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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 November 2016 (03.11.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/174123 A1

(51) International Patent Classification:
A61K 38/17 (2006.01) *A61K 48/00* (2006.01)

(21) International Application Number:
PCT/EP2016/059468

(22) International Filing Date:
28 April 2016 (28.04.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
15165741.8 29 April 2015 (29.04.2015) EP

(71) Applicant: **ETH ZURICH** [CH/CH]; Raemistrasse 101 / ETH transfer, 8092 Zurich (CH).

(72) Inventors: **ALLAIN, Frédéric**; Zelglistr. 12, 8046 Zürich (CH). **CLÉRY, Antoine**; Wehntalerstrasse 283, 8046 Zürich (CH).

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

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

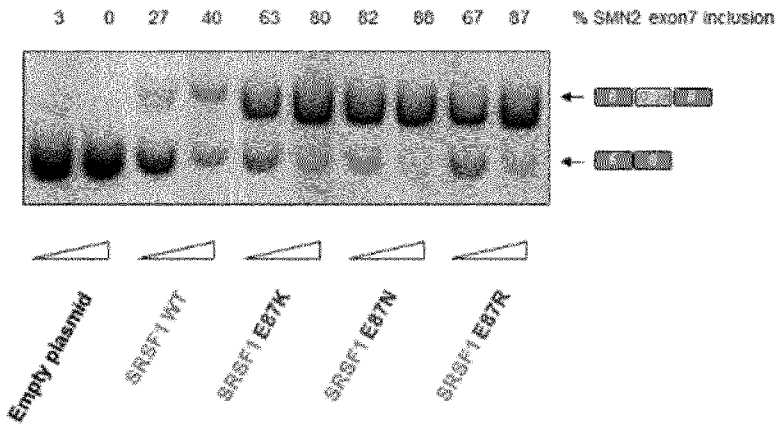
Published:

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

WO 2016/174123 A1

(54) Title: MUTANTS OF SRSF1 (ASF/SF2) FOR TREATING SPINAL MUSCULAR ATROPHY

Fig. 2



(57) Abstract: The invention relates to non-naturally occurring proteins with a splicing factor activity that are able to facilitate the inclusion of exon 7 of the SMN1 and/or SMN2 gene into the mature mRNA. The disclosed non-naturally occurring proteins share at least 70% sequence identity with hSRSF1 (ASF/SF2) and can be used in the treatment of spinal muscular atrophy.

Mutants of SRSF1 (ASF/SF2) for treating Spinal Muscular Atrophy

Description

Although spinal muscular atrophy (SMA) is considered to be a rare disease, it is the most common genetic cause of infant death, with an incidence of 1 in 10,000 live births. This disease is characterized by death of neuronal cells in the anterior horn of spinal cord and subsequent system-wide muscle wasting (atrophy). Three main degrees of severity were reported for patients with SMA: type I patients as “nonsitters”, type II patients as “sitters”, and type III patients as “walkers”. The type I, also called Werdnig-Hoffman disease, is the most severe form of SMA and by far the most common, comprising nearly 60% of all newly diagnosed cases. Infants usually develop respiratory failure and do not survive beyond age 2 years. There is to date no cure for this disease.

Spinal muscular atrophy (SMA) is an autosomal recessive disease characterized by progressive muscle wasting (atrophy) and in consequence mobility impairment or even immobility. SMA is caused by the decrease of SMN (Survival of Motor Neuron) protein levels in motor neurons.

SMN protein is encoded by two genes (SMN1 and SMN2), for which the corresponding pre-mRNAs are differentially spliced (Fig. 1). Exon 7 is very efficiently included in SMN1 mRNAs, which leads to the production of functional SMN proteins. Conversely, exon 7 is primarily skipped in SMN2 mRNAs, which are translated into unstable and therefore non-functional SMN proteins. This difference in splicing is due to a single nucleotide difference at position +6 of exon 7: in SMN1, a cytosine is present and recruits the splicing activator Serine/arginine-rich splicing factor 1 (SRSF1), whereas in SMN2 a uridine is found at this position, which prevents the binding of SRSF1 and promotes instead the recruitment of the splicing repressor hnRNP A1 (Fig. 1).

In individuals affected by SMA, the SMN1 gene is mutated resulting in the expression of non-functional SMN protein. Although the usually unaffected SMN2 gene still allows the production of small amounts of SMN protein (10-20% of the normal level), this reduced availability of the SMN protein is insufficient to maintain motor neuron viability and normally results in gradual death of motor neuron cells in the anterior horn of the spinal cord and the brain. Muscles that depend on these motor neurons for neural input now have decreased innervation. These denervated muscles undergo progressive atrophy.

The problem underlying the present invention is to provide the means to increase levels of functional SMN protein to treat or prevent SMA. This problem is solved by the subject-matter of the independent claims.

The disclosed invention consists of engineering SRSF1 to force it to recognize the uridine at the position +6 of SMN2 exon 7 instead of hnRNP A1 and reverse the splicing outcome.

Terms and definitions

Amino acid sequences are given from amino to carboxyl terminus.

Identity in the context of the present specification is a single quantitative parameter representing the result of a sequence comparison position by position. Methods of sequence comparison are known in the art; the BLAST algorithm available publicly is an example.

One such example for comparison of nucleic acid sequences is the BLASTN algorithm that uses the default settings: Expect threshold: 10; Word size: 28; Max matches in a query range: 0; Match/Mismatch Scores: 1.-2; Gap costs: Linear

One such example for comparison of amino acid sequences is the BLASTP algorithm that uses default settings such as: Expect threshold: 10; Word size: 3; Max matches in a query range: 0; Matrix: BLOSUM62; Gap Costs: Existence 11, Extension 1; Compositional adjustments: Conditional compositional score matrix adjustment.

Unless otherwise stated, sequence identity/similarity values provided herein refer to the value obtained using the BLAST suite of programs using default parameters detailed above.

In the context of the present specification, the term *spinal muscular atrophy (SMA)* is used in its meaning known in the art of medicine; it refers to a genetic disorder caused by a defect in the SMN1 (Survival motor neuron protein) gene as explained in greater detail above.

In the context of the present specification, the term *binding site* is used in its meaning known in the art of molecular biology; it refers to a specific nucleotide or amino acid sequence of a nucleic acid molecule or protein, which is the site of binding for a protein being able to specifically recognize this sequence.

In the context of the present specification, the term *non-naturally occurring protein* refers to artificially designed or engineered proteins that are different from naturally occurring proteins. The term thereby excludes all proteins that are natural products.

In the context of the present specification, the term *recombinant*, refers to a particular nucleic acid (DNA or RNA) that is the product of various combinations of cloning, restriction, polymerase chain reaction (PCR) and/or ligation steps resulting in a construct having a structural coding or non-coding sequence distinguishable from endogenous nucleic acids found in natural systems.

In the context of the present specification, the term *expression vector* is used in its meaning known in the art of molecular biology and genetics; it refers to a DNA construct designed to introduce a specific genetic sequence into a target cell in order to express the introduced

genetic sequence. An expression vector comprises a genetic sequence to be introduced into the target cells and a promoter sequence which is operably linked to the genetic sequence.

In the context of the present specification the term *viral vector* is used in its meaning known in the art of molecular biology and genetics. It includes but is not limited to expression vectors comprising sequences of retroviruses, adenoviruses, adeno-associated viruses and herpes simplex virus.

In the context of the present specification, the terms *exon* and *intron* are used in their meaning known in the art of molecular biology and genetics; the term exon refers to any nucleotide sequence encoded by a gene that remains part of the final mature RNA product of that gene. The term exon refers to both the DNA sequence within a gene and to the corresponding sequence in RNA transcripts. The nucleotide sequences removed from the pre-messenger RNA by RNA splicing are referred to as introns. In RNA splicing, introns are removed and exons are covalently joined to one another as part of generating the mature messenger RNA.

In the context of the present specification, the terms *RNA splicing* or *splicing* are used in their meaning known in the art of molecular biology and genetics; they refer to a modification of pre-messenger RNA transcripts in which non-coding introns are removed and coding exons are joined. This process gives rise to mature messenger RNA (mRNA), which is translated into proteins.

In the context of the present specification, the term *splicing factor* is used in its meaning known in the art of molecular biology and genetics; it refers to a protein with the ability to regulate the splicing outcome. Splicing factors bind the pre-messenger RNA and recruit the spliceosome, which catalyse the actual splicing/cleavage and ligation reaction. However, some splicing factors can also bind the pre-messenger RNA to prevent the recruitment of the spliceosome and force an exon to be skipped from the messenger RNA instead of being ligated with the other exons.

According to a first aspect of the invention a non-naturally occurring protein is provided. The protein has an identity of ≥70%, ≥85%, ≥90%, ≥95%, or ≥99% compared to SEQ ID NO 001 (SRSF1-WT). It is characterized by the following attributes:

The protein

i. has a splicing factor activity,

ii. can be recruited to exon 7 of the SMN1 and/or SMN2 gene, and

iii. facilitates the inclusion of exon 7 of the SMN1 and/or SMN2 gene into the mature SMN1 and/or SMN2 mRNA.

One example of a method known in the art to determine if a protein has the above mentioned attributes is an in vivo splicing assay as disclosed in the examples of this specification. Fig. 2 provides an example of this in vivo splicing assay. In this example the intensity of the upper band provides an indication of the efficiency of the splicing factor activity of the tested protein in including exon 7 in the resulting nucleic acid molecule. By using sequences that contain exon 7 of the SMN1 and/or SMN2 gene, one can simultaneously test if the splicing factor can be successfully recruited to this particular exon. The molecular weight of the resulting bands also provides an indication of the exons that have been joined together as depicted in Fig. 2.

One method to identify proteins with a certain percentage of sequence identity to a given sequence is the use of the BLAST algorithm. In other words proteins that are identified by using SEQ ID NO 001 in a BLAST search that fall within the identity limits set out above and are not synthetic would not be part of the first aspect of the invention.

In certain embodiments the non-naturally occurring protein of the invention is characterized in that the glutamate on position 87 of SEQ ID NO 001 is substituted by an asparagine, lysine, arginine, glutamine or histidine.

The glutamate located on position 87 in SEQ ID NO 001 might not be in the same position in other sequences that fall within the boundaries laid out in the first aspect of the invention and therefore the corresponding position for changes in the amino acid sequences has to be identified. A method to determine the corresponding position in these sequences would be the alignment of sequences. In aligned sequences the corresponding position of the glutamate on position 87 in SEQ ID NO 001 can be easily determined by comparison of the sequence aligned to position 87 of SEQ ID NO 001.

Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman, Adv. Appl. Math. 2:482 (1981); by the homology alignment algorithm of Needleman and Wunsch, J. Mol. Biol. 48:443 (1970); by the search for similarity method of Pearson and Lipman, Proc. Nat. Acad. Sci. 85:2444 (1988); or by computerized implementations of these algorithms.

In certain embodiments the non-naturally occurring protein is selected from SEQ ID NO 002 (E87N), SEQ ID NO 003 (E87K), SEQ ID NO 004 (E87R), SEQ ID NO 005 (E87Q) or SEQ ID NO 006 (E87H).

In certain embodiments the binding site of the non-naturally occurring protein comprises position +6 of exon 7 of the SMN1 (SEQ ID NO 007) and/or SMN2 gene (SEQ ID NO 008).

According to a second aspect of the invention a non-naturally occurring protein selected from SEQ ID NO 002 (E87N), SEQ ID NO 003 (E87K), SEQ ID NO 004 (E87R), SEQ ID NO 005 (E87Q) or SEQ ID NO 006 (E87H) is provided.

According to a third aspect of the invention a non-naturally occurring protein according to the first aspect of the invention is provided for the use as a pharmaceutical.

In certain embodiments the non-naturally occurring protein is provided for use in a treatment of spinal muscular atrophy (SMA).

In certain embodiments the nucleic acid molecule according to the fourth aspect of the invention and/or the expression vector according to the fifth aspect of the invention are provided for the use as a pharmaceutical.

According to a fourth aspect of the invention a nucleic acid molecule encoding a non-naturally occurring protein according to the first aspect of the invention is provided.

According to a fifth aspect of the invention a recombinant expression vector comprising the nucleic acid molecule according to the third aspect of the invention is provided.

In certain embodiments the nucleotide acid molecule encoding the polypeptide is operably linked to a promoter. Examples of commonly used promoters are the cytomegalovirus (CMV) and Simian virus 40 (SV40) promoters.

In certain embodiments the recombinant expression vector is a viral vector.

According to a sixth aspect of the invention a transgenic cell comprising the nucleic acid molecule according to the third aspect of the invention is provided.

According to a seventh aspect of the invention a nucleic acid molecule encoding a non-naturally occurring protein according to the first aspect of the invention for use in a treatment of spinal muscular atrophy (SMA) is provided.

According to an eighth aspect of the invention a recombinant expression vector comprising the nucleic acid molecule according to the third aspect of the invention for use in a treatment of spinal muscular atrophy (SMA) is provided.

According to a ninth aspect of the invention a transgenic cell comprising the nucleic acid molecule according to the third aspect of the invention for use in a treatment of spinal muscular atrophy (SMA) is provided.

The invention is further illustrated by the following examples and figures, from which further embodiments and advantages can be drawn. These examples are meant to illustrate the invention but not to limit its scope.

Short description of the figures

Fig. 1 shows effects of SRSF1 and hnRNP A1 on SMN exon 7 splicing. SRSF1 binds the cytosine (C) located at position +6 of SMN1 exon 7 and activates inclusion of the exon in the mature SMN1 mRNA, which is translated in stable and functional SMN protein. HnRNP A1 binds the uridine (U) located at the corresponding position of SMN2 exon 7 and primarily induces the skipping of the exon in the mature SMN2 mRNA, which is translated into a truncated, unstable and therefore non-functional SMN protein.

Fig. 2 shows effects of WT and mutated versions of SRSF1 on SMN2 exon 7 splicing. HEK293 cells were cotransfected with 1 µg of pCI-SMN2 plasmid and 1 or 3 ug of the mammalian vector pCGT7 expressing WT or mutated versions of SRSF1 full length proteins. The mutated SRSF1 proteins E87K, E87N and E87R significantly increase the inclusion of exon 7 in mature SMN2 mRNAs.

Fig. 3 shows a schematic representation of the possible involvement of Glu87 in SRSF1 RRM1 recognition of the cytosine located at position +6 of SMN1 exon 7 (left part of the panel). In this context, the mutation of Glu87 to an asparagine residue would still allow this interaction in addition to promote the recognition of the uridine present at position +6 of SMN2 exon 7 (right part of the panel).

Examples

Structural investigation of the splicing regulator SRSF1 revealed that the glutamate 87 located at the end of the beta4 strand of RRM1 is essential for the recognition of the cytosine located at the position +6 of SMN1 exon 7 (Figure 3). A single amino-acid substitution of this glutamate to an asparagine, lysine or arginine enlarges the initial SRSF1 binding specificity from a cytosine to the recognition of cytosine and uridine nucleotides. This mutation does not prevent the interaction of SRSF1 with its cytosine-containing natural targets in cells, while forcing its binding to the uridine located at position +6 of SMN2 exon 7. SRSF1 protein variants E87N, E87R and E87K were co-transfected in HEK293 cells with a synthetic SMN2 minigene and found that these proteins can promote up to 90% of exon 7 inclusion (Figure 2).

Methods

Cell culture and plasmids

Human embryonic kidney HEK293 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) (GibcoBRL) supplemented with 10% fetal calf serum (GibcoBRL). These

cells were cotransfected with 1ug of pCI-SMN2 recombinant plasmid containing the SMN2 minigene (given by Stefan Stamm's lab: Department of Molecular and Cellular Biochemistry, Biomedical Biological Sciences Research Building (BBSRB), University of Kentucky, Lexington, USA) and 1 or 3 ug of the mammalian vector pCGT7 expressing WT or mutated versions of SRSF1 full length proteins. The mammalian vector expressing pCGT7::SRSF1 WT was provided by Adrian Krainer's lab (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA) and the vectors pCGT7::SRSF1 E87N, pCGT7::SRSF1 E87K and pCGT7::SRSF1 E87R were created by site directed mutagenesis using specific primers by Dr. Antoine Cléry in Prof. Frédéric Allain's lab (ETH Zürich).

In vivo splicing assay

HEK293 cells were plated on 6-well plates. For each transfection, 1 µg of pCI-SMN2 plasmid was cotransfected with 1 or 3 ug of the mammalian vector pCGT7 expressing WT or mutated versions (E87N, E87R, E87K) of SRSF1 full length proteins. Equal amounts of DNA were transfected by addition of empty vector when necessary. 48 hours after transfection, total RNA was isolated. 1 µg total RNA was reversed transcribed using oligo(dT) and M-MuLV Reverse Transcriptase RNAseH- (Finnzyme). 1/10 of the cDNA was amplified by PCR using a vector-specific forward primer (pCI-forw: 5'-GGTGTCCACTCCCAGTTCAA; SEQ ID NO 009) and the SMN-specific reverse primer SMNex8-rev (SMNex8-rev: 5'-GCCTCACCACCGTGCTGG; SEQ ID NO 010); 24 PCR cycles using these 3 steps: 94 °C, 30 sec; 55 °C, 30 sec; 72 °C, 60 sec. PCR products were then resolved on a 1.5% agarose gel. The ratio of exon inclusion to exon skipping was determined by using the ImageQuant software.

Sequences:

SEQ ID NO 001 (hSRSF1):

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFVEFEDPRDAE
DAVYGRDGYDYDGYRLRV**E**FPRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSENRVVVSGLPPSGSWQD
LKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTYAVRKLDNTKFRSHEGETAYIRVKVDGPRSPSYGR
SRSRSRSRSRSRSNSRSRSYSPRRSRGSPRYSPRHSRSRSRT

SEQ ID NO 002 (E87N):

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFVEFEDPRDAE
DAVYGRDGYDYDGYRLRV**N**FPRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSENRVVVSGLPPSGSWQD
LKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTYAVRKLDNTKFRSHEGETAYIRVKVDGPRSPSYGR
SRSRSRSRSRSRSNSRSRSYSPRRSRGSPRYSPRHSRSRSRT

SEQ ID NO 003 (E87K):

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFVEFEDPRDAE
DAVYGRDGYDYDGYRLRV**K**FPRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSENRVVVSGLPPSGSWQD
LKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTYAVRKLDNTKFRSHEGETAYIRVKVDGPRSPSYGR
SRSRSRSRSRSRSRSNSRSRSYSPRRSRGSPRYSPRHSRSRSRT

SEQ ID NO 004 (E87R):

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFVEFEDPRDAE
DAVYGRDGYDYDGYRLRV**R**FPRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSENRVVVSGLPPSGSWQD
LKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTYAVRKLDNTKFRSHEGETAYIRVKVDGPRSPSYGR
SRSRSRSRSRSRSRSNSRSRSYSPRRSRGSPRYSPRHSRSRSRT

SEQ ID NO 005 (E87Q):

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFVEFEDPRDAE
DAVYGRDGYDYDGYRLRV**Q**FPRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSENRVVVSGLPPSGSWQD
LKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTYAVRKLDNTKFRSHEGETAYIRVKVDGPRSPSYGR
SRSRSRSRSRSRSRSNSRSRSYSPRRSRGSPRYSPRHSRSRSRT

SEQ ID NO 006 (E87H):

MSGGGVIRGPAGNNDCRIYVGNLPPDIRTKDIEDVFYKYGAIRDIDLKNRRGGPPFAFVEFEDPRDAE
DAVYGRDGYDYDGYRLRV**H**FPRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSENRVVVSGLPPSGSWQD
LKDHMREAGDVCYADVYRDGTGVVEFVRKEDMTYAVRKLDNTKFRSHEGETAYIRVKVDGPRSPSYGR
SRSRSRSRSRSRSRSNSRSRSYSPRRSRGSPRYSPRHSRSRSRT

SEQ ID NO 007 (hSMN1):

>gi|196115055|ref|NM_000344.3| Homo sapiens survival of motor neuron 1, telomeric (SMN1), transcript variant d, mRNA

CCACAAATGTGGGAGGGCGATAACCACTCGTAGAAAGCGTGAGAAGTTACTACAAGCGGTCCTCCCGG
CCACCGTACTGTTCCGCTCCCAGAAGCCCCGGGCGGCGGAAGTCGTCACTCTTAAGAAGGGACGGGGC
CCCACGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCCGGA
GCAGGAGGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGATA
CAGCACTGATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATT
TGTGAAACTTCGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAA
GAAGAATACTGCAGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAG
ACGGTTGCATTTACCCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTAC
ACTGGATATGGAAATAGAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAA
TAATATAGAACAAAATGCTCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACT
CCAGGTCTCCTGGAAATAAATCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCT
CCACCACCCCCCATGCCAGGGCCAAGACTGGGACCAGGAAAGCCAGGTCTAAAATTCAATGGCCCACC

ACCGCCACCGCCACCACCACCACCCCACTTACTATCATGCTGGCTGCCTCCATTTCCTTCTGGACCAC
CAATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGTATG
TTAATTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGGGTTT*C*AGACAAAATCAAAAA
GAAGGAAGGTGCTCACATTCCTTAAATTAAGGAGAAATGCTGGCATAGAGCAGCACTAAATGACACCA
CTAAAGAAACGATCAGACAGATCTGGAATGTGAAGCGTTATAGAAGATAACTGGCCTCATTTCTTCAA
AATATCAAGTGTTGGGAAAGAAAAAAGGAAGTGGAATGGGTAACTCTTCTTGATTAAAAGTTATGTAA
TAACCAAATGCAATGTGAAATATTTTACTGGACTCTATTTTGAAAAACCATCTGTAAAAGACTGGGGT
GGGGGTGGGAGGCCAGCACGGTGGTGAGGCAGTTGAGAAAATTTGAATGTGGATTAGATTTTGAATGA
TATTGGATAATTATTGGTAATTTTATGAGCTGTGAGAAGGGTGTTGTAGTTTATAAAAGACTGTCTTA
ATTTGCATACTTAAGCATTTAGGAATGAAGTGTTAGAGTGTCTTAAAATGTTTCAAATGGTTTAACAA
AATGTATGTGAGGCGTATGTGGCAAAATGTTACAGAATCTAACTGGTGGACATGGCTGTTCATTGTAC
TGTTTTTTTCTATCTTCTATATGTTTAAAAGTATATAATAAAAATATTTAATTTTTTTTTAAAAAAAA
AAAAAAAAAAA

Nucleotide+6 in exon 7 is marked by asterisks.

SEQ ID NO 008 (hSMN2):

>gi|196115210|ref|NM_017411.3| Homo sapiens survival of motor neuron 2, centromeric (SMN2), transcript variant d, mRNA

CCACAAATGTGGGAGGGCGATAACCACTCGTAGAAAGCGTGAGAAGTTACTACAAGCGGTCCTCCCGG
CCACCGTACTGTTCCGCTCCCAGAAGCCCCGGGCGGCGGAAGTCGTCACTCTTAAGAAGGGACGGGGC
CCCACGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCCGGA
GCAGGAGGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGATA
CAGCACTGATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATT
TGTGAAACTTCGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAA
GAAGAATACTGCAGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAG
ACGGTTGCATTTACCCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTAC
ACTGGATATGGAAATAGAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAA
TAATATAGAACAAAATGCTCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACT
CCAGGTCTCCTGGAAATAAATCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCT
CCACCACCCCCCATGCCAGGGCCAAGACTGGGACCAGGAAAGCCAGGTCTAAAATTCAATGGCCCACC
ACCGCCACCGCCACCACCACCCCACTTACTATCATGCTGGCTGCCTCCATTTCCTTCTGGACCAC
CAATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGTATG
TTAATTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGGGTTT*T*AGACAAAATCAAAAA
GAAGGAAGGTGCTCACATTCCTTAAATTAAGGAGAAATGCTGGCATAGAGCAGCACTAAATGACACCA
CTAAAGAAACGATCAGACAGATCTGGAATGTGAAGCGTTATAGAAGATAACTGGCCTCATTTCTTCAA
AATATCAAGTGTTGGGAAAGAAAAAAGGAAGTGGAATGGGTAACTCTTCTTGATTAAAAGTTATGTAA
TAACCAAATGCAATGTGAAATATTTTACTGGACTCTATTTTGAAAAACCATCTGTAAAAGACTGAGGT

GGGGGTGGGAGGCCAGCACGGTGGTGAGGCAGTTGAGAAAATTTGAATGTGGATTAGATTTTGAATGA
TATTGGATAATTATTGGTAATTTTATGAGCTGTGAGAAGGGTGTTGTAGTTTATAAAAGACTGTCTTA
ATTTGCATACTTAAGCATTTAGGAATGAAGTGTTAGAGTGTCTTAAAATGTTTCAAATGGTTTAACAA
AATGTATGTGAGGCGTATGTGGCAAAATGTTACAGAATCTAACTGGTGGACATGGCTGTTCATTGTAC
TGTTTTTTTCTATCTTCTATATGTTTAAAAGTATATAATAAAAATATTTAATTTTTTTTTAAATTAAA
AAAA

Nucleotide+6 in exon 7 is marked by asterisks.

Claims

1. A non-naturally occurring protein, wherein said protein has an identity of ≥70%, ≥85%, ≥90%, ≥95%, or ≥99% compared to SEQ ID NO 001 (SRSF1-WT), characterized in that said protein:
   i. has a splicing factor activity,
   ii. can be recruited to exon 7 of the SMN1 and/or SMN2 gene, and
   iii. facilitates the inclusion of said exon 7 of said SMN1 and/or SMN2 gene into the mature SMN1 and/or SMN2 mRNA.
2. The non-naturally occurring protein according to claim 1, characterized in that the glutamate on position 87 of SEQ ID NO 001 is substituted by an asparagine, lysine, arginine, glutamine or histidine.
3. The non-naturally occurring protein according to any one of the preceding claims, wherein said protein is selected from SEQ ID NO 002 (E87N), SEQ ID NO 003 (E87K), SEQ ID NO 004 (E87R), SEQ ID NO 005 (E87Q) or SEQ ID NO 006 (E87H).
4. A non-naturally occurring protein selected from SEQ ID NO 002 (E87N), SEQ ID NO 003 (E87K), SEQ ID NO 004 (E87R), SEQ ID NO 005 (E87Q) or SEQ ID NO 006 (E87H).
5. A non-naturally occurring protein according to any one of the preceding claims for use as a pharmaceutical.
6. A non-naturally occurring protein according to any one of claims 1 to 4 for use in a treatment of spinal muscular atrophy (SMA).
7. A nucleic acid molecule encoding a non-naturally occurring protein according to any one of the preceding claims 1 to 4.
8. A recombinant expression vector comprising the nucleic acid molecule of claim 7.
9. The recombinant expression vector of claim 8, wherein the nucleotide acid molecule encoding the polypeptide is operably linked to a promoter, particularly a promoter operable in a human cell.
10. The recombinant expression vector of claims 8 or 9, wherein the vector is a viral vector.
11. A transgenic cell comprising the nucleic acid molecule of claim 7, or the recombinant expression vector of claims 8, 9 or 10.

12. A nucleic acid molecule encoding a non-naturally occurring protein according to claim 7 for use in a treatment of spinal muscular atrophy (SMA).

13. A recombinant expression vector according to any one of claims 8 to 10 for use in a treatment of spinal muscular atrophy (SMA).

14. A transgenic cell according to claim 11 for use in a treatment of spinal muscular atrophy (SMA).

Fig. 1

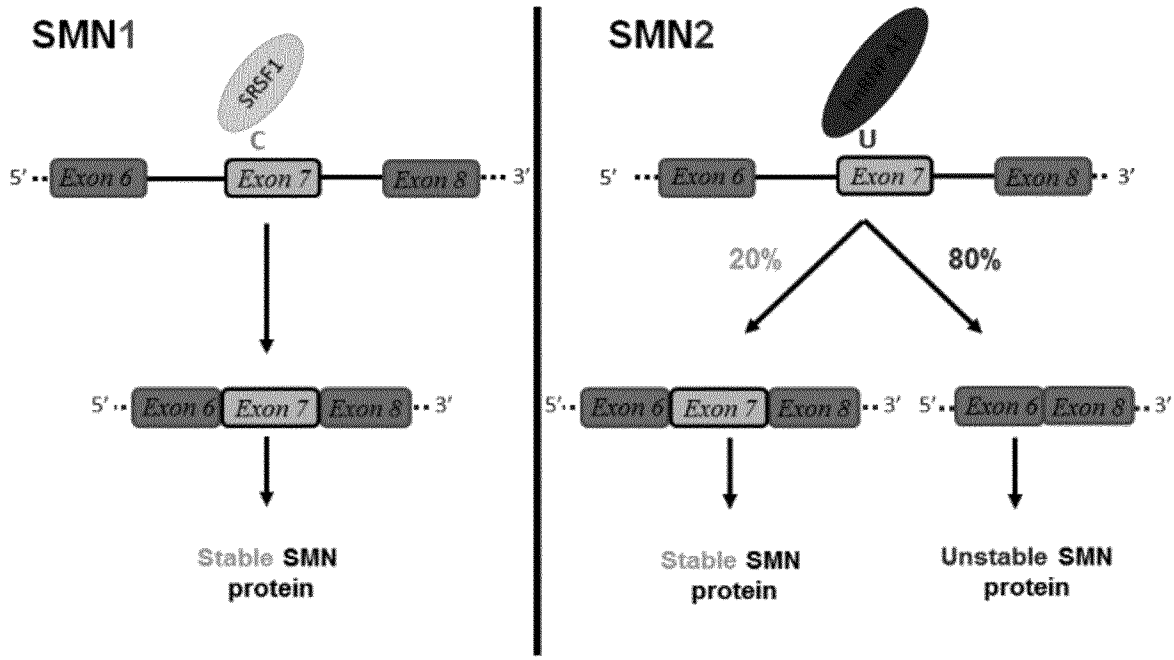

SMN1
SRSF1
C
5'
Exon 6
Exon 7
Exon 8
3'
Stable SMN protein
SMN2
U
20%
80%
Stable SMN protein
Unstable SMN protein

Fig. 2

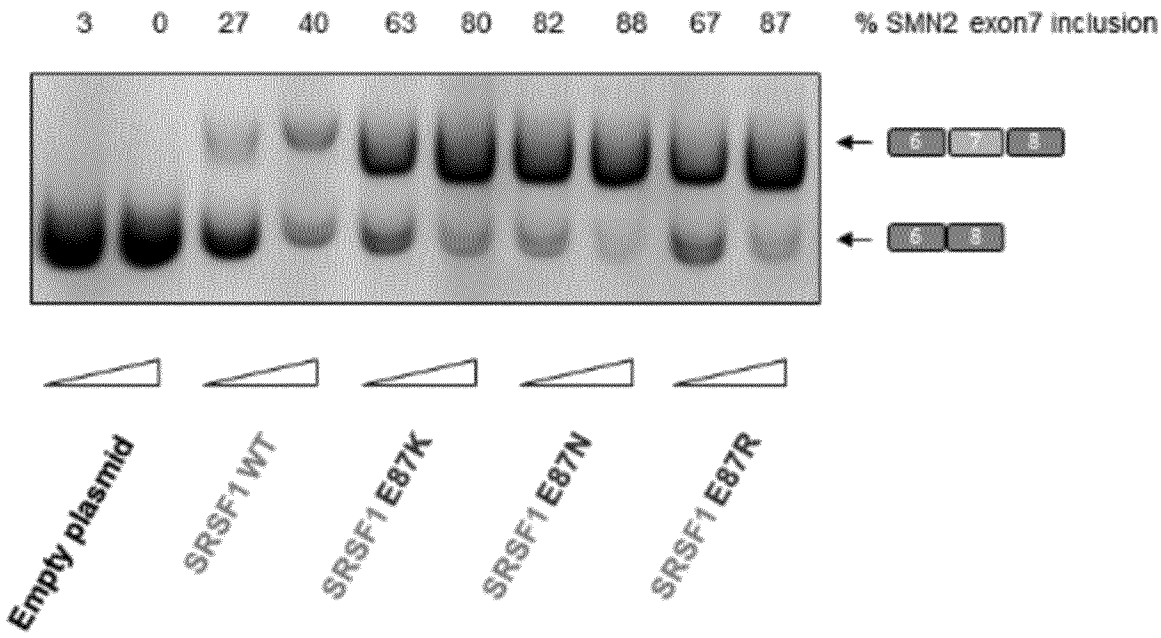
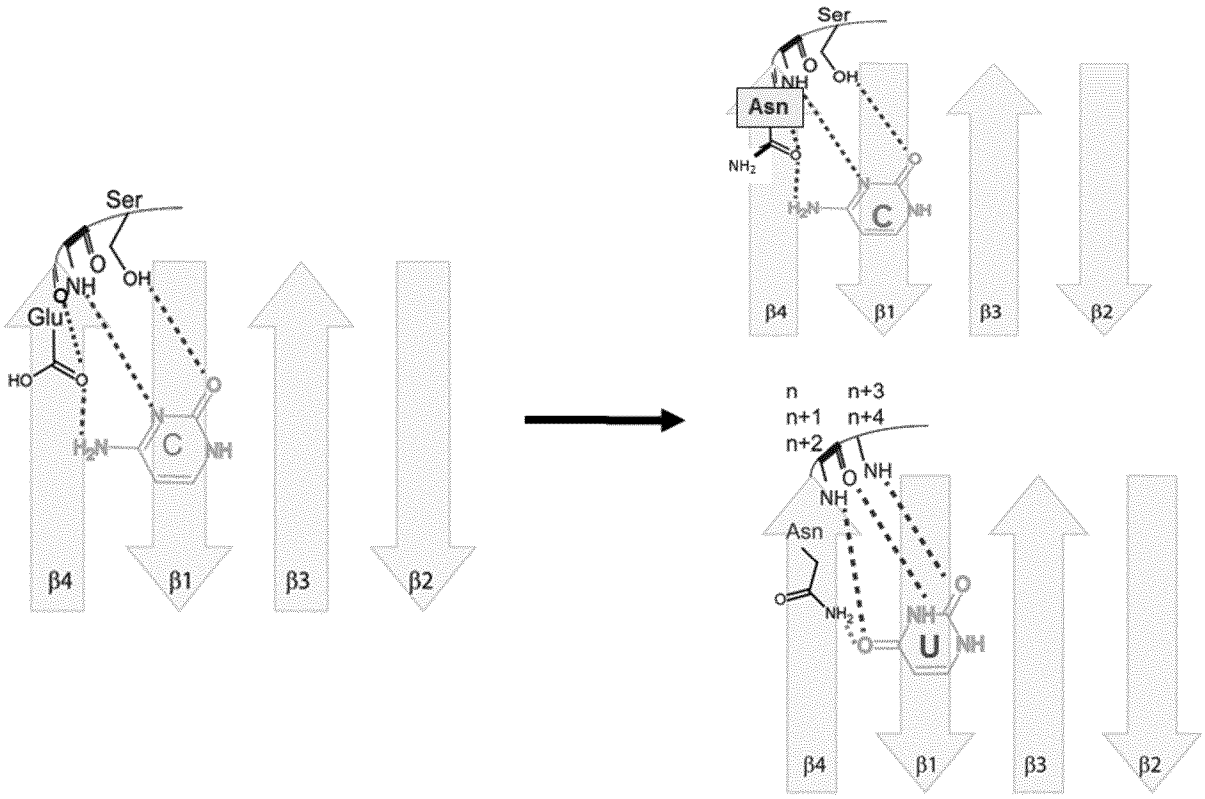

3 0 27 40 63 80 82 88 67 87
% SMN2 exon7 inclusion
Empty plasmid
SRSF1 WT
SRSF1 E87K
SRSF1 E87N
SRSF1 E87R

Fig. 3



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/059468

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61K38/17 A61K48/00
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
A61K

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, Sequence Search, WPI Data, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE UniProt [Online] 1 February 1995 (1995-02-01), "RecName: Full=Serine/arginine-rich splicing factor 1; AltName: Full=Alternative-splicing factor 1; Short=ASF-1; AltName: Full=Splicing factor, arginine/serine-rich 1; AltName: Full=pre-mRNA-splicing factor SF2, P33 subunit;", XP002743169, retrieved from EBI accession no. UNIPROT:Q07955 Database accession no. Q07955	1,5-14
A	sequence ----- -/--	2-4

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

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 July 2016	01/08/2016

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

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CARTEGNI LUCA ET AL: "Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1", NATURE GENETICS, NATURE PUBLISHING GROUP, NEW YORK, US, vol. 30, no. 4, 1 April 2002 (2002-04-01), pages 377-384, XP002289616, ISSN: 1061-4036, DOI: 10.1038/NG854	1,5-14
A	abstract	2-4

A	A. CLERY ET AL: "Isolated pseudo-RNA-recognition motifs of SR proteins can regulate splicing using a noncanonical mode of RNA recognition", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, vol. 110, no. 30, 8 July 2013 (2013-07-08) , pages E2802-E2811, XP055206992, ISSN: 0027-8424, DOI: 10.1073/pnas.1303445110 abstract	1-14

A	RAVINDRA N. SINGH: "Evolving Concepts on Human SMN Pre-mRNA Splicing", RNA BIOLOGY, vol. 4, no. 1, 1 January 2007 (2007-01-01) , pages 7-10, XP055039066, ISSN: 1547-6286, DOI: 10.4161/rna.4.1.4535 abstract	1-14

A	WANG H-Y ET AL: "Rate of evolution in brain-expressed genes in humans and other primates", PLOS BIOLOGY, PUBLIC LIBRARY OF SCIENCE, US, vol. 5, no. 2, 1 February 2007 (2007-02-01), page E13, XP003016059, ISSN: 1544-9173, DOI: 10.1371/JOURNAL.PBIO.0050013 sequence I7GLA8	1-14
	-/--	

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	-& DATABASE UniProt [Online] 3 October 2012 (2012-10-03), "SubName: Full=Macaca fascicularis brain cDNA clone: QorA-12520, similar to human splicing factor, arginine/serine-rich 1 (splicingfactor 2, alternate splicing factor) (SFRS1), mRNA, RefSeq: NM_006924.3 {ECO:0000313¦EMBL:BAE88160.1};", XP002743170, retrieved from EBI accession no. UNIPROT:I7GLA8 Database accession no. I7GLA8 sequence -----	
A	US 2002/018768 A1 (KEREM BATSHEVA [IL]) 14 February 2002 (2002-02-14) claim 1 -----	1-14

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2016/059468

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2002018768 A1	14-02-2002	NONE	