:	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 25
10	BM:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 20
	BN:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 22
	BO:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 24
	BP:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 19	SEQ ID NO: 25
	BQ:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 20
15	BR:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 22
	BS:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 24
	BT:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 21	SEQ ID NO: 25
	BU:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 20
	BV:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 22
20	BW:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 24
	BX:	SEQ ID NO: 18	SEQ ID NO: 17	SEQ ID NO: 23	SEQ ID NO: 25
	BY:	SEQ ID NO: 26	SEQ ID NO: 27	SEQ ID NO: 27	SEQ ID NO: 26
	BZ:	SEQ ID NO: 26	SEQ ID NO: 27	SEQ ID NO: 26	SEQ ID NO: 27.

25 9. The polynucleotide according to any of Claims 1 and 3 to 8 or the polypeptide according to any of Claims 2 to 8, wherein the poly-zinc finger comprises a sequence selected from one of SEQ ID NOs: 102 to 104, or a sequence having at least 90%, at least 95%, or at least 98%, at least 99% identity thereto.

30 10. The polynucleotide according to any of Claims 1 and 3 to 9 or the polypeptide according to any of Claims 2 to 9, wherein the polypeptide comprises an activation domain selected from the VP64 domain (SEQ ID NO: 125), the herpes simplex virus (HSV) VP16 domain (SEQ ID NO: 124), or the human or mouse p65-RelA activation domain (SEQ ID NO: 122 or SEQ ID NO: 123).

35 11. The polynucleotide according to any of Claims 1 and 3 to 10 or the polypeptide according to any of Claims 2 to 10, wherein the polypeptide comprises a single or double nuclear localisation signal (NLS) sequence; optionally wherein the nuclear localisation signal is selected from the SV40 NLS (SEQ ID NO: 107), the mouse primase p58 NLS (SEQ ID NO: 109), the human protein KIAA2022 NLS (SEQ ID NO: 108); the double SV40 NLS (SEQ ID NO: 113), the double human KIAA2022 NLS (SEQ ID NO: 114) or the double mouse primase p58 NLS (SEQ ID NO: 115).

40

12. A polynucleotide encoding a polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within the frataxin gene promoter of SEQ ID NO: 67, or a sequence complementary thereto, and a transcriptional activation domain.

5 13. A polypeptide comprising a poly-zinc finger peptide capable of binding to a nucleic acid target sequence within the frataxin gene promoter of SEQ ID NO: 67, or a sequence complementary thereto, and a transcriptional activation domain.

10 14. The polynucleotide according to Claim 12 or the polypeptide according to Claim 13, wherein the poly-zinc finger peptide is capable of binding to a target sequence within positions 1 to 700 of SEQ ID NO: 67, within positions 1 to 500 of SEQ ID NO: 67, within positions 1 to 400 of SEQ ID NO: 67, or within positions 20 to 380 of SEQ ID NO: 67.

15 15. The polynucleotide according to Claim 12 or Claim 14 or the polypeptide according to Claims 13 or 14, wherein the poly-zinc finger peptide is capable of binding to a target sequence within any of SEQ ID NOs: 68 to 75, or a sequence complementary thereto; particularly wherein the poly-zinc finger peptide is capable of binding to SEQ ID NOs: 72, 73, 74 or 75 or a sequence complementary thereto.

20 16. The polynucleotide according to any of Claims 12, 14 or 15 or the polypeptide according to any of Claims 13 to 15, wherein the poly-zinc finger peptide comprises 6 or more zinc finger domains; particularly wherein the poly-zinc finger peptide has 6 zinc finger domains, 9 zinc finger domains, or 11 zinc finger domains.

25 17. The polynucleotide according to any of Claims 12 or 14 to 16 or the polypeptide according to any of Claims 13 to 16, wherein the poly-zinc finger peptide comprises 6 zinc finger domains, and wherein each zinc finger domain (F1 to F6) comprises a recognition helix sequence as follows:

- (i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; and F6, SEQ ID NO: 43;
- (ii) F1, SEQ ID NO: 31; F2, SEQ ID NO: 34; F3, SEQ ID NO: 37; F4, SEQ ID NO: 40; F5, SEQ ID NO: 43; and F6, SEQ ID NO: 46;
- (iii) F1, SEQ ID NO: 34; F2, SEQ ID NO: 37; F3, SEQ ID NO: 40; F4, SEQ ID NO: 43; F5, SEQ ID NO: 46; and F6, SEQ ID NO: 49;
- (iv) F1, SEQ ID NO: 37; F2, SEQ ID NO: 40; F3, SEQ ID NO: 43; F4, SEQ ID NO: 46; F5, SEQ ID NO: 49; and F6, SEQ ID NO: 52;
- 35 (v) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; and F6, SEQ ID NO: 44;
- (vi) F1, SEQ ID NO: 32; F2, SEQ ID NO: 35; F3, SEQ ID NO: 38; F4, SEQ ID NO: 41; F5, SEQ ID NO: 44; and F6, SEQ ID NO: 47;
- (vii) F1, SEQ ID NO: 35; F2, SEQ ID NO: 38; F3, SEQ ID NO: 41; F4, SEQ ID NO: 44; F5, SEQ ID NO: 47; and F6, SEQ ID NO: 50;
- 40 (viii) F1, SEQ ID NO: 38; F2, SEQ ID NO: 41; F3, SEQ ID NO: 44; F4, SEQ ID NO: 47; F5, SEQ ID NO: 50; and F6, SEQ ID NO: 53;

- (ix) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; and F6, SEQ ID NO: 45;
- (x) F1, SEQ ID NO: 33; F2, SEQ ID NO: 36; F3, SEQ ID NO: 39; F4, SEQ ID NO: 42; F5, SEQ ID NO: 45; and F6, SEQ ID NO: 48;
- 5 (xi) F1, SEQ ID NO: 36; F2, SEQ ID NO: 39; F3, SEQ ID NO: 42; F4, SEQ ID NO: 45; F5, SEQ ID NO: 48; and F6, SEQ ID NO: 51; or
- (xii) F1, SEQ ID NO: 39; F2, SEQ ID NO: 42; F3, SEQ ID NO: 45; F4, SEQ ID NO: 48; F5, SEQ ID NO: 51; and F6, SEQ ID NO: 54.
- 10 18. The polynucleotide according to any of Claims 12 or 14 to 17 or the polypeptide according to any of Claims 13 to 17, wherein the poly-zinc finger peptide comprises 7 zinc finger domains, and wherein each zinc finger domain (F1 to F7) comprises a recognition helix sequence as follows:
- (i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; F6, SEQ ID NO: 43; and F7, SEQ ID NO: 46;
- 15 (ii) F1, SEQ ID NO: 31; F2, SEQ ID NO: 34; F3, SEQ ID NO: 37; F4, SEQ ID NO: 40; F5, SEQ ID NO: 43; F6, SEQ ID NO: 46; and F7, SEQ ID NO: 49;
- (iii) F1, SEQ ID NO: 34; F2, SEQ ID NO: 37; F3, SEQ ID NO: 40; F4, SEQ ID NO: 43; F5, SEQ ID NO: 46; F6, SEQ ID NO: 49; and F7, SEQ ID NO: 52;
- (iv) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; F6, SEQ ID NO: 44; and F7, SEQ ID NO: 47;
- 20 (v) F1, SEQ ID NO: 32; F2, SEQ ID NO: 35; F3, SEQ ID NO: 38; F4, SEQ ID NO: 41; F5, SEQ ID NO: 44; F6, SEQ ID NO: 47; and F7, SEQ ID NO: 50;
- (vi) F1, SEQ ID NO: 35; F2, SEQ ID NO: 38; F3, SEQ ID NO: 41; F4, SEQ ID NO: 44; F5, SEQ ID NO: 47; F6, SEQ ID NO: 50; and F7, SEQ ID NO: 53;
- 25 (vii) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; F6, SEQ ID NO: 45; and F7, SEQ ID NO: 48;
- (viii) F1, SEQ ID NO: 33; F2, SEQ ID NO: 36; F3, SEQ ID NO: 39; F4, SEQ ID NO: 42; F5, SEQ ID NO: 45; F6, SEQ ID NO: 48; and F7, SEQ ID NO: 51; or
- (ix) F1, SEQ ID NO: 36; F2, SEQ ID NO: 39; F3, SEQ ID NO: 42; F4, SEQ ID NO: 45; F5, SEQ ID NO: 48; F6, SEQ ID NO: 51; and F7, SEQ ID NO: 54.
- 30
19. The polynucleotide according to any of Claims 12 or 14 to 18 or the polypeptide according to any of Claims 13 to 18, wherein the poly-zinc finger peptide comprises 8 zinc finger domains, and wherein each zinc finger domain (F1 to F8) comprises a recognition helix sequence as follows:
- 35 (i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; F6, SEQ ID NO: 43; F7, SEQ ID NO: 46; and F8, SEQ ID NO: 49;
- (ii) F1, SEQ ID NO: 31; F2, SEQ ID NO: 34; F3, SEQ ID NO: 37; F4, SEQ ID NO: 40; F5, SEQ ID NO: 43; F6, SEQ ID NO: 46; F7, SEQ ID NO: 49; and F8, SEQ ID NO: 52;
- (iii) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; F6, SEQ ID NO: 44; F7, SEQ ID NO: 47; and F8, SEQ ID NO: 50;
- 40 (iv) F1, SEQ ID NO: 32; F2, SEQ ID NO: 35; F3, SEQ ID NO: 38; F4, SEQ ID NO: 41; F5, SEQ ID NO: 44; F6, SEQ ID NO: 47; F7, SEQ ID NO: 50; and F8, SEQ ID NO: 53;

- (v) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; F6, SEQ ID NO: 45; F7, SEQ ID NO: 48; and F8, SEQ ID NO: 51; or
- (vi) F1, SEQ ID NO: 33; F2, SEQ ID NO: 36; F3, SEQ ID NO: 39; F4, SEQ ID NO: 42; F5, SEQ ID NO: 45; F6, SEQ ID NO: 48; F7, SEQ ID NO: 51; and F8, SEQ ID NO: 54.

5

20. The polynucleotide according to any of Claims 12 or 14 to 19 or the polypeptide according to any of Claims 13 to 19, wherein the poly-zinc finger peptide comprises 9 zinc finger domains, and wherein each zinc finger domain (F1 to F9) comprises a recognition helix sequence as follows:

- 10 (i) F1, SEQ ID NO: 28; F2, SEQ ID NO: 31; F3, SEQ ID NO: 34; F4, SEQ ID NO: 37; F5, SEQ ID NO: 40; F6, SEQ ID NO: 43; F7, SEQ ID NO: 46; F8, SEQ ID NO: 49; and F9, SEQ ID NO: 52;
- (ii) F1, SEQ ID NO: 29; F2, SEQ ID NO: 32; F3, SEQ ID NO: 35; F4, SEQ ID NO: 38; F5, SEQ ID NO: 41; F6, SEQ ID NO: 44; F7, SEQ ID NO: 47; F8, SEQ ID NO: 50; and F9, SEQ ID NO: 53; or
- 15 (iii) F1, SEQ ID NO: 30; F2, SEQ ID NO: 33; F3, SEQ ID NO: 36; F4, SEQ ID NO: 39; F5, SEQ ID NO: 42; F6, SEQ ID NO: 45; F7, SEQ ID NO: 48; F8, SEQ ID NO: 51; and F9, SEQ ID NO: 54.

20 21. The polynucleotide according to any of Claims 12 or 14 to 20 or the polypeptide according to any of Claims 13 to 20, wherein the poly-zinc finger comprises the sequence of SEQ ID NO: 116 or a sequence having at least 90%, at least 95%, or at least 98%, at least 99% identity thereto.

25 22. A vector comprising the polynucleotide of any of Claims 1, 3 to 12 and 14 to 21, optionally which is a viral vector derived from retroviruses, adenoviruses, adeno-associated viruses (AAV) and chimeric viruses.

23. The vector according to Claim 22, which is an adeno-associated virus (AAV) vector selected from an AAV2/9 subtype vector, or an AAV2/1 subtype vector.

30 24. A pharmaceutical composition comprising a polynucleotide according to any of Claims 1, 3 to 12 and 14 to 21, the vector according to Claim 22 or Claim 23 or the polypeptide according to any of Claims 2 to 11 and 13 to 21, and a pharmaceutically acceptable carrier.

35 25. The polynucleotide according to any of Claims 1, 3 to 12 and 14 to 21, the vector according to Claim 22 or Claim 23, the polypeptide according to any of Claims 2 to 11 and 13 to 21, or the pharmaceutical composition according to Claim 24, for use in treating a disease, disorder or condition associated with pathogenic GAA-trinucleotide repeat sequences in an animal; optionally wherein the disease, disorder or condition is Friedreich's ataxia (FRDA), or a disease, disorder or condition associated therewith.

40

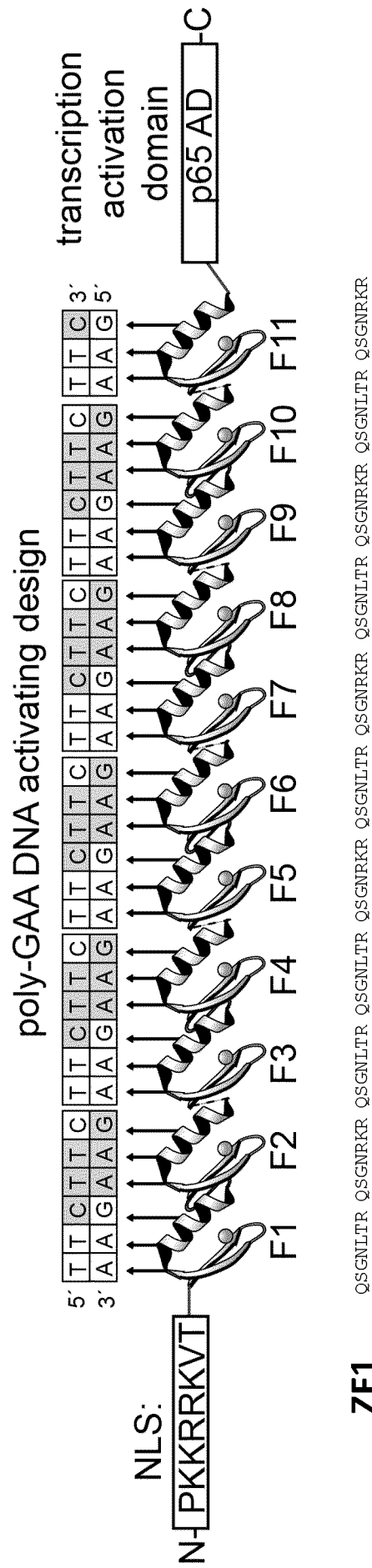
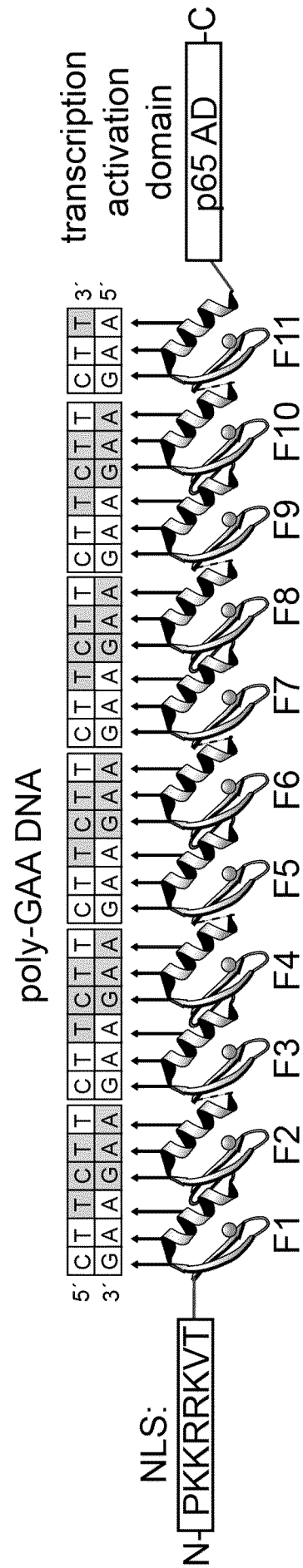


Figure 1A



ZF2

RSSNLTY RSANRKQ RSSNLSY RSANLTQ RSSNRIQ RSANLSY RSSNRTQ RSANLSQ RSSNRKY RSANLSQ RSSNRKQ

Figure 1B

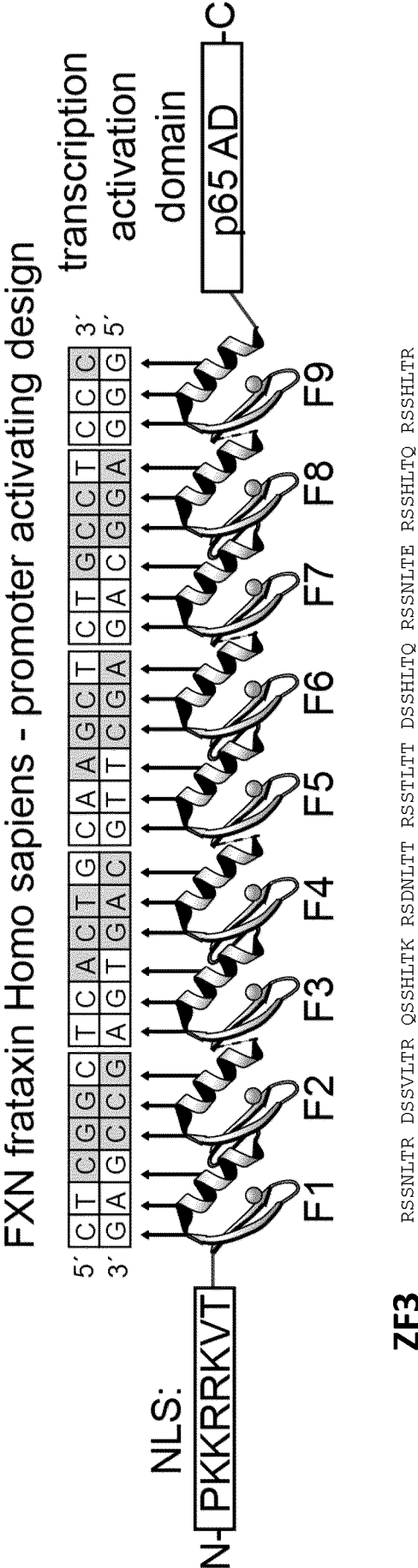


Figure 1C

FXN frataxin Homo sapiens - promoter activating design

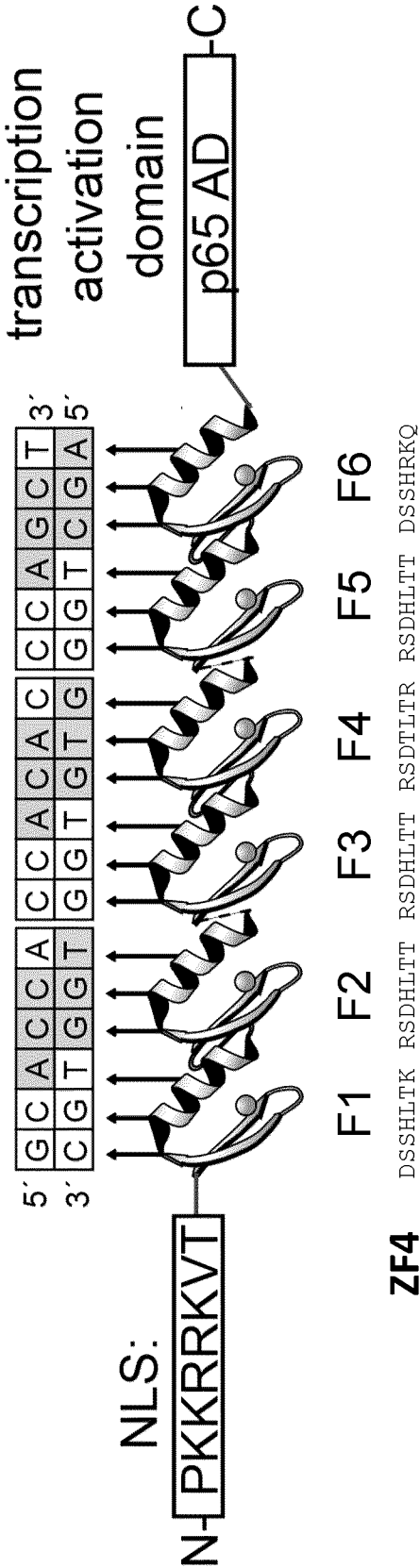


Figure 1D

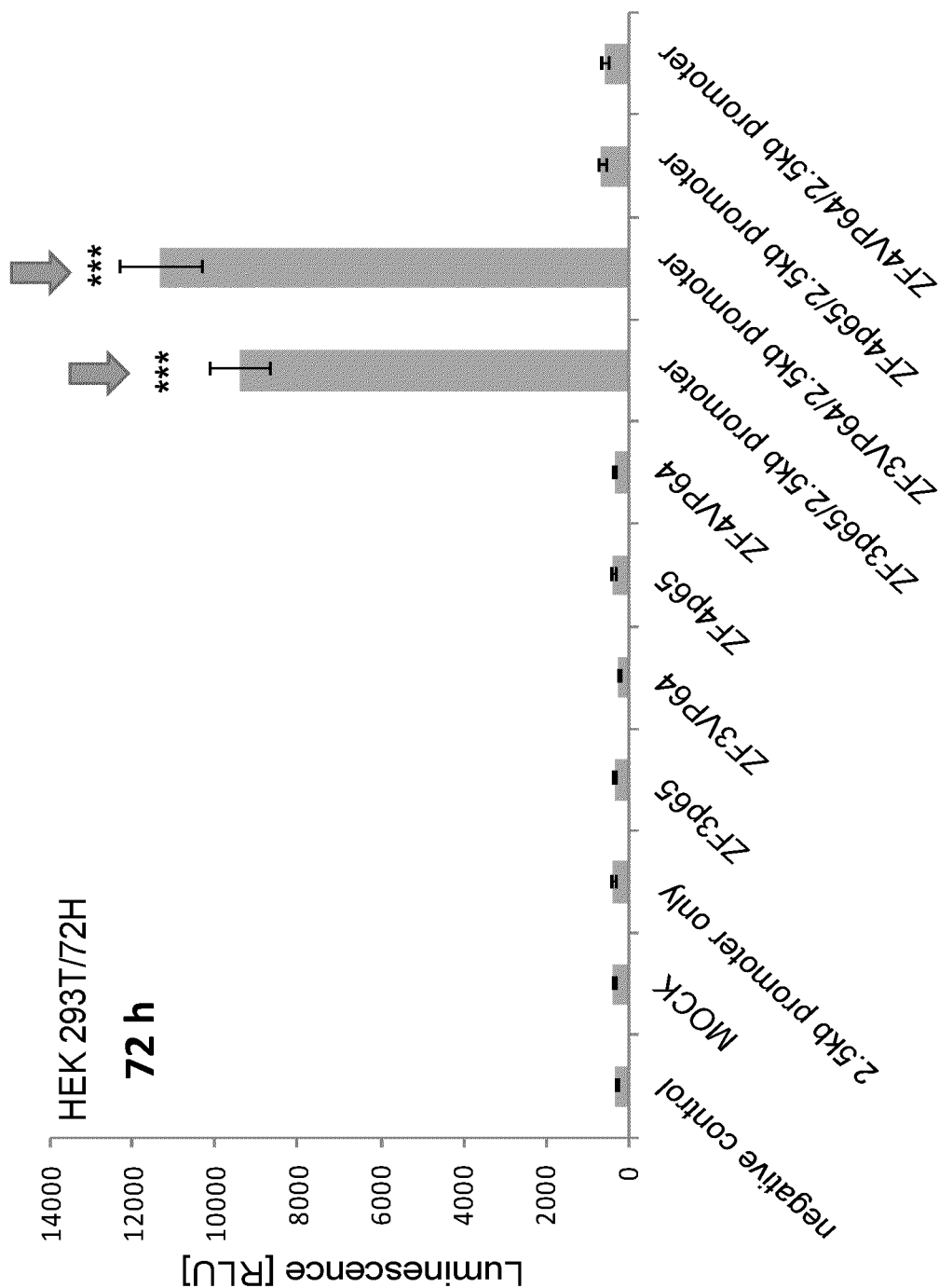


Figure 2

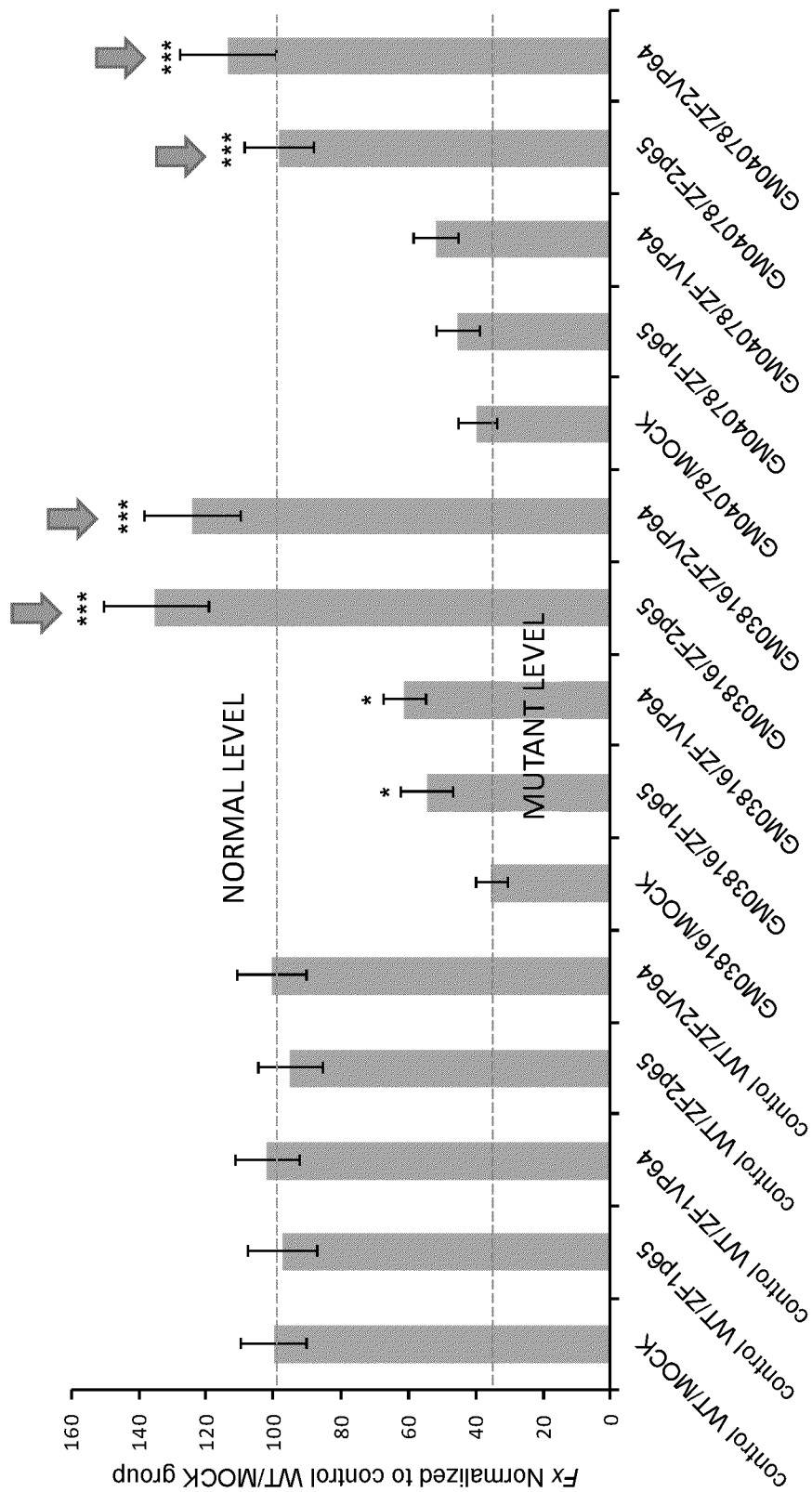


Figure 3

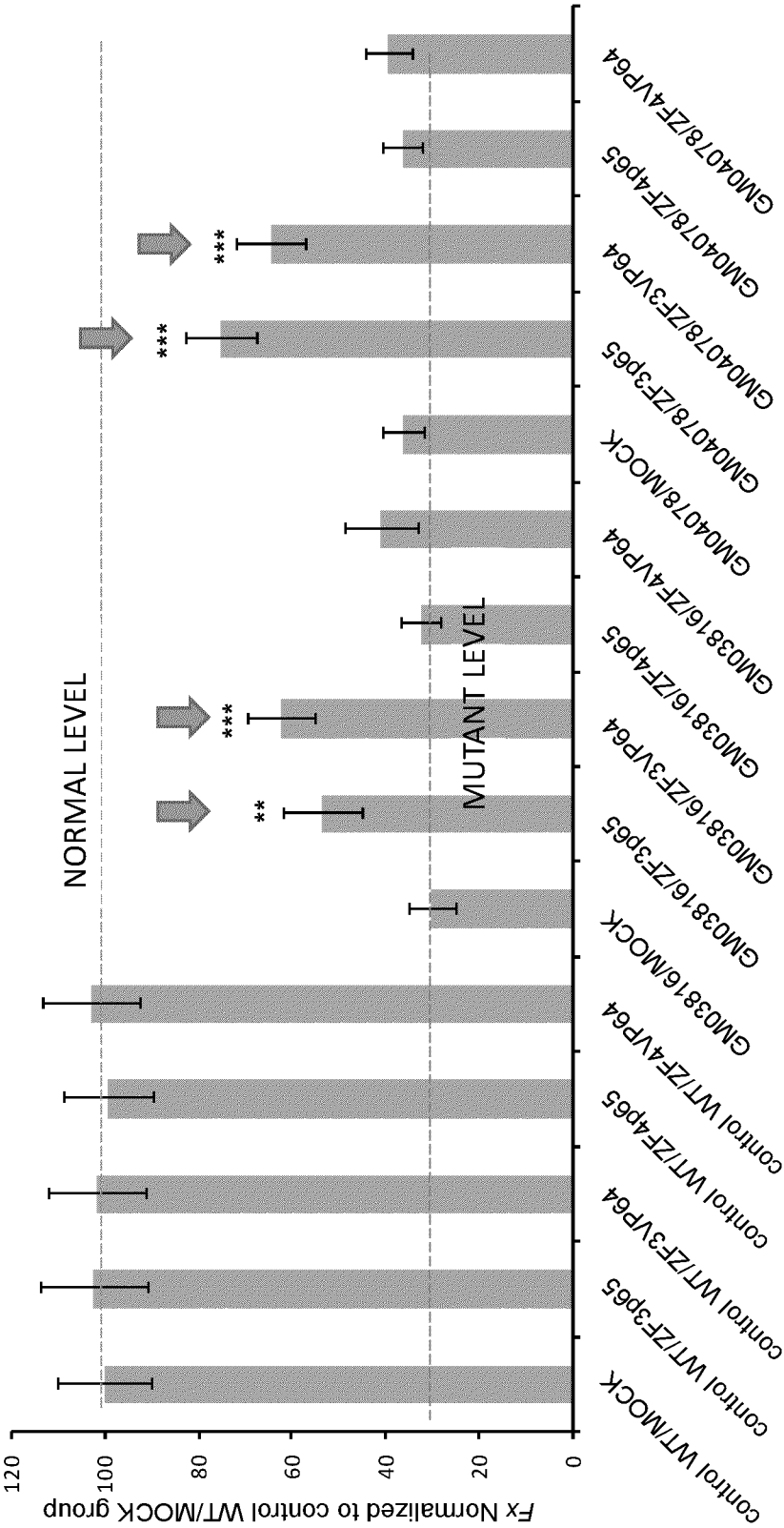


Figure 4

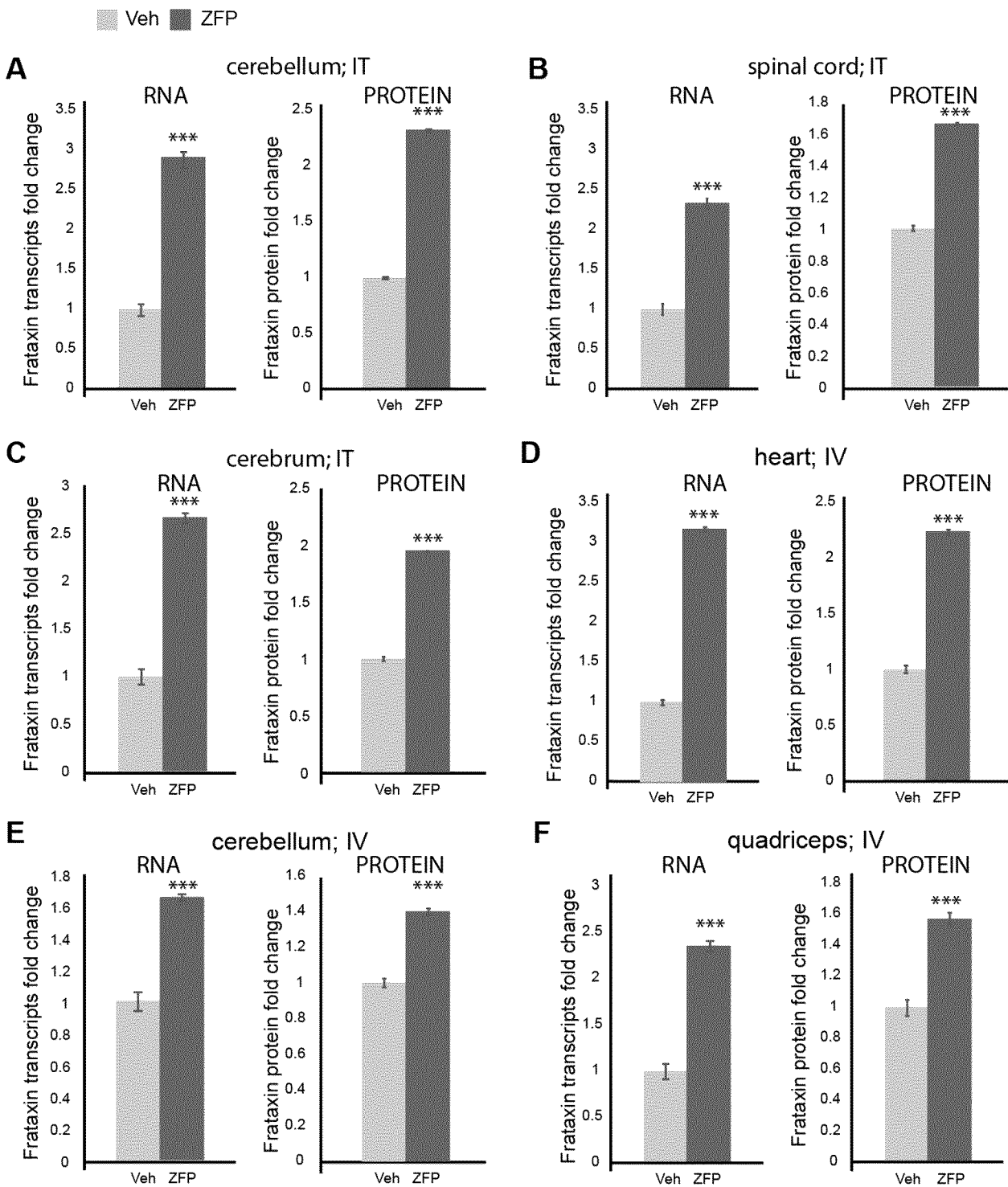


Figure 5

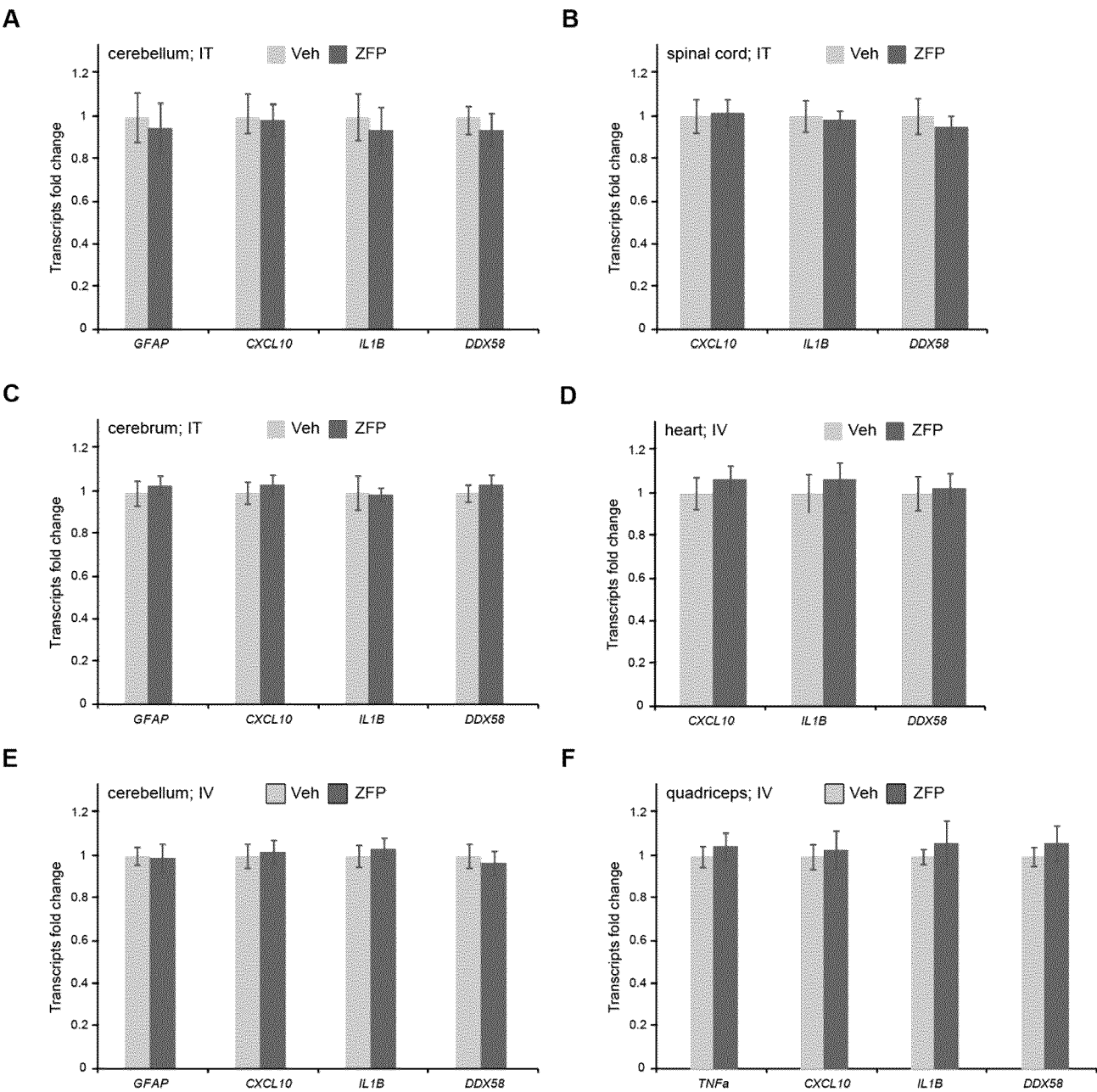


Figure 6

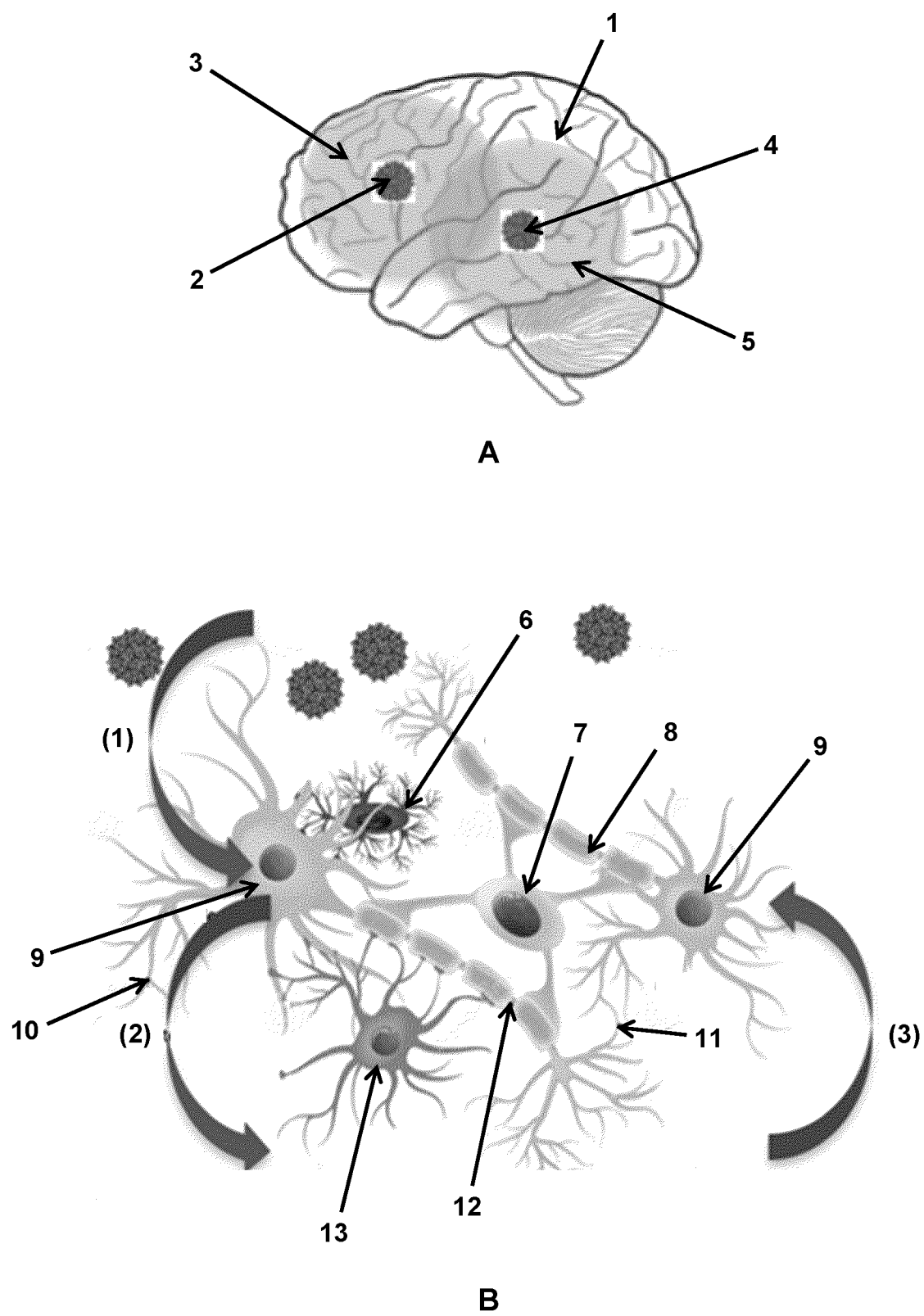


Figure 7