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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 November 2022 (10.11.2022)

WIPO | PCT

(10) International Publication Number
WO 2022/234519 A1

(51) International Patent Classification:
C12N 15/113 (2010.01) *C12N 15/10* (2006.01)
C12N 9/22 (2006.01)

(21) International Application Number:
PCT/IB2022/054177

(22) International Filing Date:
05 May 2022 (05.05.2022)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/184,461 05 May 2021 (05.05.2021) US

(71) Applicant: **CRISPR THERAPEUTICS AG** [CH/CH]; Baarerstrasse 14, 6300 Zug (CH).

(72) Inventors: **LIU, Jianming**; c/o Vertex Pharmaceuticals Incorporated, 50 Northern Avenue, Boston, Massachusetts 02210 (US). **ANDERSON, Eric Gunnar**; c/o CRISPR THERAPEUTICS, 610 Main Street, Cambridge, Massachusetts 02139 (US).

(74) Agent: **VOSSIUS & PARTNER PATENTANWÄLTE RECHTSANWÄLTE mbB**; Siebertstrasse 3, 81675 München (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, 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, 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))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*
— *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

(54) **Title:** COMPOSITIONS AND METHODS FOR USING SACAS9 SCAFFOLD SEQUENCES

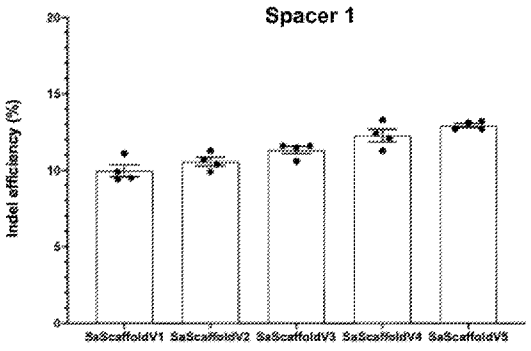

Fig. 1A

(57) **Abstract:** Provided herein are compositions and methods for gene editing using guide RNAs comprising scaffold sequences. Gene editing compositions comprising guide RNA scaffold sequences are provided for use with Staphylococcus aureus Cas9 (SaCas9). Methods of gene editing cells using Staphylococcus aureus Cas9 and guide RNAs comprising scaffold sequences are also encompassed.

WO 2022/234519 A1

COMPOSITIONS AND METHODS FOR USING SACAS9 SCAFFOLD SEQUENCES

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of U.S. Provisional Application No. 63/184,461, filed May 5, 2021, the disclosure of which is hereby incorporated by reference in its entirety.

INCORPORATION BY REFERENCE OF SEQUENCE LISTING

[0002] This application contains a Sequence Listing that has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. The ASCII copy, created on April 28, 2022 is named 100867-725781_CT164-PCT2_Sequence_Listing_ST25.txt, and is about 51,000 bytes in size.

FIELD

[0003] The present disclosure relates to the field of CRISPR-based gene editing using a Cas9 nuclease.

BACKGROUND

[0004] CRISPR-based genome editing can provide sequence-specific cleavage of genomic DNA using a Cas9 and a guide RNA. The approximately 20 nucleotides at the 5′ end of the guide RNA serves as the guide or spacer sequence that can be any sequence complementary to one strand of a genomic target location that has an adjacent protospacer adjacent motif (PAM). The PAM sequence is a short sequence adjacent to the Cas9 nuclease cut site that the Cas9 molecule requires for appropriate binding. Certain nucleotides of the guide RNA that are 3’ of the guide or spacer sequence serve as a scaffold sequence for interacting with Cas9. When expressed as a single molecule, a guide RNA is typically termed sgRNA. When expressed as more than one molecule, a guide RNA is typically termed dual guide RNA. When a guide RNA and a Cas9 are expressed, the guide RNA will bind to Cas9 and direct it to the sequence complementary to the guide sequence, where it will then initiate a double-stranded break (DSB). To repair these breaks, cells typically use an error prone mechanism of non-homologous end joining (NHEJ) which can lead to disruption of function in the target gene through insertions or deletion of codons, shifts in the reading frame, or result in a premature stop codon triggering nonsense-mediated decay.

SUMMARY

[0005] Provided herein are compositions and methods utilizing scaffold sequences for Cas9 from *Staphylococcus aureus* (SaCas9).

[0006] In some aspects the current disclosure encompasses a composition comprising a nucleic acid encoding a guide RNA comprising a scaffold sequence selected from any one of: SEQ ID NOs: 1-5. In various aspects the scaffold sequence is 3' of a guide sequence.

[0007] In some aspects the guide RNA is capable of directing a *Staphylococcus aureus* Cas9 (SaCas9) to create an edit in a target sequence.

[0008] In various aspects, another composition is provided, the composition comprising a guide RNA comprising in 5' to 3' direction: (a) a nucleic encoding a guide sequence; and (b) a nucleic acid encoding a scaffold sequence selected from any one of: SEQ ID NOs: 1-5.

[0009] Any of the compositions described herein may further comprise a nucleic acid encoding a *Staphylococcus aureus* Cas9 (SaCas9).

[0010] In any of the compositions herein, the gRNA may be an sgRNA. In various aspects, the guide RNA may be modified. In some aspects, the modification alters one or more 2' positions and/or phosphodiester linkages. In some aspects, the modification alters one or more or all of the first three nucleotides of the guide RNA. In some aspects, the modification alters one or more, or all, of the last three nucleotides of the guide RNA.

[0011] In further aspects, the modification of the guide RNA may include includes one or more of a phosphorothioate modification, a 2'-OMe modification, a 2'-O-MOE modification, a 2'-F modification, a 2′-O-methine-4′ bridge modification, a 3′-thiophosphonoacetate modification, or a 2'-deoxy modification.

[0012] In various aspects, the composition is associated with a lipid nanoparticle.

[0013] In various aspects, the nucleic acid encoding the guide RNA and the nucleic acid encoding the SluCas9 are provided in a viral vector. The viral vector may be, but is not limited to, an adeno-associated virus vector, a lentiviral vector, an integrase-deficient lentiviral vector, an adenoviral vector, a vaccinia viral vector, an alphaviral vector, or a herpes simplex viral vector.

[0014] For example, in some aspects, the viral vector may be an adeno-associated virus vector (e.g., an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh10, AAVrh74, or AAV9 vector, wherein the number following AAV indicates the AAV serotype). As an example, the AAV vector may be an AAV serotype 9 vector, an AAVrh10 vector or an AAVrh74 vector.

[0015] In any of the aspects herein, the viral vector may further comprise a tissue specific promoter. For example, the viral vector may comprise a muscle-specific promoter, optionally wherein the muscle-specific promoter is a muscle creatine kinase promoter, a desmin promoter, an MHCK7 promoter, an SPc5-12 promoter, or a CK8e promoter.

[0016] In any of the compositions provided herein, the SaCas9 comprises the amino acid sequence of SEQ ID NO: 7 or a variant thereof for example, the SaCas9 may be a variant of the amino acid sequence of SEQ ID NO: 7.

[0017] In some aspects, the SaCas9 comprises an amino acid sequence selected from any one of SEQ ID NOs: 9-11.

[0018] In some aspects the scaffold sequence may comprise SEQ ID NO: 2.

[0019] In various aspects the composition provided herein may comprise scaffold sequence comprising SEQ ID NO: 5.

[0020] In various aspects, any of the compositions provided herein may further comprise a pharmaceutically acceptable excipient.

[0021] Further provided herein is a method of gene editing, the method comprising delivering to a cell any composition provided herein.

[0022] In some aspects the current disclosure encompasses a method of gene editing, the method comprising delivering to a cell a composition comprising:

a. a guide RNA comprising in 5' to 3' direction:

i. a nucleic acid encoding a guide sequence; and

ii. a nucleic acid encoding a scaffold sequence selected from any one of SEQ ID NOs: 1-5; and

b. a nucleic acid encoding a *Staphylococcus aureus* Cas9 (SaCas9);

thereby producing a gene edit in the cell.

[0023] In some aspects of the method provided herein, the SaCas9 comprises the amino acid sequence of SEQ ID NO: 7. In various aspects the SaCas9 may be a variant of the amino acid sequence of SEQ ID NO: 7. In various aspects the the SaCas9 may comprise an amino acid sequence selected from any one of SEQ ID NOs: 9-11. In various aspects the scaffold sequence comprises SEQ ID NO: 2. In various aspects the scaffold sequence may comprise SEQ ID NO: 5.

FIGURE DESCRIPTIONS

[0024] **Figs. 1A-B** show editing efficiency of sgRNAs with different SaCas9 scaffolds in 293 T cells. **Fig. 1A** shows sgRNAs with spacer 1 and **Fig. 1B** shows sgRNAs with spacer 2.

[0025] **Fig. 2** shows editing efficiency of sgRNAs with different SaCas9 scaffolds in primary human myoblasts at three doses (30, 15, and 7.5 pmol).

DETAILED DESCRIPTION

[0026] Reference will now be made in detail to certain embodiments of the invention, examples of which are illustrated in the accompanying drawings. While the invention is described in conjunction with the illustrated embodiments, it will be understood that they are not intended to limit the invention to those embodiments. On the contrary, the invention is intended to cover all alternatives, modifications, and equivalents, which may be included within the invention as defined by the appended claims and included embodiments.

[0027] Before describing the present teachings in detail, it is to be understood that the disclosure is not limited to specific compositions or process steps, as such may vary. It should be noted

that, as used in this specification and the appended claims, the singular form "a", "an" and "the" include plural references unless the context clearly dictates otherwise. Thus, for example, reference to "a guide" includes a plurality of guides and reference to "a cell" includes a plurality of cells and the like.

[0028] Numeric ranges are inclusive of the numbers defining the range. Measured and measurable values are understood to be approximate, taking into account significant digits and the error associated with the measurement. Also, the use of "comprise", "comprises", "comprising", "contain", "contains", "containing", "include", "includes", and "including" are not intended to be limiting. It is to be understood that both the foregoing general description and detailed description are exemplary and explanatory only and are not restrictive of the teachings.

[0029] Unless specifically noted in the specification, embodiments in the specification that recite "comprising" various components are also contemplated as "consisting of" or "consisting essentially of" the recited components; embodiments in the specification that recite "consisting of" various components are also contemplated as "comprising" or "consisting essentially of" the recited components; and embodiments in the specification that recite "consisting essentially of" various components are also contemplated as "consisting of" or "comprising" the recited components (this interchangeability does not apply to the use of these terms in the claims). The term "or" is used in an inclusive sense, i.e., equivalent to "and/or," unless the context clearly indicates otherwise.

[0030] The section headings used herein are for organizational purposes only and are not to be construed as limiting the desired subject matter in any way. In the event that any material incorporated by reference contradicts any term defined in this specification or any other express content of this specification, this specification controls. While the present teachings are described in conjunction with various embodiments, it is not intended that the present teachings be limited to such embodiments. On the contrary, the present teachings encompass various alternatives, modifications, and equivalents, as will be appreciated by those of skill in the art.

I. Definitions

[0031] Unless stated otherwise, the following terms and phrases as used herein are intended to have the following meanings:

[0032] "Polynucleotide," "nucleic acid," and "nucleic acid molecule," are used herein to refer to a multimeric compound comprising nucleosides or nucleoside analogs which have nitrogenous heterocyclic bases or base analogs linked together along a backbone, including conventional RNA, DNA, mixed RNA-DNA, and polymers that are analogs thereof. A nucleic acid "backbone" can be made up of a variety of linkages, including one or more of sugar-phosphodiester linkages, peptide-nucleic acid bonds ("peptide nucleic acids" or PNA; PCT No. WO 95/32305), phosphorothioate linkages, methylphosphonate linkages, or combinations thereof. Sugar moieties of a nucleic acid can be ribose, deoxyribose, or similar compounds with substitutions, e.g., 2' methoxy or 2' halide substitutions. Nitrogenous bases can be conventional bases (A, G, C, T, U), analogs thereof (e.g.,

modified uridines such as 5-methoxyuridine, pseudouridine, or N1-methylpseudouridine, or others); inosine; derivatives of purines or pyrimidines (e.g., N^4-methyl deoxyguanosine, deaza- or aza-purines, deaza- or aza-pyrimidines, pyrimidine bases with substituent groups at the 5 or 6 position (e.g., 5-methylcytosine), purine bases with a substituent at the 2, 6, or 8 positions, 2-amino-6-methylaminopurine, O^6-methylguanine, 4-thio-pyrimidines, 4-amino-pyrimidines, 4-dimethylhydrazine-pyrimidines, and O^4-alkyl-pyrimidines; US Pat. No. 5,378,825 and PCT No. WO 93/13121). For general discussion see *The Biochemistry of the Nucleic Acids* 5-36, Adams et al., ed., 11th ed., 1992). Nucleic acids can include one or more "abasic" residues where the backbone includes no nitrogenous base for position(s) of the polymer (US Pat. No. 5,585,481). A nucleic acid can comprise only conventional RNA or DNA sugars, bases and linkages, or can include both conventional components and substitutions (e.g., conventional bases with 2' methoxy linkages, or polymers containing both conventional bases and one or more base analogs). Nucleic acid includes "locked nucleic acid" (LNA), an analogue containing one or more LNA nucleotide monomers with a bicyclic furanose unit locked in an RNA mimicking sugar conformation, which enhance hybridization affinity toward complementary RNA and DNA sequences (Vester and Wengel, 2004, *Biochemistry* 43(42):13233-41). RNA and DNA have different sugar moieties and can differ by the presence of uracil or analogs thereof in RNA and thymine or analogs thereof in DNA.

[0033] "Guide RNA", "guide RNA", "gRNA", and simply "guide" are used herein interchangeably to refer to either a crRNA (also known as CRISPR RNA), or the combination of a crRNA and a trRNA (also known as tracrRNA). The crRNA and trRNA may be associated as a single RNA molecule (single guide RNA, sgRNA) or in two separate RNA molecules (dual guide RNA, dgRNA). "Guide RNA" or "guide RNA" refers to each type. The trRNA may be a naturally-occurring sequence, or a trRNA sequence with modifications or variations compared to naturally-occurring sequences.

[0034] As used herein, a "spacer sequence," sometimes also referred to herein and in the literature as a "spacer," "protospacer," "guide sequence," or "targeting sequence" refers to a sequence within a guide RNA that is complementary to a target sequence and functions to direct a guide RNA to a target sequence for cleavage by a Cas9. A guide sequence can be 24, 23, 22, 21, 20 or fewer base pairs in length, e.g., in the case of *Staphylococcus aureus* (i.e., SaCas9) and related Cas9 homologs/orthologs. Shorter or longer sequences can also be used as guides, e.g., 15-, 16-, 17-, 18-, 19-, 20-, 21-, 22-, 23-, 24-, or 25-nucleotides in length. In some embodiments, the target sequence is in a gene or on a chromosome, for example, and is complementary to the guide sequence. In some embodiments, the degree of complementarity or identity between a guide sequence and its corresponding target sequence may be about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%. In some embodiments, the guide sequence and the target region may be 100% complementary or identical. In other embodiments, the guide sequence and the target region may contain at least one mismatch. For example, the guide sequence and the target sequence may contain 1, 2, 3, or 4

mismatches, where the total length of the target sequence is at least 17, 18, 19, 20 or more base pairs. In some embodiments, the guide sequence and the target region may contain 1-4 mismatches where the guide sequence comprises at least 17, 18, 19, 20 or more nucleotides. In some embodiments, the guide sequence and the target region may contain 1, 2, 3, or 4 mismatches where the guide sequence comprises 20 nucleotides. In some embodiments, the guide sequence and the target region do not contain any mismatches.

[0035] Target sequences for Cas9s include both the positive and negative strands of genomic DNA (i.e., the sequence given and the sequence's reverse compliment), as a nucleic acid substrate for a Cas9 is a double stranded nucleic acid. Accordingly, where a guide sequence is said to be "complementary to a target sequence", it is to be understood that the guide sequence may direct a guide RNA to bind to the reverse complement of a target sequence. Thus, in some embodiments, where the guide sequence binds the reverse complement of a target sequence, the guide sequence is identical to certain nucleotides of the target sequence (e.g., the target sequence not including the PAM) except for the substitution of U for T in the guide sequence.

[0036] As used herein, "ribonucleoprotein" (RNP) or "RNP complex" refers to a guide RNA together with a Cas9. In some embodiments, the guide RNA guides the Cas9 such as Cas9 to a target sequence, and the guide RNA hybridizes with and the agent binds to the target sequence, which can be followed by cleaving or nicking (in the context of a modified "nickase" Cas9).

[0037] As used herein, a first sequence is considered to "comprise a sequence with at least X% identity to" a second sequence if an alignment of the first sequence to the second sequence shows that X% or more of the positions of the second sequence in its entirety are matched by the first sequence. For example, the sequence AAGA comprises a sequence with 100% identity to the sequence AAG because an alignment would give 100% identity in that there are matches to all three positions of the second sequence. The differences between RNA and DNA (generally the exchange of uridine for thymidine or vice versa) and the presence of nucleoside analogs such as modified uridines do not contribute to differences in identity or complementarity among polynucleotides as long as the relevant nucleotides (such as thymidine, uridine, or modified uridine) have the same complement (e.g., adenosine for all of thymidine, uridine, or modified uridine; another example is cytosine and 5-methylcytosine, both of which have guanosine or modified guanosine as a complement). Thus, for example, the sequence 5'-AXG where X is any modified uridine, such as pseudouridine, N1-methyl pseudouridine, or 5-methoxyuridine, is considered 100% identical to AUG in that both are perfectly complementary to the same sequence (5'-CAU). Exemplary alignment algorithms are the Smith-Waterman and Needleman-Wunsch algorithms, which are well-known in the art. One skilled in the art will understand what choice of algorithm and parameter settings are appropriate for a given pair of sequences to be aligned; for sequences of generally similar length and expected identity >50% for amino acids or >75% for nucleotides, the Needleman-Wunsch algorithm with default settings of the

Needleman-Wunsch algorithm interface provided by the EBI at the www.ebi.ac.uk web server is generally appropriate.

[0038] "mRNA" is used herein to refer to a polynucleotide that is not DNA and comprises an open reading frame that can be translated into a polypeptide (i.e., can serve as a substrate for translation by a ribosome and amino-acylated tRNAs). mRNA can comprise a phosphate-sugar backbone including ribose residues or analogs thereof, e.g., 2'-methoxy ribose residues. In some embodiments, the sugars of an mRNA phosphate-sugar backbone consist essentially of ribose residues, 2'-methoxy ribose residues, or a combination thereof.

[0039] As used herein, a "target sequence" refers to a sequence of nucleic acid in a target gene that has complementarity to at least a portion of the guide sequence of the guide RNA. The interaction of the target sequence and the guide sequence directs a Cas9 to bind, and potentially nick or cleave (depending on the activity of the agent), within the target sequence.

[0040] A "pharmaceutically acceptable excipient" refers to an agent that is included in a pharmaceutical formulation that is not the active ingredient. Pharmaceutically acceptable excipients may e.g., aid in drug delivery or support or enhance stability or bioavailability.

[0041] The term "about" or "approximately" means an acceptable error for a particular value as determined by one of ordinary skill in the art, which depends in part on how the value is measured or determined.

[0042] As used herein, "*Staphylococcus aureus* Cas9" may also be referred to as SaCas9, and includes wild type SaCas9 (e.g., SEQ ID NO: 7) and variants thereof. A variant of SaCas9 comprises one or more amino acid changes as compared to SEQ ID NO: 7, including insertion, deletion, or substitution of one or more amino acids, or a chemical modification to one or more amino acids.

II. Compositions

[0043] Provided herein are guide RNA compositions comprising scaffold sequences for use with a *Staphylococcus aureus* Cas9 (SaCas9). The scaffold sequences disclosed herein may be incorporated into any guide RNA for use with a SaCas9.

[0044] In some embodiments, a composition is provided comprising: a nucleic acid encoding a guide RNA comprising a scaffold sequence selected from any one of SEQ ID NOs: 1-5. In some embodiments, a composition is provided comprising a guide RNA comprising in 5' to 3' direction: a) a nucleic acid encoding a guide sequence; and b) a nucleic acid encoding a scaffold sequence selected from any one of SEQ ID NOs: 1-5. In some embodiments, the scaffold sequence selected from SEQ ID NOs: 1-5 is for use with a *Staphylococcus aureus* Cas9 (SaCas9) or a nucleic acid encoding a *Staphylococcus aureus* Cas9 (SaCas9).

[0045] In some embodiments, a composition is provided comprising a nucleic acid encoding a guide RNA comprising a scaffold sequence selected from any one of SEQ ID NOs: 1-5, wherein the guide RNA is capable of directing a SaCas9 to edit a target sequence. In some embodiments, the scaffold

sequence selected from SEQ ID NOs: 1-5 is for use with a *Staphylococcus aureus* Cas9 (SaCas9) or a nucleic acid encoding a *Staphylococcus aureus* Cas9 (SaCas9).

[0046] In some embodiments, a composition is provided comprising: (a) a guide RNA comprising in 5' to 3' direction: i) a nucleic acid encoding a guide sequence; and ii) a nucleic acid encoding a scaffold sequence selected from any one of SEQ ID NOs: 1-5; and (b) a nucleic acid encoding a *Staphylococcus aureus* Cas9 (SaCas9). In some embodiments, the guide RNA is an sgRNA. In some embodiments, the guide RNA is modified.

[0047] A guide RNA may comprise a guide sequence and additional nucleotides to form or encode a crRNA. In some embodiments, the crRNA comprises (5' to 3') at least a spacer sequence and a first complementarity domain. The first complementary domain is sufficiently complementary to a second complementarity domain, which may be part of the same molecule in the case of an sgRNA or in a tracrRNA in the case of a dual or modular gRNA, to form a duplex. See, e.g., US 2017/0007679 for detailed discussion of crRNA and gRNA domains, including first and second complementarity domains.

[0048] A single-molecule guide RNA (sgRNA) can comprise, in the 5' to 3' direction, an optional spacer extension sequence, a spacer sequence, a minimum CRISPR repeat sequence, a single-molecule guide linker, a minimum tracrRNA sequence, a 3' tracrRNA sequence and/or an optional tracrRNA extension sequence. The optional tracrRNA extension can comprise elements that contribute additional functionality (*e.g.*, stability) to the guide RNA. The single-molecule guide linker can link the minimum CRISPR repeat and the minimum tracrRNA sequence to form a hairpin structure. The optional tracrRNA extension can comprise one or more hairpins.

[0049] Exemplary scaffold sequences suitable for use with SaCas9 to follow the guide sequence at its 3' end are shown in **Table 1** in the 5' to 3' orientation:

[0050] **Table 1: Exemplary SaCas9 Scaffold Sequences**

Scaffold ID	**SEQ ID NO**	**Scaffold Sequence (5' to 3')**
SaScaffoldV1	1	GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGC AAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGAT
SaScaffoldV2	2	GTTTAAGTACTCTGTGCTGGAAACAGCACAGAATCTACTT AAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTG GCGAGAT
SaScaffoldV3	3	GTTTAAGTACTCTGGAAACAGAATCTACTTAAACAAGGC AAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGAT
SaScaffoldV4	4	GTTTCAGTACTCTGTGCTGGAAACAGCACAGAATCTACTG AAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTG GCGAGAT
SaScaffoldV5	5	GTTTCAGTACTCTGGAAACAGAATCTACTGAAACAAGGC AAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGAT
Wildtype	6	GTTTAAGTACTCTGTGCTGGAAACAGCACAGAATCTACTT AAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTG GCGAGA

[0051] In some embodiments, the scaffold sequence is selected from any one of SEQ ID NOs: 1-5 in 5' to 3 orientation (see **Table 1**). In some embodiments, an exemplary sequence is a sequence that is at least 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one off SEQ ID NOs: 1-5, or a sequence that differs from any one of SEQ ID NOs: 1-5 by no more than 1, 2, 3, 4, 5, 10, 15, 20, or 25 nucleotides.

[0052] In some embodiments, the scaffold sequence suitable for use with SaCas9 to follow the guide sequence at its 3' end is selected from any one of SEQ ID NOs: 1-5 in 5' to 3 orientation (see **Table 1**). In some embodiments, an exemplary sequence for use with SaCas9 to follow the 3' end of the guide sequence is a sequence that is at least 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one off SEQ ID NOs: 1-5, or a sequence that differs from any one of SEQ ID NOs: 1-5 by no more than 1, 2, 3, 4, 5, 10, 15, 20, or 25 nucleotides.

[0053] In some embodiments, the nucleic acid encoding the gRNA comprises a scaffold sequence comprising SEQ ID NO: 1. In some embodiments, the nucleic acid encoding the gRNA comprises a scaffold sequence comprising SEQ ID NO: 2. In some embodiments, the nucleic acid encoding the gRNA comprises a scaffold sequence comprising SEQ ID NO: 3. In some embodiments, the nucleic acid encoding the gRNA comprises a scaffold sequence comprising SEQ ID NO: 4. In some embodiments, the nucleic acid encoding the gRNA comprises a scaffold sequence comprising SEQ ID NO: 5.

[0054] In some embodiments, the scaffold sequence comprises one or more alterations in the stem loop 1 as compared to the stem loop 1 of a wildtype SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 6) or a reference SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 5). In some embodiments, the scaffold sequence comprises one or more alterations in the stem loop 2 as compared to the stem loop 2 of a wildtype SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 6) or a reference SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 5). In some embodiments, the scaffold sequence comprises one or more alterations in the tetraloop as compared to the tetraloop of a wildtype SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 6) or a reference SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 5). In some embodiments, the scaffold sequence comprises one or more alterations in the repeat region as compared to the repeat region of a wildtype SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 6) or a reference SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 5). In some embodiments, the scaffold sequence comprises one or more alterations in the anti-repeat region as compared to the anti-repeat region of a wildtype SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 6) or a reference SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 5). In some embodiments, the scaffold sequence comprises one or more alterations in the linker region as compared to the linker region of a wildtype SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 6) or

a reference SaCas9 scaffold sequence (e.g., a scaffold comprising the sequence of SEQ ID NO: 5). See, e.g., Nishimasu et al., 2015, Cell, 162:1113-1126 for description of regions of a scaffold.

[0055] Where a tracrRNA is used, in some embodiments, it comprises (5' to 3') a second complementary domain and a proximal domain. In the case of a sgRNA, guide sequences together with additional nucleotides (e.g., SEQ ID Nos: 1-5) form or encode a sgRNA. In some embodiments, an sgRNA comprises (5' to 3') at least a spacer sequence, a first complementary domain, a linking domain, a second complementary domain, and a proximal domain. A sgRNA or tracrRNA may further comprise a tail domain. The linking domain may be hairpin-forming. See, e.g., US 2017/0007679 for detailed discussion and examples of crRNA and gRNA domains, including second complementarity domains, linking domains, proximal domains, and tail domains.

[0056] In general, in the case of a DNA nucleic acid construct encoding a guide RNA, the U residues in any of the RNA sequences described herein may be replaced with T residues, and in the case of a guide RNA construct encoded by a DNA, the T residues may be replaced with U residues.

[0057] In some embodiments, a composition is provided comprising a guide RNA, or nucleic acid encoding a guide RNA, wherein the guide RNA further comprises a scaffold sequence comprising a trRNA. In each composition and method embodiment described herein, the crRNA (comprising the spacer sequence) and trRNA may be associated as a single RNA (sgRNA) or may be on separate RNAs (dgRNA). In the context of sgRNAs, the crRNA and trRNA components may be covalently linked, e.g., via a phosphodiester bond or other covalent bond.

[0058] In any embodiment comprising a nucleic acid molecule encoding a guide RNA and/or a Cas9, the nucleic acid molecule may be a vector.

[0059] Any type of vector, such as any of those described herein, may be used. In some embodiments, the vector is a viral vector. In some embodiments, the viral vector is a non-integrating viral vector (i.e., that does not insert sequence from the vector into a host chromosome). In some embodiments, the viral vector is an adeno-associated virus vector (AAV), a lentiviral vector, an integrase-deficient lentiviral vector, an adenoviral vector, a vaccinia viral vector, an alphaviral vector, or a herpes simplex viral vector. In some embodiments, the vector comprises a muscle-specific promoter. Exemplary muscle-specific promoters include a muscle creatine kinase promoter, a desmin promoter, an MHCK7 promoter, or an SPc5-12 promoter. *See* US 2004/0175727 A1; Wang et al., *Expert Opin Drug Deliv*. (2014) 11, 345–364; Wang et al., *Gene Therapy* (2008) 15, 1489–1499, which are incorporated herein by reference in their entirety. In some embodiments, the muscle-specific promoter is a CK8 promoter. In some embodiments, the muscle-specific promoter is a CK8e promoter. In any of the foregoing embodiments, the vector may be an adeno-associated virus vector (AAV).

[0060] In some embodiments, the muscle specific promoter is the CK8 promoter. The CK8 promoter has the following sequence (SEQ ID NO. 12):

```
1      CTAGACTAGC ATGCTGCCCA TGTAAGGAGG CAAGGCCTGG GGACACCCGA GATGCCTGGT
61     TATAATTAAC CCAGACATGT GGCTGCCCCC CCCCCCCCAA CACCTGCTGC CTCTAAAAAT
121    AACCCTGCAT GCCATGTTCC CGGCGAAGGG CCAGCTGTCC CCCGCCAGCT AGACTCAGCA
```

```
181    CTTAGTTTAG GAACCAGTGA GCAAGTCAGC CCTTGGGGCA GCCCATACAA GGCCATGGGG
241    CTGGGCAAGC TGCACGCCTG GGTCCGGGGT GGGCACGGTG CCCGGGCAAC GAGCTGAAAG
301    CTCATCTGCT CTCAGGGGCC CCTCCCTGGG GACAGCCCCT CCTGGCTAGT CACACCCTGT
361    AGGCTCCTCT ATATAACCCA GGGGCACAGG GGCTGCCCTC ATTCTACCAC CACCTCCACA
421    GCACAGACAG ACACTCAGGA GCCAGCCAGC
```

[0061] In some embodiments, the muscle-cell cell specific promoter is a variant of the CK8 promoter, called CK8e. The CK8e promoter has the following sequence (SEQ ID NO. 13):

```
1      TGCCCATGTA AGGAGGCAAG GCCTGGGGAC ACCCGAGATG CCTGGTTATA ATTAACCCAG
61     ACATGTGGCT GCCCCCCCCC CCCCAACACC TGCTGCCTCT AAAAATAACC CTGCATGCCA
121    TGTTCCCGGC GAAGGGCCAG CTGTCCCCCG CCAGCTAGAC TCAGCACTTA GTTTAGGAAC
181    CAGTGAGCAA GTCAGCCCTT GGGGCAGCCC ATACAAGGCC ATGGGGCTGG GCAAGCTGCA
241    CGCCTGGGTC CGGGGTGGGC ACGGTGCCCG GGCAACGAGC TGAAAGCTCA TCTGCTCTCA
301    GGGGCCCCTC CCTGGGGACA GCCCCTCCTG GCTAGTCACA CCCTGTAGGC TCCTCTATAT
361    AACCCAGGGG CACAGGGGCT GCCCTCATTC TACCACCACC TCCACAGCAC AGACAGACAC
421    TCAGGAGCCA GCCAGC
```

[0062] In some embodiments, the nucleic acid encoding the Cas9 protein is under the control of a CK8e promoter. In some embodiments, the vector is AAV9.

[0063] In some embodiments, the nucleic acid encoding SaCas9 encodes an SaCas9 comprising an amino acid sequence that is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the sequence of SEQ ID NO: 7:

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRI
QRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEE
DTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQK
AYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYA
YNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGY
RVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEE
IEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDD
FILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVL
VKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQ
KDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYK
HHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHI
KDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSP
EKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGN
KLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKC
YEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMN
DKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG.

[0064] In some embodiments, the nucleic acid encoding SaCas9 comprises the nucleic acid of SEQ ID NO: 8:

AAGCGCAATTACATCCTGGGCCTGGATATCGGCATCACCTCCGTGGGCTACGGCATCATC
GACTATGAGACACGGGATGTGATCGACGCCGGCGTGAGACTGTTCAAGGAGGCCAACGT
GGAGAACAATGAGGGCCGGCGGAGCAAGAGGGGAGCAAGGCGCCTGAAGCGGAGAAG
GCGCCACAGAATCCAGAGAGTGAAGAAGCTGCTGTTCGATTACAACCTGCTGACCGACC

ACTCCGAGCTGTCTGGCATCAATCCTTATGAGGCCCGGGTGAAGGGCCTGTCCCAGAAGC
TGTCTGAGGAGGAGTTTTCTGCCGCCCTGCTGCACCTGGCAAAGAGGAGAGGCGTGCAC
AACGTGAATGAGGTGGAGGAGGACACCGGCAACGAGCTGAGCACAAAGGAGCAGATCA
GCCGCAATTCCAAGGCCCTGGAGGAGAAGTATGTGGCCGAGCTGCAGCTGGAGCGGCTG
AAGAAGGATGGCGAGGTGAGGGGCTCCATCAATCGCTTCAAGACCTCTGACTACGTGAA
GGAGGCCAAGCAGCTGCTGAAGGTGCAGAAGGCCTACCACCAGCTGGATCAGAGCTTTA
TCGATACATATATCGACCTGCTGGAGACCAGGCGCACATACTATGAGGGACCAGGAGAG
GGCTCCCCCTTCGGCTGGAAGGACATCAAGGAGTGGTACGAGATGCTGATGGGCCACTG
CACCTATTTTCCAGAGGAGCTGAGATCCGTGAAGTACGCCTATAACGCCGATCTGTACAA
CGCCCTGAATGACCTGAACAACCTGGTCATCACCAGGGATGAGAACGAGAAGCTGGAGT
ACTATGAGAAGTTCCAGATCATCGAGAACGTGTTCAAGCAGAAGAAGAAGCCTACACTG
AAGCAGATCGCCAAGGAGATCCTGGTGAACGAGGAGGACATCAAGGGCTACCGCGTGA
CCAGCACAGGCAAGCCAGAGTTCACCAATCTGAAGGTGTATCACGATATCAAGGACATC
ACAGCCCGGAAGGAGATCATCGAGAACGCCGAGCTGCTGGATCAGATCGCCAAGATCCT
GACCATCTATCAGAGCTCCGAGGACATCCAGGAGGAGCTGACCAACCTGAATAGCGAGC
TGACACAGGAGGAGATCGAGCAGATCAGCAATCTGAAGGGCTACACCGGCACACACAA
CCTGTCCCTGAAGGCCATCAATCTGATCCTGGATGAGCTGTGGCACACAAACGACAATCA
GATCGCCATCTTTAACAGGCTGAAGCTGGTGCCAAAGAAGGTGGACCTGAGCCAGCAGA
AGGAGATCCCAACCACACTGGTGGACGATTTCATCCTGTCCCCCGTGGTGAAGCGGAGCT
TCATCCAGAGCATCAAAGTGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAATGAT
ATCATCATCGAGCTGGCCAGGGAGAAGAACTCTAAGGACGCCCAGAAGATGATCAATGA
GATGCAGAAGAGGAACCGCCAGACCAATGAGCGGATCGAGGAGATCATCAGAACCACA
GGCAAGGAGAACGCCAAGTACCTGATCGAGAAGATCAAGCTGCACGATATGCAGGAGG
GCAAGTGTCTGTATAGCCTGGAGGCCATCCCTCTGGAGGACCTGCTGAACAATCCATTCA
ACTACGAGGTGGATCACATCATCCCCCGGAGCGTGAGCTTCGACAATTCCTTTAACAATA
AGGTGCTGGTGAAGCAGGAGGAGAACTCTAAGAAGGGCAATAGGACCCCTTTCCAGTAC
CTGTCTAGCTCCGATTCTAAGATCAGCTACGAGACCTTCAAGAAGCACATCCTGAATCTG
GCCAAGGGCAAGGGCCGCATCTCTAAGACCAAGAAGGAGTACCTGCTGGAGGAGCGGG
ACATCAACAGATTCAGCGTGCAGAAGGACTTCATCAACCGGAATCTGGTGGACACCAGA
TACGCCACACGCGGCCTGATGAATCTGCTGCGGTCCTATTTCAGAGTGAACAATCTGGAT
GTGAAGGTGAAGAGCATCAACGGCGGCTTCACCTCCTTTCTGCGGAGAAAGTGGAAGTT
TAAGAAGGAGAGAAACAAGGGCTATAAGCACCACGCCGAGGATGCCCTGATCATCGCCA
ATGCCGACTTCATCTTTAAGGAGTGGAAGAAGCTGGACAAGGCCAAGAAAGTGATGGAG
AACCAGATGTTCGAGGAGAAGCAGGCCGAGAGCATGCCCGAGATCGAGACCGAGCAGG
AGTACAAGGAGATTTTCATCACACCTCACCAGATCAAGCACATCAAGGACTTCAAGGAC
TACAAGTATTCCCACAGGGTGGATAAGAAGCCCAACCGCGAGCTGATCAATGACACCCT
GTATTCTACAAGGAAGGACGATAAGGGCAATACCCTGATCGTGAACAATCTGAACGGCC

TGTACGACAAGGATAATGACAAGCTGAAGAAGCTGATCAACAAGAGCCCCGAGAAGCT
GCTGATGTACCACCACGATCCTCAGACATATCAGAAGCTGAAGCTGATCATGGAGCAGT
ACGGCGACGAGAAGAACCCACTGTATAAGTACTATGAGGAGACCGGCAACTACCTGACA
AAGTATTCCAAGAAGGATAATGGCCCCGTGATCAAGAAGATCAAGTACTATGGCAACAA
GCTGAATGCCCACCTGGACATCACCGACGATTACCCCAACAGCCGGAATAAGGTGGTGA
AGCTGAGCCTGAAGCCATACAGGTTCGACGTGTACCTGGACAACGGCGTGTATAAGTTT
GTGACAGTGAAGAATCTGGATGTGATCAAGAAGGAGAACTACTATGAAGTGAATAGCAA
GTGCTACGAGGAGGCCAAGAAGCTGAAGAAGATCAGCAACCAGGCCGAGTTCATCGCCT
CTTTTTACAACAATGACCTGATCAAGATCAATGGCGAGCTGTATAGAGTGATCGGCGTGA
ACAATGATCTGCTGAACCGCATCGAAGTGAATATGATCGACATCACCTACCGGGAGTAT
CTGGAGAACATGAATGATAAGAGGCCCCCTCGCATCATCAAGACCATCGCCTCTAAGAC
ACAGAGCATCAAGAAGTACTCTACAGACATCCTGGGCAACCTGTATGAGGTGAAGAGCA
AGAAGCACCCTCAGATCATCAAGAAGGGC.

[0065] In some embodiments, the composition comprises a nucleic acid encoding SaCas9, the SaCas9 comprises an amino acid sequence of SEQ ID NO: 7.

[0066] In some embodiments, the SaCas9 is a variant of the amino acid sequence of SEQ ID NO: 711. In some embodiments, the SaCas9 comprises an amino acid other than an E at the position corresponding to position 781 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an amino acid other than an N at the position corresponding to position 967 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an amino acid other than an R at the position corresponding to position 1014 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises a K at the position corresponding to position 781 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises a K at the position corresponding to position 967 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an H at the position corresponding to position 1014 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an amino acid other than an E at the position corresponding to position 781 of SEQ ID NO: 7; an amino acid other than an N at the position corresponding to position 967 of SEQ ID NO: 7; and an amino acid other than an R at the position corresponding to position 1014 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises a K at the position corresponding to position 781 of SEQ ID NO: 7; a K at the position corresponding to position 967 of SEQ ID NO: 7; and an H at the position corresponding to position 1014 of SEQ ID NO: 7.

[0067] In some embodiments, the SaCas9 comprises an amino acid other than an R at the position corresponding to position 244 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an amino acid other than an N at the position corresponding to position 412 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an amino acid other than an N at the position corresponding to position 418 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an amino acid other than an R at the position corresponding to position 653 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an amino acid other than an R at the position corresponding to position 244 of SEQ ID NO:

7; an amino acid other than an N at the position corresponding to position 412 of SEQ ID NO: 7; an amino acid other than an N at the position corresponding to position 418 of SEQ ID NO: 7; and an amino acid other than an R at the position corresponding to position 653 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an A at the position corresponding to position 244 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an A at the position corresponding to position 412 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an A at the position corresponding to position 418 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an A at the position corresponding to position 653 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an A at the position corresponding to position 244 of SEQ ID NO: 7; an A at the position corresponding to position 412 of SEQ ID NO: 7; an A at the position corresponding to position 418 of SEQ ID NO: 7; and an A at the position corresponding to position 653 of SEQ ID NO: 7.

[0068] In some embodiments, the SaCas9 comprises an amino acid other than an R at the position corresponding to position 244 of SEQ ID NO: 7; an amino acid other than an N at the position corresponding to position 412 of SEQ ID NO: 7; an amino acid other than an N at the position corresponding to position 418 of SEQ ID NO: 7; an amino acid other than an R at the position corresponding to position 653 of SEQ ID NO: 7; an amino acid other than an E at the position corresponding to position 781 of SEQ ID NO: 7; an amino acid other than an N at the position corresponding to position 967 of SEQ ID NO: 7; and an amino acid other than an R at the position corresponding to position 1014 of SEQ ID NO: 7. In some embodiments, the SaCas9 comprises an A at the position corresponding to position 244 of SEQ ID NO: 7; an A at the position corresponding to position 412 of SEQ ID NO: 7; an A at the position corresponding to position 418 of SEQ ID NO: 7; an A at the position corresponding to position 653 of SEQ ID NO: 7; a K at the position corresponding to position 781 of SEQ ID NO: 711; a K at the position corresponding to position 967 of SEQ ID NO: 7; and an H at the position corresponding to position 1014 of SEQ ID NO: 7.

[0069] In some embodiments, the SaCas9 comprises an amino acid sequence that is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the sequence of SEQ ID NO: 9 (designated herein as SaCas9-KKH or SACAS9KKH):

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRI
QRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEE
DTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQK
AYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYA
YNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGY
RVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEE
IEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDD
FILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVL
VKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQ

KDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYK
HHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHI
KDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSP
EKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGN
KLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKC
YEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMN
DKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG.

[0070] In some embodiments, the SaCas9 comprises an amino acid sequence that is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the sequence of SEQ ID NO: 10 (designated herein as SaCas9-HF):

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRI
QRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEE
DTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQK
AYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELASVKYA
YNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGY
RVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEE
IEQISNLKGYTGTHNLSLKAINLILDELWHTNDAQIAIFARLKLVPKKVDLSQQKEIPTTLVDD
FILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVL
VKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQ
KDFINRNLVDTRYATAGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYK
HHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHI
KDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSP
EKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGN
KLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKC
YEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMN
DKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG.

[0071] In some embodiments, the SaCas9 comprises an amino acid sequence that is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the sequence of SEQ ID NO: 11 (designated herein as SaCas9-KKH-HF):

[0072] KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLK
RRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVH
NVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAK
QLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEE
LASVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILV
NEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNL
NSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDAQIAIFARLKLVPKKVDLSQQKE

IPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQT
NERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFD
NSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEE
RDINRFSVQKDFINRNLVDTRYATAGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFK
KERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDK
LKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPV
IKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKE
NYYEVNSKCYEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEVNMIDITY
REYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG.

Modified guide RNAs

[0073] In some embodiments, any of the guide RNA or scaffold sequences disclosed herein is chemically modified. A guide RNA or scaffold sequence comprising one or more modified nucleosides or nucleotides is called a “modified” or “chemically modified” guide RNA or scaffold sequence, to describe the presence of one or more non-naturally and/or naturally occurring components or configurations that are used instead of or in addition to the canonical A, G, C, and U residues. In some embodiments, a modified guide RNA or scaffold sequence is synthesized with a non-canonical nucleoside or nucleotide, is here called “modified.” Modified nucleosides and nucleotides can include one or more of: (i) alteration, e.g., replacement, of one or both of the non-linking phosphate oxygens and/or of one or more of the linking phosphate oxygens in the phosphodiester backbone linkage (an exemplary backbone modification); (ii) alteration, e.g., replacement, of a constituent of the ribose sugar, e.g., of the 2' hydroxyl on the ribose sugar (an exemplary sugar modification); (iii) wholesale replacement of the phosphate moiety with “dephospho” linkers (an exemplary backbone modification); (iv) modification or replacement of a naturally occurring nucleobase, including with a non-canonical nucleobase (an exemplary base modification); (v) replacement or modification of the ribose-phosphate backbone (an exemplary backbone modification); (vi) modification of the 3' end or 5' end of the oligonucleotide, e.g., removal, modification or replacement of a terminal phosphate group or conjugation of a moiety, cap or linker (such 3' or 5' cap modifications may comprise a sugar and/or backbone modification); and (vii) modification or replacement of the sugar (an exemplary sugar modification).

[0074] Chemical modifications such as those listed above can be combined to provide modified guide RNAs or scaffold sequences comprising nucleosides and nucleotides (collectively “residues”) that can have two, three, four, or more modifications. For example, a modified residue can have a modified sugar and a modified nucleobase, or a modified sugar and a modified phosphodiester. In some embodiments, every base of a guide RNA or scaffold sequence is modified, *e.g.*, all bases have a

modified phosphate group, such as a phosphorothioate group. In certain embodiments, all, or substantially all, of the phosphate groups of a guide RNA molecule or scaffold sequence are replaced with phosphorothioate groups. In some embodiments, modified guide RNAs or scaffold sequences comprise at least one modified residue at or near the 5' end of the RNA or scaffold sequence. In some embodiments, modified guide RNAs comprise at least one modified residue at or near the 3' end of the RNA or scaffold sequence.

[0075] In some embodiments, the guide RNA or scaffold sequence comprises one, two, three or more modified residues. In some embodiments, at least 5% (*e.g.*, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or 100%) of the positions in a modified guide RNA or scaffold sequence are modified nucleosides or nucleotides.

[0076] Unmodified nucleic acids can be prone to degradation by, *e.g.*, intracellular nucleases or those found in serum. For example, nucleases can hydrolyze nucleic acid phosphodiester bonds. Accordingly, in one aspect the guide RNAs or scaffold sequences described herein can contain one or more modified nucleosides or nucleotides, *e.g.*, to introduce stability toward intracellular or serum-based nucleases. In some embodiments, the modified guide RNA molecules or scaffold sequences described herein can exhibit a reduced innate immune response when introduced into a population of cells, both *in vivo* and *ex vivo*. The term "innate immune response" includes a cellular response to exogenous nucleic acids, including single stranded nucleic acids, which involves the induction of cytokine expression and release, particularly the interferons, and cell death.

[0077] In some embodiments of a backbone modification, the phosphate group of a modified residue can be modified by replacing one or more of the oxygens with a different substituent. Further, the modified residue, *e.g.*, modified residue present in a modified nucleic acid, can include the wholesale replacement of an unmodified phosphate moiety with a modified phosphate group as described herein. In some embodiments, the backbone modification of the phosphate backbone can include alterations that result in either an uncharged linker or a charged linker with unsymmetrical charge distribution.

[0078] Examples of modified phosphate groups include, phosphorothioate, phosphoroselenates, borano phosphates, borano phosphate esters, hydrogen phosphonates, phosphoroamidates, alkyl or aryl phosphonates and phosphotriesters. The phosphorous atom in an unmodified phosphate group is achiral. However, replacement of one of the non-bridging oxygens with one of the above atoms or groups of atoms can render the phosphorous atom chiral. The stereogenic phosphorous atom can possess either the "R" configuration (herein Rp) or the "S" configuration (herein Sp). The backbone can also be modified by replacement of a bridging oxygen, (*i.e.*, the oxygen that links the phosphate to

the nucleoside), with nitrogen (bridged phosphoroamidates), sulfur (bridged phosphorothioates) and carbon (bridged methylenephosphonates). The replacement can occur at either linking oxygen or at both of the linking oxygens.

[0079] The phosphate group can be replaced by non-phosphorus containing connectors in certain backbone modifications. In some embodiments, the charged phosphate group can be replaced by a neutral moiety. Examples of moieties which can replace the phosphate group can include, without limitation, *e.g.*, methyl phosphonate, hydroxylamino, siloxane, carbonate, carboxymethyl, carbamate, amide, thioether, ethylene oxide linker, sulfonate, sulfonamide, thioformacetal, formacetal, oxime, methyleneimino, methylenemethylimino, methylenehydrazo, methylenedimethylhydrazo and methyleneoxymethylimino.

[0080] Scaffolds that can mimic nucleic acids can also be constructed wherein the phosphate linker and ribose sugar are replaced by nuclease resistant nucleoside or nucleotide surrogates. Such modifications may comprise backbone and sugar modifications. In some embodiments, the nucleobases can be tethered by a surrogate backbone. Examples can include, without limitation, the morpholino, cyclobutyl, pyrrolidine and peptide nucleic acid (PNA) nucleoside surrogates.

[0081] The modified nucleosides and modified nucleotides can include one or more modifications to the sugar group, *i.e.* at sugar modification. For example, the 2' hydroxyl group (OH) can be modified, *e.g.* replaced with a number of different "oxy" or "deoxy" substituents. In some embodiments, modifications to the 2' hydroxyl group can enhance the stability of the nucleic acid since the hydroxyl can no longer be deprotonated to form a 2'-alkoxide ion.

[0082] Examples of 2' hydroxyl group modifications can include alkoxy or aryloxy (OR, wherein "R" can be, *e.g.*, alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or a sugar); polyethyleneglycols (PEG), $O(CH_2CH_2O)_nCH_2CH_2OR$ wherein R can be, *e.g.*, H or optionally substituted alkyl, and n can be an integer from 0 to 20 (*e.g.*, from 0 to 4, from 0 to 8, from 0 to 10, from 0 to 16, from 1 to 4, from 1 to 8, from 1 to 10, from 1 to 16, from 1 to 20, from 2 to 4, from 2 to 8, from 2 to 10, from 2 to 16, from 2 to 20, from 4 to 8, from 4 to 10, from 4 to 16, and from 4 to 20). In some embodiments, the 2' hydroxyl group modification can be 2'-O-Me. In some embodiments, the 2' hydroxyl group modification can be a 2'-fluoro modification, which replaces the 2' hydroxyl group with a fluoride. In some embodiments, the 2' hydroxyl group modification can include "locked" nucleic acids (LNA) in which the 2' hydroxyl can be connected, *e.g.*, by a C_{1-6} alkylene or C_{1-6} heteroalkylene bridge, to the 4' carbon of the same ribose sugar, where exemplary bridges can include methylene, propylene, ether, or amino bridges; O-amino (wherein amino can be, *e.g.*, NH_2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, or diheteroarylamino, ethylenediamine, or polyamino) and aminoalkoxy, $O(CH_2)_n$-amino, (wherein amino can be, *e.g.*, NH_2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, or diheteroarylamino, ethylenediamine, or polyamino). In some

embodiments, the 2' hydroxyl group modification can include "unlocked" nucleic acids (UNA) in which the ribose ring lacks the C2'-C3' bond. In some embodiments, the 2' hydroxyl group modification can include the methoxyethyl group (MOE), ($OCH_2CH_2OCH_3$, *e.g.*, a PEG derivative).

[0083] "Deoxy" 2' modifications can include hydrogen (*i.e.* deoxyribose sugars, *e.g.*, at the overhang portions of partially dsRNA); halo (*e.g.*, bromo, chloro, fluoro, or iodo); amino (wherein amino can be, *e.g.*, NH_2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, diheteroarylamino, or amino acid); $NH(CH_2CH_2NH)_nCH2CH_2$- amino (wherein amino can be, *e.g.*, as described herein), -NHC(O)R (wherein R can be, *e.g.*, alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or sugar), cyano; mercapto; alkyl-thio-alkyl; thioalkoxy; and alkyl, cycloalkyl, aryl, alkenyl and alkynyl, which may be optionally substituted with *e.g.*, an amino as described herein.

[0084] The sugar modification can comprise a sugar group which may also contain one or more carbons that possess the opposite stereochemical configuration than that of the corresponding carbon in ribose. Thus, a modified nucleic acid can include nucleotides containing *e.g.*, arabinose, as the sugar. The modified nucleic acids can also include abasic sugars. These abasic sugars can also be further modified at one or more of the constituent sugar atoms. The modified nucleic acids can also include one or more sugars that are in the L form, *e.g.* L- nucleosides.

[0085] The modified nucleosides and modified nucleotides described herein, which can be incorporated into a modified nucleic acid, can include a modified base, also called a nucleobase. Examples of nucleobases include, but are not limited to, adenine (A), guanine (G), cytosine (C), and uracil (U). These nucleobases can be modified or wholly replaced to provide modified residues that can be incorporated into modified nucleic acids. The nucleobase of the nucleotide can be independently selected from a purine, a pyrimidine, a purine analog, or pyrimidine analog. In some embodiments, the nucleobase can include, for example, naturally-occurring and synthetic derivatives of a base.

[0086] In embodiments employing a dual guide RNA, each of the crRNA and the tracr RNA can contain modifications. Such modifications may be at one or both ends of the crRNA and/or tracr RNA. In embodiments comprising sgRNA, one or more residues at one or both ends of the sgRNA may be chemically modified, and/or internal nucleosides may be modified, and/or the entire sgRNA may be chemically modified. Certain embodiments comprise a 5' end modification. Certain embodiments comprise a 3' end modification.

[0087] Modifications of 2'-O-methyl are encompassed.

[0088] Another chemical modification that has been shown to influence nucleotide sugar rings is halogen substitution. For example, 2'-fluoro (2'-F) substitution on nucleotide sugar rings can increase oligonucleotide binding affinity and nuclease stability. Modifications of 2'-fluoro (2'-F) are encompassed.

[0089] Phosphorothioate (PS) linkage or bond refers to a bond where a sulfur is substituted for one nonbridging phosphate oxygen in a phosphodiester linkage, for example in the bonds between

nucleotides bases. When phosphorothioates are used to generate oligonucleotides, the modified oligonucleotides may also be referred to as S-oligos.

[0090] Abasic nucleotides refer to those which lack nitrogenous bases.

[0091] Inverted bases refer to those with linkages that are inverted from the normal 5' to 3' linkage (i.e., either a 5' to 5' linkage or a 3' to 3' linkage).

[0092] An abasic nucleotide can be attached with an inverted linkage. For example, an abasic nucleotide may be attached to the terminal 5' nucleotide via a 5' to 5' linkage, or an abasic nucleotide may be attached to the terminal 3' nucleotide via a 3' to 3' linkage. An inverted abasic nucleotide at either the terminal 5' or 3' nucleotide may also be called an inverted abasic end cap.

[0093] In some embodiments, one or more of the first three, four, or five nucleotides at the 5' terminus, and one or more of the last three, four, or five nucleotides at the 3' terminus are modified. In some embodiments, the modification is a 2'-O-Me, 2'-F, inverted abasic nucleotide, PS bond, or other nucleotide modification well known in the art to increase stability and/or performance.

[0094] In some embodiments, the first four nucleotides at the 5' terminus, and the last four nucleotides at the 3' terminus are linked with phosphorothioate (PS) bonds.

[0095] In some embodiments, the first three nucleotides at the 5' terminus, and the last three nucleotides at the 3' terminus comprise a 2'-O-methyl (2'-O-Me) modified nucleotide. In some embodiments, the first three nucleotides at the 5' terminus, and the last three nucleotides at the 3' terminus comprise a 2'-fluoro (2'-F) modified nucleotide.

Ribonucleoprotein complex

[0096] In some embodiments, a composition is encompassed comprising: a) one or more guide RNAs comprising a scaffold sequence comprising any one of SEQ ID NOs: 1-5 and b) SaCas9, or any of the variant Cas9 proteins disclosed herein. In some embodiments, the guide RNA together with a Cas9 is called a ribonucleoprotein complex (RNP).

[0097] In some embodiments, chimeric Cas9 (SaCas9) nucleases are used, where one domain or region of the protein is replaced by a portion of a different protein. In some embodiments, a Cas9 nuclease domain may be replaced with a domain from a different nuclease such as Fok1. In some embodiments, a Cas9 nuclease may be a modified nuclease.

[0098] In some embodiments, the Cas9 is modified to contain only one functional nuclease domain. For example, the agent protein may be modified such that one of the nuclease domains is mutated or fully or partially deleted to reduce its nucleic acid cleavage activity.

[0099] In some embodiments, a conserved amino acid within a Cas9 protein nuclease domain is substituted to reduce or alter nuclease activity. In some embodiments, a Cas9 nuclease may comprise an amino acid substitution in the RuvC or RuvC-like nuclease domain. Exemplary amino acid substitutions in the RuvC or RuvC-like nuclease domain include D10A (based on the S. *pyogenes* Cas9 protein). *See*, *e.g.*, Zetsche et al. (2015) *Cell* Oct 22:163(3): 759-771. In some embodiments, the Cas9

nuclease may comprise an amino acid substitution in the HNH or HNH-like nuclease domain. Exemplary amino acid substitutions in the HNH or HNH-like nuclease domain include E762A, H840A, N863A, H983A, and D986A (based on the S. *pyogenes* Cas9 protein). *See*, *e.g.*, Zetsche et al. (2015). Further exemplary amino acid substitutions include D917A, E1006A, and D1255A (based on the *Francisella novicida* U112 Cpf1 (FnCpf1) sequence (UniProtKB - A0Q7Q2 (CPF1_FRATN)). Further exemplary amino acid substitutions include D10A and N580A (based on the *S. aureus* Cas9 protein). *See, e.g.*, Friedland et al., 2015, Genome Biol., 16:257. Corresponding substitutions to any of those preceding substitutions may be made to the SaCas9 protein.

[00100] In some embodiments, the Cas9 lacks cleavase activity. In some embodiments, the Cas9 comprises a dCas DNA-binding polypeptide. A dCas polypeptide has DNA-binding activity while essentially lacking catalytic (cleavase/nickase) activity. In some embodiments, the dCas polypeptide is a dCas9 polypeptide. In some embodiments, the Cas9 lacking cleavase activity or the dCas DNA-binding polypeptide is a version of a Cas nuclease (e.g., a Cas9 nuclease discussed above) in which its endonucleolytic active sites are inactivated, e.g., by one or more alterations (e.g., point mutations) in its catalytic domains. See, e.g., US 2014/0186958 A1; US 2015/0166980 A1 which are incorporated herein by reference in their entirety.

[00101] In some embodiments, the Cas9 comprises one or more heterologous functional domains (e.g., is or comprises a fusion polypeptide).

[00102] In some embodiments, the heterologous functional domain may facilitate transport of the Cas9 into the nucleus of a cell. For example, the heterologous functional domain may be a nuclear localization signal (NLS). In some embodiments, the Cas9 may be fused with 1-10 NLS(s). In some embodiments, the Cas9 may be fused with 1-5 NLS(s). In some embodiments, the Cas9 may be fused with one NLS. Where one NLS is used, the NLS may be attached at the N-terminus or the C-terminus of the Cas9 sequence, and may be directly attached or attached via a linker. It may also be inserted within the Cas9 sequence. In other embodiments, the Cas9 may be fused with more than one NLS. In some embodiments, the Cas9 may be fused with 2, 3, 4, or 5 NLSs. In some embodiments, the Cas9 may be fused with two NLSs. In certain circumstances, the two NLSs may be the same (*e.g.*, two SV40 NLSs) or different. In some embodiments, the Cas9 protein is fused with an SV40 NLS. In some embodiments, the SV40 NLS comprises the amino acid sequence of SEQ ID NO: 14 (PKKKRKV). In some embodiments, the Cas9 protein (e.g., the SaCas9 protein) is fused to a nucleoplasmin NLS. In some embodiments, the nucleoplasmin NLS comprises the amino acid sequence of SEQ ID NO: 15 (KRPAATKKAGQAKKKK). In some embodiments, the Cas9 is fused to two SV40 NLS sequences linked at the carboxy terminus. In some embodiments, the Cas9 may be fused with two NLSs, one linked at the N-terminus and one at the C-terminus. In some embodiments, the Cas9 may be fused with 3 NLSs. In some embodiments, the Cas9 may be fused with no NLS. In some embodiments, the Cas9 protein is fused to an SV40 NLS and to a nucleoplasmin NLS. In some embodiments, the SV40 NLS is fused to the C-terminus of the Cas9, while the nucleoplasmin NLS is fused to the N-terminus of the

Cas9 protein. In some embodiments, the SV40 NLS is fused to the N-terminus of the Cas9, while the nucleoplasmin NLS is fused to the C-terminus of the Cas9 protein. In some embodiments, the SV40 NLS is fused to the Cas9 protein by means of a linker. In some embodiments, the nucleoplasmin NLS is fused to the Cas9 protein by means of a linker.

[00103] In some embodiments, the heterologous functional domain may be capable of modifying the intracellular half-life of the Cas9. In some embodiments, the half-life of the Cas9 may be increased. In some embodiments, the half-life of the Cas9 may be reduced. In some embodiments, the heterologous functional domain may be capable of increasing the stability of the Cas9. In some embodiments, the heterologous functional domain may be capable of reducing the stability of the Cas9. In some embodiments, the heterologous functional domain may act as a signal peptide for protein degradation. In some embodiments, the protein degradation may be mediated by proteolytic enzymes, such as, for example, proteasomes, lysosomal proteases, or calpain proteases. In some embodiments, the heterologous functional domain may comprise a PEST sequence. In some embodiments, the Cas9 may be modified by addition of ubiquitin or a polyubiquitin chain. In some embodiments, the ubiquitin may be a ubiquitin-like protein (UBL). Non-limiting examples of ubiquitin-like proteins include small ubiquitin-like modifier (SUMO), ubiquitin cross-reactive protein (UCRP, also known as interferon-stimulated gene-15 (ISG15)), ubiquitin-related modifier-1 (URM1), neuronal-precursor-cell-expressed developmentally downregulated protein-8 (NEDD8, also called Rub1 in *S. cerevisiae*), human leukocyte antigen F-associated (FAT10), autophagy-8 (ATG8) and -12 (ATG12), Fau ubiquitin-like protein (FUB1), membrane-anchored UBL (MUB), ubiquitin fold-modifier-1 (UFM1), and ubiquitin-like protein-5 (UBL5).

[00104] In some embodiments, the heterologous functional domain may be a marker domain. Non-limiting examples of marker domains include fluorescent proteins, purification tags, epitope tags, and reporter gene sequences. In some embodiments, the marker domain may be a fluorescent protein. Non-limiting examples of suitable fluorescent proteins include green fluorescent proteins (*e.g.*, GFP, GFP-2, tagGFP, turboGFP, sfGFP, EGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), yellow fluorescent proteins (*e.g.*, YFP, EYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), blue fluorescent proteins (*e.g.*, EBFP, EBFP2, Azurite, mKalamal, GFPuv, Sapphire, T-sapphire,), cyan fluorescent proteins (*e.g.*, ECFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), red fluorescent proteins (*e.g.*, mKate, mKate2, mPlum, DsRed monomer, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRasberry, mStrawberry, Jred), and orange fluorescent proteins (mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato) or any other suitable fluorescent protein. In other embodiments, the marker domain may be a purification tag and/or an epitope tag. Non-limiting exemplary tags include glutathione-S-transferase (GST), chitin binding protein (CBP), maltose binding protein (MBP), thioredoxin (TRX), poly(NANP), tandem affinity purification (TAP) tag, myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1,

T7, V5, VSV-G, 6xHis, 8xHis, biotin carboxyl carrier protein (BCCP), poly-His, and calmodulin. Non-limiting exemplary reporter genes include glutathione-S-transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT), beta-galactosidase, beta-glucuronidase, luciferase, or fluorescent proteins.

[00105] In additional embodiments, the heterologous functional domain may target the Cas9 to a specific organelle, cell type, tissue, or organ. In some embodiments, the heterologous functional domain may target the Cas9 to muscle.

[00106] In further embodiments, the heterologous functional domain may be an effector domain. When the Cas9 is directed to its target sequence, *e.g.*, when a Cas9 is directed to a target sequence by a guide RNA, the effector domain may modify or affect the target sequence. In some embodiments, the effector domain may be chosen from a nucleic acid binding domain or a nuclease domain (e.g., a non-Cas nuclease domain). In some embodiments, the heterologous functional domain is a nuclease, such as a FokI nuclease. See, e.g., US Pat. No. 9,023,649, which are incorporated herein by reference in their entirety.

III. Methods of Gene Editing

[00107] The disclosure provides methods for gene editing. In some embodiments, any of the compositions described herein may be administered to a subject in need thereof. In some embodiments, the composition is administered for the purpose of making a double strand break in a target sequence.

[00108] In some embodiments, a method of gene editing is provided, the method comprising delivering to a cell any one or more of the compositions described herein.

[00109] In some embodiments, a method of gene editing is provided, the method comprising delivering to a cell a composition comprising: (a) a guide RNA comprising in 5' to 3' direction: (i) a nucleic acid encoding a guide sequence; and (ii) a nucleic acid encoding a scaffold sequence selected from any one of SEQ ID NOs: 1-5; and (b) a nucleic acid encoding a *Staphylococcus aureus* Cas9 (SaCas9); thereby producing a gene edit in the cell. In some embodiments, the nucleic acid encoding SaCas9 comprises the amino acid sequence of SEQ ID NO: 7. In some embodiments, the nucleic acid encoding SaCas9 is a variant of the amino acid sequence of SEQ ID NO: 7. In some embodiments, the nucleic acid encoding SaCas9 comprises an amino acid sequence selected from any one of SEQ ID NOs: 9-11. In some embodiments, the nucleic acid encoding the gRNA comprises a scaffold sequence comprising SEQ ID NO: 1. In some embodiments, the nucleic acid encoding the gRNA comprises a scaffold sequence comprising SEQ ID NO: 2. In some embodiments, the nucleic acid encoding the gRNA comprises a scaffold sequence comprising SEQ ID NO: 3. In some embodiments, the nucleic acid encoding the gRNA comprises a scaffold sequence comprising SEQ ID NO: 4. In some embodiments, the nucleic acid encoding the gRNA comprises a scaffold sequence comprising SEQ ID NO: 5.

[00110] An "edit", such as a "gene edit", as used herein refers to any insertion, deletion, or substitution in the target sequence. A gene edit includes various types of indels, e.g., indels disclosed herein.

IV. Delivery of Guide RNA Compositions

[00111] The methods and uses disclosed herein may use any suitable approach for delivering the guide RNAs and compositions described herein. Exemplary delivery approaches include vectors, such as viral vectors; lipid nanoparticles; transfection; and electroporation. In some embodiments, vectors or LNPs associated with the single-vector guide RNAs/Cas9's disclosed herein are for use in preparing a medicament for treating DM1.

[00112] Where a vector is used, it may be a viral vector, such as a non-integrating viral vector. In some embodiments, the viral vector is an adeno-associated virus vector, a lentiviral vector, an integrase-deficient lentiviral vector, an adenoviral vector, a vaccinia viral vector, an alphaviral vector, or a herpes simplex viral vector. In some embodiments, the viral vector is an adeno-associated virus (AAV) vector. In some embodiments, the AAV vector is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh10 (*see, e.g.*, SEQ ID NO: 81 of US 9,790,472, which is incorporated by reference herein in its entirety), AAVrh74 (*see, e.g.*, SEQ ID NO: 1 of US 2015/0111955, which is incorporated by reference herein in its entirety), or AAV9 vector, wherein the number following AAV indicates the AAV serotype. Any variant of an AAV vector or serotype thereof, such as a self-complementary AAV (scAAV) vector, is encompassed within the general terms AAV vector, AAV1 vector, etc. See, e.g., McCarty et al., *Gene Ther*. 2001;8:1248–54, Naso et al., *BioDrugs* 2017; 31:317-334, and references cited therein for detailed discussion of various AAV vectors.

[00113] In some embodiments, the vector (e.g., viral vector, such as an adeno-associated viral vector) comprises a tissue-specific (e.g., muscle-specific) promoter, e.g., which is operatively linked to a sequence encoding the guide RNA. In some embodiments, the muscle-specific promoter is a muscle creatine kinase promoter, a desmin promoter, an MHCK7 promoter, or an SPc5-12 promoter. In some embodiments, the muscle-specific promoter is a CK8 promoter. In some embodiments, the muscle-specific promoter is a CK8e promoter. Muscle-specific promoters are described in detail, e.g., in US2004/0175727 A1; Wang et al., *Expert Opin Drug Deliv*. (2014) 11, 345–364; Wang et al., *Gene Therapy* (2008) 15, 1489–1499. In some embodiments, the tissue-specific promoter is a neuron-specific promoter, such as an enolase promoter. See, e.g., Naso et al., *BioDrugs* 2017; 31:317-334; Dashkoff et al., *Mol Ther Methods Clin Dev*. 2016;3:16081, and references cited therein for detailed discussion of tissue-specific promoters including neuron-specific promoters.

[00114] In some embodiments, in addition to guide RNA and Cas9 sequences, the vectors further comprise nucleic acids that do not encode guide RNAs. Nucleic acids that do not encode guide RNA and Cas9 include, but are not limited to, promoters, enhancers, and regulatory sequences. In some

embodiments, the vector comprises one or more nucleotide sequence(s) encoding a crRNA, a trRNA, or a crRNA and trRNA.

[00115] Lipid nanoparticles (LNPs) are a known means for delivery of nucleotide and protein cargo, and may be used for delivery of the guide RNAs, compositions, or pharmaceutical formulations disclosed herein. In some embodiments, the LNPs deliver nucleic acid, protein, or nucleic acid together with protein.

[00116] Electroporation is a well-known means for delivery of cargo, and any electroporation methodology may be used for delivering the single vectors disclosed herein.

[00117] In some embodiments, the invention comprises a method for delivering any one of the single vectors disclosed herein to an ex vivo cell, wherein the guide RNA is encoded by a vector, associated with an LNP, or in aqueous solution. In some embodiments, the guide RNA/LNP or guide RNA is also associated with a Cas9 or sequence encoding Cas9 (e.g., in the same vector, LNP, or solution).

V. Specific Compositions and Methods of the Disclosure

[00118] Accordingly, the present disclosure relates, in particular, to the following non-limiting compositions and methods.

[00119] In a first composition, Composition 1, the present disclosure provides a composition comprising a nucleic acid encoding a guide RNA comprising a scaffold sequence selected from any one of: SEQ ID NOs: 1-5.

[00120] In another composition, Composition 2, the present disclosure provides acomposition of according to composition 1, wherein the sequence is 3' of a guide sequence.

[00121] In another composition, Composition 3, the present disclosure provides a composition of composition 1 or composition 2, wherein the guide RNA is capable of directing a *Staphylococcus aureus* Cas9 (SaCas9) to create an edit in a target sequence.

[00122] In another composition, Composition 4, the present disclosure provides a composition comprising a guide RNA comprising in 5' to 3' direction:

a. a nucleic encoding a guide sequence; and

b. a nucleic acid encoding a scaffold sequence selected from any one of: SEQ ID NOs: 1-5.

[00123] In another composition, Composition 5, the present disclosure provides a composition of any one of Compositions 1-4, further comprising a nucleic acid encoding a *Staphylococcus aureus* Cas9 (SaCas9).

[00124] In another composition, Composition 6, the present disclosure provides a composition of any one of Compositions 1-5, wherein the guide RNA is an sgRNA.

[00125] In another composition, Composition 7, the present disclosure provides a composition of any one of Compositions 1-6, wherein the guide RNA is modified.

[00126] In another composition, Composition 8, the present disclosure provides a composition of Composition 7, wherein the modification alters one or more 2' positions and/or phosphodiester linkages.

[00127] In another composition, Composition 9, the present disclosure provides a composition of any one of Compositions 7 or 8, wherein the modification alters one or more, or all, of the first three nucleotides of the guide RNA.

[00128] In another composition, Composition 10, the present disclosure provides a composition of any one of Compositions 7-9, wherein the modification alters one or more, or all, of the last three nucleotides of the guide RNA.

[00129] In another composition, Composition 11, the present disclosure provides a composition of any one of Compositions 7-10, wherein the modification includes one or more of a phosphorothioate modification, a 2'-OMe modification, a 2'-O-MOE modification, a 2'-F modification, a 2′-O-methine-4′ bridge modification, a 3′-thiophosphonoacetate modification, or a 2'-deoxy modification.

[00130] In another composition, Composition 12, the present disclosure provides a composition of any one of Compositions 1-11, wherein the composition is associated with a lipid nanoparticle (LNP).

[00131] In another composition, Composition 13, the present disclosure provides a composition of any one of Compositions 1-11, wherein the nucleic acid encoding the guide RNA and the nucleic acid encoding the SaCas9 is in a viral vector.

[00132] In another composition, Composition 14, the present disclosure provides a composition of Composition 13, wherein the viral vector is an adeno-associated virus vector, a lentiviral vector, an integrase-deficient lentiviral vector, an adenoviral vector, a vaccinia viral vector, an alphaviral vector, or a herpes simplex viral vector.

[00133] In another composition, Composition 15, the present disclosure provides a composition of Composition 14, wherein the viral vector is an adeno-associated virus (AAV) vector.

[00134] In another composition, Composition 16, the present disclosure provides a composition of Composition 15, wherein the AAV vector is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh10, AAVrh74, or AAV9 vector, wherein the number following AAV indicates the AAV serotype.

[00135] In another composition, Composition 17, the present disclosure provides a composition of Composition 16, wherein the AAV vector is an AAV serotype 9 vector.

[00136] In another composition, Composition 18, the present disclosure provides a composition of Composition 16, wherein the AAV vector is an AAVrh10 vector.

[00137] In another composition, Composition 19, the present disclosure provides a composition of Composition 16, wherein the AAV vector is an AAVrh74 vector.

[00138] In another composition, Composition 20, the present disclosure provides a composition of any one of Compositions 13-19, wherein the viral vector comprises a tissue-specific promoter.

[00139] In another composition, Composition 21, the present disclosure provides a composition of any one of Compositions 13-20, wherein the viral vector comprises a muscle-specific promoter, optionally wherein the muscle-specific promoter is a muscle creatine kinase promoter, a desmin promoter, an MHCK7 promoter, an SPc5-12 promoter, or a CK8e promoter.

[00140] In another composition, Composition 22, the present disclosure provides a composition of any one of Compositions 1-21, wherein the SaCas9 comprises the amino acid sequence of SEQ ID NO: 7.

[00141] In another composition, Composition 23, the present disclosure provides a composition of any one of Compositions 1-21, wherein the SaCas9 is a variant of the amino acid sequence of SEQ ID NO: 7.

[00142] In another composition, Composition 24, the present disclosure provides a composition of any one of Compositions 1-21, wherein the SaCas9 comprises an amino acid sequence selected from any one of SEQ ID NOs: 9-11.

[00143] In another composition, Composition 25, the present disclosure provides a composition of any one of Compositions 1-24, wherein the scaffold sequence comprises SEQ ID NO: 2.

[00144] In another composition, Composition 26, the present disclosure provides a composition of any one of Compositions 1-24, wherein the scaffold sequence comprises SEQ ID NO: 5.

[00145] In another composition, Composition 27, the present disclosure provides a composition of any one of Compositions 1-24, further comprising a pharmaceutically acceptable excipient.

[00146] In a first method of gene editing provided herein, Method 28, the present disclosure provides a method comprising delivering to a cell the composition of any one of Compositions 1-27.

[00147] In another method of gene editing as provided herein, Method 29, the present disclosure provides a method comprising delivering to a cell a composition comprising:

a. a guide RNA comprising in 5’ to 3’ direction:

i. a nucleic acid encoding a guide sequence; and

ii. a nucleic acid encoding a scaffold sequence selected from any one of SEQ ID NOs: 1-5; and

b. a nucleic acid encoding a *Staphylococcus aureus* Cas9 (SaCas9);

thereby producing a gene edit in the cell.

[00148] In another method provided herein, Method 30, the present disclosure provides a method according to Method 29, wherein the SaCas9 comprises the amino acid sequence of SEQ ID NO: 7.

[00149] In another method provided herein, Method 31, the present disclosure provides a method according to any one of Methods 29-30, wherein the SaCas9 is a variant of the amino acid sequence of SEQ ID NO: 7.

[00150] In another method provided herein, Method 32, the present disclosure provides a method according to any one of Methods 29-31, wherein the SaCas9 comprises an amino acid sequence selected from any one of SEQ ID NOs: 9-11.

[00151] In another method provided herein, Method 33, the present disclosure provides a method according to any one of Methods 29-32, wherein the scaffold sequence comprises SEQ ID NO: 2.

[00152] In another method provided herein, Method 34, the present disclosure provides a method according to any one of Methods 29-32, wherein the scaffold sequence comprises SEQ ID NO: 5.

EXAMPLES

[00153] The following examples are provided to illustrate certain disclosed embodiments and are not to be construed as limiting the scope of this disclosure in any way.

Example 1: Editing Efficiency with Different Scaffolds

[00154] To assess editing efficiencies with different scaffolds, plasmids carrying SaCas9 and spacer 1 (**Fig. 1A**) or spacer 2 (**Fig. 1B**) were transfected into HEK293T cells with Lipofectamine 2000. Scaffold sequences are shown in **Table 1**. Genomic DNA was extracted at 48 h post transfection, and a 1174 bp sequence covering the CTG repeat expansion and the sgRNA target sites was amplified by PCR. Sanger sequencing and TIDE analysis were then used to quantify the frequency of indels generated by each sgRNA. Results are shown as mean ± standard error (n = 4).

[00155] Editing efficiencies were further assessed in primary human myoblasts. SaCas9 protein and synthetic sgRNAs (with spacer 1 or spacer 2) with different scaffolds (see **Table 1** for scaffold sequences) were nucleofected into primary human myoblasts at a ratio of 1:3. Three doses were evaluated: high dose, 30 pmol SaCas9 protein; medium dose, 15 pmol SaCas9 protein; and low dose, 7.5 pmol SaCas9 protein. Genomic DNA was extracted at 72 h post nucleofection, and a 1195 bp sequence covering the CTG repeat expansion and the sgRNA target sites was amplified by PCR. Sanger sequencing and ICE analysis were then used to quantify the frequency of indels generated by each

sgRNA. Results are shown in **Fig. 2** as mean ± standard error (n = 4). Bars with "*" above contain some data points with lower R^2.

[00156] This description and exemplary embodiments should not be taken as limiting. For the purposes of this specification and appended claims, unless otherwise indicated, all numbers expressing quantities, percentages, or proportions, and other numerical values used in the specification and claims, are to be understood as being modified in all instances by the term "about," to the extent they are not already so modified. Accordingly, unless indicated to the contrary, the numerical parameters set forth in the following specification and attached claims are approximations that may vary depending upon the desired properties sought to be obtained. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should at least be construed in light of the number of reported significant digits and by applying ordinary rounding techniques.

[00157] It is noted that, as used in this specification and the appended claims, the singular forms "a," "an," and "the," and any singular use of any word, include plural referents unless expressly and unequivocally limited to one referent. As used herein, the term "include" and its grammatical variants are intended to be non-limiting, such that recitation of items in a list is not to the exclusion of other like items that can be substituted or added to the listed items.

What is claimed is:

1. A composition comprising a nucleic acid encoding a guide RNA comprising a scaffold sequence selected from any one of: SEQ ID NOs: 1-5.

2. The composition of claim 1, wherein the scaffold sequence is 3' of a guide sequence.

3. The composition of claim 1 or claim 2, wherein the guide RNA is capable of directing a *Staphylococcus aureus* Cas9 (SaCas9) to create an edit in a target sequence.

4. A composition comprising a guide RNA comprising in 5' to 3' direction:

    a. a nucleic encoding a guide sequence; and

    b. a nucleic acid encoding a scaffold sequence selected from any one of: SEQ ID NOs: 1-5.

5. The composition of any one of claims 1-4, further comprising a nucleic acid encoding a *Staphylococcus aureus* Cas9 (SaCas9).

6. The composition of any one of claims 1-5, wherein the guide RNA is an sgRNA.

7. The composition of any one of claims 1-6, wherein the guide RNA is modified.

8. The composition of claim 7, wherein the modification alters one or more 2' positions and/or phosphodiester linkages.

9. The composition of any one of claims 7-8, wherein the modification alters one or more, or all, of the first three nucleotides of the guide RNA.

10. The composition of any one of claims 7-9, wherein the modification alters one or more, or all, of the last three nucleotides of the guide RNA.

11. The composition of any one of claims 7-10, wherein the modification includes one or more of a phosphorothioate modification, a 2'-OMe modification, a 2'-O-MOE modification, a 2'-F modification, a 2′-O-methine-4′ bridge modification, a 3′-thiophosphonoacetate modification, or a 2'-deoxy modification.

12. The composition of any one of claims 1-11, wherein the composition is associated with a lipid nanoparticle (LNP).

13. The composition of any one of claims 1-11, wherein the nucleic acid encoding the guide RNA and the nucleic acid encoding the SaCas9 is in a viral vector.

14. The composition of claim 13, wherein the viral vector is an adeno-associated virus vector, a lentiviral vector, an integrase-deficient lentiviral vector, an adenoviral vector, a vaccinia viral vector, an alphaviral vector, or a herpes simplex viral vector.

15. The composition of claim 14, wherein the viral vector is an adeno-associated virus (AAV) vector.

16. The composition of claim 15, wherein the AAV vector is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh10, AAVrh74, or AAV9 vector, wherein the number following AAV indicates the AAV serotype.

17. The composition of claim 16, wherein the AAV vector is an AAV serotype 9 vector.

18. The composition of claim 16, wherein the AAV vector is an AAVrh10 vector.

19. The composition of claim 16, wherein the AAV vector is an AAVrh74 vector.

20. The composition of any one of claims 13-19, wherein the viral vector comprises a tissue-specific promoter.

21. The composition of any one of claims 13-20, wherein the viral vector comprises a muscle-specific promoter, optionally wherein the muscle-specific promoter is a muscle creatine kinase promoter, a desmin promoter, an MHCK7 promoter, an SPc5-12 promoter, or a CK8e promoter.

22. The composition of any one of claims 1-21, wherein the SaCas9 comprises the amino acid sequence of SEQ ID NO: 7.

23. The composition of any one of claims 1-21, wherein the SaCas9 is a variant of the amino acid sequence of SEQ ID NO: 7.

24. The composition of any one of claims 1-21, wherein the SaCas9 comprises an amino acid sequence selected from any one of SEQ ID NOs: 9-11.

25. The composition of any one of claims 1-24, wherein the scaffold sequence comprises SEQ ID NO: 2.

26. The composition of any one of claims 1-24, wherein the scaffold sequence comprises SEQ ID NO: 5.

27. The composition of any one of claims 1-24, further comprising a pharmaceutically acceptable excipient.

28. A method of gene editing, the method comprising delivering to a cell the composition of any one of claims 1-27.

29. A method of gene editing, the method comprising delivering to a cell a composition comprising:

a. a guide RNA comprising in 5’ to 3’ direction:

i. a nucleic acid encoding a guide sequence; and

ii. a nucleic acid encoding a scaffold sequence selected from any one of SEQ ID NOs: 1-5; and

b. a nucleic acid encoding a *Staphylococcus aureus* Cas9 (SaCas9);

thereby producing a gene edit in the cell.

30. The method of claim 29, wherein the SaCas9 comprises the amino acid sequence of SEQ ID NO: 7.

31. The method of any one of claims 29-30, wherein the SaCas9 is a variant of the amino acid sequence of SEQ ID NO: 7.

32. The method of any one of claims 29-31, wherein the SaCas9 comprises an amino acid sequence selected from any one of SEQ ID NOs: 9-11.

33. The method of any one of claims 29-32, wherein the scaffold sequence comprises SEQ ID NO: 2.

34. The method of any one of claims 29-32, wherein the scaffold sequence comprises SEQ ID NO: 5.

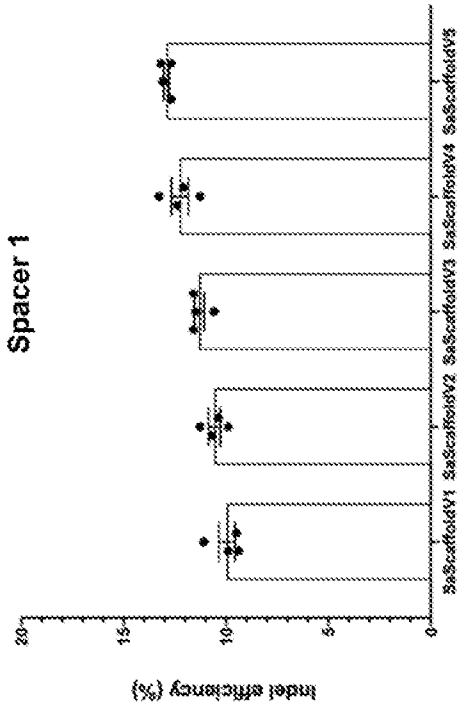
Spacer 1
Indel efficiency (%)
20
15
10
5
0
SaScaffoldV1
SaScaffoldV2
SaScaffoldV3
SaScaffoldV4
SaScaffoldV5

Fig. 1A

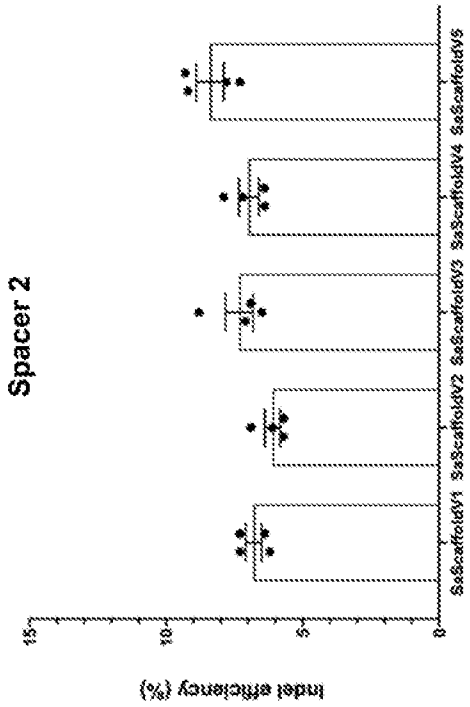
Spacer 2
Indel efficiency (%)
0
5
10
15
SaScaffoldV1
SaScaffoldV2
SaScaffoldV3
SaScaffoldV4
SaScaffoldV5

Fig. 1B

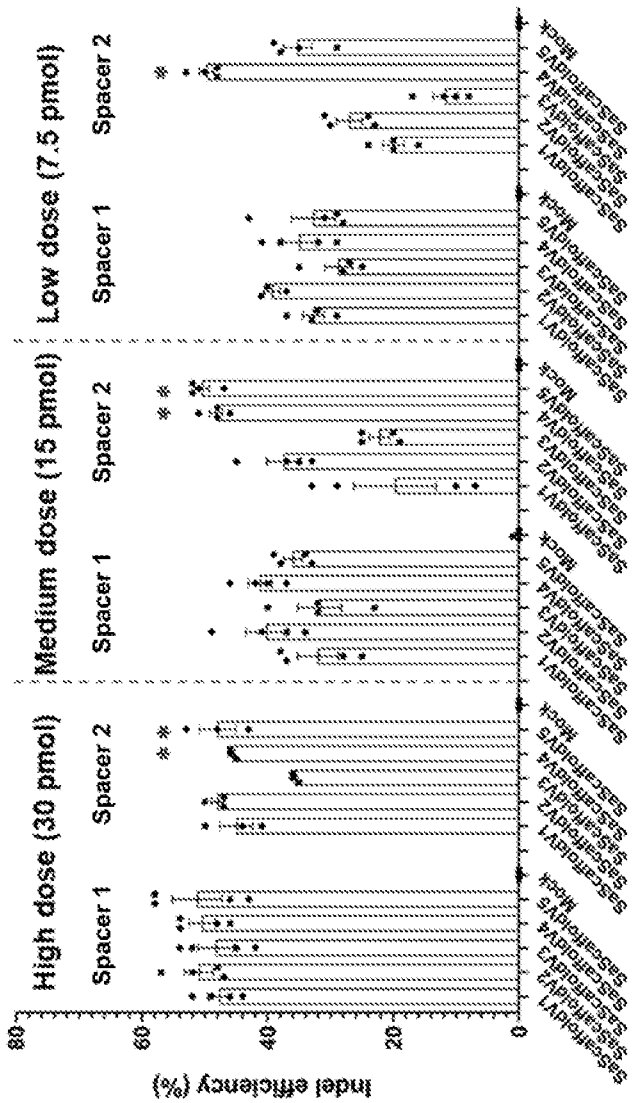
High dose (30 pmol)
Medium dose (15 pmol)
Low dose (7.5 pmol)
Spacer 1
Spacer 2
Mock
Indel efficiency (%)
0
20
40
60
80

Fig. 2

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2022/054177

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/113 C12N9/22 C12N15/10
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)
C12N C40B

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, BIOSIS, EMBASE, WPI Data, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 2020/027982 A1 (EDITAS MEDICINE INC [US]) 6 February 2020 (2020-02-06) figures 18B,25B; sequences 418,2779,2785	1-6, 12-16, 20,22, 27-30 17-19, 21,23, 24,31,32
	----- WO 2021/041546 A1 (VERTEX PHARMA [US]) 4 March 2021 (2021-03-04) pages 306,307, paragraph 513; figures 67,68; sequence 97	
X	----- WO 2020/118073 A1 (VERTEX PHARMA [US]) 11 June 2020 (2020-06-11) pages 11,12,29; sequence 12	1-6,22, 27-30 1-11,22, 27-30
	----- -/--	

[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
15 July 2022	26/09/2022

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	Romano, Alper

1

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

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2022/054177

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2020/241903 A1 (ASTELLAS PHARMA INC [JP]; MODALIS THERAPEUTICS CORP [JP]) 3 December 2020 (2020-12-03) page 42, line 26 - line 29 pages 45,46,51	17-19,21
Y	WO 2018/119359 A1 (HARVARD COLLEGE [US]) 28 June 2018 (2018-06-28) sequences 331,318	23,24, 31,32
Y	US 2021/062169 A1 (ZHENG ZONGLI [HK] ET AL) 4 March 2021 (2021-03-04) sequences 3,4	24,32
A	US 2021/015869 A1 (BURLEIGH STEPHEN MICHAEL [US] ET AL) 21 January 2021 (2021-01-21) paragraphs [0307], [0641]; sequence 158	1-24, 27-32

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2022/054177

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed:

      [X] in the form of an Annex C/ST.25 text file.

      [ ] on paper or in the form of an image file.

   b. [ ] furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. [ ] furnished subsequent to the international filing date for the purposes of international search only:

      [ ] in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).

      [ ] on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).

2. [ ] In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2022/054177

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims;; it is covered by claims Nos.:

1-24, 27-32(all partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-24, 27-32(all partially)

A composition comprising a nucleic acid encoding a guide RNA comprising a scaffold sequence of SEQ ID NO: 1 wherein the guide RNA is especially capable of directing a Staphylococcus aureus Cas9 (SaCas9) to create an edit in a target sequence.

---

2. claims: 25, 33(completely); 1-24, 27-32(partially)

A composition comprising a nucleic acid encoding a guide RNA comprising a scaffold sequence of SEQ ID NO: 2 wherein the guide RNA is especially capable of directing a Staphylococcus aureus Cas9 (SaCas9) to create an edit in a target sequence.

3-4. claims: 1-24, 27-32(all partially)

A composition comprising a nucleic acid encoding a guide RNA comprising a scaffold sequence selected from SEQ ID NOs: 3 or 4 wherein the guide RNA is especially capable of directing a Staphylococcus aureus Cas9 (SaCas9) to create an edit in a target sequence.

---

5. claims: 26, 34(completely); 1-24, 27-32(partially)

A composition comprising a nucleic acid encoding a guide RNA comprising a scaffold sequence of SEQ ID NO: 5 wherein the guide RNA is especially capable of directing a Staphylococcus aureus Cas9 (SaCas9) to create an edit in a target sequence.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/IB2022/054177

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2020027982 A1	06-02-2020	AU 2019314208 A1	18-02-2021
		CA 3108376 A1	06-02-2020
		EP 3830264 A1	09-06-2021
		WO 2020027982 A1	06-02-2020
WO 2021041546 A1	04-03-2021	AU 2020337919 A1	24-03-2022
		BR 112022003505 A2	24-05-2022
		CA 3152288 A1	04-03-2021
		EP 4022057 A1	06-07-2022
		IL 290575 A	01-04-2022
		KR 20220070443 A	31-05-2022
		TW 202118873 A	16-05-2021
		US 2022186216 A1	16-06-2022
		WO 2021041546 A1	04-03-2021
WO 2020118073 A1	11-06-2020	AU 2019391114 A1	24-06-2021
		CA 3121781 A1	11-06-2020
		CO 2021007320 A2	21-06-2021
		EA 202191555 A1	03-09-2021
		EP 3891283 A1	13-10-2021
		IL 283631 A	29-07-2021
		US 2021403906 A1	30-12-2021
		WO 2020118073 A1	11-06-2020
WO 2020241903 A1	03-12-2020	AU 2020283372 A1	18-11-2021
		BR 112021023899 A2	18-01-2022
		CA 3135271 A1	03-12-2020
		CN 113785066 A	10-12-2021
		CO 2021014746 A2	19-11-2021
		EP 3976785 A1	06-04-2022
		IL 287178 A	01-12-2021
		JP 2022534627 A	02-08-2022
		KR 20220012926 A	04-02-2022
		SG 11202111427R A	29-11-2021
		TW 202111122 A	16-03-2021
		US 2022233721 A1	28-07-2022
		WO 2020241903 A1	03-12-2020
WO 2018119359 A1	28-06-2018	US 2018179503 A1	28-06-2018
		US 2020399619 A1	24-12-2020
		WO 2018119359 A1	28-06-2018
US 2021062169 A1	04-03-2021	NONE	
US 2021015869 A1	21-01-2021	AU 2019247199 A1	15-10-2020
		BR 112020020073 A2	05-01-2021
		CA 3095084 A1	10-10-2019
		CN 112585277 A	30-03-2021
		EP 3775237 A1	17-02-2021
		IL 277704 A	30-11-2020
		JP 2021520211 A	19-08-2021
		KR 20210020873 A	24-02-2021
		MA 52207 A	17-02-2021
		RU 2020136054 A	06-05-2022
		SG 11202009446T A	29-10-2020
		US 2021015869 A1	21-01-2021
		WO 2019195491 A1	10-10-2019